Full Supervised Learning for Osteoporosis Diagnosis Using Micro-CT Images

YAN XU,[1] DIANSHI LI,[2] QINLANG CHEN,[1] AND YUBO FAN[1]*
[1]*State Key Laboratory of Software Development Environment, Key Laboratory of Biomechanics and Mechanobiology of Ministry of Education, Beihang University, Beijing, 100191, China*
[2]*Department of Electrical and Computer, School of Engineering, Northeastern University, Massachusetts 02115*

KEY WORDS osteoporosis; trabecular bone; micro-CT images; volumetric topological analysis

ABSTRACT Early osteoporosis diagnosis is of important significance for reducing fracture risk. Image analysis provides a new perspective for noninvasive diagnosis in recent years. In this article, we propose a novel method based on machine-learning method performed on micro-CT images to diagnose osteoporosis. The aim of this work is to find a way to more effectively and accurately diagnose osteoporosis on which many methods have been proposed and practiced. In this method, in contrast to the previously proposed methods in which features are analyzed individually, several features are combined to build a classifier for distinguishing osteoporosis group and normal group. Twelve features consisting of two groups are involved in our research, including bone volume/total volume (BV/TV), bone surface/bone volume (BS/BV), trabecular number (Tb.N), obtained from the software of micro-CT, and other four features from volumetric topological analysis (VTA). Support vector machine (SVM) method and k-nearest neighbor (kNN) method are introduced to create classifiers with these features due to their excellent performances on classification. In the experiment, 200 micro-CT images are used in which half are from osteoporosis patients and the rest are from normal people. The performance of the obtained classifiers is evaluated by precision, recall, and F-measure. The best performance with precision of 100%, recall of 100%, and F-measure of 100% is acquired when all the features are included. The satisfying result demonstrates that SVM and kNN are effective for diagnosing osteoporosis with micro-CT images. *Microsc. Res. Tech. 76:333–341, 2013.*

INTRODUCTION

Osteoporosis is one of the commonest medical conditions which troubles thousands of people in the world (National Institute of Health, 1984). It has been reported that 15–30% men and 30–50% women in the world face suffering from osteoporosis in their life (He et al., 2007). In the aged group, osteoporosis is the major reason resulting in fractures, which usually cause paralysis, even death. With the advent of the aging society, it is of great importance to find an effective way for early osteoporosis detection to decrease the fracture risk. In recent decades, many efforts have been taken for this goal. But the diagnosis accuracy is still not satisfying, which underlines the need for more advanced methods. The most obvious characteristics of osteoporosis are bone mineral density (BMD) reduction and trabecular bone (TB) structure deterioration (Wehrli et al., 2001), which lead to increase in fracture risk. Micro-CT images of TB from osteoporosis patients and normal people are illustrated in Figure 1, respectively (Li and Xu, 2012). In osteoporosis group, BMD reduces due to loss of minerals. BMD is related to bone strength and can be used to predict fracture risk (Baum et al., 2010). In normal aging process, BMD can decrease by 50%, followed by decrease in bone strength by up to 70–80% (Mosekilde et al., 2000). This strong correlation between the decrease in BMD and bone strength directly implies the great influence BMD has on bone strength. Peacock et al. (1995) introduced BMD to discriminate hip fracture. In their research, testing of BMD of fractured and nonfractured women showed that decreases in BMD strongly increase the hip fracture risk. Though BMD is seen as an important factor to assess bone strength, it can only on the average explain about 60% of the bone strength (Wehrli et al., 2002). The rest 40% is related to TB structure, which also plays an important role in bone strength. TB is a complex network of plates and rods and makes up the major parts of cancellous bone (Saha et al., 2010), the structure of which usually has a clear difference between osteoporosis people and normal people. Thus, it has been used as a key indicator for osteoporosis detection. A few related researches have been done in recent years. Pothuaud et al. (2002) combined BV/TV and topological features of TB to predict the mechanical properties of TB.

*Correspondence to: Yubo Fan, School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China. E-mail: yubofan@buaa.edu.cn

Received 14 November 2012; accepted in revised form 21 December 2012

Contract grant sponsor: National Science Foundation of China; Contract grant number: 61073077; Contract grant sponsor: State Key Laboratory of Software Development Environment, Key Laboratory of Biomechanics and Mechanobiology of Ministry of Education, Beihang University, China; Contract grant number: SKLSDE-2011ZX-13.

DOI 10.1002/jemt.22171
Published online 17 January 2013 in Wiley Online Library (wileyonlinelibrary.com).

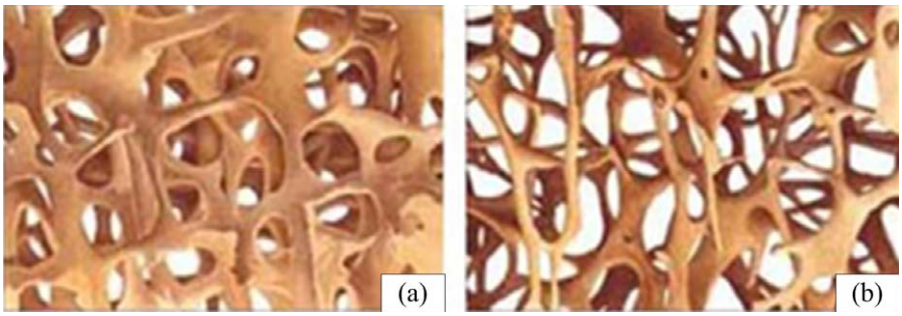

Fig. 1. Micro-CT images of TB. (a) TB of normal people. (b) TB of osteoporosis patients. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Researches on osteoporosis diagnosis in early stage are majorly based on histomorphometry analysis (Parfitt et al., 1987). In these researches, many features were proposed to describe the TB's structure and strength. Parfitt et al. (1987) referred to some two-dimensional (2D) features for bone measurement, like area, length, and distance between points. However, these features are only profiles of three-dimensional (3D) structures, and are not able to reflect the accurate 3D characters of TB. Though some 3D features are also proposed in his researches, like BV/TV and trabecular thickness, these measurements are not the same as current ones, as they are measured by manual labor, instead of real 3D analysis. Due to the lack of direct 3D analysis, these methods cannot accurately describe TB's inner structure and the result is usually not sufficiently convincing.

In later researches, image processing and analysis based on X-ray and CT became the mainstream (Cervinka et al., 2010; Lee et al., 2008). These methods do not rely on manual measurement but on image processing and analysis. All features are calculated by programs based on corresponding algorithms, which can provide analysis with higher accuracy. Lee et al. (2008) used plain X-ray images to distinguish the osteoporosis group and nonosteoporosis group. In his research, 94 women were studied in which 47 belonged to the nonfractured group and 47 experienced osteoporotic fractures. All features are extracted from these X-ray images, and further used to create a classifier, the accuracy of which is about 75%. Cervinka et al. (2010) also proposed an approach to improve bone structure analysis by pQCT image preprocessing. Even though such methods provide direct 3D analysis, the resolution is limited. Micro-CT changes the situation with a maximum resolution of 8 μmm and largely improves the image quality (Saha et al., 2010). Hence, high resolution 3D images can be obtained providing deeper insight into the micro-structure of TB. Following that, relevant 3D analysis has been developed and many effective features related to TB structure could be calculated more accurately, namely, trabecular separation (Tb.Sp), trabecular pattern factor (Tb.Pf), structure model index (SMI) (Baum et al., 2010), etc. These features have good performance on distinguishing between osteoporosis group and normal group in many situations. Ulrich et al. (1999) used micro-CT to scan 237 cancellous bone samples from lumbar spine, femoral head, etc. Features like BV/TV, BS/BV, Tb.Sp, Tb.Th were calculated from 3D segmented images, of which the mean, SD minimum, and maximum values were also obtained. Laib and Ruegsegger (1999) have done a research to compare the performance of pQCT and micro-CT on TB structure measurements. Micro-CT could produce a better result with higher resolution. Besides micro-CT, MRI is also applied to analyze TB's structure, which has been used in many researches (Gamio et al., 2006; Petrantonaki et al., 2005; Phan et al., 2006).

In many researches, features for osteoporosis diagnosis are usually analyzed individually ignoring the inner relationship between them. This may be problematic as a single feature cannot be effective in all situations. For example, BMD has a higher average value in normal people, which is considered as a golden standard for osteoporosis prediction. But in fact, this value has an overlap between osteoporosis and normal people, as factors like age, gender, and height have an impact on it (Lee et al., 2008). As a result, accurate diagnosis cannot be usually achieved by such a single feature. Similar cases also exist in other features. Hence, these features should be combined. But different features may yield different results, and their importance in osteoporosis diagnosis may not be the same. How to analyze these results from different features comprehensively is the key point. However, in traditional methods, there is no effective way to do such work.

To solve such problems, machine learning was introduced in some researches (Gao et al., 2010; Gregory et al., 1999; Jennane et al., 2010; Rae et al., 1999; Wang et al., 1998). Through enough training, it can produce desirable results. In earlier researches, some machine-learning methods were used for osteoporosis diagnosis, yielding good results. Jennane et al. (2010) applied genetic algorithm on 3D images of TB for osteoporosis diagnosis, which obtained a high accuracy in recognizing osteoporosis samples. Artificial neural network (ANN) combining image analysis or other methods also has been extensively applied to predict osteoporosis risk (Gregory et al., 1999; Rae et al., 1999; Wang et al., 1998). Gregory et al. (1999) applied Fourier transforms to TB's images from osteoporosis group and normal group and extracted some features, with which neural network was used to classify the two groups. The accuracy was about 77–84%. Rae et al. (1999) used ANN to classify a group of 274 women consisting of osteoporosis and normal people, and obtained

the best prediction accuracy of 73.1%. Wang et al. (1998) developed a multiversion system based on neural networks, which was able to achieve 80% correct classification. Another machine-learning method, C4.5 decision tree, is also frequently used. In Gao's (2010) research, micro-CT images of TB from 14 patients with osteoporosis and from 14 normal people were used. He used C4.5 to classify the two groups with an accuracy of 92.9%. In the previous work (Li and Xu, 2012), the authors have studied some topological features, with which SVM was used to classify the osteoporosis and nonosteoporosis group. In this research, more features are introduced and SVM, a supervised machine-learning method, is used to create classifiers with such features for osteoporosis diagnosis.

Our method originally brought in new features extracted by VTA, which enabled perfect description of topological characters of TB. High-resolution 3D micro-CT images in this research made more accurate analysis of TB possible. Compared to previous works, we also combined features together from which we are able to establish classifiers with more satisfying performances.

In our experiment, features from two groups are used. The first group is calculated by the software Micro-CT skyscan 1076, including the most frequently-used features, like BV/TV, BS/BV, and SMI. The second group is obtained by volumetric topological analysis, a novel topological way to analyze TB structure, which can map the TB structure to the continuum of perfect rods and perfect plates (Saha et al., 2010). These features are curve percentage, surface percentage, surface edge percentage, and average shape distance transform value, which are able to describe TB's local topological characters.

All the features are extracted from 3D micro-CT images. The high resolution of micro-CT largely improves the image resolution for more accurate analysis. Hence, the features obtained from these images can deeply reflect the inner structure of TB.

According to the analysis above, features should be combined together for more accurate analysis for osteoporosis diagnosis. To find an effective method to combine these features, SVM is used to train these data and find the best classifier. To evaluate the performance of the obtained classifiers, precision, recall, and F-measure are used.

This article is organized as follows: Materials and Methods are described in the next section followed by the Experiment Results and Discussion sections. The conclusions are made in the last section.

MATERIALS AND METHODS

Materials

In our experiments, 200 micro-CT images are used, each of which has a size of 337 × 360 × 30. These images are obtained by scanning distal tibia specimens with micro-CT skyscan 1076. With its high resolution, TB structure can be clearly seen. In these 200 images, 100 are from osteoporosis patients and the other 100 are from normal people. In our experiment, age, gender, etc. are not considered. All features used in our experiment are calculated from these images by 3D image analysis.

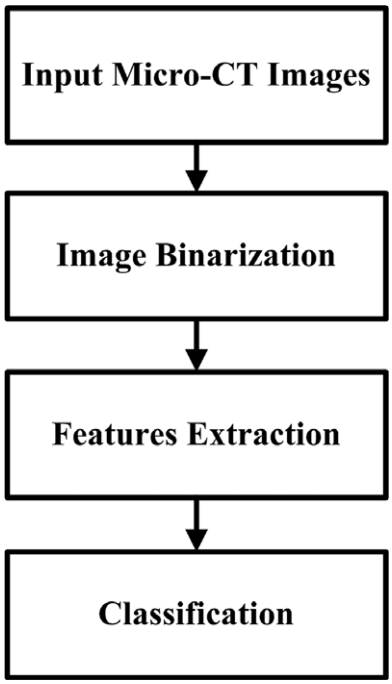

Fig. 2. Method flow chart. a: original images. b: binary images.

In order to create a dataset with high accuracy and completeness, the ground truth classification is generated by three researchers. Two researchers (R1 and R2) with the background of biomedical engineering make the first round of classification in which they classify images individually. The results of R1 and R2 are sent to a third researcher (R3) who is a doctor. R3 checks these classifications and if there is disagreement, a discussion will be held by the three researchers together to make a final decision.

Methods Overview

Our method mainly includes the following three steps: image preprocessing, features extraction, and classification. In image preprocessing, image data are binarized to remove nonbone tissues. In feature analysis, we will discuss their characters and functions, as well as how to calculate them. The last part is about how to create classifiers with such features. The brief introduction of support vector machine is also included. The flow chart of the method is shown in Figure 2.

Image Preprocessing

Before processing these image data, preprocessing is necessary. The original images not only contain TB, but also contain marrow and some other soft tissues. As these soft tissues may negatively influence the following analysis, they should be removed before being processed. As soft tissues usually have a lower gray level than TB, we set a threshold which is assigned a value of 30 to binarize the images. With this threshold, most of the soft tissues can be removed. Binary images also simplify the following topological analysis. Figures 3a and b show the contrast between the original image and the binary image.

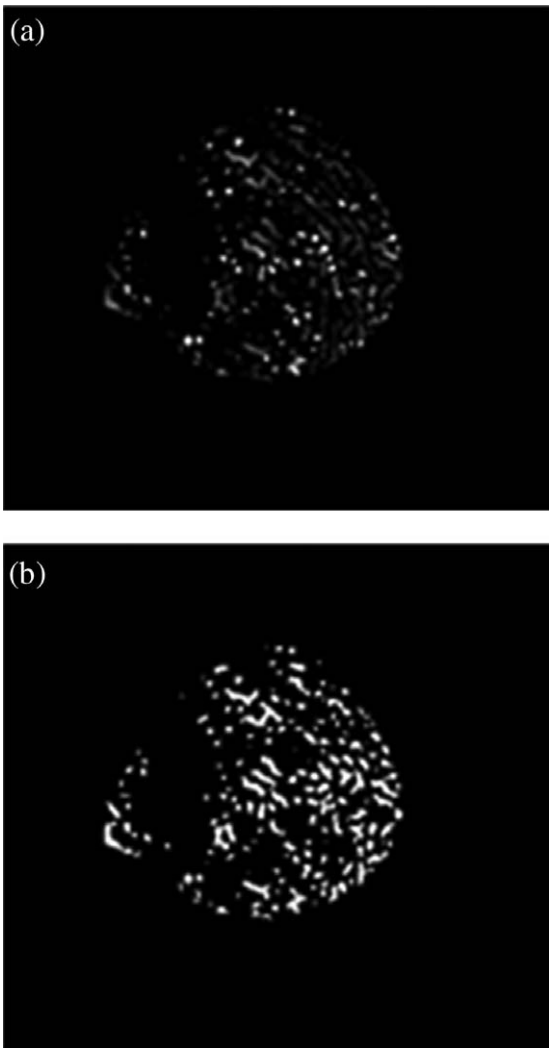

Fig. 3. Contrast between the original image and the binary image.

Features Extraction

The features used in our experiments include general features and topological features. The first group is obtained from the software of micro-CT skyscan 1076. The other group is calculated from volumetric topological analysis, a new topological analysis method, which can map the TB structure to the continuum on perfect plates and perfect rods. With these features, the spatial structure of TB can be clearly quantized.

General Features

The general features are all based on 3D image analysis and contain much structure information of TB. These features include bone volume/total volume (BV/TV) (used to calculate the BMD), bone surface/bone volume (BS/BV), trabecular number (Tb.N), trabecular separation (Tb.Sp), trabecular thickness (Tb.Th), structure model index (SMI), trabecular patter factor (Tb.Pf), and fracture dimension (FD). Other relevant features, like text analysis, Euler number, and connection density, are not included in this research.

BMD Analysis. BMD is an important feature to evaluate the TB condition. It has a close relationship with bone strength. BMD reduction will lead to fracture risk. In traditional methods, its value is usually measured by dual energy X-ray absorptiometry (DEXA), quantitative computerized technology (QCT), and single X-ray absorptiometry (SRA), etc. (Blake and Fogelman, 2007; Chiu et al., 2006; Grados et al., 2009; Hong et al., 2008). In our research, as BMD has a positive correlation with BV/TV, BT/TV is used as the substitute value for it (Jennane et al., 2009). This way is totally based on image analysis without extra labor. The formulation is shown below (Li and Xu, 2012). White pixels are the bone pixels in the binary images. X, Y, and Z are the width, height, and slice number of the micro-CT images. In our experiment, they are 337,360 and 30, respectively. WP stands for numbers of white pixels.

$$\mathrm{BMD}=\frac{\mathrm{BV}}{\mathrm{TV}}=\frac{\mathrm{WP}}{X\cdot Y\cdot Z} \tag{1}$$

As is mentioned above, TB is a complex network consisting of plates and rods. Thus, a feature that is able to describe how TB is like rods or plates can be useful. SMI is one of such features, which defines the perfect plates as 0 and the perfect rods as 3 (Lespessialles et al., 2006). Usually, the plate-like TB has a SMI value of 0–1, and rod-like TB has a value of 2–3 (Wehrli et al., 2001). The formulation for SMI is shown below, in which BV is the bone volume, BS is the bone surface and r is the radius. What is more, another important character of TB is its complex texture, like roughness and self-similarity degree which can be described by fractal dimension (FD) (Lemineur et al., 2007). In osteoporosis, due to the erosion of TB, the texture will change; thus, FD can reflect this variance. In addition, TB has many convex and curve parts, and Tb.Pf is used to represent the ratio of the convex to curve surfaces (Hahn et al., 1992), which is also included in this research.

$$\mathrm{SMI}=6\times\frac{\mathrm{BV}\cdot\frac{\mathrm{dBS}}{\mathrm{d}r}}{\mathrm{BS}^2}, \tag{2}$$

where $\frac{\mathrm{dBS}}{\mathrm{d}r}$ is the increment of bone surface area/radius increment.

Topological Features

Besides the features obtained from the software of Micro-CT, four other features gotten from volumetric topological analysis provide considerable detailed information about the TB's structure (Saha et al., 2010). TB is a complex network of plates and rods. However, in the traditional 3D analysis, the different parts of TB are not taken into consideration. In this article, we use volumetric topological analysis to distinguish TB's different parts, and map every region to the continuum of plates and rods. Hence, this kind of analysis can deeply expose intrinsical structure characters of TB.

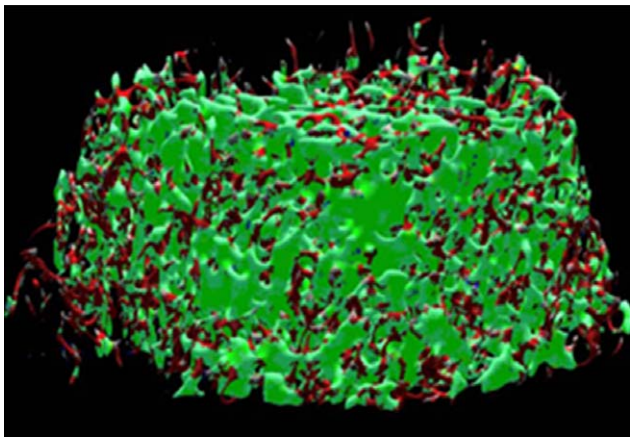

Fig. 4. The major three parts of TB: curve (red), surface (green), surface edge (light green). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

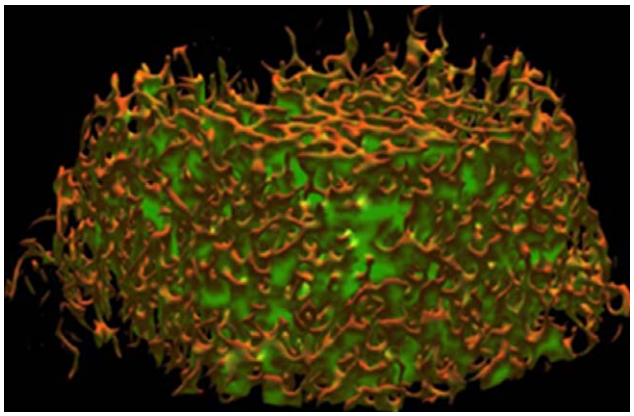

Fig. 5. Osteoporosis TB. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Curve, Surface and Surface Edge Percentage

Based on the above analysis, we first use digital topological analysis (one stage of volumetric topological analysis) to divide the TB majorly into three parts: curve, surface, and surface edge. Curve is the slender part of TB, surface is the wider part, and surface edge is the edge between surfaces. Figure 4 shows the three parts in which curve, surface, and face edge is colored in red, green, and light green, respectively (Li and Xu, 2012). As the main part of TB, the changes in their certain characters can be used as good indicators for osteoporosis detection. When osteoporosis occurs, TB is gradually eroded. As a result, its structure inevitably changes. The wider parts will become slender due to erosion, so the surfaces tend to be curve. Hence, the percentage of surfaces and surface edge decrease; meanwhile the percentage of curve increases. Therefore, the values of the three parts' percentages can be powerful features to diagnose osteoporosis. The calculation of the percentage is based on the digital topological analysis. The formulation is shown below (Li et al., 2012). The numerator is the number of their pixel, and the denominator is the total pixel number of the bone. Curve percentage (CP), surface percentage (SP), and surface edge percentage (SEP) are obtained from the equations below, where BN, CN, SN, and SEN stand for bone pixels number, curve pixels number, surface pixels number and surface edge pixels number, respectively.

$$\mathrm{CP}=\frac{\mathrm{CN}}{\mathrm{BN}}$$

$$\mathrm{SP}=\frac{\mathrm{SN}}{\mathrm{BN}}$$

$$\mathrm{SEP}=\frac{\mathrm{SEN}}{\mathrm{BN}} \tag{3}$$

Average Shape Distance Transform Value

Another powerful feature is the average shape distance transform value. Shape distance transform (SDT) is a new kind of distance transform which was proposed in (Saha et al., 2010). The structure of TB consists of complex plates and rods. However, in fact, there is no absolute standard to distinguish the rods and plates because the local structure of TB changes from plates to rods gradually. Hence, it may be difficult to distinguish the slender plates and rods. Volumetric topological analysis takes into consideration the continuous distributions of TB's structure between the perfect rods and perfect plates, and tries to map all of the local structure of TB to the continuum of rods and plates. SDT value can evaluate how the local structure resembles rods and plates. In Figures 5 and 6, TB from osteoporosis patients and normal people is colored in the range of pure red and pure green according to its SDT value (Saha et al., 2010). The pure red is for perfect rods and pure green for perfect plates. As normal TB consists of more plates than rods, it looks greener on the whole, and in contrast, osteoporotic TB looks redder, which can be seen in Figures 5 and 6. SDT value is effective to reflect TB is plate-like or rod-like. In our research, we use the average SDT value as a TB-related feature for osteoporosis diagnosis. In osteoporosis, the TB will be eroded, the curve with a lower SDT value increases and surface with a high SDT value decreases. So, the average SDT value will decrease when osteoporosis occurs. The formulation is below (Li and Xu 2012). N is the total number of the local regions.

$$\bar{\mathrm{SDT}}=\sum_{i=0}^{N}\mathrm{SDT}(i)/N \tag{4}$$

Classification

After obtaining the features, the next important work is to create the classifiers to classify the normal and abnormal TB. In many machine-learning methods, like artificial neural network and decision tree, we choose SVM and kNN to accomplish this task for their excellent performance on two-class classification. The specific descriptions of SVM and kNN are presented below.

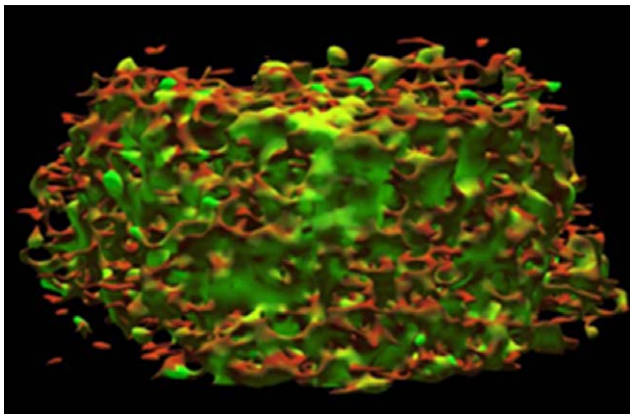

Fig. 6. Normal TB. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Brief Introduction of SVM

Support vector machine is a supervised machine-learning method and has good performance on data classification (Akgundogdu et al., 2010). For a training dataset with two classes, SVM tries to find the best hyperplane to separate each class. In an example for two-class classification, SVM tries to find the linear equation $w \cdot x - b = 0$ to max the margins of the two classes. If the classes cannot be linearly separated, SVM introduces the kernel function to transform the feature space in low dimension to that in high dimension, where the classes can be separated linearly. When the hyperplane cannot completely separate the groups due to some certain mislabeled data, SVM uses the penalty factors C and slack variables ζ_i to measure the degree of deviation of the data from its belonged group and find a relatively optimized hyperplane (Ordonez et al., 2009). We use X to denote the vector of features from the training set. X can belong to the negative class or positive class. y_i is the class label. 1 represents the positive class and -1 for negative class. w is the desired hyperplane which is also a vector. b is an unknown constant. The goal of SVM is to find the best w and b to maximize the margin of the two classes. This is a quadratic programming optimization problem, which can be seen in formulation 5. More details about SVM can be seen in (Christopher and Burges, 1998).

$$\min_{w,\xi,b} \left\{ \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{n} \xi_i \right\}$$

Subject to:

$$y_i (w \cdot x_i - b) \geq 1 - \xi_i, \xi_i \geq 0 \quad (5)$$

Brief Introduction of *k*NN

*k*NN algorithm is a classification method which classifies objects based on their nearest training examples (neighbors). This method is simple but effective. The neighbors of the object to be classified are taken from a training dataset of which the labels are already given. A user defined instance k is used to indicate how many neighbors of the object are retrieved when it is assigned to a class. Majority voting among the object's neighborhood is usually used to decide which class the object is assigned to (Guo et al., 2003; Zhang and Zhou, 2007).

Classifier Generation

In the above statement, we have described how to obtain these power features used to diagnose osteoporosis. In this section, we will state how to use these features to create a classifier. By micro-CT and volumetric topological analysis, we get such features from 200 micro-CT image data, 100 of which are from normal people and 100 are from people with osteoporosis. In our experiment, we use SVM as well as *k*NN to create the classifiers, and radial function is used as kernel function for SVM. Among the 12 features, the first eight features are obtained from the software of micro-CT, and the other four features are calculated from a program used for volumetric topological analysis, which runs on Linux. For better comparison of the performance of every feature, each one is used to create a classifier and tested individually, and three combinations of the 12 features are also used to create classifiers. A 10-fold cross validation is used in the experiment. The 200 image data is divided into 10 groups, and each group consists of 10 image data from osteoporosis group and 10 from nonosteoporosis group. Cross-validation is a method used to assess the generalization ability of a statistical analysis, which is also called rotation estimation (Devijver and Kittler 1982; Kohavi 1995). In many kinds of cross validation, K-fold cross validation is commonly used. In K-fold cross validation, the dataset is divided equally into K subsets, each of which contains roughly the same proportions of the two-group. Each time, the K-1 subsets are used for training and the rest ones for test. This process repeats K times, in which every subset is used for test one time. Each process will obtain a result and the final result is the average of the K results. This method eliminates the bias of random selection of training data and test data. Generally, K is chosen as 10 (McLachlan et al., 2004). Classifier generation is accomplished by LIBSVM software package due to its easy access and efficiency which runs on Linux in our experiment. (http://www.csie.ntu.edu.tw/~cjlin/libsvm/). Experiments were conducted on a computer with 2.13 GHz Core2. Both classifiers spent less than 3 s to train and test in cross validation.

RESULTS

Result is obtained by the method above. The performance of the classifiers is evaluated by precision, recall, and F-measure. We use TP and FN to denote the number of correct and false classifications of the osteoporosis group respectively, and TN and FP to denote the number of correct and false classifications of the normal group. The formulation 6 gives the specific definition for the three evaluators. The result is shown in Tables 1 and 2. Figures 7 and 8 show the F-measures of the two classifiers. $1 \rightarrow 4$, $1 \rightarrow 8$, and $1 \rightarrow 12$ represent the combinations of the first 4, 8, and 12 features, respectively.

TABLE 1. Performance of classifiers using individual feature

Feature	SVM			*k*NN		
	Precision (%)	Recall (%)	F-measure (%)	Precision (%)	Recall (%)	F-measure (%)
1.BV/TV	74.5	91.6	82.2	75.8	91.5	82.9
2.Tb.N	67.0	79.4	73.4	65.8	77.9	71.3
3.SMI	68.5	64.1	66.2	68.0	62.8	65.3
4.Tb.Pf	90.2	73.7	81.1	88.6	72.5	79.7
5.BS/BV	71.3	74.5	72.9	77.4	62.3	69.0
6.Tb.Sp	77.6	73.5	75.5	80.3	65.5	72.1
7.Tb.Th	82.2	92.3	87.0	85.6	82.4	84.0
8.Fd	89.7	83.1	86.3	81.7	78.8	82.7
9.Curve	79.3	80.9	80.1	82.4	73.7	77.9
10.SurfaceEdge	77.3	73.6	75.4	76.5	67.0	71.3
11.Surface	81.5	90.6	85.8	80.7	88.6	84.5
12.Average SDT	90.2	91.4	90.8	87.2	90.3	88.7

TABLE .2. Performance of classifiers using feature combinations

Feature Combinations	SVM			*k*NN		
	Precision (%)	Recall (%)	F-measure (%)	Precision (%)	Recall (%)	F-measure (%)
1→4	91.3	90.5	90.9	92.4	87.8	90.0
1→8	91.9	100	95.8	95.3	93	94.1
1→12	100	100	100	100	100	100

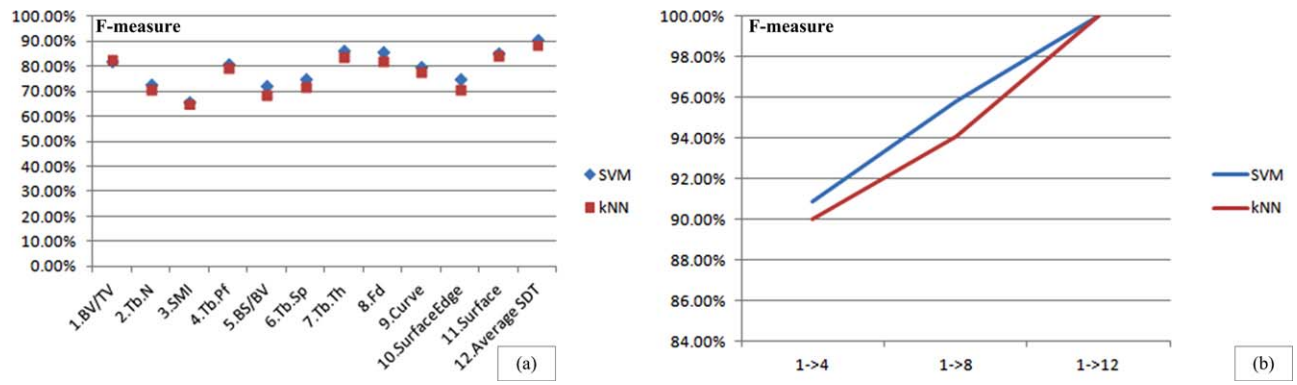

Fig. 7. F-measures of different classifiers. a: displays comparison of F-measures of different classifiers using individual features. b: displays comparison of F-measures of different classifiers using features combinations. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

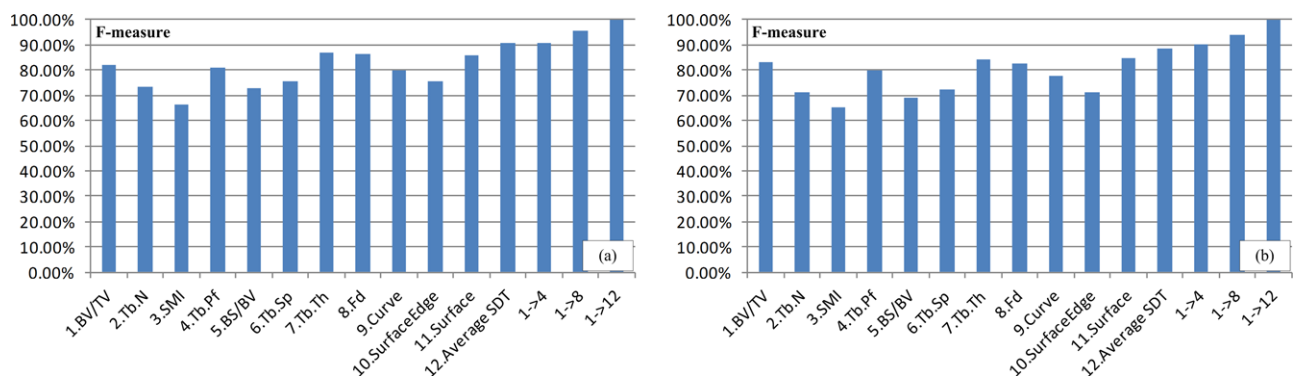

Fig. 8. F-measures of different classifiers. a: displays F-measures of classifier based on SVM. b: displays F-measures of classifier based on *k*NN. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

$$\text{precision} = \frac{\text{TP}}{\text{TP}+\text{FP}}$$

$$\text{recall} = \frac{\text{TP}}{\text{TP}+\text{FN}}$$

$$F-\text{measure} = \frac{2\times\text{precision}\cdot\text{recall}}{\text{precision}+\text{recall}} \quad (6)$$

From the result of the experiment, we are able to get the following conclusions. First, in the individual features, SMI and BS/BV have a relatively weak performance, and Tb.Th, surface percentage, and Average SDT values have a relatively good performance. Second, the combination of feature 1 to feature 4 has the best performance than any one individual in the four features. It can be seen from the performances of the three combinations that the performance becomes better with more features involved. Third, the best one is obtained when all features are included with precision of 100%, recall of 100%, and F-measure of 100%. To make the result more robust, more data numbers should be included in the future work. From the tendency of the distributions of the data, it may be safe to say that SVM has strong power in classification. The result also confirms the supposition that the combinations of features will have a better performance than individual ones because different features can be complementary to other ones. By comparison of the performances of the two classifiers, it can be concluded that SVM performs better in this task in general as more evident improvement is acquired by employing features.

CONCLUSION AND DISCUSSION

In this article, we propose a new method for osteoporosis diagnosis based on micro-CT images and machine learning. Many of the former researches have been done focusing on individual features, which, however, lack global optimization. In our research, we try to combine these features together to obtain a global optimized result. We use SVM and *k*NN to create the classifiers because they have good performances on classification. To improve the performance of the classifier, the feature selection is very important. In our experiment, we use some frequently-used and effective features obtained from the software of micro-CT, and include four topological features obtained from a new kind of topological analysis. Each feature and three feature combinations are used to create classifiers, respectively. From the results, we can see that average SDT value has the best performance compared with other individual features, and the best result is obtained when combining all features together with precision of 100%, recall of 100%, and F-measure of 100%. This result proves that by combining features together including new features obtained from VTA, we are able to generate more comprehensive information of TB from micro-CT images, which enables more accurate classification than using features independently.

Though the two classifiers based on SVM model and *k*NN model both obtained satisfying performances, SVM proved to be more effective for our task. SVM model outperforms *k*NN model as it enables features to work better from the experimental results.

Even though the result is satisfying, it should be admitted that more data should be used in future work to make the result more robust. What is more, some nonbone factors like age and gender are not considered, which may have an influence on the final results. For example, the BV/TV may have a higher average value in the young than in the old. Third, in our experiment, we only use SVM and *k*NN to create classifiers. Many other machine-learning methods, which are not included in this research, can also accomplish such task. In future work, more methods may be used and compared with each other to find the best ones for specific purpose.

In addition, besides the 12 features included in our experiment, some other effective features may exist that we did not introduce in, such as texture features which are used in Ghalila et al. (2004) and Valentinitsch et al. (2010). And also, some features in our experiment may not be sufficiently strong or universal. The optimization of the feature combinations will be studied in future work. Moreover, if the proposed method can be used with clinical CT images, the value of this article would be further enhanced.

In this article, we propose a new method for osteoporosis diagnosis. We use SVM and *k*NN to train our data to obtain classifiers. Compared with the traditional methods, this method considers the inner relationship of these features and gives out a global optimized result. Twelve features related to the TB structure are used. In this research, a new kind of 3D analysis method volumetric topological analysis proposed recently is used to get four topological features. The best results are obtained when combining all features together with precision of 100%, recall of 100%, and F-measure of 100%. This result is very satisfying, convincing us that SVM and *k*NN combining micro-CT images have strong potentials for osteoporosis diagnosis. By using this method, a well-trained classifier can give an accurate decision by only inputting image data, which is very convenient and noninvasive. Thus, osteoporosis based on machine learning is very promising.

REFERENCES

Akgundogdu A, Jennane R, Aufort G, Benhamou CL, Ucan ON. 2010. 3D image analysis and artificial intelligent for bone disease classification. J Med Syst 34:815–828.

Baum T, Gamio JC, Huber MB, Muller D, Monetti R, Rath C, Eckstein F, Lochmuller EM, Majumdar S, Majumdar S, Rummeny EJ, Link TM, Bauer JS. 2010. Automated 3D trabecular bone structure analysis of the proximal femur-prediction of biomechanical strength by CT and DXA. Osteoporosis Int 21:1553–1564..

Blake GM, Fogelman I. 2007. Role of dual-energy X-ray absorptiometry in the diagnosis and treatment of osteoporosis. J Clin Densitometry 10:102–110.

Cervinka T, Hyttinen J, Sievanen H. 2010. Enhanced bone structural analysis through pQCT image preprocessing. Med Eng Phys 32:398–406.

Chiu JS, Li YC, Yu FC, Wang YF. 2006. Applying an artificial neural network to predict osteoporosis in the elderly. In: A. Hasman et al., editors, Ubiquity: Technologies for better health in aging societies. pp 609–614.

Christopher JC, Burges. 1998. A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov 2:121–167.

Devijver PA, Kittler J. 1982. Pattern recognition: A statistical approach. London: Prentice-Hall.
Gamio JC, Phan C, Link TM, Majumdar S. 2006. Characterization of trabecular bone structure from high-resolution magnetic resonance images using fuzzy logic. Magn Resonan Imaging 24:1023–1029.
Gao Z, Hong WX, Xu YH, Zhang T, Song Z, Liu J. 2010. Osteoporosis diagnosis based on the multifractal spectrum features of micro-CT images and C4.5 decision tree. 2010. In: First International Conference on Pervasive Computing, Signal Processing and Applications, pp 1043–1047.
Ghalila SS, Benyahia AB, Ricordeau A, Mellouli N, Chappard C, Laurent B. 2004. Texture image analysis for osteoporosis detection with morphological tools. In: Medical Image Computing and Computer Assisted Intervention-MICCAI 2004. Barillot C, Haynor DR, Hellier P, editors. Springer, Berlin Heidelberg, Berlin, Vol. 3216. pp 87–94.
Grados F, Fechtenbaum J, Flipon E, Kolta S, Roux C, Fardellone P. 2009. Radiographic methods for evaluating osteoporotic vertebral fractures. Joint Bone Spine 76:241–247.
Gregory JS, Junold RM, Undrill PE, Aspden RM. 1999. Analysis of trabecular bone structure using Fourier transforms and neural networks. IEEE Trans Inf Technol Biomed 3:289–294.
Guo G, Wang H, Bell D, Bi Y, Greer K. 2003. KNN model-based approach in classification. Lect Notes Comput Sci 2888:986–996.
Hahn M, Vogel M, Kempa MP, Delling G. 1992. Trabecular bone pattern factor-a new parameter for simple quantification of bone microarchitecture. Bone 13:327–330.
He JC, Leow WK, Howe TS. 2007. Hierarchical classifiers for detection of fractures in X-Ray images. In: CAIP'07 Proceedings of the International Conference on Computer Analysis of Images and Patterns. Berlin, Heidelberg: Springer-Verlag. pp 962–969.
Hong CM, Lin CT, Huang CY, Lin YM. 2008. An intelligent fuzzy-neural diagnosis system for osteoporosis risk assessment. World Acad Sci Eng Technol 42:597–602.
Jennane R, Aufort G, Lheritier H, Benhamou CL. 2009. Multivariable statistical classification of 3D bone microarchitecture using morphological and mechanical features. In: 17th European Signal Processing Conference, Glasgow, Scotland, 2009. Hindawi Publishing Corporation, Nasr City. pp 1557–1561.
Jennane R, Imjabber AA, Hambli R, Ucan ON, Benhamou CL. 2010. Genetic algorithm and image processing for osteoporosis diagnosis. In: 32nd Annual International Conference of the IEEE EMBS, 2010, Buenos Aires, Argentina, pp 5597–5600.
Kohavi R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, Montreal, 1995. Morgan Kaufmann Publishers Inc. San Francisco, Vol. 2:1137–1143.
Laib A, Ruegsegger P. 1999. Calibration of trabecular bone structure measurements of in vivo three-dimensional peripheral quantitative computed tomography with 28-microm-resolution microcomputed tomography. Bone 24:35–39.
Lee S, Lee JW, Jeong JW, Yoo DS, Kim S. 2008. A preliminary study on discrimination of osteoporotic fractured group from nonfractured group using support vector machine. In: 30th Annual International IEEE EMBS Conference, Vancouver, British Columbia, Canada, 2008. IEEE Engineering in Medicine and Biology Society, Piscataway, pp 474–477.
Lemineur G, Harba R, Kilic N, Ucan ON, Osman O, Benhamou L. 2007. Efficient estimation of osteoporosis using artificial neural networks. In: The 33rd Annual Conference of the IEEE Industrial Electronics Society, Taipei, Taiwan. 2007. IEEE Industrial Electronics Society, Piscataway, pp 3039–3044.
Lespessialles E, Chappard C, Bonnet N, Benhamou CL. 2006. Imaging techniques for evaluating bone microarchitecture. Joint Bone Spine 73:254–261.
Li DS, Xu Y. 2012. An approach for osteoporosis diagnosis based on support vector machine using Micro-CT images. Medical Physics and Biomedical Engineering, Beijing, China.
McLachlan GJ, Do KA, Ambroise C. 2004. Analyzing microarray gene expression data. Wiley, Hoboken.
Mosekilde L, Ebbesen EN, Tornvig L, Thomsen JS. 2000. Trabecular bone structure and strength-remodelling and repair. J Musculoskelet Neuron Interact 1:25-30.
National Institute of Health. 1984. Consensus conference: osteoporosis. J Am Med Assoc 252:799–802.
Ordonez C, Matias JM, Cos Juez JF, Garcia PJ. 2009. Machine learning techniques applied to the determination of osteoporosis incidence in post-menopausal women. Math Comput Modell 50:673–679.
Parfitt AM, Drezner MK, Glorieux FH, Kanis JA, Malluche H, Meunier PJ, Ott SM. 1987. Bone histomorphometry: Standardization of nomenclature, symbols, and units. J Bone Mineral Res 2:595–610.
Peacock M, Turner CH, Liu G, Manatunga AK, Timmerman L, Johnston CC. 1995. Better discrimination of hip fracture using bone density, geometry and architecture. Osteoporosis Int 5:167–173.
Petrantonaki M, Maris T, Damilakis J. 2005. MRI techniques for the examination of trabecular bone structure. Curr Med Imaging Rev 1:35–41.
Phan CM, Matsuura M, Bauer JS, Dunn TD, Newitt D, Lochmueller EM, Eckstein F, Majumdar S, Link TM. 2006. Trabecular bone structure of the calcaneus: Comparison of MR imaging at 3.0 and 1.5 T with micro-CT as the standard of reference. Radiology 239:488–496.
Pothuaud L, Rietbergen BV, Mosekilde L, Beuf O, Levitz P, Benhamou CL, Majumdar S. 2002. Combination of topological parameters and bone volume fraction better predicts the mechanical properties of trabecular bone. J Biomech 35:1091–1099.
Rae SA, Wang WJ, Partridge D. 1999. Artificial neural networks: a potential role in osteoporosis. J R Soc Med 92: 119–122.
Saha PK, Xu Y, Duan H, Heiner A, Liang GY. 2010. Volumetric topological analysis: A novel approach for trabecular bone classification on the continuum between plates and rods. IEEE Trans Med Imaging 29:1821–1838.
Ulrich D, Rietbergen BV, Laib A, Ruegsegger P. 1999. The ability of three dimensional structure indices to reflect mechanical aspects of trabecular bone. Bone 25:55–60.
Valentinitsch A, Patsch J, Mueller D, Kainberger F, Langs G. 2010. Texture analysis in quantitative osteoporosis assessment: Charactering microarchitecture. In: The IEEE International Symposium on Biomedical Imaging, Rotterdam, 2010. Institute of Electrical and Electronics Engineers, Inc., New York. pp 1361–1364.
Wang WJ, Partridge D, Rae S. 1998. Multi-version systems of neural networks for predicting the risk of osteoporosis. In: Proceedings of the Eleventh International FLAIRS Conference, Sanibel Island, 1998.
Wehrli FW, Gomberg BR, Saha PK, Song HK, Hwang SN, Snyder PJ. 2001. Digital topological analysis of in vivo magnetic resonance microimages of trabecular bone reveals structural implications of osteoporosis. J Bone Min Res 16:1520–1531.
Wehrli FW, Saha PK, Gomberg BR, Song HK, Snyder PJ, Benito M, Wright A, Weening R. 2002. Role of magnetic resonance for assessing structure and function of trabecular bone. Top Magn Reson Imaging 13:335–356.
Zhang ML, Zhou ZH. 2007. ML-KNN: A lazy learning approach to multi-label learning. Pattern Recognit 40:2038–2048.