Histogram of Oriented Gradients and Texture Features for Bone Texture Characterization

Hany M. Harb
Faculty of Engineering, Al-Azhar University, Cairo, Egypt

Abeer S. Desuky
Faculty of Engineering, Al-Azhar University, Cairo, Egypt

Asmaa Mohammed
Faculty of Engineering, Al-Azhar University, Cairo, Egypt

Rachid Jennane
I3MTO Laboratory, University of Orleans, Orleans, France

ABSTRACT

Texture Characterization of Bone radiograph images (TCB) is a challenge in the osteoporosis diagnosis organized for the International Society for Biomedical Imaging. The objective of this paper is to distinguish osteoporotic cases from healthy controls on 2D bone radiograph images, using texture analysis. In this paper, we propose a Bone Texture Characterization method based on texture features (Segmentation-based Fractal Texture Analysis (SFTA), Basic Texture and Gabor filters) and compare these resulted features with HOG features for 2D bone structure evaluation. The classification experiments are tested with linear SVM and decision tree classifiers. The classification performance for HOG features are always higher than other texture features, and show excellent classification performance compared to other existing methods.

General Terms

Image mining, Artificial Intelligence.

Keywords

Texture, HOG, SFTA, Gabor filter, Bone, Osteoporosis, Classification.

1. INTRODUCTION

The texture is one of the significant characteristics used in identifying objects or regions of interest (ROI) of an image.

Osteoporosis is defined as a skeletal disorder characterized by compromised bone strength predisposing to an increased risk of fracture [1]. The most widespread method for osteoporosis diagnosis is to calculate Bone Mineral Density (BMD) by dual-energy X-ray absorptiometry [2]. However, BMD alone performs only 60% of fracture prediction. The characterization of trabecular bone micro architecture has been recognized as an important factor and completes the osteoporosis diagnosis using BMD [3], but it cannot be periodically obtained by noninvasive methods and requires a bone biopsy with histomorphometric analysis. 2D texture analysis displays a simple way to estimate bone structure on conventional radiographs. The evaluation of osteoporotic disease from bone radiograph images makes the main challenge for pattern recognition and medical applications. Textured images from the bone micro architecture of osteoporotic and healthy subjects display a high degree of similarity, thus significantly grow the difficulty of classifying such textures. Figure 1 shows the bone texture similarities of control and osteoporotic images.

In this paper, we use two techniques to extract features, texture features and the histogram of oriented gradients (HOG) features.

In order to deal with classification qualities, we suggest new features for these bone textures. The histogram of oriented gradients (HOG) is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique computes occurrences of gradient orientation in localized portions of an image. This method is similar to that of edge orientation histograms, scale-invariant feature transform descriptors, and shape contexts, but differs in that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy [4].

Navneet Dalal and Bill Triggs, researchers for the French National Institute for Research in Computer Science and Automation (INRIA), first described HOG descriptors at the 2005 Conference on Computer Vision and Pattern Recognition (CVPR). In their work, they focused on pedestrian detection in static images, although since then they extended their tests to include human detection in videos, as well as to a variety of common animals and vehicles in static imagery.

The essential thought behind the histogram of oriented gradients descriptor is that local object appearance and shape within an image can be characterized by the distribution of intensity gradients or edge directions. The image is partitioned into small connected regions called cells, and for the pixels within each cell, a histogram of gradient directions is compiled. The descriptor is the sequences of these histograms. For improved accuracy, the local histograms can be contrast-normalized by calculating a measure of the intensity across a larger region of the image, called a block, and then using this value to normalize all cells within the block. This normalization results in preferable invariance to modification in illumination and shadowing.

The HOG descriptor has a few key advantages over other descriptors [5]. Since it works on local cells, it is invariant to geometric and photometric transformations, except for object orientation. Such changes would only show in larger spatial regions. Moreover, as Dalal and Triggs discovered, coarse spatial sampling, fine orientation sampling, and strong local photometric normalization authorizes the individual body movement of pedestrians to be ignored so long as they maintain a roughly upright position.

2D Texture takes the important part in pattern recognition. Textures can be considered as patterns in which the statistics such as mean, standard deviation, angular second moment, contrast, correlation, variance (sum of squares), inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, information measures of correlation coefficient, and others can be used for characterization.

Recently 2D texture analysis using X-ray imaging has shown its potential in providing a cost-effective and efficient way to detect and evaluate osteoporosis [6]. The current research thus
focuses on designing and evaluating various feature descriptors for characterizing the bone textures [7,8,9,10].

Salmi et al. [11] developed a texture analysis method for the trabecular bone X-ray images. The main goal of this project was to study the effect of preprocessing the data of bone radiograph images for the diagnosis of osteoporosis. In the preprocessing step, they enhanced the image by using anisotropic morlet wavelet. The exploitation of the fully anisotropic morlet enabled solving the problem of orientation which is caused by the non-uniform changes. In the third step, the Renyi entropy was used for the anisotropic description of the bone textures.

Materka [12] made an attempt to apply the digital image analysis technique for the detection of bone mass and its structure. Here the distal forearm bones were investigated. They included a calibration phantom to improve the image intensity. They extracted first order texture parameters and fractal dimensions were evaluated. These derived texture features were correlated with the bone mineral density by using DEXA (Dual Energy X-ray Absorptiometry). In the methodology, they used image preprocessing to remove the noise as well as to extract the region of interest (ROI). Results obtained showed that by measuring the changes in statistical texture parameters and fractal dimensions of X-ray images it is possible to monitor changes in calcium contents and internal structure of the bone. Texture analysis showed potential usefulness in the diagnosis of skeletal diseases. This initial research was carried out by using first-order texture features only.

Yger [13] proposed a new method for texture analysis based on covariance matrices and wavelet marginal. In their work, they focused mainly on covariance matrices and wavelet marginal rather than the complicated features. Covariance matrices have been studied as image descriptor in the wide variety of applications from license plate detection to pedestrian detection, but in this case, samples were taken are more compared to the parameters (features) and they are more sensitive to its outliers. In order to overcome this issue, they have used minimum covariance determinant and aims in giving the lower determinant for their experiment, they have used two variants of features they are, gradient based and Gabor-based and these covariance matrices belongs to a non-Euclidean space where distances are not computed on straight line rather they computed on curves.

2. DATA

Dataset consists of 58 control, 58 osteoporotic and 58 blind images retrieved from worldwide challenge IEEE-ISBI 2014; texture characterization [14]. Each image is of 400 x 400 pixels showing the region of bone textures only. We have 116 instances and applied four methods: HOG algorithm, SFTA algorithm, Gabor filter, basic texture.

3. METHOD

3.1 Texture Features

For extracting texture features from the images, three texture classification methods are used:

SFTA, Basic texture, and Gabor texture filters.

First: SFTA as local features (Segmentation-based Fractal Texture Analysis) algorithm. the authors in [15], proposed the SFTA algorithm for texture classification. The algorithm can be divided into two steps: input gray scale image is segmented to set of binary images based on Otsu method; fractal features are extracted for every binary image. More particular, the Two-Threshold Binary Decomposition (TTBD) estimates a set of thresholds automatically which is based on multi-level Otsu algorithm. From these estimated thresholds called, (lower thresholds), decomposition step in gray scale image I (x, y) is executed to generate a set of binary images through two threshold methods.

\[ I_b(x, y) = \begin{cases} 1 & t_1 < I(x, y) \leq t_u \\ 0 & \text{otherwise} \end{cases} \]

where \( t_1 \) and \( t_u \) are adjacent lower and upper thresholds.

For each generated binary image, the SFTA feature vector is constructed with 3 components: size of the binary image (number of foreground pixels), mean gray level, and boundaries’ fractal dimension (using the Box Counting Algorithm).

The binary images segmented from SFTA algorithm for control and osteoporotic (from Fig. 1) also show the high similarities, so the features such as fractal dimension, mean gray level, and size of binary cannot distinguish the two populations.

Fig. 1: The similarity of control and osteoporotic images.

Second: Gabor filter. Gabor filters are band pass filters which are used in image processing for feature extraction, texture analysis [16], and stereo disparity estimation [17-23]. The impulse response of these filters is created by multiplying a Gaussian envelope function with a complex oscillation. Gabor [24] showed that these elementary functions minimize the space (time)-uncertainty product. By extending these functions to two dimensions it is possible to create filters which are selective for orientation [25].

Third: Basic texture feature, Basic Image Features [26] are defined by a partition of the filter-response space (jet space) of a set of six Gaussian derivative filters. These filters provide an uncommitted front-end to describe an image locality fully up to second order at some scale.

Finally: HOG features as global features.

The goal of this paper is to provide an enhancement method of how texture information can be used to classify images from the two populations. In the proposed method 116 2D radiographic images have been used. In that 58 are normal subject images and another 58 are images of a patient with osteoporotic fractures. Basically, a texture classifier will be learned from a set of labeled images depicting textures. Then, the learned classifier will be used to provide a class label for an unlabeled image.

Texture classification can be divided into three phases which are discussed in the following:

I. Extracting texture features;
II. Training a classifier;
III. Classification of an unlabeled texture image.
IV. Selecting optimal features to enhance the performance of classification.

In this work, a 10-fold Cross-validation was used.

Cross-validation is a technique to evaluate predictive models by partitioning the original sample into a training set to train the model, and a test set to evaluate it.

3.2 HOG Features

Apply classification on HOG features as global features.

First: features by using HOG algorithm is extracted. Two algorithms are applied on 116 instances.

First: HOG features that extracted 6887 features(attributes).

Second: the data is filtered, firstly instances are filtered to remove un-useful instances (this is done by using Resample), and attributes(features) are filtered by using Discretize.

Third: the more effective optimal set of features (attributes) are selected by using CfsSubsetEval as Attribute Evaluator and BestFirst as a search method.

Finally: two different classifier methods SVM and J 48 are used.

![Figure (2) Image Mining Process](image)

The steps of image mining are illustrated in figure 2, we applied these steps in our method, as we discussed above.

3.3 Support Vector Machine (SVM)

Support Vector Machine (SVM), is one of best machine learning algorithms, which was proposed in 1990’s and used mostly for pattern recognition. This has also been applied to many pattern classification problems such as image recognition, speech recognition, text categorization, face detection and faulty card detection, etc. Pattern recognition aims to classify data based on either a priori knowledge or statistical information extracted from raw data [27], which is a powerful tool in data separation in many disciplines. SVM is a supervised type of machine learning algorithm in which, given a set of training examples, each marked as belonging to one of the many categories, an SVM training algorithm builds a model that predicts the category of the new example. SVM has the greater ability to generalize the problem, which is the goal in statistical learning. Support Vector Machine (SVM) is a learning classification algorithm that learns from a training data set and attempt to generalize and make accurate predictions on new data sets. It is used for classification problems like binary classification [28].

Table (1) - TP Rate

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Extracted Features</th>
<th>Original Features</th>
<th>After select attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>SFTA</td>
<td>0.48</td>
<td>0.80</td>
</tr>
<tr>
<td>Basic Texture</td>
<td>0.53</td>
<td></td>
<td>0.57</td>
</tr>
<tr>
<td>Gabor</td>
<td>0.49</td>
<td>0.49</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Table (1) represents results TP – It stands for true positive: The number of subjects with Osteoporosis that are correctly identified.

Where, features extracted from images using SFTA features and Instance filtered by Resample and then attributes selected by using (principle component & Ranker) and images

4. EXPERIMENT RESULTS

In literature, different performance measures have been proposed to evaluate the learning models. Between them the most popular performance measures are following:

1) Sensitivity, 2) Specificity and 3) Accuracy.

Sensitivity (True positive fraction/recall) is the proportion of actual positives which are predicted positive. [29] Mathematically, Sensitivity can be defined as

\[
\text{Sensitivity} = \frac{TP}{TP+FN}
\]  

(2)

Specificity (True negative fraction) is the proportion of actual negatives which are predicted negative. [27] It can be defined as

\[
\text{Specificity} = \frac{TN}{TN+FP}
\]  

(3)  

Accuracy is the probability to correctly identify individuals. i.e. it is the proportion of true results, either true positive or true negative [29]. It is computed as

\[
\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}
\]  

(4)  

In general, sensitivity points out, how well model characterizes positive cases and specificity computes how well it identifies the negative cases. While accuracy is predicted to measure how well it characterizes both categories. Therefore, if both sensitivity and specificity are high (low), accuracy will be high (low). But, if any one of the measures, sensitivity or specificity is high and other is low, then accuracy will be prejudiced towards one of them. For this reason, accuracy single cannot be a good performance measure.
classified by two methods (SVM, Decision Tree). These results show SVM algorithm achieves the best accuracy.

Table (2) - FP Rate

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Extracted Features</th>
<th>Original features</th>
<th>After select attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>SFTA</td>
<td>0.52</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Basic Texture</td>
<td>0.47</td>
<td>0.43</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>SFTA</td>
<td>0.51</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>Basic Texture</td>
<td>0.54</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>Gabor</td>
<td>0.45</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Table (2) represents results FP – It stands for false positive: The number of Control subjects which are incorrectly identified. These results show SVM with SFTA algorithm achieves the best accuracy. This method achieves the least false identification.

Table (3) - precision

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Extracted Features</th>
<th>Original features</th>
<th>After select attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>SFTA</td>
<td>0.48</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>Basic Texture</td>
<td>0.54</td>
<td>0.57</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>SFTA</td>
<td>0.49</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>Basic Texture</td>
<td>0.55</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>Gabor</td>
<td>0.57</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Table (3) represents results precision (also called positive predictive value) is the fraction of retrieved instances that are relevant. It is representing the accuracy rate.

\[
\text{Precision} = \frac{TP}{TP+FP} \quad (5)
\]

Table (4) - Recall

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Extracted Features</th>
<th>Original features</th>
<th>After select attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>SFTA</td>
<td>0.48</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>Basic Texture</td>
<td>0.53</td>
<td>0.57</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>SFTA</td>
<td>0.49</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>Basic Texture</td>
<td>0.54</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>Gabor</td>
<td>0.56</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Table (4) represents results Recall (also known as sensitivity) is the fraction of relevant instances that are retrieved.

\[
\text{Recall} = \frac{TP}{TP+FN} \quad (6)
\]

Table (5) - F-Measure

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Extracted Features</th>
<th>Original features</th>
<th>After select attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>SFTA</td>
<td>0.48</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>Basic Texture</td>
<td>0.53</td>
<td>0.57</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>SFTA</td>
<td>0.49</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Table (5) represents results F – Measure A measure that combines precision and recall is the harmonic mean of precision and recall.

Table (6) - HOG (confusion matrix)

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Features</th>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>SN</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>Original</td>
<td>20</td>
<td>25</td>
<td>33</td>
<td>38</td>
<td>0.44</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>After select attribute</td>
<td>56</td>
<td>5</td>
<td>53</td>
<td>2</td>
<td>0.92</td>
<td>0.91</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>Original</td>
<td>33</td>
<td>31</td>
<td>27</td>
<td>25</td>
<td>0.52</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>After select attribute</td>
<td>58</td>
<td>7</td>
<td>51</td>
<td>0</td>
<td>0.89</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Tables (6,7) represent results when features extracted from images using HOG features as global features (6887 features “attributes”) and classification applied using two techniques: SVM and Decision Tree.

Preprocessing is one of the important steps in texture analysis, the raw images are not free from the noise, this may cause variation in the statistical measures derived from the 2D radiographic images. In order to make the image free from noise and to enhance the quality of the image, preprocessing is the initial procedure in the development of image analysis algorithm.

In the proposed algorithm, Attribute filter “Discretize”, Instance filter “Resample” has been used. And features are
selected by Attribute Evaluator using (CfsSubsetEval), Search Method by (BestFirst) algorithm. And finally, images are classified with two algorithms (SVM, Decision Tree).

Histograms of Oriented Gradients (HOG) are one of the well-known features for object recognition. HOG features are calculated by taking orientation histograms of edge intensity in a local region.

Principal Component Analysis (PCA) is applied to these HOG feature vectors to obtain the score (PCA-HOG) vectors.

Tables (8,9) represent results Where, features extracted from images using SFTA features (25 features “attributes”) and classification applied using two techniques: (SVM, Decision Tree). Instance filtered by Resampled and then attributes selected by using (principle component & Ranker). These results show LibSVM algorithm and J48 achieve the same accuracy.

Sensitivity, Specificity calculated as follows:

\[ \text{Sn - Sensitivity: defined as } \text{Sn} = \frac{TP}{TP+FN} \quad (7) \]
\[ \text{Sp - Specificity: defined as } \text{Sp} = \frac{TN}{FP+TN} \quad (8) \]

### Table (8) - Confusion matrix

<table>
<thead>
<tr>
<th>Classifier Techniques</th>
<th>Features</th>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>SN</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>Original Features</td>
<td>34</td>
<td>36</td>
<td>22</td>
<td>24</td>
<td>0.49</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>After select attribute</td>
<td>46</td>
<td>11</td>
<td>47</td>
<td>12</td>
<td>0.81</td>
<td>0.81</td>
</tr>
<tr>
<td>Basic Texture</td>
<td>Original Features</td>
<td>24</td>
<td>20</td>
<td>38</td>
<td>34</td>
<td>0.55</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>After select attribute</td>
<td>32</td>
<td>24</td>
<td>34</td>
<td>26</td>
<td>0.57</td>
<td>0.59</td>
</tr>
<tr>
<td>Gabor</td>
<td>Original Features</td>
<td>32</td>
<td>33</td>
<td>25</td>
<td>26</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>After select attribute</td>
<td>42</td>
<td>17</td>
<td>41</td>
<td>16</td>
<td>0.71</td>
<td>0.72</td>
</tr>
</tbody>
</table>

### Table (9) - Confusion matrix

<table>
<thead>
<tr>
<th>Classifier Techniques</th>
<th>Features</th>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>SN</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>J48</td>
<td>Original Features</td>
<td>32</td>
<td>33</td>
<td>25</td>
<td>26</td>
<td>0.49</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Table (10) combining different methods and show that the accuracy of HOG features as Global Features is the best. As represented in section 4 that discussed the different evaluation matrices that measure the accuracy.

Texture features are designed to capture the granularity and repetitive patterns of regions within an image. From a statistical point of view, textures can be seen as complicated pictorial patterns from which sets of statistics can be achieved for characterization purposes [30].

### 5. CONCLUSION

In this paper, we have presented the texture features for bone texture characterization. The SFTA, Basic texture and Gabor are based on Texture Features. HOG Features as global features, this is the first result of applying HOG features for classification of 2D bone texture images. For evaluation, the performance of the HOG method, we compared HOG performance against some popular texture such as Gabor filters, SFTA algorithm, Basic Textures on the dataset used in the worldwide challenge in bone texture characterization. It is observed that the average accuracy is higher than that in the related work [10]. we think that it is the potential direction of using 2D textures to solve this problem. Texture analysis plays a supportive rather than a comprehensive role in the future of medical image interpretation. The robustness of texture analysis makes it particularly attractive for monitoring disease progression or treatment response with time, as demonstrated with Bone Osteoporosis. Support Vector Machines using the Polynomial kernel and RBF kernel play an important role in this application and give satisfactory results.
6. REFERENCES