Application of Image Processing Techniques to Bone Radiograph Images for Osteoporosis Diagnosis

Anirudha Sundaresan
School of Electrical and
Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332
Email: asundaresan7@gatech.edu

Krishna Kumar Balakrishnan
School of Electrical and
Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332
Email: krishna.kumar@gatech.edu

***Abstract*—In this paper, we use image processing and machine learning techniques as a tool for 2D texture analysis of bone radiograph images for osteoporosis detection. Automated diagnosis presents a major challenge for pattern recognition and medical applications due to high degree of resemblance between osteoporotic and healthy subjects. We focus on the use of various feature extraction methods with different classification techniques. We further investigate the performance of our method when subject to noise and distortion.**

Keywords—Texture analysis, feature extraction, osteoporosis, classification, machine learning.

I. Introduction

Osteoporosis is a disease characterized by low bone mass and loss of bone tissue that may lead to weak and fragile bones. About 54 million Americans have low bone mass, placing them at increased risk of osteoporosis [1]. A low bone mineral density (BMD) is presently regarded as the biggest factor for the development of osteoporosis. However, BMD can represent only 60% fracture prediction rate [2], estimated by the dual-energy X-ray absorptiometry. 2D texture analysis for evaluating bone structure on radiographs, as studied in [2], [3], [4] has been shown to improve prediction rates. Texture features are effective in capturing the degradation of bone microarchitecture.

Our approach, thus, deals with extracting spatial and transform based texture features that can be used to differentiate between the two subjects. Features based on gabor filters, local binary patterns, covariance, fractal dimension and statistical features constitute the spatial feature domain. Transform features are based on the wavelet and curvelet decomposition, and discrete cosine transforms. Using these features, support vector machines (SVM), boosting algorithms (Adaboost) and ensemble methods (random forests) have been employed for classification. We further investigate the performance of our proposed method in the presence of noise and distortion.

II. Methods

A. Texture Features

1) First Order and Second Order Statistical features:

Statistical methods do not attempt to understand explicitly the hierarchical structure of the texture. Instead they represent the texture indirectly by properties that govern the distributions and relationships between the grey scale levels of the image. [5]

The texture of an image can be characterized by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image. Thus, a GLCM was constructed and the statistical measures were extracted from this matrix. These statistical measures include correlation, contrast, homogeneity and energy of the GLCM. We also compute the mean, standard deviation and entropy of the images, and finally obtain a 7 dimensional feature vector.

2) Local Binary Pattern (LBP) based features:

Local binary pattern based approaches are another important class of statistical features that are highly useful for texture analysis. [6]. LBP features are simple and efficient for grey scale and rotation invariant texture classification [7]. Here, a 3x3 mask compares intensity value of each pixel with that of its neighbours and assigns 8 digit binary numbers for each bin of the histogram. These numbers constitute a 59 dimensional feature vector.

3) Fractal Dimensional features:

A fractal is defined as a mathematical set whose Hausdorff dimension exceeds its topological dimension [8]. Fractal dimensions are important characteristics of fractals that reveal their geometric structure. We have used Segmentation-based Fractal Texture Analysis (SFTA) as described in [9]. The input image is decomposed into a set of binary images from which the fractal dimensions of the resulting regions are computed to describe segmented texture patterns. The fractal dimensions, area and mean intensity for each point constitute the feature vector.

4) Gabor Filter:

Gabor filter is a linear filter that analyses particular frequency content in an image in specific directions in a localized region around the point or region of analysis. They have been found to be particularly appropriate for texture representation and discrimination. A Gabor filter is essentially a sinusoidal signal with a given frequency and orientation, modulated by a Gaussian [10].

A set of Gabor filters with different frequencies and orientations may help in extracting useful features from an image

[7]. In the discrete domain, two-dimensional Gabor filters are given by:

$$G_c[i,j] = Be^{\frac{-(i^2+j^2)}{2\sigma^2}} cos(2f\pi(icos\theta + jsin\theta))$$
$$G_s[i,j] = Ce^{\frac{-(i^2+j^2)}{2\sigma^2}} cos(2f\pi(icos\theta + jsin\theta))$$

where B and C are the normalizing factors to be determined. f defines the frequency being looked for in the texture. By varying θ, we can look for texture oriented in a direction. By varying σ, we change the support of the basis or the size of the image region being analysed. 2-D Gabor filters have rich applications in image processing, especially in feature extraction for texture analysis and segmentation.

5) Covariance based features:

Covariance matrices have been extensively used for texture analysis [11]. For a region of an image $I\varepsilon R^{d_1 x d_2}$ we compute the local features (usually statistical) for every pixel p_{ij}. The covariance matrix is then computed using the following equation:

$$C = \frac{1}{n-1}\sum_{i,j}(f_{ij} - \bar{f})(f_{ij} - \bar{f})^T$$

with $n = d_1$ x d_2 and $\bar{f}$ being the empirical mean of f.

We use gradient based features for computing the local features on each pixel.

$$f_{ij} = \Bigg[I^{ij}, |I_x^{ij}|, |I_y^{ij}|, |I_{xx}^{ij}|, |I_{yy}^{ij}|, \sqrt{(I_x^{ij})^2 + (I_{xy}^{ij})^2}, arctan\frac{|I_x^{ij}|}{|I_y^{ij}|}\Bigg]^T$$

where I^{ij} is the intensity of pixel (i,j) and I_x, I_{xx},.. are the intensity derivatives (first and second order along x and y axis) and the last term represents edge orientation leading to a 7x7 covariance matrix [11].

6) Non-directional transform based features:

In frequency based approach, we have utilized the discrete cosine transform (DCT) coefficients as they provide information about texture specific spectral characteristics. Lower frequencies indicate coarse textures whereas higher frequency represent fine textures [2].

In wavelet based approach, each image has been decomposed using the Haar wavelets. For extracting the Haar wavelet features we need to have images with dyadic dimensions. Thus, each image has been resized to a 64x64 image. Mean of the Haar high frequency wavelet coefficients along the horizontal and vertical direction (HH block) for every level of decomposition constitutes the wavelet feature vector [11].

7) Directional transform based features (Curvelets):

Curvelets are a non-adaptive technique for multi-scale object representation. Curvelets are an extension of wavelets where the basis functions are localized in orientation as well. Due to this directionality, the curvelet coefficients are good texture descriptors [12].

We first resize the images to fit dyadic dimensions. Curvelet coefficients are generated using the 2D FFT. FFT is first applied to an image. Next, the 2D Fourier plane is divided into different scales. Each scale is further divided into many different tiles [13].The feature vector in this case consists of the mean of the curvelet coefficients in the higher frequency sub-bands.

B. Classifiers

1) Support Vector Machines (SVMs):

SVMs are supervised learning models with associated learning algorithms that analyse data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier. In our approach, we have used a linear SVM kernel.

2) Adaptive Boosting (AdaBoost):

Boosting is an approach to machine learning based on the idea of creating a highly accurate prediction rule by combining many relatively weak and inaccurate rules. The AdaBoost algorithm of Freund and Schapire [14] consists of two parts, a simple weak classifier and a boosting part. The weak classifier tries to find the best threshold in one of the data dimensions to separate the data into two classes 1 for osteoporotic and -1 for control. The boosting part calls the classifier iteratively and after every classification step, it changes the weights of mis-classified examples. This creates a cascade of "weak classifiers" which behaves like a "strong classifier". We have used 50 training iterations for this algorithm.

3) Ensemble Method (Random Forest):

Random forests are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes or mean prediction of the individual trees. Random decision forests correct for decision trees' habit of overfitting to their training set. They also show robustness with respect to noise. We use an ensemble of 30 bagged classification trees for training.

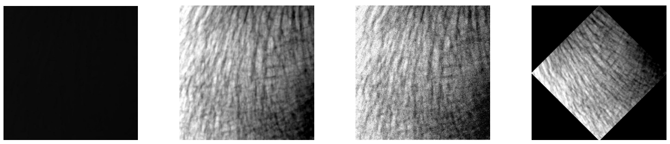
4) Deep Learning Method (CNNs):

CNN is a class of deep, feed-forward artificial neural networks that has successfully been applied to analyzing visual imagery. CNNs use a variation of multilayer perceptrons designed to require minimal preprocessing. CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were hand-engineered. This independence from prior knowledge and human effort in feature design is a major advantage. Steepest gradient descent algorithm is used to minimize the cross entropy cost function.

III. Experiment

This is an IEEE-ISBI: Bone Texture Characterization open challenge (2014) and the database used in this research work is publicly available online at http://www.univ-orleans.fr/i3mto/data. Our database contains annotated images of 58 osteoporotic and 58 healthy subjects that display a region of interest in the trabecular bone. The images are 16-bit TIFF format (400x400 pixels size). We split the dataset into 80 images for training and 36 for testing.

Pre-processing is a pivotal step in texture analysis of medical images, since the images obtained are prone to noise. Our first step involves adjusting the intensity values to enhance contrast. We first test our algorithm assuming noiseless images. We then evaluate our methods in the case of noisy images (Gaussian Noise with 0 mean and 0.01 variance) and images with distortion (Rotation of 45^o). In the latter cases, wiener filter and rotation correction measures have been implemented in the pre-processing steps as well.



(a) Original (b) Enhanced (c) Noisy (d) Distorted

Fig. 1: Example of an image in our dataset.

Due to the low number of images in our dataset, partitioning has been done. It was observed that this led to improvement in the prediction rates of the classifiers. Thus, for the training set, each of the 80 images has been partitioned into 16 overlapping blocks of dimensions 100x100. This generates a total of 1280 images (640 for osteoporotic and 640 for control subjects) for training. The remaining 18 images from both the control and osteoporotic cases are used for testing.

As mentioned earlier, we broadly divide our feature set based on 5 categories, namely, Spatial domain, Transform domain, Directional transform, Non-directional transform and a combination of all the features. To analyse their performance, each feature set has been fed to the classifier separately. We further investigate the effect of noise and distortion on the classifiers that we have used. These results have been recorded in TABLE I and TABLE II.

Noiseless Partition		All	Transform	Spatial	Directional	Non-Directional
SVM	**TPR**	0	0	22.22	5.56	0
	TNR	5.56	0	50	11.11	0
	ACC	47.2	**50**	36.11	47.22	**50**
Boosting Technique (Adaboost)	**TPR**	77.8	72.22	77.78	72.22	38.89
	TNR	66.7	55.56	66.66	27.78	44.44
	ACC	55.6	58.33	55.56	**72.22**	47.22
Ensemble Method (Random Forests)	**TPR**	50	38.89	55.56	66.67	11.11
	TNR	22.2	77.78	38.89	22.22	38.89
	ACC	63.9	30.56	58.33	**72.22**	36.11

TABLE I: Classification results for sets of features

From TABLE I we observe that, for a linear SVM classifier, the transform based features contributes the most towards prediction. On the other hand, for AdaBoost and Random Forest classification, curvelet's performance was far superior when compared to other feature sets. This indicates that directionality is a good measure of the textural differences between visually similar images.

The observations in TABLE II show the accuracy of the classifiers in the presence of noise and distortion. The

All Features		Noise	Angular Distortion	Noise & Distortion
SVM	**TPR**	0	0	5.56
	TNR	5.56	5.56	5.56
	ACC	47.22	**47.22**	50
Boosting Technique (Adaboost)	**TPR**	94.44	100	100
	TNR	88.89	100	100
	ACC	**52.78**	50	50
Ensemble Method (Random Forests)	**TPR**	50	83.33	50
	TNR	50	88.89	33.33
	ACC	50	**47.22**	**58.33**

TABLE II: Classification results using all features for Noisy & Distorted Images

ineffectiveness of the SVM classifier is inherent from the fact that there is no change in accuracy despite disturbances. It can also be seen that, Random Forest and Adaboost tend to perform well even in the case of noise and distortion.

We also train the partitioned data using CNN and include the results in TABLE III. The CNN is trained with 15 filters of dimensions 15x15, gradient descent learning rate of 10^{-6}, mini-batch size of 70 and with 6 epochs.

Epoch	Iteration	Mini-batch Loss	Mini-batch Accuracy
1	1	0.6895	48.44%
3	50	0.4935	82.81%
5	100	0.3566	93.75%
6	120	0.3045	93.75%

TABLE III: Training Results for CNN

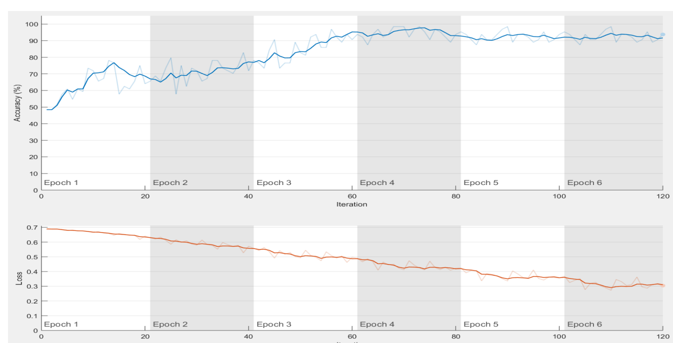


Fig. 2: Training Progress of CNN indicating accuracy and batch loss for each epoch

A prediction rate of 63.89% was achieved using CNN. This would increase with the number of layers as well as the number of neurons in each hidden layer.

IV. Conclusion

Our results show that degradation of the bone architecture in bone radiograph images can be identified by texture analysis. A prediction rate of 72.22% was achieved despite high visual similarity between the two classes, indicating that our proposed system is efficient in discriminating osteoporotic subjects from control subjects. Furthermore, we plan to use feature selection techniques for more robust classification.

REFERENCES

[1] Klibanski, Anne, et al. ”Osteoporosis prevention, diagnosis, and therapy.” *Journal of the American Medical Association* 285.6 (2001): 785-795.

[2] Zheng, Keni, and Sokratis Makrogiannis. ”Bone texture characterization for osteoporosis diagnosis using digital radiography.” *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the. IEEE*, 2016.

[3] Ahmadv, A., and M. R. Daliri. ”A review on texture analysis methods in biomedical image processing.” *OMICS J Radiol* (2016).

[4] Brodatz, Phil. Textures: a photographic album for artists and designers.*Dover Pubns*, 1966.

[5] A. Materka, M. Strzelecki, Texture Analysis Methods A Review, Technical University of Lodz, Institute of Electronics, *COST B11 report, Brussels*, 1998.

[6] T. Ojala, M. Pietikainen, and T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 24, no. 7, pp. 971987*, 2002.

[7] Y. Hu, Z. Long, G. AlRegib, Scale Selective Extended Local Binary Pattern for Texture Classification, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2017), New Orleans*, March 2017.

[8] A. P. Pentland, Fractal-Based Description of Natural Scenes, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, 1984, pp. 661-674.]

[9] Costa, A. F., G. E. Humpire-Mamani, A. J. M. Traina. 2012. ”An Efficient Algorithm for Fractal Analysis of Textures.” *In SIBGRAPI 2012 (XXV Conference on Graphics, Patterns and Images)*, 39-46, Ouro Preto, Brazil.

[10] M. Haghighat, S. Zonouz, M. Abdel-Mottaleb, ”CloudID: Trustworthy cloud-based and cross-enterprise biometric identification,” *Expert Systems with Applications*, vol. 42, no. 21, pp. 7905-7916, 2015.

[11] Yger, Florian,”Challenge IEEE-ISBI/TCB: application of covariance matrices and wavelet marginals.” *arXiv preprint arXiv:1410.2663* (2014).

[12] Arivazhagan, S., L. Ganesan, and TG Subash Kumar. ”Texture classification using curvelet statistical and co-occurrence features.” *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*. Vol. 2. IEEE, 2006.

[13] H. Al-Marzouqi and G. AlRegib, Curvelet Transform with Learning-based Tiling, *Elsevier Signal Processing: Image Communication Journal*, 53 (2017) 24-39, January 2017.

[14] Freund, Yoav, and Robert E. Schapire. ”Experiments with a new boosting algorithm.” *Icml*. Vol. 96. 1996. APA.