An Efficient Clustering based Texture Feature Extraction for Medical Image

Marghny.H.Mohamed
Faculty of computers and Information
Assiut University, Assiut, Egypt,
marghny@aun.edu.eg

Mohammed. M. Abdelsamea
Mathematics Department,
Assiut University, Egypt,
mm_abdelsamea@yahoo.com

Abstract

In some medical applications where a tissue of interest covers a large fraction of the image or a prior knowledge on the region of interest is available, extracting features by fixed blocs in the image is sufficient. However in the general case, one would like to identify features for each tissue in the image. This would require prior image segmentation. Medical image segmentation is one of the most challenging problems in medical image analysis and a very active research topic. Therefore, there is no algorithm available in the general case for isolating medical image regions [1].

This paper presents an accurate method for extracting texture features from medical image for classification. It is based on bloc wise clustering of medical images. The proposed technique extracts accurate and general set of texural features. Experimental result showed the high accuracy of the extracted textural features. Experiments held on Mammographic Image Analysis Society M(*IAS*) *dataset.*

1. Introduction

Image analysis techniques have played an important role in several medical applications. In general, the applications involve the automatic extraction of features from the image which is then used for a variety of classification tasks, such as distinguishing normal tissue from abnormal tissue. Chabat [2] used 13 texture parameters, derived from the histogram, co-occurrence matrix and run-length matrix categories, to differentiate between a variety of obstructive lung diseases in thin-section CT images. Kovalev [3] used texture parameters derived from gradient vectors and from generalized co-occurrence matrices for the characterization of texture of some MR-T2 brain images. Herlidou [4] used texture parameters based on the histogram, co-occurrence matrix, gradient and run-length matrix for the characterization of healthy and pathological human brain tissues (white matter, grey matter, cerebrospinal fluid, tumours and oedema). Mahmoud [5] used the texture analysis approach based on a three-dimensional co-occurrence matrix in order to improve brain tumour characterization. Du-Yih Tsai [6] used four texture features derived from the co-occurrence matrix was used for classification of the heart disease. H.S. Sheshadri [7] used Six textural features for mammogram images derived from the histogram categories was used as a part of developing a computer aided decision system for early detection of breast cancer. Maria-Luiza [8] used texture parameters based on the histogram for tumour classification in mammograms.

Feature extraction [9] is a vital component of the Computer Aided Diagnosis (CAD) System that can discriminate between medical tissues to serve as a second reader to aid radiologists. The feature extraction unit is used to prepare data in a form that is easy for a decision support system or a classification unit to use. Compared to the input, the output data from the feature extraction unit is usually of a much lower dimension as well as in a much easier form to classify. Medical images possess a vast amount of texture information relevant to clinical practice [10]. Hence texture is the most promising feature to work on. Texture analysis gives information about the arrangement and spatial properties of fundamental image elements [11]. Coggins [12] has compiled a catalogue of texture definitions.

One of the most commonly used texture parameters come from Co-occurrence matrix as a statistical approach [10] which represents texture in an image using properties governing the distribution and relationships of grey-level values in the image methods normally achieve higher discrimination indexes than the structural or transform methods.

This paper introduces a bloc wise clustering as a region of interest selection (ROI) approach which provides a more accurate extraction of textural features. We use four texture features measured from a gray-level co-occurrence matrix generated from the breast images for classification of the images. A statistical discrimination method (fisherfaces algorithm) [13, 14] for feature selection algorithm is also used for extracting discriminative information from extracted feature of medical images to be used as inputs to our classification system.

2. Data collection

The data collection which has been used in our experiments was taken from the MIAS [15]. The same collection has been used in other studies of automatic mammography classification. Its corpus consists of 322 images, which belong to three big categories: normal, benign and malign. There are 208 normal images, 63 benign and 51 malign, which are considered abnormal. In addition, the abnormal cases are further divided in six categories: microcalcification, circumscribed masses, speculated masses, ill-defined masses, architectural distortion and asymmetry.

3. Pre-processing phase

Mammograms are images difficult to interpret, and a preprocessing phase of the images is necessary to improve the quality of the images and make the feature extraction phase more reliable. We applied to the images two techniques: a cropping operation and an image enhancement one. The first one was employed in order to cut the black parts of the image as well as the existing. Image enhancement helps in qualitative improvement of the image with respect to a specific application. In order to diminish the effect of over brightness or over darkness in the images and accentuate the image features, we applied a widely used technique in image enhancement to improve visual appearance of images known as Histogram Equalization. This process equalizes the illumination of the image and accentuates the features to be extracted [8].

4. Feature extraction

A major component in analyzing images involves data reduction which is accomplished by intelligently modifying the image from the lowest level of pixel data into higher level representations. From these higher level representations we can gather useful information; a process called feature extraction [9]. Extracting features by fixed blocs in the image has been considered to be sufficient as an ROI selection method in some medical applications where a large fraction of the image is covered by tissue of interest.

4.1. Measurements of texture features

The gray-level co-occurrence matrix (GLCM) is a matrix used to express the correlation of spatial location and gray-level distribution of an image [16]. From it, the local variation of gray levels on an image or sub-image can be statistically investigated and in turn, enable us to know the manner of change in gray level as a whole. In the current application, we used the following conditions to generate gray-level co-occurrence matrices.

- Direction: In general the gray-level co-occurrence matrices from 0, 45, 90, and 135 directions are used. Only the direction of 0 was used in the study.
- Distance: The length of 1-pixel was used.

Of the 14 original statistics developed by Haralick et al. co-occurrence y used features

ıst are often used in practice. The following four statistics will be used exclusively in this paper:

1. Dissimilarity $=\sum_{i,j=1}^{G} C_{ij}\left|i-j\right|$.
2. Uniformity $=\sum_{i,j=1}^{G} C_{ij}^{2}$.
3. Entropy $=-\sum_{i,j=1}^{G} C_{ij}\lg C_{ij}$.
4. Contrast $=\sum_{i,j=1}^{G} C_{ij}(i-j)^{2}$.

Where, Cij represents co-occurring probabilities stored inside GLCM. G represents number of grey level available.

Now, we introduce and study two ROI selection methods that give us better texture feature extraction when co-occurrence matrix method is used and hence a more accurate texture feature (see section 7).

4.1.1. Extraction based on pixel wise segmentation approach. For a more accurate feature extraction, apply extraction based pixel wise partitioning method as follows (see Figure.1):

Step1: divide the entire image into SN non-overlapping sub-images SI= $\{I_1,I_2,\ldots,I_{SN}\}$.

Step2: use the k-means algorithm [17] to cluster the sub-images (SI) into several classes based on pixel intensity for each I_i, i=1,2,...,SN independently.

Step3: for each cluster in I_i, i=1,2,...,SN, construct a sub-image representing set of texture feature vectors $F_K=\{f_1,f_2,\ldots,f_X\}$, k=1,2,...,L; where L is the number of classes each of which contains X texture features.

Step4: build the final set of texture features representing the overall image in the form of a single transaction of the final dataset (set of images), $T=\{t_1,t_2,\ldots,t_c\}$, where c is the number of images, t_i is a vector of the size (SN×L×X), i=1, 2, ..., c.

Step 5: for each t_i, i=1,2,...,c add the class label of its image.

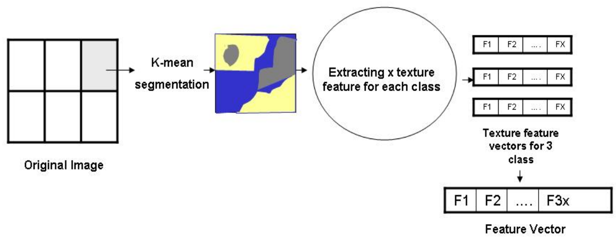

Figure.1 Illustrate simple example describes feature extraction phase: where SN=6 and L=3.

4.1.2. Extraction method based on proposed ROI selection approach. For a more accurate feature extraction and a further investigation of the localization, apply extraction based bloc wise partitioning method as follows (see Figure.2):

Step1: divide the entire image into SN non-overlapping sub-images SI= $\{I_1,I_2,\ldots,I_{SN}\}$.

Step2: split each of these SN sub-image into other M blocs $I_j=\{B_1,B_2,\ldots B_M\}$, j=1,2,…,SN.

Step3: for each bloc B_i, i=1,2,…,M, construct a bloc representing set of texture feature vectors.

Step4: use the k-means algorithm to cluster the feature vectors into several classes for each sub-image I independently.

Step5: for each cluster in I_i, i=1,2,…,SN, construct a sub-image representing set of texture feature vectors $F_K=\{f_1,f_2,\ldots,f_X\}$, k=1,2,...,L; where L is the number of classes each of which contains X texture features.

Step6: build the final set of texture features representing the overall image in the form of a single transaction of the final dataset (set of images), $T=\{t_1,t_2,\ldots,t_c\}$, where c is the number of images, t_i is a vector of the size (SN×L×X), i=1, 2, ..., c.

Step7: for each t_i, i=1,2,...,c add the class label of its image.

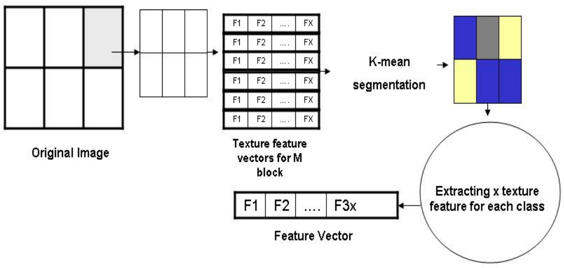

Figure.2 Illustrate simple example for describe feature extraction phase: where SN=6, M=6 and L=3.

5. Feature selection

The statistical discrimination methods are suitable not only for classification but also for characterization of differences between a reference group of patterns and the population under investigation.

For the image classification case, without class labels, Principle Component Analysis (PCA) can be applied, when some class labels are available, linear discriminant analysis (LDA) is applied to transform the features into the most discriminating Feature space [18-21].

Let us consider a set of N sample images $\{x_1,x_2,\ldots,x_N\}$ taking values in an n-dimensional image space, and assume that each image belongs to one of c classes. Let us also consider a linear transformation mapping the original n-dimensional image space into an m-dimensional feature space, where m < n. The new feature vectors Y are defined by the following linear transformation:

$$Y_k = W^T x_k, \qquad k=1,N. \qquad (1)$$

5.1. Fisher's Linear Discriminant Analysis (FLD)

In this method, W in (1) can be selected in such a way that the ratio of the between-class scatter and the within-class scatter is maximized. Let the between-class scatter matrix be defined as

$$S_b = \sum_{i=1}^{c} N_i(\mu_i-\mu)(\mu_i-\mu)^T,$$

and the within-class scatter matrix be defined as

$$S_w = \sum_{i=1}^{c} \sum_{x_k \in X_i} (x_k-\mu_i)(x_k-\mu_i)^T,$$

where μ_i is the mean image of class X_i, and N_i is the number of samples in class X_i. If S_w is nonsingular, the optimal projection W_{opt} is chosen as the matrix with orthonormal columns which maximizes the ratio of the determinant of the between-class scatter matrix of the projected samples to the determinant of the within-class scatter matrix of the projected samples, i.e.,

$$W_{opt} = \arg\max_W \frac{\left|W^T S_b W\right|}{\left|W^T S_w W\right|}. \qquad (2)$$

In this paper, we can not apply (FLD) directly to solve the recognition problem since the dimension of the sample space is typically larger than the number of samples in the training set. As a consequence, S_w is singular in this case.

5.1.1. Fisherfaces. Swets and Weng [22] proposed a two stage PCA+LDA method, also known as the Fisherfaces method, in which PCA is first used for dimension reduction so as to make S_W nonsingular before the application of LDA. In this method an alternative to the criterion in (2) is applied. This method avoids this problem by projecting the image data set to a lower dimensional space so that the resulting within-class scatter matrix S_W is nonsingular. This is achieved by using PCA to reduce the dimension of the feature space to N-c, and then applying the standard FLD defined by (2) to reduce the dimension to (c-1) [20]. More formally, W_{opt} is given by

$$W_{opt}^T = W_{fld}^T W_{pca}^T,$$

where

$$W_{pca} = \arg\max_W \left|W^T S_T W\right|,$$

$$W_{fld} = \arg\max_W \frac{\left|W^T W_{pca}^T S_b W_{pca} W\right|}{\left|W^T W_{pca}^T S_w W_{pca} W\right|},$$

$$S_T = \sum_{k=1}^{N} (x_k-\mu)(x_k-\mu)^T.$$

Note that the optimization for W_{opt} is performed over n$\times$(N-c) matrices with orthonormal columns, while the optimization for W_{fld} is performed over (N-c)$\times$m matrices with orthonormal columns. In computing W_{opt}, we have thrown away only the smallest (c-1) principal components.

6. Measures for performance evaluation

We evaluated the performance of the proposed methods in terms of sensitivity, specificity and overall accuracy [6], Such that k-fold cross validation process is used.

K-fold cross validation is used for model testing and evaluation to determine how accurately a learning algorithm will be able to predict data that it was not trained on. In this method it is not important how the data is divided. Every data point appears in a test set exactly once, and appears in a training set k-1 times. We can independently choose the size of the each test and the number of trials (Kohavi & Provost, 1998).

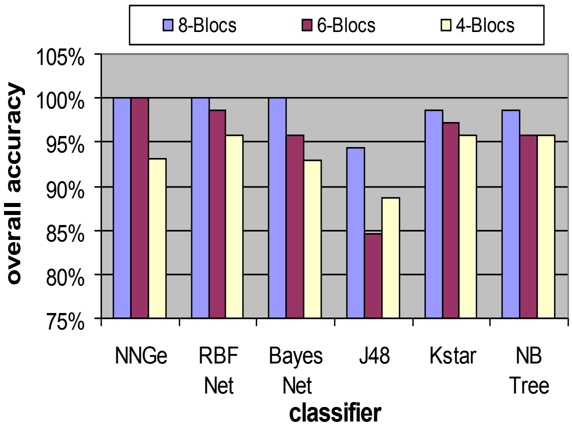

Figure.3 Success rates of classifiers with decreasing of bloc size

7. Experimental results

In our experiment a sample of 22% from the MIAS data set is selected randomly for model testing and evaluation. The sample is distributed among classes as follows:
Normal class (n=30), Microcalcification class (n=10), Circumscribed Masses class (n=7), Spiculated Masses class (=8), Ill-Defined Masses class (n=5), Architectural Distortion class (n=7) and Asymmetry class (n=4).

We evaluated the performance of the proposed methods described in sections 4.1.1 and 4.1.2 with standard classifiers. Using weka experimenter [23], Such that 10-fold cross-validation process is used.

Table.1 present the overall accuracy of textural feature set extracted by the extraction method based on bloc wise segmentation as a ROI selection approach (see section 4.1.2) for SN=6 and L=3. Several values of M were checked for the possibility of improving the results and increasing features localization. See table. 1.

Table.1 Overall accuracy for the proposed method with M=4, 6 and 8.

Classifier	Overall accuracy		
	4-blocs	6-blocs	8-blocs
NNGe	93.10%	100%	100%
RBF net	95.77%	98.59%	100%
Bayes net	92.95%	95.77%	100%
J48	88.73%	84.5%	94.37%
Kstar	95.77%	97.18%	98.59%
NB tree	95.77%	95.77%	98.59%

From results in the above table, we can notice that, using a certain feature selection algorithm (Fisherfaces), classifiers accuracy is dependent on number of blocs regardless of classification algorithms used. The overall accuracy with different bloc size is presented in Figure.3.

Table.2 illustrates comparison between the following textural feature sets based on different ROI selection methods:

1. Textural feature sets based on fixed bloc segmentation (simply divide the whole image into 6 non-overlapping sub-images then split each of these 6 sub-images into other 8 non-overlapping blocs).
2. Textural feature sets based on pixel intensity segmentation described in section 4.1.1 with SN=6 and L=3.
3. Textural feature sets based on our bloc wise segmentation described in section 4.1.2 with SN=6, M=8 and L=3.

Table.2 The classification rate comparison.

Classifier	ROI	Sensitivity	Specificity	Overall accuracy
NNGe		100%	100%	100%
RBF net		100%	100%	100%
Bayes net	Fixed bloc	96.77%	100%	98.59%
J48		85.71%	95.35%	90.14%
Kstar		96.77%	100%	98.59%
NB tree		93.75%	100%	97.18%
NNGe		100%	100%	100%
RBF net		100%	100%	100%
Bayes net	Pixel wise	96.77%	100%	98.59%
J48		88.24%	97.62%	92.96%
Kstar		100%	100%	100%
NB tree		96.77%	100%	98.59%
NNGe		100%	100%	100%
RBF net		100%	100%	100%
Bayes net	Bloc wise	100%	100%	100%
J48		90.9%	97.62%	94.37%
Kstar		96.77%	100%	98.59%

NB tree		96.77%	100%	98.59%

From results shown in the table.2, we can notice that, using Fisherfaces feature selection algorithm, our proposed extraction methods gave better accuracy than extraction method based on fixed bloc partitioning as a ROI selection method, which means a more accurate extraction of textural feature and hence a more efficient image representation. The classifier NNGe and RBF-Network gave the highest performance when compared to the others which implies that NNGe and RBF-Network are much more suitable for such application. Figure.4. presents a comparison between classification results.

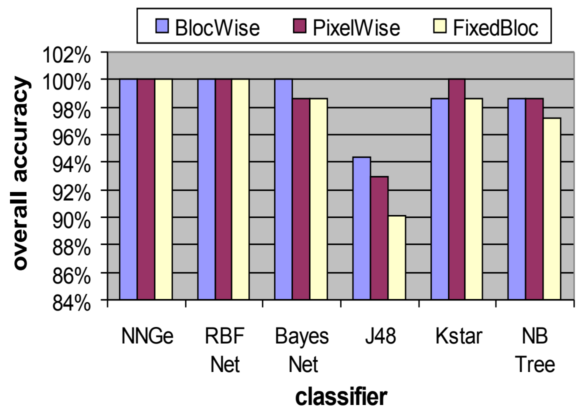

Figure.4 Success rates of classifiers with our proposed methods.

8. Conclusions

In this work, four texture features derived from the co-occurrence matrix was used as a part of developing CAD system for early detection of breast cancer.

This paper present textural extraction method based on three different ROI selection methods for obtaining efficient image representation. We proposed an ROI selection method based on bloc wise segmentation designed to obtain more accurate extraction of textural feature from medical images. It was shown that this approach improves the localization of texture features and produce general texture features. A fisherfaces feature selection algorithm is used for extracting discriminative information from extracted feature of medical images to be used as inputs to our classification system.

Our experiments confirm the importance of our proposed ROI selection approaches as a primary step of feature extraction in the final decision of the CAD.

9. Reference

[1] Tristan Glatard, Johan Montagnat, and Isabelle E.Magnin, "Texture based medical image indexing and retrieval: application to cardiac imaging", Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval, October 15-16, 2004, New York, NY, USA.

[2] Chabat F, Yang G-Z, and Hansell DM, "Obstructive lung diseases: texture classification for differentiation at CT", Radiology2003;228:871-7.

[3] Kovalev VA, Petrou M, and Bondar YS, "Texture anisotropy in 3D images", IEEE Trans Image Process 1999; 8:346-60.

[4] Herlidou-Meˆme S, Constans JM, and Carsin B, "MRI texture analysis on texture test objects, normal brain and intracranial tumours", Magn Reson Imaging 2003;21:989-93.

[5] Mahmoud-Ghoneim D, Toussaint G, Constans JM, and De Certaines JD, "Three dimensional texture analysis in MRI: apreliminary evaluation in gliomas", Magn Reson Imaging 2003; 21: 983-7.

[6] Du-Yih Tsai and Katsuyuki Kojima, "Measurements of texture features of medical images and its application to computer-aided diagnosis in cardiomyopathy", Measurement 37 (2005) 284–292.

[7] H.S.Sheshadri and A. Kandaswamy, "Experimental investigation on breast tissue classification based on statistical feature extraction of mammograms", Computerized Medical Imaging and Graphics 31 (2007) 46-48.

[8] Maria-Luiza Antonie, Osmar R. Zaiane, and Alexandru Coman, "Application of data mining techniques for medical image classification", In Proc. of Second Intl. Workshop on Multimedia Data Mining (MDM/KDD'2001) in conjunction with Seventh ACM SIGKDD, pages 94-101, San Francisco, USA, 2001.

[9] V.S.Kodogiannisa, M.Boulougourab, E.Wadgea and J.N.Lygourasc, "The usage of soft-computing methodologies in interpreting capsule endoscopy", Engineering Applications of Artificial Intelligence 20 (2007) 539-553.

[10] G.Castellano, L.Bonilha, L.M.Li, and F.Cendes, "Texture analysis of medical images", Clinical Radiology (2004) 59, 1061-1069.

[11] V.S.Kodogiannisa, M.Boulougourab, J.N.Lygourasc, and I.Petrouniasd, "A neuro-fuzzy-based system for detecting abnormal patterns in wireless-capsule endoscopic images", Neurocomputing 70 (2007) 704-717.

[12] Coggins, J. M, "A Framework for Texture Analysis Based on Spatial Filtering", Ph.D. Thesis, Computer Science Department, Michigan State University, East Lansing, Michigan, 1982.

[13] Carlos E. Thomaz, Nelson A. O. Aguiar, Sergio H. A. Oliveira, Fabio L. S. Duran, Geraldo F. Busatto, Duncan Fyfe Gillies, and Daniel Rueckert, "Extracting Discriminative Information from Medical Images: A Multivariate Linear Approach", SIBGRAPI 2006: 113-120.

[14] I.Cohen, Q.Tian, X.Zhou, and T.S.Huang, "Feature Selection and Dimensionality Reduction Using Principle Feature Analysis", submitted to the Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI-2000) Workshop on Fusion of Domain Knowledge with Data For Decision Support, Stanford University, Stanford, CA, June 30, 2000.

[15] http://www.wiau.man.ac.uk/services/MIAS/MIASweb.html

[16] David A.Clausi, and Yongping Zhao, "Rapid extraction of image texture by co-occurrence using ahybrid data structure", Computers and Geosciences 28 (2002) 763-774.

[17] Jose Marroquin, and Federico Girosi, "Some extensions of the K-means algorithm for image segmentation and pattern classification", Technical report, Massachusetts Institute of Technology, January 1993.

[18] K.Fukunaga, Introduction to Statistical Pattern Recognition, second ed., Academic Press, Boston, 1990.

[19] C.E.Thomaz, J.P.Boardman , S.Counsell , D.L.G.Hill, J.V.Hajnal, A.D.Edwards, M.A.Rutherford, D.F.Gillies, and D. Rueckert, "A multivariate statistical analysis of the developing human brain in preterm infants", Image and Vision Computing 25 (2007) 981–994.

[20] Peter N.Belhumeur, Joao P.Hespanha, and David J.Kriegman. "Eigenfaces vs. Fisherfacess: Recognition Using Class Specific Linear Projection", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 19, NO. 7, JULY 1997.

[21] P.A.Devijver, and J.Kittler, "Pattern Classification: A Statistical Approach", Prentice-Hall, Englewood Cliffs, N. J., 1982.

[22] D.L.Swets, and J.Weng, "Using discriminant eigenfeatures for image retrieval", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 18, no. 8, pp. 831-836, August 1996.

[23] http://www.cs.waikato.ac.nz/ml/weka/.