Bone mineral density value evaluation based on photoacoustic spectral analysis combined with deep learning method

Xue Zhou (周 雪)[1,3], Zhibin Jin (金志斌)[2], Ting Feng (封 婷)[4], Qian Cheng (程 茜)[4], Xueding Wang (王学鼎)[4], Yao Ding (丁 尧)[1], Hongchen Zhan (詹洪陈)[1,], and Jie Yuan (袁 杰)[1,3,*]**

[1]*Jinling College, Nanjing University, Nanjing 210089, China*
[2]*Nanjing Drum Tower Hospital, Nanjing 210008, China*
[3]*School of Electronic Science and Engineering, Nanjing University, Nanjing 210093, China*
[4]*Institution of Acoustics, Tongji University, Shanghai 200092, China*
**Corresponding author: yuanjie@nju.edu.cn; **corresponding author: hongchen_zhan@jlxy.nju.edu.cn*
Received October 1, 2019; accepted December 19, 2019; posted online March 31, 2020

The diagnosis of osteoporosis is eventually converted to the measurement of bone mineral density (BMD) in clinical trials. Since our previous work had proved the ability of using photoacoustic spectral analysis (PASA) to efficiently detect osteoporosis, in this contribution, we proposed a fully connected multi-layer deep neural network combined with PASA to semi-quantify BMD values corresponding to varying degrees of bone loss and to further evaluate the degree of osteoporosis. Experiments were carried out on swine femur heads, and the performance of our proposed method is satisfying for future clinical screening.

Keywords: photoacoustics; osteoporosis; neural network.
doi: 10.3788/COL202018.041701.

Osteoporosis[1] is a skeletal disease characterized by abnormal microarchitectural arrangement of bone tissue as well as a decrease in bone mass. It is responsible for increased bone fragility and consequent high incidence of low-trauma hip, spine, and other fractures[2]. The fracture risk of patients with osteoporosis is up to 40% during their lifetime. From the perspective of patients, fractures may bring a lot of inconvenience to life and severe symptoms may even be life-threatening[3]. Moreover, viewed from a social standpoint, osteoporosis has already become an enormous public health problem, and the associated medical and socioeconomic costs have risen sharply[4] for the acceleration of the aging process and the arrival of the old-age society. Hence, timely diagnosis and effective treatment are of great significance to ease patients' suffering and reduce relevant costs.

The diagnosis methods can be generally divided into two categories—radiography methods and ultrasound methods[5]. Dual X-ray absorptiometry (DXA) is currently the most widely used radiography technique to assess mineral content in specified sites or in the whole skeleton. It can derive value of bone mineral density (BMD), which is the worldwide accepted gold standard for clinical diagnosis of osteoporosis. However, it only explains 60%–80% of bone strength, while other factors that also play important roles in determining fracture risk are not taken into account, such as bone microarchitecture, bone turnover rate, micro-damage, and bone mineralization. In addition to DXA, quantitative computed tomography (QCT)[6] is also one representative radiography method that can be available for BMD measurement of cortical or cancellous bone discretely. Moreover, the value measured by this method is three-dimensional volumetric BMD, which improves the sensitivity and accuracy of the measurement result. In comparison with DXA, a higher radiation exposure and a more expensive expense make QCT not conducive to clinical application. Non-ionizing and non-invasive quantitative ultrasound (QUS) technology is a typical representative of the ultrasound method. In both cross-sectional and prospective researches[7–9], QUS seems to be a predictor of osteoporotic fractures because its parameters are closely associated with the degree of bone loss. However, QUS can only measure BMD at peripheral sites such as the finger, wrist, and heel, and the measurement result is not that precise.

Photoacoustic (PA) imaging, also known as thermoacoustic imaging, is a hybrid imaging modality combining the advantage of both optics imaging and ultrasound imaging[10]. The attribute of high contrast derived from the sensitive absorption of optics in different tissues, the features of high spatial resolution, and great penetration depth are superiorities of ultrasound. It is proved that PA measurement possesses the unique capability of evaluating tissue microscopic architecture information at ultrasonic resolution, so its application field has rapidly been expanded to biomedicine[11–13]. Several researches aimed at studying frequency domain analysis of the broadband PA signals, also called PA spectral analysis (PASA), have already been done in the past few years to validate its potential ability to characterize bone microstructures[14–16]. In our previous study[17], the efficiency of the PASA method in osteoporosis assessment was proved by comparing the

quantified parameter slope from the PASA and broadband ultrasound attenuation (BUA) from QUS among different bone models. Both the simulation and *ex-vivo* experiment results showed that osteoporotic bone corresponded to a higher slope value, which meant the variation trend of bone loss was consistent with that of the slope derived from the PA spectrum. However, the above research still stayed in quantitative analysis and was not able to give quantitative diagnosis for varying degrees of osteoporosis. In practical terms, specific criteria about the ranges of slope values for varying degrees of osteoporosis were not formally defined and widely accepted. This time we wanted to go one step further to realize a semi-quantitative diagnosis for varying degrees of osteoporosis. Considering that the standard for BMD has already been well established and widely acknowledged, we put forward an idea to combine PASA with the deep learning network, where the BMD value was taken as the network output. To verify the feasibility of the proposed semi-quantitative method, we conducted *ex-vivo* experiments on swine femur heads in this study.

The relationship between the PA spectrum and corresponding BMD value is nonlinear, complicated, and difficult to be modeled or to be denoted by a mathematical expression precisely. The deep learning method provides one solution to fit the complex mapping relationship between network input and output with a high accuracy by offering sophisticated nonlinear models. Therefore, the neural network was introduced to this study. Specifically, the fitting process can be divided into three parts: neural network construction, network training, and network testing. In the network construction stage, both the layer number and layer node number can be customized by specific needs. In the process of training, the PA spectra and corresponding BMD values in training set were fed into the network as input and output, respectively. Then, the deep network continuously renewed the weights and bias in each layer through iteration to minimize the error between network output and the ground truth of BMD values until it achieved the presupposed precision. In the testing process, by comparing the predicted BMD values of the PA spectra in the test set with actual ones, the network performance could be evaluated, and the feasibility of the proposed method could be judged. In this study, we chose relative error as the performance evaluation index.

The N-layer network contains $N-1$ hidden layers and one output layer, where the output of each node in the previous layer acts as input of the next layer. What we processed was a one-dimensional signal of length 125, so a whole conjunction neural network, which meant that neighboring layers were fully connected by weighting values instead of a convolutional neural network, was adopted in this study.

Each layer node except the ones in the input layer mapped the sum of linear transformation to one output through the selected activation function. In our study, tan-sigmoid and purlin were chosen as the activation functions in the hidden layers and output layer, respectively.

To validate the idea that PA signals can contribute to the realization of semi-quantitative evaluation of osteoporosis, we performed *ex-vivo* experiments on swine femur heads. As shown in the diagram of the experimental system (Fig. 1), the specimen with an average thickness of around 5 mm was put vertically between a Q-switched neodymium-doped yttrium aluminum garnet (Nd:YAG) pulse laser and an ultrasonic transducer. The laser whose pulse width was about 8 ns was transmitted at a fixed wavelength of 532 nm to irradiate one side of the specimen. Then, the ultrasonic waves derived from the PA effect were detected by a line focused ultrasonic transducer (V310-SM, PANAMETRICS-NDT) put on the other side. The center frequency and −6 dB bandwidth of the transducer are 5 MHz and 4.4 MHz, respectively. After being amplified (SA-230F5, NF) and sampled (PCI-5101, NI) at a sampling rate of 60 MHz, the complicated, raw PA signals were acquired.

We prepared 30 swine specimen slices and collected two signals per slice; each of the signals was acquired by the laser's irradiation to one or the other side of the specimen. The corresponding BMD values were obtained by DXA. After the first round of signal acquisition, the 30 swine slices were soaked in an ethylene diamine tetraacetic acid (EDTA) solution (0.5 mol/L) for 24 h to imitate the process of bone loss (Fig. 2). After fetching out and going through the same acquisition process described above, another 60 raw PA signals and corresponding BMD values were obtained. Eventually, the number of raw signals added up to 120.

After eliminating the low signal-to-noise ratio (SNR) ones, the eventual number of PA signals used for building

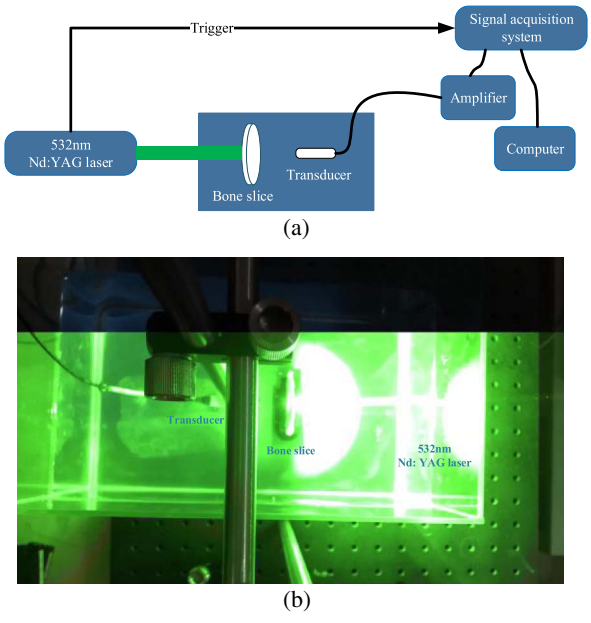

Fig. 1. (a) Schematic diagram of the experimental setup. (b) The image of actual experimental installation.

Fig. 2. Step of soaking swine slices into EDTA solution to imitate the process of bone loss.

data sets became 99; 87 of them were for the training set, and the remaining 12 were for the test set. Through careful observation and analysis of the time domain graph of PA signals, we found there existed a redundant component. After calculating the start and end points of the actual desired part according to the distances between the transducer and specimen, the thickness of the specimen, and the sound speed in the specimen, we cut the raw PA signals into a length of 1000. Considering that the maximum frequency of the PA signal is around 7 MHz, the sampling rate of 15 MHz is adequate, so we sampled each signal into four signals with a length of 250. This step was equivalent to converting the sampling rate from 60 MHz to 15 MHz and could function as the signal length reduction and data augmentation at the same time. As a result, 348 samples for the training set and 48 samples for the test set were obtained. Besides, for signals in the training set, we went on doing augmentation by adding Gaussian noise twice so that each signal obtained another two signals, and the samples for training were 1044 in total. Training of a network was essentially a procedure to adjust the parameters in it so as to make our model convergent. Considering the huge scale of network parameters, a proportional amount of samples were demanded. In order to make the network get the best performance, data augmentation was taken on the original dataset.

By means of Fourier transform, we got the signals in the frequency domain and went further to cut the length from 250 to 125 for signal spectrum symmetry. Moreover, the decrease in signal length is of benefit to the subsequent network training.

The nodes in the input and output layers were decided once the data sets were determined and were 125 and 1, respectively, in our study. The vital factors that had an effect on the performance of the neural network were the number of hidden layers and nodes in each hidden layer. The network could not possess the necessary capacity to learn and process information if the network depth (depending on the number of layers) or the number of nodes in each hidden layer was too small. Conversely, too large number would not only greatly increase the complexity of the network structure that may make the learning speed slower, but also make the network be more prone to causing problems such as gradient vanishing and overfitting. Eventually, a fully connected seven-layer network was selected as the optimal one through a preliminary experiment, and its specific structure is shown in Fig. 3. One diagram of the network input is shown in Fig. 4.

The BMD values measured by DXA were distributed in a range from 0.06 g/cm^2 to 0.22 g/cm^2. The comparison results calculated in test set between the network

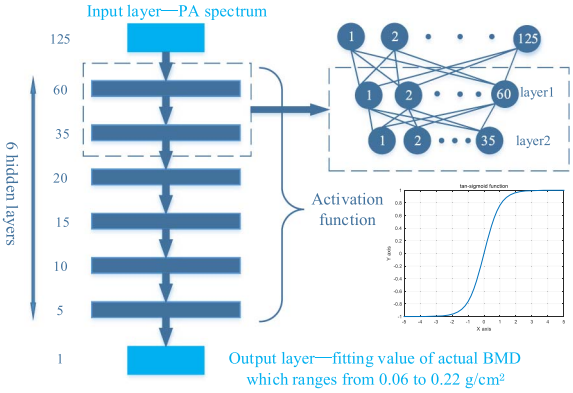

Fig. 3. Specific structure of adopted fully connected seven-layer network, and the details of layer 1 and layer 2 are presented in the right as an example.

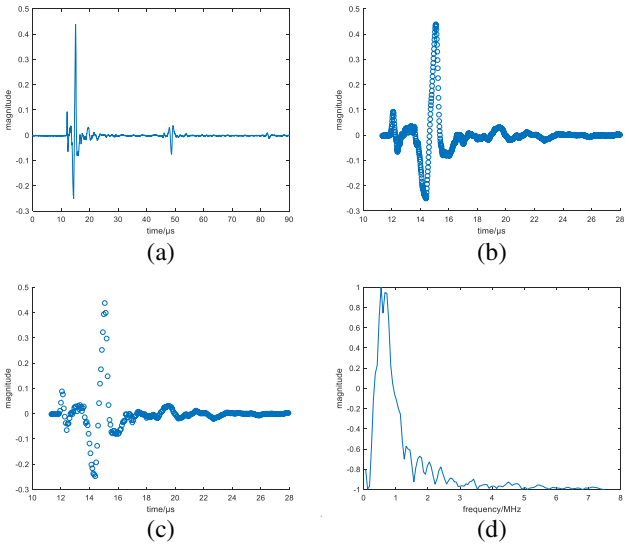

Fig. 4. Each diagram corresponds to one stage of signal processing. (a) The raw time domain PA signal. (b) Truncated time domain PA signal. (c) Time domain PA signal of further sampling in (b); it retains the information of (b) but has a larger interval between adjacent points (scatter diagrams are utilized to discriminate them). (d) The result of (c) after Fourier transform as an example illustration of the eventual normalized input of the network.

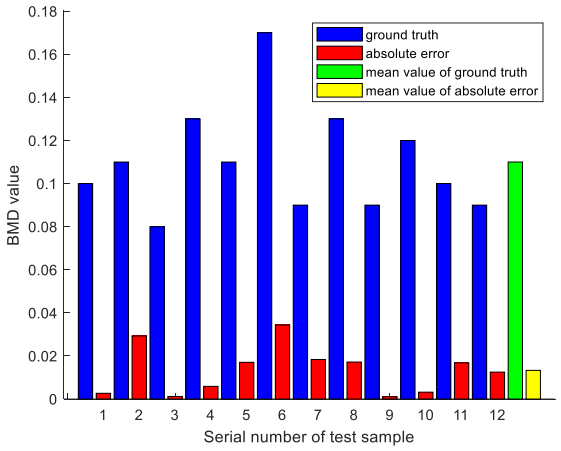

Fig. 5. Diagram of actual BMD value and its corresponding absolute error compared with the network fitting one.

predicted values and actual ones are shown in Fig. 5. The absolute errors ranged from 0.0011 g/cm^2 to 0.0344 g/cm^2, and the average value was 0.0133 g/cm^2. The relative error that could give a more intuitive cognition of the network performance for the actual BMD values was small, and the mean relative error was around 14%.

In this research, we verified the feasibility of utilizing a multi-layer neural network to fit BMD values based on the PA spectra. Compared with PASA method, which can only reveal a variation trend of bone loss, our method can realize semi-quantitative diagnosis for varying degrees of osteoporosis. Raw PA signals were collected by experiments conducted on swine femur heads. After various processing, the obtained PA spectra were input into a fully connected deep neural network to continuously update the parameters until the network was stably convergent. The eventual fitting results were satisfying, and our method's potential application value in clinical screening was proved by professional doctors. In our further study, a large data set that includes more samples of each individual BMD value and covers a larger range of BMD values will be built to improve the prediction accuracy and comprehensiveness.

This work was supported by the National Key Research and Development Program of China (No. 2017YFC0111402), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (No. 19KJB510031), and the Natural Science Foundation of Jiangsu Province (No. BK20181256).

References

1. T. D. Rachner, S. Khosla, and L. C. Hofbauer, Lancet **377**, 1276 (2011).
2. N. Harvey, E. Dennison, and C. Cooper, Nat. Rev. Rheumatol. **6**, 99 (2010).
3. C. Cooper, E. J. Atkinson, S. J. Jacobsen, W. M. O'Fallon, and L. J. Melton, III, Am. J. Epidemiol. **137**, 1001 (1993).
4. J. Leal, A. M. Gray, D. Prieto-Alhambra, N. K. Arden, C. Cooper, M. K. Javaid, and A. Judge, and The REFReSH study group, Osteoporosis Int. **27**, 549 (2016).
5. J. A. Kanis, Lancet **359**, 1929 (2002).
6. 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, J. Clin. Densitom. **11**, 123 (2008).
7. D. Hans and M. A. Krieg, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **55**, 1529 (2008).
8. C. C. Glüer and The International Quantitative Ultrasound Consensus Group, J. Bone Miner. Res. **12**, 1280 (1997).
9. C. C. Gluer, C. Y. Wu, M. Jergas, S. A. Goldstein, and H. K. Genant, Calcified Tissue Int. **55**, 46 (1994).
10. P. Beard, Interface Focus **1**, 602 (2011).
11. M. Xu and L. V. Wang, Rev. Sci. Instrum. **77**, 041101 (2006).
12. K. Y. Xia, X. H. Zhai, Z. H. Xie, K. Zhou, Y. T. Feng, G. J. Zhang, and C. H. Li, Chin. Opt. Lett. **16**, 121701 (2018).
13. Z. W. Cheng, H. G. Ma, Z. Y. Wang, and S. H. Yang, Chin. Opt. Lett. **16**, 081701 (2018).
14. G. Xu, I. A. Dar, C. Tao, X. J. Liu, C. X. Deng, and X. D. Wang, Appl. Phys. Lett. **101**, 221102 (2012).
15. B. Lashkari and A. Mandelis, J. Biomed. Opt. **19**, 036015 (2014).
16. T. Feng, J. E. Perosky, K. M. Kozloff, G. Xu, Q. Cheng, S. D. Du, J. Yuan, C. X. Deng, and X. D. Wang, Opt. Express, **23**, 25217 (2015)
17. W. Z. He, Y. H. Zhu, T. Feng, J. Yuan, Q. Cheng, G. Xu, and X. D. Wang, Chin. Opt. Lett. **15**, 111101 (2017).