Acoustic simulation for transcranial focused ultrasound using GAN-based synthetic CT

Heekyung Koh, Tae Young Park, Yong An Chung, Jong-Hwan Lee* and Hyungmin Kim*

***Abstract*—Transcranial focused ultrasound (tFUS) is a promising non-invasive technique for treating neurological and psychiatric disorders. One of the challenges for tFUS is the disruption of wave propagation through the skull. Consequently, despite the risks associated with exposure to ionizing radiation, computed tomography (CT) is required to estimate the acoustic transmission through the skull. This study aims to generate synthetic CT (sCT) from T1-weighted magnetic resonance imaging (MRI) and investigate its applicability to tFUS acoustic simulation. We trained a 3D conditional generative adversarial network (3D-cGAN) with 15 subjects. We then assessed image quality with 15 test subjects: mean absolute error (MAE) = 85.72± 9.50 HU (head) and 280.25±24.02 HU (skull), dice coefficient similarity (DSC) = 0.88±0.02 (skull). In terms of skull density ratio (SDR) and skull thickness (ST), no significant difference was found between sCT and real CT (rCT). When the acoustic simulation results of rCT and sCT were compared, the intracranial peak acoustic pressure ratio was found to be less than 4%, and the distance between focal points less than 1 mm.**

Index Terms— Transcranial focused ultrasound, acoustic simulation, single-element transducer, synthetic CT, MRI-only, generative adversarial network, conditional GAN

I. Introduction

Transcranial focused ultrasound (tFUS), in which acoustic energy is delivered to a small and localized area in the brain, has been widely applied as a non-invasive therapeutic tool. The tissue ablation using the thermal and mechanical energy generated from the acoustic focus is one of the basic applications of tFUS. In addition, a number of previous studies have also shown that brain stimulation using the low-intensity tFUS can evoke neuromodulatory effects in healthy humans without any thermal or non-thermal damages on the brain tissue [1]–[3].

To attain more accurate targeting and higher spatial resolution of acoustic focus, the multi-array transducer system which can implement the phase correction scheme was preferred in the thermal ablation field. This equipment is, however, generally expensive and complex. As an alternative to a multi-array system, a single-element transducer has been widely used because of its cost efficiency and lower complexity. Although lacking an aberration correction mechanism, this particular type of transducer successfully modulated the neuronal activity in the brain [4]–[6].

Regardless of the type of transducer, the biggest challenge to all field of tFUS application is the presence of the skull, which can shift or distort acoustic focus in the brain due to the large mismatch of acoustic properties between different media (i.e., speed of sound, density, and attenuation coefficient) [7]. Therefore, computed tomography (CT) is adopted to precisely estimate acoustic wave propagation through the skull, using the Hounsfield unit (HU) of the CT scan to convert to the acoustic properties of the skull [8]–[10]. However, exposure to ionizing radiation during CT scanning may increase patients' lifetime risk of developing cancer [11], [12].

As an alternative to CT, efforts have been made to extract skull information from magnetic resonance imaging (MRI). Previous studies suggested that ultrashort echo time (UTE) and zero echo time (ZTE) MRI sequences can contrast the bone against the soft-tissue [10], [13]. Deep learning-based methods have also been proposed for synthesizing CT from MRI, such as by using a convolutional neural network (CNN) or a generative adversarial network (GAN). Han et al. proposed the CNN-based model for generating synthetic CT (sCT) from T1-weighted (T1w) MRI [14]. Emami et al. demonstrated a conditional-GAN (cGAN)-based model featuring less noise during detailed reconstruction at bone/air interfaces than the CNN-based method [15]. CycleGAN was also adopted by some studies and showed better performance in terms of misalignment errors [16], [17].

There have been studies on multi-array tFUS treatment planning using UTE-MRI data, which focus on thermal ablation. Guo et al. showed acoustic and thermal distributions in the brain using a UTE-based skull structure [10]. Su et al. also indicated the feasibility of UTE-based sCT for tFUS treatment planning [18]. Since these studies focused on the tFUS application of

This research was supported by the Technology Innovation Program (20012396) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea) and the National Research Council of Science & Technology (NST) grant by the Korea government (MSIT) (No. CAP-18-01-KIST).

H. Koh (heekko@kist.re.kr), T.Y. Park (pty0220@kist.re.kr) and H. Kim (hk@kist.re.kr) are with the Bionics Research Center, Biomedical Research Division, Korea Institute of Science and Technology, Seoul 02792, Republic of Korea.

Y.A. Chung (yongan@catholic.ac.kr) is with Department of Radiology and Nuclear Medicine, Incheon St. Mary's Hospital, The Catholic University of Korea, Seoul 21431, Republic of Korea.

T.Y. Park and H. Kim are also with the Division of Bio-Medical Science & Technology, KIST School, Korea University of Science and Technology, Seoul 02792, Republic of Korea.

H. Koh and J.-H. Lee (jonghwan_lee@korea.ac.kr) are also with the Department of Brain and Cognitive Engineering, Korea University, Seoul 136-713, Republic of Korea

H. Koh and T.Y. Park contributed equally to this work.

thermal ablation, the peak acoustic intensity and temperature change between sCT and rCT were mainly considered, rather than the location and shape of the acoustic focus.

Unlike the previous studies, this study aims to generate sCT from T1w MRI and investigate its feasibility in acoustic simulation for tFUS neuromodulation with a single-element transducer. We suggested a 3D-cGAN featuring residual blocks and a combination of multiple loss functions to enhance the quality of sCT. In addition to evaluating the image quality and associated skull characteristic measurements of the proposed method, acoustic simulation results of sCT were compared to those of real CT (rCT) under the same sonication configuration. We have validated the results with three brain regions, including the primary motor cortex (M1), primary visual cortex (V1), and dorsal anterior cingulate cortex (dACC), which have shown the neuromodulatory effects by tFUS stimulation [19]–[21].

II. METHODS

A. Data acquisition

We used the brain CT and MRI data acquired from 33 subjects. The CT images were scanned on a Toshiba scanner (Aquilion ONE, Otawara, Japan) with 120 kVp, with a range of 0.4-0.53 × 0.4-0.53 × 0.3-1.0mm^3 resolutions, in a 512 × 512 in-plane matrix. The MRI data were acquired using a 3-T Siemens scanner (Skyra, Erlangen, Germany) with a T1-weighted Magnetization Preparation Rapid Acquisition Gradient Echo (MPRAGE) pulse sequence (2.46 ms echo time [TE], 1900 ms repetition time [TR], 900 ms inversion time [TI], 9° flip angle, 0.94 × 0.94 × 0.94 mm^3 resolution, and 256 × 256 in-plane matrix). Typically, MPRAGE sequences consist of a non-selective (180º) inversion pulse with TI of 800-1200 ms, which is followed by rapidly acquired gradient echoes at short TE of 2-4 ms and flip angles of 8-12º [22]. The TR is relatively long, approximately 2000 ms. We used the imaging parameters of the MPRAGE-based T1w MRI that have been utilized for the tFUS neuromodulation studies [23], [24]. The Institutional Review Board (IRB) of Incheon St. Mary's Hospital, South Korea, approved the entire study protocol.

B. Image preprocessing

N4 bias field correction was applied to the MRI data to correct intensity inhomogeneity, and rigid registration was performed to align the corresponding CT to the MRI, both performed by the 'N4ITKbiasfieldcorrection' and 'Elastix' modules of 3D Slicer (version 4.11.0) [25]–[27]. The MRI and CT data were resampled and padded with air values (0 for MRI; -1000 HU for CT) to a 256 × 256 × 256 volume with a spacing of 0.94 mm. The image backgrounds were removed with a binary mask of each subject's brain by selecting the largest continuous voxel region above a specific threshold value (100) of the MRI data [15], [28]. Holes in the masks were filled by a binary fill hole operation with dilation and erosion using the 'ndimage' module of SciPy [29]. These procedures ensured that the region of interest was not excluded from the registered CT. The voxel regions of MRI and CT outside the mask were respectively filled with air values. The intensities of MRI were clipped to a range of [0, 2500] which is the [0, 99.5th] percentile intensity value of the bias field corrected MRI [30]. After hyperparameter tuning (details in II.C.3), we investigated the effect of MRI intensity normalization for sCT generation by applying two different approaches (i.e., Z-score, and Piecewise linear Histogram matching [HM]) and compared the performance with raw MRI. The Z-score method normalizes the MRI using the mean (μ) and the standard deviation (σ) of MRI intensities within the head : $\mathrm{MRI}_{\mathrm{Z\_score}} = \frac{\mathrm{MRI} - \mu}{\sigma}$. This method was applied to each of MRI volumes. The HM normalization linearly maps the intensity of each MRI to align with the standard histogram, which was learned from a set of reference MRI volumes [31]. The mean intensities histogram at the landmarks were defined as the standard histogram. The locations of landmarks that were determined within [1st, 99th] percentiles with a step of 10 (i.e., [1st, 10th], [10th, 20th], …, [90th, 99th]) were used [32]. The standard histogram was acquired from MRIs in our training set and applied to all datasets, including training, validation, and test set. The detailed methods of MRI normalizations are available in previous studies [31], [32] with code implementations [33]. Both the raw and normalized MRI were linearly scaled to [-1, 1] and used as input of the cGAN model. The HU intensities of CT were clipped to a range of [-1000, 3000] followed by a linear scaling to [-1, 1].

C. Network overview

In this study, we were motivated to use the conditional generative adversarial network (cGAN) to generate sCT [34]. The proposed network consists of a generator and a discriminator. The schematic diagram of the proposed network is illustrated in Fig. 1.

1) cGAN architecture

As shown in Fig. 1b, the generator is an encoder-decoder type network featuring nine residual blocks. The encoder section consists of three convolutional layers with down-sampling operations to reduce the size of the feature map. After the encoder, the nine continuous residual blocks are implemented to mitigate the level of the vanishing gradient. Each residual block is composed of two convolutional layers with a dropout layer to prevent overfitting (Fig. 1c). Consecutive residual blocks are concatenated. In the decoder section, two up-sampling layers with transposed convolutional layers are utilized to reconstruct the size of the feature map. Reflection paddings are applied in the middle of the layers to preserve the size of the output. All convolutional layers are followed by instance normalization and ReLU activation except for the last layers of the residual block and decoder. Lastly, tanh activation is employed to make the generator generalize images in the range of [-1, 1].

We adopted the discriminator network based on PatchGAN [34], consisting of five consecutive convolutional layers, which, except for the first and last layer, are followed by instance normalization and LeakyReLU (slope=0.2) (Fig. 1d). The node in the last layer with the sigmoid activation function determines the probabilities of rCT or sCT in the range of [0, 1].

2) Loss functions

The generator (G) and discriminator (D) are trained by optimizing parameters (θ_G and θ_D), which reflect the adversarial goals of each network:

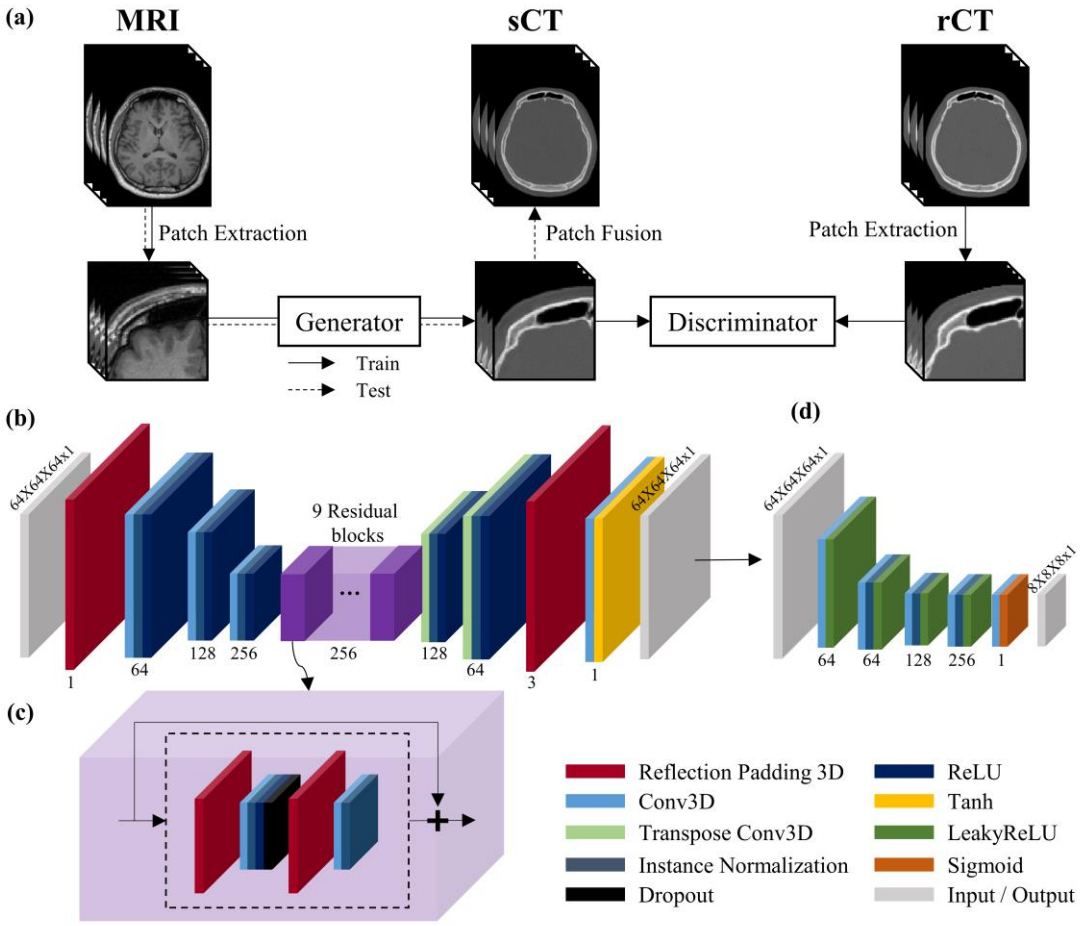

Fig. 1. a) Schematic diagram of training and test flow. b) The proposed generator. c) The residual blocks. d) The proposed discriminator. The number of filters are described below each layer in (b) and (d).

$$\theta_G = \arg\min_{\theta_G}(\lambda_1 L_{G,ls}(G,D) + \lambda_2 L_{mpd}(G) + \lambda_3 L_{gdl}(G)) \quad (1)$$

$$\theta_D = \arg\min_{\theta_D}(L_{D,ls}(G,D)) \quad (2)$$

where λ_1, λ_2, and λ_3 indicate the weight factors of each loss function. The loss functions of the generator are composed of three losses: least-squares loss (LS), mean p distance loss (MPD), and image gradient difference loss (GDL).

LS is known to enhance the stability of GAN [35], and is defined as:

$$L_{G,ls} = (D(G(X)) - 1)^2 \quad (3)$$

where X is the source MRI, and $G(X)$ is the sCT. $D(G(X))$ is the output of discriminator which indicates judgment of sCT. This enhances the misjudgment of the discriminator by minimizing the difference between $D(G(X))$ and 1, where 1 indicates a label for rCT.

MPD is known to overcome the blurry image and misclassification issues that occur when using traditional distance loss functions [16] (p = 1.5):

$$L_{mpd} = \|G(X) - Y\|_p^p \quad (4)$$

where Y is the rCT of the corresponding MRI.

GDL is known to help generate more realistic and sharp sCT by minimizing differences between the gradients [16], [36]:

$$L_{gdl} = ||\nabla G(X)_x| - |\nabla Y_x||^2 + \left||\nabla G(X)_y| - |\nabla Y_y|\right|^2 + ||\nabla G(X)_z| - |\nabla Y_z||^2 \quad (5)$$

These three losses were combined and minimized to optimize the generator.

The discriminator loss is defined as:

$$L_{D,ls} = (D(Y) - 1)^2 + (D(G(X)))^2 \quad (6)$$

where D(Y) and D(G(X)) are the decisions of the discriminator using rCT and sCT. The output of the discriminator is the probability to be rCT (1 being rCT and 0 being sCT). The discriminator is trained to predict the correct label (1 for rCT, 0 for sCT) by minimizing the LS cost functions.

3) Training

Paired CT and MRI data from 15 pseudo-randomly selected subjects were used to train the proposed model. In the training phase, data augmentation was performed at every epoch to provide diverse data to the generator and prevent overfitting problems. The applied factors are as follows: rotation in the range [-3.5°, 3.5°], shearing in the range [0.97, 1.03], zoom in the range [0.7, 1.3], and horizontal flip with a 0.5 probability [30], [37].The same factors were applied to the MRI and CT

data of each pair. Additionally, from each 256 × 256 × 256 MRI and CT volume, we extracted 3D patches of 64 × 64 × 64 voxels using a 16 × 16 × 16 stride and fed them to the generator to synthesize CT patches. We obtained 2,200 sets of the paired MRI and CT patches for each subject. Thus, the total number of paired patches used for training was 33,000 (2,200 × 15 subjects). We utilized 2,200 randomly selected paired patches of MRI and CT for each epoch with an additional augmentation scheme. The MRI patches were fed to the generator to synthesize the corresponding CT patches, and the paired patches of rCT and sCT were fed to the discriminator (Fig. 1a). We set a mini-batch size of 1 to maintain the size of patches to 64 × 64 × 64 rather than smaller patches and to prevent GPU memory crashes that can break down the process of cGAN training. The size of patches was small enough to fit on the physical memory of GPU but large enough to provide continuous anatomical structures. Using the 3D patches rather than a whole brain volume, we were able to increase the size of the training dataset. The initial learning rate of both generator and discriminator was 0.0002, and this linearly decayed after epoch 100 [30], [38]. The total number of epochs was 400. The Adam optimizer was used for optimization with $\beta_1 = 0.5$ and $\beta_2 = 0.9$. Five possible scenarios for weight factors of generator loss λ_1, λ_2 and λ_3 were considered to evaluate the performance depending on these hyperparameters (Table II): 1) MPD [λ_1=1, λ_2=40, λ_3=0], 2) GDL [λ_1=1, λ_2=0 , λ_3=40], 3) MPD = GDL [λ_1=1, λ_2=20 , λ_3=20], 4) MPD > GDL [λ_1=1, λ_2=30 , λ_3=10], and 5) MPD < GDL [λ_1=1, λ_2=10 , λ_3=30]. The λ_1 remained at 1 to set a baseline magnitude level for λ_2 and λ_3. Three subjects separated from the training (n = 15) and test (n = 15) subjects were used to estimate validation performance. The training and validation/test of the model were performed using our GPU server (i.e., Intel Xeon Gold 5120, 8 NVIDIA GeForce RTX 2080 Ti GPUs 11 GB; 512 GB RAM).

D. Target selection

Before the acoustic simulation, three brain regions (M1, V1, and dACC) were selected as potential targets for tFUS application. Due to limitations of focal length, deep brain areas (e.g., the thalamus) were excluded. In order to obtain the subject-specific coordinates of these target regions, the T1w template of ICBM152 [39] was co-registered to the T1w images of each subject utilizing the 'Elastix' module in 3D Slicer. The obtained transformation matrices were used to transform the target regions in MNI to subject-specific coordinates.

E. Acoustic simulation

We performed an acoustic simulation to assess the feasibility of sCT for tFUS application. The simulation was implemented in both rCT and sCT images under the same sonication conditions. A 200 kHz single-element bowl-shaped transducer (GPS 200, Ultran Group, USA) with a focal length of 55.2 mm was used. The 50 locations of the transducer were selected using the average reflection coefficient (ARC) method [9]. Since the result of acoustic propagation depends on spatial information of the transducer, various locations were considered as similarly in the realistic treatment planning procedure. The ARC is the mean ratio of the amplitude of the reflected acoustic wave to the incident acoustic wave at each transducer location. The lower ARCs indicate the more efficient locations that deliver acoustic energy to the brain target. By selecting transducer locations with the lowest 50th ARC among every possible location, we could avoid transducer locations that might induce total reflection of simulated acoustic waves caused by the skull and which make it difficult to compare intracranial acoustic fields using rCT and sCT. We performed 4,500 acoustic simulations in total (i.e., 50 transducer locations × 3 target regions × 15 subjects × 2 CT data).

TABLE I
ACOUSTIC PROPERTIES

Speed of sound (m/s)	Density (kg/m^3)	Attenuation coefficient (Np/MHz/m)
$c_{bone} = 3100$	$\rho_{bone} = 2200$	$\alpha_{min,skull} = 21.5$, $\alpha_{max,skull} = 208.9$
$c_{water} = 1482$	$\rho_{water} = 1000$	$\alpha_{water} = 0.0253$

The open-source k-Wave acoustics toolbox was used to obtain a time-independent solution to the wave equation [40]. The acoustic simulation domain was 160 × 160 × 160 grid points with a grid spacing of 1.02 mm, i.e., 7 points per acoustic wavelength (PPW) [7]. The simulation ended at 100 μs with a 0.1 Courant-Friedrichs-Lewy (CFL) number. In the acoustic simulation, the CT images were used to obtain the subject's skull structure. The CT images were cropped and resampled to the same size as the acoustic simulation domain. The acoustic properties (speed of sound, density, and attenuation coefficient) of the skull were calculated using HU to describe the porous and inhomogeneous characteristics of the skull. Based on the porosity of the skull (ψ) [10], the acoustic properties were obtained from (7) ~ (10). The acoustic properties of skull and water are summarized in Table I [7].

$$\psi = 1 - \frac{HU}{Max(HU)} \tag{7}$$

$$c_{skull} = c_{water}\psi + c_{bone}(1-\psi) \tag{8}$$

$$\rho_{skull} = \rho_{water}\psi + \rho_{bone}(1-\psi) \tag{9}$$

$$\alpha_{skull} = \alpha_{water}\psi + (\alpha_{max,skull} - \alpha_{min,skull})\psi^{0.5} \tag{10}$$

F. Evaluation

The paired CT and MRI data from the remaining 15 subjects were used as test data to evaluate the proposed method. The same preprocessing pipeline used for the training data was also applied to the paired CT and MRI test data. Since the input and output size of the generator is 64 × 64 × 64, the MRI patches of one subject were fed into the model to generate the sCT patches. The patches of sCT were fused by averaging overlapped voxel regions of patches using a 16 × 16 × 16 stride (Fig. 1a). The HU intensity values of the final sCT were rescaled from [-1, 1] to [-1000 HU, 3000 HU], the same range as the training CT images.

1) Image quality

The mean absolute error (MAE) represents the absolute HU intensity difference between rCT and sCT (11) where *n* is the number of voxels in regions of interest, and $rCT(i)$ and $sCT(i)$ represent the intensities at voxel index i.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|rCT(i) - sCT(i)| \quad (11)$$

$$DSC_{skull} = \frac{2|V_{rCT} \cap V_{sCT}|}{|V_{rCT}| + |V_{sCT}|} \quad (12)$$

The MAEs were calculated for the whole head (MAE_{head}) and skull (MAE_{skull}) region. The regions were defined as the intersected voxels with binary masks of rCT and sCT meeting specific thresholds, i.e., ≥–100HU and ≥250HU, respectively. The MAE_{head} was evaluated within only the head contour, in which the corresponding mask was defined to include air, soft and bony tissue by binary fill hole operation.

The dice similarity coefficient of the skull (DSC_{skull}) was also calculated to describe the overlapping of the skull region (12), where $|V_{rCT}|$ and $|V_{sCT}|$ represent the number of voxels in the skull area of each CT.

2) *Skull characteristics*

The skull density ratio (SDR) and the skull thickness (ST) were used to compare skull characteristics between rCT and sCT. These are known to be relevant factors for assessing acoustic transmission efficiency through the skull in clinical tFUS applications [41], [42]. The SDR is the mean ratio of the minimum to the maximum HU intensity along the skull paths of ultrasound rays. The ST is the distance between the first and last skull intersection coordinates of the ray paths. Both factors were calculated using the open-source software Kranion [43]. The SDR and ST were calculated for each of the three target regions (i.e., M1, V1, and dACC). Two-tailed paired *t*-test and Pearson's correlation analysis were performed between the skull characteristics of rCT and sCT.

3) *Acoustic simulation*

We compared the acoustic simulation results from rCT and sCT, and three physical quantities were obtained: 1) intracranial peak pressure ratio (Δ_{PR}), 2) dice similarity coefficient of the acoustic focuses ($DSC_{A90\%}$), 3) distance between the acoustic focuses (Δ_F) from (13) ~ (15).

$$\Delta_{PR} = \frac{|P - \hat{P}|}{P} \quad (13)$$

$$DSC_{A90\%} = \frac{2|A_{90\%} \cap \hat{A}_{90\%}|}{|A_{90\%}| + |\hat{A}_{90\%}|} \quad (14)$$

$$\Delta_F = \sqrt{(C_x - \hat{C}_x)^2 + (C_y - \hat{C}_y)^2 + (C_z - \hat{C}_z)^2} \quad (15)$$

The peak pressure in the brain was indicated as P. We defined the 3D region of acoustic focus as 90%-maximum of intracranial acoustic pressure (denoted as $A_{90\%}$ in Fig. 7) [8]. The peak intracranial pressures of the rCT and sCT were calculated separately to define acoustic focus respectively. The centroid of the $A_{90\%}$ region was used as the location of the acoustic focus, and labelled C. All values obtained from the simulation result using sCT are denoted by a caret (i.e., $\hat{P}$, $\hat{A}_{90\%}$, and $\hat{C}$).

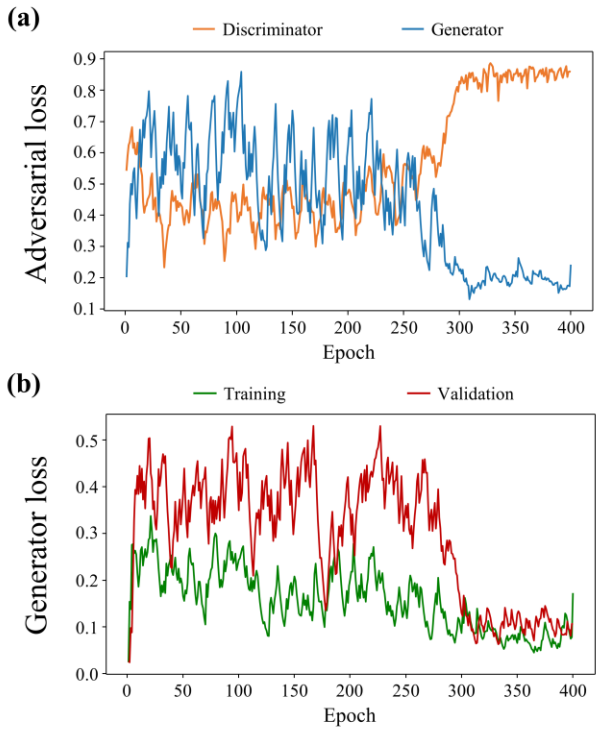

Fig. 2. Loss curves during the cGAN model training using Z-scored T1w MRI. (a) LS losses of generator and discriminator. (b) Training and validation loss of the generator. All three losses in the generator as represented in Eqs. (3) to (5) were combined in (b).

III. RESULTS

A. *Image quality*

The image qualities of the following models were reported from a model of best validation performance among all epochs. The example of training and validation losses are depicted in Fig. 2. We conjectured that the total epoch of 400 is appropriate as the oscillations in LS losses of generator and discriminator diminished, and the validation loss of generator converged to the level of the training loss after 300. The image qualities of test subjects depending on weight factors of the generator are summarized in Table II. It is notable that performances are better or at least similar overall when the combination of three losses was applied. The model trained using three losses with

TABLE II
IMAGE QUALITIES ON WEIGHT FACTORS

Model	MAE_{head}	MAE_{skull}	DSC_{skull}
MPD [λ_1=1, λ_2=40 , λ_3=0]	98.62±15.89	303.27±40.86	0.86±0.04
GDL [λ_1=1, λ_2=0 , λ_3=30]	108.14±11.63	310.56±30.75	0.86±0.02
MPD = GDL [λ_1=1, λ_2=20 , λ_3=20]	94.95±13.5	299.95±41.68	0.87±0.03
MPD > GDL [λ_1=1, λ_2=30 , λ_3=10]	93.80±13.40	299.26±41.20	0.87±0.03
MPD < GDL [λ_1=1, λ_2=10 , λ_3=30]	94.65±16.71	303.71±46.40	0.87±0.03

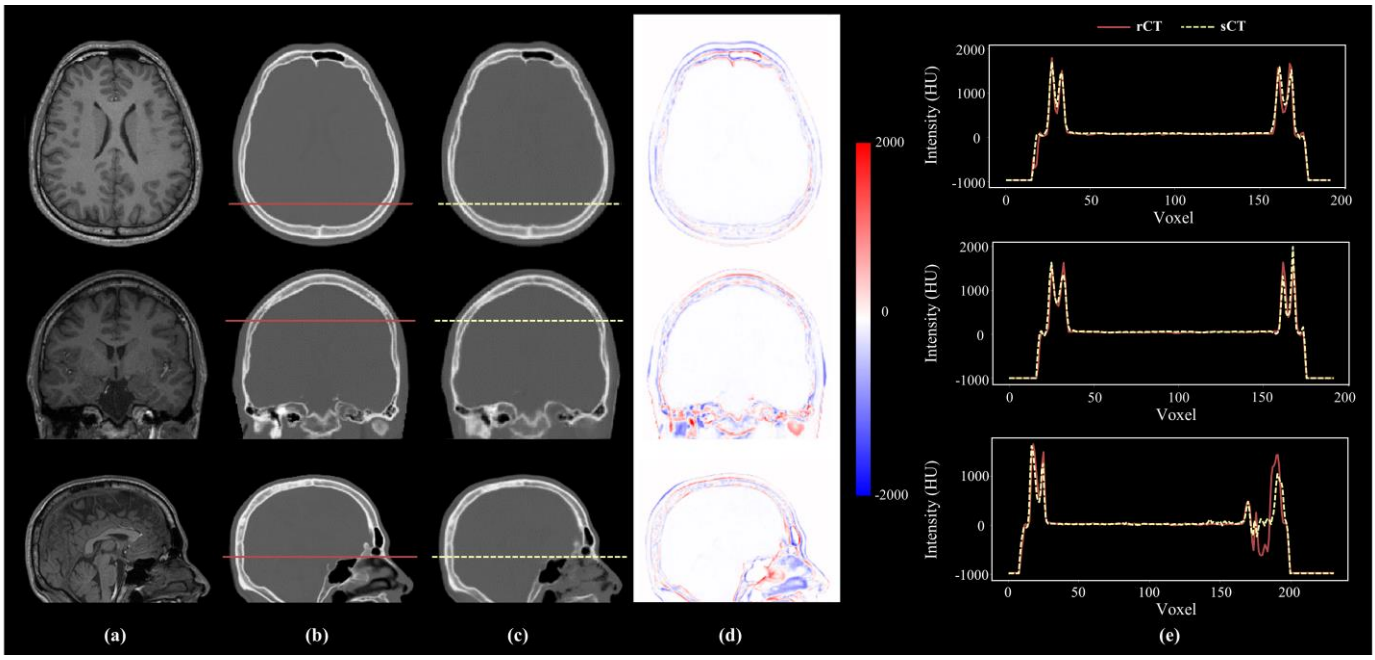

Fig. 3. Axial, coronal, and sagittal view of (a) MRI, (b) rCT, and (c) sCT data from a representative test subject (S9). (d) Difference maps in HU intensity produced by subtracting rCT from sCT. (e) The profile plot along the red line in (b) and yellow dotted line in (c).

TABLE III
IMAGE QUALITIES ON MRI NORMALIZATION

Model	MAE_{head}	MAE_{skull}	DSC_{skull}
Raw	93.80±13.40	299.26±41.20	0.87±0.03
Z-score	85.72±9.50	280.25±24.02	0.88±0.02
HM	91.02±17.18	300.81±48.06	0.87±0.04

the greater MPD than GDL showed the lowest MAEs and the highest DSC_{skull}. The corresponding model was adapted to investigate the effect of MRI normalization, in which the Z-score normalization outperformed raw or HM normalization (Table III). The image qualities from the Z-score normalization model were MAE_{head} = 85.72 ± 9.50, MAE_{skull} =280.25 ± 24.02, and DSC_{skull} = 0.88 ± 0.02 across 15 test subjects. Fig. 3 shows that the model can generate sCT similar to rCT. Particularly, soft tissues were generated with virtually no error (i.e., a difference close to 0 HU; Fig. 3d). Although the model can reconstruct anatomical structures comparable to rCT, there are HU intensity discrepancies in the cancellous bone of the skull. Most of the discrepancies exist in the interface regions between different media, especially between the air and bone. The first

TABLE IV
COMPARISON OF SKULL CHARACTERISTICS

Target	SDR			ST		
	rCT	sCT	*p*	rCT	sCT	*p*
M1	0.54±0.17	0.55±0.14	0.56	6.31±1.2	6.57±0.75	0.16
V1	0.58±0.14	0.60±0.11	0.08	6.27±1.12	6.5±0.82	0.07
dACC	0.55±0.14	0.55±0.12	0.58	6.55±1.13	6.86±0.73	0.07

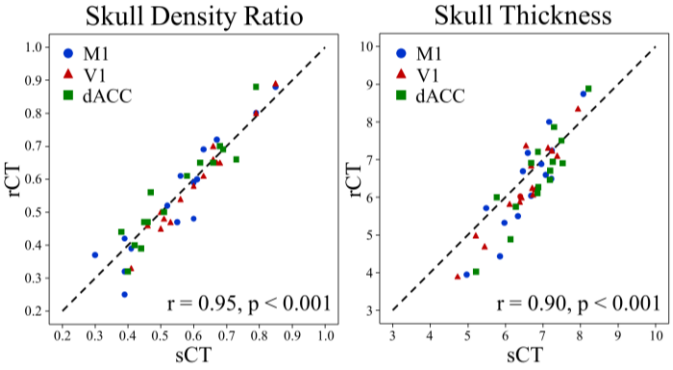

Fig. 4. Correlations of skull characteristics between rCT and sCT for the three target regions (M1, V1 and dACC).

(axial) and second row (coronal) in Fig. 3e showed similar tendencies across the soft tissue and skull. There were mismatches across the nasal regions in the third row (sagittal), which resulted in relatively blurry images. The estimation of skull characteristics and the acoustic simulation were performed using the Z-score based sCT.

B. Skull characteristics

As shown in Fig. 4, significant correlations between sCT and rCT in terms of skull characteristics were observed (r = 0.95, p < 0.001 for SDR; r = 0.90, p < 0.001 for ST). Furthermore, as Table IV shows, there was no statistically significant difference between sCT and rCT in terms of skull characteristics.

C. Acoustic simulation

Fig. 5 shows an example of the 2D normalized acoustic pressure distribution when targeting dACC. The absolute difference in acoustic simulation between rCT and sCT is shown in the last column. Fig. 6 is the 3D version of acoustic simulation from another subject. The simulated 3D focal region (i.e., $A_{90\%}$) and overlapped area (i.e., $A_{90\%} \cap \hat{A}_{90\%}$) are depicted in the last row. The overall mean and standard

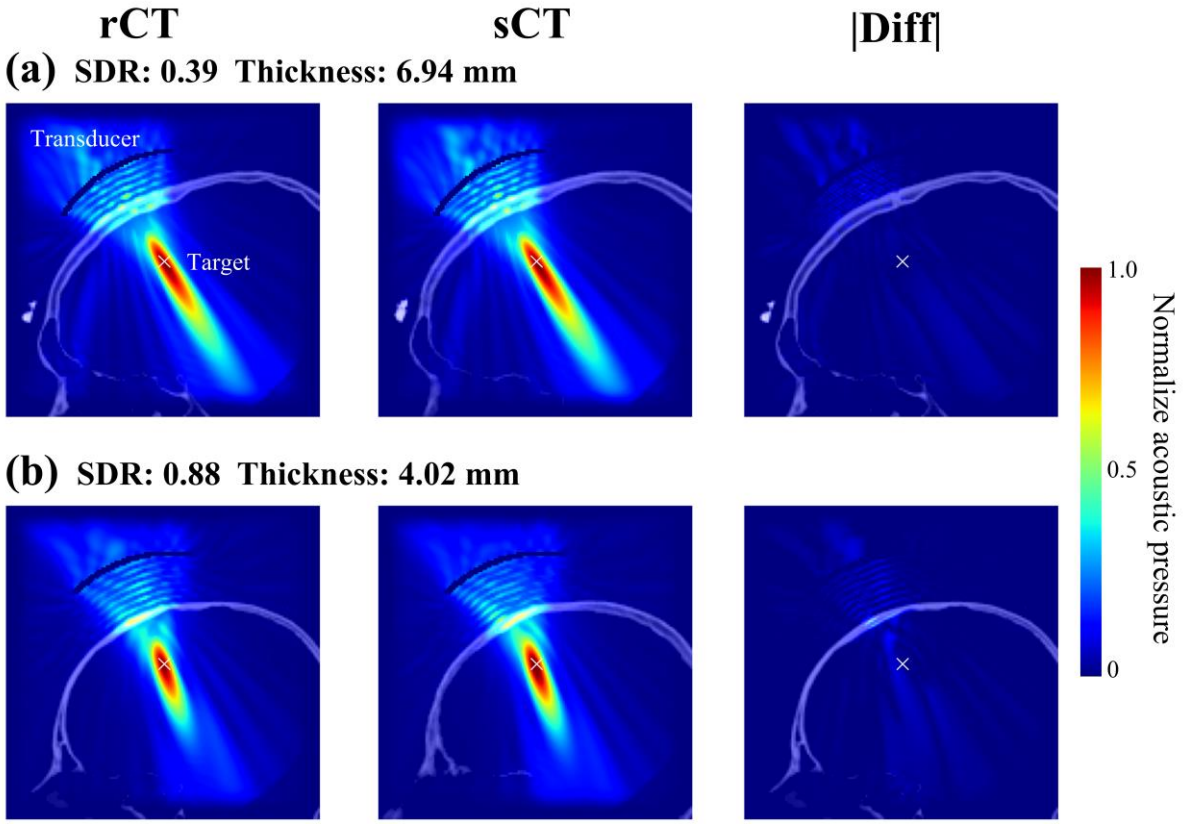

Fig. 5. Acoustic simulation results comparing sCT and rCT when targeting dACC in test subjects (a for S15 in the low SDR group with high skull thickness; b for S1 in the high SDR group with low skull thickness). The absolute difference between the pressure maps is indicated in the last column. The point (i.e., '×') in the brain is the intended target.

TABLE V
EVALUATION OF ACOUSTIC SIMULATION RESULTS

Target	Δ_{PR} (%)	$DSC_{A90\%}$	Δ_F (mm)
M1	3.84±2.97	0.81±0.09	0.90±0.55
V1	2.05±2.19	0.86±0.07	0.75±0.50
dACC	3.45±2.84	0.82±0.09	0.94±0.51
Mean±STD	3.11±2.79	0.83±0.08	0.86±0.52

deviation (STD) for acoustic simulation evaluation were Δ_{PR} = 3.11 ± 2.97 %, $DSC_{A90\%}$ = 0.83 ± 0.08, and Δ_F = 0.86 ± 0.52 mm. The evaluation result for each of the three target regions is summarized in Table V and Fig. 7.

IV. DISCUSSION

In the present study, we investigated the feasibility of acoustic simulation using cGAN-generated sCT for tFUS applications. We suggested a 3D-cGAN model incorporating various loss functions that enhanced the quality of the synthetic images. The CT and MRI data from 15 subjects were used to train the proposed model, and the data from the remaining 15 subjects were used for evaluation of the trained model. The performance depending on weight factors in loss functions and normalization schemes of MRI were also investigated. The quality of the generated sCT was overall best from the cGAN model with weight factors, [λ_1=1, λ_2=30, λ_3=10] using Z-scored MRI with average MAE_{skull} of 280 HU and DSC_{skull} of 0.88 across the 15 test subjects. In terms of skull characteristics (SDR and ST), there was no significant difference between rCT and sCT for all target regions. The acoustic simulation results also showed a high similarity between rCT and sCT (Δ_{PR}=3.11 %, $DSC_{A90\%}$=0.83, and Δ_F =0.86 mm).

We showed that the cGAN model performed best when three losses (LS, MPD, and GDL) were combined together compared to if only one of either MPD or GDL was added to the LS loss. This finding is in line with the previous studies that models with GDL alleviated the blurry issue of synthetic images compared to the models without GDL [36], [44]. The cGAN model trained without MPD, however, resulted in the largest error, and the cGAN with three losses combined and with highly weighted MPD showed the best performance (Table II). This indicates that minimizing direct difference (i.e., MPD) plays an important role in image generation tasks. An additional cost function to minimize derivative values via GDL helped to further enhance the quality of generated CT images. Future work is warranted to systematically change the weight factors in generator loss and consequently optimize these hyperparameters for a broader range of candidate values. In addition to the performance varying depending on the weight factors in the loss function, the normalization scheme of MRI also affected the quality of the generated CT images. Z-score normalized MRI substantially improved the sCT (Table III). The HM normalization showed similar results with the non-normalized method. Previous studies reported that image synthesis results

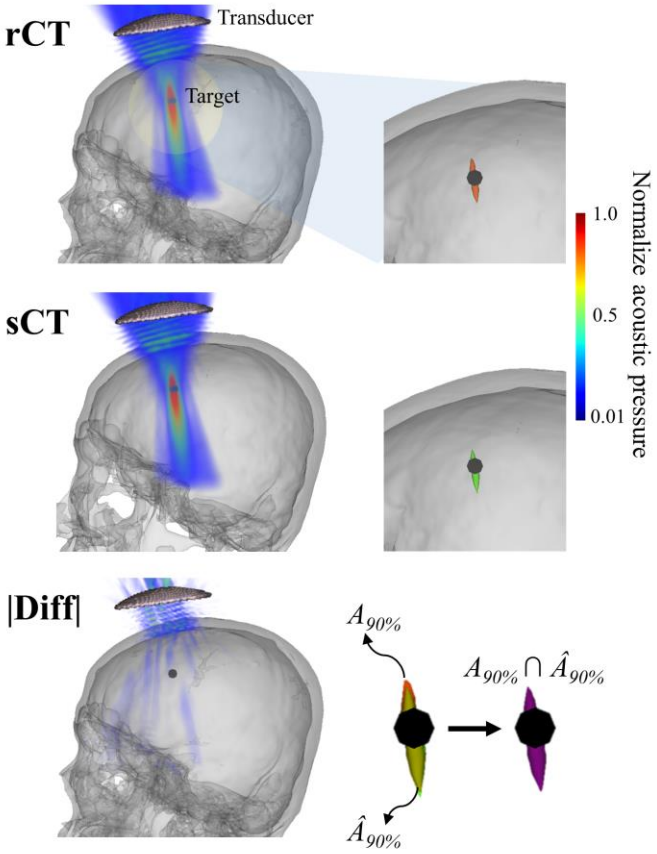

Fig. 6. Example of 3D acoustic simulation results (S13). The simulated 3D focal region (i.e. $A_{90\%}$) and overlapped area are indicated in the second column.

were stable across MRI normalization methods, however, the optimal scheme for MRI intensity normalization also depends on various conditions such as raw MRI intensity ranges and adopted models [30], [33]. We have also demonstrated that the MRI normalization scheme altered the quality of the synthetic CT, stressing the importance of the systematic investigation across various normalization methods for MRI in a future study.

The image quality evaluation results (MAE_{head} = 85.72 ± 9.50, MAE_{skull} = 280.25±24.02, and DSC_{skull} = 0.88±0.02) outperform the previous studies based on atlas (MAE_{head} = 113.4±17.6 HU) [45], classification (MAE_{head} = 147.5±8.3 HU, MAE_{skull} = 422.5±33.4 HU) [46], and deep learning (MAE_{head} = 120.1±20.4, MAE_{skull} = 399.4±51.8) [37]. In addition, a skull characteristics evaluation was performed since the inhomogeneity and thickness of the skull also crucially affect acoustic simulation results . The SDRs determined by rCT and sCT showed high correlations for all targets (r = 0.95), suggesting that the HU intensity distribution of the skull derived from sCT is similar to that from rCT (Fig. 4). This result is comparable to the correlation of SDR (r = 0.88) in the previous study [10]. The ST values also showed high correlation (r = 0.90) with no significant difference for any target region tested ($p > 0.05$). The ST result is congruent with the outperformance in DSC_{skull} (0.88±0.02), suggesting that sCT can effectively capture the skull structure from the corresponding MRI.

As shown in Fig. 3d, the errors have mainly occurred at the tissue interfaces. Similar issues have been repeatedly reported in previous studies [16], [30], [37], [47]. The ambiguous boundary of the air region in MRI due to artifacts from bony tissue could be responsible for the generation of sCT images with interface discrepancies. The rigid image registration method that we used might also contribute to this issue. For the acoustic simulation, rigid registration was preferred over non-rigid registration to preserve the skull structure. However, Zheng et al. has shown that, compared to non-rigid registration, rigid registration provides poor overlays of the air region mask between brain CT and MRI [46]. This indicates that rigid registration could be insufficient to make perfectly paired training data for cGAN. We chose an averaging strategy to generate the full volume of sCT from patches, and this might be associated with the blurry tissue interfaces affected by outlier HU [30]. Thus, alternative registration methods and patch fusion strategies (e.g., median and voting) should be tested in future work.

The overall mean of Δ_{PR} was 3.11%, which shows that the acoustic simulation using sCT can accurately predict the intracranial acoustic pressure. The average of $DSC_{A90\%}$ and ΔF were 0.83 and 0.86 mm, respectively, reflecting small discrepancies in the position and shape of the acoustic focus between sCT and rCT. We conjecture that the use of the 3D-cGAN model is the major contributing factor to the simulation accuracy. Since ultrasound usually penetrates the skull along

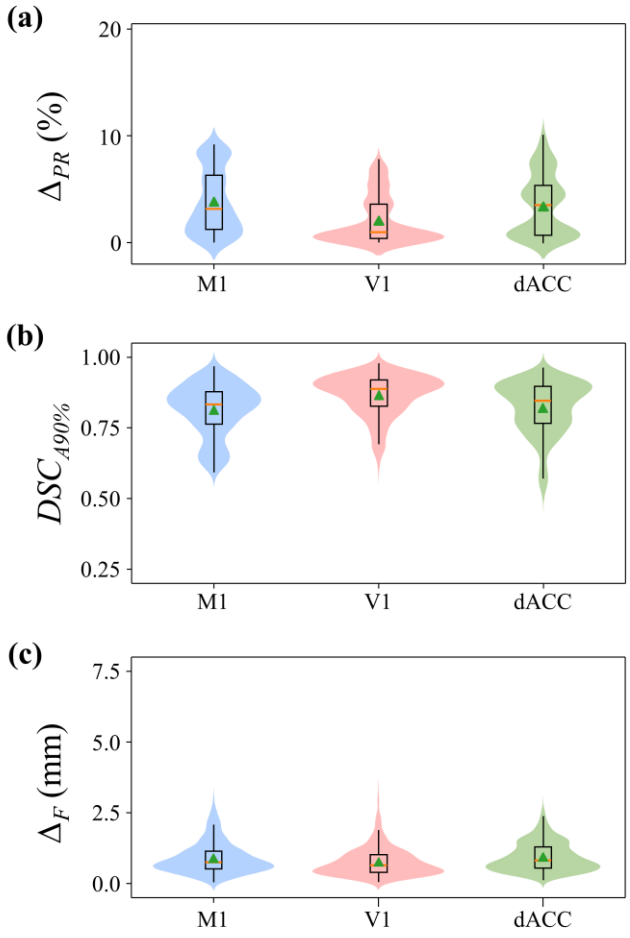

Fig.7. Violin plot showing (a) the peak pressure ratio (Δ_{PR}), b) the dice similarity coefficient of acoustic focus ($DSC_{A90\%}$), and c) the distance between the acoustic focuses (Δ_F). Inside each violin plot is a box plot summarizing ranges. The orange central line and green triangle in each box indicate the median and mean value, respectively.

the oblique direction, the acoustic simulation result is vulnerable to 3D reconstruction errors such as the stair-step artifact. The 3D-cGAN method that preserves the context information of the anatomical structure minimizes this reconstruction error, leading to accurate simulation results when using sCT. Moreover, as we mentioned above, the main error of sCT was observed at the skull interface, and skull thickness errors were less than 0.3 mm (Table IV). Since the wavelength of the ultrasound that we used (i.e., 7.5 mm) was much greater than these errors, the simulation result might not be substantially affected.

Although the sCT showed a promising possibility for MRI-only tFUS neuromodulation, additional evaluation is required before using sCT in real clinical practice. Because the errors of sCT have greater impacts at higher fundamental frequencies (i.e., shorter wavelengths), we could expect that the accuracy of acoustic simulation with sCT would gradually decrease with increasing frequency. Therefore, an assessment of sCT at ultrasonic conditions higher than 200 kHz needs to be performed. Due to the limited focal length of the transducer, only cortical and subcortical regions were tested in this study. Thus, the feasibility of sCT when targeting deep brain areas need to be the subject of future work. As illustrated in Fig. 5, SDR did not have a significant impact on the accuracy of the acoustic simulation. Since the number of subjects with low SDRs (i.e., < 0.5) was relatively small in this study (7 for M1; 6 for each of V1 and dACC), we did not conduct a statistical analysis. Future studies are warranted by including the increased number of subjects with various skull characteristics in this context.

Similarly with our work, previous studies have also shown the possibility of using synthetic CT (i.e., sCT) in the multi-array transducer condition [10], [18]. However, the sonication environment of our study with a single-element transducer for neuromodulation and previous studies using multi-array transducer for thermal ablation is very different. A relatively small area of the skull affects ultrasound waves when using the single-element transducer (i.e., 32 cm^2 in this study), while when using the multi-array transducer, the overall shape of the skull may affect wave propagation. Moreover, since the intensity of tFUS neuromodulation is mainly less than 1MPa [48], which makes only negligible temperature change, the thermal simulation was not conducted in this study. In order to extend the potential applications of the proposed method, additional evaluation such as acoustic and thermal simulation on the multi-array transducer system is needed.

The proposed model might be advantageous for preserving context information using 3D-cGAN due to its incorporation of multiple loss functions (i.e., LS, MPD, and GDL), which helped stabilize training and retain image details. Moreover, the residual blocks embedded in the generator concatenate feature maps from neighboring blocks. This structure focuses on the residuals, which indicate the differences between the two feature maps, and lead to an accurate mapping from MRI to CT [16]. The previous study reported that the GAN-based model accurately distinguishes the blurry images as fake, which can further enhance the quality of the synthetic image with a similar level of resolution as the ground truth image [49]. Indeed, Emami et al. claimed that cGAN generates higher resolution sCTs preventing ambiguous boundaries of soft tissue and skull than ResNet CNN (i.e., cGAN with discriminator removed) [15]. Therefore, we believe that the performance might be degraded when the generator is used alone without a discriminator, although we combined additional loss (i.e., GDL) in the generator. Furthermore, by utilizing a data augmentation strategy per epoch and patch extraction, the model could be trained with diverse data despite the training data being limited to 15 subjects. The volume of training data used in this study was relatively small compared to previous studies (i.e., > 25 subjects) [17], [18], [30] Given that a previous study using cGAN [50] showed decreasing MAEs with increasing training data volumes (34-135 subjects), future investigation testing the proposed model with more training data sets is warranted.

Recent studies have developed sCT generation techniques u sing various MRI sequences (i.e., T1w, T2w, Flair, Dixon, and UTE) [18], [37], [51], [52]. However, clinical situations frequ ently feature restrictions preventing the acquisition of multiple or specialized MRI sequences. Previous studies have demonst rated that T1w performs the best of the conventional MRI sequ ences when generating sCT [51], [52]. This might be associate d with the superior anatomical structure representation of T1w MRI. With the evidence and its accessibility, we adopted the st andard T1w MRI as the modality to investigate the feasibility of generating sCT for single-element tFUS applications. Since the potential issue of many MRI sequences is the poor contrast of bone which might induce the misclassification of surroundi ng tissues or air, applications of bone specialized MRI such as UTE or ZTE have also demonstrated the feasibility in CT gene ration [10], [13]. To the best of our knowledge, it has not yet b een directly compared the performance of the deep learning-ba sed method to generate CT images using bone specialized MR I sequence and conventional MRI sequence. Investigating the a dvantages of each sequence as inputs to cGAN would be intere sting future work. In addition, one potential issue regarding the learning-based sCT is generality in the performance of the mo del when an MRI from a different scanner is fed despite havin g an equivalent type of sequence. Our sCT generator, like any other model, may be sensitive to alternative T1w images acqui red from different pulse sequences and/or scanner vendors com pared to a training set. This is because the intensity configurati on across different anatomical structures of gray matter, white matter, cerebrospinal fluid (CSF), and non-brain areas includin g the skull may vary depending on the imaging pulse sequence s and/or MRI scanner vendors. Although some MRI intensity n ormalization approaches may alleviate this issue, it has not bee n evaluated clearly. As it's common to collect T1w MRI image s from several candidate pulse sequences or from different sca nner vendors in clinical practice, it would be important future work to investigate the robustness in the performance of the m odel in the perspective of MRI data harmonization [53], [54].

In this study, we examined the feasibility of generating sCT from T1w MRI for acoustic simulation in single-element conditions. Skull characteristics revealed by rCT and sCT demonstrated high correlations, and the result of acoustic simulation showed clinically tolerable levels of difference. This indicates the possibility of sCT replacing rCT for MRI-only tFUS treatment. The proposed use of sCT could simplify clinical workflow and prevent the exposure of patients and healthcare workers to radiation.

REFERENCES

[1] H. Baek, K. J. Pahk, and H. Kim, "A review of low-intensity focused ultrasound for neuromodulation," *Biomed. Eng. Lett.*, vol. 7, no. 2, pp. 135–142, 2017, doi: 10.1007/s13534-016-0007-y.

[2] C. Pasquinelli, L. G. Hanson, H. R. Siebner, H. J. Lee, and A. Thielscher, "Safety of transcranial focused ultrasound stimulation: A systematic review of the state of knowledge from both human and animal studies," *Brain Stimul.*, vol. 12, no. 6, pp. 1367–1380, 2019, doi: 10.1016/j.brs.2019.07.024.

[3] W. Legon *et al.*, "A retrospective qualitative report of symptoms and safety from transcranial focused ultrasound for neuromodulation in humans," *Sci. Rep.*, vol. 10, no. 1, pp. 1–10, 2020, doi: 10.1038/s41598-020-62265-8.

[4] B. W. Badran *et al.*, "Sonication of the anterior thalamus with MRI-Guided transcranial focused ultrasound (tFUS) alters pain thresholds in healthy adults: A double-blind, sham-controlled study," *Brain Stimul.*, vol. 13, no. 6, pp. 1805–1812, 2020, doi: 10.1016/j.brs.2020.10.007.

[5] W. Lee, D. S. Weisholtz, G. E. Strangman, and S. Yoo, "Safety Review and Perspectives of Transcranial Focused Ultrasound Brain Stimulation," *Brain & NeuroRehabilitation*, vol. 14, no. 1, 2020.

[6] S.-S. Yoo, "Technical Review and Perspectives of Transcranial Focused Ultrasound Brain Stimulation for Neurorehabilitation," *Brain & Neurorehabilitation*, vol. 11, no. 2, 2018, doi: 10.12786/bn.2018.11.e16.

[7] G. Pinton, J. F. Aubry, E. Bossy, M. Muller, M. Pernot, and M. Tanter, "Attenuation, scattering, and absorption of ultrasound in the skull bone," *Med. Phys.*, vol. 39, no. 1, pp. 299–307, 2012, doi: 10.1118/1.3668316.

[8] K. Yoon, W. Lee, P. Croce, A. Cammalleri, and S. S. Yoo, "Multi-resolution simulation of focused ultrasound propagation through ovine skull from a single-element transducer," *Phys. Med. Biol.*, vol. 63, no. 10, 2018, doi: 10.1088/1361-6560/aabe37.

[9] T. Y. Park, K. J. Pahk, and H. Kim, "Method to optimize the placement of a single-element transducer for transcranial focused ultrasound," *Comput. Methods Programs Biomed.*, vol. 179, p. 104982, 2019, doi: 10.1016/j.cmpb.2019.104982.

[10] S. Guo *et al.*, "Feasibility of ultrashort echo time images using full-wave acoustic and thermal modeling for transcranial MRI-guided focused ultrasound (tcMRgFUS) planning," *Phys. Med. Biol.*, vol. 64, no. 9, 2019, doi: 10.1088/1361-6560/ab12f7.

[11] S. J. Tepper, "Computed tomography - An increasing source of radiation exposure: Commentary," *Headache*, vol. 48, no. 4, p. 657, 2008, doi: 10.1111/j.1526-4610.2008.01071.x.

[12] P. K. Nguyen and J. C. Wu, "Radiation exposure from imaging tests: Is there an increased cancer risk?," *Expert Rev. Cardiovasc. Ther.*, vol. 9, no. 2, pp. 177–183, 2011, doi: 10.1586/erc.10.184.

[13] F. Wiesinger *et al.*, "Zero TE-based pseudo-CT image conversion in the head and its application in PET/MR attenuation correction and MR-guided radiation therapy planning," *Magn. Reson. Med.*, vol. 80, no. 4, pp. 1440–1451, 2018, doi: 10.1002/mrm.27134.

[14] X. Han, "MR-based synthetic CT generation using a deep convolutional neural network method:," *Med. Phys.*, vol. 44, no. 4, pp. 1408–1419, 2017, doi: 10.1002/mp.12155.

[15] H. Emami, M. Dong, S. P. Nejad-Davarani, and C. K. Glide-Hurst, "Generating synthetic CTs from magnetic resonance images using generative adversarial networks," *Med. Phys.*, vol. 45, no. 8, pp. 3627–3636, 2018, doi: 10.1002/mp.13047.

[16] Y. Lei *et al.*, "MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks," *Med. Phys.*, vol. 46, no. 8, pp. 3565–3581, 2019, doi: 10.1002/mp.13617.

[17] G. Shafai-Erfani *et al.*, "MRI-based proton treatment planning for base of skull tumors," *Int. J. Part. Ther.*, vol. 6, no. 2, pp. 12–25, 2020, doi: 10.14338/IJPT-19-00062.1.

[18] P. Su *et al.*, "Transcranial MR Imaging–Guided Focused Ultrasound Interventions Using Deep Learning Synthesized CT," *Am. J. Neuroradiol.*, pp. 1–8, 2020, doi: 10.3174/ajnr.a6758.

[19] W. Legon, P. Bansal, R. Tyshynsky, L. Ai, and J. K. Mueller, "Transcranial focused ultrasound neuromodulation of the human primary motor cortex," *Sci. Rep.*, vol. 8, no. 1, pp. 1–14, 2018, doi: 10.1038/s41598-018-28320-1.

[20] W. Lee *et al.*, "Transcranial focused ultrasound stimulation of human primary visual cortex," *Sci. Rep.*, vol. 6, no. May, pp. 1–12, 2016, doi: 10.1038/srep34026.

[21] M. Fini and W. J. Tyler, "Emotional Processing , Physiology , and Performance I : Dorsal Anterior Cingulate Cortex Targeting .," *medRxiv*, pp. 1–29, 2020.

[22] J. Wang, L. He, H. Zheng, and Z. L. Lu, "Optimizing the Magnetization-Prepared Rapid Gradient-Echo (MP-RAGE) sequence," *PLoS One*, vol. 9, no. 5, 2014, doi: 10.1371/journal.pone.0096899.

[23] W. Lee, H. Kim, Y. Jung, I. U. Song, Y. A. Chung, and S. S. Yoo, "Image-guided transcranial focused ultrasound stimulates human primary somatosensory cortex," *Sci. Rep.*, vol. 5, pp. 1–10, 2015, doi: 10.1038/srep08743.

[24] W. Lee, Y. A. Chung, Y. Jung, I. U. Song, and S. S. Yoo, "Simultaneous acoustic stimulation of human primary and secondary somatosensory cortices using transcranial focused ultrasound," *BMC Neurosci.*, vol. 17, no. 1, pp. 1–11, 2016, doi: 10.1186/s12868-016-0303-6.

[25] N. J. Tustison *et al.*, "N4ITK: Improved N3 bias correction," *IEEE Trans. Med. Imaging*, vol. 29, no. 6, pp. 1310–1320, 2010, doi: 10.1109/TMI.2010.2046908.

[26] S. Klein, M. Staring, K. Murphy, M. a. Viergever, and J. Pluim, "elastix: A Toolbox for Intensity-Based Medical Image Registration," *IEEE Trans. Med. Imaging*, vol. 29, no. 1, pp. 196–205, 2010.

[27] S. Pieper, M. Halle, and R. Kikinis, "3D Slicer," *2004 2nd IEEE Int. Symp. Biomed. Imaging Macro to Nano*, vol. 1, no. May, pp. 632–635, 2004, doi: 10.1109/isbi.2004.1398617.

[28] M. Ferreira, "A generative adversarial network approach to synthetic-CT creation for MRI-based radiation therapy," 2019.

[29] P. Virtanen *et al.*, "SciPy 1.0: fundamental algorithms for scientific computing in Python," *Nat. Methods*, vol. 17, no. 3, pp. 261–272, 2020, doi: 10.1038/s41592-019-0686-2.

[30] P. Klages *et al.*, "Patch-based generative adversarial neural network models for head and neck MR-only planning," *Med. Phys.*, vol. 47, no. 2, pp. 626–642, 2020, doi: 10.1002/mp.13927.

[31] J. K. Udupa, X. Zhang, and A. Notation, "New Variants of a Method of MRI Scale Standardization," *IEEE Trans. Med. Imaging*, vol. 19, no. 2, pp. 143–150, 2000.

[32] M. Shah *et al.*, "Evaluating intensity normalization on MRIs of human brain with multiple sclerosis," *Med. Image Anal.*, vol. 15, no. 2, pp. 267–282, 2011, doi: 10.1016/j.media.2010.12.003.

[33] J. C. Reinhold, B. E. Dewey, A. Carass, and J. L. Prince, "Evaluating the impact of intensity normalization on MR image synthesis," p. 126, 2019, doi: 10.1117/12.2513089.

[34] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 5967–5976, 2017, doi: 10.1109/CVPR.2017.632.

[35] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least Squares Generative Adversarial Networks," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2017-Octob, pp. 2813–2821, 2017, doi: 10.1109/ICCV.2017.304.

[36] D. Nie *et al.*, "Medical Image Synthesis with Deep Convolutional Adversarial Networks," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 12, pp. 2720–2730, 2018, doi: 10.1109/TBME.2018.2814538.

[37] Y. Koike *et al.*, "Feasibility of synthetic computed tomography generated with an adversarial network for multi-sequence magnetic resonance-based brain radiotherapy," *J. Radiat. Res.*, vol. 61, no. 1, pp. 92–103, 2019, doi: 10.1093/jrr/rrz063.

[38] C. Bin Jin *et al.*, "Deep CT to MR synthesis using paired and unpaired data," *Sensors (Switzerland)*, vol. 19, no. 10, pp. 1–19, 2019, doi: 10.3390/s19102361.

[39] V. Fonov, A. C. Evans, K. Botteron, C. R. Almli, R. C. McKinstry, and D. L. Collins, "Unbiased average age-appropriate atlases for pediatric studies," *Neuroimage*, vol. 54, no. 1, pp. 313–327, 2011, doi: 10.1016/j.neuroimage.2010.07.033.

[40] B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *J. Biomed. Opt.*, vol. 15, no. 2, p. 021314, 2010, doi: 10.1117/1.3360308.

[41] M. D'Souza *et al.*, "Impact of skull density ratio on efficacy and safety of magnetic resonance–guided focused ultrasound treatment of essential tremor," *J. Neurosurg.*, vol. 132, no. 5, pp. 1392–1397, May 2020, doi: 10.3171/2019.2.JNS183517.

[42] J. A. Pineda-Pardo *et al.*, "Letter: The Role of Skull Thickness beyond the Skull Density Ratio on MRgFUS Thalamotomy Feasibility: Which Patients Should We Exclude?," *Neurosurgery*, vol. 86, no. 5. Oxford University Press, pp. E477–E479, May 01, 2020, doi: 10.1093/neuros/nyz566.

[43] F. Sammartino, D. W. Beam, J. Snell, and V. Krishna, "Kranion, an open-source environment for planning transcranial focused ultrasound surgery: Technical note," in *Journal of Neurosurgery*, Mar. 2020, vol. 132, no. 4, pp. 1249–1255, doi: 10.3171/2018.11.JNS181995.

[44] M. Mathieu, C. Couprie, and Y. LeCun, "Deep multi-scale video prediction beyond mean square error," *4th Int. Conf. Learn. Represent. ICLR 2016 - Conf. Track Proc.*, no. 2015, pp. 1–14, 2016.

[45] J. Sjölund, D. Forsberg, M. Andersson, and H. Knutsson, "Generating patient specific pseudo-CT of the head from MR using atlas-based regression," *Phys. Med. Biol.*, vol. 60, no. 2, pp. 825–839, Jan. 2015, doi: 10.1088/0031-9155/60/2/825.

[46] W. Zheng, J. P. Kim, M. Kadbi, B. Movsas, I. J. Chetty, and C. K. Glide-Hurst, "Magnetic Resonance-Based Automatic Air Segmentation for Generation of Synthetic Computed Tomography Scans in the Head Region," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 93, no. 3, pp. 497–506, 2015, doi: 10.1016/j.ijrobp.2015.07.001.

[47] Y. Peng *et al.*, "Magnetic resonance-based synthetic computed tomography images generated using generative adversarial networks for nasopharyngeal carcinoma radiotherapy treatment planning," *Radiother. Oncol.*, vol. 150, pp. 217–224, Sep. 2020, doi: 10.1016/j.radonc.2020.06.049.

[48] J. K. Mueller, L. Ai, P. Bansal, and W. Legon, "Numerical evaluation of the skull for human neuromodulation with transcranial focused ultrasound," *J. Neural Eng.*, vol. 14, no. 6, 2017, doi: 10.1088/1741-2552/aa843e.

[49] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 105–114, 2017, doi: 10.1109/CVPR.2017.19.

[50] J. M. Edmund, H. M. Kjer, K. Van Leemput, R. H. Hansen, J. A. L. Andersen, and D. Andreasen, "A voxel-based investigation for MRI-only radiotherapy of the brain using ultra short echo times," *Phys. Med. Biol.*, vol. 59, no. 23, pp. 7501–7519, 2014, doi: 10.1088/0031-9155/59/23/7501.

[51] M. Qi *et al.*, "Multi-sequence MR image-based synthetic CT generation using a generative adversarial network for head and neck MRI-only radiotherapy," *Med. Phys.*, vol. 47, no. 4, pp. 1880–1894, 2020, doi: 10.1002/mp.14075.

[52] A. Larroza *et al.*, "Deep learning for MRI-based CT synthesis: A comparison of MRI sequences and neural network architectures," *2019 IEEE Nucl. Sci. Symp. Med. Imaging Conf. NSS/MIC 2019*, pp. 2019–2022, 2019, doi: 10.1109/NSS/MIC42101.2019.9060051.

[53] R. Pomponio *et al.*, "Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan," *Neuroimage*, vol. 208, no. July 2019, 2020, doi: 10.1016/j.neuroimage.2019.116450.

[54] J. Radua *et al.*, "Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA," *Neuroimage*, vol. 218, no. May, 2020, doi: 10.1016/j.neuroimage.2020.116956.