IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 23 April 2024, accepted 15 May 2024, date of publication 23 May 2024, date of current version 10 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3404644

RESEARCH ARTICLE

A Consensus Mechanism to Improve Prediction of Cortical Bone Properties Using Ultrafast Ultrasound Acquisition

HOSSAM H. SULTAN[1], ENRICO GRISAN[1], (Senior Member, IEEE), PAUL DRYBURGH[2], LAURA PERALTA[2], (Member, IEEE), AND SEVAN HARPUT[1], (Member, IEEE)

[1]South Bank Applied Bioengineering Research, London South Bank University, SE1 0AA London, U.K.
[2]Department of Surgical and Interventional Engineering, School of Biomedical Engineering and Imaging Sciences, King's College London, WC2R 2LS London, U.K.

Corresponding authors: Sevan Harput (harputs@lsbu.ac.uk) and Hossam H. Sultan (abdelh13@lsbu.ac.uk)

ABSTRACT Human bone microarchitecture is complex and density-based bone assessment modalities cannot fully capture bone strength or health. Ultrasound can be used to assess bone microstructure, but it is hindered by the dense and acoustically diverse nature of cortical bone. This study proposes a methodology for predicting cortical bone thickness and porosity through a novel approach utilizing convolutional neural networks (CNNs), processing multi-frequency radiofrequency (RF) data obtained from ultrafast ultrasound, and implementing a consensus mechanism to enhance reliability. Received ultrasound RF signals are processed using a CNN with a mutual consensus mechanism, which is used to discard received RF data when measurement variation is over a certain threshold. The feasibility of the proposed method is demonstrated through realistic simulations and an ex vivo animal bone study using an ultrafast ultrasound scanner. The preliminary findings of this study demonstrate an enhancement in overall accuracy, with an increase from 92% to 95.6% for thickness and an increase from 73.4% to 88.4% for porosity classification, without and with consensus respectively. The implemented mutual consensus mechanism increases the accuracy of the thickness and porosity estimations both in silico and ex vivo. Ultrafast ultrasound scanners can capture thousands of RF signals within seconds, which results in availability of large datasets for implementation of artificial intelligence and machine learning algorithms. Here, we propose a new approach for ultrafast ultrasound data processing that values data quality over quantity by discarding noisy measurements using a consensus mechanism to improve the final estimation reliability.

INDEX TERMS Bone characterization, CNN, consensus mechanism, multi-frequency ultrasound, ultrafast ultrasound.

I. INTRODUCTION

Bone is a living tissue that is constantly being remodelled, involving the breakdown of old bone tissue and replacement with new bone tissues [1], [2], [3]. As people age, the rate of bone breakdown can exceed the rate of bone formation, leading to medical conditions that cause bones to become weak, increasing the risk of fractures [4]. Cortical bone is the dense outer layer of all bones [5], [6], and has an important role in determining bone strength while withstanding various mechanical loads and stresses [7]. Bone health assessment is important to evaluate the risk of fractures or monitor the effectiveness of osteoporosis treatment plans [8], [9], [10], [11].

The gold standard technique used in clinical routine exams is dual-energy X-ray absorptiometry (DXA), which measures bone mineral density (BMD) in the spine, hips, and other bones [12], [13]. However, DXA lacks detailed information about cortical thickness and porosity, which are important indicators of bone strength and fracture risk in bone health diseases and osteoporosis [14]. Quantitative computed tomography (QCT) and high-resolution peripheral quantitative computed tomography (HR-pQCT) can also be used to

The associate editor coordinating the review of this manuscript and approving it for publication was Riccardo Carotenuto.

assess bone health, while measuring BMD [15], [16], as well as other bone parameters, such as cortical thickness and trabecular bone microarchitecture [17], [18], [19]. However, all of these modalities employ ionizing radiation, which may pose risks, particularly with repeated exposures.

Bone strength refers to the ability of a bone to resist fracture. It is determined by various factors, including BMD, bone microarchitecture, and bone geometry. Ultrasound has emerged as a promising technique for assessing cortical bone microstructure [20], [21], [22], and can effectively identify bone abnormalities such as erosion, joint space narrowing, new bone formation, periarticular soft tissue swelling, and periarticular osteoporosis [23], [24], [25], [26]. It also can be used to assess cortical bone properties at various skeletal sites, including the tibia, radius, and phalanx [21], [27], [28]. Ultrasound offers a non-invasive approach, eliminating the need for ionizing radiation exposure, making it a safe and cost-effective alternative to X-ray-based techniques. Ultrasound portability enables convenient use at the bedside and in clinical settings. Ultrasound considerations that have been shown to be useful in assessing cortical bone include the speed of sound (SOS) [29], [30], [31], [32], [33], [34], broadband ultrasound attenuation (BUA) [35], and other quantitative ultrasound (QUS) parameters such as stiffness and acoustic impedance [20], [36]. These parameters reflect various aspects of cortical bone microstructure, such as thickness, porosity, and collagen content, which are not fully captured by BMD [37], [38]. Therefore, the importance of ultrasound in cortical bone assessment lies in its ability to provide information about bone geometry and strength via parameters like cortical thickness and porosity.

Cortical bone assessment with ultrasound has several challenges due to the unique characteristics of cortical bone tissue. Unlike soft tissues, cortical bone is dense and acoustically heterogeneous, making it challenging for ultrasound waves to penetrate effectively [39], [40]. The SOS within bone, influenced by porosity, adds complexity to accurate measurements [31], and the high attenuation and scattering of ultrasound in cortical bone limit the penetration depth, which can result in poor image resolution and reduced sensitivity to subtle changes within the bone structure. Additionally, the anisotropic nature of cortical bone, with varying orientations of collagen fibres and mineralized matrix, contributes to anisotropic acoustic properties that complicate ultrasound signals' interpretation. Relying on a single measurement to determine both thickness and porosity may lead to insufficient results as the relationship between porosity and the SOS in bone introduces variability, particularly as porosity changes, impacting the precision of assessment [41]. Moreover, the high variation between signals received from consecutive elements of the probe adds complexity to data interpretation.

Ultrafast ultrasound imaging enables the capture of images at frame rates up to 100 times faster than conventional imaging. This is achieved through the use of plane or divergent waves instead of focused beams, allowing for rapid acquisitions [42], [43]. Additionally, parallel processing techniques enable ultrafast ultrasound systems to compute in parallel as many lines as requested, resulting in the capability to compute a full acquisition rapidly [44], [45]. This can be advantageous in clinical settings where quick and accurate assessments are essential. Furthermore, multi-frequency ultrasound has emerged as a promising technique for assessing cortical bone properties, particularly for the evaluation of bone strength and fracture risk. The importance of multi-frequency ultrasound in cortical bone assessment lies in its ability to provide information about bone properties at multiple depth scales. Different frequencies of ultrasound waves penetrate to different depths in the bone tissue and interact with different structural components, such as pores with various sizes, and hence provide information about different aspects of bone microstructure, which can be used to estimate bone thickness and porosity more accurately than single-frequency ultrasound measurements [46].

The assessment of cortical bone health has witnessed significant advancements through the integration of deep learning and CNNs. In dental panoramic radiographs, where CNNs have demonstrated high accuracy for osteoporosis screening [47], and in applications such as predicting BMD and bone microarchitecture from non-contrast CT examinations [48], the adaptability of CNNs has been distinct. Further applications include the screening of primary osteopenia and osteoporosis based on lumbar radiographs [49], and predicting lumbar vertebrae BMD from unenhanced abdominal CT scans [50]. Moreover, a few recent studies have explored deep learning applications in ultrasound cortical bone analysis. A multichannel CNN (MCC-CNN) has been used to analyse ultrasonic guided waves pseudo images [51]. A real-time CNN architecture has been developed for ultrasound computed tomography (USCT) bone images segmentation [52], while a deep learning-based ultrasound computed tomography (UCT) approach has been introduced to reconstruct ultrasound images using backscattered information [53].

Despite these advancements, the direct processing of ultrasound RF data using deep learning techniques remains relatively unexplored as most existing techniques focus on image-based approaches. This study introduces a novel approach for predicting cortical bone thickness and porosity, avoiding the necessity for ultrasound images beamforming and computationally intensive image-based tasks by directly processing multi-frequency RF data using CNNs with attention mechanism. To improve the CNNs accuracy, a consensus approach is employed to choose or discard RF datasets acquired from sequential ultrasound acquisitions. Thanks to ultrafast ultrasound, it is possible to record thousands of RF data acquisitions every second, which are later processed with the CNN and only the most appropriate datasets are selected for estimation of cortical bone thickness and porosity. The consensus uses cross-correlation analysis, which not only considers the inter-channel correlations but also reflects

signal-to-noise ratio (SNR) variations. By integrating these factors, our approach surpasses simplistic majority voting methods, ensuring the retention of only the most reliable and informative data for accurate cortical bone thickness and porosity estimation. The feasibility of the proposed method is first demonstrated with a set of realistic simulations using Finite-Difference Time-Domain (FDTD) method and then an ex vivo animal bone study is performed using an ultrafast ultrasound system.

II. METHODOLOGY

A. SIMULATION FRAMEWORK FOR ULTRASOUND PROPAGATION

Accurately simulating ultrasound propagation in cortical bone requires a complex environment which includes realistic geometries with varying thickness and porosity, which mimics most of human range of healthy and osteoporotic bones [54], [55], [56]. This allows for the accurate modelling of the interaction between ultrasound waves and different tissues. This environment begins with soft tissue modelling, simulating the nature of biological tissues and has similar acoustic impedance, followed by a highly detailed cortical bone model with a grid size of 5 micrometres, including bone boundaries curvature, pore weighting to simulate real human bone [57], and surface irregularities extracted from micro-CT scans of human bones as described in [58]. After that, these two layers are followed by a bone marrow layer to mimic long bone shafts such as tibia as shown in FIGURE 1. Then, a 2D FDTD method, Simsonic [59], has been used for simulating ultrasound propagation in cortical bones.

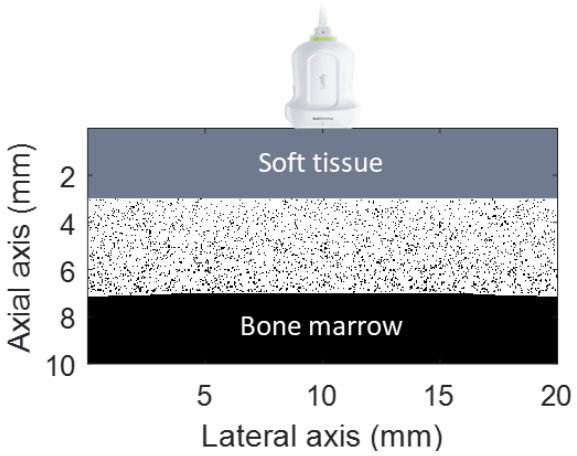


FIGURE 1. A simulation environment featuring the transducer positioning and the bone model that includes realistic pores obtained from ex vivo bone specimens, curvature, pore weighting, and surface irregularities.

B. WAVEFORM AND MULTI-FREQUENCY ACQUISITIONS

A 128 elements linear transducer array with 100 μm pitch was used in the simulations. To facilitate multi-frequency ultrasound measurements, chirp coded signals were designed as an excitation waveform featuring a frequency range that varies from 1 and 8 MHz, 60% fractional bandwidth (BW), and 2.5 μs pulse duration, which offer an improved SNR when compared to non-coded pulses. Additionally, a Tukey apodization was applied to mitigate edge wave effects. Plane-wave scanning has been employed to acquire echo patterns. FIGURE 2 shows a representation of RF data acquired from single channel at different frequencies. Each frequency reveals a specific layer of information, shaped significantly by numerous variables such as porosity, thickness, and frequency.

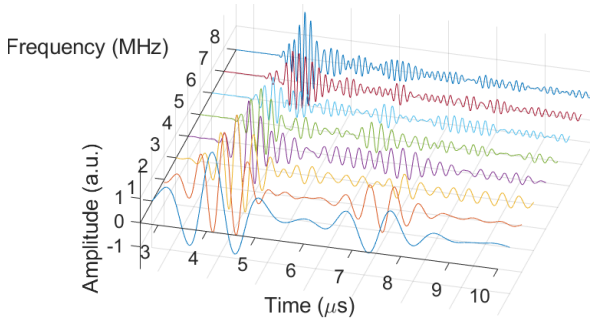


FIGURE 2. An example received RF data from the same channel at different frequencies from 1-8 MHz shows how the frequency affects the distal side echo.

C. SIMULATION DATASET OVERVIEW

A dataset consisting of 10240 RF scans of cortical bone samples was used with diverse collection of RF signals obtained from tibial cortical bone ultrasound simulations with varying thickness (1-8 mm), porosity levels (1-20%) using the 128 elements of the ultrasound probe. Porosity and thickness values incremented in 1% and 1 mm steps, respectively, and the porosity levels were categorized into four groups, porosities ranging from 1 to 5% were classified as low, those from 6 to 10% as moderate, porosities within the range of 11 to 15% were considered elevated, and porosities falling between 16 and 20% were categorized as high porosity.

D. EX VIVO VALIDATION AND POROSITY ASSESSMENT

An ex vivo validation study was conducted using bovine and swine bone samples with varying thickness and porosity levels. The sample set has 14 samples extracted from tibias with range of 3.1-9.3 mm thickness and 5-9.23% porosity values. Thickness measurements were obtained manually using a calliper, as the cortical bone layer was only backed with bone marrow, and the experiments were conducted ex vivo. Microscopic imaging techniques were employed to capture high-resolution 2D images of cross-sectional bone samples with a resolution of 2.26 μm to capture low pores sizes, and to estimate the average porosity percentage within the bone samples after thresholding and applying k-clustering to detect objects [60]. In addition, different morphological operations were applied to remove large voids originating from the specimen's extraction process and carefully removing sample cracks and minor scratches caused by bone dust during sample polishing. This estimation process was iteratively repeated 50 times for each sample to ensure robustness in porosity calculation.

FIGURE 3 shows a transverse cortical bone section scanned using a Meiji Techno MT9930L microscope equipped with a 10x objective lens. Ultrasound measurements have been conducted using a ULA-OP 256 ultrasound system [61] with 144 transducer elements, Esaote LA332. Chirp signals with spanning frequencies of 3-7 MHz were utilized, with a 3 MHz BW and a 2.5 μs duration.

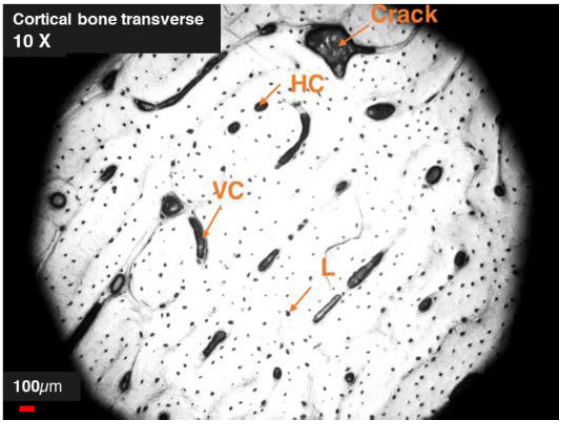


FIGURE 3. Microscopic image of transverse section of cortical bovine bone with an average porosity of 7.3%. The image shows bone microstructure including Haversian canals (HC), Volkmann's canals (VC), and Lacunae (L).

E. CONVOLUTIONAL NEURAL NETWORKS

Advances in neural networks, particularly attention mechanisms, enhance deep learning's appeal. These mechanisms dynamically adjust the significance of spectral ranges and spatial locations, improving the model's ability to capture discriminative features [62], [63], [64]. FIGURE 4 shows the proposed CNN architecture, which features distinct paths for dot product attention, each comprising convolutional layers with diverse kernel sizes, followed by batch normalization and ReLU activation. Outputs from the convolutional final stages are concatenated, subsequently processed through a max-pooling layer and a flatten layer. The flattened output is then directed into a multi-head self-attention layer [46], to discriminate endosteal and periosteal features inherent within the multi-dimensional multi-frequency RF data [65]. To mitigate overfitting, a dropout layer with a rate of 50% is applied between the last two fully connected layers.

The given CNN input size is 8429 $\times$ 1 $\times$ 8 $\times$ 10240, indicating 8429 $\times$ 1 data points samples, and the data is organized in an 8 $\times$ 1 grid where each element represents a frequency, the fourth dimension is the dataset size where each element represents a single RF response from a single channel. The mathematical notation is expressed as follows:

$$X \in \mathbb{R}^{8429x1\ x8\ x10240} \tag{1}$$

where, X represents the input data, $\mathbb{R}$ denotes the real numbers. The convolutional layers have different sizes of [1024,8,32], [512,8,32], [256,8,32], [32,1,32] which are used for building the structure. The parallel paths are used with different kernel sizes for the convolutional operation in each path, enabling the model to focus on different aspects simultaneously. A dot product attention is performed using the output of the four convolutional layers. All outputs are combined using concatenation to get a single output:

$$C = L1.L2.L3.L4 \tag{2}$$

where $L1.L2.L3.L4$ are the independent outputs from pathway 1, 2, 3, and 4 respectively. The concatenated output passes through a series of transformations throughout various layers, each serving a distinct and essential role in enhancing the model's capabilities. Max pooling aids in capturing the most salient features from the input data by selecting the maximum values within defined regions as:

$$\text{MaxPooling}(x) = \max(x) \tag{3}$$

Following max pooling, the processed information is forwarded to fully connected (fc) layers, to learn advanced complex hierarchical representations. The output of the fully connected layers is often passed through dropout layers, introducing a regularization mechanism that lessens overfitting. The dropout operation, denoted by $Dropout(x)$, randomly sets a fraction of input units to zero during training:

$$Dropout(x) = x \odot Bernoulli(p) \tag{4}$$

where $\odot$ represents element-wise multiplication, and $Bernoulli(p)$ is a binary mask with values sampled from a Bernoulli distribution with probability p. Subsequently, the flattened output joins a multi-head self-attention layer, a key component for capturing long-range dependencies within the data. This multi-head self-attention functions as an ensemble of attention mechanisms which calculates attention scores based on the input's own content [62]. The attention mechanism employed in the convolutional model for the input V can be mathematically described using the following equations for the query (Q), key (K), and value (V) transformations:

$$Attention\,(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{5}$$

where, d_k is the dimension of the key vectors. Additionally, multi-head self-attention involves linearly projecting the input query, key, and value vectors multiple times and then concatenating the results as a further step. Given h different sets of learned linear projections W_q^i, W_k^i, W_v^i for the i^{th} head, the multi-head self-attention is calculated as follows:

$$MultiHead\,(Q, K, V) = Concat\,(Head_1, \ldots, Head_h)\,.W_o \tag{6}$$

where, $Head_i = AttentionMechansim(Q.W_q^i, K.W_k^i, V.W_v^i)$, and W_o is the learned output projection and h is the total number of heads.

F. CONSENSUS-BASED DECISION-MAKING

A mutual agreement, referred to as a consensus mechanism, has been used to increase the collective agreement within

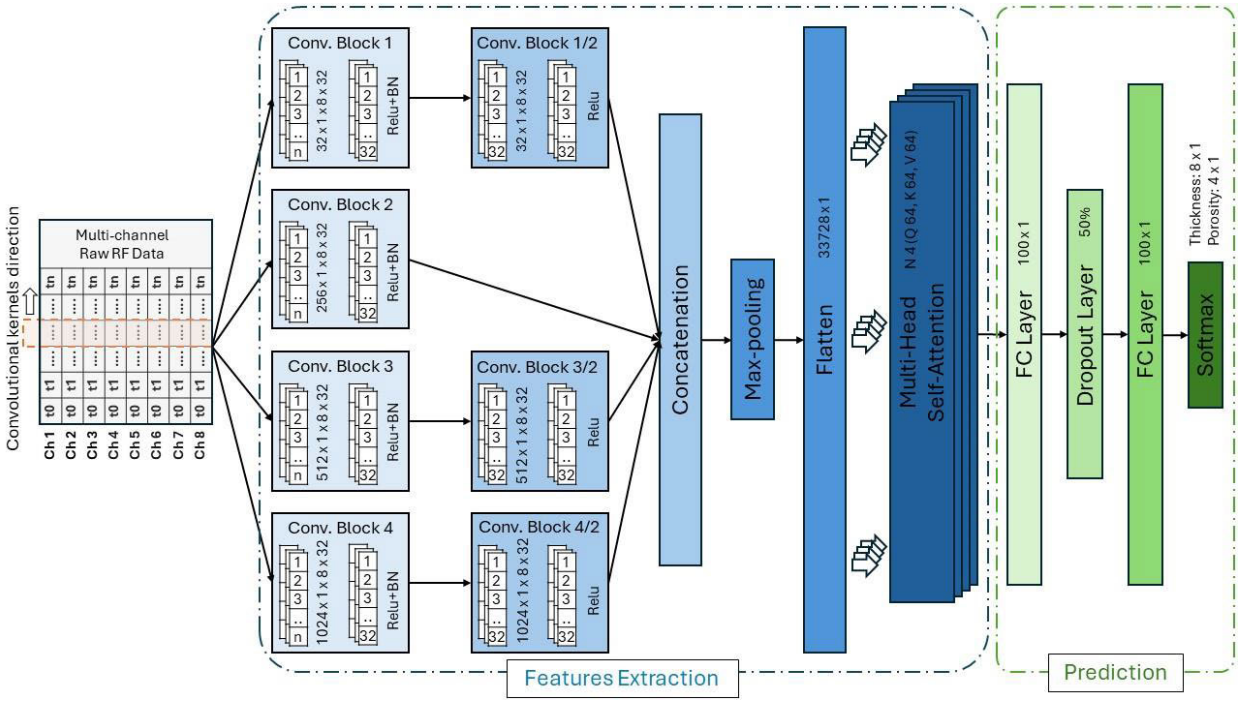


FIGURE 4. The proposed CNN architecture used to extract features from echoes between the endosteum and periosteum for thickness and porosity classification. The structure includes convolutional layers with kernel sizes of 32, 256, 512, and 1024, followed by batch normalization and ReLU layers. Outputs from the second convolutional stages are concatenated and processed through pooling and flatten layers. The flattened output then through a multi-head self-attention layer, followed by fully connected layers, and a dropout layer with a rate of 50% to avoid overfitting.

a system, involving selected channels based on the calculated values in the decision-making process. The consensus approach in the context of this study is a method of combining information from multiple channels to improve the accuracy and reliability of the prediction. It is based on the principle that the transducer elements or channels, determined by calculated values, agree on the same class or category of bone thickness or porosity. For example, if 64 channels are used during the scan, and the calculated threshold value is 0.75, then at least 48 channels need to agree on a specific decision (e.g., 2 mm thickness and low porosity), then the consensus mechanism would classify the bone sample into this category. However, if the level of agreement is below this certain threshold, the consensus mechanism would not assign a class, and this measurement would be discarded. Assuming having N channels of transducer array involved in decision making, and each channel provides decision D_i, the final decision $D_{consensus}$ is determined based on a majority vote, but with a threshold value for agreement.

$$D_{consensus} = \begin{cases} argmax_d \sum_{i=1}^{N} \partial\,(D_i, d), \\ if \ \dfrac{\sum_{i=1}^{N} \partial\,(D_i, D_{argmax})}{N} \geq Threshold \\ \\ Undecided, \quad otherwise \end{cases} \tag{7}$$

where, $\partial\,(D_i, d)$ is the Kronecker delta function, equal to 1 if $D_i = d$ and 0 otherwise. $argmax_d$ finds the decision d that maximizes the sum of votes. D_{argmax} is the decision that received the most votes. *Threshold* is the calculated tolerance for agreement as shown in the following equation:

$$Threshold = 1 - \frac{mean\ cross\ correlation}{2} \tag{8}$$

The *Threshold* value balances the *mean cross correlation* between channels to derive a suitable threshold value. For instance, if the cross-correlation between channels is high, indicating strong similarity, the threshold value will be relatively low. Conversely, if the cross-correlation is low, suggesting less similarity, the threshold will be increased to ensure robust decision-making. This approach allows to filter and select relevant data points for further analysis or classification tasks, ensuring the reliability of the results. Therefore, the threshold value can be adjusted for accuracy and consistency; a higher threshold increases precision but may require more measurements for confidence. Additionally, the consensus mechanism adopted by the model is characterized by a conscious selection of measurements, rather than blindly following the majority. This strategy ensures that the model's decision-making process is based on cross-correlation between the received signals acquired from different channels.

G. MODEL TRAINING AND OPTIMIZATION TECHNIQUES

The models are trained on Intel i7-7700HQ CPU at 2.80GHz with 16 GB of RAM and a NVIDIA GeForce GTX 1060 GPU. To ensure generalization and prevent

TABLE 1. Hyperparameters and fine-tuning options for CNN training.

Hyperparameter	Options	Description
Learning Rate	0.001, 0.0001, 0.00001	Initial learning rate for optimizers
Optimizer	Adam, SGD, RMSprop	Optimization algorithm used during training
Batch Size	8, 16, 32, 64	Number of samples per training iteration
Weight Decay	0.001, 0.0001, 0.00001	Coefficient for L2 regularization
Epochs	20, 30, 40, 50, 100	Number of complete passes through the training set
Activation Function	ReLU, Leaky ReLU, Tanh	Activation function applied to hidden layers
Dropout	0.1, 0.3, 0.5	Dropout rate applied to prevent overfitting
Kernel Size	[32x8], [64x8], [128x8], [256x8], [512x8], [1024x8]	Size of convolutional kernels
Stride	8, 32, 64, 256	Step size for moving the kernel during convolution
Padding	Same, Valid, Zero	Padding method to handle edge effects during convolution
Pooling	Max Pooling, Average Pooling	Pooling method to down sample feature maps
Initialization	Zero, He, Glorot	Method for initializing weights in the network
Early Stopping	Patience 5, 10	Criteria for terminating training early to prevent overfitting after some epochs without loss change

overfitting, an independent set of simulations was generated for testing, with 12800 data points for training and validation and 10240 for testing. Training spanned 40 epochs to allow the models to converge with a batch size of 16. ADAM optimizer with a dynamic learning rate of 0.0001, adjusted during training based on the gradients of the cross-entropy loss. It helped to mitigate the risk of overshooting or slow convergence for effective training dynamics. Additionally, L2 regularization was applied with a weight decay coefficient of 0.001 to prevent overfitting and improve model generalization. Table 1 shows hyperparameters used to fine-tune the CNN during training.

III. RESULTS

A. CHANNEL TO CHANNEL RF DATA VARIATION

This section explains the main motivation behind the consensus mechanism. Consider FIGURE 5, which shows an example of the correlation coefficient (r) of all channels (from 1 to 128) at 5 MHz with respect to the received signal from the central element. The similarities between the signals received from different channels of the ultrasound array begins to decrease after a few wavelengths and consequently deteriorate at the border regions of the transducer array, as shown in FIGURE 5. This decay can be attributed to the physical properties of the medium through which the ultrasound signals propagate. As the signals traverse the medium, interference patterns may develop, leading to a loss of correlation among channels. This consistency between channels could be quantified using cross-correlation coefficients.

Additionally, the choice of using an apodized ultrasound signal is another factor that reduced the correlation for the elements at the edge of transducer's aperture during simulation. For this reason, a strategic decision was made to use only 64 middle channels out of 128 available channels for data analysis and implement consensus mechanism for data received from these channels. For the signals within this subarray, between channels 33-96 (vertical dashed

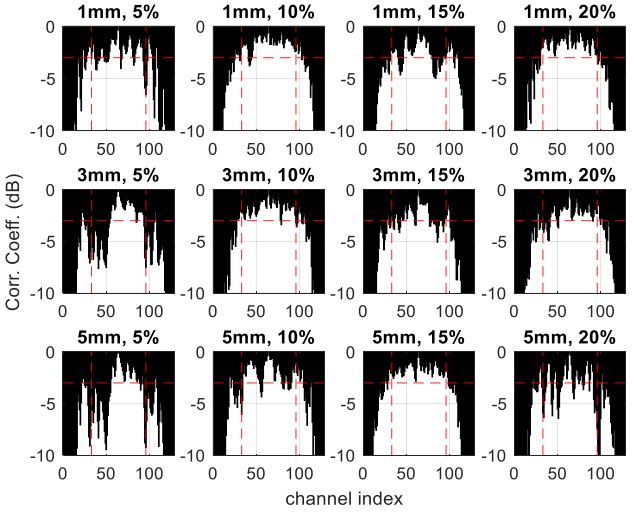


FIGURE 5. An example of the cross-correlation at 5 MHz for all 128 channels with respect to the middle channel. The horizontal red line represents the −3dB level and the vertical red lines represent the apodization roll-off. The correlation (dB) starts to decay, and gradually deteriorate until reach low correlations at the boundaries (r < 0.1).

boundaries), the correlation between received signals are usually larger than -3 dB (horizontal dashed levels), which shows that the received signals have similar information (r >0.7). However, decorrelation of the signal is observed for many cases, such as FIGURE 5 (bottom-right). The dissimilarity between the signals, indicated by low coherence, may potentially lead to inaccurate results when using a CNN or other processing methods. This issue was addressed by implementation of the consensus mechanism in this study.

B. QUANTITATIVE ASSESSMENT AND PREDICTION OF THICKNESS

1) THICKNESS CLASSIFICATION WITHOUT CONSENSUS

FIGURE 6 shows the confusion matrix that reveals the model performance in classifying critical bone thickness values. For low cortical bone thickness values, 1-2 mm, the classifier

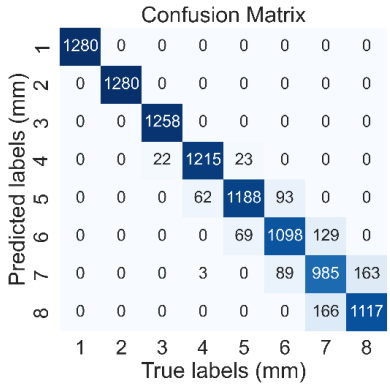


FIGURE 6. Confusion matrix showing the classification results of a cortical bone thickness classifier for thicknesses ranging from 1 to 8 mm without consensus.

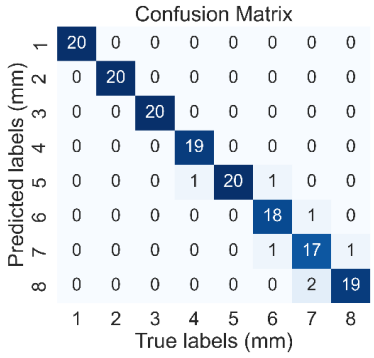


FIGURE 7. Confusion matrix showing the classification results of a cortical bone thickness classifier for thicknesses ranging from 1 to 8 mm with consensus.

exhibited a high accuracy, while reaching 100% in some cases. However, as the bone thickness increases, the classifier's accuracy decreases. While most of instances were correctly classified as 7 mm, some were slightly under or overestimated as 6 or 8 mm.

2) THICKNESS CLASSIFICATION WITH CONSENSUS

In the consensus approach, the results in FIGURE 7 shows the aggregation after combining information from all 64 channels for each scan with a threshold value of 58% to make a decision. The threshold value is calculated from equation 8 based on the global cross-correlation average which is 0.844. The diagonal elements represent the true positives for each class, while off-diagonal elements indicate misclassifications. In the confusion matrix given in FIGURE 7, most of the diagonal values are near 20, indicating that the consensus decision is consistent with the ground truth for the corresponding classes. The off-diagonal values, though present, are generally low, indicating minimal confusion between adjacent thickness classes of ± 1 mm. The overall accuracy of consensus is 95.6%. Comparing this to the results before consensus, where the overall accuracy was 92%, The consensus approach helped to reduce misclassifications and to increase the true positive counts. With consensus, the overall accuracy is 95.6% and 153 true positives out of 160, in which

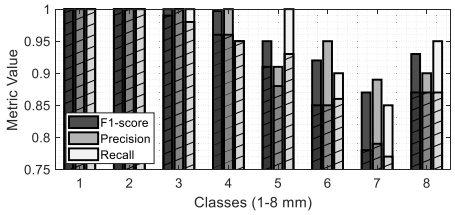


FIGURE 8. Comparison of metrics for consensus and non-consensus (hatched) models for thickness classification.

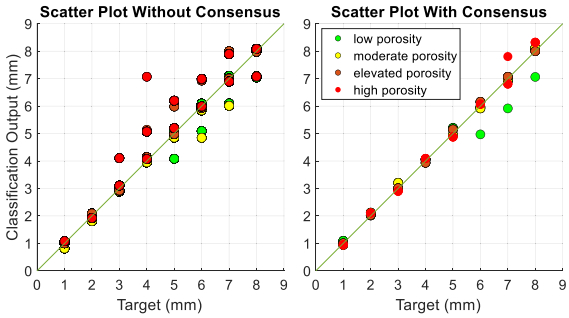


FIGURE 9. (Left) the output of the CNN model against the actual target cortical bone thicknesses for all porosity levels without consensus, (Right) the output of the CNN model against the actual target cortical bone thicknesses for all porosity levels with consensus.

consensus has an impact on classification accuracy specifically for lower accuracies as for 7mm class.

3) COMPARING CONSENSUS VERSUS NON-CONSENSUS

FIGURE 8 presents a comparison between performance metrics for both consensus and non-consensus models. For consensus, the overall accuracy is 92% with 9421 true positives. Classes 1 and 2 stand out with perfect precision, recall, and F1 Scores, signifying high accuracy. Class 3 has a high accuracy of 98.3%, with a relatively balanced precision and recall, indicating its ability to correctly identify most instances of this class but with occasional misclassifications. In contrast, Classes 4-7 show varying degrees of misclassification, suggesting a potential challenge in distinguishing. Class 7 suffers from the lowest recall of 0.77, indicating frequent instances missed by the model, resulting in a relatively low F1 Score of 0.87. However, it is significant that even instances classified incorrectly exhibit a margin of error within the range of approximately ± 1 mm. This variability happens due to the complex relationship of factors such as the speed of sound and porosity combinations within the dataset. The overall accuracy without consensus is 92%.

The scatter plot presented in FIGURE 9 visually demonstrates the association between the cortical bone thickness values predicted by the CNN model and their corresponding target values. Each scatter plot involves four distinct porosity levels (low, moderate, elevated, and high porosity) for various thicknesses. The dataset without consensus approach includes 10240 data points, 1280 for each thickness value, while the dataset with consensus includes 160 data points which is the combination of 64 channels and the unique number of thickness vs porosity combination. The scatter plot

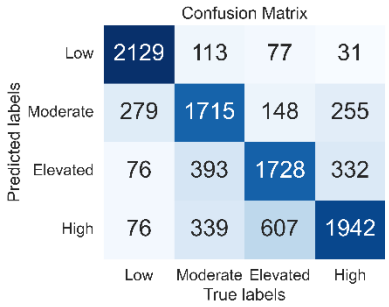


FIGURE 10. Confusion matrix illustrating the classification results of a cortical bone porosity classifier for different porosity levels.

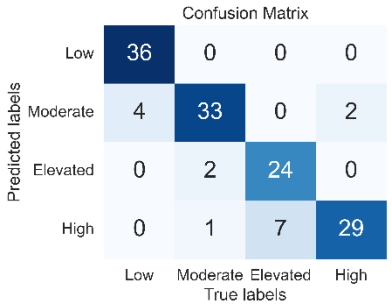


FIGURE 11. Confusion matrix illustrating the classification results of a cortical bone porosity levels classifier for with consensus mechanism.

highlights the model's performance across different thickness categories and porosity levels. For certain values, particularly for thickness classes 6, 7, and 8 mm, shows an underestimation trend. This trend is particularly evident for instances characterized by low porosity levels, suggesting a potential influence of the faster speed of sound associated with lower porosity.

C. QUANTITATIVE ASSESSMENT AND PREDICTION OF POROSITY

1) POROSITY CLASSIFICATION WITHOUT CONSENSUS

FIGURE 10 presents the confusion matrix illustrating the model's performance in classifying various porosity levels without consensus being used. For low porosity values, the classifier demonstrated an accuracy of nearly 83%. However, for moderate and elevated porosities, the classifier achieved an accuracy of around 67%. For high porosity, the classifier attained a relatively higher results compared to the former two classes with an accuracy of nearly 76%.

2) POROSITY CLASSIFICATION WITH CONSENSUS

In FIGURE 11, the model's performance is evaluated based on the diagonal elements representing true positive counts for each porosity class, while off-diagonal elements indicate instances of misclassification. The primary focus is on the model's ability to accurately identify samples within each porosity category by using consensus. Measurements that do not pass the threshold of the consensus mechanism are excluded to increase the reliability of the system. Upon closer examination, the classifier showed a significant increase in

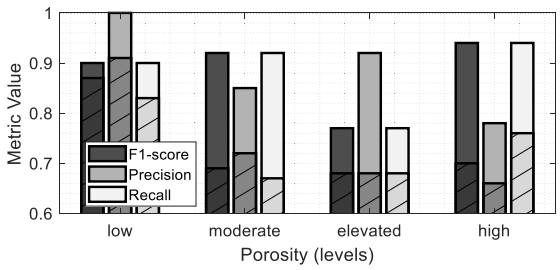


FIGURE 12. Comparison of metrics for consensus and non-consensus (hatched) models for porosity classification, with the most significant impact of consensus observed at higher porosity levels.

classifying all porosities levels with overall accuracies of 90.0%, 91.66%, 77.42%, 93.55% compared to 83.2%, 67.0%, 67.5%, and 75.9% for all levels with and without consensus respectively, with an overall accuracy of 88.4%.

3) COMPARING CONSENSUS VERSUS NON-CONSENSUS

Further details the classification performance metrics for consensus is shown in FIGURE 12, where consensus mechanism outperforms the CNN-only results.

FIGURE 13 shows a scatter plot of the predicted porosity values using the consensus mechanism. The results include all porosity levels from 1-20% where each level also acquired from all thicknesses (1-8 mm), and each scatter is the average output of a group of 64 channels of the same sample. The vertical distances between each data point and the trend line represent the absolute residuals. Smaller residuals indicate a closer match between predictions and ground truth, while larger residuals suggest greater deviations. The average percentage error is 11.6%, and the percentage errors associated with each thickness are as follows: 0.0%, 11.2%, 27.28%, 6.67%, 5.89%, 15.8%, 5.0%, and 10.0% for thicknesses from 1-8mm respectively. However, it shows an overall accuracy of 73.4% and 7514 true positive samples for classification without consensus. Class 1 stand out with higher precision, recall, and F1 Scores of 0.91, 0.83, and 0.87 respectively. Classes 2-4 have a comparable accuracy of around 84%, with a relatively imbalanced precision and recall, suggesting its suffering to correctly identify instances of this class.

D. CNN TRAINING AND VALIDATION LOSSES

FIGURE 14 shows the convergence behaviour of the CNN model during training for thickness and porosity classification tasks. The top plot shows the training and validation loss over 40 epochs for thickness classification. Both the training and validation losses exhibit an initial sharp decrease, indicating effective learning from the training data. However, as training progresses, the training loss continues to decrease steadily, reaching a minimum value of 0.0026. The validation loss demonstrates a stable plateau between epochs 12 and 26, accompanied by occasional fluctuations thereafter. These fluctuations closely mirror the trends observed in the training loss, indicating a coherent learning process. This synchronization suggests that the model effectively adapts to

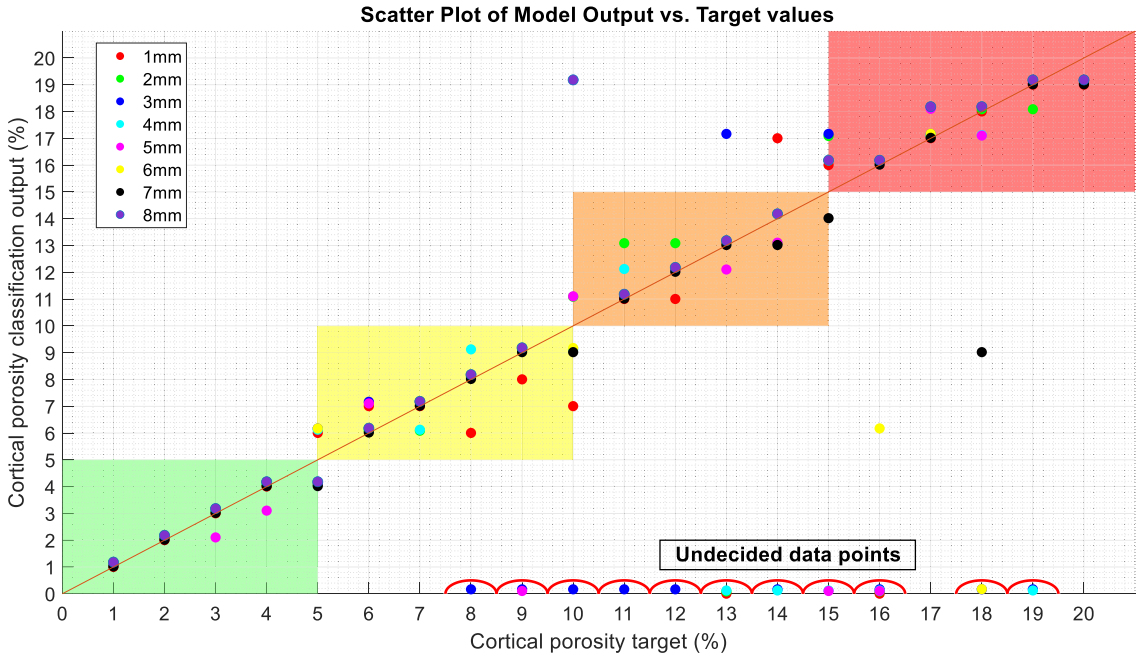


FIGURE 13. The classification output of the CNN model against the actual target cortical bone porosity. The coloured groups represent different levels of porosity: low, moderate, elevated, and high for green, yellow, orange, and red regions respectively. Estimated porosity levels falling within these coloured regions are considered correct. Undecided data points (22/160) are the results of the consensus mechanism.

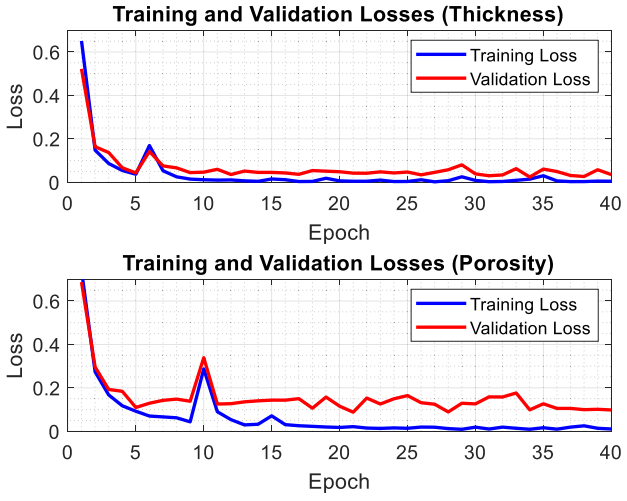


FIGURE 14. Training and validation losses for thickness and porosity classification tasks over the same number of epochs to monitor the model's performance during training. (Up) the loss curves for thickness classification, (bottom) the loss curves for porosity classification.

the training data, maintaining robust performance even as it encounters new patterns during validation. The training and validation losses for porosity classification, FIGURE 14 (bottom), show a different progression throughout the 40 epochs. Similarly, the validation loss matches, initially decreasing from 0.687 to 0.088. However, after epoch 17, occasional slight fluctuations in the validation loss become apparent, though it remains relatively stable after epoch 35. In comparison to thickness classification, the gap between training and validation losses appears more obvious for porosity classification. This divergence could be attributed to the inherent complexity of cortical bone, where capturing and interpreting porosity information poses greater challenges. Furthermore, the close alignment between training and validation losses throughout epochs highlights a well-balanced training regimen. The slight divergence observed between these losses implies minimal overfitting for both thickness and porosity classification.

E. EMPIRICAL VERIFICATION THROUGH EX VIVO SPECIMENS

First, observations are done on ex vivo RF data to identify the variation between received signals, similar to section III-A. FIGURE 15 shows the channel correlation for an ex vivo sample with a 6.33±0.01 mm thickness and 5.45% porosity. Although, this sample had low thickness and porosity variation, the received RF data starts decorrelating after a few wavelengths.

In FIGURE 16, each point on the plot corresponds to a specific acquisition, with the actual measured thickness predicted value. The results show a reasonable alignment between the actual and classified thickness values for consensus results (blue). The error distribution graph highlights two scenarios: one with consensus (blue) and another without consensus (black). For consensus, the mean error, μ_1, is -0.7 mm with a standard deviation, σ_1, of 1.15 mm. On the contrary, in the absence of consensus, the distribution has a $\mu 2$ of -1.15 mm and a $\sigma 2$ of 2.06 mm. Additionally, in FIGURE 17, the results for porosity classification reveal

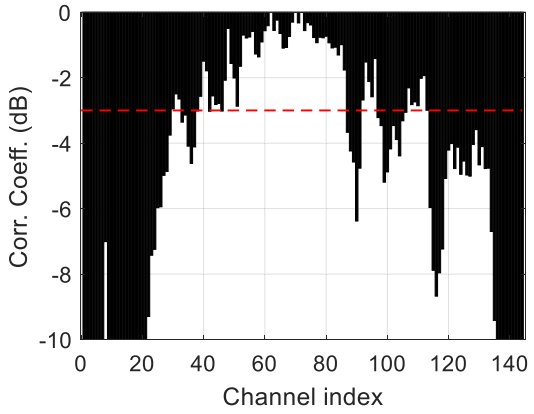


FIGURE 15. Cross-correlation between received RF signals with respect to the middle channel is shown for an ex vivo bone sample scanned at 3 MHz using ULA-OP system. The sample thickness is 6.33±0.01 mm with a porosity of 5.45%.

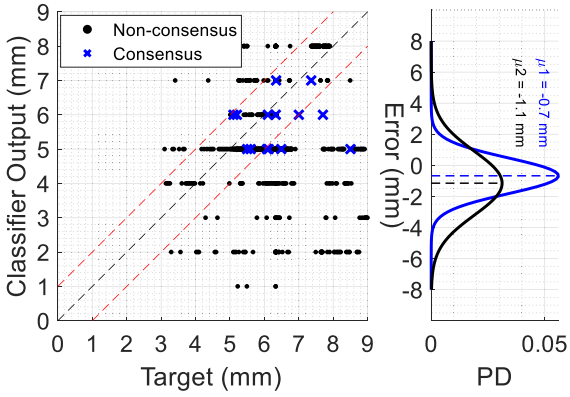


FIGURE 16. (Left) Thickness classification output against the actual target thicknesses for the ex vivo bone measurements. Red diagonal lines show the confidence interval with a classifier error of one unit. (Right) Probability distribution (PD) of error values for consensus (blue) and non-consensus (black) cases, derived from the fitted histograms.

a more varied distribution. For consensus, μ_1 is 0.49 on a scale from 1 to 4 representing 4 distinct porosity levels, with σ_1 of 0.82. On the other hand, for non-consensus, μ_2 is 1.2, with a σ_2 of 1.38. The results demonstrate an improvement in thickness and porosity estimation using the consensus mechanism.

IV. DISCUSSION

In the simulation study, the distribution and sizes of bone pores have been carefully engineered to align with the authentic human ranges observed in cortical bone with 25th and 75th percentiles ranging for different age groups [56]. In addition, covering both normal and abnormal cortical bone conditions [54], [55], [56], [66]. Our previous investigations employed various CNN models for the thickness classification [67], by applying continuous wavelet transformation of the received RF echoes. While these models have demonstrated good accuracy, it is important to note that their design was primarily oriented towards image classification. This includes computational expenses and long processing times. Furthermore, the utilization of these models

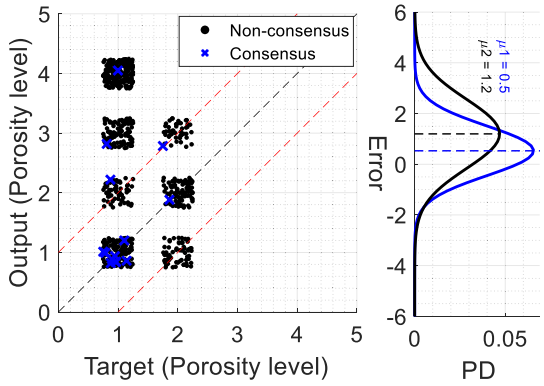


FIGURE 17. (Left) Porosity classification output against the actual level for the ex vivo measurements. Red diagonal lines show the confidence interval with a classifier error of one unit. (Right) Probability distribution (PD) of error values for consensus, blue, and non-consensus, black, cases. Overlapping data points were jittered to reduce overplotting for better visualization.

required pre-processing to convert signals into images. In the light of these considerations, we have developed a novel model customised to seamlessly accommodate and adapt to multi-frequency RF signals with varying parameters [46]. We found that Adam consistently outperformed other optimization algorithms in terms of convergence speed and final model accuracy. Specifically, compared to stochastic gradient descent (SGD), Adam reduced the training time by approximately 20% while achieving a 3% improvement in validation accuracy. The training time for our model is approximately 30 times shorter than the reported time in [67]. Moreover, it is needed to acknowledge the limitations stemming from the dataset's generation process. The comprehensive coverage of thickness and porosity levels needed significant time, highlighting the challenges associated with data partitioning during training.

The integration of multi-frequency ultrasound represents a progress in overcoming limitations associated with single-frequency ultrasound [67]. The sensitivity of single-frequency ultrasound to variations in soft tissue and the impact of surface roughness on measurements have been well-documented challenges in the ultrasound bone imaging field [68]. By employing multi-frequency ultrasound, these limitations can be mitigated, leading to more robust diagnostic capabilities. The best overall accuracy for classification prediction is 95.6%. Some misclassified data points fall within the range of ±1 mm, which is likely to be the results of varying speed of sound values in bone tissue at different porosity levels or variations across different section of the same bone sample [69]. For porosity classification, precision in numerical measurements may not always be the primary goal. Instead, the significance often lies in the ability to classify bones into distinct categories [70], where the proposed model showed an overall accuracy of 88.4%.

A novel concept proposed in this study is the implementation of a collective consensus mechanism for ultrafast ultrasound scans, which has shown to enhance the reliability

and accuracy of the overall decision-making process. Specifically, our findings indicate a reduction in the variability of ultrasound interpretation, coupled with an increase in confidence and an improvement in diagnostic quality. For thickness classification, the consensus approach led to an accuracy increasing by approximately 3.5%. For porosity classification, the influence of consensus was even more substantial, raising the overall accuracy from 73.4% to 88.4%. These specific numerical improvements highlight the impact of the collective decision strategy. This emphasizes the strategy's pivotal role not only in enhancing the accuracy of detecting true positives but also in providing a valuable assessment of data sufficiency and determining the adequacy of data for decision making as in FIGURE 13. In a dataset of 160 cases, 22 inaccurate predictions were discarded by the consensus mechanism, which did not meet the preset consensus threshold value.

While previous studies have made significant developments in utilizing deep learning for cortical bone ultrasound scanning and thickness classification, they primarily focus on image-based approaches [51], [52], [53]. The study in [51] used a directly contacted transmitters and receivers on bone surfaces, making it unsuitable for in vivo human measurements and neglecting the influence of soft tissue. Moreover, it fails to account for irregularities in bone geometry and surface topology, potentially limiting its applicability in real-world scenarios. The absence of porosity variation assessment in the same study complicates the accurate characterization of cortical bone properties.

The experimental results showed that the transformation from in silico study to real-life measurements is possible, FIGURE 16 and FIGURE 17. The decorrelation observed in FIGURE 15 corresponds with the simulated data illustrated in FIGURE 5, both measured at the -3 dB level. Moreover, our current study is limited to experimental samples categorized as having low and medium porosities due to unavailability of bone samples with higher porosities. However, extending real measurements and re-training, would increase the portability of the system to be further examined on human measurements.

Integration of multi-frequency ultrasound and the adoption of a consensus approach represent promising opportunities for refining ultrasound bone characterisation. It is essential to acknowledge that the determination of the consensus threshold level remains an area not well explored, and the lack of a standardized method poses challenges in its implementation. These factors include the number and type of transducers or channels, the establishment of thresholds or criteria for agreement, tissue properties, the choice of algorithm or technique for information integration. Despite the challenges associated with standardization and implementation, the benefits in terms of improved reliability, accuracy, and diagnostic quality are evident. Future research should address these challenges to promote the wider adoption of the consensus approach, ultimately contributing to enhanced clinical outcomes.

V. CONCLUSION

This study presented a consensus mechanism to classify cortical bone thickness and porosity using muti-frequency RF data acquisitions processed by a CNN. Consensus mechanism was used to accept or discard the recorded ultrasound acquisitions. The utilization of ultrasound plane waves in both simulation and experimental studies, proved instrumental in overcoming challenges related to practicality of the consensus mechanism, where thousands of measurements can be performed within a second. The effectiveness of the proposed method was verified using in silico and ex vivo studies across various cortical bone thickness (1-8 mm) and porosity values (1-20%). In addition, the integration of a consensus strategy in decision-making, coupled with access to a large simulation dataset, resulted in improvements in the cortical bone assessment with ultrasound.

REFERENCES

[1] D. J. Hadjidakis and I. I. Androulakis, "Bone remodeling," *Ann. New York Acad. Sci.*, vol. 1092, no. 1, pp. 385–396, 2006.
[2] N. R. Fuggle, L. D. Westbury, G. Bevilacqua, P. Titcombe, M. Ó Breasail, N. C. Harvey, E. M. Dennison, C. Cooper, and K. A. Ward, "Level and change in bone microarchitectural parameters and their relationship with previous fracture and established bone mineral density loci," *Bone*, vol. 147, Jun. 2021, Art. no. 115937.
[3] X. Feng and J. M. McDonald, "Disorders of bone remodeling," *Annu. Rev. Pathol., Mech. Disease*, vol. 6, pp. 121–145, Feb. 2011.
[4] O. Demontiero, C. Vidal, and G. Duque, "Aging and bone loss: New insights for the clinician," *Therapeutic Adv. Musculoskeletal Disease*, vol. 4, no. 2, pp. 61–76, Apr. 2012.
[5] P. Augat and S. Schorlemmer, "The role of cortical bone and its microstructure in bone strength," *Age Ageing*, vol. 35, no. 2, pp. ii27–ii31, Sep. 2006.
[6] F. Pouresmaeili, B. K. Dehghan, M. Kamarehei, and G. Y. Meng, "A comprehensive overview on osteoporosis and its risk factors," *Therapeutics Clin. Risk Manage.*, vol. 14, pp. 2029–2049, Nov. 2018.
[7] Y. Bala, R. Zebaze, and E. Seeman, "Role of cortical bone in bone fragility," *Current Opinion Rheumatol.*, vol. 27, no. 4, pp. 406–413, 2015.
[8] G. Osterhoff, E. F. Morgan, S. J. Shefelbine, L. Karim, L. M. McNamara, and P. Augat, "Bone mechanical properties and changes with osteoporosis," *Injury*, vol. 47, pp. S11–S20, Jun. 2016.
[9] E. J. Samelson, "Cortical and trabecular bone microarchitecture as an independent predictor of incident fracture risk in older women and men in the bone microarchitecture international consortium (BoMIC): A prospective study," *Lancet Diabetes Endocrinol.*, vol. 7, no. 1, pp. 34–43, Jan. 2019.
[10] G. Jonasson, A. Hassani-Nejad, and M. Hakeberg, "Mandibular cortical bone structure as risk indicator in fractured and non-fractured 80-year-old men and women," *BMC Oral Health*, vol. 21, no. 1, pp. 1–8, Dec. 2021.
[11] E. Seeman, "Growth and age-related abnormalities in cortical structure and fracture risk," *Endocrinol. Metabolism*, vol. 30, no. 4, p. 419, 2015.
[12] X.-G. Cheng, D.-Z. Yang, Q. Zhou, T.-J. Zhuo, H.-C. Zhang, J. Xiang, H.-F. Wang, P.-Z. Ou, J.-L. Liu, L. Xu, G.-Y. Huang, Q.-R. Huang, H. S. Barden, L. S. Weynand, K. G. Faulkner, and X.-W. Meng, "Age-related bone mineral density, bone loss rate, prevalence of osteoporosis, and reference database of women at multiple centers in China," *J. Clin. Densitometry*, vol. 10, no. 3, pp. 276–284, Jul. 2007.
[13] P. Clavert, R.-M. Javier, J. L. Charrissoux, L. Obert, L. Pidhorz, F. Sirveaux, P. Mansat, and T. Fabre, "How to determine the bone mineral density of the distal humerus with radiographic tools?" *Surgical Radiologic Anatomy*, vol. 38, no. 4, pp. 389–393, May 2016.
[14] O. R. Boughton, S. Ma, X. Cai, L. Yan, L. Peralta, P. Laugier, J. Marrow, F. Giuliani, U. Hansen, R. L. Abel, Q. Grimal, and J. P. Cobb, "Computed tomography porosity and spherical indentation for determining cortical bone millimetre-scale mechanical properties," *Sci. Rep.*, vol. 9, no. 1, p. 7416, May 2019.
[15] H. Fuller, R. Fuller, and R. M. R. Pereira, "High resolution peripheral quantitative computed tomography for the assessment of morphological and mechanical bone parameters," *Revista Brasileira de Reumatologia English Ed.*, vol. 55, no. 4, pp. 352–362, Jul. 2015.

[16] K. Engelke, J. E. Adams, G. Armbrecht, P. Augat, C. E. Bogado, M. L. Bouxsein, D. Felsenberg, M. Ito, S. Prevrhal, D. B. Hans, and E. M. Lewiecki, "Clinical use of quantitative computed tomography and peripheral quantitative computed tomography in the management of osteoporosis in adults: The 2007 ISCD official positions," *J. Clin. Densitometry*, vol. 11, no. 1, pp. 123–162, Jan. 2008.

[17] J. P. van den Bergh, P. Szulc, A. M. Cheung, M. Bouxsein, K. Engelke, and R. Chapurlat, "The clinical application of high-resolution peripheral computed tomography (HR-pQCT) in adults: State of the art and future directions," *Osteoporosis Int.*, vol. 32, no. 8, pp. 1465–1485, Aug. 2021.

[18] K. K. Nishiyama and E. Shane, "Clinical imaging of bone microarchitecture with HR-pQCT," *Current Osteoporosis Rep.*, vol. 11, no. 2, pp. 147–155, Jun. 2013.

[19] C. J. Collins, P. R. Atkins, N. Ohs, M. Blauth, K. Lippuner, and R. Müller, "Clinical observation of diminished bone quality and quantity through longitudinal HR-pQCT-derived remodeling and mechanoregulation," *Sci. Rep.*, vol. 12, no. 1, p. 17960, Oct. 2022.

[20] Q. Grimal and P. Laugier, "Quantitative ultrasound assessment of cortical bone properties beyond bone mineral density," *IRBM*, vol. 40, no. 1, pp. 16–24, Feb. 2019.

[21] J. Karjalainen, O. Riekkinen, J. Toyras, H. Kroger, and J. Jurvelin, "Ultrasonic assessment of cortical bone thickness in vitro and in vivo," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 55, no. 10, pp. 2191–2197, Oct. 2008.

[22] J. Foiret, J.-G. Minonzio, M. Talmant, and P. Laugier, "Cortical bone quality assessment using quantitative ultrasound on long bones," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 1121–1124.

[23] M. Hughes, C. Bruni, G. Cuomo, A. D. Sedie, L. Gargani, M. Gutierrez, G. Lepri, B. Ruaro, T. Santiago, Y. Suliman, S. Watanabe, A. Iagnocco, D. Furst, and S. Bellando-Randone, "The role of ultrasound in systemic sclerosis: On the cutting edge to foster clinical and research advancement," *J. Scleroderma Rel. Disorders*, vol. 6, no. 2, pp. 123–132, Jun. 2021.

[24] G. S. Kaeley, C. Bakewell, and A. Deodhar, "The importance of ultrasound in identifying and differentiating patients with early inflammatory arthritis: A narrative review," *Arthritis Res. Therapy*, vol. 22, no. 1, pp. 1–10, Dec. 2020.

[25] A. Iagnocco, "Imaging the joint in osteoarthritis: A place for ultrasound?" *Best Pract. Res. Clin. Rheumatol.*, vol. 24, no. 1, pp. 27–38, Feb. 2010.

[26] A. S. Zayat, K. Ellegaard, P. G. Conaghan, L. Terslev, E. M. A. Hensor, J. E. Freeston, P. Emery, and R. J. Wakefield, "The specificity of ultrasound-detected bone erosions for rheumatoid arthritis," *Ann. Rheumatic Diseases*, vol. 74, no. 5, pp. 897–903, May 2015.

[27] J. Damilakis, G. Papadokostakis, H. Vrahoriti, I. Tsagaraki, K. Perisinakis, A. Hadjipavlou, and N. Gourtsoyiannis, "Ultrasound velocity through the cortex of phalanges, radius, and tibia in normal and osteoporotic postmenopausal women using a new multisite quantitative ultrasound device," *Investigative Radiol.*, vol. 38, no. 4, pp. 207–211, Apr. 2003.

[28] A. Tatarinov, V. Egorov, N. Sarvazyan, and A. Sarvazyan, "Multifrequency axial transmission bone ultrasonometer," *Ultrasonics*, vol. 54, no. 5, pp. 1162–1169, Jul. 2014.

[29] E. C. Rose, M. Hagenmüller, I. E. Jonas, and B. A. Rahn, "Validation of speed of sound for the assessment of cortical bone maturity," *Eur. J. Orthodontics*, vol. 27, no. 2, pp. 190–195, Jan. 2005.

[30] H. Nguyen Minh, J. Du, and K. Raum, "Estimation of thickness and speed of sound in cortical bone using multifocus pulse-echo ultrasound," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 3, pp. 568–579, Mar. 2020.

[31] C. T. M. Eneh, M. K. H. Malo, J. P. Karjalainen, J. Liukkonen, J. Töyräs, and J. S. Jurvelin, "Effect of porosity, tissue density, and mechanical properties on radial sound speed in human cortical bone," *Med. Phys.*, vol. 43, no. 5, pp. 2030–2039, Apr. 2016.

[32] J. Fincke, X. Zhang, B. Shin, G. Ely, and B. W. Anthony, "Quantitative sound speed imaging of cortical bone and soft tissue: Results from observational data sets," *IEEE Trans. Med. Imag.*, vol. 41, no. 3, pp. 502–514, Mar. 2022.

[33] H. N. Minh, M. M'uller, and K. Raum, "Estimation of thickness and speed of sound for transverse cortical bone imaging using phase aberration correction methods: An in silico and ex vivo validation study," *Appl. Sci.*, vol. 12, no. 10, p. 5283, May 2022.

[34] A. Omar, S. Turan, and A. Bereket, "Reference data for bone speed of sound measurement by quantitative ultrasound in healthy children," *Arch. Osteoporosis*, vol. 1, nos. 1–2, pp. 37–41, Dec. 2006.

[35] Y. Tasinkevych, K. Falińska, P. A. Lewin, and J. Litniewski, "Improving broadband ultrasound attenuation assessment in cancellous bone by mitigating the influence of cortical bone: Phantom and in-vitro study," *Ultrasonics*, vol. 94, pp. 382–390, Apr. 2019.

[36] K. Chiba, R. Suetoshi, D. Cretin, T. Arai, T. Kawajiri, A. Okayama, S. Tsuji, N. Okazaki, M. Osaki, and K. Yoh, "Development of a QUS device to evaluate deterioration of cortical bone: Verification by HR-pQCT and measurements in healthy individuals and dialysis patients," *J. Clin. Densitometry*, vol. 24, no. 1, pp. 94–105, Jan. 2021.

[37] P. Ammann and R. Rizzoli, "Bone strength and its determinants," *Osteoporosis Int.*, vol. 14, no. S3, pp. 13–18, Mar. 2003.

[38] A. Dhainaut, M. Hoff, U. Syversen, and G. Haugeberg, "Technologies for assessment of bone reflecting bone strength and bone mineral density in elderly women: An update," *Women's Health*, vol. 12, no. 2, pp. 209–216, Mar. 2016.

[39] G. Rosi, V.-H. Nguyen, and S. Naili, "Numerical investigations of ultrasound wave propagating in long bones using a poroelastic model," *Math. Mech. Solids*, vol. 21, no. 1, pp. 119–133, Jan. 2016.

[40] O. Yousefian, Y. Karbalaeisadegh, and M. M'uller, "Modeling ultrasound attenuation in porous structures with mono-disperse random pore distributions using the independent scattering approximation: A 2D simulation study," *Phys. Med. Biol.*, vol. 64, no. 15, Aug. 2019, Art. no. 155013.

[41] Y. J. Yoon, J.-P. Chung, C.-S. Bae, and S.-Y. Han, "The speed of sound through trabecular bone predicted by biot theory," *J. Biomechanics*, vol. 45, no. 4, pp. 716–718, Feb. 2012.

[42] M. Tanter and M. Fink, "Ultrafast imaging in biomedical ultrasound," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 61, no. 1, pp. 102–119, Jan. 2014.

[43] J. Souquet and J. Bercoff, "Ultrafast ultrasound imaging," *Ultrasound Med. Biol.*, vol. 37, no. 8, p. S17, Aug. 2011.

[44] A. Petrescu, J. D'hooge, and J. Voigt, "Concepts and applications of ultrafast cardiac ultrasound imaging," *Echocardiography*, vol. 38, no. 1, pp. 7–15, Jan. 2021.

[45] H. Hasegawa and C. de Korte, "Special issue on ultrafast ultrasound imaging and its applications," *Appl. Sci.*, vol. 8, no. 7, p. 1110, Jul. 2018.

[46] H. H. Sultan, E. Grisan, P. Dryburgh, L. Peralta, and S. Harput, "Cortical bone thickness assessment from multi-frequency ultrasound RF data using a convolutional architecture with multi-head attention," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2023, pp. 1–4.

[47] K.-S. Lee, S.-K. Jung, J.-J. Ryu, S.-W. Shin, and J. Choi, "Evaluation of transfer learning with deep convolutional neural networks for screening osteoporosis in dental panoramic radiographs," *J. Clin. Med.*, vol. 9, no. 2, p. 392, Feb. 2020.

[48] K. Yoshida, Y. Tanabe, H. Nishiyama, T. Matsuda, H. Toritani, T. Kitamura, S. Sakai, K. Watamori, M. Takao, E. Kimura, and T. Kido, "Feasibility of bone mineral density and bone microarchitecture assessment using deep learning with a convolutional neural network," *J. Comput. Assist. Tomogr.*, vol. 47, no. 3, pp. 467–474, 2023.

[49] L. Mao, Z. Xia, L. Pan, J. Chen, X. Liu, Z. Li, Z. Yan, G. Lin, H. Wen, and B. Liu, "Deep learning for screening primary osteopenia and osteoporosis using spine radiographs and patient clinical covariates in a Chinese population," *Frontiers Endocrinol.*, vol. 13, Sep. 2022, Art. no. 971877.

[50] K. Yasaka, H. Akai, A. Kunimatsu, S. Kiryu, and O. Abe, "Prediction of bone mineral density from computed tomography: Application of deep learning with a convolutional neural network," *Eur. Radiol.*, vol. 30, no. 6, pp. 3549–3557, Jun. 2020.

[51] Y. Li, K. Xu, Y. Li, F. Xu, D. Ta, and W. Wang, "Deep learning analysis of ultrasonic guided waves for cortical bone characterization," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 68, no. 4, pp. 935–951, Apr. 2021.

[52] M. Fradi, E.-H. Zahzah, and M. Machhout, "Real-time application based CNN architecture for automatic USCT bone image segmentation," *Biomed. Signal Process. Control*, vol. 71, Jan. 2022, Art. no. 103123.

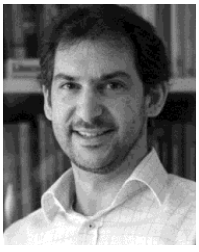
[53] C. Zhou, T. Jiang, K. Xu, and D. Ta, "Deep learning-based ultrasound computed tomography for cortical bone imaging," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2023, pp. 1–4.

[54] G. M. Treece, A. H. Gee, P. M. Mayhew, and K. E. S. Poole, "High resolution cortical bone thickness measurement from clinical CT data," *Med. Image Anal.*, vol. 14, no. 3, pp. 276–290, Jun. 2010.

[55] S. L. Croker, W. Reed, and D. Donlon, "Comparative cortical bone thickness between the long bones of humans and five common non-human mammal taxa," *Forensic Sci. Int.*, vol. 260, pp. 104.e1–104.e17, Mar. 2016.

[56] V. Bousson, A. Meunier, C. Bergot, É. Vicaut, M. A. Rocha, M. H. Morais, A.-M. Laval-Jeantet, and J.-D. Laredo, "Distribution of intracortical porosity in human midfemoral cortex by age and gender," *J. Bone Mineral Res.*, vol. 16, no. 7, pp. 1308–1317, Jul. 2001.
[57] C. S. Rajapakse, M. Bashoor-Zadeh, C. Li, W. Sun, A. C. Wright, and F. W. Wehrli, "Volumetric cortical bone porosity assessment with MR imaging: Validation and clinical feasibility," *Radiology*, vol. 276, Mar. 2015, Art. no. 141850.
[58] H. H. Sultan, E. Grisan, L. Peralta, and S. Harput, "Estimation of cortical bone strength using CNN-based regression model," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2022, pp. 1–4.
[59] E. Bossy, M. Talmant, and P. Laugier, "Three-dimensional simulations of ultrasonic axial transmission velocity measurement on cortical bone models," *J. Acoust. Soc. Amer.*, vol. 115, no. 5, pp. 2314–2324, May 2004.
[60] N. M. Salem, "Segmentation of white blood cells from microscopic images using K-means clustering," in *Proc. 31st Nat. Radio Sci. Conf. (NRSC)*, Apr. 2014, pp. 371–376.
[61] E. Boni, L. Bassi, A. Dallai, F. Guidi, V. Meacci, A. Ramalli, S. Ricci, and P. Tortoli, "ULA-OP 256: A 256-channel open scanner for development and real-time implementation of new ultrasound methods," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 63, no. 10, pp. 1488–1495, Oct. 2016.
[62] A. Vaswani, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–12.
[63] X. Yang, "An overview of the attention mechanisms in computer vision," *J. Phys., Conf. Ser.*, vol. 1693, no. 1, Dec. 2020, Art. no. 012173.
[64] F. Tian, L. Wang, and M. Xia, "Signals recognition by CNN based on attention mechanism," *Electronics*, vol. 11, no. 13, p. 2100, Jul. 2022.
[65] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, Sep. 2021.
[66] X. Wang and Q. Ni, "Determination of cortical bone porosity and pore size distribution using a low field pulsed NMR approach," *J. Orthopaedic Res.*, vol. 21, no. 2, pp. 312–319, Mar. 2003.
[67] H. H. Sultan, E. Grisan, L. Peralta, and S. Harput, "Classification of cortical bone thicknesses based on RF signal spectral analysis," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2022, pp. 1–4.
[68] K. Cho, Y. Lee, S. Lee, M. U. Shahid, K. J. Suh, and J. H. Choi, "Sonography of bone and bone-related diseases of the extremities," *J. Clin. Ultrasound*, vol. 32, no. 9, pp. 511–521, Nov. 2004.
[69] A. M. Richards, N. W. Coleman, T. A. Knight, S. M. Belkoff, and S. C. Mears, "Bone density and cortical thickness in normal, osteopenic, and osteoporotic sacra," *J. Osteoporosis*, vol. 2010, pp. 1–5, Jun. 2010.
[70] S. K. Ramchand and E. Seeman, "The influence of cortical porosity on the strength of bone during growth and advancing age," *Current Osteoporosis Rep.*, vol. 16, no. 5, pp. 561–572, Oct. 2018.

HOSSAM H. SULTAN received the B.Sc. and M.Sc. degrees in biomedical engineering from Helwan University, Cairo, Egypt, in 2012 and 2019, respectively. He is currently pursuing the Ph.D. degree in electrical and electronics engineering with the South Bank Applied Bioengineering Research, London South Bank University, London, U.K. He joined as a Postdoctoral Research Assistant with London South Bank University. His research interests include image processing, medical imaging, ultrasound bone characterization, and machine learning for medical diagnosis.



ENRICO GRISAN (Senior Member, IEEE) received the M.Sc. degree in electrical engineering from the University of Padua, Italy, in 2000, and the joint Ph.D. degree in bioengineering from the University of Padua and City University, London, U.K., in 2004, defending a thesis on automatic analysis of retinal images. In 2005, he was an Intern with Siemens Corporate Research, Princeton, NJ, USA, and then a Postdoctoral Fellow with the University of Padua, where he has been an Assistant Professor of biomedical engineering, since 2008. In 2019, he joined London South Bank University, as an Associate Professor of artificial intelligence. His current research interests include the understanding of medical imaged and identification of relevant biomarkers from medical data, either through classical image processing and analysis or through machine learning, with applications to neuroimaging, confocal microscopy and microendoscopy, and ultrasound. He has been a member of the IEEE Technical Committee in Biomedical Imaging and Image Processing, since 2015. He has served as the General Chair for IEEE ISBI 2019. He was an Associate Editor of IEEE ISBI and IEEE EMBC conferences.

PAUL DRYBURGH received the M.Eng. degree in mechanical engineering from The University of Edinburgh, U.K., in 2016, flowed by the University of Nottingham, U.K., in 2016. In 2022, he joined King's College London, London, U.K., where he currently holds the position of a Research Associate with the Department of Surgical and Interventional Engineering. His current research interests include biomedical applications of ultrasound, the development of multi-transducer modalities, and ultrasonic materials characterization.

LAURA PERALTA (Member, IEEE) received the B.Sc. degree in civil engineering, the dual M.Sc. degree in civil engineering and in structural mechanics, and the Ph.D. degree from the University of Granada, Granada, Spain, in 2008, 2012, and 2015, respectively. As a part of the Ph.D. program, she spent some time with the Institut Langevin, Paris, France, and University College London, London, U.K. From 2015 to 2017, she was with the Laboratoire d'Imagerie Biomédicale, Paris, as a Postdoctoral Researcher, working on quantitative ultrasound bone assessment. In 2017, she joined King's College London, London, as a Research Fellow, where she has been a Lecturer in physics and engineering of medical ultrasound with the Department of Surgical and Interventional Engineering, since 2022. Her current research interests include biomedical applications of ultrasound, beamforming methods, tissue characterization, and ultrasound propagation in complex media. She received the Wellcome/EPSRC Centre for Medical Engineering Fellowship, in 2021. In 2022, she became a Royal Society University Research Fellow.

SEVAN HARPUT (Member, IEEE) received the B.Sc. degree in microelectronics engineering and the M.Sc. degree in electronic engineering and computer sciences from Sabanci University, Turkey, in 2005 and 2007, respectively, and the Ph.D. degree from the University of Leeds, U.K., in 2013. After the Ph.D. study, he was a Researcher with Imperial College London, King's College London, and the University of Leeds. He is currently an Associate Professor with the Division of Electrical and Electronic Engineering, London South Bank University. His research aim is to develop new imaging and sensing technologies using acoustic waves at ultrasonic frequencies. His current research projects focus on high frame-rate ultrasound imaging, super-resolution methods, hard tissue characterization, ultrasonic non-destructive testing, and signal processing for biomedical imaging. He was an Associate Editor of IEEE TRANSACTIONS ON ULTRASONICS, FERROELECTRICS, AND FREQUENCY CONTROL, from 2018 to 2022.