



OPEN ACCESS

EDITED BY

Ridha Hambli,
Polytech Orléans, France

REVIEWED BY

Zhan Liu,
Sichuan University, China
Qiang Chen,
Southeast University, China

*CORRESPONDENCE

Songtao Ai,

✉ ai.songtao@qq.com
Tsung-Yuan Tsai,
✉ tytsai@sjtu.edu.cn

[†]These authors have contributed equally to this work

SPECIALTY SECTION

This article was submitted to Biomechanics, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 27 September 2022

ACCEPTED 20 March 2023

PUBLISHED 19 May 2023

CITATION

Cui J, Liu CL, Jennane R, Ai S, Dai K and Tsai T-Y (2023), A highly generalized classifier for osteoporosis radiography based on multiscale fractal, lacunarity, and entropy distributions.
Front. Bioeng. Biotechnol. 11:1054991.
doi: 10.3389/fbioe.2023.1054991

COPYRIGHT

© 2023 Cui, Liu, Jennane, Ai, Dai and Tsai. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A highly generalized classifier for osteoporosis radiography based on multiscale fractal, lacunarity, and entropy distributions

Jingnan Cui^{1†}, Cheng Lei Liu^{2†}, Rachid Jennane³, Songtao Ai^{2*}, Kerong Dai^{1,4} and Tsung-Yuan Tsai^{1*}

¹School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China, ²Department of Radiology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, ³IDP Institute, UMR CNRS 7013, University of Orléans, Orléans, France, ⁴Department of Orthopaedic Surgery, Shanghai Ninth People's Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China

Background: Osteoporosis is a common degenerative disease with high incidence among aging populations. However, in regular radiographic diagnostics, asymptomatic osteoporosis is often overlooked and does not include tests for bone mineral density or bone trabecular condition. Therefore, we proposed a highly generalized classifier for osteoporosis radiography based on the multiscale fractal, lacunarity, and entropy distributions.

Methods: We collected a total of 104 radiographs (92 for training and 12 for testing) of lumbar spine L4 and divided them into three groups (normal, osteopenia, and osteoporosis). In parallel, 174 radiographs (116 for training and 58 for testing) of calcaneus from health and osteoporotic fracture groups were collected. The texture feature data of all the radiographs were pulled out and analyzed. The Davies–Bouldin index was applied to optimize hyperparameters of feature counting. Neighborhood component analysis was performed to reduce feature dimension and increase generalization. A support vector machine classifier was trained with only the most effective six features for each binary classification scenario. The accuracy and sensitivity performance were estimated by calculating the area under the curve.

Results: Interpretable feature trends of osteoporotic pathological changes were depicted. On the spine test dataset, the accuracy and sensitivity of binary classifiers were 0.851 (95% CI: 0.730–0.922), 0.813 (95% CI: 0.718–0.878), and 0.936 (95% CI: 0.826–1) for osteoporosis diagnosis; 0.721 (95% CI: 0.578–0.824), 0.675 (95% CI: 0.563–0.772), and 0.774 (95% CI: 0.635–0.878) for osteopenia diagnosis; and 0.935 (95% CI: 0.830–0.968), 0.928 (95% CI: 0.863–0.963), and 0.910 (95% CI: 0.746–1) for osteoporosis diagnosis from osteopenia. On the calcaneus test dataset, they were 0.767 (95% CI: 0.629–0.879), 0.672 (95% CI: 0.545–0.793), and 0.790 (95% CI: 0.621–0.923) for osteoporosis diagnosis.

Conclusion: This method showed the capacity of resisting disturbance on lateral spine radiographs and high generalization on the calcaneus dataset. Pixel-wise texture features not only helped to understand osteoporosis on radiographs better but also shed new light on computer-aided osteopenia and osteoporosis diagnosis.

KEYWORDS

osteoporosis, trabecular degeneration, fractal dimension, entropy, lacunarity, neighborhood component analysis, image dehazing

1 Introduction

Nearly 50% of women and 20% of men suffer from osteoporotic fractures in their lifetime (Office of the Surgeon General (US), 2004; Yamamoto et al., 2020). Asymptomatic osteoporosis often remains overlooked or misdiagnosed until a low-trauma trunk fracture occurs, which leads to hospitalization (Tu et al., 2018). A total of 8.9 million fractures inflicted by osteoporosis occur worldwide each year, causing a significant burden on social health and the economy (Johnell and Kanis, 2006). The widely accepted method of screening for osteoporosis, which involves assessing bone mineral density (BMD) through dual-energy X-ray absorptiometry (DXA), is unable to reach the majority of the high-risk population. This is not only due to DXA's low availability and cost but also the concern of the X-ray dosage and the low screening rate of <10% in the US (Amarnath et al., 2015). While in China, only 4.3% of women aged ≥ 50 years have undergone testing, even less women in rural areas are subjected to test, accounting for only 1.9% (National Health Commission of China, Epidemiological investigation of osteoporosis in China, 2018). In addition, acute respiratory distress syndrome such as COVID-19 and glucocorticoid therapy could lead to osteoporosis directly and indirectly (Tang, 2022). Overall, there is a need for an easy-to-use screening tool for osteoporosis.

With the advance in machine learning (ML) (Hassouni et al., 2017), especially deep learning (DL) (Areeckal et al., 2018a; Teclé et al., 2020; Yamamoto et al., 2020; Zhang et al., 2020), radiographic images can be used to diagnose osteoporosis, whereas the multilayer architecture of DL models result in several shortcomings. First, the interpretability of a model is expected. Whether we can replace physician intelligence with a DL model trained from a labeled dataset largely depends on its interpretability for clinicians (Smets et al., 2021). Nevertheless, the state-of-the-art DL studies lack solid clinical implications (Smets et al., 2021). Second, the reliability of multilayer DL models relies on the quantity and the quality of the whole labeled dataset. For complex DL model training, thousands of well-defined images are commonly required to avoid overfitting and inconsistent prediction (Scanlan et al., 2018; Xiao et al., 2020). In practice, diagnosing and disclosing the essential difference between health and osteoporotic bone trabecular structure on a limited sample size is a big challenge.

Indeed, diagnostic reasoning is derived from the first principle (Reiter, 1987). According to pathophysiology, the osteoporotic fracture risk is aroused by the loss of bone tissue (Cosman et al., 2014; Smets et al., 2021). From the microarchitectural point of view, individual 3D trabecular plates of bone are lost as a microscopic representation of bone loss, leaving an architecturally weakened structure with significantly reduced mass (Smets et al., 2021). Referring to histomorphometry, the 3D trabecular bone structure changes correlate with the 2D fractal descriptor of their projections (Scanlan et al., 2018). Further research has demonstrated that the self-similarity parameter of the 3D fractal character holds an influencing factor of 0.5 on its 2D projection. In prior studies (Jennane et al., 2001; Yger, 2014; Palanivel et al., 2020), textual

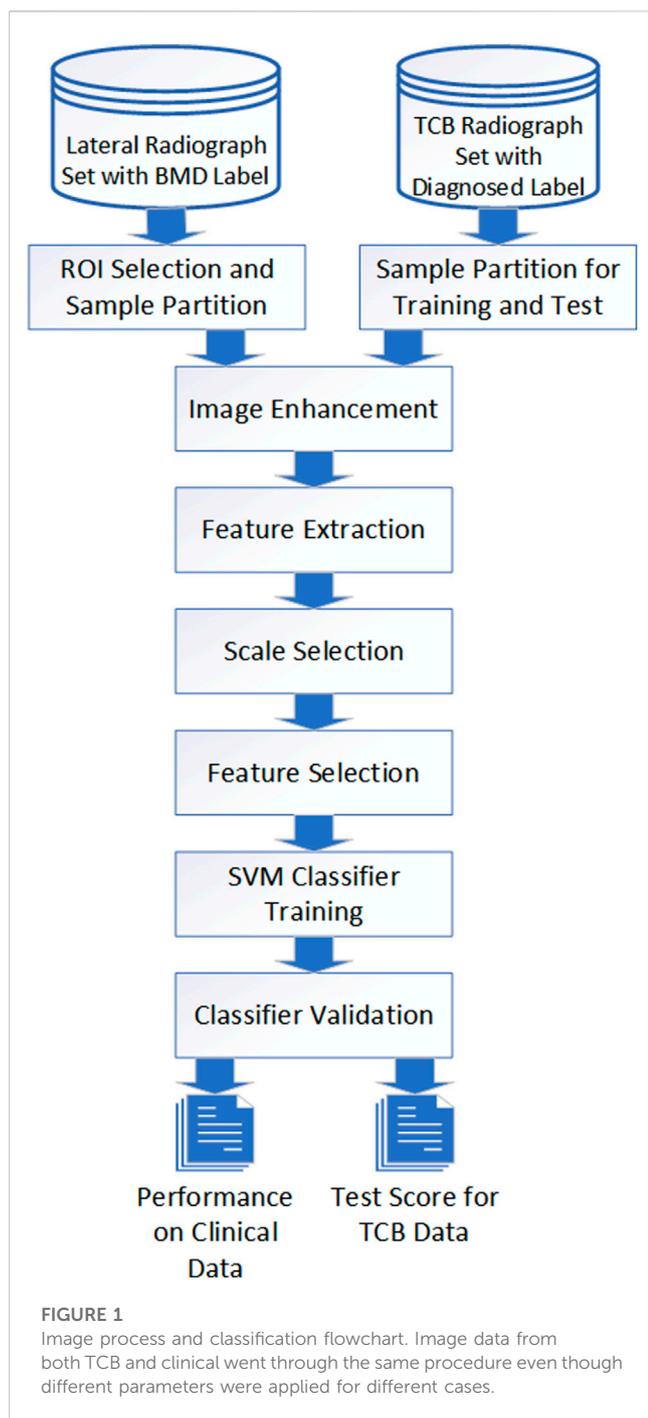
features such as fractal or lacunarity inferred from histomorphometry were successfully accounted for osteoporosis. However, most of them focused on homogeneous features such as fractal dimension measured by fractal Brownian motion (fBm) (Harrar et al., 2013; Hassouni et al., 2017), which could not describe a heterogeneous texture pattern. Studies on microarchitecture have shown the episodic nature of osteoporotic lesions (Greenwood et al., 2018; Hussain and Han, 2019). The subtle trabecular architectural changes induce a negligible difference in spatial dispersion patterns, especially in the osteopenia stage (Jennane et al., 2001). Therefore, an interpretable and pixel-level detection method can be a reliable tool to identify osteoporotic characterization on radiographs at an early stage, providing a better understanding of the pathological process of osteoporosis *via* texture features. Thus, to the best of our knowledge, no work has been devoted to pairwise feature comparison and classification across normal, osteopenia, and osteoporosis on spine X-ray images.

In this study, we analyzed clinical lateral vertebral lumbar spine radiographs. The whole soft tissue around the tummy, intestinal gas, and even the patient's waist belt are all superimposed in spine radiographs. In parallel, calcaneus radiographs from the textural characterization of a bone (TCB) challenge dataset were also investigated since these images have a clean bone texture with less routine clinical exam coverage. To distinguish osteoporotic radiographs from healthy ones in two different lesion sites, we customized the image enhancement process for those special artifacts and synthesized multiscale pixel-wise fractal distribution, multiscale pixel-wise entropy distribution, and global lacunarity. The six most influential features out of 11 were selected to train support vector machine (SVM) classifiers to achieve high generalization, which could differentiate normal, osteopenia, and osteoporosis cases. In the process of classifier construction, we aimed to

- i) Interpret pixel-wise feature trends to better understand osteoporosis's pathogenesis and train a feature-based SVM classifier for each osteoporotic binary classification scenario
- ii) Evaluate the classification performance on a triple-class problem, normal, osteopenia, and osteoporosis, for the clinical dataset, and a binary class problem for the TCB dataset
- iii) Compare the proposed classifier to the other two kinds of models, the other feature-based classifiers in the TCB ranking, and the latest DL-based classifiers

2 Methodology

The description of the entire operating procedure from X-rays to diagnosis can be represented by flowchart modules (Figure 1). For different image sources, identical manipulation steps were taken. Those steps included image preparation, enhancement, feature extraction, scale selection, feature selection, SVM training, and validation. However, due to different intrinsic artifacts introduced



in image acquisition, differential processing parameters and hyperparameters were applied.

2.1 Data acquisition and partition

Two different kinds of radiograph data were collected. One was the lateral radiograph of the lumbar spine from clinical dataset (IRB code: SH9H-2021-T94-2.) of Shanghai Ninth People's Hospital. Informed consent from patients was exempted due to the retrospective nature of this study. The inclusion criteria were 1)

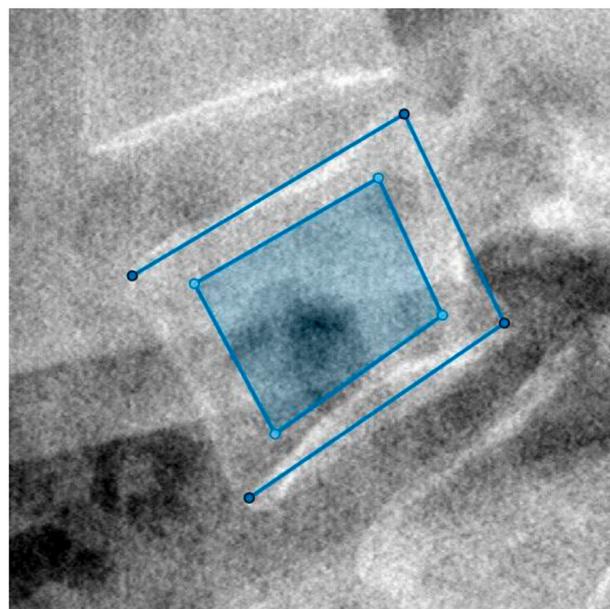


FIGURE 2
The four vertexes of the outer open quadrilateral selected manually were the four endpoints of endplates' projection. Then, the dark region cut from the central 2/3 part of the two diagonal of the outer quadrilateral was computed as ROI.

osteopenia and osteoporosis in women aged ≥ 50 years. They were identified as postmenopausal by radiologists; 2) all subjects had undergone lateral lumbar spine radiography and DXA examinations within 1 month without any other bone treatments in between. According to BMD results (Cosman et al., 2014), all subjects' radiographs were categorized into three groups which were the normal group ($n = 44$) with $-1 < T\text{-score}$, the osteopenia group ($n = 28$) with $-2.4 \leq T\text{-score} \leq -1$, and the osteoporosis group ($n = 32$) with $T\text{-score} \leq -2.5$.

The other was a lateral calcaneus radiograph from the TCB challenge dataset (Overview, 2022). It was a region of interest (ROI) image set in the trabecular bone of the calcaneus for the evaluation of osteoporosis diagnostic methods. The data used in this study included those of 87 healthy and 87 osteoporotic (with fracture) subjects.

For clinical data, we semi-manually labeled the ROI for the following texture analysis. We utilized MathWorks MATLAB 9.11.0.1769968 to program our labeling tool. Operation steps were as follows: 1) We selected the approximate center of L4 in the original DICOM image and took the point of central focus chosen as the center to crop a sub-view (512*512) of the whole image. 2) In the new cropped lateral view of L4, we manually selected the four corners of vertebral laminae to define the L4 posture. 3) We intercepted the central 2/3 part of the two diagonal. The four breakpoints were taken as the ROI vertexes. The selection and the final ROI of an L4 lateral radiograph are shown in Figure 2.

For the TCB data, 29 subjects were randomly selected from both the control and osteoporosis group. A total of 58 subjects were combined and subjected to the blind test set, in which the health to osteoporosis ratio was 1: 1. For TCB challenges, the ratio of the

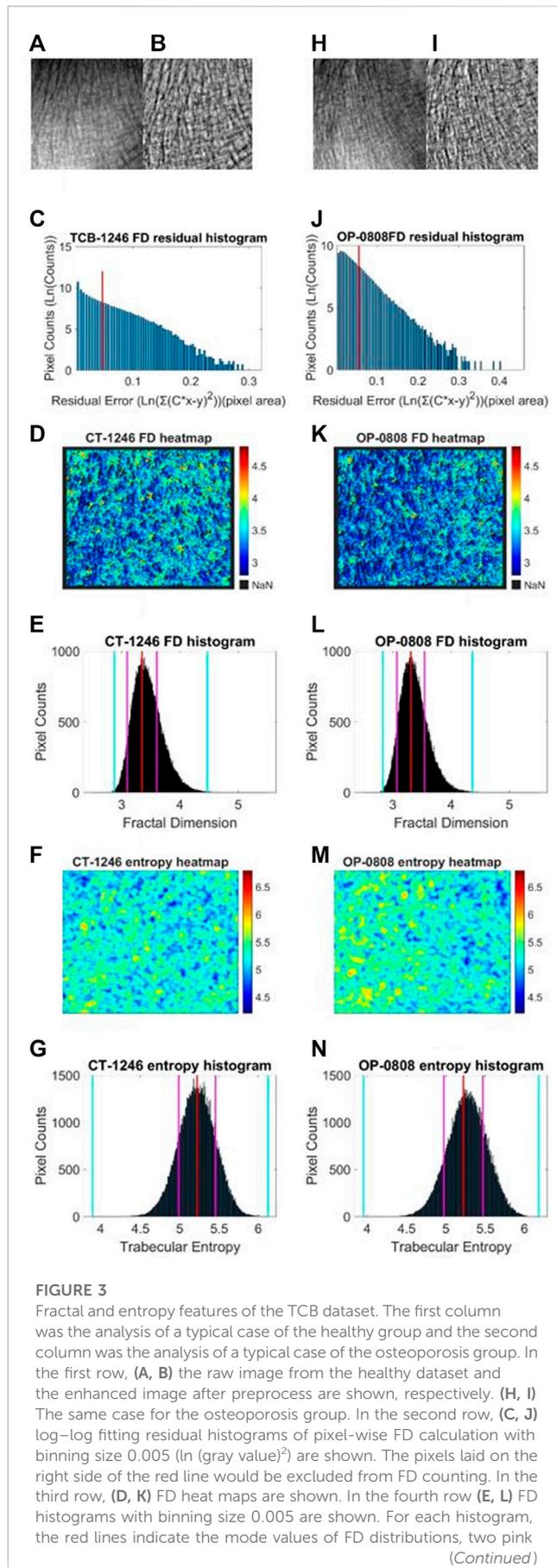


FIGURE 3 (Continued)

lines indicate the \pm standard deviation, and two cyan lines indicate the superior and inferior 0.1% quantiles. In the fifth row (F, M) entropy heat maps are shown. In the sixth row (G, N) entropy histograms with binning size 0.001 are shown; the red lines indicate the mode values of FD distributions, two pink lines indicate the \pm standard deviation, and two cyan lines indicate the superior and inferior limits of the whole distribution.

training set to test set was 2: 1. The mapping between image ID and subject group in the blind test set was saved for classification result assessment.

2.2 Image enhancement

We performed a similar image enhancement operation for both TCB and clinical datasets using different parameters. Due to bone thickness variation in an image or heterogeneous soft tissue projection, X-ray exposure imbalance occurred in a single image of trabecular bone. First, we flattened the ROI with the MATLAB function *imflatfield* with radii 20 for the TCB set and radii 10 for the clinical set. Second, considering the X-ray scatter caused by soft tissue, we used the dehazing method *imreducehaze* to suppress this blurry artifact for bone imaging. The intensity was 0.05 for the TCB set and 0.1 for the clinical set. Third, we adjusted the gray value window width and window level for both TCB and clinical set to make the highest 0.1% pixels saturate and the lowest 0.1% cut-off in each image. Those original images were named as raw images (Figures 3A, H and Figures 4A, H, O), while those enhanced images were named as clean images (Figures 3B, I and 4B, I, P).

Because of clinical images, different projection conditions were applied to different patients, i.e., each radiograph had its respective source-to-image distance (SID) and source-to-object distance (SOD). We unified the image scale diversity. There were two brands of digital radiography (DR), GE Healthcare, and Carestream Healthcare, from which these clinical images were taken. Those DICOM images from the GE Healthcare had both SID and SOD, whereas images from the Carestream Healthcare only had SID. We draw those scatter points from SID to SOD of all GE Healthcare images and linearly interpolate missing values between them. We applied interpolation values for those images with no SOD parameters (Figure 5).

2.3 Feature extraction

2.3.1 Pixel-wise fractal dimension (FD)

As Peleg et al.'s (1984) extension on Mandelbrot's (1994) definition of fractal dimension, when a gray-level image was viewed as a hilly terrain surface, whose height from the ground level is proportional to the image gray value, the area of the fractal surface is displayed according to the power function of measure scale or rather scaled pixel size. Then, the hilly terrain surface was characterized by textural fractal dimension (Creutzberg and Ivanov, 1989). The transformed relationship is shown as follows (Creutzberg and Ivanov, 1989; Enkins et al., 2018):

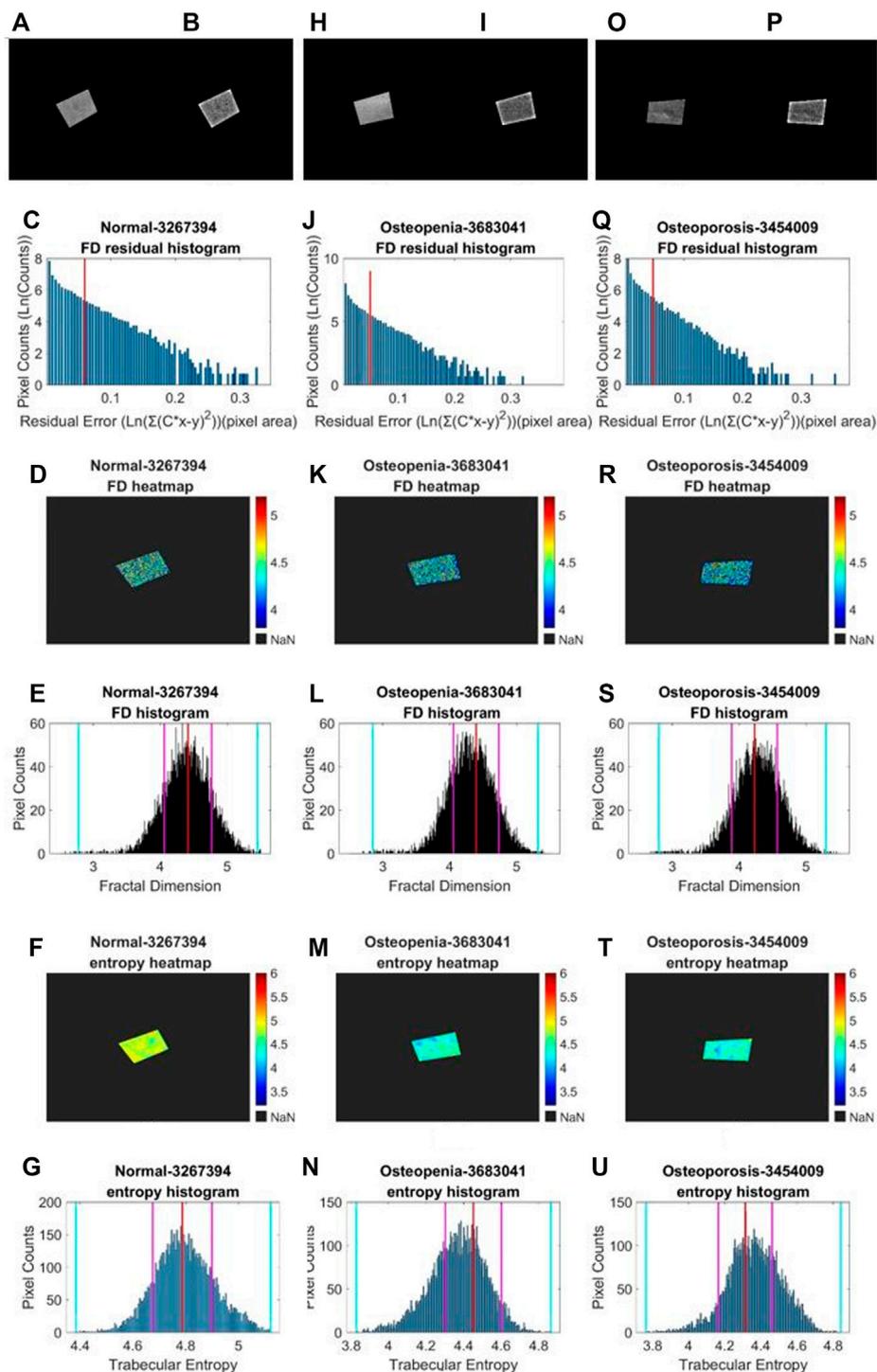


FIGURE 4

Fractal and entropy features of clinical L4 lateral radiograph dataset. Three columns from (A, B–G), from (H, I–N), and from (O, P–U) depict typical cases in normal, osteopenia, and osteoporosis individually. In the first row, (A, B), (H, I), (O, P) a comparison between the raw image and clean image from each group is shown. In the second row, (C, J, Q) log–log fitting residual histograms with binning size 0.005 $[\ln(\text{gray value})]^2$ are shown. In the third row, (D, K, R) FD heat maps are shown. In the fourth row, (E, L, S) FD histograms with binning sizes of 0.005 are shown. For each histogram, the red lines indicate the mode values of FD distributions, two pink lines indicate the \pm standard deviation, and two cyan lines indicate the superior and inferior 0.1% quantiles. In the fifth row, (F, M, T) entropy heat maps are shown. In the sixth row (G, N, U) entropy histograms with binning size 0.005 are shown; the red lines indicate the mode values of FD distributions, two pink lines indicate the \pm standard deviation, and two cyan lines indicate the superior and inferior limits of the whole distribution.

$$FD = 2 - \frac{\Delta \ln(SA)}{\Delta \ln(R)}. \quad (1)$$

In Eq. 1, FD means the fractal dimension. SA means the surface area of the image gray value, and it was the mass we wanted to measure in various scales. R stands for the rescaled pixel size. The mechanism of a typical radiographic image mainly is about X-ray attenuation projection on a flat panel detector. All anatomies involving bone and soft tissues that X-ray passed through would be superimposed into the image. The superimposed soft tissue visibly affected the bone texture gray value. So, we employed the blanket method, which was robust against gray-level changes (Lopes and Betrouni, 2009). To further acquire pixel-wise local FD distribution in a single image, we adapted the modified blanket method (MBM) (Enkins et al., 2018) as our final implementation.

We used a mathematical model to disclose the relationship between the surface area and rescaled pixel size rather than changing the whole image resolution. Former studies have shown that we could take pixels as columns and their side area could be equivalent to the image surface area (Peleg et al., 1984; Enkins et al., 2018). Horizontal and vertical gradient were calculated with convolution kernel $[1 \ -1, 0]$ and $[0, -1, 1]^T$, respectively. When we got those two perpendicular direction gradients, it was easy to express the surface area as follows:

$$SA(p, r) = \sum_{\|p-q\|_2 < r} |G_x(q)| + \sum_{\|p-q\|_2 < r} |G_y(q)|. \quad (2)$$

In this equation, $SA(p, r)$ represents the 3D topological surface area of an image at current scale. The parameter p represents the p -th pixel. The parameter r means the radii of a measuring disk in which a local FD value was counted. The physical size of a measuring disk was constant for all rescaled pixel sizes. So, when the rescaled pixel size turned to be bigger, the corresponding r expressed with rescaled pixel would be smaller. $|G_x(q)|$ and $|G_y(q)|$ represent the absolute values of the horizontal and vertical gradient at pixel q . Here, q represents any pixel whose distance to the pixel p is less than r . As described in the original MBM, the whole image would resemble at a different resolution to rescale the pixel size. For example, all pixels in the same binning patch would have the same pixel value when physical pixel-wise FD was counted. Namely, in Figure 6, all physical pixels inside the pink double-line square region had the same resample value. Nevertheless, in this study, we have made several improvements to the original MBM procedure to obtain a more reliable FD.

First, local surface areas were measured in different binning patches which were centralized at each pixel. When we calculated the local FD of the p -th pixel, the surface of a disc-shaped neighborhood of the p -th pixel, shown in Figure 6, would be measured on different scales. After the physical radii of the neighborhood disk were predefined, a serial of continuously changing rescaled disc radii meant a serial of corresponding pixel binning was taken for a constant physical neighborhood size. In Figure 7, the disk radii changed from 0 to 4. Here, all different sizes of pixel binning centralized at the p -th pixel were taken in the side neighborhood disk. In Figure 6, different color cells distributing around the p -th pixel demonstrated how to centralize at the p -th pixel. The gray cell stood for the physical pixels. The blue cells were rescaled pixels with the binning size 3.

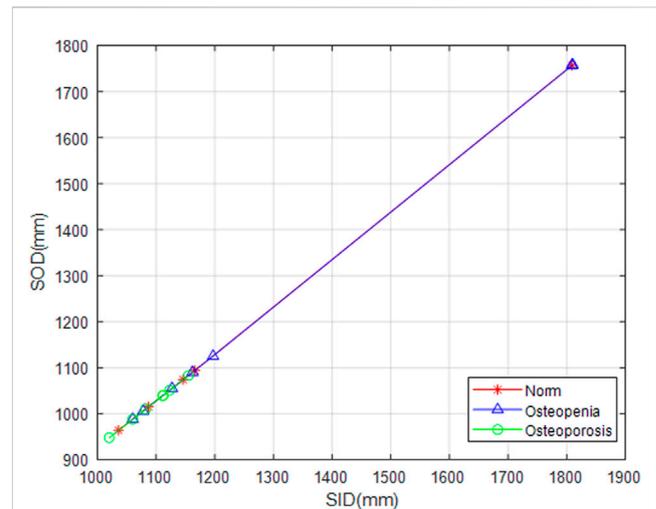


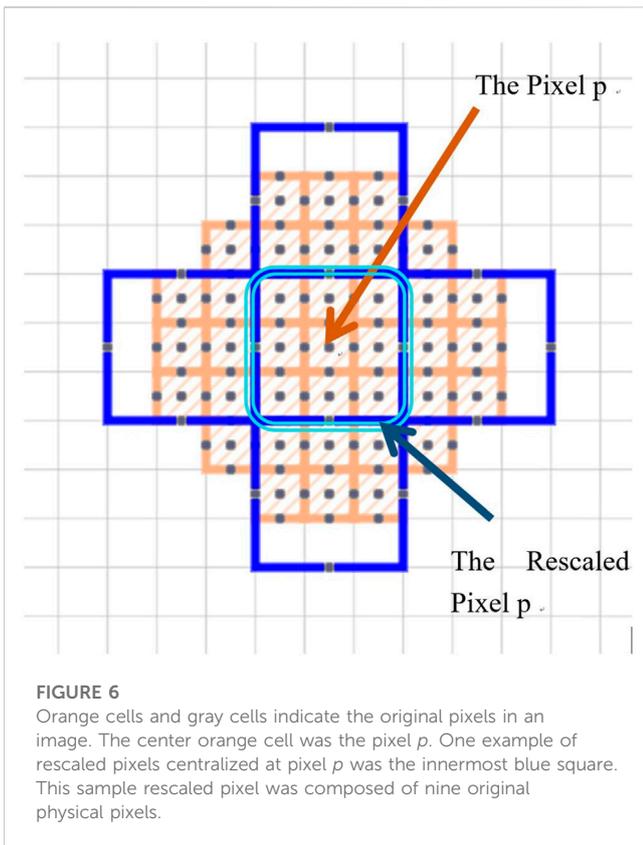
FIGURE 5

SOD value fitting. Snowflakes, triangles, and circles marked SID against SOD from normal, osteopenia, and osteoporosis groups, respectively. It is a ground breaking uniformization of X-ray projection scale transformation for scale-related feature counting.

Blue cells and orange cells indicated the disc neighborhood of the pixel p on a different scale, and they had approximately equal measure regions. So, instead of the original MBM using the whole image multi-resolution, here, different rescaled pixel sizes were all centralized at each pixel and generated pixel-wise FD precisely located at each pixel.

Second, in an image, sporadic pixels which did not characterize fractal significantly would be excluded for FD counting. Due to the influence of imaging noise and superimposed artifacts of radiographs, sporadic pixels' neighborhoods did not characterize fractal features so well that a few pixels had a relatively high residual value in the residual histograms in Figures 3C, J for the TCB dataset and Figures 4C, J, Q for the clinical dataset. To figure out those pixels, all residuals from the straight-line fitting of the surface area against scaling factors were cached and sorted. The residual distribution could be checked in a log-log histogram with binning size $0.005 [\ln(\text{gray value})^2]$ in Figures 3C, J, 4C, J, Q. The absolute majority of fitting residuals was close to zero, so we remained with those pixel FDs with 80% of the smallest residuals for further statistics and abolished the other pixels. In Figures 3C, J, 4C, J, Q, those FD values corresponding to the left side of the red line remained, and values on the right side were abolished. We could draw the FD heat maps from these reserved pixel-wise FDs, Figures 3D, K, 4D, K, R, and histogram, Figures 3E, L, 4E, L, S, with binning size 0.005.

Third, more FD distribution features were detected by introducing the FD histogram of the whole image. In Figures 3E, L, 4E, L, S, it is easy to find that the FD distribution shape was basically a single peak but asymmetric in both the control and osteoporosis groups, so we calculated the mode value instead of the mean. Regarding the hackly distribution around the peak, we applied a sliding window (window width = 0.011, moving step = 0.001) to calculate an average value on the histogram and took the



maximum as the mode value. In addition, FD standard deviation, FD skewness, FD kurtosis, and the Shannon entropy of FD rescaled to 16-bit unsigned integer were all counted as features.

2.3.2 Differential box-counting (DBC) lacunarity

Although some studies suggested that FD and lacunarity may be correlated in some cases (Smith et al., 1996), lacunarity is by no means redundant. Some patterns are more distinguishable by their lacunarity (Karperien et al., 2011). Especially in our study, the image texture is neither an ideal fractal topological structure nor a disclosed graph, so the lacunarity feature deserves to be tried. The lacunarity of a set implicates the distribution character of mass. A well-accepted approach to calculating lacunarity is as follows:

$$\Lambda(r) = \frac{\sum_M M^2 Q(M, r)}{[\sum_M M Q(M, r)]^2} = \frac{\sigma^2(r)}{\mu^2(r)} + 1, \quad (3)$$

$$\text{where } Q(M, r) = \frac{n(M, r)}{N}.$$

In Eq. 3, $\Lambda(r)$ is the lacunarity, a function of measure unit r . On the right of the first equal sign, M means the mass of a square area $r \times r$. The probability $Q(M, r)$ means the number of boxes $n(M, r)$ with size r and mass M dividing the total number of boxes N in the whole image described in the equation. The mean-square of M divided by its square mean was the lacunarity. The simplified expression is on the right.

In this study, we took the differential box-counting (DBC) algorithm to calculate M at scale unit r (Sarkar and Chaudhuri, 1994; Conci and Campos, 1996; Li et al., 2009). Yet, considering the

influence of the initial position of the array of counting boxes, we took an average of all possible initial positions. The average lacunarity, $\bar{\Lambda}$, is calculated in the following equation. γ implicates the possible initial position, and G equals the number of all possible initial positions, and Λ_γ is the whole image lacunarity Λ at the initial position γ .

$$\bar{\Lambda} = \frac{\sum_{\gamma=1}^G \Lambda_\gamma}{G}. \quad (4)$$

2.3.3 Pixel-wise entropy

Shannon entropy quantifies the amount of uncertainty involved in the value of a random variable or the outcome of a random process (Shannon, 2001). Numerous studies validated that focus on measuring the complexity of a chaotic system using information entropy were reliable (Costa and Goldberger, 2015; Humeau-Heurtier, 2016). Nevertheless, most of them were limited in studying unidimensional datasets. Here, we attempted to extend the entropy calculation to conduct a 2D osteoporotic radiography texture analysis.

$$v_p = \max(I_q) - \min(I_q), \quad (5)$$

where $q \in \{q \mid \|p - q\|_2 \leq r\}$,

$$H_l = -\sum_{i=1}^n P(v_i) \log_2 P(v_i). \quad (6)$$

To get robust random value observations, we reused the DBC method in computing lacunarity to count differential value instead of absolute value. First, we set a measuring scale representing a small neighborhood of the p -th pixel in the image ROI with radii r . The difference between the maximum and minimum gray level inside the pixel's small neighborhood was taken as an observation of the local random measure at that pixel. In Eq. 5, v_p is the measure observation at the p -th pixel and q is a pixel inside the small neighborhood of the p -th pixel. Second, we set a big neighborhood, a disc-shaped statistic field, as the local entropy evaluation range, in which observations at every pixel's small neighborhood were counted to calculate Shannon entropy. When we took the disc-shape statistic field with n pixels inside and moved it as a sliding window to cover the whole image. Then, the local entropy H_l was calculated by Eq. 6, in which n is the total number of pixels in the big neighborhood of the p -th pixel and $P(v_i)$ is the probability of the observation v_i . In Figures 3F, M, and Figures 4F, M, T the pixel-wise entropy heatmap of the whole image were shown. Furthermore, for each of them, a histogram can be generated in Figures 3G, N and Figures 4G, N, U demonstrating osteoporotic pathological progress.

After the statistic slide window moved over the whole image, an entropy heat map and its histogram (binning size 0.005) were achieved. The entropy histogram extracted four related features involving mode value, standard deviation, skewness, and kurtosis.

2.4 Scale parameter selection

Any classification problem can be solved as a pairwise binary classification problem. In the following section, we only discuss the

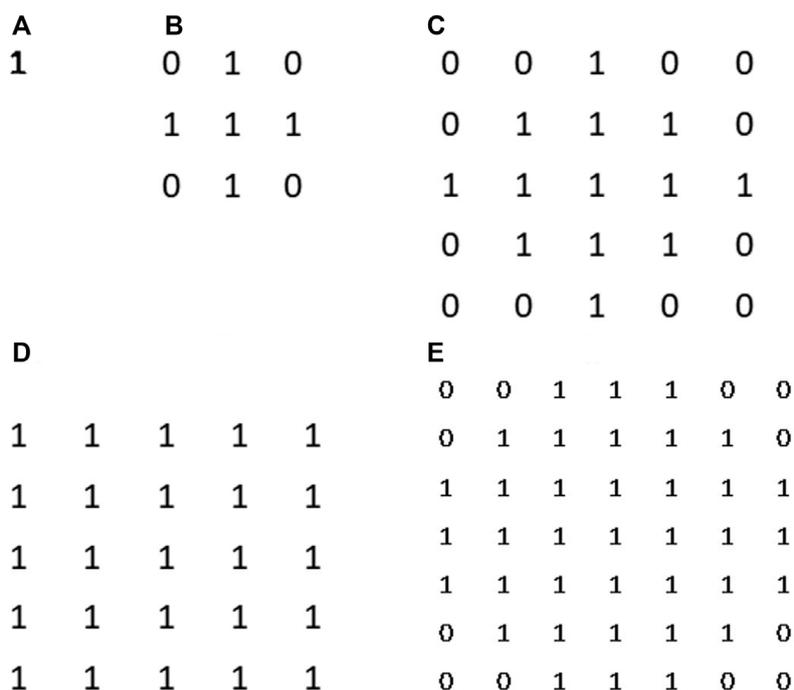


FIGURE 7
Surface area counting disks. From (A–D) they are disk kernel with corresponding radii from 0 to 4. Those disk kernels can help us to restrict a specific neighbourhood around counting pixel for measurement.

binary classification performance, in both the TCB challenge case and our clinical bone fragility grading case.

Concerning the scale factor’s impact on the outcome of classification, an optimized scale had to be selected rationally. For FD calculation, we had to decide the finest measure resolution (the smallest scaled pixel size) and the maximum measure range (the biggest scaled pixel size) pixel scale to fit a straight line from the scaled pixel to the surface area. The optimized FD fitting scale was expected to best divide the normal group and the other observed groups. The diversity scale size r in lacunarity calculation, together with the various measure scales and counting scale in entropy statistic, would all lead to a different performance in the classification job.

We employed the Davies–Bouldin index (DBI) (Davies and Bouldin, 1979), a cluster distance index, to evaluate scale factors’ performance in class separations. DBI estimated the class cohesion based on the distance from the feature points in a class to its centroid and the separation based on the distance between centroids. The ratio of cohesion to separation was the DBI value (Arbelaitz et al., 2013). The smaller the DBI, the better the class was separated. The definition formula of the DBI is as follows:

$$DBI(C) = \frac{1}{K} \sum_{c_k \in C} \max_{c_l \in C, c_l \neq c_k} \left\{ \frac{S(c_k) + S(c_l)}{\| \bar{c}_k - \bar{c}_l \|} \right\}, \tag{7}$$

where $S(c_k) = \frac{1}{|c_k|} \sum_{x_i \in c_k} \| x_i - \bar{c}_k \|$.

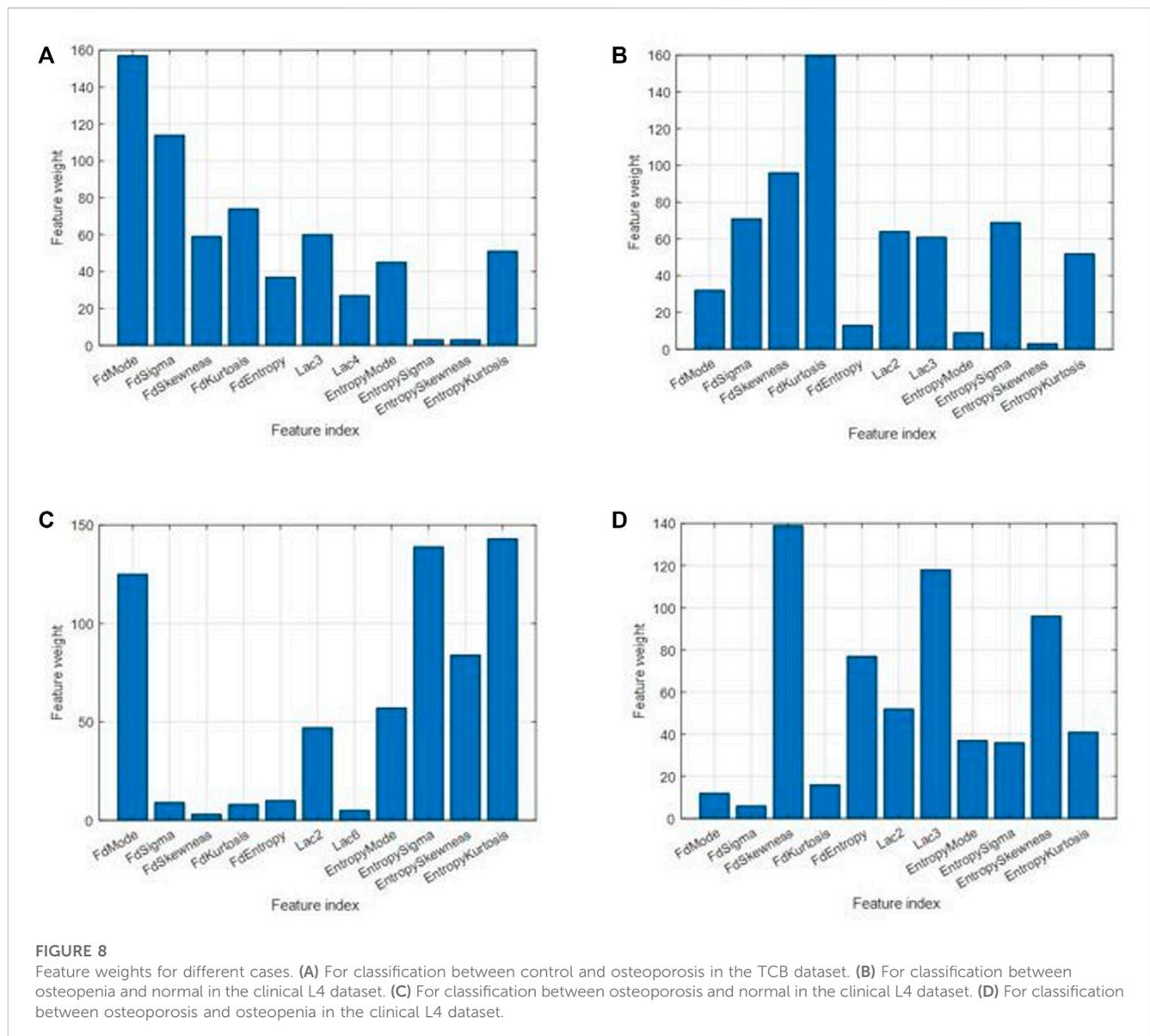
The parameter C of DBI was the full set of classes to be separated. K is the number of classes in the full set. For this

study $K = 2$. $S(c_k)$ is the average distance from each scatter point to the centroid of K and x_i is an n -dimensional feature vector assigned to any point i in class k . \bar{c}_k is the centroid of K . $\| \cdot \|$ is the euclidean distance.

Via the DBI, we could iterate and clarify which scale size outperformed in this classification job. For two-class classification, we iterated some empirical value scale factors and figured out one or two locally optimized scale sizes for each of those three feature categories. So far, we had 11 features, which are FD mode, FD sigma, FD skewness, FD kurtosis, FD entropy, lacunarity A, lacunarity B, entropy mode, entropy sigma, entropy skewness, and entropy kurtosis with definitive scale size. Lacunarity A and B were the best two distinguishable scale sizes.

2.5 Feature selection

We selected six feature dimensions from 11. Feature selection can not only mean the most effective features to classify bone fragility grade but also overcome overfitting and level up generalization performance. Via trialing on the TCB dataset, we found six feature dimensions that could provide comparable performance on both the blind test set and the training test. The best six features were selected via the popular neighborhood component analysis (NCA) method (Yang et al., 2012). NCA, a non-parametric method, differs from principle component analysis (PCA) and linear discriminant analysis (LDA). PCA was influenced by the scaling of the variables. LDA works only if all class



distributions are Gaussian with one single shared covariance (Raghu and Sriram, 2018). In NCA, the goal was to maximize the prediction accuracy of classification algorithms. On the other hand, in learning feature weights, a regularization parameter λ was introduced to learn feature weights to minimize an objective function that measured the average leave-one-out classification loss over the training data to expect a better generalization performance.

In NCA, 10 steps were taken. i) We partitioned the labeled sample set with 11 dimension features into five folders. In each fold, 4/5 of data were assigned to a training set, and the other 1/5 of data were assigned to a test set; ii) generated a regularization parameter vector, λ , using the MATLAB command λ values = *linspace* (0, 2, 20)/n where n is the number of samples of the whole training set; iii) trained the NCA model with the MATLAB function *fscnca* for each λ using the training set; iv) computed the classification loss for the corresponding test set in the fold using the NCA model and recorded the loss value; v) repeated (iii) for all folds and all λ values; vi) found the λ value that corresponds to the minimum average loss; vii) fitted

the NCA model with *fscnca* to all of the data using the best λ value. Meanwhile, we got the fitting loss and feature weights of this round; viii) repeated from (i) to (vii) 60 times and recorded all those fitting losses and corresponding feature weights; ix) sorted the fitting losses and retained the best-fitting ones and their corresponding feature weight rankings. The other half of them was abolished; and x) took each weight ranking as a ballot. Each feature weight ranking had its ballot score [0, 0, 0, 0, 0, 1, 2, 3, 4, 5, and 6] from the least to the most significant and summed up the remaining 30 time ballot scores for each feature. We got the overall ranking for all features. Feature weight bar charts are shown in Figure 8.

2.6 SVM training

On the training sets for TCB and clinical L4, the six most significant labeled features of each pair of classes were input into the SVM classification function *fitsvm*. Configure parameters were

TABLE 1 Scale size assessment *via* DBI for TCB dataset classification. The red and yellow numbers also in bold stand for the minimum and subminimum DBI of each kind of feature scale parameter list. The corresponding scale parameters of the red and yellow DBI are the optimized ones.

Scale size assessment <i>via</i> DBI for TCB classification								
Radii range	1–3	1–4	1–5	1–6	1–7			
FD	4.664	3.26	3.589	5.361	5.245			
Measure radii, count radii	1,4	1,5	1,6	1,7	1,8	2, 5	2, 6	2, 7
Entropy	7.741	6.535	6.428	7.035	6.960	10.168	12.968	17.269
Size	2*2	3*3	4*4	5*5	6*6	7*7		
Lacunarity	4.045	2.890	2.939	3.151	3.478	3.877		

TABLE 2 Scale size assessment *via* DBI for L4 dataset classification. The red and yellow numbers also in bold stand for the minimum and subminimum DBI of each kind of feature scale parameter list. The corresponding scale parameters of the red and yellow DBI are the optimized ones.

Fractal dimension											
Radii	1–3	1–4	1–5	1–6	1–7	1–8	1–9	2–8	0–3	0–4	0–7
Osteopenia vs. normal	6.215	7.179	7.789	11.006	8.094	7.724	9.552	15.059	8.620	11.044	9.475
Osteoporosis vs. normal	9.524	13.472	11.821	10.014	7.297	7.098	7.455	8.203	10.419	9.178	6.774
Osteoporosis vs. osteopenia	18.335	7.360	14.312	13.597	7.364	5.400	9.681	10.425	13.035	15.143	10.153
Lacunarity											
Size	2*2	3*3	4*4	5*5	6*6	7*7					
Osteopenia vs. normal	2.915	7.826	34.919	27.391	23.874	22.148					
Osteoporosis vs. normal	7.408	10.758	9.849	10.310	7.548	7.900					
Osteoporosis vs. osteopenia	2.358	4.367	7.229	7.026	5.512	5.564					
Entropy											
Measure radii, count radii	1,4	1,5	1,6	1,7	1,8	1,9	1,10	1,11	2,6	2,9	
Osteopenia vs. normal	19.557	20.941	14.669	11.817	10.407	9.910	10.471	10.625	10.747	14.478	
Osteoporosis vs. normal	6.676	4.697	4.225	4.289	4.496	4.679	5.060	5.728	5.639	6.352	
Osteoporosis vs. osteopenia	7.015	5.204	5.117	5.242	5.242	5.937	6.123	6.704	7.070	6.753	

listed as follows: “OptimizeHyperparameters,” “all;” “Standardize,” true; “HyperparameterOptimizationOptions,” struct (“AcquisitionFunctionName,” “probability-of-improvement,” UseParallel’, true, “Repartition,” true, “Kfold,” 10, “MaxObjectiveEvaluations,” 800). Considering that the learning process is time consuming, we used the command pair *parpool* and *delete [gcp(“nocreate”)]* to embrace the *fitsvm* function, which would be executed in a multi-core CPU in parallel.

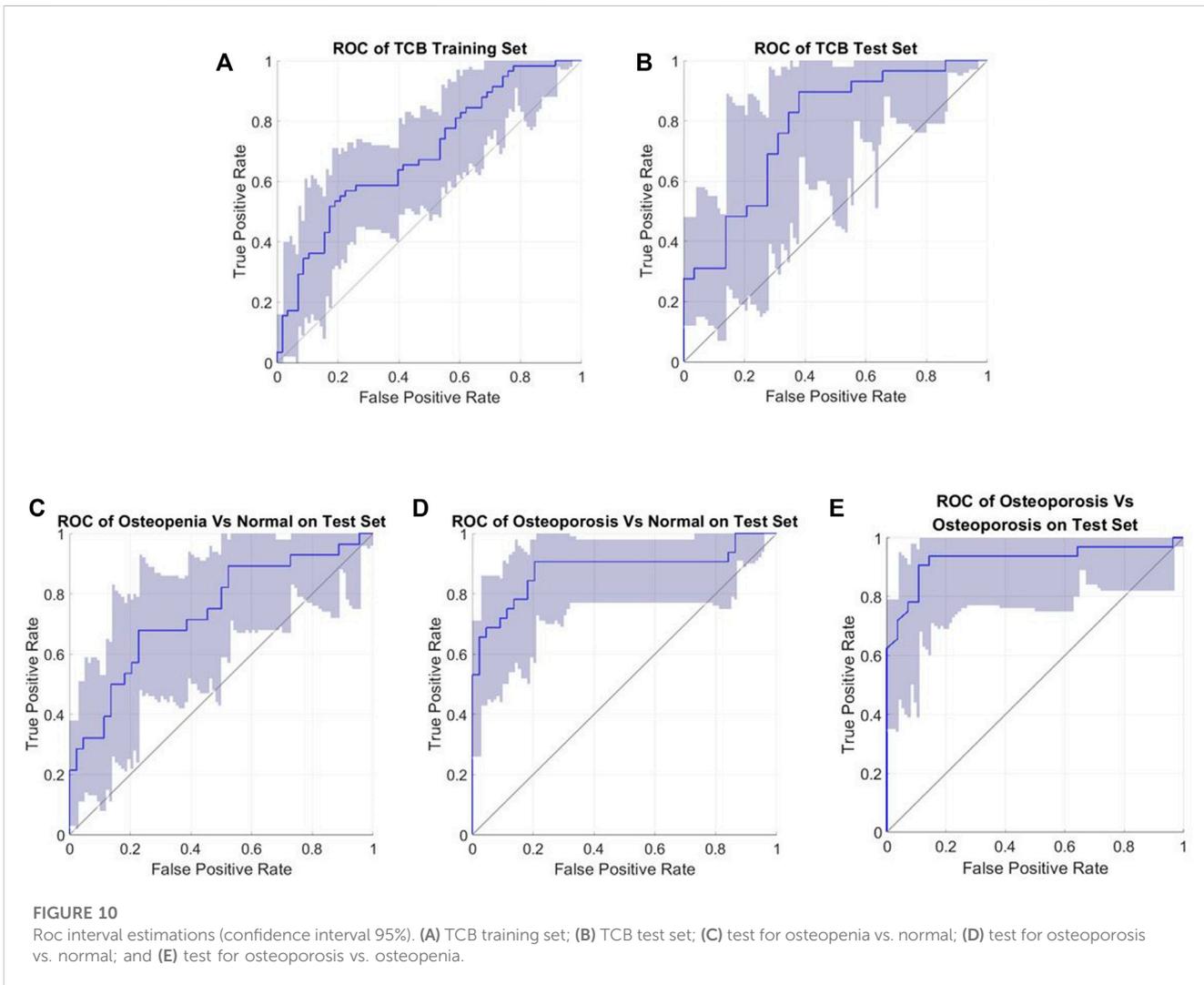
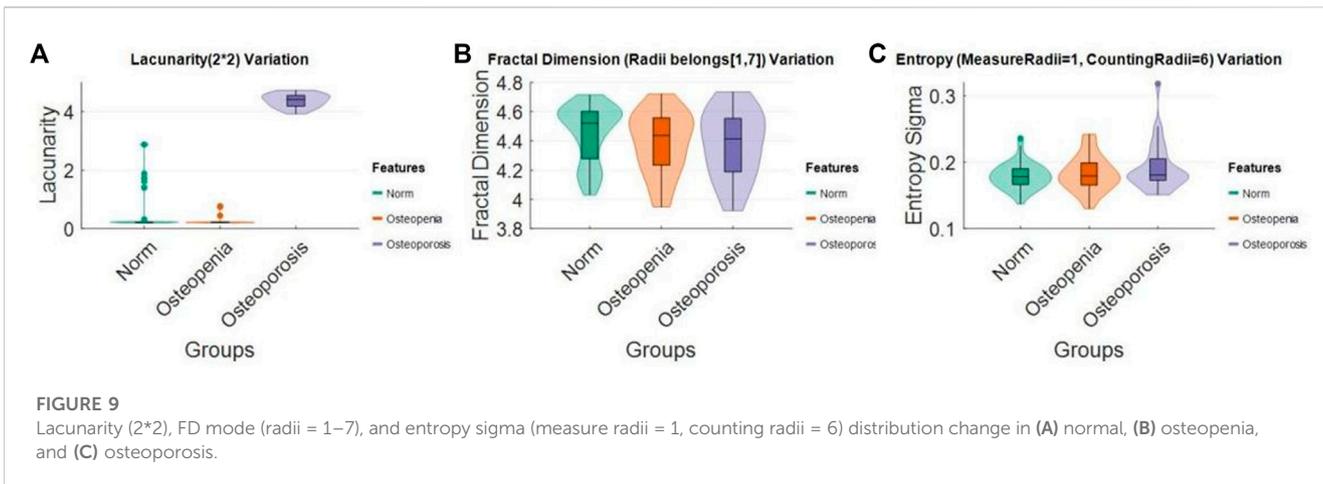
2.7 Classifier validation

For the TCB blind test, we estimated receiver operating characteristic (ROC), area under the curve (AUC), and Matthews correlation coefficient (MCC) with a 95% confidence interval and 1,000 times bootstrap. In addition, to use the TCB ranking for reference, we also calculated accuracy (Acc), sensitivity (Sn), and specificity (Sp).

For the clinical L4 test set, due to the limitation of sample size on the full dataset, we performed 1,000 times bootstrap. We classified normal, osteopenia, and osteoporosis in pairs. Interval estimates of ROC and AUC in between each pair of them were calculated.

3 Results

As hyperparameters, feature counting scales in pixels were optimized by choosing the one with the smallest DBI of pixel-wise FD, global lacunarity, and pixel-wise entropy. For the TCB dataset, the FD radii range was from 1 to 4, and the entropy measure radii and count radii were 1 and 6, respectively. Also, the two optimized lacunarity sizes were 3*3 and 4*4 (Table 1). For the clinical dataset (Table 2), there were three scenarios. In the scenario osteopenia vs. normal, the FD radii range was from 1 to 3, the entropy measure radii and count radii were 1 and 9,



and the two optimized lacunarity sizes were 2*2 and 3*3, respectively. In the scenario osteoporosis vs. normal, the FD radii range was from 0 to 7, the entropy measure radii and count radii were 1 and 6, and the two optimized lacunarity sizes were

2*2 and 6*6, respectively. In osteoporosis vs. osteopenia, the radii range was from 1 to 4, the entropy measure radii and count radii were 1 and 7, and the two optimized lacunarity sizes were 2*2 and 3*3, respectively. In tables, those preferred scales were marked in

TABLE 3 Presented classifier with the Acc marked in red was compared with the TCB ranking.

Rank	Author	TP	FP	TN	FN	Sn	Sp	Acc	F1 score
	Presented classifier ^a	16	6	23	13	0.552	0.793	0.672	0.628
1	Yger (2014)	19	10	19	10	0.655	0.655	0.655	0.655
2	Palanivel et al. (2020)	17	12	17	12	0.586	0.586	0.586	0.586
3	Christoph et al. (Ranking, 2022)	24	22	7	5	0.828	0.241	0.535	0.640
4	Jerome et al. (Ranking, 2022)	21	17	12	8	0.724	0.414	0.569	0.62

^aThe whole image dataset was repartitioned into a train set and blind test set with the prescribed ratio of 2:1.

TABLE 4 Classification performance of L4 on the clinical dataset.

Scenario	AUC	Accuracy	Precision	Sensitivity	Specificity	F1 score	MCC
	(95% CI)						
Osteopenia vs. normal	0.721	0.675	0.732	0.774	0.537	0.757	0.342
	(0.578, 0.824)	(0.563, 0.772)	(0.592, 0.833)	(0.635, 0.878)	(0.351, 0.720)	(0.659, 0.855)	(0.114, 0.547)
Osteoporosis vs. normal	0.851	0.813	0.807	0.936	0.690	0.874	0.681
	(0.730, 0.922)	(0.718, 0.878)	(0.679, 0.898)	(0.826, 1)	(0.483, 0.829)	(0.783, 0.932)	(0.501, 0.811)
Osteoporosis vs. osteopenia	0.935	0.928	0.964	0.910	0.967	0.943	0.899
	(0.830, 0.968)	(0.863, 0.963)	(0.810, 1)	(0.746, 1)	(0.829, 1)	(0.849, 0.986)	(0.699, 0.967)

red with a straight underscore and orange with a wave underscore.

Distribution densities of some selected features measured with a fix scale in normal, osteopenia, and osteoporosis were visualized (Figure 9). The lacunarity (2*2) of osteoporosis jumped absolutely out of the range of its value in other cases. However, for the FD mode value and entropy standard deviation, we observed the changing distribution trend, but it was not significant.

The six most effective features were selected, and the effectiveness for classification was ranked by weight (Figure 8). The six most significant features were prioritized: FD mode, FD sigma, FD kurtosis, lacunarity (3*3), FD skewness, and entropy kurtosis for the TCB task. The priority order was FD kurtosis, FD skewness, FD sigma, entropy sigma, lacunarity (2*2), and lacunarity (3*3) for osteopenia vs. normal from the clinical L4 dataset. The priority order was entropy kurtosis, entropy sigma, FD mode, entropy skewness, entropy mode, and lacunarity (2*2) for osteoporosis vs. normal from the clinical L4 dataset. The priority order was FD skewness, lacunarity (3*3), entropy skewness, FD entropy, lacunarity (2*2), and entropy kurtosis for osteoporosis vs. osteopenia from the clinical L4 dataset.

All related ROC interval estimations on the test set are depicted in Figure 10. The interval estimation classification on L4 is listed in Table 4. For osteoporosis diagnosis, the AUC, Acc, and Sn (recall) were 0.851 (95% CI: 0.730–0.922), 0.813(95% CI: 0.718–0.878), and 0.936 (95% CI: 0.826–1), respectively. For osteopenia diagnosis, the AUC, Acc, and Sn were 0.721 (95% CI: 0.578–0.824), 0.675 (95% CI: 0.563–0.772), and 0.774 (95% CI: 0.635–0.878), respectively. For osteoporosis diagnosis from osteopenia, the AUC, Acc, and Sn were

0.935(95% CI: 0.830–0.968), 0.928 (95% CI: 0.863–0.963), and 0.910 (95% CI: 0.746–1), respectively.

Our TCB training set and test set were repartitioned, but the performance was comparable with the other feature-based and SVM classifiers within the TCB ranking (Table 3). For the training dataset, the AUC, Acc, and Sn were 0.694 (95% CI: 0.592–0.778), 0.663 (95% CI: 0.569–0.741), and 0.742 (95% CI: 0.607–0.839), respectively. For the blind test set the AUC, Acc, and Sn were 0.767 (95% CI: 0.629–0.879), 0.672 (95% CI: 0.545–0.793), and 0.790 (95% CI: 0.621–0.923), respectively. Here, the Acc was slightly higher than that of the method from the work of Florian et al.

4 Discussion

Unlike most of the previously reported classifiers (Zheng and Makrogiannis, 2016; Oulhaj et al., 2017; Palanivel et al., 2020; Omiotek et al., 2021), our study exhibited that a novel classifier synthesizing pixel-wise fractal, entropy, and global lacunarity can distinguish three levels of bone mass loss of lumbar spine into normal, osteopenia, and osteoporosis. In addition, it also showed high generalization on osteoporosis diagnosis using calcaneus according to lateral radiographs, which are both routine physical examinations in a broader sense. For a typical DXA examination focus, such as the vertebral spine, or non-DXA examination, such as the calcaneus, the proposed classifier can be successfully applied. The performance of the proposed classifier was reliable, indicating whatever the T-score or the fractured state as the golden standard. In all aforementioned scenarios, Acc is higher than 0.66 and Sn is

higher than 0.74. This method can be a promising tool to diagnose osteopenia and osteoporosis on X-ray images for further clinical application.

Several features trended from normal health to osteopenia and osteoporosis. We can try to partially explain the physio-pathological change process of a bone trabecular structure. From [Figure 9A](#), lacunarity drastically rose in the process of pathological changes from normal to osteoporosis. It means the image intensity turns more heterogeneous in the whole image. This phenomenon suggests some lacunarity measure patches of bone degeneration far different from the other groups. In [Figure 9B](#), the FD population distribution of each group mildly declined. This trend could be checked with median values and lower bounds. Lower FD means the bone mass loss happened on a considerable scale range. From [Figure 9C](#), we found that each group's pixel-wise entropy population's standard deviation sigma slightly increased. Median values and upper bound depicted the increase. From [Figures 3G, N, 4G, N, U](#), we knew the pixel-wise entropy distribution was unimodal. According to information theory ([Chung et al., 2017](#)), the sigma of unimodal distribution indicates the population entropy. So, in other words, the worse the osteoporosis, the higher the global entropy, which is generally following the entropy law, also named the second law of thermodynamics ([Schneider and Kay, 1994](#)). The osteoporotic pathogenesis was unprecedentedly depicted through the 2D projection entropy. The entropy could be taken into account as a new osteoporosis biomarker.

The classifier performance was comparable to other latest reports. The state-of-the-art DL-based classification models held an average accuracy of 90.1% ranging from 70.0% to 98.9% and a mean AUC of 0.90 ranging from 0.74 to 1.00 ([Zhang et al., 2020](#)) for only osteoporosis diagnosis from healthy subjects. As shown in [Table 4](#), roughly speaking, from the performance point of view, our feature-based model matched with those DL-based models. For osteopenia diagnosis, as we know, the latest DL-based model ([Zhang et al., 2020](#)) acquired an AUC of 0.629 with 95% CI: (0.545, 0.707) and sensitivity of 50% with 95% CI: (32.4, 67.6). As in [Table 4](#), our performance, with an AUC of 0.741 with 95% CI: (0.598, 0.844) and sensitivity of 79.4% with 95% CI: (0.655, 0.898), was comparable. High-sensitivity results in a low fraction of false-negative classifications, in other words, a low missed diagnosis rate. Apart from the aforementioned, we also classified osteoporosis from osteopenia. The AUC and Acc are even better than the osteoporosis diagnosis in [Table 4](#). In summary, it is possible to screen osteoporosis from the lateral radiograph of L4 with interpretable features matched with DL-based classifiers.

The TCB challenge dataset showing characteristic texture had been held since 2014, and here, it was also utilized to show the high generalization of our classifier. The present TCB challenge requires participants to use the SVM classifier, and so, only feature-based methods were taken into account in the ranking list according to their accuracy. In this dataset, the number of osteoporosis cases is identical to the number of healthy cases, and the case ratio for train and test was 2 to 1. To the best of our knowledge, counting follow-up studies are listed in [Table 3](#). On the TCB challenge, no report has a better score than ours yet, regardless of our random dataset repartition. The latest hit list models ([Palanivel et al., 2019; 2020](#)) have paid attention to the heterogeneity of the texture and combined multifractal features with lacunarity to gain improvement. However,

at the early stages of osteoporosis, bone loss happens at individual trabecular plates ([Cosman et al., 2014; Golob and Laya, 2015](#)), so our pixel-wise feature counting has the potential for early diagnosis of osteoporosis.

The proposed image preprocess is advantageous to the following texture feature extraction. The fractal dimension, lacunarity, and entropy are scale sensitive to all three feature categories. They all depicted one facet of trabecular structure changes, whereas the source-to-image distance (SID) and the source-to-object distance (SOD) were not uniformed among clinical radiographs. We carried out imaging scale (pixel size) normalization by SOD vs. SID line fitting for radiographs. Concerning the curved shape of human bones whether lumbar spine or calcaneus, the radiography image intensity was naturally biased. Especially for lateral radiographs of the lumbar spine, there was the thickness of fat, organs, and even clothes having an overlay projection on the image. To limit those negative impacts, we compensated for the bias with Gaussian smoothing. Due to the influence of the thickest soft tissue on the human body, lateral radiographs of the lumbar spine were a blurred with scatters. We utilized the dehaze algorithm to enhance the detailed information on bone texture.

A robust measure, the image surface area, was adopted to overcome the superposition imaging of other human organs and tissues rather than the image intensity analyzed by previous studies ([Zheng and Makrogiannis, 2016; Hassouni et al., 2017](#)). Lateral radiographs of the lumbar spine are widely available with low cost. However, subjects' abdominal organs, cavities, and intestinal gas even waist belts could be all projected in radiographs. Image surface area counting was practiced to overcome those complex impacts ([Lopes and Betrouni, 2009; Areeckal et al., 2018b](#)). Fractal dimension features were taken with the local surface area as a measure. Lacunarity and local entropy features were taken based on intensity difference which was a derivative method of image surface measure.

In our study, the proposed model capacity was well restrained to only six feature dimensions not only for better generalization but also as a good foundation for further regressive analysis. In the field of ML, a principle of parsimony ([Vapnik, 1999; Vapnik and Chervonenkis, 2015](#)) named Occam's razor is widely acknowledged; among competing hypotheses that explain known observations equally well, we should choose the "simplest" one. Furthermore, the proposed hyperparameter optimization, feature selection, and classifier training process arguably follow the concept of ensemble learning which had the property of improving generalization ([Hansen and Salamon, 1990](#)). Compared with the latest studies for osteoporotic diagnosis, a convolutional neural network (CNN) ([Yamamoto et al., 2020; Zhang et al., 2020](#))-based model had dozens or even hundreds of non-explanatory and abstract features, and real feature-based models ([Oulhaj et al., 2017; Palanivel et al., 2020](#)) also collected dozens of feature dimensions to obtain an ideal performance. What we gained from restrained feature dimensions was a better generalization. The AUC of 0.767 with 95% CI: (0.622, 0.874) of the TCB test dataset was not worse than the AUC of 0.694 with 95% CI: (0.591, 0.778) in the training dataset.

There are still several limitations of this study. First, as the golden standard, the BMD value to diagnose osteoporosis is still questionable and challenged by bone histomorphometry studies ([Wainwright et al., 2005; Greenwood et al., 2018](#)), so a high-resolution CT exam is suggested as a powerful complement to

the DXA exam for osteoporotic diagnosis (Greenwood et al., 2018). Second, the labeled sample size is restricted to a bit more than one hundred. The data scale is not sufficient and diverse to bring this method to real clinical trials, especially for a triple-classification problem, in which osteopenia and osteoporosis are both analyzed. Third, in this study, for vertebral spine radiography analysis, only L4 was analyzed, which is similar to a recent study (Lee et al., 2020). Nevertheless, a DXA scan is from L1 to L4. A corresponding skeleton feature extraction could be a better way to obtain classification results matching T-scores.

This method exhibited the capacity of resisting disturbance on lateral spine radiographs and high generalization on the calcaneus dataset. Interpretable pixel-wise features not only helped to better understand osteoporosis on radiographs but also shed new light on computer-aid diagnosis for osteopenia and osteoporosis.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by the Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. The patients/participants provided their written informed consent to participate in this study.

Author contributions

The authors confirm contribution to the paper as follows: study conception and design: T-YT and JC; data collection: CL

References

- Amarnath, A. L. D., Franks, P., Robbins, J. A., Xing, G., and Fenton, J. J. (2015). Underuse and overuse of osteoporosis screening in a regional health system: A retrospective cohort study. *J. Gen. Intern. Med.* 30, 1733–1740. doi:10.1007/s11606-015-3349-8
- Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., and Perona, I. (2013). An extensive comparative study of cluster validity indices. *Pattern Recognit.* 46, 243–256. doi:10.1016/j.patcog.2012.07.021
- Areeckal, A. S., Jayasheelan, N., Kamath, J., Zawadynski, S., Kocher, M., David S, S., et al. (2018a). Early diagnosis of osteoporosis using radiogrammetry and texture analysis from hand and wrist radiographs in Indian population. *Osteoporos. Int.* 29, 665–673. doi:10.1007/s00198-017-4328-1
- Areeckal, A. S., Kocher, M., and David, S. (2018b). Current and emerging diagnostic imaging-based techniques for assessment of osteoporosis and fracture risk. *IEEE Rev. Biomed. Eng.* 12, 254–268. doi:10.1109/RBME.2018.2852620
- Chung, H. W., Sadler, B. M., and Hero, A. O. (2017). Bounds on variance for unimodal distributions. *IEEE Trans. Inf. Theory* 63, 6936–6949. doi:10.1109/TIT.2017.2749310
- Conci, A., and Campos, C. F. J. (1996). An efficient box-counting fractal dimension approach for experimental image variation characterization. *Proceedings IWISP'96*, 665–668.
- Cosman, F., de Beur, S. J., LeBoff, M. S., Lewiecki, E. M., Tanner, B., Randall, S., et al. (2014). Clinician's guide to prevention and treatment of osteoporosis. *Osteoporos. Int.* 25, 2359–2381. doi:10.1007/s00198-014-2794-2
- Costa, M. D., and Goldberger, A. L. (2015). Generalized multiscale entropy analysis: Application to quantifying the complex volatility of human heartbeat time series. *Entropy* 17, 1197–1203. doi:10.3390/e17031197
- Creutzberg, R., and Ivanov, E. (1989). "Computing fractal dimension of image segments," in *Proceedings of the 3rd international conference of computer analysis of images and patterns (CAIP'89)*. (Berlin, Germany: Springer-Verlag).
- Davies, D. L., and Bouldin, D. W. (1979). A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-1, 224–227. doi:10.1109/tpami.1979.4766909
- Enkins, J., Riffin, R. O. J. G., Ings, R. U. U. D. P. M. D., Arasimhan, N., Ajaram, R., Uinn, K. Y. L. E. P. Q., et al. (2018). Rapid quantification of mitochondrial fractal dimension in individual cells. *Biomed. Opt. Express* 9, 5269–5279. doi:10.1364/boe.9.005269
- Golob, A. L., and Laya, M. B. (2015). Osteoporosis: Screening, prevention, and management. *Med. Clin. North Am.* 99, 587–606. doi:10.1016/j.mcna.2015.01.010
- Greenwood, C., Clement, J., Dicken, A., Evans, P., Lyburn, I., Martin, R. M., et al. (2018). Age-related changes in femoral head trabecular microarchitecture. *Aging Dis.* 9, 976–987. doi:10.14336/AD.2018.0124
- Hansen, L. K., and Salamon, P. (1990). Neural network ensembles. *IEEE Trans. Pattern Anal. Mach. Intell.* 12, 993–1001. doi:10.1109/34.58871
- Harrar, K., Hamami, L., Lespessailles, E., and Jennane, R. (2013). Piecewise Whittle estimator for trabecular bone radiograph characterization. *Biomed. Signal Process. Control* 8, 657–666. doi:10.1016/j.bspc.2013.06.009

and RJ; analysis and interpretation of results: JC and KD; and draft manuscript preparation: JC and T-YT. All authors reviewed the results and approved the final version of the manuscript.

Funding

This research was supported by the Pudong New Area Science & Technology and Development Fund under award number 210H1147900 and by the Fundamental Research Funds for the Central Universities under award number AF0820055.

Acknowledgments

Researcher Layli Turak provided great assistance in modifying and polishing the literature expression. We are very grateful for her efforts. All our authors would like to thank the anonymous reviewer for their valuable and constructive comments.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Hassouni, M. El, Tafaouti, A., Toumi, H., Lespessailles, E., and Jennane, R. (2017). Fractional brownian motion and rao geodesic distance for bone X-ray image characterization. *IEEE J. Biomed. Heal. Inf.* 21, 1347–1359. doi:10.1109/JBHI.2016.2619420
- Humeau-Heurtier, A. (2016). Multivariate refined composite multiscale entropy analysis. *Phys. Lett. A* 380, 1426–1431. doi:10.1016/j.physleta.2016.02.029
- Hussain, D., and Han, S.-M. (2019). Computer-aided osteoporosis detection from DXA imaging. *Comput. Methods Programs Biomed.* 173, 87–107. doi:10.1016/j.cmpb.2019.03.011
- Jennane, R., Ohley, W. J., Majumdar, S., and Lemineur, G. (2001). Fractal analysis of bone X-ray tomographic microscopy projections. *IEEE Trans. Med. Imaging* 20, 443–449. doi:10.1109/42.925297
- Johell, O., and Kanis, J. A. (2006). An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporos. Int.* 17, 1726–1733. doi:10.1007/s00198-006-0172-4
- Karperien, A., Jelinek, H. F., and Milosevic, N. (2011). “Reviewing lacunarity analysis and classification of microglia in neuroscience,” in *Proceedings of the 8th European conference on mathematical and theoretical biology*. (Kraków, Poland: Springer). Available at: http://www.researchgate.net/profile/Audrey_Karperien/publication/230857319_Lacunarity_Analysis_and_Classification_of_Microglia_in_Neuroscience/links/0c9605288742c529d8000000.pdf.
- Lee, S., Choe, E. K., Kang, H. Y., Yoon, J. W., and Kim, H. S. (2020). The exploration of feature extraction and machine learning for predicting bone density from simple spine X-ray images in a Korean population. *Skelet. Radiol.* 49, 613–618. doi:10.1007/s00256-019-03342-6
- Li, J., Du, Q., and Sun, C. (2009). An improved box-counting method for image fractal dimension estimation. *Pattern Recognit.* 42, 2460–2469. doi:10.1016/j.patcog.2009.03.001
- Lopes, R., and Betrouni, N. (2009). Fractal and multifractal analysis: A review. *Med. Image Anal.* 13, 634–649. doi:10.1016/j.media.2009.05.003
- Mandelbrot, B. B. (1994). “A Fractal’s Lacunarity, and how it can be Tuned and Measured,” in *Fractals in biology and medicine* (Berlin, Germany: Springer), 8–21.
- National Health Commission of China, Epidemiological investigation of osteoporosis in China (2018). Peoples health. Network. Available at: <https://www.phsciencedata.cn/Share/jsp/PublishManager/foregroundView/1/9eaadbb3-bd64-4531-9d9f-753ec183f26d.html> (Accessed August 1, 2022).
- Office of the Surgeon General (US) (2004). “Bone health and osteoporosis: A report of the Surgeon general,” in *US health and human services*. (Rockville, MD) 5. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Bone+Health+and+Osteoporosis+A+Report+of+the+Surgeon+General#0>.
- Omiotek, Z., Dzierzak, R., and Kepa, A. (2021). Fractal analysis as a method for feature extraction in detecting osteoporotic bone destruction. *Fractals* 29, 2150095. doi:10.1142/S0218348X2150095X
- Oulhaj, H., Rziza, M., Amine, A., Toumi, H., Lespessailles, E., Jennane, R., et al. (2017). Trabecular bone characterization using circular parametric models. *Biomed. Signal Process. Control* 33, 411–421. doi:10.1016/j.bspc.2016.10.009
- Overview (2022). Orleans university. Available at: <https://www.univ-orleans.fr/en/overview> (Accessed June 5, 2022).
- Palanivel, D. A., Natarajan, S., Gopalakrishnan, S., and Jennane, R. (2020). Multifractal-based lacunarity analysis of trabecular bone in radiography. *Comput. Biol. Med.* 116, 103559. doi:10.1016/j.compbiomed.2019.103559
- Palanivel, D. A., Natarajan, S., Gopalakrishnan, S., and Jennane, R. (2019). “Trabecular bone texture characterization using regularization dimension and box-counting dimension,” in *TENCON 2019-2019 IEEE region 10 conference (TENCON)*. (Kochi, India: IEEE), 1047–1052.
- Peleg, S., Naor, J., Hartley, R., and Avnir, D. (1984). Multiple resolution texture analysis and classification. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI- 6, 518–523. doi:10.1109/TPAMI.1984.4767557
- Raghu, S., and Sriraam, N. (2018). Classification of focal and non-focal EEG signals using neighborhood component analysis and machine learning algorithms. *Expert Syst. Appl.* 113, 18–32. doi:10.1016/j.eswa.2018.06.031
- Ranking (2022). Orleans university. Available at: <https://www.univ-orleans.fr/fr/i3mto/innovation/challenge-ieee-isbi-bone-texture-characterization/resultsw> (Accessed March 7, 2022).
- Reiter, R. (1987). A theory of diagnosis from first principles. *Artif. Intell.* 32, 57–95. doi:10.1016/0004-3702(87)90062-2
- Sarkar, N., and Chaudhuri, B. B. (1994). An efficient differential box-counting approach to compute fractal dimension of image. *IEEE Trans. Syst. Man. Cybern.* 24, 115–120. doi:10.1109/21.259692
- Scanlan, J., Li, F. F., Umnova, O., Rakoczy, G., Lövey, N., and Scanlan, P. (2018). Detection of osteoporosis from percussion responses using an electronic stethoscope and machine learning. *Bioengineering* 5, 107. doi:10.3390/bioengineering5040107
- Schneider, E. D., and Kay, J. J. (1994). Life as a manifestation of the second law of thermodynamics. *Math. Comput. Model.* 19, 25–48. doi:10.1016/0895-7177(94)90188-0
- Shannon, C. E. (2001). A mathematical theory of communication. *ACM Sigmob. Mob. Comput. Commun. Rev.* 5, 3–55. doi:10.1145/584091.584093
- Smets, J., Shevroja, E., Hügle, T., Leslie, W. D., and Hans, D. (2021). Machine learning solutions for osteoporosis—a review. *J. Bone Min. Res.* 36, 833–851. doi:10.1002/jbmr.4292
- Smith, T. G., Lange, G. D., and Marks, W. B. (1996). Fractal methods and results in cellular morphology - dimensions, lacunarity and multifractals. *J. Neurosci. Methods* 69, 123–136. doi:10.1016/S0165-0270(96)00080-5
- Tang, J. (2022). COVID-19 pandemic and osteoporosis in elderly patients. *Aging Dis.* 13, 960–969. doi:10.14336/AD.2021.1201
- Teclé, N., Teitel, J., Morris, M. R., Sani, N., Mitten, D., and Hammert, W. C. (2020). Convolutional neural network for second metacarpal radiographic osteoporosis screening. *J. Hand Surg. Am.* 45, 175–181. doi:10.1016/j.jhsa.2019.11.019
- Tu, K. N., Lie, J. D., Wan, C. K. V., Cameron, M., Austel, A. G., Nguyen, J. K., et al. (2018). Osteoporosis: A review of treatment options. *P T* 43, 92–104.
- Vapnik, V. N., and Chervonenkis, A. Y. (2015). “On the uniform convergence of relative frequencies of events to their probabilities,” in *Measures of complexity* (Berlin, Germany: Springer), 11–30.
- Vapnik, V. (1999). *The nature of statistical learning theory*. Berlin, Germany: Springer science & business media.
- Wainwright, S. A., Marshall, L. M., Ensrud, K. E., Cauley, J. A., Black, D. M., Hillier, T. A., et al. (2005). Hip fracture in women without osteoporosis. *J. Clin. Endocrinol. & Metab.* 90, 2787–2793. doi:10.1210/jc.2004-1568
- Xiao, P., Zhang, T., Dong, X. N., Han, Y., Huang, Y., and Wang, X. (2020). Prediction of trabecular bone architectural features by deep learning models using simulated DXA images. *Bone Rep.* 13, 100295. doi:10.1016/j.bonr.2020.100295
- Yamamoto, N., Sukegawa, S., Kitamura, A., Goto, R., Noda, T., Nakano, K., et al. (2020). Deep learning for osteoporosis classification using hip radiographs and patient clinical covariates. *Biomolecules* 10, 1534. doi:10.3390/biom10111534
- Yang, W., Wang, K., and Zuo, W. (2012). Neighborhood component feature selection for high-dimensional data. *J. Comput.* 7, 161–168. doi:10.4304/jcp.7.1.161-168
- Yger, F. (2014). *Application of covariance matrices and wavelet marginals*. *arXiv Prepr. arXiv1410.2663*.
- Zhang, B., Yu, K., Ning, Z., Wang, K., Dong, Y., Liu, X., et al. (2020). Deep learning of lumbar spine X-ray for osteopenia and osteoporosis screening: A multicenter retrospective cohort study. *Bone* 140, 115561. doi:10.1016/j.bone.2020.115561
- Zheng, K., and Makrogiannis, S. (2016). Bone texture characterization for osteoporosis diagnosis using digital radiography. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2016, 1034–1037. doi:10.1109/EMBC.2016.7590879