Bone Reports 13 (2020) 100295

Contents lists available at ScienceDirect

Bone Reports

journal homepage: www.elsevier.com/locate/bonr

Prediction of trabecular bone architectural features by deep learning models using simulated DXA images

Pengwei Xiao[a], Tinghe Zhang[b], Xuanliang Neil Dong[c], Yan Han[a], Yufei Huang[b,*], Xiaodu Wang[a,**]

[a] Mechanical Engineering, University of Texas at San Antonio, United States of America
[b] Electrical and Computer Engineering, University of Texas at San Antonio, United States of America
[c] Health and Kinesiology, University of Texas at Tyler, United States of America

ARTICLE INFO

Keywords:
Trabecular bone microarchitecture
Deep learning
DXA
Histomorphometric parameters

ABSTRACT

Dual-energy X-ray absorptiometry (DXA) is widely used for clinical assessment of bone mineral density (BMD). Recent evidence shows that DXA images may also contain microstructural information of trabecular bones. However, no current image processing techniques could aptly extract the information. Inspired by the success of deep learning techniques in medical image analyses, we hypothesized in this study that DXA image-based deep learning models could predict the major microstructural features of trabecular bone with a reasonable accuracy. To test the hypothesis, 1249 trabecular cubes (6 mm × 6 mm × 6 mm) were digitally dissected out from the reconstruction of seven human cadaveric proximal femurs using microCT scans. From each cube, simulated DXA images in designated projections were generated, and the histomorphometric parameters (*i.e.*, BV/TV, BS, Tb.Th, DA, Conn. D, and SMI) of the cube were determined using Image J. Convolutional neural network (CNN) models were trained using the simulated DXA images to predict the histomorphometric parameters of trabecular bone cubes. The results exhibited that the CNN models achieved high fidelity in predicting these histomorphometric parameters (from $R = 0.80$ to $R = 0.985$), showing that the DL models exhibited the capability of predicting the microstructural features using DXA images. This study also showed that the number and resolution of input simulated DXA images had considerable impacts on the prediction accuracy of the DL models. These findings support the hypothesis of this study and indicate a high potential of using DXA images in prediction of osteoporotic bone fracture risk.

1. Introduction

Approximately 53.6 million US adults had osteoporosis or osteopenia in 2010, representing 54% of the US adult population aged 50 years or older (Wright et al., 2014). Osteoporosis can cause bone fragility fractures, thus significantly increasing the risk of morbidity and even mortality to the patients (Campion and Maricic, 2003). The direct cost of treating the skeletal disease has tripled from $28.1 billion between 1998 and 2000 to $73.6 billion between 2012 and 2014 in US alone (Wright et al., 2019). Hence, early prognosis of osteoporotic bone fractures not only helps prevent such fractures but also helps ease the financial burden to the patients and society.

Osteoporotic bone fractures are commonly associated with both bone mass loss and bone quality deterioration (*Am J Med*, 1991). Currently, bone mineral density (BMD) is usually used to predict the risk of osteoporotic bone fractures. Since BMD is only a measure of bone mass loss, use of BMD alone could only describe 50–60% of osteoporotic bone fractures (Siris et al., 2004), thus intensifying the necessity to take into account bone quality in prediction of bone fracture risks. It is well documented that bone quality is mainly dependent on the microstructural and material properties of the tissue (Seeman and Delmas, 2006). However, the advanced methodologies for assessing microstructural and material properties of bone *in vivo* are yet to emerge for clinical applications.

Clinically, several biomedical imaging modalities (*e.g.* QCT, hr-pQCT, MRI, and DXA) have been employed to assess bone fracture risks. Among them, dual-energy X-ray absorptiometry (DXA) is the most affordable and convenient biomedical imaging modality for the purpose. DXA not only gives rise to an accurate assessment of areal BMD (Dong and Wang, 2013), but also exhibits a potential in revealing bone microstructural features. DXA image texture analyses (Le Corroller et al., 2012), including trabecular bone score (TBS) (Pothuaud et al., 2007),

* Correspondence to: Y. Huang, Electrical and Computer Engineering, University of Texas at San Antonio, TX 78249, United States of America.
** Correspondence to: X. Wang, Mechanical Engineering, University of Texas at San Antonio, TX 78249, United States of America.
E-mail addresses: yufei.huang@utsa.edu (Y. Huang), xiaodu.wang@utsa.edu (X. Wang).

https://doi.org/10.1016/j.bonr.2020.100295
Received 25 June 2020; Accepted 2 July 2020
Available online 08 July 2020

fractal analysis (Chappard et al., 2005), and correlation length analysis (Dong et al., 2015; Dong et al., 2013), have shown strong correlations between the DXA image texture and microstructural features of trabecular bone. Although it is presumable that DXA images contain the information of microstructural features, the efficacy of the current DXA image-based techniques in extracting the microstructural properties is still debatable. Hence, it necessitates the use of advanced image processing techniques to extract the microstructural information from DXA images.

In recent years, deep learning (DL) has achieved tremendous success in the field of biomedical image processing, such as image classification, object or lesion detection, organ and substructure segmentation, and registration (Litjens et al., 2017). Moreover, researchers have successfully employed DL techniques in image-based diagnosis/prognosis of different diseases, such as skin cancer, Alzheimer disease, and glaucoma (Esteva et al., 2017; Ding et al., 2019; Chen et al., 2015). These successes have motivated us to investigate the possibility of applying DL techniques to extract the microstructural information of trabecular bone from DXA images.

In this study, we hypothesized that DL models could be trained using DXA images to predict major microstructural features of trabecular bone. To test this hypothesis, we trained DL (convolutional neural network) models using simulated DXA images of actual bone samples and verified its efficacy in predicting the histomorphometric properties of trabecular bone. We also investigated the effect of the resolution and number of input simulated DXA images on the prediction accuracy of the DL models.

2. Materials and methods

2.1. Preparation of trabecular bone specimens and microCT image-based reconstruction

Seven human cadaver proximal femurs were acquired from a diverse pool of donors with different age, gender, and microstructural features (Table 1). Among the femurs, six were scanned using a SkyScan 1173 microCT system with a resolution of 35 μm, and the remaining one was scanned using a Scanco ViVaCT 40 microCT system with a resolution of 21 μm. One thousand two hundred and forty nine ($N = 1249$) trabecular bone cubes (6 mm × 6 mm × 6 mm) were digitally dissected out from femoral head, femoral neck, and greater trochanter regions of the reconstructed proximal femurs from microCT scans (Fig. 1).

2.2. Measurement of histomorphometric parameters

To capture the microstructural features of trabecular bone, six histomorphometric parameters were measured from the trabecular bone cubes using ImageJ (1.52 h) and BoneJ. These parameters include the degree of anisotropy (DA), connectivity density (Conn.D), bone surface (BS), structure model index (SMI), bone volume fraction (BV/TV), and trabecular thickness (Tb.Th) (Table 2). Among the parameters, BV/TV reflects the bone mass per volume, Tb.Th defines the size of trabeculae, BS is an indicator of both trabecular size and number, Conn.D is an indicator of the trabecular number, SMI defines whether the trabecular structure of the bone sample is either plate-like or rod-like, and DA measures the trabecular orientation. The Pearson correlation coefficients between these histomorphometric parameters measured from the trabecular bone cubes were calculated to illustrate the interdependence of the parameters (Table 3). Tb.Th, BS, and SMI were strongly correlated with BV/TV, whereas DA and Conn.D exhibited much weaker correlation with BV/TV.

2.3. Simulated DXA images

In this study, DXA images were not obtained from experimental measurements, but from *in silico* simulations. Simulated DXA images were digitally generated for each trabecular bone cube by projecting the voxel-based model of trabecular cubes onto a plane perpendicular to a designated projection axis using custom MATLAB (MathWorks, Natick, MA) scripts (Patil and Ravi, 2005) (Fig. 1). Briefly, the voxel-based model was converted into a stack of plane layers with a thickness of one voxel. Then, the simulated DXA image was generated by summing the binarized values of voxels in each layer that fall onto each bin (pixel) in the projection plane (Fig. 1). The greyscale value (Z) of each pixel at the location (x_i, y_i) in the simulated DXA image (projection plane) was determined by the following equation (Dong et al., 2013):

$$Z(x_i, y_i) = \frac{1}{N}\sum_{k=1}^{N} V(x_i, y_i, z_k) \quad (1)$$

where, $Z(x_i, y_i)$ represents the greyscale value of the ith bin at the location (x_i, y_i) in the projection plane; $V(x_i, y_i, z_k)$ is the summation of binarized voxels in the bin at the location (x_i, y_i, z_k); N is the number of stacks of layers in the projection direction (z). The resolution of DXA images could be readily adjusted by altering the pixel (bin) size in the projection plane.

To investigate the effect of the resolution of simulated DXA images on the prediction accuracy of the DL models, five different resolutions of simulated DXA images were generated from coarse to fine by changing the bin size. The resolutions were 3.0 mm (2 × 2 pixels), 2.0 mm (3 × 3 pixels), 1.2 mm (5 × 5 pixels), 0.3 mm (20 × 20 pixels), and 0.075 mm (80 × 80 pixels), respectively.

This study also intended to investigate the effect of the number of simulated DXA images on the prediction accuracy of the DL models. Nine projection directions, including x, y, z, x ± 45°, y ± 45°, and z ± 45°, were used to generate the simulated DXA images. Specifically, as the input of the DL models, simulated DXA images were obtained from a single projection (along either x, or y, or z-axis), triple-projections (along x, y, and z-axis), sextuple-projections (along x, y, z, x + 45°, y + 45°, and z + 45°), and nonuple-projections (along x, y, z, x ± 45°, y ± 45°, and z ± 45°), respectively (Fig. 1).

2.4. Training of the deep learning model

A popular DL model called convolutional neural network (CNN) (Kandi et al., 2017) was implemented in this study because CNN is the

Table 1
General information of trabecular bone samples from human cadaveric proximal femurs ($N = 1249$).

Age (years)	Sex	MicroCT resolution (μm)	No. of cubes	DA	Conn.D (mm^{-3})	BS (mm^2)	SMI	BV/TV	Tb.Th (mm)
22	F	35	141	0.47 ± 0.11	3.31 ± 1.38	600 ± 217	2.29 ± 2.22	0.32 ± 0.21	0.32 ± 0.14
41	F	35	137	0.60 ± 0.16	2.14 ± 1.42	447 ± 299	3.76 ± 2.24	0.21 ± 0.17	0.25 ± 0.08
82	F	35	79	0.60 ± 0.12	2.94 ± 1.21	523 ± 186	3.44 ± 1.22	0.20 ± 0.10	0.25 ± 0.05
42	F	21	234	0.67 ± 0.15	2.12 ± 1.36	346 ± 177	5.72 ± 2.75	0.08 ± 0.05	0.16 ± 0.03
24	M	35	239	0.51 ± 0.14	3.72 ± 1.35	641 ± 166	2.51 ± 1.64	0.29 ± 0.15	0.28 ± 0.13
43	M	35	174	0.49 ± 0.13	3.37 ± 1.21	589 ± 212	2.71 ± 1.55	0.27 ± 0.15	0.27 ± 0.07
79	M	35	245	0.59 ± 0.15	2.63 ± 1.76	499 ± 287.	2.70 ± 2.44	0.26 ± 0.22	0.29 ± 0.12

Note: The parameters are presented as mean ± S.D.

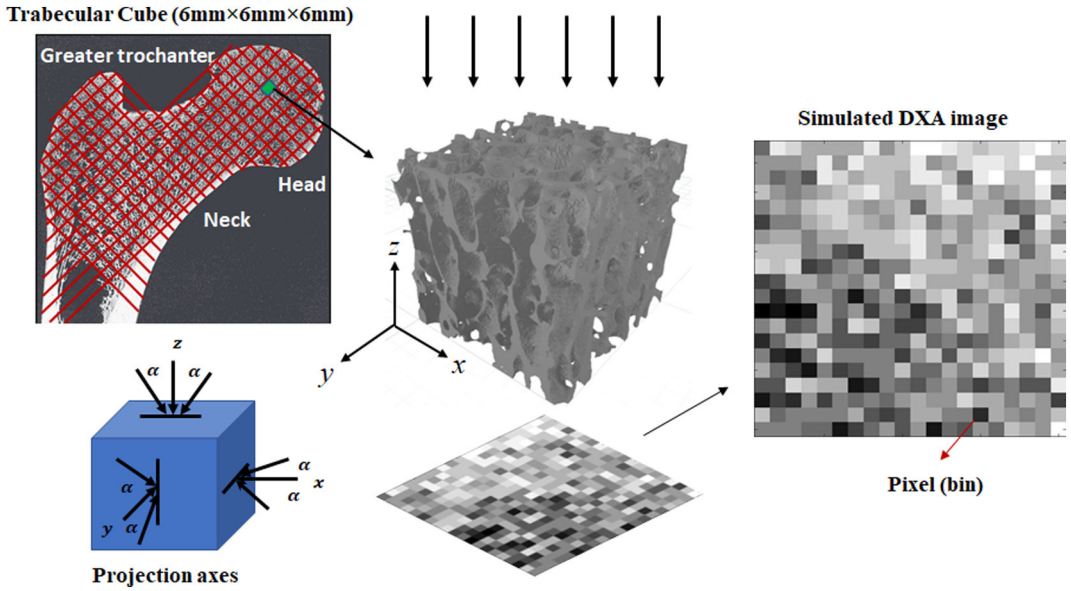

Fig. 1. Schematic representation of dissecting out trabecular bone cubes (6 mm × 6 mm × 6 mm) from the femoral head, neck and greater trochanter regions and acquisition of simulated DXA images from the cubes. The projection directions were designated in x, y, z, x ± 45°, y ± 45°, and z ± 45°. The pixel size of the simulated DXA images varied from 0.075 to 3.0 mm by changing the bin size for the projections.

Table 2
Trabecular microstructural properties and parameters.

Microstructural features	Histomorphometric parameters	
Bone mass	BV/TV (bone volume fraction)	
Trabecular size	Tb.Th (trabecular thickness)	BS (bone surface)
Trabecular number	Conn.D (connectivity density)	
Trabecular type	SMI (structure model index)	
Trabecular orientation	DA (degree of anisotropy)	

Table 3
Pearson's correlation matrix for the six-histomorphometric parameters ($N = 1000$).

Histomorphometric parameters	DA	Conn.D	BS	SMI	BV/TV	Tb.Th
DA	1	−0.426	−0.420	0.366	−0.430	−0.321
Conn.D	–	1	0.824	−0.497	0.575	0.281
BS	–	–	1	−0.775	0.876	0.595
SMI	–	–	–	1	−0.862	−0.737
BV/TV	–	–	–	–	1	0.823
Tb.Th	–	–	–	–	–	1

Note: Correlation is significant at the 0.01 level (2-tailed).

most widely used DL model in image related predictions. The architecture of the proposed CNN model is illustrated in Fig. 2. It comprised of three convolutional layers, three max-pooling layers, one fully connected neural network layer with 64 neurons, followed by the output layer (Table 4). In the CNN model, we used unpadded 3 × 3 convolutions and 2 × 2 max-pooling with stride 1 to decrease the dimension of the feature maps, while maintaining the most important features. During the training process, the simulated DXA images were used as the input and each of the six histomorphometric parameters (*i.e.*, BV/TV, BS, Tb.Th, DA, Conn.D, and SMI) measured from the trabecular bone cubes were used as the output (ground true). Thus, six CNN models were trained for these six parameters separately. The mean square error (MSE) was employed as loss function, which is defined as

$$MSE = \frac{1}{2}\sum_{i=i}^{n}(y_i - \hat{y}_i)^2 \quad (2)$$

where, y_i is the measured histomorphometric parameter and $\hat{y}_i$ is the predicted parameter, and n is the total number of samples (or sample size). The stochastic gradient algorithm with the ADAM optimizer was used to train the CNN models. Moreover, hyperparameter optimization was performed to tune the architecture of the CNN model for the optimal performance. The evaluated parameters of CNN architecture included the number of convolutional layers, the number of fully

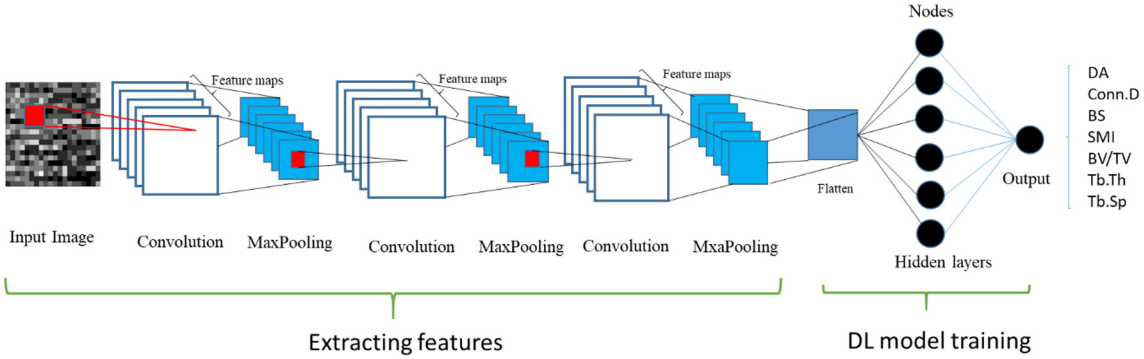

Fig. 2. The schematic architecture of CNN model for prediction of microstructural parameters of the trabecular bone cubes. Simulated DXA images were used as input and the histomorphometric parameters as the ground truth. The convolution layers were used to extract input features and the hidden layers were used to train the DL models with ground truth-values.

Table 4
CNN architecture features for predicting different histomorphometric parameters.

Model	Kernel size	Pool size	Convolutional layers	No of hidden layers	Learning rates	No. of epochs	No. of filters	Dropout	Output
1	3 × 3	2 × 2	(8,16,32)	1	0.0001	250	64	0.2	BV/TV Tb.Th BS SMI
2	5 × 5	2 × 2	(16,16,16)	1	0.0001	150	64	0.4	Conn.D
3	3 × 3	2 × 2	(8,8,8)	1	0.0001	200	64	0.5	DA

connected layers, the number of filters and kernel size, the optimizer functions, the learning rates, the number of epochs, and the value of dropout (Table 4). In the end, the best parameters for the architecture of CNN model were finalized.

The six histomorphometric parameters were normalized by using rescaling (min-max normalization) before training the CNN models. Rescaling is an efficient method for normalizing the data and scales the range of trained features in [0, 1]. In addition, the rescaling keeps the relationship among original data, which works well for the prediction model. The rescaling was calculated by the following formula:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (3)$$

where x is an original value, x' is the normalized value.

The performance of CNN model was verified by using the cross-validation approach during the training process. 80% of the simulated DXA images and the associated histomorphometric parameters were randomly selected from the entire dataset ($N = 1249$) as the *training data* to determine the parameters of the CNN model. Of the *training data*, 80% were randomly selected to train the CNN model, whereas the remaining 20% were used as the validation set to learn the model hyperparameters. Finally, the remaining 20% of data were used as the *testing data* to determine the performance of the CNN models in predicting the microstructural features of trabecular bone cubes (Table 5).

All CNN models were programmed in Python using Keras library with a TensorFlow backend and were trained on a Dell desktop computer (XPS 8930, Intel Core i9-9900k 8-Core Processor, 64GB Memory, NVIDIA® GeForce(R) GTX 1080 with 8GB GDDR5X Graphic Memory).

2.5. Data analysis

Statistical analyses were performed using JMP PRO 15 (SAS software, United States). Linear regression analyses were implemented between the DL model-predicted and measured histomorphometric parameters. The prediction accuracy of the DL models was assessed using Pearson correlation coefficient (R) of the regressions. The statistical significance was considered only if $p < 0.05$.

3. Results

3.1. Comparison between the predicted and measured microstructural features

The linear regression analyses indicated that strong correlations existed between the histomorphometric parameters predicted by the DL models using simulated DXA images and those measured directly from the trabecular bone cubes (Fig. 3). The Pearson correlation coefficient (R) was 0.998, 0.971, 0.976, 0.873, 0.833, and 0.802 for BV/TV, BS, Tb.Th, SMI, Conn.D and DA, respectively, with all the p-values being smaller than 0.0001. Using R-values as the measure of prediction accuracy of the DL models, the results suggested that the DL models had very high accuracy in predicting BV/TV, BS, and Tb.Th, whereas its prediction accuracy was relatively lower for SMI, Conn.D and DA. In addition, BV/TV, BS, Tb.Th, Conn.D, and DA values predicted by the DL models appeared to have strong linear relationships with those measured directly from the trabecular bone cubes (Fig. 3). Only SMI demonstrated a nonlinear relationship between the model-predicted and measured values. Nonetheless, a linear portion could be observed on the regression curve when the SMI values ranged between *0* and *4*, suggesting that the DL models could predict SMI only within this range.

3.2. The effect of the resolution of images

Simulated DXA images with five resolutions, *i.e.*, 3 mm/pixel, 2 mm/pixel, 1.2 mm/pixel, 0.3 mm/pixel, and 0.075 mm/pixel, were used to train the DL models. The results indicated that the prediction accuracy of the DL models increased with increasing resolutions for all histomorphometric parameters (BV/TV, BS, Tb.Th, Conn.D, SMI and DA) (Fig. 4). In addition, the prediction accuracy of the DL models converged as the simulated DXA image resolution increased, and eventually leveled off after the resolution approached 0.3 mm/pixel (for all parameters except for Conn.D) or even 1.0 mm/pixel (for BV/TV and DA). This observation suggests that image resolution would have a limited effect on the prediction accuracy after it reaches to a certain level (Fig. 4). In contrary, the prediction accuracy of the DL model for Conn.D increased almost linearly with the image resolution.

3.3. The effect of the number of simulated DXA images on the prediction accuracy

The prediction accuracy of the DL models increased as the input number of simulated DXA images increased (Fig. 4). In addition, its effect was dependent on the resolution of the simulated DXA images, showing that increasing the number of input simulated DXA images had very limited effects on the prediction accuracy of the DL models when the simulated DXA image resolution was at or smaller than 1.2 mm/pixel. Moreover, the resolution of simulated DXA images appeared to have very limited effects on the prediction accuracy of the DL model in predicting SMI if more than six DXA images were used as input.

4. Discussion

This study investigated the possibility of using DL models to predict

Table 5
Datasets used for training and testing the CNN model.

Training phase	Inputs		Outs	
Training	80% randomly selected DXA images		Corresponding histomorphometric parameters	
	80% for training	20% for verifying	80% for training	20% for verifying
Testing	The remaining 20% DXA images		Predicted histomorphometric parameters	

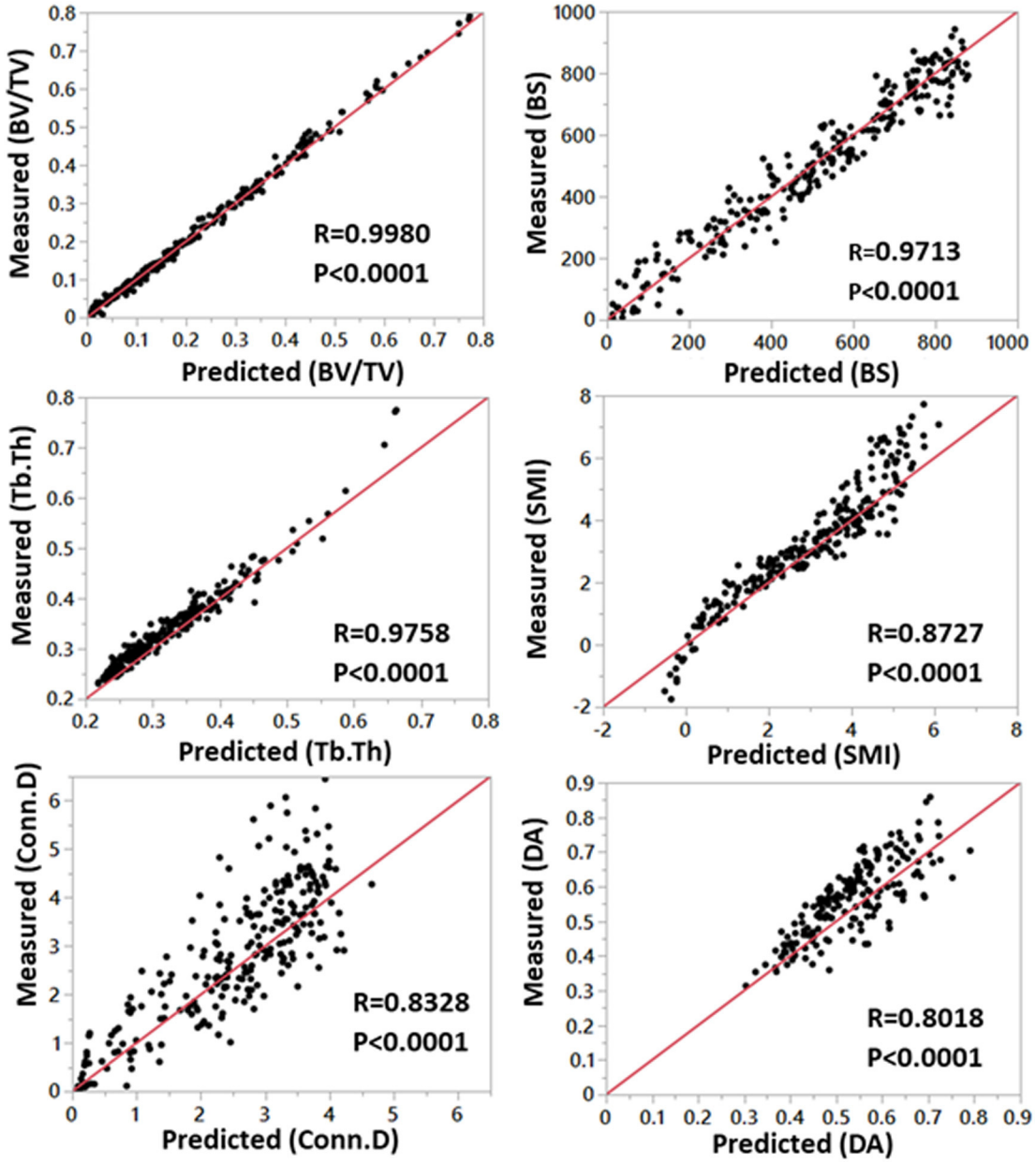

Fig. 3. The regression plots of histomorphometric parameters between the DL model-predicted and measured values. High prediction accuracy was obtained for BV/TV, Tb.Th, and BS (R = 0.971–0.998), whereas the prediction accuracy was relatively lower for SMI, Conn.D and DA (R = 0.802–0.873). The total number of samples was N = 1,249, with 80% as training and 20% as testing data. The resolution of DXA images was 0.3 mm. Nine simulated DXA images were used for BV/TV, Tb.Th, BS, SMI, and Conn.D, whereas two simulated DXA images were used for DA.

the microstructural features (*i.e.*, BV/TV, Tb.Th, BS, Conn.D, SMI, and DA) of trabecular bone using simulated DXA images. CNN models trained using the simulated DXA images of 1249 trabecular bone cubes from seven human cadaveric femurs exhibited high fidelities (R = 0.8–0.998) in predicting the microstructural features, thus supporting the hypothesis that DXA images could be used to train DL models for predicting the microstructural features of trabecular bone.

Our hypothesis was based on a large number of evidence reported in the literature (Chappard et al., 2005; Pothuaud et al., 2000; Benhamou et al., 2001). Pothuaud et al. (2000) reported strong correlations between the fractal dimension of two-dimensional projections from the 3D MRI image and the microstructural features of a trabecular bone. These projections are similar to the simulated DXA images used in this study. However, their study was limited only to predicting porosity and connectivity, without taking into account other microstructural parameters (*e.g.* DA, SMI, Tb.Th, and BS). Chappard et al. (2005) carried out texture analyses of X-ray radiographs and found significant correlations between the texture parameters and BV/TV, Tb.N (trabecular number), and Tb.Sp. Apostol et al. also performed texture analysis of 2D projection images and used multiple regression analyses to predict 3D micro-architecture from the 2D texture patterns (Apostol et al., 2006). They found strong correlations between the estimated and measured microstructural parameters of trabecular bone. Other previous results also showed that trabecular bone score (TBS) (Pothuaud et al., 2007; Muschitz et al., 2015; Hans et al., 2011; Winzenrieth et al., 2013) and variogram based stochastic analyses (Dong et al., 2015; Dong et al., 2013) of DXA images could also provide assessments of microstructural parameters of trabecular bone. This study provides the first evidence to show that DL models could be trained using DXA images to assess microstructural properties of trabecular bone. More importantly,

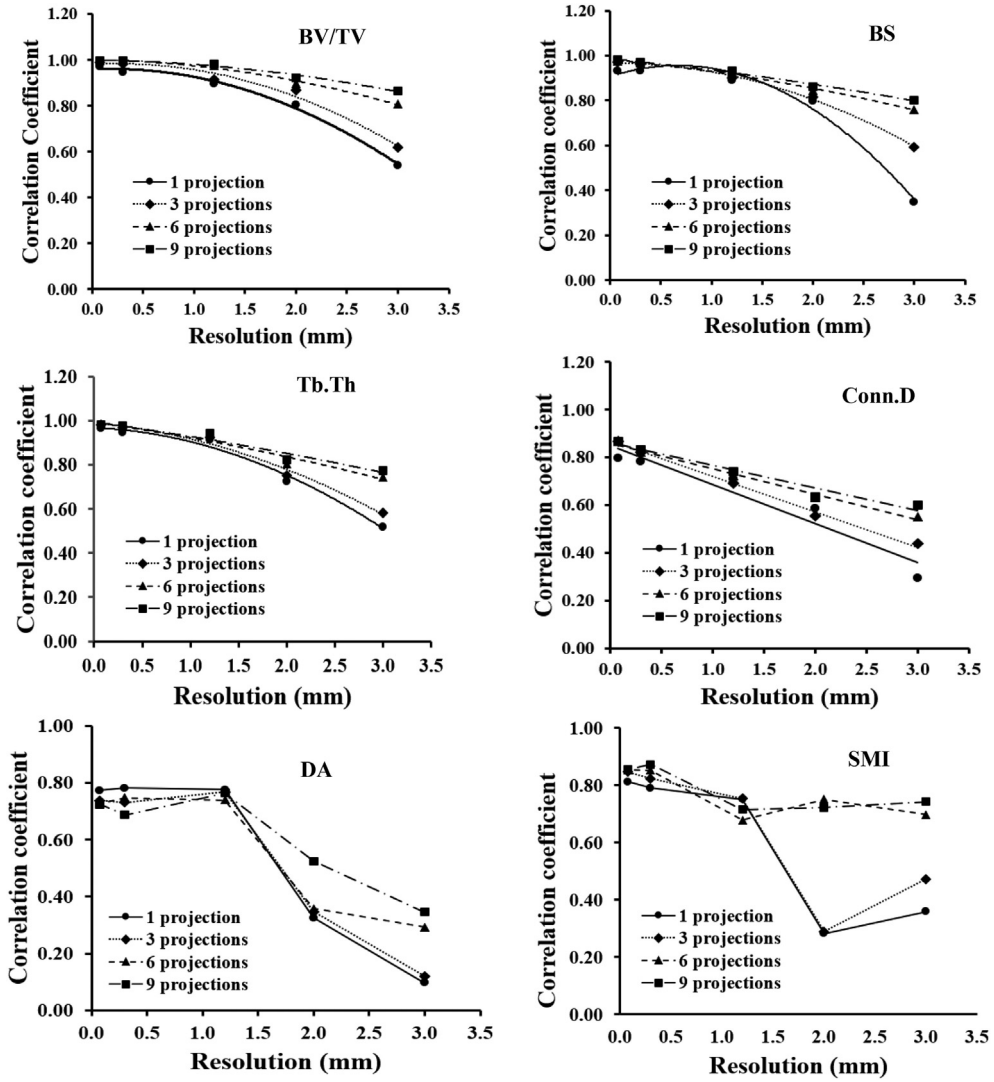

Fig. 4. Effect of the number and resolution of DXA images on the prediction accuracy of the DL models. The total number of samples was $N = 1249$ for developing the DL models for all the parameters, with 80% as training and 20% as testing data.

the DL models are DXA image based, thus making it more suitable for clinical applications.

The results of this study demonstrate that the low-resolution DXA image-based DL techniques could predict major histomorphometric parameters of trabecular bone with much better accuracy than those estimated using the BV/TV based regression models (Table 6). Thus, it is presumable that the microstructural features of trabecular bone can be extracted not only from the inherent correlations of these parameters with BV/TV, but also from DXA image textures. Among the histomorphometric parameters investigated in this study, BV/TV, BS, and Tb.Th are predicted very accurately by the DL models ($R = 0.976–0.998$). For BV/TV, it is not surprising because the simulated DXA images themselves could directly provide an accurate assessment of BV/TV. For BS and Tb.Th, since they are strongly correlated to BV/TV (Parkinson and Fazzalari, 2003; Boutroy et al., 2005; Greenwood et al., 2015) it is anticipated that the prediction accuracy of the DL models for these parameters is augmented by taking into account the inherent correlation of the parameters with BV/TV.

In contrast, the prediction accuracy of the DL models for Conn.D, DA, and SMI is relatively lower ($R = 0.802–0.873$). Among the parameters, Conn.D is estimated based on the Euler characteristics of the trabecular structure and is significantly correlated with the number of trabeculae in the structure (Odgaard and Gundersen, 1993). The large scattering of data around the regression curve (Fig. 3) and hampered prediction accuracy (around 0.83) suggest that DXA images may not provide sufficient information for the DL model to fully capture the topological characteristics of Conn.D in the trabecular bone samples.

Next, DA is an overall measure of trabecular orientation in bone, which is estimated using Mean Intercept Length (MIL) method (Odgaard, 1997; Harrigan and Mann, 1984). This study shows that the prediction accuracy of the DL model for DA is reasonably good but

Table 6
Comparison of prediction accuracy (R) between BV/TV based regression models and DXA image based DL models.

Trabecular microstructural features	Regression models with BV/TV as independent variable				DXA image based DL models
	Greenwood et al. (2015)	Kim and Henkin (2015)	Kabel et al. (1999)	This study	
Tb.Th	0.760	0.18	–	0.823	0.976
SMI	0.648	0.74	–	0.876	0.873
Conn.D	–	0.76	0.200	0.575	0.832
DA	0.265	–	–	0.430	0.802
BS	–	–	0.842	0.876	0.971

limited to around 0.8, suggesting that the DL model can predict, to some but not full extent, DA in the trabecular bone samples (Fig. 3). A possible reason is that since trabecular bone usually has an orthogonally anisotropic structure, it necessitates the alignment of the primary axes of the samples with a common coordinate system in order to make DXA projections consistent with the primary directions for all samples. If mismatch of the primary axes of trabecular cubes with the cube coordinates exists, increases of sample size would tend to intensify the uncertainties and blunt the sensitivity of the DL model to predict DA.

In addition, the nonlinear relationship between the DL model predicted and the measured SMI values suggests that the prediction by the DL model would be inaccurate when SMI values are outside the range of 0 to 4 (Fig. 3). SMI is a structural index showing the plate- or rod-like structures, which varies from 0 (plate-like), to 3 (rod like), and to 4 (sphere like) based on the assumption that trabecular surfaces are convex (Salmon et al., 2015). Due to this assumption, SMI becomes negative when the surfaces are concave, thus making it difficult to define the trabecular components with concave surfaces. Therefore, SMI values outside the range of 0 to 4 actually do not have physical meaning in defining trabecular structures. These intrinsic limitations of SMI could be a major reason for loss of prediction accuracy by the DL models. However, a strong linear relationship between the DL model predicted and measured values does exist when SMI is within the range from 0 to 4, thus suggesting that the DL model still can be used to predict SMI if the outliers outside the range are ignored. Moreover, it is noteworthy that the prediction accuracy of the DL model for SMI was similar to that obtained from the BV/TV based regression model (Table 6). This observation suggests that the prediction of SMI by the DL model is most likely depending solely on the inherent correlation of SMI with BV/TV.

This study shows that the resolution of simulated DXA images is directly related to the prediction accuracy of the DL models for assessment of all six histomorphometric parameters, exhibiting increased improvements in the prediction accuracy as the resolution of simulated DXA images is increased. However, no significant improvements would be expected after the image resolution is better than 0.3 mm/pixel and high prediction accuracies are still retained as long as the image resolution is less than 1.0 mm/pixel. This observation suggests that the sensitivity of DXA image based DL model may saturate after the image resolution exceeds a certain level. Since the resolution of clinical DXA images is usually around 1.0 mm/pixel and could be up to 0.3 mm/pixel (Boudousq et al., 2005), the results of this study suggest that the clinical DXA images may have the required image resolution for the DL models to capture the major microstructural features of trabecular bones. This finding is important because it ratifies, for the first time, the use of clinical DXA images in predicting the microstructural features of trabecular bones. Since clinical DXA can also provide accurate assessment of BMD, such understanding would facilitate development of advanced DL technologies for accurate prediction of osteoporotic bone fractures using DXA, the most affordable and convenient imaging modality for the purpose.

Moreover, this study also shows that the number of input simulated DXA images has a remarkable effect on DL model prediction of microstructural features of trabecular bone, especially when the simulated DXA image resolution is low (Fig. 4). From the computer graphics perspective, more images in different projections would provide more information on the spatial characteristics of trabecular structures, thus making it easier for the DL models to capture the features. However, its effect appears to diminish when the simulated DXA image resolution improves. This observation implies that the higher the simulated DXA image resolution, the more details of trabecular microstructure would be exposed, which would cancel out the information that could be provided to the DL models by increasing the number of input simulated DXA images.

There are several limitations of this study. Firstly, the sample population was limited in this study. We collected proximal femurs only from seven donors. Although different ages and genders are included, this small donor population is not necessarily representative of the entire human population. Nonetheless, the results of this study are still valid to prove the hypothesis that DXA image based DL models can assess the major histomorphometric properties of trabecular bone with high accuracies. Secondly, only six histomorphometric parameters (*i.e.*, BV/TV, Tb.Th, BS, Conn.D, SMI, and DA) were examined in this preliminary study, which are not inclusive of all histomorphometric parameters, such as Tb.Sp, Tb.N, and fractal dimensions. However, we believe that the selected parameters are representative of major microstructural features of trabecular bone, thus sufficient for testing the hypothesis of this study. Thirdly, the input data (simulated DXA images) used in this study for the DL models were not acquired from actual DXA measurements, but *via* computational simulations. In addition, only trabecular bone cubes were examined in this study, without taking into account the effect of cortical bone on DXA measurements. Thus, the simulated DXA images may not be fully representative of clinical cases.

5. Conclusion

This study is the first attempt ever to use DL techniques to predict the microstructural features of trabecular bone using simulated DXA images. The DL models demonstrated a high fidelity in predicting microstructural features of trabecular bone, thus verifying the hypothesis of this study. In addition, the results of this study indicate: (1) Even with the simulated DXA images of a resolution of 1.0 mm/pixel the DL models could predict the major microstructural features of trabecular bone with reasonable accuracies. (2) The number of input simulated DXA images has very limited effects on the prediction accuracy of the DL models if the resolution of the DXA images is better than 0.3 mm/pixel. These findings would help develop DXA image based DL techniques to assess both bone quantity and microstructural features, thus ensuring more accurate prognosis of bone fracture risks.

CRediT authorship contribution statement

Pengwei Xiao: Conceptualization, Methodology, Investigation, Software, Formal analysis, Writing - original draft, Visualization. **Tinghe Zhang:** Software, Formal analysis. **Xuanliang Neil Dong:** Methodology, Writing - review & editing. **Yan Han:** Writing - review & editing. **Yufei Huang:** Methodology, Formal analysis, Resources, Writing - review & editing. **Xiaodu Wang:** Conceptualization, Methodology, Formal analysis, Writing - review & editing, Project administration, Supervision.

Transparency document

The Transparency document associated with this article can be found, in online version.

Declaration of competing interest

All authors of this work claim no conflict of interests.

References

Consensus development conference: prophylaxis and treatment of osteoporosis. Am J Med 90 (1), 107–110.
Apostol, L., et al., 2006. Relevance of 2D radiographic texture analysis for the assessment of 3D bone micro-architecture. Med. Phys. 33 (9), 3546–3556.
Benhamou, C.L., et al., 2001. Fractal analysis of radiographic trabecular bone texture and bone mineral density: two complementary parameters related to osteoporotic fractures. J. Bone Miner. Res. 16 (4), 697–704.
Boudousq, V., et al., 2005. Image resolution and magnification using a cone beam densitometer: optimizing data acquisition for hip morphometric analysis. Osteoporos. Int. 16 (7), 813–822.
Boutroy, S., et al., 2005. In vivo assessment of trabecular bone microarchitecture by high-

resolution peripheral quantitative computed tomography. J. Clin. Endocrinol. Metab. 90 (12), 6508–6515.
Campion, J.M., Maricic, M.J., 2003. Osteoporosis in men. Am. Fam. Physician 67 (7), 1521–1526.
Chappard, D., et al., 2005. Texture analysis of X-ray radiographs is correlated with bone histomorphometry. J. Bone Miner. Metab. 23 (1), 24–29.
Chen, X.Y., et al., 2015. Automatic feature learning for glaucoma detection based on deep learning. In: Medical Image Computing and Computer-assisted Intervention, Pt III. vol. 9351. pp. 669–677.
Ding, Y., et al., 2019. A deep learning model to predict a diagnosis of Alzheimer disease by using (18)F-FDG PET of the brain. Radiology 290 (2), 456–464.
Dong, X., Wang, X., 2013. Assessment of bone fragility with clinical imaging modalities. Hard Tissue 2 (1), 7.
Dong, X.L.N., Shirvaikar, M., Wang, X.D., 2013. Biomechanical properties and micro-architecture parameters of trabecular bone are correlated with stochastic measures of 2D projection images. Bone 56 (2), 327–336.
Dong, X.N., et al., 2015. Stochastic predictors from the DXA scans of human lumbar vertebrae are correlated with the microarchitecture parameters of trabecular bone. J. Biomech. 48 (12), 2968–2975.
Esteva, A., et al., 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542 (7639), 115.
Greenwood, C., et al., 2015. The micro-architecture of human cancellous bone from fracture neck of femur patients in relation to the structural integrity and fracture toughness of the tissue. Bone Rep. 3, 67–75.
Hans, D., et al., 2011. Correlations between trabecular bone score, measured using anteroposterior dual-energy X-ray absorptiometry acquisition, and 3-dimensional parameters of bone microarchitecture: an experimental study on human cadaver vertebrae. J. Clin. Densitom. 14 (3), 302–312.
Harrigan, T., Mann, R., 1984. Characterization of microstructural anisotropy in orthotropic materials using a second rank tensor. J. Mater. Sci. 19 (3), 761–767.
Kabel, J., et al., 1999. Connectivity and the elastic properties of cancellous bone. Bone 24 (2), 115–120.
Kandi, H., Mishra, D., Gorthi, S.R.K.S., 2017. Exploring the learning capabilities of convolutional neural networks for robust image watermarking. Comp. Secur. 65, 247–268.
Kim, Y.J., Henkin, J., 2015. Micro-computed tomography assessment of human alveolar bone: bone density and three-dimensional micro-architecture. Clin. Implant. Dent. Relat. Res. 17 (2), 307–313.
Le Corroller, T., et al., 2012. Combination of texture analysis and bone mineral density improves the prediction of fracture load in human femurs. Osteoporos. Int. 23 (1), 163–169.
Litjens, G., et al., 2017. A survey on deep learning in medical image analysis. Med. Image Anal. 42, 60–88.
Muschitz, C., et al., 2015. TBS reflects trabecular microarchitecture in premenopausal women and men with idiopathic osteoporosis and low-traumatic fractures. Bone 79, 259–266.
Odgaard, A., 1997. Three-dimensional methods for quantification of cancellous bone architecture. Bone 21 (2), 191–199.
Odgaard, A., Gundersen, H.J.G., 1993. Quantification of connectivity in cancellous bone, with special emphasis on 3-D reconstructions. Bone 14 (2), 173–182.
Parkinson, I.H., Fazzalari, N.L., 2003. Interrelationships between structural parameters of cancellous bone reveal accelerated structural change at low bone volume. J. Bone Miner. Res. 18 (12), 2200–2205.
Patil, S., Ravi, B., 2005. Voxel-based representation, display and thickness analysis of intricate shapes. In: Ninth International Conference on Computer Aided Design and Computer Graphics, (Hong Kong, China).
Pothuaud, L., et al., 2000. Fractal dimension of trabecular bone projection texture is related to three-dimensional microarchitecture. J. Bone Miner. Res. 15 (4), 691–699.
Pothuaud, L., et al., 2007. Correlations between grey level variations on 2D DXA-images (TBS) and 3D microarchitecture in human cadaver bone samples. Bone 40 (6), S248.
Salmon, P.L., et al., 2015. Structure model index does not measure rods and plates in trabecular bone. Front. Endocrinol. 6.
Seeman, E., Delmas, P.D., 2006. Bone quality–the material and structural basis of bone strength and fragility. N. Engl. J. Med. 354 (21), 2250–2261.
Siris, E.S., et al., 2004. Bone mineral density thresholds for pharmacological intervention to prevent fractures. Arch. Intern. Med. 164 (10), 1108–1112.
Winzenrieth, R., Michelet, F., Hans, D., 2013. Three-dimensional (3D) microarchitecture correlations with 2D projection image gray-level variations assessed by trabecular bone score using high-resolution computed tomographic acquisitions: effects of resolution and noise. J. Clin. Densitom. 16 (3), 287–296.
Wright, N.C., et al., 2014. The recent prevalence of osteoporosis and low bone mass in the United States based on bone mineral density at the femoral neck or lumbar spine. J. Bone Miner. Res. 29 (11), 2520–2526.
Wright, N., et al., 2019. The burden of osteoporosis in the United States - a US Bone and Joint Initiative report. J. Bone Miner. Res. 34, 28.