

Multi-parameter ultrasound imaging for musculoskeletal tissues based on a physics informed generative adversarial network

Pengxin Wang(王鹏鑫), Heyu Ma(马贺雨), Tianyu Liu(刘天宇), Chengcheng Liu(刘成成), Dan Li(李旦), and Dean Ta(他得安)

Citation: Chin. Phys. B, 2025, 34 (**4**): 044301. DOI: 10.1088/1674-1056/adb390
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Scheme of negative acoustic radiation force based on a multiple-layered spherical structure

Menyang Gong(宫门阳), Xin Xu(徐鑫), Yupei Qiao(乔玉配), Jiehui Liu(刘杰惠), Aijun He(何爱军), and Xiaozhou Liu(刘晓宙)

Chin. Phys. B, 2024, 33 (**1**): 014302. DOI: 10.1088/1674-1056/acfa80

Wideband frequency-dependent dielectric properties of rat tissues exposed to low-intensity focused ultrasound in the microwave frequency range

Xue Wang(王雪), Shi-Xie Jiang, Lin Huang(黄林), Zi-Hui Chi(迟子惠), Dan Wu(吴丹), and Hua-Bei Jiang

Chin. Phys. B, 2023, 32 (**3**): 034305. DOI: 10.1088/1674-1056/ac7453

Quantitative ultrasound brain imaging with multiscale deconvolutional waveform inversion

Yu-Bing Li(李玉冰), Jian Wang(王建), Chang Su(苏畅), Wei-Jun Lin(林伟军), Xiu-Ming Wang(王秀明), and Yi Luo(骆毅)

Chin. Phys. B, 2023, 32 (**1**): 014303. DOI: 10.1088/1674-1056/ac6dad

Acoustic radiation force on a rigid cylinder near rigid corner boundaries exerted by a Gaussian beam field

Qin Chang(常钦), Yuchen Zang(臧雨宸), Weijun Lin(林伟军), Chang Su(苏畅), and Pengfei Wu(吴鹏飞)

Chin. Phys. B, 2022, 31 (**4**): 044302. DOI: 10.1088/1674-1056/ac2d1f

High-resolution bone microstructure imaging based on ultrasonic frequency-domain full-waveform inversion

Yifang Li(李义方), Qinzhen Shi(石勤振), Ying Li(李颖), Xiaojun Song(宋小军), Chengcheng Liu(刘成成), Dean Ta(他得安), and Weiqi Wang(王威琪)

Chin. Phys. B, 2021, 30 (**1**): 014302. DOI: 10.1088/1674-1056/abc7aa

---

Multi-parameter ultrasound imaging for musculoskeletal tissues based on a physics informed generative adversarial network

Pengxin Wang(王鹏鑫)[1], Heyu Ma(马贺雨)[1], Tianyu Liu(刘天宇)[1], Chengcheng Liu(刘成成)[1,2,†], Dan Li(李旦)[3], and Dean Ta(他得安)[2,4]

[1] *Institute of Biomedical Engineering & Technology, Academy for Engineering and Technology, Fudan University, Shanghai* 200433, *China*
[2] *State Key Laboratory of Integrated Chips and Systems, Fudan University, Shanghai* 201203, *China*
[3] *Department of Electronic Engineering, School of Information Science and Technology, Fudan University, Shanghai* 200438, *China*
[4] *Department of Biomedical Engineering, School of Information Science and Technology, Fudan University, Shanghai* 200438, *China*

(Received 7 January 2025; revised manuscript received 27 January 2025; accepted manuscript online 7 February 2025)

Full waveform inversion (FWI) has showed great potential in the detection of musculoskeletal disease. However, FWI is an ill-posed inverse problem and has a high requirement on the initial model during the imaging process. An inaccurate initial model may lead to local minima in the inversion and unexpected imaging results caused by cycle-skipping phenomenon. Deep learning methods have been applied in musculoskeletal imaging, but need a large amount of data for training. Inspired by work related to generative adversarial networks with physical informed constrain, we proposed a method named as bone ultrasound imaging with physics informed generative adversarial network (BUIPIGAN) to achieve unsupervised multi-parameter imaging for musculoskeletal tissues, focusing on speed of sound (SOS) and density. In the in-silico experiments using a ring array transducer, conventional FWI methods and BUIPIGAN were employed for multi-parameter imaging of two musculoskeletal tissue models. The results were evaluated based on visual appearance, structural similarity index measure (SSIM), signal-to-noise ratio (SNR), and relative error (RE). For SOS imaging of the tibia–fibula model, the proposed BUIPIGAN achieved accurate SOS imaging with best performance. The specific quantitative metrics for SOS imaging were SSIM 0.9573, SNR 28.70 dB, and RE 5.78%. For the multi-parameter imaging of the tibia–fibula and human forearm, the BUIPIGAN successfully reconstructed SOS and density distributions with SSIM above 94%, SNR above 21 dB, and RE below 10%. The BUIPIGAN also showed robustness across various noise levels (i.e., 30 dB, 10 dB). The results demonstrated that the proposed BUIPIGAN can achieve high-accuracy SOS and density imaging, proving its potential for applications in musculoskeletal ultrasound imaging.

Keywords: ultrasound image, physics informed, generative adversarial network, musculoskeletal imaging

PACS: 43.60.Lq, 43.80.Qf, 43.35.Wa, 87.63.dh

DOI: 10.1088/1674-1056/adb390 **CSTR:** 32038.14.CPB.adb390

1. Introduction

Bone constitutes a fundamental element of the human skeletal system, exerting a pivotal function in the physiological activities of the human body.[1] The integrity of bone health is vital for the maintenance of normal physiological functions in humans. However, with the advancement of age and the alteration of specific environmental factors, there is an accelerated loss of bone mass, which consequently elevates the risk of developing osteoporosis.[2] Osteoporosis is a metabolic bone disorder characterized by changes in bone parameters, including a reduction in bone mass, a decrease in bone thickness, and an increase in bone porosity.[3,4] These changes signify a degradation of bone architecture. Additionally, research has demonstrated that the disease is often accompanied by chronic pain, severely impacting the patient's quality of life.[5] In recent years, the aging of the population has led to a rise of skeletal cases. Fractures caused by these diseases and additional complications not only affect the quality of life for individuals but also place a significant financial burden on public health care.[6,7] Therefore, routine detection plays an essential role in the diagnosis and treatment of skeletal diseases.

The World Health Organization (WHO) has recommended dual-energy x-ray absorptiometry (DXA) as the gold standard for measuring bone mineral density (BMD). The DXA quantifies the bone status based on BMD, but it may be insufficient in providing information on bone geometry or mechanical properties,[8,9] Additionally, high spatial resolution imaging techniques such as micro-computed tomography (micro-CT) and magnetic resonance imaging (MRI) can acquire detailed geometric structures of the bone. Micro-CT carries the risk of ionizing radiation, and the high cost of MRI equipment makes it less suitable for routine detection use.[10] Ultrasound imaging, characterized by its non-invasive nature, cost-effectiveness, and portability, is recognized as a widely used method for the medical diagnosis.[11,12] Currently, ultrasonic beamforming imaging is mainly based on a uniform assumption for speed of sound (i.e., 1500 m/s) in biological tissues.[13] However, bone is a complex tissue with a big vari-

†Corresponding author. E-mail: chengchengliu@fudan.edu.cn

© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

ation in the speed of sound (SOS) distribution. Ultrasonic beamforming imaging shows limitations in displaying the microstructure of the bone.[14]

Ultrasound computed tomography (USCT) is a potential method for imaging complex medium. The USCT employing inverse algorithms associated with the wave equation to reconstruct ultrasonic SOS across the measured region. The travel-time tomography inversion method utilizes a ray approximation of the wave equation, focusing solely on refraction. Consequently, the resolution of travel-time tomography is constrained by the first Fresnel zone.[15] The diffraction tomography (DT) method, based on the Born approximation of the wave equation, takes into account diffraction and scattering,[16] but fails to image quantitative information, such as the distribution of SOS, particularly for bone tissues.[17] The Full Waveform Inversion (FWI) using the full wave equation and considers higher-order diffraction and scattering of the wave equation.[18] Studies showed that FWI can provide promising reconstruction of SOS for musculoskeletal tissues.[19–21] However, FWI iteratively updates the SOS distribution by minimizing the difference between the simulated data and the observed data. The FWI is an ill-posed inverse problem and has a high requirement on the initial model during the imaging process. An inaccurate initial model may lead to local minima in the inversion and unexpected imaging results caused by circle-skipping phenomenon.[19,22] An effective improvement is the application of a multi-scale frequency band inversion strategy.[23] Inversion is performed from low to high frequencies, which effectively avoids cycle jumping and enhances the stability of the process.

Additionally, in order to simplify the problem, most FWI methods utilize the constant-density acoustic wave equation, which only considers longitudinal wave propagation in the medium and is insufficient to be consistent with the actual situation. In recent years, studies have increasingly applied the elastic wave equation in FWI to achieve joint inversion of SOS and density, and some preliminary results have been obtained.[24,25] However, due to parameter coupling, the joint inversion strategy is relatively complex. Overall, achieving multi-parameter inversion in medical ultrasound remains challenging.

Deep learning (DL) is experiencing a surge in development, with applications across various domains consistently demonstrating state-of-the-art performance. The field of ultrasonic imaging also has witnessed a multitude of applications.[26–28] Liu *et al*. had employed an improved diffusion model to achieve super-resolution (SR) enhancement in ultrasound imaging.[26] Shi *et al*. achieved the automatic extraction of time-of-flight (TOF) by utilizing a convolutional neural network-based segmentation method.[27] Suo *et al*. implemented data-driven FWI for quantitative ultrasound bone imaging by utilizing an enhanced dual-encoder approach.[28] The success of these researches can be attributed to the extensive development of universal programming frameworks for DL, such as PyTorch[29] and TensorFlow.[30] However, the implementation of the aforementioned methods dominantly relies on supervised learning, which necessitates the prior acquisition of massive datasets. Meanwhile, to our best knowledge, current deep learning methods have not focused on the study of multi-parameter imaging in medium. Moreover, in practical application scenarios like ultrasonic musculoskeletal imaging, it is not always feasible to obtain a substantial amount of high-quality data, or the data acquisition process can be challenging.

Recent studies have proposed the concept of physics-informed embedding, which holds promise for the improvement of deep learning performance and interpretability.[31] Generally, this approach integrates human observation, experience, and prior understanding of physics or mathematics into data augmentation, loss functions, and architecture within deep learning to enhance the performance of learning algorithms. For instance, some physics-informed DL frameworks have been provided for solving the problems associated with nonlinear partial differential equations in mathematical applications.[32,33] In the domain of acoustics, a physics-informed neural network (PINN) method has facilitated the forward solution of wave equations.[34] In the field of non-destructive testing, an optimized physics-informed method has enabled the quantitative identification of surface crack defects in metal plates.[35] In terms of biomolecular imaging, a method based on embedding measurement model physical information has been employed to achieve protein structure reconstruction.[36] The PINN method has also solved wave propagation problems and achieved geophysical SOS imaging in the field of geophysics.[37,38] Although PINN has made preliminary progress in various fields, multi-parameter imaging of medium has still not been realized in the field of medical ultrasound.

Inspired by these PINN applications and incorporating the acoustic wave equation, we propose a method called Bone Ultrasound Imaging with Physics Informed Generative Adversarial Network (BUIPIGAN) to achieve unsupervised multi-parameter imaging of musculoskeletal tissue using ultrasound, focusing on SOS and density. The BUIPIGAN consists of two main parts, the generator and the discriminator. The generator optimizes the current distributions of SOS and density, producing generated data based on the physics-informed acoustic wave equation. The discriminator uses neural network composed of depthwise separable convolution (DSC) blocks, residual connections, and convolution blocks to quantify the difference between the generated data and the observed data. At the same time, the discriminator provides feedback to the

generator to optimize the SOS and density of musculoskeletal tissue. To the best of our knowledge, this work represents the first report of utilizing an unsupervised neural network with embedded physical information for multi-parameter imaging of musculoskeletal tissue. The main contributions of this work can be summarized as the following:

(i) A deep learning framework based on generative adversarial network, BUI-PIGAN, is proposed for multi-parameter imaging of musculoskeletal tissue.

(ii) A module based on depthwise separable convolution and residual connection is proposed to enhance feature extraction and filtering capabilities of the discriminator.

(iii) A multi-parameter optimization generator is presented to achieve high-quality ultrasound imaging of SOS and density.

(iv) The experiments on two categories of musculoskeletal models, tibia–fibula model and human forearm model, are implemented to evaluate the performance. The experimental results demonstrate that the proposed framework can achieve the high-accuracy and robust multi-parameter imaging of musculoskeletal tissues.

The structure of this paper is outlined as follows. The second section provides the theoretical principles of BUI-PIGAN. The third section details the in-silico experiments and parameter settings. The fourth section presents the analysis of the results. The fifth section discusses the highlights and limitations. Finally, the concluding section summarizes this research.

2. Method

This study primarily focuses on the multi-parameter imaging for musculoskeletal tissue and aims to reconstruct the SOS and density from the observed data. The reconstruct process is obviously an ill-posed inverse problem. There are two common methods can be used for this type of problem. The first method (i.e., traditional FWI) is a regularization-based approach that uses multiple approximability constraints to make the model convergence. The FWI method is mathematically interpretive, but the inversion process tends to fall into local minima, i.e., the cycle-skipping phenomenon.[39] The second method (i.e., deep learning) is a learning-based approach that implements nonlinear mapping through training. It has achieved a success across various application scenarios but is limited by the dataset sometime. The physics-informed networks combine the advantages of high interpretability and strong nonlinear mapping ability. Therefore, the physics-informed generative adversarial network might have potential for multi-parameter reconstruction in musculoskeletal ultrasound imaging. The BUIPIGAN method is based on an iterative learning approach embedding with physical information and does not require massive dataset.

2.1. Wasserstein generative adversarial network with gradient penalty (WGAN-GP)

Generative adversarial networks (GANs) possess the advantage of adversarial learning compared to other generative algorithms. The essence of adversarial learning is that the generator produces data distributed identically to real data, making it indistinguishable to the discriminator. Meanwhile, the discriminator strives to differentiate between real and generated data. Both improve their performance through the process of adversarial learning. This method can be directly parallelized for generation on a single large image, and the design of the generator is highly flexible. It is capable of accommodating various tasks.[40] For the original GANs, both the generator and the discriminator are essentially neural networks. The objective function of this learning task can be formulated as follows:

$$\min_G \max_D V(D,G) = E_{x\sim \mathrm{P}_{\mathrm{real}}(x)}\left[\log\left(D(x)\right)\right] + E_{z\sim P_n(z)}\left[\log\left(1-D\left(G(z)\right)\right)\right], \quad (1)$$

where G and D represent the mapping of the generator and the discriminator, respectively. $P_{\mathrm{real}}(x)$ is the probability distribution of the real data and $P_n(z)$ is the input data distribution of G. Through adversarial training, the generator is enabled to produce data that reflects the distribution of real data.

Due to certain difficulties encountered during the training process of the original GANs, the results can be unstable. Additionally, adjusting the hyperparameters can make the model difficult to converge, presenting other challenges.[41] To enhance the controllability of model training and achieve superior performance, researchers have proposed Wasserstein GAN (WGAN).[42] This network utilizes the Wasserstein distance to measure the difference between two distributions, rather than relying on the Kullback–Leibler (KL) divergence or the Jensen–Shannon (JS) divergence. Meanwhile, in order to avoid the gradient explosion or vanishing issues caused by gradient clipping, the WGAN-GP method has been proposed, which incorporates a gradient penalty operation under the condition of satisfying Lipschitz constraints.[43] A detailed derivation of the methodology can be found in Ref. [40].

2.2. The principle of SOS imaging with BUI-PIGAN

Incorporating the principles of WGAN-GP, the physical information of ultrasound propagation within the medium is embedded into the generator to realize data generation. In order to validate the effectiveness of the BUIPIGAN method with low computational cost, a two-dimensional (2D) constant-density acoustic wave equation in the medium is utilized below:

$$\frac{1}{v(\boldsymbol{r})^2}\frac{\partial^2 u(\boldsymbol{r},t)}{\partial t^2} = \nabla^2 u(\boldsymbol{r},t) + s(\boldsymbol{r},t)\,\delta(\boldsymbol{r}-\boldsymbol{r}_s), \quad (2)$$

where $\boldsymbol{r} = M \times N$ denotes the discrete medium grid coordinates, t represents the temporal coordinate, $v(\boldsymbol{r})$ signifies the SOS distribution defined by the coordinate $\boldsymbol{r}$, $u(\boldsymbol{r},t)$ denotes the wavefield amplitude at position $\boldsymbol{r}$ at time t, and $s(\boldsymbol{r},t)$ indicates a source term constrained by a Dirac function at position $\boldsymbol{r}_s$, with ∇^2 representing the Laplacian operator.

The ultimate task of SOS reconstruction can be expressed by the definition of Wasserstein distance as follows:

$$\begin{aligned} \boldsymbol{v}_{\text{pre}} &= \arg\min_{v} \inf_{\gamma \in \Pi\left(P_{\text{gen}}, P_{\text{true}}\right)} E_{\left(\boldsymbol{u}_{\text{gen}}, \boldsymbol{u}_{\text{true}}\right) \sim \gamma} \\ &\quad \times \left[\| \boldsymbol{u}_{\text{gen}} - \boldsymbol{u}_{\text{true}} \| \right], \end{aligned} \tag{3}$$

where $\boldsymbol{v}$ denotes the discretized SOS distribution, $\boldsymbol{v}_{\text{pre}}$ is the result of reconstruction, $\boldsymbol{u}_{\text{gen}}$ represents the wavefield amplitude data signal produced by the generator using the current SOS distribution, $\boldsymbol{u}_{\text{true}}$ is the real data signal received by the transducer using the true SOS distribution $\boldsymbol{v}_{\text{true}}$. P_{gen} is the probability function $P(\boldsymbol{u}_{\text{gen}}|\boldsymbol{v})$ of the generated data and P_{true} is the probability function $P(\boldsymbol{u}_{\text{true}}|\boldsymbol{v}_{\text{true}})$ of the observed data. The $\Pi\left(P_{\text{gen}}, P_{\text{true}}\right)$ denotes the set of adjoint distributions. E represents the weighted average of the distance function $\| \boldsymbol{u}_{\text{gen}} - \boldsymbol{u}_{\text{true}} \|$ under the joint distribution γ. Specifically, it reflects the average value of the minimum transportation cost required to move from distribution P_{gen} to distribution P_{true}. This imaging process involves adjusting the current SOS distribution $\boldsymbol{v}$ to minimize the Wasserstein distance between the two distributions. The adjusted SOS distribution is used as the final result.

For the SOS imaging of BUI-PIGAN, the objective function for adversarial learning of generators and discriminators can be described as follows:

$$\begin{aligned} &\min_{G_{\text{P}}} \max_{D\phi} V\left(D_\phi, G_{\text{P}}\right) \\ &= E_{\boldsymbol{u}_i \sim P_{\text{true}}}\left[D_\phi\left(\boldsymbol{u}_i\right)\right] - E_{\boldsymbol{u}_j \sim P_{\text{gen}}}\left[D_\phi\left(\boldsymbol{u}_j\right)\right] \\ &\quad + \eta E_{\boldsymbol{u}_{\text{s}} \sim P_{\text{uni}}}\left[\left(\left\|\nabla_{\boldsymbol{u}_{\text{s}}} D_\phi\left(\boldsymbol{u}_{\text{s}}\right)\right\|_2 - 1\right)^2\right], \end{aligned} \tag{4}$$

where G_{P} denotes the physics-informed generator that embeds the 2D constant-density acoustic wave equation, while D_ϕ signifies the discriminator with neural network parameters represented by ϕ. The first two terms correspond to the objective function of the WGAN. The final term is the gradient penalty term with η as the hyperparameter. P_{uni} represents the uniform distribution along the straight line between sampled points from P_{true} and P_{gen}, i.e., $u_{\text{s}} = \theta \boldsymbol{u}_{\text{true}} + (1 - \theta \boldsymbol{u}_{\text{gen}})$, where θ is uniformly sampled between 0 and 1.

2.3. The principle of multi-parameter imaging with BUI-PIGAN

In practical ultrasound imaging, in addition to the SOS parameter, the medium has other physical properties that affect the propagation of ultrasound waves, such as density, impedance, and modulus. Therefore, relying solely on SOS parameters for imaging is difficult to fully and accurately describe the physical information inside the medium. Previous studies have shown that SOS-density parametric models are widely applied and demonstrate well applicability and effectiveness.[25,44] In order to achieve multi-parameter imaging of SOS and density, the 2D elastic acoustic wave equation in the medium is utilized. In the 2D coordinate system (x–z plane), assuming the medium is isotropic, the elastic wave equation can be described using the momentum conservation equation and the stress-strain relationship. The momentum conservation equation can be formulated as

$$\begin{aligned} \boldsymbol{\rho} \frac{\partial^2 \boldsymbol{u}_x}{\partial t^2} &= \frac{\partial \boldsymbol{\sigma}_{xx}}{\partial x} + \frac{\partial \boldsymbol{\sigma}_{xz}}{\partial z}, \\ \boldsymbol{\rho} \frac{\partial^2 \boldsymbol{u}_z}{\partial t^2} &= \frac{\partial \boldsymbol{\sigma}_{xz}}{\partial x} + \frac{\partial \boldsymbol{\sigma}_{zz}}{\partial z}. \end{aligned} \tag{5}$$

Here, $\boldsymbol{\rho}$ represents the density of the medium. $\boldsymbol{u}_x$ and $\boldsymbol{u}_z$ represent the displacement field components in the x and z directions, respectively. $\boldsymbol{\sigma}_{xx}$, $\boldsymbol{\sigma}_{xz}$, $\boldsymbol{\sigma}_{zz}$ are the stress components, corresponding to the normal stress in the x direction, the shear stress, and the normal stress in the z direction, respectively.

The stress–strain relationship can be formulated as

$$\begin{aligned} \boldsymbol{\sigma}_{xx} &= \lambda \left(\frac{\partial \boldsymbol{u}_x}{\partial x} + \frac{\partial \boldsymbol{u}_z}{\partial z} \right) + 2\mu \frac{\partial \boldsymbol{u}_x}{\partial x}, \\ \boldsymbol{\sigma}_{zz} &= \lambda \left(\frac{\partial \boldsymbol{u}_x}{\partial x} + \frac{\partial \boldsymbol{u}_z}{\partial z} \right) + 2\mu \frac{\partial \boldsymbol{u}_z}{\partial z}, \\ \boldsymbol{\sigma}_{xz} &= \mu \left(\frac{\partial \boldsymbol{u}_x}{\partial z} + \frac{\partial \boldsymbol{u}_z}{\partial x} \right). \end{aligned} \tag{6}$$

Here, λ is the first Lamé parameter, and μ is the second Lamé parameter (shear modulus). The SOS for longitudinal waves can be expressed as follows:

$$\boldsymbol{v} = \sqrt{\frac{\lambda + 2\mu}{\boldsymbol{\rho}}}. \tag{7}$$

Combining Eqs. (3) and (4), by incorporating both SOS and density into the optimization process of the generator, the multi-parameter reconstruction task using the BUIPIGAN can be described as follows:

$$\begin{aligned} &\min_{G_{\text{P}}} \max_{D\phi} L\left(D_\phi, G_{\text{P}}\right) \\ &= E_{\boldsymbol{u}'_i \sim P'_{\text{true}}}\left[D_\phi\left(\boldsymbol{u}'_i\right)\right] - E_{\boldsymbol{u}'_j \sim P'_{\text{gen}}}\left[D_\phi\left(\boldsymbol{u}'_j\right)\right] \\ &\quad + \eta E_{\boldsymbol{u}'_{\text{s}} \sim P'_{\text{uni}}}\left[\left(\left\|\nabla_{\boldsymbol{u}'_{\text{s}}} D_\phi\left(\boldsymbol{u}'_{\text{s}}\right)\right\|_2 - 1\right)^2\right]. \end{aligned} \tag{8}$$

In contrast to Eq. (4), the optimization task $L\left(D_\phi, G_{\text{P}}\right)$ involves the optimization of both SOS and density. Here, P'_{true} is the probability distribution of the displacement field under the true SOS and density conditions. P'_{gen} is the probability distribution of the displacement field produced by the physics-informed generator. P'_{uni} represents the uniform distribution

along the straight line between sampled points from P'_{true} and P'_{gen}.

The schematic diagram of the BUIPIGAN is shown in Fig. 1. The overall structure is based on the WGAN-GP as the baseline. Unlike traditional approaches, the discriminator's objective is to quantify the difference between the generated data and the observed data, aiming to maximize the distinction between the two by adjusting the parameters of the discriminator network. The physics-informed generator (G_P) minimizes the difference between the generated data and the observed data by optimizing the SOS and density distribution with feedback from the quantified gradients of the discriminator. Through adversarial learning and optimization, the method ultimately produces excellent and stable multi-parameter (SOS and density) imaging results.

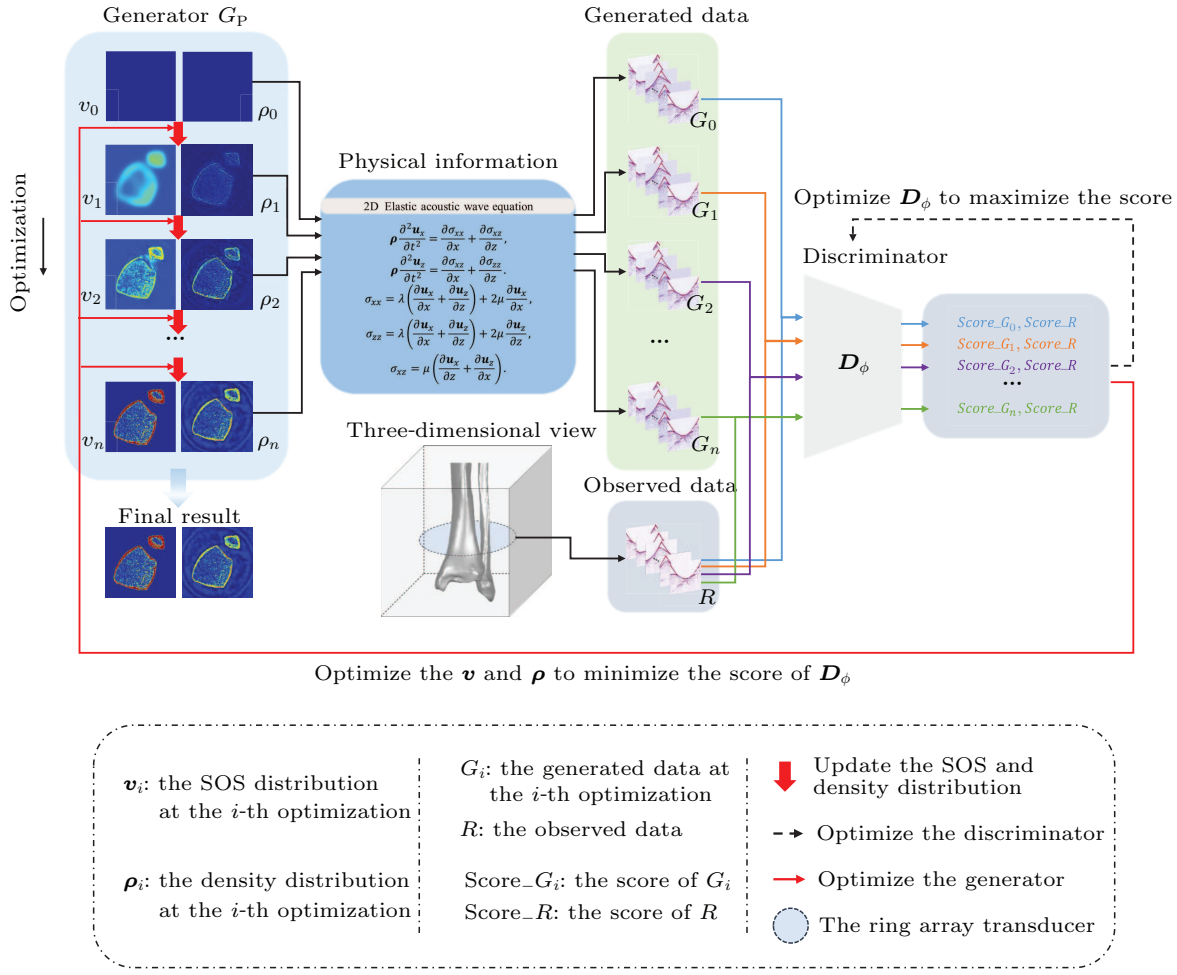


Fig. 1. Schematic diagram of the multi-parameter BUIPIGAN method.

2.4. The physics-informed generator of BUI-PIGAN

In SOS imaging, the generator embedded physical information from the 2D constant-density equation and produced the generated data based on the SOS distribution. For multi-parameter imaging, the generator embedded the 2D elastic wave equation, producing the generated data based on both the SOS and density distributions. Both the SOS and density distributions were discretized into a 210×210 grids, with a spatial length of 0.5×0.5 mm per grid cell, covering a region of 10.5×10.5 cm. This part was implemented with a sampling precision of 20 MHz under the condition of satisfying CFL (Courant–Friedrichs–Lewy). Perfectly matched layer was applied in the simulation. To test the performance with noise, certain signal-to-noise ratio (SNR) generated data with noise was obtained by adding Gaussian noise to the noise-free data.

2.5. The discriminator network of BUI-PIGAN

To enable the generator to effectively learn the distribution of real observed data, the discriminator must accurately quantify the differences between the generated data and the observed data, as well as distinguish them as much as possible. However, an overpowered discriminator might make it difficult for the generator to obtain effective gradient feedback, causing the generator to fail to produce data that is consistent with the actual distribution. Therefore, the enhancement of the discriminator's capability must be balanced with the optimization of the generator to ensure the stability of adversarial learning.[42] Many studies use standard convolutional neu-

ral networks (CNNs) for feature extraction in generative adversarial networks (GANs). Some have introduced attention mechanisms,[45] but this significantly increases computational complexity and may disrupt the balance between the generator and discriminator.

In this study, depthwise separable convolution (DSC) and residual connections are used to enhance the network structure of the discriminator, improving the effectiveness of adversarial learning without introducing high computational complexity. DSC decomposes standard convolutions into depthwise and pointwise operations, reducing computational cost and parameters while preserving feature extraction ability, thus improving model efficiency.[46] Additionally, residual connections help maintain gradient flow and prevent vanishing gradients in deeper networks, thereby improving learning stability.[47] In this way, the discriminator is able to provide more effective feedback to the generator, enhancing optimization and improving model performance.

The structure is shown in Fig. 2. The DSC block consists of convolutional layers matching the number of channels (size 7×7, padding size 3×3), two 1×1 convolutional layers, a GELU (Gaussian Error Linear Unit) activation, and a skip connection. The convolutional block consists of a convolutional layer (size 3×3, padding size 1×1), a max pooling layer (size 2×2), and a leaky RELU (Rectified Linear Unit negative slope 0.1). The fully connected layer first flattens the output of the last convolutional block into a 2000-dimensional feature vector, and then outputs this feature vector as a scalar that quantifies the difference.

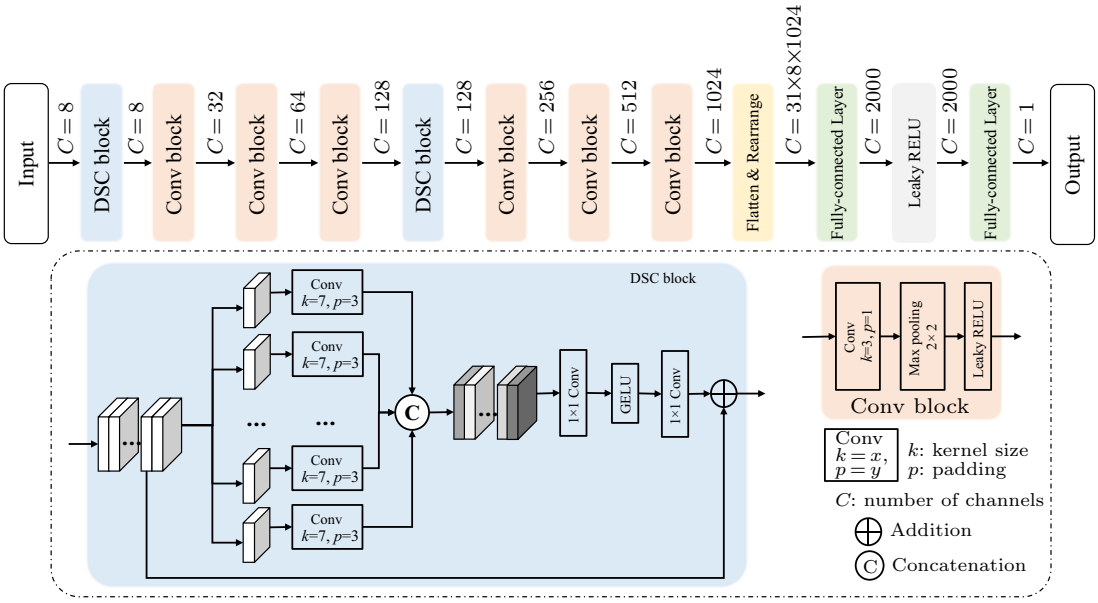


Fig. 2. Network of the discriminator in the BUI-PIGAN.

3. Experimental design and parameter configuration

3.1. Musculoskeletal tissue and ring array transducer setup

This study employed two different types of musculoskeletal tissue models. The first model was a tibia–fibula model. The structure of tibia–fibula was obtained using high-resolution peripheral quantitative computed tomography (HR-pQCT). In this model, the soft tissues were assumed to be uniformly distributed, and only the SOS and density were defined. The SOS and density values of the tibia–fibula were obtained by the grayscale mapping of CT image. The SOS values was ranged from 1600 m/s to 2800 m/s. The density values was ranged from 1000 kg/m^3 to 2000 kg/m^3.

This study compared several SOS imaging methods. These methods included conventional FWI (C-FWI), conventional FWI with an improved accurate initial model (C-FWI-I), an enhanced FWI method using a multiscale-frequency inversion strategy (FWI-MS),[23] and FWIGAN,[38] a newly emerging method in geophysics. For multi-parameter imaging, FWI-MP,[25] a joint inversion FWI method for SOS and density, was also used for comparison. Since these methods required an initial parameter distribution, all methods except C-FWI-I and FWI-MP used a constant SOS and density distribution. C-FWI-I and FWI-MP incorporated the accurate initial SOS distribution as their initial parameter distribution. As shown in Fig. 3, The accurate initial SOS distribution was obtained by applying Gaussian smoothing to the true SOS distribution (Fig. 3(c)). The constant SOS distribution was set to 1600 m/s (Fig. 3(b)), and the constant density distribution was set to 1080 kg/m^3 (Fig. 3(e)).

In order to more accurately represent the actual ultrasound imaging object, the human forearm model was used as the second musculoskeletal tissue model to demonstrate

the imaging performance of the multi-parameter BUIPIGAN method. This model was obtained using 5T-MRI and includes bone tissue and various soft tissues. As shown in Fig. 4. and Table 1, different SOS and density values were assigned based on the actual properties of each tissue. For comparison, the FWI-MP method used an accurate initial SOS distribution, which was extracted by contouring the muscle regions (Fig. 4(c)). The constant SOS distribution was set to 1500 m/s (Fig. 4(b)), and the constant density distribution was set to 1000 kg/m^3 (Fig. 4(e)).

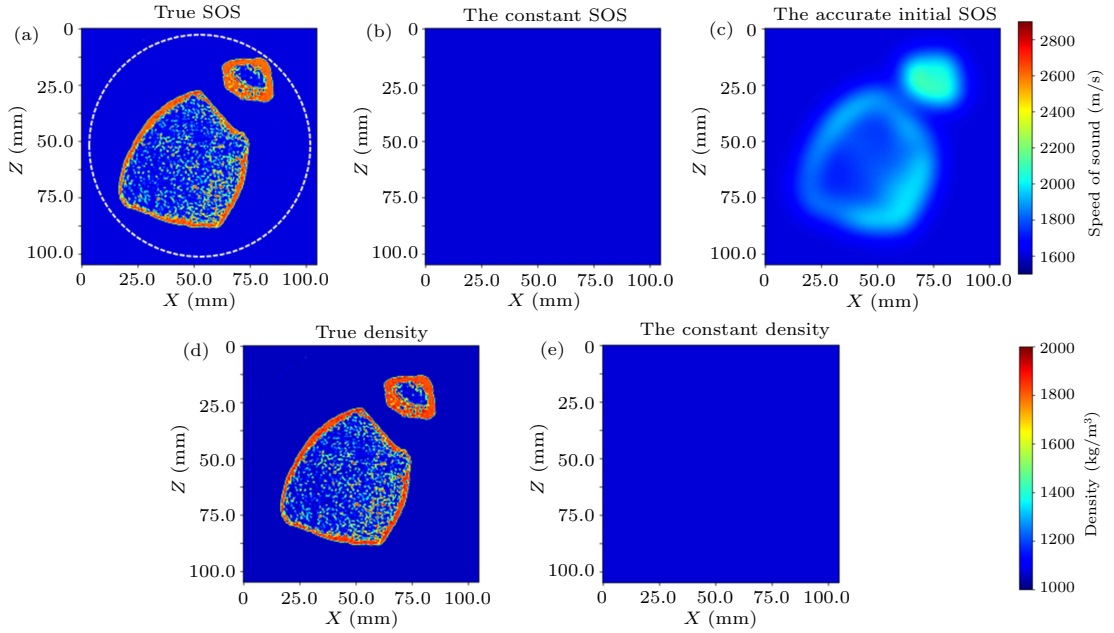


Fig. 3. Tibia–fibula musculoskeletal tissue model. (a) True SOS distribution of the tibia–fibula model. (b) An initial distribution with a constant SOS of 1600 m/s. (c) An accurate initial SOS distribution obtained through Gaussian smoothing of the true SOS. (d) True density distribution of the tibia–fibula model. (e) An initial distribution with a constant density of 1080 kg/m^3.

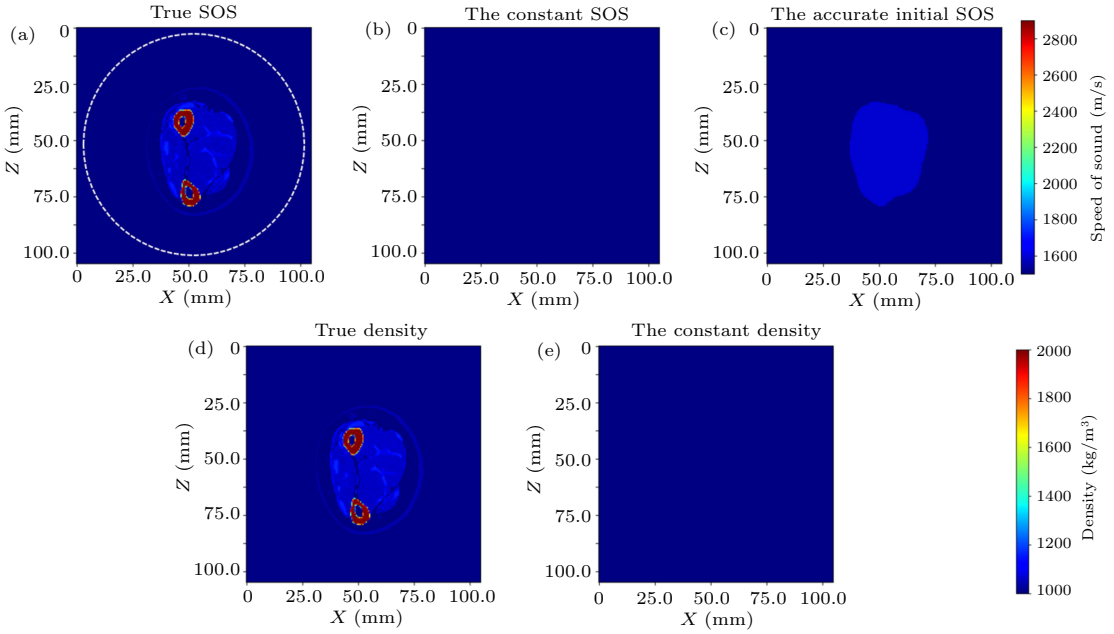


Fig. 4. Human forearm musculoskeletal tissue model. (a) True SOS distribution of the human forearm model. (b) An initial distribution with a constant SOS of 1500 m/s. (c) An accurate initial SOS distribution obtained by contouring the muscle regions. (d) True density distribution of the human forearm model. (e) An initial distribution with a constant density of 1000 kg/m^3.

Table 1. Acoustic properties of different tissues.[48,49]

Tissue type	SOS (m/s)	Density (kg/m^3)
Water	1500	1000
Fat	1480	950
Muscle	1560	1080
Tendon	1680	1200
Bone	2900	2000

The ring array transducer consists of 128 elements which could be used both as emitter or receiver. The observed data are acquired in the full-matrix capture model through numerical simulation. The diameter of the ring array is 10 centimeters. Considering the size of the ring array and computational costs, this study employed 128 elements for the experiments.

A 0.5 MHz Ricker wave is utilized as the source signal for tibiofibular and human forearm imaging. The signal and its spectrum from the sound source are illustrated in the Fig. 5.

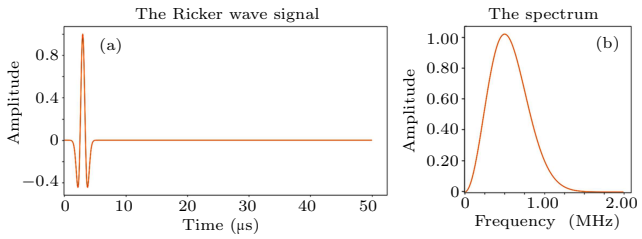


Fig. 5. Ricker wave source: (a) time-domain signal, (b) amplitude spectrum with a central frequency $f_c = 0.5$ MHz.

3.2. Setting of learning parameters and platform configuration

The BUIPIGAN employed an adversarial learning strategy, indicating that the generator and discriminator were trained separately during the process. We adopt a policy of learning the discriminator 6 times before learning the generator once in a single update cycle. In order to accelerate computation and reduce runtime memory usage, data was fed in batches, with the observed data from 128 transducers being divided into 8 batches. Numerous trial results demonstrated that the size of the batch does not significantly affect the performance of the results when learning converges. The optimizer used Adam ($\beta_1 = 0.5$, $\beta_2 = 0.9$, $\varepsilon = 10^{-8}$). The entire implementation process utilized the Ubuntu operating system, Nvidia GPU computational resources, and the PyTorch deep learning framework. Gradients were computed using backward automatic differentiation, and hyperparameters were fine-tuned for optimal performance. Specifically, the physics-informed generator used learning rates of 30 and 10 to optimize the SOS and density distribution, respectively. Meanwhile the discriminator used a learning rate of 10^{-3} to optimize network parameters, with the gradient penalty weight set at $\eta = 10$. In the noise experiments, the SNR of the noise is also optimized through the gradients computed by the discriminator, with the learning rate set to 10^{-2}.

3.3. Quantitative metrics

In addition to visual comparison, three metrics were selected to quantify the similarity between the imaging results and the actual parameter distribution. The region of interest (ROI) for the quantitative comparison was located inside the ring array transducer. The structural similarity index measure (SSIM) was calculated as follows:

$$\begin{aligned}&\mathrm{SSIM}\left(\boldsymbol{y}_{\mathrm{pre}}, \boldsymbol{y}_{\mathrm{true}}\right)\\ &= \frac{\left(2\mu_{\boldsymbol{y}_{\mathrm{pre}}}\mu_{\boldsymbol{y}_{\mathrm{true}}} + c_1\right)\left(2\sigma_{\boldsymbol{y}_{\mathrm{pre}}\boldsymbol{y}_{\mathrm{true}}} + c_2\right)}{\left(\mu^2_{\boldsymbol{y}_{\mathrm{pre}}} + \mu^2_{\boldsymbol{y}_{\mathrm{true}}} + c_1\right)\left(\sigma^2_{\boldsymbol{y}_{\mathrm{pre}}} + \sigma^2_{\boldsymbol{y}_{\mathrm{true}}} + c_2\right)},\end{aligned} \tag{9}$$

where $\boldsymbol{y}_{\mathrm{pre}}$ and $\boldsymbol{y}_{\mathrm{true}}$ denote the reconstructed result and the true distribution of the corresponding parameters, which can refer to SOS or density. The local means of these distributions are denoted by $\mu_{\boldsymbol{y}_{\mathrm{pre}}}$ and $\mu_{\boldsymbol{y}_{\mathrm{true}}}$, while their standard deviations are represented by $\sigma_{\boldsymbol{y}_{\mathrm{pre}}}$ and $\sigma_{\boldsymbol{y}_{\mathrm{true}}}$. The covariance between $\boldsymbol{y}_{\mathrm{pre}}$ and $\boldsymbol{y}_{\mathrm{true}}$ is given by $\sigma_{\boldsymbol{y}_{\mathrm{pre}}\boldsymbol{y}_{\mathrm{true}}}$, c_1 and c_2 are regularization stability constants.

The definition of SNR was also employed to describe the effectiveness of the reconstructed result relative to the true distribution

$$\mathrm{SNR}\left(\boldsymbol{y}_{\mathrm{pre}}, \boldsymbol{y}_{\mathrm{true}}\right) = 10\log_{10}\frac{\left\|\boldsymbol{y}_{\mathrm{true}}\right\|_2^2}{\left\|\boldsymbol{y}_{\mathrm{pre}} - \boldsymbol{y}_{\mathrm{true}}\right\|_2^2}. \tag{10}$$

In contrast to SNR, the relative error (RE) quantifies the numerical deviation between the reconstructed result and the true distribution. It is defined as

$$\mathrm{RE}\left(\boldsymbol{y}_{\mathrm{pre}}, \boldsymbol{y}_{\mathrm{true}}\right) = \frac{\left\|\boldsymbol{y}_{\mathrm{pre}} - \boldsymbol{y}_{\mathrm{true}}\right\|_F}{\left\|\boldsymbol{y}_{\mathrm{true}}\right\|_F}. \tag{11}$$

Here, $\|\cdot\|_F$ represents the Frobenius norm, the square root of the sum of the squared elements.

In the quantitative analysis of different results, SSIM measures the similarity between the imaging result and the true distribution with respect to structure, brightness, and contrast, reflecting the overall visual characteristics of the image. SNR indicates the effectiveness of the imaging result compared to the true distribution, reflecting the image quality. RE quantifies the deviation between the reconstruction result and the true distribution at both pixel and numerical levels, reflecting the accuracy of the reconstruction. Higher SSIM and SNR values, along with lower RE, signify the superior imaging performance.

4. Results

In this section, the performance of the BUIPIGAN in SOS imaging was firstly validated. The SOS imaging of the tibia–fibula was performed using five different methods, and the results were compared. Then, the multi-parameter imaging capabilities of BUIPIGAN were evaluated on the tibia–fibula model. Lastly, a human forearm model, which more consistent with real-world conditions, was used to further assess the effectiveness of BUIPIGAN in multi-parameter imaging. Additionally, the robustness of the BUIPIGAN was validated by using data with different noise levels in each part.

4.1. SOS imaging of the tibia–fibula

To demonstrate the performance of the proposed method, a comparison was made with the existing methods in terms of SOS imaging. As described in Subsection 3.1, four SOS imaging methods were used for comparison. The first method was the conventional FWI method, named as C-FWI in this study. The second method was an improvement of the conventional FWI by using an accurate initial SOS distribution as the initial

model to reduce cycle skipping, named as C-FWI-I. The third method was another improvement to reduce cycle skipping, the FWI method using a multiscale-frequency band inversion strategy, named as FWI-MS.[23] The fourth method was the recently emerging FWIGAN,[38] which achieved high-quality imaging in the field of geophysics. During the imaging process, except for the C-FWI-I method, which used an accurate initial SOS distribution (Fig. 3(c)), all other methods used a constant SOS distribution as the initial model (Fig. 3(b)).

As shown in Fig. 6 and Table 2, the C-FWI method tended to get stuck in local minima during the imaging process. This method successfully imaged only the boundary region of the cortical bone, while the cancellous bone area could not be effectively imaged, resulting in severe blurring. In contrast, the C-FWI-I and FWI-MS methods improved the imaging quality, especially in the cancellous bone region. Although the cancellous bone contour was imaged, there were still significant artifacts. The FWIGAN was able to effectively image the cancellous bone region and significantly improved image quality without introducing significant artifacts. However, the SOS values of FWIGAN still exhibited some deviations, especially in the cancellous bone region. Despite the reduction in artifacts, the FWIGAN failed to fully reflect the actual physical properties accurately. In comparison, the proposed method in this study, the BUIPIGAN, achieved high-quality imaging in both the cortical and cancellous bone regions without introducing noticeable artifacts. Furthermore, the BUIPIGAN effectively reduced deviations in SOS values, enhancing both the realism and accuracy of the image. Comparative analysis showed that BUIPIGAN achieved the best performance, outperforming other imaging methods, specifically, SSIM (0.9537), SNR (28.70 dB), RE (5.78%).

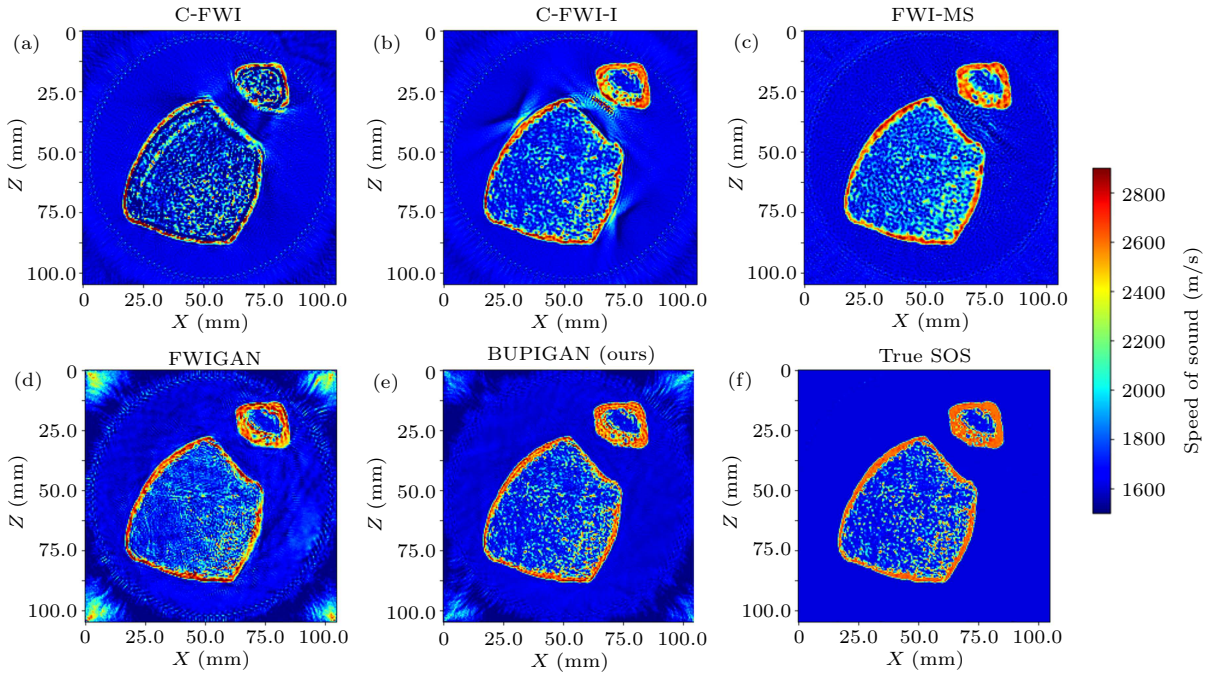


Fig. 6. The SOS imaging results of tibia–fibula obtained by different methods: (a) C-FWI, (b) C-FWI-I, (c) FWI-MS, (d) FWIGAN, (e) BUIPIGAN (ours), (f) true SOS distribution of tibia–fibula.

Table 2. Quantitative metrics of SOS imaging obtained by different methods.

Metric	C-FWI	C-FWI-I	FWI-MS	FWIGAN	BUIPIGAN
SSIM ↑	0.4957	0.9226	0.9350	0.9387	**0.9573**
SNR (dB) ↑	14.01	21.52	22.08	21.95	**28.70**
RE ↓	17.09%	7.39%	7.00%	7.60%	**5.78%**

The bold values are the best results.

To further evaluate the imaging performance of different methods on the tibia–fibula, quantitative analyses were conducted on the SOS values at horizontal positions of 75 mm (Fig. 7, $z = 75$ mm). The relative error between the imaging results of each method and the true SOS values was calculated. As shown in Fig. 7, the comparative results demonstrated that the proposed BUIPIGAN accurately reconstructed the SOS distribution of the tibia and fibula. Compared to other methods, the BUIPIGAN exhibited the smallest relative error. The SOS comparison results indicated that the BUIPIGAN not only improved the physical consistency of the imaging but also significantly reduced the deviation from the true SOS distribution.

To assess the robustness of the BUIPIGAN in SOS imaging, especially under noisy conditions, Gaussian white noise of different intensities was added to the noise-free ultrasound signals to simulate real imaging noise. The SNR of the received data was set to 30 dB and 10 dB, respectively. Figure 8 illustrates the multi-channel ultrasound signals at different noise levels.

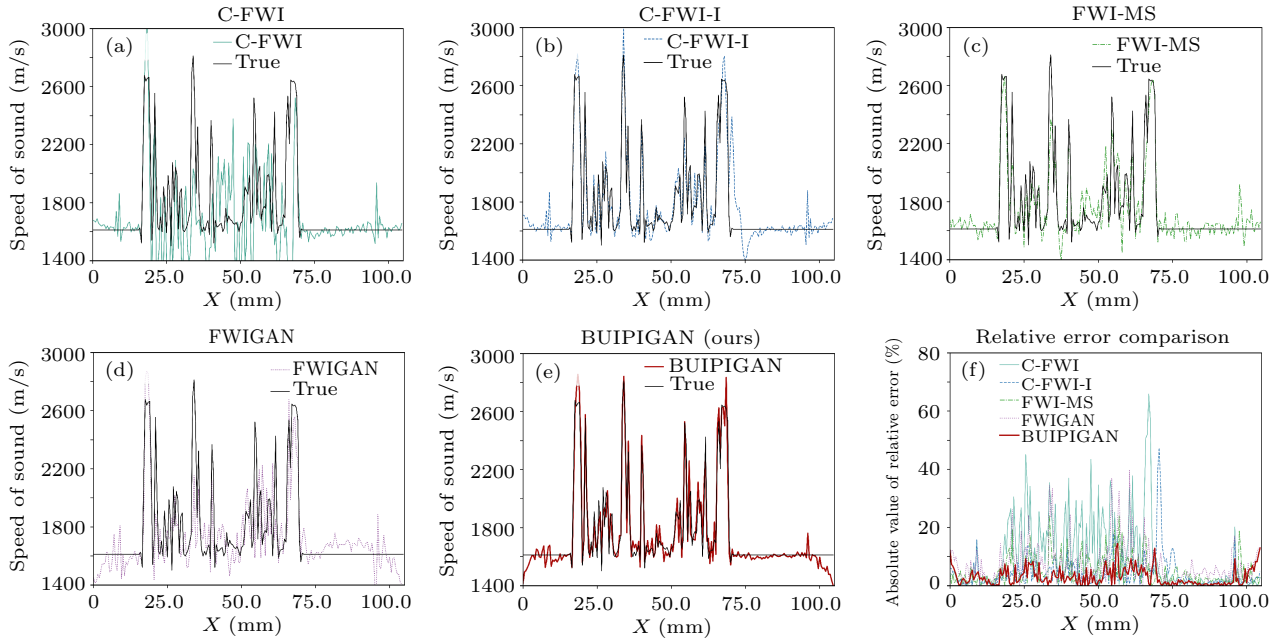


Fig. 7. Comparisons of profiles for SOS imaging results at $z = 75$ mm with the true SOS distribution using different methods: (a) C-FWI, (b) C-FWI-I, (c) FWI-MS, (d) FWIGAN, (e) BUIPIGAN (ours), (f) comparison of relative error for different methods.

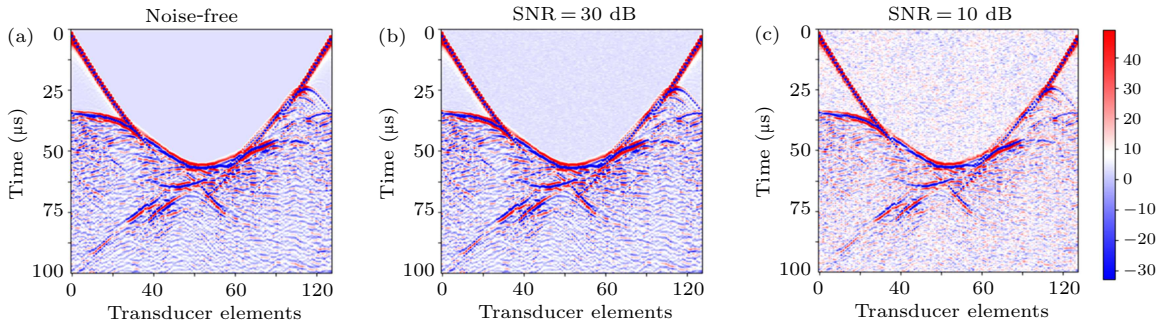


Fig. 8. The multi-channel ultrasound signals at different noise levels: (a) noise-free, (b) SNR = 30 dB, (c) SNR = 10 dB.

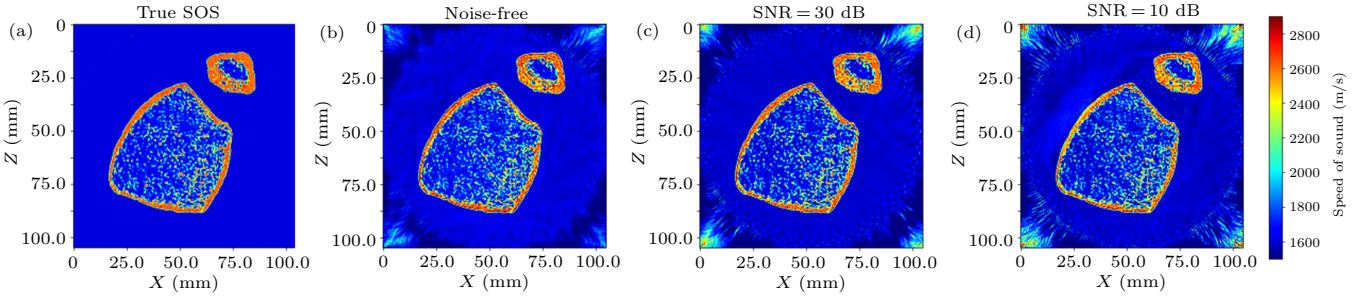


Fig. 9. The SOS imaging results of tibia–fibula at different noise levels using BUIPIGAN: (a) true SOS distribution of tibia–fibula, (b) noise-free, (c) SNR = 30 dB, (d) SNR = 10 dB.

Table 3. Quantitative metrics of SOS imaging for the tibia–fibula model at different noise levels.

Metric	Noise-free	SNR = 30 dB	SNR = 10 dB
SSIM ↑	0.9572	0.9552	0.9528
SNR (dB) ↑	28.70	28.52	27.75
RE ↓	5.77%	5.89%	6.28%

Figure 9 and Table 3 presented the SOS imaging results obtained with the BUIPIGAN under noisy conditions. Despite the noise interference, BUIPIGAN maintained high imaging accuracy without significant artifacts. These results demonstrated the robustness of the BUIPIGAN in noisy environments.

4.2. Multi-parameter imaging of the tibia–fibula

As described in Subsection 3.1, for comparative analysis, this study employed the FWI method based on SOS-density joint inversion, named as FWI-MP.[25] The proposed multi-parameter BUIPIGAN is denoted as BUIPIGAN-MP. FWI-MP used an accurate SOS distribution (Fig. 3(c)) and constant density (Fig. 3(e)) as the initial model, while BUIPIGAN-MP

used constant SOS (Fig. 3(b)) and constant density (Fig. 3(e)) as the initial model.

As shown in Fig. 10 and Table 4, the visual performance indicated that the FWI-MP method successfully imaged the tibia structure, but significant artifacts were present. And the fibula could not be effectively imaged, resulting in overall poor image quality. In contrast, the BUIPIGAN-MP accurately reconstructed the tibia and fibula structures, although the reconstruction deviation in density distribution was higher than in SOS distribution. This was because, in elastic medium, the influence of SOS on wave propagation was more significant than that of density. Therefore, during multi-parameter imaging with BUIPIGAN-MP, the SOS reconstruction was prioritized, leading to higher accuracy in SOS imaging compared to density imaging. For the quantitative metrics, the FWI-MP method reached more than 0.7 in terms of SSIM for both SOS and density inversion, indicating that the method could tolerably reflect the general distribution of the tibia and fibula. However, the SNR was below 20 dB, indicating that the effective signal in the images was limited, leading to poor image quality. Additionally, the RE exceeded 15%, indicating substantial deviation between the reconstructed values and the true values. In contrast, the BUIPIGAN-MP achieved more than 0.9 in terms of SSIM for both SOS and density distribution, especially 0.95 for SOS, demonstrating accurate structural reconstruction of the tibia and fibula. Furthermore, the SNR was above 20 dB, indicating superior image quality. The RE for SOS was 5.39%, indicating minimal deviation from the true values and high reconstruction accuracy. Although the RE for density was relatively higher, it remained below 8%, also indicating a small deviation from the true values.

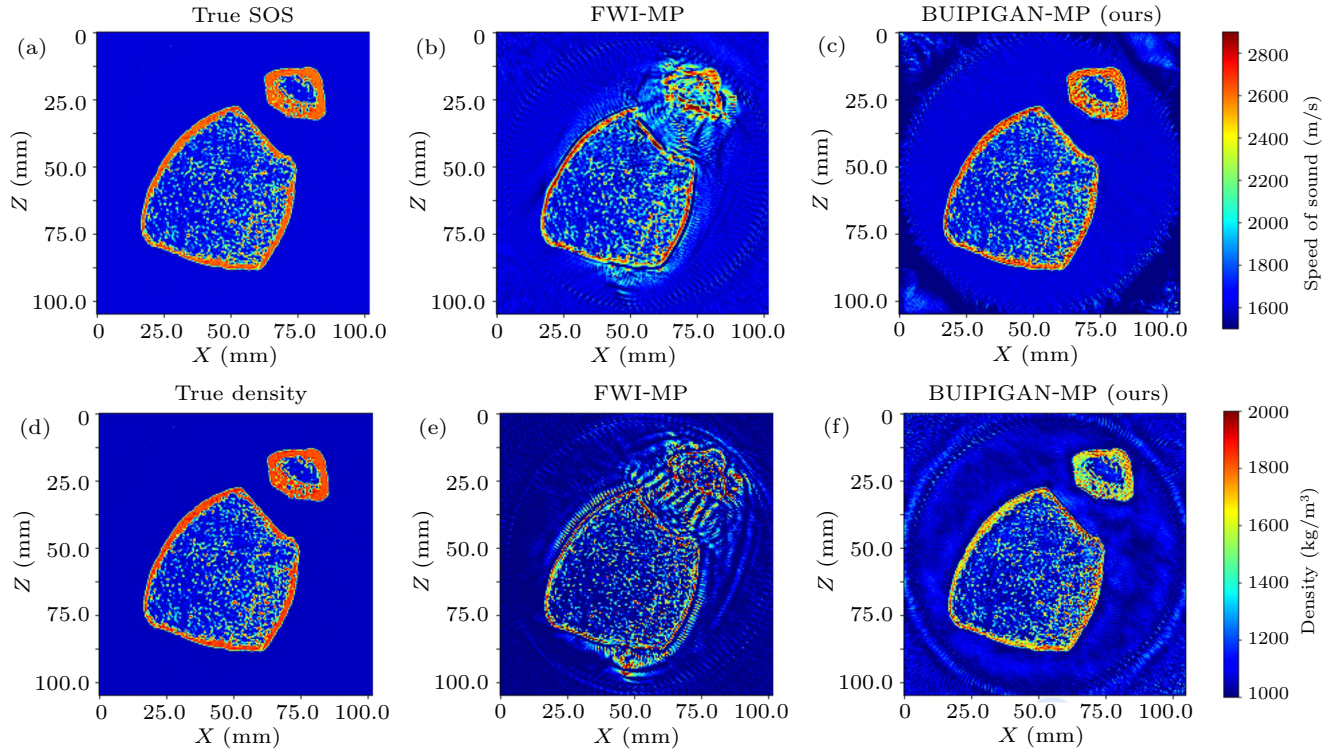


Fig. 10. The SOS and density imaging results of tibia–fibula obtained by different methods. (b) SOS of FWI-MP, (c) SOS of BUIPIGAN-MP (ours), (e) density of FWI-MP, and (f) density of BUIPIGAN-MP (ours).

Table 4. Quantitative metrics of SOS and density imaging for the tibia–fibula model obtained by different methods.

Metric	FWI-MP		BUIPIGAN-MP	
	SOS	Density	SOS	Density
SSIM ↑	0.7584	0.7257	**0.9530**	**0.9498**
SNR (dB) ↑	17.48	13.66	**27.15**	**22.31**
RE ↓	15.36%	20.75%	**5.39%**	**7.66%**

The bold values are the best results.

Overall, the FWI-MP method could reasonably reflect the structure of the tibia and fibula, but the image quality was poor, with significant deviations. In contrast, the BUIPIGAN-MP not only accurately reconstructed the tibia and fibula structures but also demonstrated significant advantages in both image quality and deviation, which could significantly improve imaging accuracy and quality.

Figure 11 compared the reconstructed SOS and density values with the true values and calculated the corresponding relative error. The results showed that the BUIPIGAN-MP significantly reduced the deviation between the reconstructed values and the true values. This indicated that the BUIPIGAN-MP effectively improved reconstruction accuracy in multi-parameter imaging tasks.

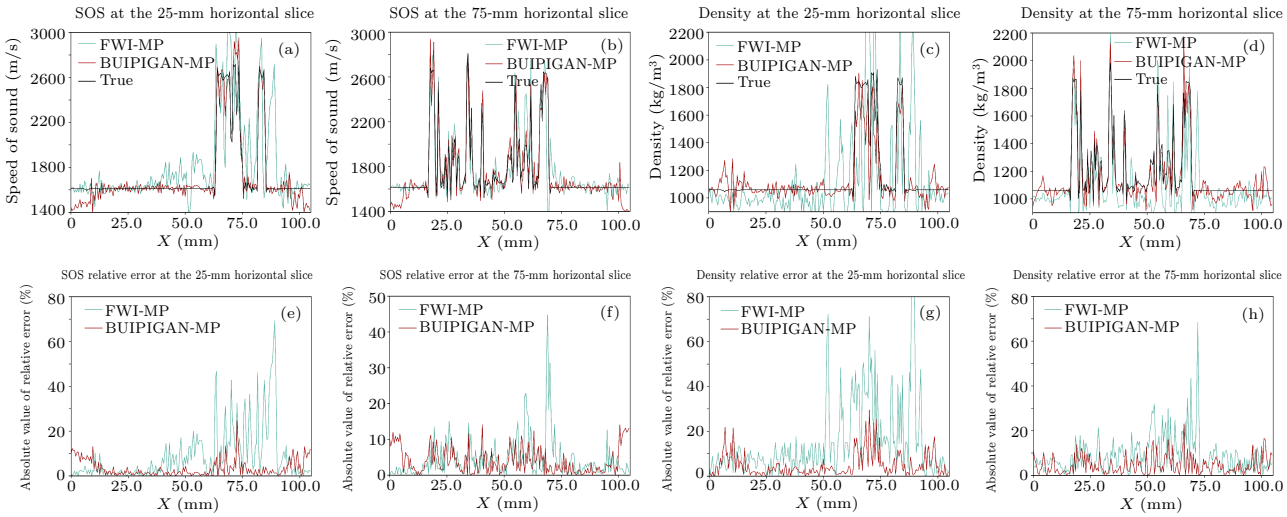


Fig. 11. (a)–(d) Comparisons of profiles for SOS and density imaging results at $z = 25$ mm and $z = 75$ mm with the true SOS and density distribution using different methods. (e)–(h) Comparison of relative error for SOS and density imaging results at $z = 25$ mm and $z = 75$ mm with the true SOS and density distribution using different methods.

4.3. Multi-parameter imaging of the human forearm

To more accurately represent the actual distribution of human musculoskeletal tissues in ultrasound imaging, the FWI-MP and BUIPIGAN-MP were applied to the human forearm model for multi-parameter imaging. In the FWI-MP method, an accurate SOS distribution (Fig. 4(c)) and a constant density distribution (Fig. 4(e)) were used as the initial model. In contrast, the BUIPIGAN-MP used constant SOS distribution (Fig. 4(b)) and constant density distribution (Fig. 4(e)) as the initial model.

Figure 12 and Table 5 presented the multi-parameter imaging results of the human forearm. A comparison of the imaging performance showed that the FWI-MP method exhibited significant reconstruction errors in high-SOS and high-density bone regions, failing to accurately reconstruct the details of musculoskeletal tissue. This problem is attributed to the fact that the FWI-MP method tends to fall into local minima when dealing with musculoskeletal tissue characterized by strong SOS and density differences, leading to significant deviations in the reconstruction results. In contrast, the BUIPIGAN-MP demonstrated higher accuracy in musculoskeletal tissue reconstruction and achieved the best performance in quantitative metrics.

Table 5. Quantitative metrics of SOS and density imaging for the human forearm model obtained by different methods.

Metric	FWI-MP		BUIPIGAN-MP	
	SOS	Density	SOS	Density
SSIM ↑	0.8873	0.8752	**0.9698**	**0.9498**
SNR (dB) ↑	20.62	19.30	**30.04**	**22.75**
RE ↓	10.32%	10.83%	**3.13%**	**8.17%**

The bold values are the best results.

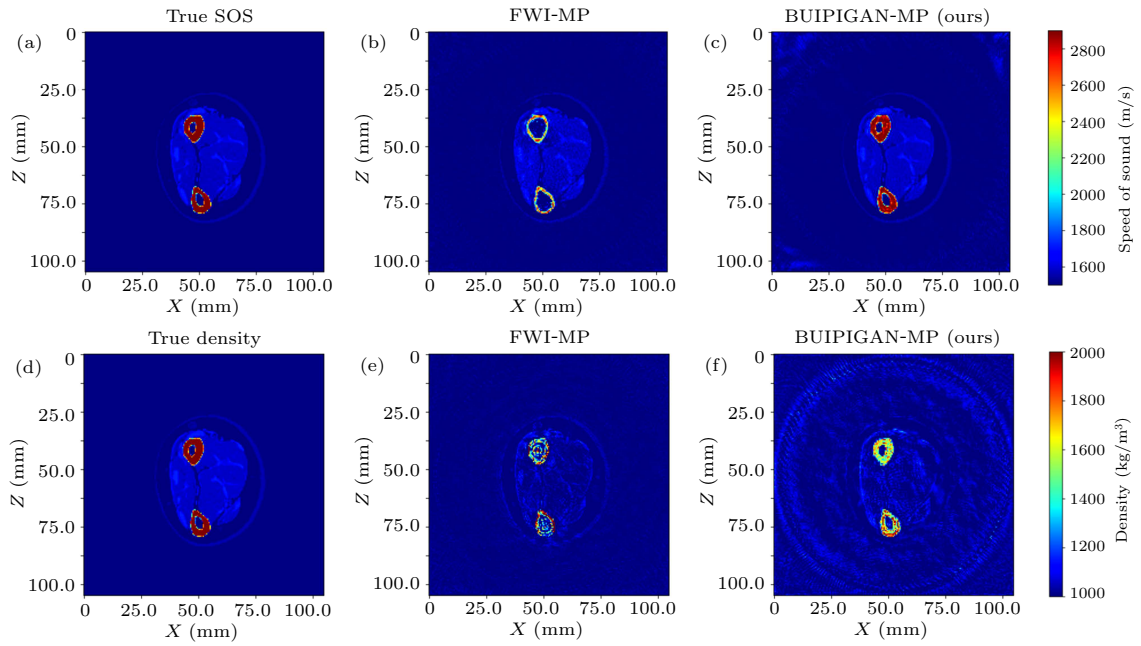


Fig. 12. The SOS and density imaging results of human forearm obtained by different methods. (b) SOS of FWI-MP, (c) SOS of BUIPIGAN-MP (ours), (e) density of FWI-MP, and (f) density of BUIPIGAN-MP (ours).

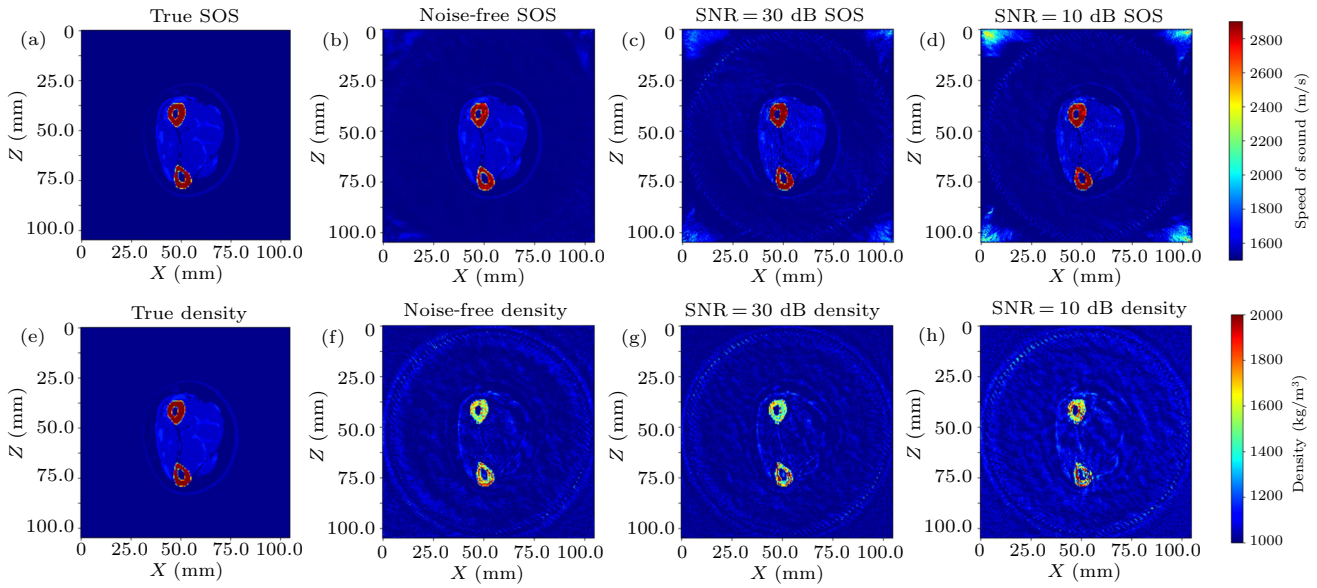


Fig. 13. The SOS and density imaging results of human forearm at different noise levels using BUIPIGAN. (a) True SOS distribution of human forearm. (b)–(d) The SOS imaging results at different noise levels. (e) True density distribution of human forearm. (f)–(h) The density imaging results at different noise levels.

Table 6. Quantitative metrics of SOS and density imaging for the human forearm model at different noise levels.

Metric	Noise-free		SNR = 30 dB		SNR = 10 dB	
	SOS	Density	SOS	Density	SOS	Density
SSIM ↑	0.9698	0.9498	0.9573	0.9453	0.9562	0.9434
SNR (dB) ↑	30.04	22.75	29.47	22.59	29.13	21.86
RE ↓	3.13%	8.17%	5.08%	8.85%	5.95%	9.38%

Figure 13 and Table 6 presented the multi-parameter imaging results of the BUIPIGAN-MP under different noise levels. Comparison of SOS imaging at different noise levels showed that noise had a limited impact on SOS reconstruction, with image quality remaining stable. This indicated that the BUIPIGAN-MP was robust in SOS reconstruction. The density reconstruction was more sensitive to noise. Despite this, the density distribution still accurately represented bone structure, and under high noise conditions (i.e., SNR = 10 dB), the SNR remained above 20 dB, with RE kept below 10%, indicating stable performance even in challenging noise environments.

5. Discussion

This study proposes the Bone Ultrasound Imaging Physical Information Generative Adversarial Network (BUIPIGAN) as an innovative multi-parameter ultrasound imaging approach. The BUIPIGAN effectively reduces the local minima problem through adversarial learning. By combining physical information with generative adversarial networks, BUIPIGAN is capable of achieving accurate multi-parameter imaging. On one hand, physical information enhances the interpretability and physical consistency of the imaging results. On the other hand, the powerful nonlinear mapping capability of GANs enables high-precision multi-parameter reconstruction simultaneously. In imaging of tibia–fibula and human forearm models, the BUIPIGAN significantly outperforms other methods, successfully achieving high-precision and stable multi-parameter SOS and density imaging.

In this work, firstly, embedding physical information into the generative adversarial network enhanced the physical accuracy of reconstruction. From embedding a 2D constant-density wave equation for SOS imaging to embedding a 2D elastic wave equation for multi-parameter imaging of SOS and density, the BUIPIGAN more accurately described ultrasound wave propagation, thus providing imaging results that are more consistent with the distribution of actual medium parameters. Secondly, the introduction of the DSC block in the discriminator not only improved algorithm performance but also significantly enhanced computational efficiency. Compared to the FWIGAN, the effective parameter size of the BUIPIGAN discriminator was reduced from 508.38 MB to 266.71 MB, leading to a substantial improvement in computational efficiency and further optimization of imaging results. Additionally, this study used more realistic tibia–fibula and human forearm models for multi-parameter imaging, achieving more realistic results compared to simplified models used in previous studies. Overall, the BUIPIGAN achieved high accuracy and robustness in multi-parameter imaging of musculoskeletal tissue, outperforming other methods in quantitative metrics.

Nevertheless, the BUIPIGAN still faces challenges with high storage and computational costs. Although the number of parameters in the discriminator has been reduced, the wavefield and gradient information still need to be computed and stored. Furthermore, this study employs ideal point sources and assumes that the actual ultrasound propagation conforms to the elastic wave equation, which still differs from real-world conditions. Future research could embed more realistic physical information (i.e., real transducer element size or ultrasound propagation equations consistent with the real-world conditions) for practical applications. Additionally, with advancements in computational hardware, extending the BUIPIGAN to 3D imaging or clinical applications holds great potential.

6. Conclusion

This study proposed a physical informed generative adversarial network to achieve multi-parameter imaging of musculoskeletal tissue. The in-silico experiment demonstrated that the BUIPIGAN could accurately reconstruct the SOS and density distribution. In addition, the BUIPIGAN could also exhibits robustness under different noise levels. The BUIPIGAN shows potential for the assessment of bone structure and the diagnosis of skeletal disease.

Acknowledgements

Project supported by the National Natural Science Foundation of China (Grant Nos. 12122403 and 12327807). The computations in this research were performed using the CFFF platform of Fudan University.

References

[1] Su N, Yang J, Xie Y, Du X, Chen H, Zhou H and Chen L 2019 *International Journal of Biological Sciences* **15** 776
[2] Mirza F and Canalis E 2015 *European Journal of Endocrinology* **173** R131
[3] Leslie W D and Morin S N 2014 *Current Opinion in Rheumatology* **26** 440
[4] Compston J, McClung M and Leslie W 2019 *The Lancet* **393** 364
[5] Cauley J 2013 *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* **68** 1243
[6] Si L, Winzenberg T, Jiang Q, Chen M and Palmer A 2015 *Osteoporosis International* **26** 1929
[7] Organization W H 1997 *The world health report* 1997*: conquering suffering; enriching humanity* pp. 162–162
[8] Guglielmi G, Grimston S, Fischer K and Pacifici R 1994 *Radiology* **192** 845
[9] Liu C, Li B, Diwu Q, Li Y, Zhang R, Ta D and Wang W 2018 *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **65** 2311
[10] Wang F, Zheng L, Theopold J, Schleifenbaum S, Heyde C E and Osterhoff G 2022 *Journal of Orthopaedic Surgery and Research* **17** 174
[11] Huddleston A, Wishko D, Dunn G, Metcalf D, Craven J, Stern R, Friedman R and Greenfield M A 1974 *Physics in Medicine & Biology* **19** 251
[12] Li B, Liu C, Liu X, Tran T, Li Y, Li D, Bi D, Liu D and Ta D 2022 *Chin. Phys. B* **31** 114303
[13] Garcia D, Le Tarnec L, Muth S, Montagnon E, Poree J and Cloutier G 2013 *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **60** 1853
[14] Jensen J, Nikolov S, Gammelmark K and Pedersen M 2006 *Ultrasonics* **44** e5
[15] Rao J, Ratassepp M and Fan Z 2016 *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **63** 737
[16] Zheng R, Le L, Sacchi M and Lou E 2015 *Ultrasound in Medicine & Biology* **41** 2955
[17] Guillermin R, Lasaygues P, Rabau G and Lefebvre J P 2013 *The Journal of the Acoustical Society of America* **134** 1001
[18] Pérez-Liva M, Herraiz J, Udías J, Miller E, Cox B and Treeby B 2017 *The Journal of the Acoustical Society of America* **141** 1595
[19] Suo M, Zhang D, Yang H and Yang Y 2023 *Computer Methods and Programs in Biomedicine* **231** 107404
[20] Wu Y, Yan W, Liu Z, Zhang Q, Zhou L, Song J, Qiu W, Ding M and Yuchi M 2024 *Physics in Medicine & Biology* **69** 105024
[21] Li Y, Wang J, Su C, Lin W, Wang X and Luo Y 2023 *Chin. Phys. B* **32** 014303
[22] Da Silva S, Karsou A, de Souza A, Capuzzo F, Costa F, Moreira R and Cetale M 2022 *Geophysical Journal International* **231** 1363
[23] Zhang W, Luo J and Teng J 2015 *Chinese Journal of Geophysics* **58** 216
[24] Dong L, Chi B, Tao J and Liu Y 2013 *Chinese Journal of Geophysics* **56** 685
[25] Yang J, Liu Y and Dong L 2014 *Chinese Journal of Geophysics* **57** 628
[26] Liu T, Han S, Xie L, Xing W, Liu C, Li B and Ta D 2024 *Physics in Medicine & Biology* **69** 125026
[27] Shi Q, Zhou T, Li Y, Liu C and Ta D 2022 *IEEE Transactions on Computational Imaging* **8** 1063
[28] Suo M, Zhang D, Yang H and Yang Y 2023 *Neural Computing and Applications* **35** 25027
[29] Paszke A, Gross S, Massa F, *et al.* 2019 *Advances in Neural Information Processing Systems* **2019** 32
[30] Pang B, Nijkamp E and Wu Y 2020 *Journal of Educational and Behavioral Statistics* **45** 227
[31] Karniadakis G, Kevrekidis I, Lu L, Perdikaris P, Wang S and, Yang L 2021 *Nat. Rev. Phys.* **3** 422
[32] Raissi M, Perdikaris P and Karniadakis G 2019 *Journal of Computational Physics* **378** 686
[33] Guo J, Wang H, Gu S and Hou C 2024 *Chin. Phys. B* **33** 050701
[34] Moseley B, Markham A and Nissen-Meyer T 2020 arXiv: 2006.11894
[35] Shukla K, Di Leoni P, Blackshire J, Sparkman D and Karniadakis G 2020 *Journal of Nondestructive Evaluation* **39** 61
[36] Gupta H, McCann M, Donati L and Unser M 2021 *IEEE Transactions on Computational Imaging* **7** 759
[37] Rasht-Behesht M, Huber C, Shukla K and Karniadakis G 2022 *Journal of Geophysical Research: Solid Earth* **127** e2021JB023120
[38] Yang F and Ma J 2023 *Journal of Geophysical Research: Solid Earth* **128** e2022JB025493
[39] Hu W, Chen J, Liu J and Abubakar A 2018 *IEEE Signal Processing Magazine* **35** 132
[40] Gui J, Sun Z, Wen Y, Tao D and Ye J 2021 *IEEE Transactions on Knowledge and Data Engineering* **35** 3313
[41] Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A and Chen X 2016 *Advances in Neural Information Processing Systems* **29**
[42] Arjovsky M, Chintala S and Bottou L 2017 *Proceedings of the* 34*th International Conference on Machine Learning*, 2017, Sydney, Australia, p. 214
[43] Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V and Courville A 2017 *Advances in Neural Information Processing Systems* **30**
[44] Guitton A and Alkhalifah T 2017 *Geophysics* **82** R299
[45] Alexey D 2020 arXiv: 2010.11929
[46] Chollet F 2017 *Proceedings of the IEEE conference on computer vision and pattern recognition,* 2017, pp. 1251–1258
[47] He K, Zhang X, Ren S and Sun J 2016 *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778
[48] Culjat M, Goldenberg D, Tewari P and Singh R 2010 *Ultrasound in medicine & biology* **36** 861
[49] Li Y, Shi Q, Li Y, Song X, Liu C, Ta D and Wang W 2021 *Chin. Phys. B* **30** 014302