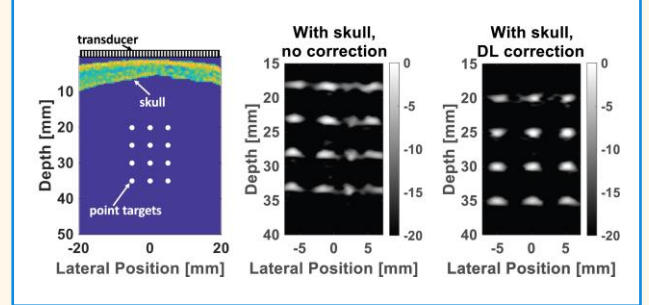


Transcranial Phase Correction Using Pulse-echo Ultrasound and Deep Learning: A 2D Numerical Study

Zixuan Tian, *Student Member, IEEE*, Matthew Olmstead, Yun Jing, *Senior Member, IEEE*, and Aiguo Han, *Member, IEEE*

Abstract—Phase aberration caused by human skulls severely degrades the quality of transcranial ultrasound images, posing a major challenge in the practical application of transcranial ultrasound techniques in adults. Aberration can be corrected if the skull profile (i.e., thickness distribution) and speed of sound (SOS) are known. However, accurately estimating the skull profile and SOS using ultrasound with a physics-based approach is challenging due to the complexity of the interaction between ultrasound and the skull. A deep learning approach is proposed herein to estimate the skull profile and SOS using ultrasound radiofrequency (RF) signals backscattered from the skull. A numerical study was performed to test the approach's feasibility. Realistic numerical skull models were constructed from computed tomography (CT) scans of five *ex vivo* human skulls in this numerical study. Acoustic simulations were performed at various skull locations ($N=3,595$) to generate array-based ultrasound backscattered signals. A deep learning model was developed and trained to estimate skull thickness and SOS from RF channel data. The trained model was shown to be highly accurate. The mean absolute error was 0.15 mm (2% error) for thickness estimation and 13 m/s (0.5% error) for SOS estimation. The Pearson correlation coefficient between the estimated and ground-truth values was 0.99 for thickness and 0.95 for SOS. Aberration correction performed using deep-learning-estimated skull thickness and SOS values yielded significantly improved beam focusing (e.g., narrower beams) and transcranial imaging quality (e.g., improved spatial resolution and reduced artifacts) compared with no aberration correction. The results demonstrate the feasibility of the proposed approach for transcranial phase aberration correction.

Index Terms—Deep learning, phase aberration correction, pulse-echo ultrasound, transcranial ultrasound.



I. INTRODUCTION

TRANSCRANIAL ultrasound could enable a wide variety of applications, such as brain imaging [1, 2], stroke diagnosis [3], intracerebral hemorrhage detection [4], brain perfusion evaluation [5], and endonasal trans-sphenoidal surgery [6]. Compared with other imaging modalities (e.g., computed tomography [CT] and magnetic resonance imaging), ultrasound imaging has the intrinsic advantages of being real-time, affordable, portable, noninvasive, and nonionizing. However, a large speed of sound (SOS) mismatch between the human skull and the brain [7, 8] combined with the strong scattering and absorption [9] severely distort the ultrasound wavefront, leading to highly degraded ultrasound images. In

particular, skull phase aberration is a major limiting factor for the transcranial application of ultrasound.

Many phase correction techniques have been developed for transcranial ultrasound. Phase correction can be accomplished by using the experimental reference signals or a known skull profile and speed of sound (SOS). The experimental signal-guided techniques include the representative time-reversal method [10-12], which was initially presented as a wave-mirroring method by Fink *et al.* [10]. In this method, the transcranial temporal pressure waveform is stored, reversed, and re-emitted, providing both phase and amplitude correction. However, this particular time reversal method is not practical because it requires a sensor inside the brain. Vignon *et al.* [13] made further improvements by placing two identical linear arrays, one on each side of a skull and using a spatiotemporal

Research reported in this publication was supported in part by the National Institutes of Health (R21EB032638) and the National Center for Supercomputing Applications (NCSA) Faculty Fellowship Program. The work made use of NCSA's computing resources, including GPU clusters in the Innovative Systems Laboratory and the Delta Supercomputer supported by the National Science Foundation (award OCI 2005572) and the State of Illinois.

Z. Tian is with Department of Electrical and Computer Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801 USA (e-mail: zixuant5@illinois.edu).

M. Olmstead and Y. Jing are with the Graduate Program in Acoustics, Penn State University, University Park, PA 16802 USA (e-mail: mjo5585@psu.edu, jing.yun@psu.edu).

A. Han is with the Department of Biomedical Engineering and Mechanics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061 USA, and holds an affiliate appointment with Department of Electrical and Computer Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801 USA (e-mail: aiguo@han@vt.edu).

Highlights

- A deep learning method was proposed to estimate the skull profile and speed of sound using ultrasound echo signals from the skull, based on which a phase aberration correction approach was developed.
- Aberration correction using deep-learning-estimated skull profile and speed of sound showed significantly improved quality for transcranial imaging and beam focusing in this numerical study.
- This study demonstrated the feasibility of the new skull aberration correction approach, providing a potentially practical method to improve transcranial ultrasound imaging and therapy.

inverse filter to estimate the phase aberration. Some experimental signal-guided methods [14-18] used a near-field phase-screen model coupled with a multi-lag, least-squares, correlation-based phase error algorithm. These methods assume a negligibly thin aberrator, which serves as an approximation for the more complex aberration profiles such as a skull. Thus, the application of these methods is mainly limited to the temporal bone, where the skull is thin (the mean thickness reported to be 3-4 mm). In addition, Clement et al. [19] induced a shear mode for brain imaging as the shear wave speed in the skull is close to the longitudinal sound speed of brain tissues and provides less refraction, but its applicability has been restricted by extremely high attenuation of shear waves in the skull [20].

Approaches modeling refraction based on the accurate geometry and SOS of the skull have been proposed. Some studies [21, 22] have extracted skull information from CT images based on the empirical correlation between Hounsfield units and SOS. Although this CT-based approach has demonstrated success in brain therapy, it is less suitable for ultrasound brain imaging due to the added ionizing radiation associated with additional CT scans, which contradicts the goal of ultrasound imaging. There is a need to develop a robust approach to extract the skull profile and SOS by directly using ultrasound, preferably with the same imaging probe. Wydra et al. [23] introduced the variable focus technique and measured the thickness and SOS by searching for maximum reflected amplitude. This method requires clearly distinct reflected echoes from the near and far surfaces of the skull. Mozaffarzadeh et al. [24] used a single probe and the bidirectional headwave technique to estimate the compressional wave speed in the skull, focusing specifically on the human temporal window. Existing ultrasound-based methods for extracting the skull profile and SOS explicitly model the interactions between ultrasound and the skull. While promising, these methods often require simplified assumptions such as the skull being a homogenous aberrator or distinct reflected echoes from skull surfaces being available. These simplified assumptions may limit the ultimate capability of the methods that are based on explicit modeling of physics. The interactions between ultrasound and the skull are complex, and explicitly modeling the physical process is challenging. Data-driven approaches such as deep learning may be a promising alternative to avoid explicit modeling of complex physics while achieving accurate results.

A deep learning method is proposed herein to estimate the skull profile and SOS with ultrasound radiofrequency (RF) signals backscattered from the skull. Various studies have

shown that deep learning can effectively extract tissue properties from raw RF data (e.g., [25, 26]). In this paper, a deep learning model is developed to estimate skull thickness and average SOS along each scan line of an ultrasound scan. The resulting skull thickness and SOS values will allow the reconstruction of the skull profile and the lateral distribution of the SOS that can be subsequently used for phase aberration correction using a fast marching method (FMM) [27].

This numerical study aims to test the feasibility of our deep-learning-based phase correction approach for pulse-echo ultrasound. The paper is organized as follows: Sec. II describes our proposed approach and the methodology used for this numerical study. In Sec. III, phase correction results of focusing and B-mode imaging are presented in detail. Sec. IV and V are the discussion and conclusion of this paper, respectively.

II. METHODOLOGY

Our approach consists of the following steps:

- 1) *Skull outer surface detection*: A standard synthetic aperture scan [28] is performed on the skull using an array transducer to detect the skull's outer surface.
- 2) *Skull thickness and SOS scan*: Without changing the transducer location, a second ultrasound scan is performed to acquire RF signals backscattered from the skull for skull thickness and SOS estimation. The ultrasound beam is focused inside the skull for this scan. The acquired RF data are to be used by a deep learning model to estimate the skull thickness and SOS. The deep learning model needs to be trained *a priori*.
- 3) *Skull SOS map reconstruction*: Combining the results from steps 1) and 2), a skull SOS map is reconstructed to represent the skull's inner and outer surfaces and the lateral distribution of skull SOS. Our phase aberration correction method does not require the axial distribution of skull SOS.
- 4) *Time delay computation*: The FMM is applied to the skull SOS map reconstructed in step 3) to compute the time delay to be applied to each transducer element for aberration correction. The time delay information will be used subsequently to deliver focused acoustic energy to the brain for therapy or to perform aberration-corrected transcranial imaging scans of the brain.

All ultrasound data acquisitions involved in the above steps can be performed using the same array transducer scanning the same skull location, allowing ease of use and potential real-time implementation of this approach.

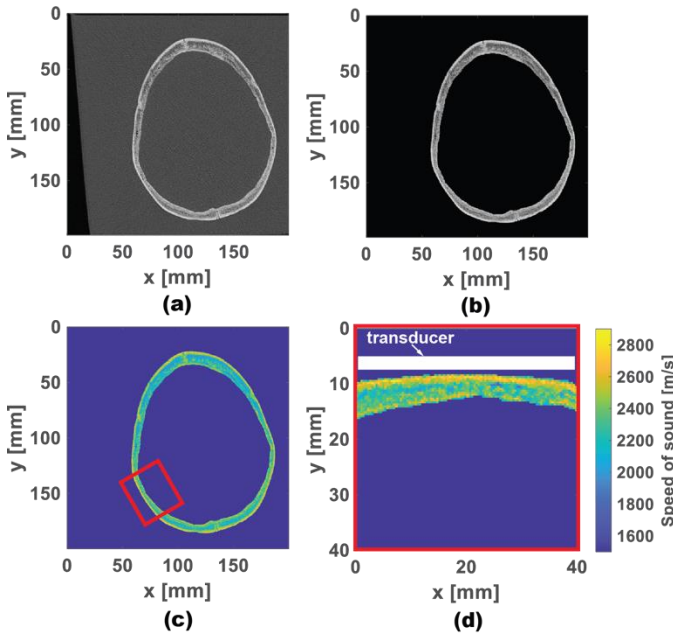


Fig. 1. (a) A representative slice of the CT scan of an *ex vivo* human skull. (b) A processed CT image with the skull segmented from the background. (c) The SOS map derived from (b) with a 40-mm-long skull segment indicated by the red rectangular box. (d) The 40-mm-long segment after rotation.

The feasibility of this approach is evaluated herein through a numerical study, where acoustic simulations are performed on realistic skull models derived from the CT scans of 5 *ex vivo* skulls. The simulations were performed using the k-Wave toolbox [29], which has been shown to be accurate and efficient for transcranial ultrasound simulations [30]. The shear wave was not considered because the ultrasound beam was mostly perpendicular to the skull surface in our simulations and shear waves are strongly attenuated in the skull at the frequency range considered in this work. It has been demonstrated that skull phase errors due to shear waves can be negligible on the focusing for angles of incidence up to 20° [31]. Two-dimensional simulations were performed for computational efficiency. The performance of the proposed phase aberration correction approach was demonstrated in terms of the accuracy in skull thickness and SOS estimates, and the effectiveness of phase aberration correction in transcranial beam focusing and imaging.

A. Numerical Skull Models

Numerical skull models were derived from archived diagnostic CT scan data of 5 *ex vivo* human skulls. CT scans were conducted using a Siemens SOMATOM CT Scanner, utilizing a bone reconstruction kernel (AH82) to obtain image intensities directly proportional to bone density. The acquired CT images featured a slice thickness of 1 mm, a spatial resolution of 0.39 mm/pixel, and a size of 200 mm \times 200 mm (Fig. 1a) [32]. The CT images were processed to segment the skull from the background (Fig. 1b). Density and SOS maps of the skull were created from the segmented CT images by assuming linear relationships [33, 34] between the CT Hounsfield Unit values and SOS and density values. The

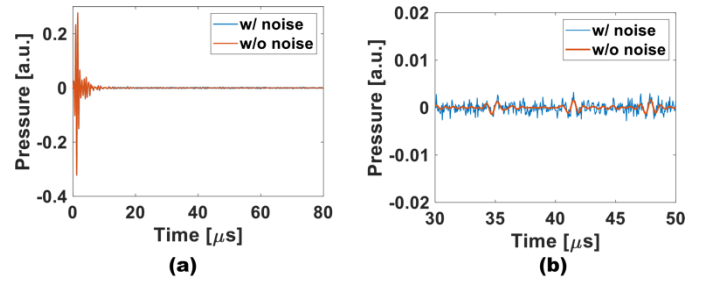


Fig. 2. The noise levels used in the simulations. Representative RF signals obtained by simulating transcranial imaging of point targets, without noise and with noise shown in different temporal windows of (a) [0 μ s, 80 μ s], and (b) [30 μ s, 50 μ s]. Echoes from the point targets appear around 35, 41, and 48 μ s.

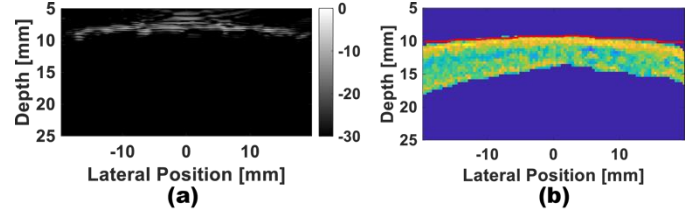


Fig. 3. (a) A B-mode image of a skull segment generated by the synthetic aperture approach. (b) The corresponding ground truth SOS map of the skull segment, superimposed with a red line representing the outer surface of the skull detected by the synthetic aperture approach.

density and SOS were set to be 1000 kg/m³ and 1500 m/s, respectively, for the background.

Simulations were performed on skull segments rather than the entire skull to mimic the conditions of real applications where the ultrasound transducer has a finite footprint covering a small region of the skull, as well as to minimize the computation. The skull density and SOS maps were segmented into 40 mm long segments (Fig. 1c), with a 20-mm overlap between adjacent segments. The transducer was placed at a fixed location in the simulation, with the transducer surface being parallel to the horizontal direction. Each skull segment was rotated to make its outer surface as parallel to the horizontal transducer as possible (Fig. 1d). We obtained 3,595 2D segments from 5 *ex vivo* human skulls.

The spatial resolution of the skull density and SOS maps were increased from 0.39 mm/pixel to 0.10 mm/pixel *via* image interpolation for more accurate acoustic simulations at the frequency of interest. A nearest neighbor interpolation was used to preserve the discontinuity of the density and SOS values at the edges of the skull. Acoustic simulations were performed on each segment after interpolation.

The skull attenuation was considered to better model the realistic conditions in the simulation. A spatially homogeneous skull attenuation coefficient of 17 dB/cm at 1.5 MHz and a power law frequency dependence with an exponent of 2 was used in the simulation. Spatially homogeneous attenuation for the skull has been shown to result in a good agreement between simulations and experiments [35]. A study [9] measured the attenuation of the skull bone with a center frequency of 1 MHz and determined the average attenuation coefficient to be 13.3 ± 0.97 dB/cm. Another group [36] investigated the attenuation coefficient of freshly excised human skulls and found the attenuation coefficient to be 27 dB/cm at 1.4 MHz for cortical bones.

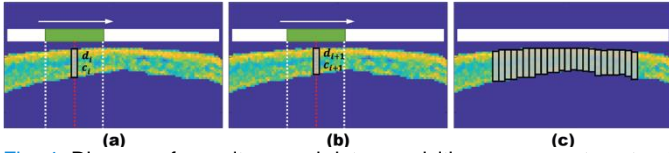


Fig. 4. Diagram of our ultrasound data acquisition sequence to extract the skull thickness and sound speed. The white bar represents the array, and active elements are marked in green. (a) The ground-truth local thickness and SOS of the scanline (the narrow black rectangle) were calculated by spatially averaging the thickness and SOS of the skull below the active elements. (b) Shifting the active elements by an index of 1 and repeating the procedure in (a) yielded the local thickness and SOS of the next scanline. (c) Combining the results from all acquisitions yielded the skull profile and SOS lateral distribution.

B. Simulation Settings

A linear array was used as the transducer to transmit and receive ultrasound waves in all of our k-Wave simulations. The central frequency was 1.5 MHz, and the pitch was 0.8 mm, equivalent to 4/5 of the wavelength at the central frequency. The frequency response of the array was modeled as a Gaussian function with an 80% bandwidth for transmission and reception.

Noise was considered in our simulation to investigate its effects on the performance of our proposed phase aberration correction approach. Simulations were performed to image point targets through the skull. Therefore, two noise levels were considered: 1) zero noise, and 2) a noise level that allowed visualization of point targets through the skull. The signal-to-noise ratio (SNR) was 10 dB with noise, calculated by assuming the “signals” to be those from the point targets we simulated.

C. Ultrasound Data Acquisitions for Skull Outer Surface Detection and Skull Thickness and SOS Estimation

A standard synthetic aperture approach [28] was applied to detect the outer surface of the skull. In each emission of synthetic aperture imaging, a single element was used to transmit a wave covering the full image region, and all the elements were to receive the echoes. After all emissions, the delay and sum (DAS) algorithm [37] was used to generate B-mode images (Fig. 3a) that allowed us to identify the outer surface of the skull (Fig. 3b).

The ultrasound data acquisition process for skull thickness and SOS estimation is described as follows. The simulated array transducer consisted of a total of 50 elements, with 15 active elements simultaneously excited to form a beam focused at 5 mm below the transducer surface (inside the skull), generating a virtual scanline (red dashed line in Fig. 4a). Electronic scanning was performed by continuously shifting active elements by an index of 1 to move the scanline. For each scanline, the backscattered RF signals at all active elements were recorded, yielding a total of 15 channels of RF signals to be used by a deep learning model to estimate the local skull thickness and SOS for that scanline. The width of the virtual scanline (the narrow black rectangle in Fig. 4a) was equal to the pitch of the transducer by definition. However, the ground truth values of the local thickness and SOS for each scanline were calculated by spatially averaging the thickness and SOS of the

entire skull region below the active elements (i.e., the skull region within the two white dashed lines in Fig. 4a). This spatial averaging procedure served to reduce variance, resulting in a homogenized ground truth SOS map.

In real-world scenarios, the transducer would touch the human scalp (typical thickness, 1 mm [38]) during ultrasound data acquisition. Therefore, the transducer was placed 1 mm above the skull in the simulations to account for the thickness of the scalp.

D. Deep-Learning-Based Skull SOS and Thickness Estimation

A deep learning model was developed to estimate the local skull thickness and SOS using the frequency-domain representation of the 15-channels of RF signals. The neural network architecture (Fig. 5a) consisted of multiple layers of operations, starting with two linear layers, followed by two convolutional blocks, and ending with two fully connected layers. The first convolutional block consisted of a one-dimensional (1-D) convolutional layer (convolution along the frequency axis), an activation layer and a pooling layer. The second convolutional block consisted of a 1-D convolutional layer, a batch normalization layer, an activation layer, a pooling layer, and a dropout layer.

Fast Fourier transform was applied to the 15 channels of RF data to yield the frequency domain representation of the signals. The deep neural network used the signals in the frequency range of 0-4 MHz because signal energy outside this frequency range was negligible. The real and imaginary parts of the frequency-domain RF data were both considered, resulting in 30 channels of data (15 channels for the real part and 15 channels for the imaginary part).

Two linear layers were used before the first convolutional layer to take the physics of RF channel data into account (Supplementary Materials Fig. S5). The input data were represented by a 30×64 matrix, corresponding to 30 channels \times 64 data points per channel. Without linear layers, the convolutional layer would treat each channel equally and ignore the phase information and the correlation between signals received by array elements. However, the complexity of the scattering of skull in the near field made it difficult to explicitly find the physical connection between input channel signals. The two trainable linear layers addressed this issue by linearly transforming the channel data to implicitly model the physical connection between signals from the 15 array elements. The first linear layer rearranged the 64 data points and the second linear layer processed the 30 channels. Matrix multiplications in the frequency domain enabled the computation of signals in the time domain (such as linear combination, time shift, phase shift, and convolution). The two linear layers served as a preprocessing step before the first convolutional layer and were expected to improve model performance by identifying the optimal linear transformation of the channel data.

Batch normalization (BN) [39] and dropout [40] were used in the second convolutional block. The BN layer was used to reduce the shift of internal covariance and accelerate the training process of the deep neural network. Dropout was applied to prevent the network from overfitting.

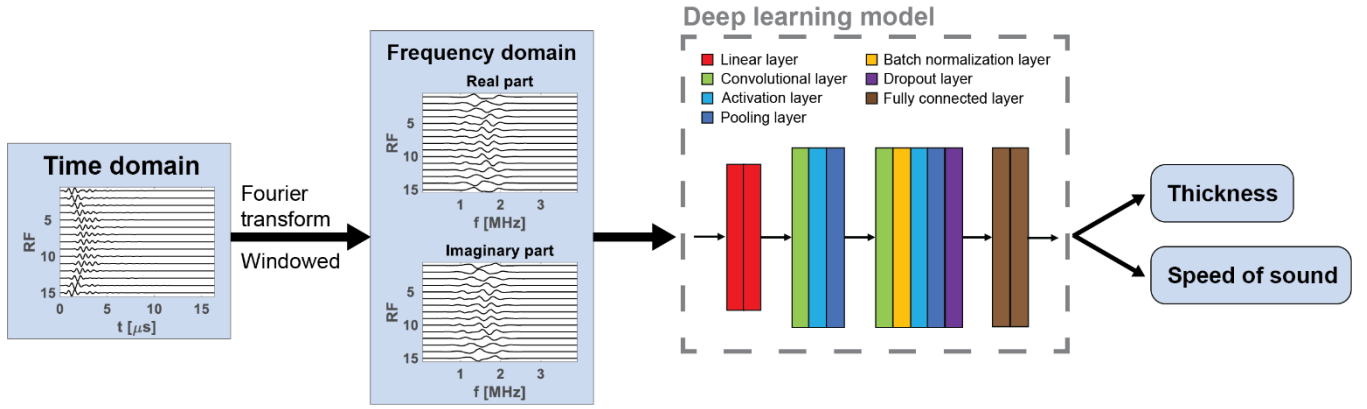


Fig. 5. Diagram illustrating the use of neural network to predict the skull thickness and SOS based on backscattered RF signals.

For deep learning model training and evaluation, the 3,595 skull segments were split into three sets, the training set, the validation set, and the test set, which were correspondingly 60%, 20%, and 20% of the entire skull segments. To prevent data leakage, skull segments from the same CT image were assigned to the same set.

To evaluate the skull thickness and SOS estimation performance of the neural network model, Pearson correlation coefficient [41] and mean absolute error (MAE) between the ground truth and estimated values were calculated. Pearson correlation coefficient describes the strength and direction of the linear relationship between two sets of data and equals unity for perfect estimates. Mean absolute errors represented the quantity difference and allocation difference between the ground truth and estimated values [42].

E. Reconstruction of Skull SOS Maps

The SOS map was reconstructed for each skull segment to represent the skull shape and the lateral distribution of the SOS. The outer skull surface was determined by the synthetic aperture scan as described in Section II-C. Adding the deep-learning-estimated thickness to the outer surface yielded the inner surface of the skull. The lateral distribution of the SOS was obtained from the deep learning estimates. The deep learning model yielded 36 local thickness values and 36 SOS values that could be used to construct 36 consecutive scanlines for each skull segment, representing the lateral distribution of thickness and SOS. Considering the width of the scanline ($=0.8$ mm), 36 consecutive scanlines spanned a total length of 28.8 mm, which was shorter than the full length of the skull segment. Therefore, the thickness and SOS estimates of the leftmost (rightmost) scanline were extrapolated using the same estimated values to the left (right) edge of the skull segment to cover the full length.

F. Phase Aberration Correction

Transcranial focusing and imaging were simulated to assess the performance of phase aberration correction.

1) Focusing and B-mode imaging simulation

Transcranial focusing and B-mode imaging were simulated with the array placed 1 mm above the skulls, the same as the simulations of acquiring RF signals for SOS and thickness estimation. While the beams were focused at 5 mm for thickness and SOS estimation, the focus was 30 mm for

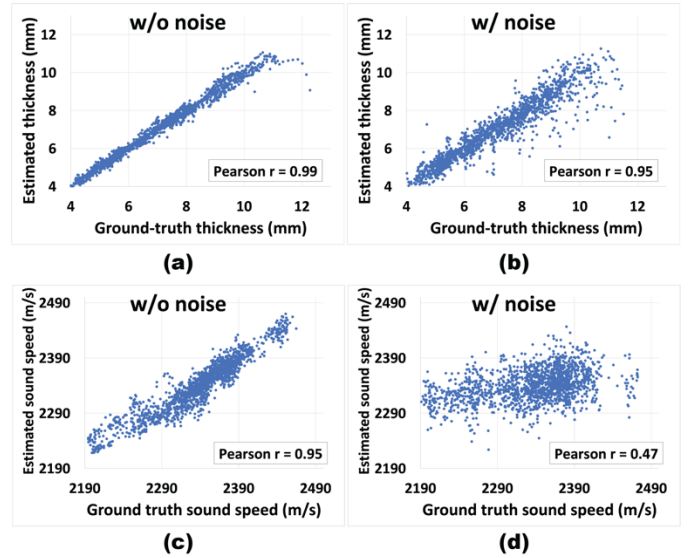


Fig. 6. Scatter plots between deep-learning-estimated thickness values and ground truth, obtained from (a) noise-free RF signals and (b) noisy RF signals in the test set. Scatter plot between deep-learning-estimated SOS values and ground truth, obtained from (c) noise-free RF signals and (d) noisy RF signals in the test set.

transcranial focusing and imaging. For imaging purposes, each focused beam was generated using 32 active consecutive elements (instead of 15 active elements as used in SOS and thickness estimation) to have a better lateral resolution. In the imaging process, twelve point targets with a radius of 0.3 mm were placed 20, 25, 30, and 35 mm below the array (Fig. 10a). The density of point targets was 4000 kg/m^3 , and the SOS was 2000 m/s .

2) The fast marching method

The FMM [43] was used to correct the phase aberration caused by skulls. The FMM computed the numerical solutions of the eikonal equation that can be considered as an approximation to the wave equation:

$$|\nabla T|^2 = 1/c^2 \quad (1)$$

where T is the arrival time of the acoustic wave at a point and c is the spatially varying SOS. The arrival times computed from the eikonal equation were directly linked to the desired time delay of the transducer element.

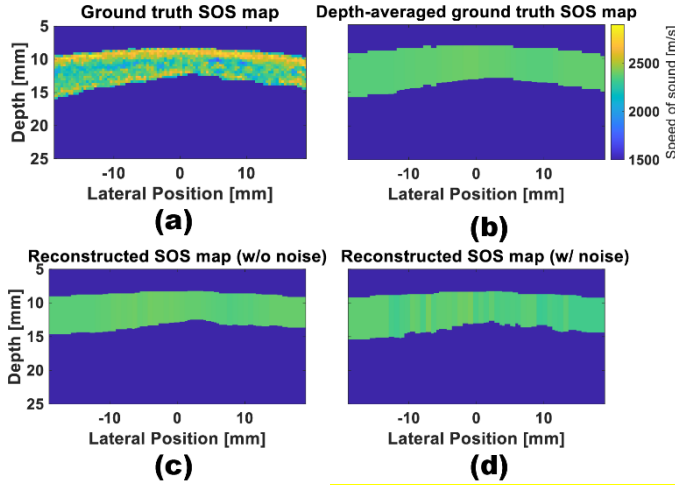


Fig. 7. (a) Ground truth SOS map and (b) depth-averaged ground truth SOS map of the skull segment. Reconstructed SOS maps of the skull segment based on detected outer surface of the skull and thickness and SOS values estimated from (c) noise-free RF signals and (d) noisy RF signals.

We used FMM with the reconstructed SOS maps to calculate accurate transmit time delays to correct phase aberration for better focusing. Point targets were imaged through the skull with the corrected beam. FMM was then applied to calculate accurate receive time delays in the DAS algorithm during B-mode image construction, which yielded phase corrected B-mode images.

The FMM was performed using the open-source implementation in MATLAB and C++ [44] (available at <https://github.com/patnr/FM>). The version used in this study was version 1.0, and the code was executed without modification.

III. RESULTS

A. Deep-Learning-Based Skull SOS and Thickness Estimation

When the RF signals were noise-free, our network yielded a MAE of 0.15 mm and a Pearson correlation coefficient of 0.99 for thickness estimation in the test set. The MAE was 13 m/s, and Pearson correlation coefficient was 0.95 for SOS estimation in the test set. For comparison, the mean value of ground-truth skull thickness was 6.3 mm, and mean value of ground-truth skull SOS was 2341 m/s. Scatter plots between deep-learning-estimated values and the ground truth of the test set are shown in Fig. 6 a and c.

The linear layers added to the network improved the model performance as shown in Table I. For instance, the linear layers decreased the MAE from 0.24 mm to 0.15 mm for thickness estimation, and from 22 m/s to 13 m/s for SOS estimation.

The SOS maps of skull segments were reconstructed using the deep-learning-estimated thickness and SOS (Fig. 7c and 7d). The profiles of reconstructed maps were close to the ground truth maps. Two more examples of reconstructed SOS maps are presented in Supplementary Materials (Fig. S1 and S3). Although the reconstructed SOS maps only provided the lateral distribution of skull SOS without considering the axial distribution, the phase correction performance was comparable to that of the ground truth SOS maps, as presented in Sections III-B and III-C.

TABLE I
PEARSON CORRELATION COEFFICIENTS AND MEAN ABSOLUTE ERRORS FOR THICKNESS AND SOS ESTIMATION

		Thickness estimation performance		SOS estimation performance	
		r	MAE (mm)	r	MAE (m/s)
w/o noise	w/o linear layers	0.98	0.24	0.85	22
	w/ linear layers	0.99	0.15	0.95	13
w/ noise	w/o linear layers	0.93	0.40	0.44	39
	w/ linear layers	0.95	0.35	0.47	39

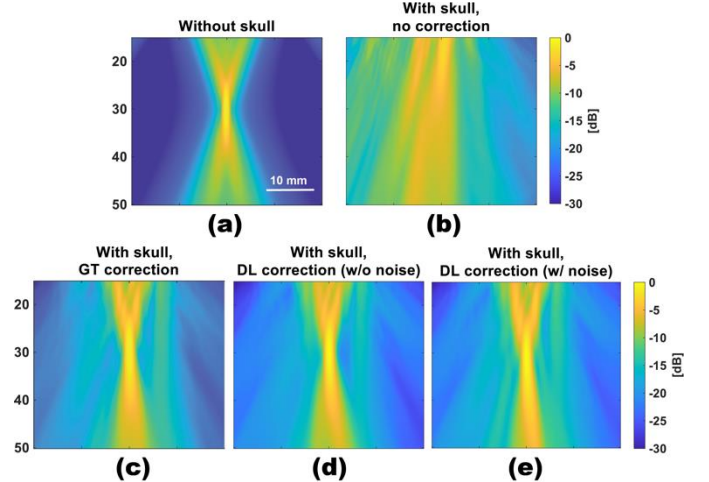


Fig. 8. Profiles of beams focused at 30 mm (a) without skull, (b) with skull but no correction, (c) with skull and corrected by the ground truth SOS map, (d) with skull and corrected by the deep-learning-estimated SOS map when no noise was present in RF signals, and (e) with skull and corrected by the SOS map reconstructed by our method when noise was present in RF signals. Relative levels are consistent.

When noise was present in the RF signals, the MAE was 0.35 mm (5.6% of the mean ground truth) and Pearson correlation coefficient was 0.95 in the test set for thickness estimation. For SOS estimation, the MAE was 39 m/s (1.7% of the mean ground truth), and Pearson correlation coefficient was 0.47. Scatter plots between deep-learning-estimated values and the ground truth are shown in Fig. 6 b and d.

Similar to the noise-free case, the linear layers also improved the deep learning performance when noise was present (Table I). For thickness estimation, the linear layers reduced the MAE from 0.40 mm to 0.35 mm and increased the Pearson correlation coefficient from 0.93 to 0.95. For SOS estimation, although the MAE remained unchanged (39 m/s), Pearson correlation coefficient increased from 0.44 to 0.47.

B. Transcranial Ultrasound Focusing

Correcting the phase aberration using the reconstructed SOS maps and the FMM significantly improved the focusing of ultrasound beams through the skull. A noise-free example is shown in Fig. 8, which compares the beam profiles of focused beams generated by using the central 32 active elements of the array under various conditions, including focusing without the skull (Fig. 8a), focusing through a skull segment without aberration correction (Fig. 8b), focusing through the same skull segment with aberration correction using ground-truth skull thickness and SOS values (Fig. 8c), and focusing using deep-

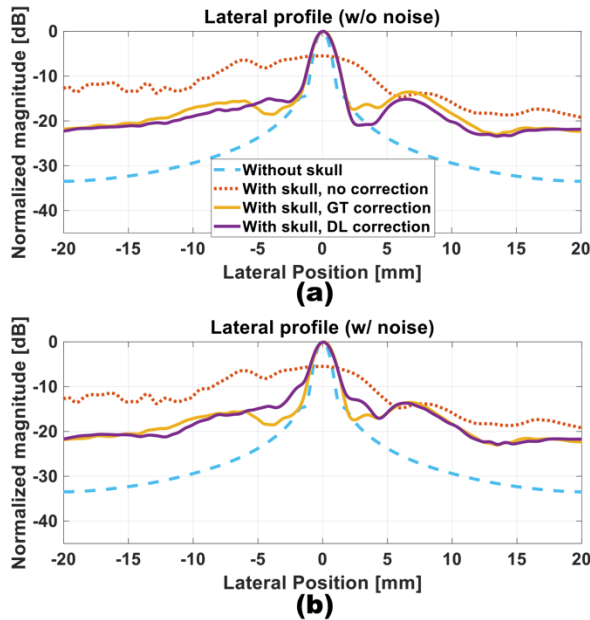


Fig. 9. Lateral profiles of beams at the focal depth of 30 mm without skull (blue dashed curves), with skull but no correction (red dotted curves), with skull and corrected by the ground truth SOS map (orange solid curves), and with skull and corrected by the deep-learning-estimated SOS map (purple solid curves) obtained when (a) no noise was present in RF signals, and (b) noise was present in RF signals.

learning-estimated skull thickness and SOS values (Fig. 8d). The focused beams were characterized using root mean square pressure values. The target focusing depth was 30 mm in this example. This focusing depth was accurately achieved when the skull was absent (Fig. 8a). When the skull was present as shown in Fig. 8b, the focal depth was shifted up and the quality of the focus was severely degraded without phase aberration correction, highlighting the severe distortion to the beam caused by skull phase aberration. These aberration effects were largely eliminated when phase aberration correction was performed (Fig. 8c-d). The performance was visually similar between correction using ground-truth thickness and SOS and correction using deep-learning-estimated thickness and SOS (Fig. 8c-d). The full widths at half maximum (FWHM) along the lateral direction at target focal depth of 30 mm were 1.50 mm, 2.08 mm, and 2.09 mm, for focused beams generated without the skull, with the skull and corrected using the ground truth SOS map, and with the skull and corrected by deep-learning-estimated thickness and SOS, respectively, demonstrating the effectiveness of phase aberration correction, although perfect phase aberration correction was not achieved (e.g., slightly wider beam after correction than the focused beam without the skull, as well as mild but noticeable side lobes in the beams after correction [Fig. 9a]).

Similar to the noise-free cases, phase aberration correction improved the quality of transcranial focusing with the presence of noise. The lateral profile of the focused beam with the skull and corrected by deep-learning-estimated thickness and SOS had an FWHM of 2.16 mm at focal depth, close to the profile corrected using the ground-truth SOS map (Fig. 9b).

C. Transcranial Ultrasound Imaging

Phase aberration correction using the reconstructed SOS

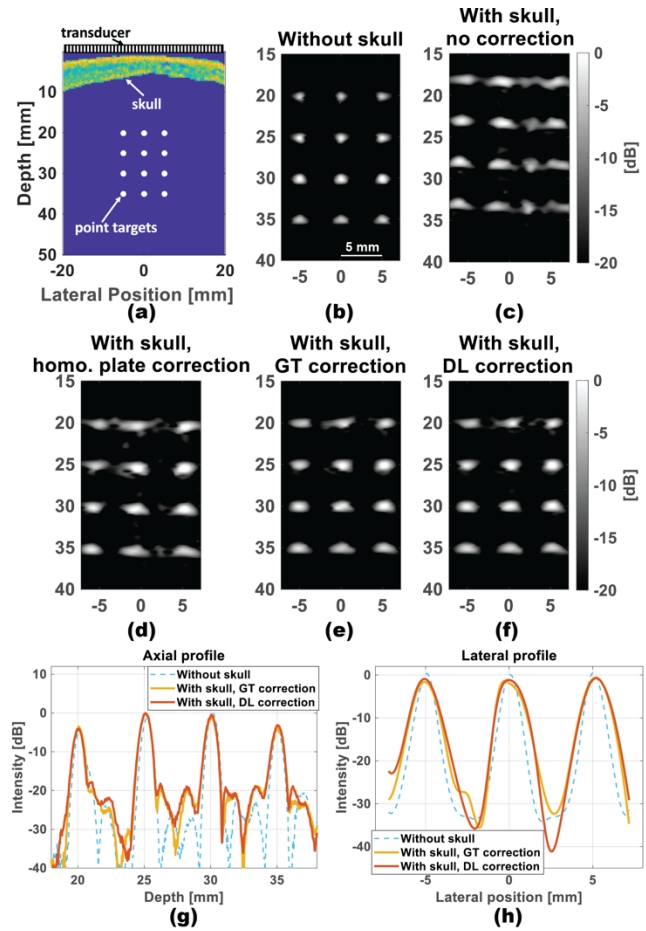


Fig. 10. (a) Schematic illustration of imaging simulations. B-mode images of twelve point targets (b) without skull, (c) with skull but no correction, (d) with skull and corrected using the SOS map of a rectangular homogeneous skull with mean ground-truth thickness and SOS, (e) with skull and corrected using the ground truth SOS map, and (f) with skull and corrected using the deep-learning-estimated SOS map. (g) Axial profiles of the four targets on the central column. (h) Lateral profiles of the three targets on the third row from top. All were obtained from noise-free RF signals.

maps and FMM significantly improved the transcranial ultrasound imaging. A representative example when no noise was present in RF signals is shown in Fig. 10, comparing the B-mode images of twelve point targets under various conditions, including imaging without the skull (Fig. 10b), imaging through a skull segment without aberration correction (Fig. 10c), imaging through the same skull segment with aberration correction using a homogenous rectangular plate with mean ground-truth skull thickness (6.3 mm) and SOS values (2341 m/s) (Fig. 10d, hereinafter referred to as homogenous plate correction), with aberration correction using the ground-truth skull SOS map (Fig. 10e), and with aberration correction using the SOS map reconstructed by deep-learning-estimated thickness and SOS values (Fig. 10f). Targets at various depths and horizontal positions were accurately imaged without the skull (Fig. 10b). When the skull was present, images of targets deviated from their original positions and were significantly distorted (Fig. 10c). The homogenous plate correction corrected the depth shift of targets, but not the lateral shift or the degraded lateral resolution, indicating limited correction performance without knowledge of the SOS map for

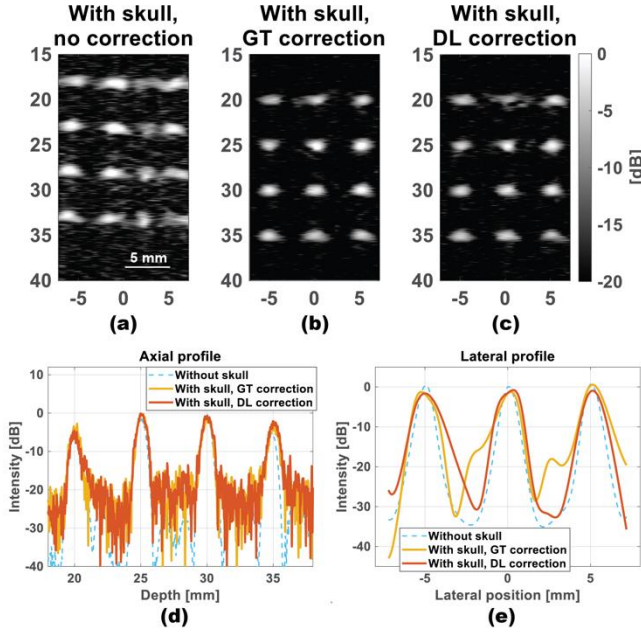


Fig. 11. B-mode images of twelve point targets constructed by RF signals with noise (a) with aberration but no correction, (b) with skull and corrected using the ground truth SOS map, and (c) with skull and corrected using the deep-learning-estimated SOS map. (d) Axial profiles of the four targets on the central column. (e) Lateral profiles of the three targets on the third row from top.

the specific skull segment. The correction using the ground-truth SOS map and the SOS map reconstructed by deep-learning-estimated thickness and SOS values largely eliminated aberration effects (Fig. 10 e-f), with similar performances, especially for the axial profile of point targets in the central column (Fig. 10g) and lateral profile of point targets at the focal depth (Fig. 10h).

The quality of transcranial imaging in the presence of noise was also improved by our approach. B-mode images through the skull segment without aberration correction (Fig. 11a), corrected using the ground-truth SOS map (Fig. 11b), and corrected using deep-learning (Fig. 11c) were constructed with RF signals with noise. Twelve targets were clearly visible in the B-mode image with the deep-learning correction. Two more examples of phase aberration corrected images are shown in Supplementary Materials (Fig. S2 and S4).

IV. DISCUSSION

We developed a deep learning method to estimate the skull thickness and SOS using RF signals backscattered from the skull. The deep learning method accurately yielded the skull thickness and SOS estimates. The deep-learning-estimated skull thickness and SOS values were shown to be effective for phase aberration correction, as demonstrated through two potential use case scenarios, transcranial focusing and transcranial imaging. Phase aberration correction in transcranial focusing is important for a wide range of focused ultrasound applications, such as ultrasound neuromodulation [45], sonogenetics [46], sonothermogenetics [47], and transcranial high-intensity focused ultrasound (HIFU) ablation [48]. Phase aberration correction in transcranial imaging is critical for applications such as brain anatomical imaging, brain

vascular imaging, functional ultrasound, Doppler imaging, and photoacoustic tomography.

Several studies measured skull thickness and SOS using pulse-echo ultrasound. Wydra et al. [23] used a variable focus technique to measure the thickness of a flat thin phantom with an array probe at 2.25 MHz, achieving an error rate of 4.9%. Due to their singular ground truth value, MAE comparison is not directly applicable. Minh et al. [49] used a similar refraction-corrected multi-focus imaging method and performed measurements on plate-shaped bovine bone samples at a higher frequency (5 MHz). They had a root-mean-square error (RMSE) of 0.94% for compressional sound velocity and 1.09% for thickness. For comparison, we recalculated our own RMSEs, which yielded 0.01% for SOS and 0.09% for thickness estimations. Wang and Jing [27] performed a numerical study using Wydra's approach to estimate the thickness and SOS of realistic skull models generated from CT scans of ex vivo human skulls. They assumed the skull to an SOS-homogeneous medium and had an error of 2% in estimating the average SOS with no noise was present in RF signals. They also measured the varying thickness of a skull segment, resulting in a maximum error of 10% for thickness estimation with an average thickness of 4.5 mm. Utilizing the same ultrasound frequency of 1.5 MHz, our approach showed lower MAEs. Nonetheless, it is noteworthy that outliers exist in our thickness estimation, with a maximum relative error of 26.1% when the ground truth thickness is 12.28 mm. This error can be attributed to the ground truth thickness being much thicker than our focal depth of 5 mm.

We also demonstrated the capability of our method to correct phase aberration caused by the skull in ultrasound focusing and imaging. However, it was observed that the corrected beams always had a wider beamwidth than the beam without skull. This phenomenon might be caused by the high attenuation and the power law frequency-dependence of the attenuation in the skull. With a power law exponent of 2, the high-frequency components in transmitted waves were more strongly attenuated, which led to a wider beamwidth and introduced deteriorated lateral resolution and reduced the SNR. To address these issues, signal intensity enhancement methods such as coded excitation [50] can be incorporated with our proposed phase aberration correction method to further recover the focusing quality. Furthermore, the amplitude inverting modulation method can be used by assuming the skull as an attenuating layer close to the transducer array [51]. Since the skull attenuation is frequency-dependent, corresponding amplitude correction could also be frequency specific (e.g., frequency modulated coded excitation [52]), and low frequency transmission can be employed to reduce the attenuation ratio.

Our DL-based approach exhibits sensitivity to noise. The MAE for thickness estimation increased from 0.15 mm (noise free) to 0.35 mm with 10 dB SNR, while the MAE for SOS estimation increased from 13 m/s (noise free) to 39 m/s with 10 dB SNR. However, the performance of phase aberration correction in transcranial focusing and transcranial imaging appeared to be robust under the impact of noise.

The proposed method could be practical for clinical applications due to several advantages. First, it is relatively easy to implement because it utilizes the same ultrasound array

for both transcranial focusing/imaging and acquiring data for skull thickness and SOS estimation. This eliminates the need for additional CT scans of the skull. Second, our method is not location dependent and can be applied to various skull locations, not limited to the thin temporal window. Simulations performed on 3,595 locations of 5 ex vivo human skulls with ground-truth thickness ranging from 2.8 to 11.4 mm demonstrated that our DL-based method provided accurate estimation for a wide range of skull profiles. Third, our neural network does not require large-scale data for training, as demonstrated by the sufficient RF data from 5 ex vivo skulls used in this study.

This study has several limitations:

- 1) The numerical skull models were derived from diagnostic CT images at a resolution (0.39 mm/pixel) that required image interpolation prior to acoustic simulations. Future studies could adopt higher-resolution CT images (e.g., micro CT at 0.05 mm/pixel) to reveal finer structure of the skull in order to more accurately model the scattering occurring in the skull. Also, this study did not account for the elevation interpolation of the CT images, which could be investigated in subsequent work.
- 2) The method generates quasi-homogeneous SOS maps rather than providing voxel-level granularity. Taking into account the axial inhomogeneity of the SOS maps may further improve the effectiveness of phase aberration correction.
- 3) Two-dimensional simulations were performed in this study for computational cost considerations. Two-dimensional simulations are intrinsically less realistic than 3-D simulations due to the lack of consideration of out-of-plane refraction.
- 4) As this study is a numerical-only feasibility proof-of-concept study, experimental validation is needed to further evaluate the potential of this method.

V. CONCLUSION

This study demonstrated the feasibility of the deep-learning-based pulse-echo ultrasound method for transcranial phase aberration correction. The aberration correction using deep-learning-estimated skull thickness and SOS values was shown to significantly improve the quality of transcranial focusing and imaging.

ACKNOWLEDGMENT

The authors wish to express their grateful appreciation to Drs. Volodymyr Kindratenko, Dawei Mu, and Alexander Schwing for their helpful discussions, and to Mr. Churan He for his assistance with the work.

REFERENCES

- [1] E. Macé, G. Montaldo, I. Cohen, M. Baulac, M. Fink, and M. Tanter, "Functional ultrasound imaging of the brain," *Nature Methods*, vol. 8, no. 8, pp. 662-664, 2011/08/01 2011, doi: 10.1038/nmeth.1641.
- [2] E. Tiran *et al.*, "Transcranial Functional Ultrasound Imaging in Freely Moving Awake Mice and Anesthetized Young Rats without Contrast Agent," *Ultrasound in Medicine & Biology*, vol. 43, no. 8, pp. 1679-1689, 2017/08/01/ 2017, doi: <https://doi.org/10.1016/j.ultrasmedbio.2017.03.011>.
- [3] T. Hölscher *et al.*, "Transcranial Ultrasound from Diagnosis to Early Stroke Treatment," *Cerebrovascular Diseases*, vol. 26, no. 6, pp. 659-663, 2008, doi: 10.1159/000166844.
- [4] C.-H. Fan, H.-L. Liu, C.-Y. Huang, Y.-J. Ma, T.-C. Yen, and C.-K. Yeh, "Detection of Intracerebral Hemorrhage and Transient Blood-Supply Shortage in Focused-Ultrasound-Induced Blood-Brain Barrier Disruption by Ultrasound Imaging," *Ultrasound in Medicine & Biology*, vol. 38, no. 8, pp. 1372-1382, 2012/08/01/ 2012, doi: <https://doi.org/10.1016/j.ultrasmedbio.2012.03.013>.
- [5] M. Wiesmann and G. Seidel, "Ultrasound Perfusion Imaging of the Human Brain," *Stroke*, vol. 31, no. 10, pp. 2421-2425, 2000, doi: 10.1161/01.STR.31.10.2421.
- [6] M. A. Lediju Bell, A. K. Ostrowski, K. Li, P. Kazanzides, and E. M. Boctor, "Localization of Transcranial Targets for Photoacoustic-Guided Endonasal Surgeries," *Photoacoustics*, vol. 3, no. 2, pp. 78-87, 2015/06/01/ 2015, doi: <https://doi.org/10.1016/j.pacs.2015.05.002>.
- [7] F. J. Fry and J. E. Barger, "Acoustical properties of the human skull," *The Journal of the Acoustical Society of America*, vol. 63, no. 5, pp. 1576-1590, 1978, doi: 10.1121/1.381852.
- [8] T. S. Riis, T. D. Webb, and J. Kubanek, "Acoustic properties across the human skull," *Ultrasonics*, vol. 119, p. 106591, 2022/02/01/ 2022, doi: <https://doi.org/10.1016/j.ultras.2021.106591>.
- [9] G. Pinton, J. F. Aubry, E. Bossy, M. Muller, M. Pernot, and M. Tanter, "Attenuation, scattering, and absorption of ultrasound in the skull bone," *Medical physics*, vol. 39, no. 1, pp. 299-307, 2012.
- [10] M. Fink, "Time reversal of ultrasonic fields. I. Basic principles," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 39, no. 5, pp. 555-566, 1992, doi: 10.1109/58.156174.
- [11] J. L. Thomas and M. A. Fink, "Ultrasonic beam focusing through tissue inhomogeneities with a time reversal mirror: application to transskull therapy," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 43, no. 6, pp. 1122-1129, 1996, doi: 10.1109/58.542055.
- [12] M. Fink, G. Montaldo, and M. Tanter, "Time-reversal acoustics in biomedical engineering," *Annual review of biomedical engineering*, vol. 5, no. 1, pp. 465-497, 2003.
- [13] F. Vignon, J. Aubry, M. Tanter, A. Margoum, and M. Fink, "Adaptive focusing for transcranial ultrasound imaging using dual arrays," *The Journal of the Acoustical Society of America*, vol. 120, no. 5, pp. 2737-2745, 2006.

- [14] N. M. Ivancevich, J. J. Dahl, G. E. Trahey, and S. W. Smith, "Phase-aberration correction with a 3-D ultrasound scanner: Feasibility study," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 53, no. 8, pp. 1432-1439, 2006.
- [15] S. Flax and M. O'Donnell, "Phase-aberration correction using signals from point reflectors and diffuse scatterers: Basic principles," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 35, no. 6, pp. 758-767, 1988.
- [16] G. C. Ng, S. S. Worrell, P. D. Freiburger, and G. E. Trahey, "A comparative evaluation of several algorithms for phase aberration correction," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 41, no. 5, pp. 631-643, 1994.
- [17] D. L. Liu and R. C. Waag, "Correction of ultrasonic wavefront distortion using backpropagation and a reference waveform method for time-shift compensation," *The Journal of the Acoustical Society of America*, vol. 96, no. 2, pp. 649-660, 1994, doi: 10.1121/1.410304.
- [18] D. L. Liu and R. C. Waag, "Time-shift compensation of ultrasonic pulse focus degradation using least-mean-square error estimates of arrival time," *The Journal of the Acoustical Society of America*, vol. 95, no. 1, pp. 542-555, 1994, doi: 10.1121/1.408348.
- [19] G. T. Clement, P. J. White, and K. Hynynen, "Enhanced ultrasound transmission through the human skull using shear mode conversion," *The Journal of the Acoustical Society of America*, vol. 115, no. 3, pp. 1356-1364, 2004.
- [20] P. J. White, G. T. Clement, and K. Hynynen, "Longitudinal and shear mode ultrasound propagation in human skull bone," *Ultrasound in medicine & biology*, vol. 32, no. 7, pp. 1085-1096, 2006.
- [21] F. Marquet *et al.*, "Non-invasive transcranial ultrasound therapy based on a 3D CT scan: protocol validation and in vitro results," *Physics in Medicine & Biology*, vol. 54, no. 9, p. 2597, 2009.
- [22] S. Almquist, D. L. Parker, and D. A. Christensen, "Rapid full-wave phase aberration correction method for transcranial high-intensity focused ultrasound therapies," *Journal of therapeutic ultrasound*, vol. 4, no. 1, pp. 1-11, 2016.
- [23] A. Wydra, E. Malyarenko, K. Shapoori, and R. G. Maev, "Development of a practical ultrasonic approach for simultaneous measurement of the thickness and the sound speed in human skull bones: a laboratory phantom study," *Physics in Medicine & Biology*, vol. 58, no. 4, p. 1083, 2013.
- [24] M. Mozaffarzadeh, E. Verschuur, M. D. Verweij, V. Daeichin, N. De Jong, and G. Renaud, "Refraction-Corrected Transcranial Ultrasound Imaging Through the Human Temporal Window Using a Single Probe," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 4, pp. 1191-1203, 2022.
- [25] A. Han *et al.*, "Noninvasive diagnosis of nonalcoholic fatty liver disease and quantification of liver fat with radiofrequency ultrasound data using one-dimensional convolutional neural networks," *Radiology*, vol. 295, no. 2, p. 342, 2020.
- [26] M. Feigin, D. Freedman, and B. W. Anthony, "A deep learning framework for single-sided sound speed inversion in medical ultrasound," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 4, pp. 1142-1151, 2019.
- [27] T. Wang and Y. Jing, "Transcranial ultrasound imaging with speed of sound-based phase correction: a numerical study," *Physics in Medicine & Biology*, vol. 58, no. 19, p. 6663, 2013.
- [28] J. A. Jensen, S. I. Nikolov, K. L. Gammelmark, and M. H. Pedersen, "Synthetic aperture ultrasound imaging," *Ultrasonics*, vol. 44, pp. e5-e15, 2006.
- [29] B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of biomedical optics*, vol. 15, no. 2, p. 021314, 2010.
- [30] J. L. Robertson, B. T. Cox, J. Jaros, and B. E. Treeby, "Accurate simulation of transcranial ultrasound propagation for ultrasonic neuromodulation and stimulation," *The Journal of the Acoustical Society of America*, vol. 141, no. 3, pp. 1726-1738, 2017.
- [31] M. Hayner and K. Hynynen, "Numerical analysis of ultrasonic transmission and absorption of oblique plane waves through the human skull," *The Journal of the Acoustical Society of America*, vol. 110, no. 6, pp. 3319-3330, 2001.
- [32] G. T. Clement and K. Hynynen, "A non-invasive method for focusing ultrasound through the human skull," *Physics in Medicine & Biology*, vol. 47, no. 8, p. 1219, 2002.
- [33] J.-F. Aubry, M. Tanter, M. Pernot, J.-L. Thomas, and M. Fink, "Experimental demonstration of noninvasive transskull adaptive focusing based on prior computed tomography scans," *The Journal of the Acoustical Society of America*, vol. 113, no. 1, pp. 84-93, 2003.
- [34] Y. Jing, F. C. Meral, and G. T. Clement, "Time-reversal transcranial ultrasound beam focusing using a k-space method," *Physics in Medicine & Biology*, vol. 57, no. 4, p. 901, 2012.
- [35] H. Montanaro *et al.*, "The impact of CT image parameters and skull heterogeneity modeling on the accuracy of transcranial focused ultrasound simulations," *Journal of Neural Engineering*, vol. 18, no. 4, p. 046041, 2021.
- [36] S. Pichardo, V. W. Sin, and K. Hynynen, "Multi-frequency characterization of the speed of sound and attenuation coefficient for longitudinal transmission of freshly excised human skulls," *Physics in Medicine & Biology*, vol. 56, no. 1, p. 219, 2010.
- [37] V. Perrot, M. Polichetti, F. Varray, and D. Garcia, "So you think you can DAS? A viewpoint on delay-and-sum beamforming," *Ultrasonics*, vol. 111, p. 106309, 2021.
- [38] K. Chopra *et al.*, "A Comprehensive Examination of Topographic Thickness of Skin in the Human Face,"

- Aesthetic Surgery Journal*, vol. 35, no. 8, pp. 1007-1013, 2015, doi: 10.1093/asj/sjv079.
- [39] S. Santurkar, D. Tsipras, A. Ilyas, and A. Madry, "How does batch normalization help optimization?," *Advances in neural information processing systems*, vol. 31, 2018.
- [40] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929-1958, 2014.
- [41] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise reduction in speech processing*: Springer, 2009, pp. 1-4.
- [42] R. G. Pontius, O. Thontteh, and H. Chen, "Components of information for multiple resolution comparison between maps that share a real variable," *Environmental and Ecological Statistics*, vol. 15, no. 2, pp. 111-142, 2008/06/01 2008, doi: 10.1007/s10651-007-0043-y.
- [43] J. A. Sethian, "Fast marching methods," *SIAM review*, vol. 41, no. 2, pp. 199-235, 1999.
- [44] *patricknraanes/FM: Version 1.0*. (2011). Accessed: July 5, 2022. [Online]. Available: <https://doi.org/10.5281/zenodo.2025811>
- [45] J. Blackmore, S. Shrivastava, J. Sallet, C. R. Butler, and R. O. Cleveland, "Ultrasound neuromodulation: a review of results, mechanisms and safety," *Ultrasound in medicine & biology*, vol. 45, no. 7, pp. 1509-1536, 2019.
- [46] S. Ibsen, A. Tong, C. Schutt, S. Esener, and S. H. Chalasani, "Sonogenetics is a non-invasive approach to activating neurons in *Caenorhabditis elegans*," *Nature communications*, vol. 6, no. 1, p. 8264, 2015.
- [47] Y. Yang *et al.*, "Sonothermogenetics for noninvasive and cell-type specific deep brain neuromodulation," *Brain stimulation*, vol. 14, no. 4, pp. 790-800, 2021.
- [48] J. W. Jenne, "Non-invasive transcranial brain ablation with high-intensity focused ultrasound," *Translational Neurosurgery*, vol. 36, pp. 94-105, 2015.
- [49] H. N. Minh, J. Du, and K. Raum, "Estimation of thickness and speed of sound in cortical bone using multifocus pulse-echo ultrasound," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 67, no. 3, pp. 568-579, 2019.
- [50] E. P. Vienneau and B. C. Byram, "A Coded Excitation Framework for High SNR Transcranial Ultrasound Imaging," *IEEE Transactions on Medical Imaging*, 2023.
- [51] J.-L. Thomas and M. A. Fink, "Ultrasonic beam focusing through tissue inhomogeneities with a time reversal mirror: application to transskull therapy," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 43, no. 6, pp. 1122-1129, 1996.
- [52] A. Yousefi, D. E. Goertz, and K. Hynynen, "Transcranial shear-mode ultrasound: assessment of imaging performance and excitation techniques," *IEEE transactions on medical imaging*, vol. 28, no. 5, pp. 763-774, 2009.