Statistical Evaluation of Law’s Mask Texture Analysis for Osteoporosis Detection

GAN HONG SENG, HUM YAN CHAI, TAN TIAN SWEE
Department of Biomechanics and Biomedical Materials
Medical Implant Technology Group, Material and Manufacturing Research Alliance
Faculty of Biosciences and Medical Engineering
Universiti Teknologi Malaysia
UTM skudai, 81310 Johor
MALAYSIA
hongseng1008@gmail.com, ychum2@live.utm.my, meditagutm@gmail.com

Abstract: - The ineffectiveness of bone mineral density (BMD) to demonstrate the bone quality has led to steady increase of osteoporosis population. Although newer assessment has been made possible through quantitative CT (QCT), quantitative ultrasound (QUS) and magnetic resonance imaging (MRI), interest on developing texture analysis of X-ray bone image using Laws’ mask for direct evaluation of the bone quality is gaining momentum. However, the reliability of the system has been relatively untested. In this paper, we will introduce the confusion matrix to unveil the reliability of the system. By using standard deviation statistical descriptor, the system is 91.30% and 76.92% precise in detecting healthy and osteoporosis images respectively. In conclusion, we are encouraged by the high accuracies delivered by standard deviation descriptor and will extend the research by including more statistical descriptors in future research.

Key-Words: - Osteoporosis, Texture Analysis, Confusion Matrix, Accuracy, Precision

1 Introduction
Bone mineral density (BMD) is the indirect measurement approach for osteoporosis based on the absolute value called T score. T score is the number of standard deviation at which the indicated values reflect the mass density condition of the bone. According to World Health Organization (WHO), there are three classes of T score evaluations. The T score of normal population should lies above the value of -1. On the other hand, osteopenia (prelude to osteoporosis) ranges between -1 and -2.5 while osteoporosis population would have a T score below -2.5 [1].

The disadvantage of BMD, although widely accepted in medical diagnosis technologies, lies with this approach inability to demonstrate the bone quality directly. Current BMD has been practiced by using Dual Energy X-ray Absorptiometry (DXA). However, DXA is both inconvenient and expensive due to the large dimension of the machine and advanced technologies integrated into the machine. The indirect BMD, coupled with the inconvenient DXA, poses serious flaw in nowadays medical osteoporosis diagnosis services, and resembles a failure in combating the skeletal disease. Thus, other diagnosis methods such as Quantitative CT (QCT), Quantitative Ultrasound (QUS) and Magnetic Resonance Imaging (MRI) have been promoted as alternatives to counter the exponential growth of osteoporosis population. Nevertheless, neither the former nor the latter has been proven to be effective in deterring osteoporosis [2, 3].

Computerized image processing in assessing osteoporosis X-ray images has been touted as potential due to several advantages. Firstly, the computerized osteoporosis evaluation is derived based on accurate features representing osteoporosis. Therefore, we could eliminate the error of subjectivity reported in current manual assessment of osteoporosis. Secondly, image
processing in osteoporosis could render another new breakthrough in medical diagnostic technologies at which the evaluation is swift, convenient and low cost.

In 1980 Kenneth Ivan Laws brought forward the Laws’ masks idea, where the main contribution of his approach has been the filtering of images with specific masks created from the combination of one dimensional Kernel vector in order to assess the texture properties. The Laws’ masks match the pixel neighborhood to the set of standard masks to compute the texture properties of images.

There are 5 types of masks, namely level (L), edge (E), spot (S), ripple (R) and wave (W). In addition, this image processing technique could be further divided into 5x5 dimensions and 3x3 dimensions. The level vector demonstrates the average grey level or center weighted local average; edge vector resembles the gradient operator, responds to the column and row stepped edges in an images; spot vector represents the spot extraction; ripples vector detects ripple from the image while wave vector responds to any image pixel changes.

In this paper, we propose to investigate reliability of the 5x5 dimension Laws’ Mask texture analysis using the confusion matrix. The result of this experiment will contribute to the characterization accuracy of osteoporosis feature through appropriate classification statistical criterions. Two statistical descriptors namely mean and standard deviation will be tested using confusion matrix.

2 Overview of the Evaluation of System Reliability

In this experiment, we have acquired 6 normal X-ray and 1 osteoporosis X-ray image. Those X-ray images are depicting the knee joint section. Next, we enhance the contrast of each image through histogram equalization and resize the image before the desired features are extracted and parameterized through texture analysis.

We apply 5x5 dimension Laws’ Mask, producing a total of 25 masks. 4 regions of interest of great significance have been identified in order to decrease the total computational time. Through mean statistical descriptor and standard deviation statistical descriptor, we have conducted the statistical evaluation to examine the reliability of the system for future osteoporosis detection using Laws’ masks technique. The mathematical formulas for the mean and standard deviation statistical descriptor computed in this experiment have been illustrated in Eq. (1) and Eq. (2). The overall experiment flow to evaluate the reliability of the system is illustrated by the flow diagram shown in Fig 1.

\[
Mean = \frac{\sum_{i=0}^{M} \sum_{j=0}^{N}[TR_{ij}]}{M \times N} \tag{1}
\]

\[
SD = \sqrt{\frac{\sum_{i=0}^{M} \sum_{j=0}^{N}[TR_{ij} - Mean]^2}{M \times N}} \tag{2}
\]

Fig 1: Flow diagram of the evaluation of system reliability
2.1 Histogram Equalization of X-ray Images

Prior to the texture analysis of X-ray knee joint images, we process the X-ray images by using the global histogram equalization approach to better delineate the bone structure of X-ray images. The histogram with intensity level \([0, L - 1]\) of the image is adjusted by normalizing the number of pixel with intensity \(r_k, n_k\) with the total number of pixel, \(N\) in the X-ray image, where the mathematical expression of the histogram normalization, \(p(r_k)\) is shown in Eq. (3). The normalized histogram represents an estimate of the likelihood of intensity level \(r_k\) to occur in the X-ray image [4].

\[
p(r_k) = \frac{n_k}{N} \quad (3)
\]

In this experiment, we view the input intensity variable, \(r\) in the X-ray image as random variables in the interval of \([0, L - 1]\). The output of the intensity variable, \(s\) can be produced through transformational function, \(T(r)\). The transformation is illustrated in Eq. (4). We would descript the random variable through probability density function (PDF), in which the input intensity variable \(r\) and the transformed intensity \(s\) are contributing to the PDF of the transformed variable \(s\) shown in Eq. (5). Then, Leibniz’s rule is introduced into the PDF that the derivative of a definite integral with respect to its upper limit is the integrand evaluated at the limit as shown in Eq. (6). The final \(p_s(s)\) is obtained in the form of uniform probability density function, in which the resulting intensity does not depend on the form of the PDF of \(p_r(r)\). The independence of \(p_s(s)\) is proven in Eq. (7).

\[
s = T(r)
\]

\[
= (L - 1) \int_0^r p_r(w)dw \quad (4)
\]

\[
p_s(s) = p_r(r)\left|\frac{dr}{ds}\right| \quad (5)
\]

2.2 Laws’ Masks Texture Analysis

Derived from the concept of texture energy defined at each pixel after a series of particular convolution with selected mask, Laws’ Mask can produce the texture energy measurement for the analysis of the texture property of each pixel [5, 6]. The two 2-dimension convolution kernels, generated from different combinations of the 5 masks, are applied onto the converted gray scale image. The image \(I_{i,j}\) is assumed to have a size of \(N\) rows and \(M\) columns. After the filtration, say by a 2D mask \(L_5L_5\), the image is recognized as a texture image, \(TI\), and will have an image dimension of \(N + 1 \times M + 1\). The texture image is described in Eq. (8).

\[
TI_{L_5L_5} = I_{i,j} \times (L_5L_5) \quad (8)
\]

According to Laws, only texture image \(TI_{L_5L_5}\) and \(TI_{E_3E_3}\) can normalize the contrast of all texture image \(TI_{i,j}\) because these two 2D mask combinations have zero mean compare to other [7]. As a result, the result produced by these mask combination will depend heavily on image intensity rather than texture.

We will perform the windowing operation. The image passes through the texture energy measure (TEM) to replace every pixel of the image by comparing the pixel with its local neighborhood and subsequently sum up the absolute values from the neighboring pixel. The operation will lead to the creation of TEM image and described in Eq. (9).
After the windowing operation, we will normalize the features for the contrast of all the obtained images in order to be presented appropriately as image. All convolution kernels are zero-mean except for $L_5$, where the normalized kernels are described in Eq. (10).

\[
\text{Normalised (} TI_{\text{mask}} \text{)} = \frac{TI_{(i,j)\text{mask}}}{TI_{(i,j) L_5 L_5}}
\]  

Lastly, we will combine the TEM descriptors to eliminate a dimensionality bias from features. As a result, there will be only 14 masks left.

### 2.3 Euclidean Distance-based Clustering

We have conducted cluster analysis in pattern recognition stage. Cluster analysis focuses primarily on assigning different set of objects with certain degree of similarities into their assigned group (cluster) [8, 9]. Cluster analysis enables us to identify the relationship between different clusters. Euclidean distance remains the subset of the connectivity based clustering or hierarchical clustering. This type of cluster analysis derived from the idea of relating a reference object with two objects with different distances. The nearer object is more related to the reference object [10]. Different clusters formed based on distance measurements will be represented by dendogram.

The Euclidean distance based clustering represents the grouping of a set of data based on the distance measurement. The definition of Euclidean distance measurement is illustrated in Eq. (11), where the $x_i$ is the mean value of testing input image and $x$ is the standard mean value of the training set images. For instance, two data point is computed using square root of the sum of the squares for the difference between these two corresponding values based on the Pythagoras theorem.

\[
d(x_i, x) = \sqrt{\sum_{i=1}^{n} (x_i - x)^2}
\]  

In our experiment, we assume the output distance that is shorter than our threshold distance to be the same cluster and the classification is divided into 2 classes i.e. healthy and osteoporosis. Thus, an image is classified as healthy if the distance value computed is smaller compared to that of the osteoporosis bone image.

We repeat the procedure when measuring the standard deviation descriptor. Nevertheless, the mean value of healthy bone image is lower than osteoporosis image while in the case of standard deviation, the osteoporosis image will render a higher value compared to healthy image. The higher standard deviation demonstrated by osteoporosis image is largely due to the reason that the pixel value of osteoporosis image is lower, contributing to larger deviation from its mean value.

### 2.4 Confusion Matrix

Conventional reliability of the developed system has been determined by using the simple yet unreliable classification metric such as accuracy. For instance, the accuracy fails in the presence of large number of varying samples derived from different classes. The confusion matrix, on the other hand, offers much detailed statistical analysis on the true reliability of the system. The matrix discloses the number of actual results that match the predicted results.

Fig 2 depicts the general formulation of confusion matrix. These 4 outcomes are defined as True positive (TP), true negative (TN), false positive (FP), and false negative (FN). If the positive and negative predictions are correctly matched by the actual incidents, that the situation is recognized as true positive and true negative respectively. In the case where the actual class is a “yes” but the predicted class has been wrongly classified as “no”, then the result is considered false negative. The same case applies for false positive.
Fig 2: General confusion matrix

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Negative</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Negative</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Actual Positive</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

Meanwhile, the precision and the accuracy of the system can be computed through confusion matrix as well, in which the mathematical expression of these two elements are illustrated in Eq. (12) and Eq. (13).

\[
\text{Accuracy} = \frac{a + d}{a + b + c + d} \quad (12)
\]

\[
\text{Precision} = \frac{a}{a + c} \quad (13)
\]

3 System Reliability based on Confusion Matrix Precision

Table 1: Confusion matrix for mean statistical analysis with the overall accuracy and precision for the system

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Positive</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Actual Negative</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Accuracy</td>
<td>66.67%</td>
<td></td>
</tr>
<tr>
<td>Precision</td>
<td>83.33%</td>
<td>50%</td>
</tr>
</tbody>
</table>

Table 2: Confusion matrix for standard deviation statistical analysis with the overall accuracy and precision for the system

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Positive</td>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>Actual Negative</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Accuracy</td>
<td>86.11%</td>
<td></td>
</tr>
<tr>
<td>Precision</td>
<td>91.30%</td>
<td>70.92%</td>
</tr>
</tbody>
</table>

Laws’ mask is useful because this technique eliminates the pre-processing, thus minimizing the loss of information. This technique has been proven to be reliable after substantial amount of researches have been conducted to improve this technique.

The confusion matrix has been introduced in order to assess the reliability of the texture analysis system for osteoporosis detection. In this paper, we treat positive as healthy image class while negative as osteoporosis image class. The results are compared by using the mean and standard deviation statistical analysis in order to determine a suitable statistical analysis descriptor for the system.

From Table 1, the recorded reliability of the system based on mean statistical analysis is rendering mediocre assessments. The system is capable of recognizing the osteoporosis image by 50% and healthy image by 83.33% while the accuracy of the system stands at 66.67%.

Meanwhile, from Table 2, the standard deviation statistical analysis is considered a better selection with the precision of identifying healthy images of 91.30% and osteoporosis of 76.92%. The accuracy of the overall system is 86.11%. The standard deviation statistical analysis improves the classification statistical descriptor for the system in both the healthy and osteoporosis cases.
4 Conclusion
In this paper, we have evaluated the reliability of using the mean statistical descriptor and the standard deviation statistical descriptor to enhance the reliability of the system and the reliability of the overall texture analysis system for osteoporosis detection. The mean statistical descriptor has been demonstrating lower accuracy in identifying both the healthy and osteoporosis images when compared with standard deviation statistical descriptor.

Therefore, we are strongly recommending the application of the standard deviation statistical descriptor in future researches for developing this system compared with mean statistical descriptor. Besides, we will suggest that future researchers need to increase the number of images included in their experiment. The small number of acquired images in this paper represents a drawback to our experiment. In future, we will extend our experiment by analyzing more statistical analysis parameters such as entropy, skewness and kurtosis to develop the texture analysis for osteoporosis detection.

Reference: