Estimation of Cortical Bone Strength Using CNN-based Regression Model

**Hossam H. Sultan1, Enrico Grisan1, Laura Peralta2, Sevan Harput1**

1South Bank Applied Bioengineering Research, London South Bank University, London, UK

2School of Biomedical Engineering & Imaging Sciences, King’s College London, London, UK

E-mail: [abdelh13@lsbu.ac.uk](mailto:abdelh13@lsbu.ac.uk)

***Abstract*— Cortical bone is a compact layer that acts as a protective surface and forms an external layer of all bones. With osteoporosis, imbalance between bone formation and bone loss occurs, and this leads to a deterioration of bone microstructure including cortical bone thinning. Therefore, there is a clinical need to estimate and assess bone strength and quality. The detection of bone cortical thickness is still challenging due to the high variance in the speed of sound in the cortical bone.**

**The main aim of this study is to develop an accurate ultrasound method to estimate cortical bone thickness that could be used as a proxy of bone quality by using CNN-based regression models. To achieve this, pulse-echo measurements are performed at multiple ultrasound frequencies and the continuous wavelet transformations (CWT) of the acquired data was used as an input to the CNN. The maximum observed percentages were in 1 mm and 2 mm with an average error of 5.57%, and the minimum error was in group 7 (7mm) with a percentage of 1.6%. The preliminary results showed that combination of multi-frequency RF signals has potential to be used for cortical thickness estimation.**

*Keywords—* Bone characterization, Continuous wavelet transformation, Deep learning, Chirp signal, Regression models.

# Introduction

Bone is a mineralized tissue that could be classified into two main types, cortical (dense) and cancellous (trabecular or spongy) bone. It forms around 80% of the skeletal mass and is imperative to body structure and weight-bearing due to its resistance to bending and torsion, and it contains less than 10% of soft tissue. Bone as a living tissue is constantly levels between being broken down (resorption) and formed (rebuilding) [1, 2]. Furthermore, many risk factors could affect this balance, including gender, BMI, race, hormones levels,.. etc. For example, the bone loss process is usually faster than bone formation as people age [3]. Hence, the cortical bone porosity assessment and thickness measurement are crucial stages in evaluating bone strength and predicting bone diseases.

Many different imaging techniques could be used for this assessment. However, ultrasound has many advantages over other techniques. Ultrasound machines are relatively inexpensive, and patients are not exposed to ionizing radiation. Also, it is a non-invasive technique and can capture real-time images [4].

However, ultrasound bone strength assessment is an important and challenging criterion in bone disorders’ evaluation. As bone porosity increases, the reflected echoes from distal bone boundary heavily attenuate. This phenomenon affects speed of sound measurement, which is commonly used as a prior step to thickness estimation or beamforming data in traditional ultrasound bone assessment methods.

We hypothesize that multi-frequency ultrasound can play an essential role in bone strength assessment, exploring all responses as the ultrasound wavelength becomes comparable to both bone macro and micro-structures. In this study, to enable multi-frequency ultrasound imaging, coded signals (up-chirp) were used as an excitation pulse. They provide better SNR compared to non-coded excitation methods, such as a Gaussian pulse.

The proposed approach is an end-to-end technique to avoid the previously-mentioned problems and uses plane wave imaging that is suitable for fast acquisition. It comprises two major phases; the first phase is the simulation of ultrasound propagation in the cortical bone and then the transformation from RF signal to wavelet representation, while the second one involves hierarchal feature extraction based on different deep neural networks (Alexnet, Googlenet, Inception\_v3, Resnet, VGG16, and Densenet).

# Methods

## Simulation data

First, a 2-D finite-difference time-domain approach (Simsonic) [5] was used for cortical bone realistic simulation with varying porosity (0-20%) and thickness levels (1-8 mm) to mimic different bone sites and states. Figure 1 shows the simulation environment. The effective range of frequencies for each bone sample is a function of both porosity and thickness. Based on these ranges, a typical 128 elements linear transducer was employed, and chirp signals with multiple center frequencies (1-8 MHz) were used as excitation pulses to increase SNR [6].

Second, CNN-based network was trained with the continuous wavelet transformation coefficients (scalograms) [7, 8] (Figure 2) due to their ability to provide better time localization for short-term high-frequency events. CNNs can interpret the visual input by extracting hierarchical features and finally predict cortical bone thickness as continuous data.

A screenshot of a computer

Description automatically generated with medium confidence

Figure 1: Simulation environment shows the bone model with realistic pores extracted from ex vivo bone samples, curvature, pores weighting, and surface roughness.

Graphical user interface, application

Description automatically generated

Figure 2: (a) Ultrasound wave propagation shows the reflection from both proximal and distal bone interfaces, (b) received RF-data from a single channel showing pulse compressed signals coming from bone surfaces, and (c) a scalogram that reveals the absolute values of the continuous wavelet transform coefficients for an extracted RF signal.

## Convolutional neural networks

The CWT transformed RF signals coming from all channels are fed into a regression CNN. Different networks have been used as backbone for comparison. Table 1 shows the different topologies in terms of networks complexity and learnable parameters. The convolutional blocks are followed by a flattening and fully connected layers producing a high-level abstract value of the input scalogram.

During training, Mean Square Error (MSE) loss is used to optimize the network output with respect to the real bone thickness. For each observation, the MSE is given by the following equation:

,where  is the number of samples,  is the target value, and  is the system’s prediction for response . MSE is also known as L2 loss and used because most variables could be modeled into a Gaussian distribution. For image to one regression, the loss function of a single regression output is calculated by half of the mean squared error of the predictions, not normalized by as follows:

# results and discussion

The proposed models are trained on Intel i7-7700HQ CPU @ 2.80GHz , NVIDIA GeForce GTX 1060 (6 GB) GPU, 16.0 GB RAM. All models are validated using the simulated data (6,400 data points), resulted from 20 different channels. All scalograms, which are plotted as a function of time and frequency, have been resized as shown in Table 1. Then, the data have been partitioned into 20% and 80% for testing and training respectively. Based on the system convergence, 20 epochs have been trained with initial learning rate of 0.001, mini-batch size of 16, and SGDM optimizer. Root MSE has been used to estimate the difference between the predicted and actual output.

Table 1: Comparison between networks in terms of complexity and learnable parameters.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model/**  **Parameter** | **Alexnet** | **Google**  **net** | **Inception**  **V3** | **Resnet18** | **VGG**  **16** | **Dense**  **net** |
| **Depth** | 8 | 22 | 48 | 18 | 16 | 201 |
| **Input size** | 227 | 224 | 299 | 224 | 224 | 224 |
| **parameters (M)** | 61 | 7 | 23.9 | 11.7 | 138 | 20 |
| **Number of layers** | 25 | 144 | 315 | 71 | 47 | 707 |

When all the networks are compared, Alexnet showed a good compromise between the overall accuracy and training speed. The results showed that the proposed approach achieved an overall relative RMSE of 4.25%.

Figure 3 represents a box-plot of the network responses for all of the simulation data. Every data point reflects the response from different 32 test set (4 X 8 frequencies) channels for each specific porosity and bone thickness values. Numerically distant observations from the rest data appears on the plot which are overestimated in low thicknesses or underestimated in high thicknesses.

For low or high thickness values, some frequencies might not be suitable due to large wavelength or high attenuation constrains, respectively. However, combining multiple frequency acquisitions helped to increase the robustness of the model by reducing the outliers, which can be observed as reduction in the dispersion of the estimation around the predicted values in Figure 3. Moreover, in order to further reduce these outliers, different models have been applied and the median values showed the best fit within the expected bone thicknesses as demonstrated in Figure 4. Filtered data distribution has been able to decrease the whiskers of the boxplot and significantly reduce overlapping in terms of mean averages as shown in Figure 5.



Figure 3: Output distribution using Alexnet. Colour bands represent different thickness values. Every thickness covers porosity values of 0, 5, 10, 15, and 20%.

**

Figure 4: Comparison between system output based on different outliers-removal approaches.



Figure 5: Reproduction of Figure 4 after removing outliers using “Median” method.

Figure 6 (left) bar chart shows the absolute errors in mm. As the bone thickness is changing over the x-axis. The predicted values distribution demonstrates the ability of the proposed system to accurately predict cortical bone thickness with minimal error. In order to make the representation more uniform, Figure 6 (right) represents the error in percentage based on the cortical bone thickness of the sample. The maximum observed percentages were in 1 mm and 2 mm with an average error of 5.57%, and the minimum error was in group 7 (7mm) with a percentage of 1.6%.

Figure 6: The error values given in Figure 5 is replotted for a better visualisation , where (Left) is the absolute error in mm, (Right) is the error in percentage.

# Conclusion

This work presents a numerical study to predict cortical bone thickness based on a deep learning technique and the spectral analysis of ultrasound wave propagation in healthy and osteoporotic bone models. Multi-frequency pulse-echo approach enabled to acquire large dataset (a total number of 6,400 RF data points) which is a combination of 8 different thicknesses, 5 different porosities, and 8 frequencies at 20 different channels. Including real scanned data will help to increase the system reliability. Preliminary results showed that CWT of RF signals has potential for a segmentation free data analysis.

# References

[1] P. Augat, S. J. A. Schorlemmer, and ageing, "The role of cortical bone and its microstructure in bone strength," vol. 35, no. suppl\_2, pp. ii27-ii31, 2006.

[2] D. Cooper, C. Kawalilak, K. Harrison, B. Johnston, and J. Johnston, "Cortical bone porosity: what is it, why is it important, and how can we detect it?," *Current osteoporosis reports,* vol. 14, no. 5, pp. 187-198, 2016.

[3] F. Pouresmaeili, B. Kamalidehghan, M. Kamarehei, and Y. M. Goh, "A comprehensive overview on osteoporosis and its risk factors," *Therapeutics clinical risk management,* vol. 14, p. 2029, 2018.

[4] P. R. Hoskins, K. Martin, and A. Thrush, *Diagnostic ultrasound: physics and equipment*. CRC Press, 2019.

[5] E. Bossy and Q. Grimal, "Numerical methods for ultrasonic bone characterization," *Bone quantitative ultrasound,* pp. 181-228, 2011.

[6] H. Sultan, E. Grisan, L. Peralta, and S. Harput, "Classification of Cortical Bone Thicknesses Based on RF Signal Spectral Analysis," 2022.

[7] L. Aguiar-Conraria and M. J. Soares, "The continuous wavelet transform: A primer," NIPE-Universidade do Minho2011.

[8] J. Slavič, I. Simonovski, and M. Boltežar, "Damping identification using a continuous wavelet transform: application to real data," *J Journal of Sound Vibration,* vol. 262, no. 2, pp. 291-307, 2003.