Classification of Medical Images Using Data Mining Techniques

B.G. Prasad[1] and Krishna A.N.[2]

[1] Department of Computer Science and Engineering,
B.N.M. Institute of Technology, Bangalore 560 070
drbgprasad@gmail.com
[2] Department of Computer Science and Engineering,
S.J.B. Institute of Technology, Bangalore 560 060
krishna12742004@yahoo.co.in

Abstract. Automated classification of medical images is an increasingly important tool for physicians in their daily activity. This paper proposes data mining classifiers for medical image classification. In this study, we have used J48 decision tree and Random Forest (RF) classifiers for classifying CT scan brain images into three categories namely inflammatory, tumor and stroke. The proposed classification system is based on the effective use of texture information of images. Three different methods implemented are: Haralick (H), Tamura (T) and Wold (W) texture features. All three texture features and the classification methods are compared based on Precision and Recall. The experimental result on pre-diagnosed database of brain images showed Haralick features combined with Random Forest classifier is found to give best results for classification of CT-scan brain images.

Keywords: CAD, Texture features, Data Mining classifiers.

1 Introduction

Advances in medical imaging technology and computer science have greatly enhanced interpretation of medical images, and contributed to early diagnosis. The development of Computer Aided Diagnosis (CAD) systems to assist physicians in making better decisions has been the area of interest in the recent past. CAD systems aim to provide computer output as a second opinion in order to assist physicians in the detection of abnormalities, quantification of disease progress and alternate diagnosis of lesions [1]. Recently CAD systems that use CBIR to search for clinically relevant and visually similar images (regions) depicting suspicious lesions has also been attracting research interest. CBIR-based CAD schemes have potential to provide radiologists with visual aid and increase their confidence in accepting CAD-cued results in the decision making [2]. In either conventional or CBIR-based CAD, automated classification techniques are needed to facilitate physician's diagnosis of complex diseases.

V.V. Das and J. Stephen (Eds.): CNC 2012, LNICST 108, pp. 54–59, 2012.

The remainder of the paper is organized as follows. In Section 2, we give a brief review of the classifiers available and the proposed classifiers. In Section 3, 4 and 5, we described the extraction of Co-occurrence Matrix, Tamura and Wold texture features. Section 6 gives a brief review of classifiers used for our experimentation. An evaluation study, including data set acquisition is described in Section 7. Finally conclusions are drawn in Section 8.

2 Related Work

Extraction of suitable features is an important step in any image classification. Medical images are often highly textured and hence texture analysis becomes crucial in medical image analysis. Texture perception plays an important role in the human visual system of recognition and interpretation. After feature extraction, the next important step is building, training and assessing the classifier. Researchers have tried to solve image classification problems using typical pattern recognition methods such as Support Vector Machine (SVM), Artificial Neural Networks (ANN) and Bayesian Networks (BN). Recently many data mining techniques [3] [4] have proven to be good classifiers for medical images. A number of data mining methods are implemented in the WEKA software, which contains tools for data pre-processing, classification, regression, clustering, association rules and visualization.

Proposed Work: In this paper, we classify medical images using data mining classifiers. The CT-scan brain images are classified as *inflammatory*, *stroke* or *tumor* using J48 and Random Forest classifiers based on texture features.

3 Statistical Co-occurrence Matrix

Texture information is specified by a set of gray-tone spatial-dependence matrices which are computed for various angular relationships and distances between neighboring resolution cell pairs on the image. All the textural features are derived from these angular nearest-neighbor gray-tone spatial-dependence matrices. We compute four closely related measures $P(i, j, d, \theta)$ quantized to 45^o intervals with $d=1$ from which all of our texture features are derived. Out of the equations which define a set of 14 measures of textural features [5], we have used the three most distinguishing parameters to describe the texture of an image for classification as depicted below

$$Energy = \sum_i \sum_j P(i,j)^2 \tag{1}$$

$$Entropy = -\sum_i \sum_j P(i,j) \log(P(i,j)) \tag{2}$$

$$Contrast = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j) \right\} where \mid i-j \mid = n \tag{3}$$

4 Tamura Features

Tamura [6] proposes six texture features corresponding to human visual perception: coarseness, contrast, directionality, line-likeness, regularity and roughness. Among them, it is found that the first three features correlate strongly with the human perception and are defined as follows:

Coarseness: The coarseness gives information about the size of the texture elements. The coarseness measure is calculated as follows:

1. Take averages at every point over neighborhoods of sizes of powers of two. The average over the neighborhood of size $2^k x 2^k$ at the point (x, y) is given by

$$A_k(x,y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} f(i,j)/2^{2k} \tag{4}$$

   where *f(i, j)* is the gray-level at (x, y).

2. For each point, take differences between pairs of averages corresponding to pairs of non-overlapping neighborhoods just on opposite sides of the point in both horizontal and vertical orientations given by

$$E_{k,h}(x,y) = | A_k(x+2^{k-1}, y) - A_k(x-2^{k-1}, y) | \tag{5}$$

$$E_{k,v}(x,y) = | A_k(x, y+2^{k-1}) - A_k(x, y-2^{k-1}) | \tag{6}$$

3. At each point, pick the best size which gives the highest output value: $S_{best}(x, y) = 2^k$, where k maximizes E in either direction, i.e., $E_k = E_{max} = Max(E_1, E_2, ..., E_L)$.

4. Finally, take the average of S_{best} to be a coarseness measure, given by

$$F_{crs} = \frac{1}{m \times n} \sum_{i}^{m} \sum_{j}^{n} S_{best}(i,j) \tag{7}$$

where m and n are the effective width and height of the image.

Contrast: In the narrow sense, contrast stands for picture quality. The contrast of an image is calculated by $F_{con} = \sigma/(\alpha_4)^n$ with $\alpha_4 = \mu_4/\sigma^4$ where μ_4 is the fourth moment about the mean and σ^2 is the variance and n has been experimentally determined to be 1/4.

Directionality: To calculate the directionality, the horizontal and vertical derivatives Δ_H and Δ_V are measured by Prewitt operator. The magnitude $|\Delta G|$ and the local edge direction θ are approximated as

$$|\Delta G| = \left(|\Delta_H| + |\Delta_V|\right)/2 \tag{8}$$

$$\theta = \tan^{-1}(\Delta_V / \Delta_H) + \pi/2 \tag{9}$$

The resultant θ is a real number $(0 \le \theta < \pi)$ measured counterclockwise so that the horizontal direction is 0. The desired histogram H_D can be obtained by quantizing θ and counting the points with the magnitude $|\Delta G|$ over the threshold t given by

$$H_D(k) = N_\theta(k) / \sum_{i=0}^{n-1} N_\theta(i), k = 0,1,\cdots,n-1 \tag{10}$$

where $N_\theta(k)$ is the number of points at which $(2k - 1)\pi/2n \le \theta < (2k + 1)/2n$ and $|\Delta G| \le t$. Thresholding $|\Delta G|$ by t is aimed at preventing counting of unreliable directions which cannot be regarded as edge points. In our experiments, we have used $n = 16$ and $t = 12$.

A way of measuring the directionality quantitatively from H_D is to compute the sharpness of the peaks. The approach adopted is to sum the second moments around each peak from valley to valley, if multiple peaks are determined to exist in order to compute the directionality from H_D. This measure can be defined by

$$F_{dir} = 1 - r \cdot n_p \cdot \sum_{p}^{n_p} \sum_{\phi \in w_p} \left(\phi - \phi_p\right)^2 \cdot H_D(\phi) \tag{11}$$

where n_p number of peaks,
ϕ_p p^{th} peak position of H_D
w_p range of p^{th} peak between valleys,
r normalizing factor related to quantizing levels of ϕ
ϕ quantized direction code (cyclically in modulo 180º).

5 Wold Features

The 2-D Wold theory [7] allows an image pattern to be decomposed into two mutually orthogonal components: deterministic and nondeterministic. The deterministic component is further divided into a harmonic component and an evanescent component. Let $y(m, n)$ be a real valued, regular, and homogeneous random field uniquely represented by the decomposition given by

$$y(m,n) = w(m,n) + p(m,n) + g(m,n) \tag{12}$$

where $w(m, n)$ is purely nondeterministic,
$p(m, n)$ is the harmonic random field and
$g(m, n)$ is generalized evanescent field.

To extract the feature set characterizing the harmonic structure of an image, first the image Discrete Fourier Transform magnitudes are computed. Then, the local maxima of the magnitudes are found by searching 8x8 neighborhood of each frequency sample. Next, the local maxima are examined for harmonic peaks. A local maximum is a harmonic peak if and only if its frequency is either fundamental or harmonic. A fundamental is defined as a frequency which can be used to linearly express the frequencies of some other local maxima. A harmonic is a frequency which can be expressed as a linear combination of some fundamentals. Starting from the one with the lowest frequency and in ascending order of their frequencies, each local maximum is then checked first for its harmonicity if its frequency can be expressed as a linear combination of the existing fundamentals. Each local maximum is then for its

fundamentality if the multiples of its frequency, combined with the multiples of existing fundamentals, coincide with the frequency of another local maximum. In our work, only the ten largest ones are considered for each image.

6 Classifiers

Classification is a data mining or machine learning technique used to predict group membership for data instances. Machine learning refers to a system that has the ability to automatically learn knowledge from experience and other ways. Decision trees are supervised algorithms which recursively partition the data based on its attributes, until some stopping condition is reached. This recursive partitioning gives rise to a tree-like structure. In Decision trees, classification rules learned by them can be easily obtained by tracing the path from the root node to each leaf node in the tree. Decision trees are very efficient even with large volumes of data. This is due to the partitioning nature of the algorithm. The two data mining classifiers used in our experiment are: J48 and Random Forest of WEKA data mining tool.

J48 Decision Tree Classifier: J48 is a java implementation of C4.5 algorithm, which does not require discretization of numeric attributes. The Decision tree classifies the given data item using the values of its attributes. The decision tree is initially constructed from a set of pre-classified data. The main approach is to select the attributes, which best divides the data items into their classes. According to the values of these attributes the data items are partitioned. This process is recursively applied to each partitioned subset of the data items. The process terminates when all the data items in the current subset belongs to the same class.

Random Forest Classifier: Random Forest is a machine learning technique that builds a forest of classification trees wherein each tree is grown on a bootstrap sample of the data, and the attribute at each tree node is selected from a random subset of all attributes. The final classification of an individual is determined by voting over all trees in the forest.

7 Results and Discussions

The aim is to determine discriminating features and a suitable classifier for classification of CT-scan brain images. We have considered the J48 Decision Tree and Random Forest of WEKA classification methods to classify CT scan brain images into inflammatory, stroke and tumor. Our work is carried out on a database of size 184 samples of 256x256 GIF format and is obtained from brain atlas of Harvard University. Among them, 60 samples are used as training set, 20 from each class and the rest 124 samples are used for testing purpose. Among 124 samples used for testing, 28 are inflammatory, 55 are stroke, and 41 are tumor. For each image in the database, the Haralick, Tamura and Wold texture features are extracted. The extracted features are transformed into the format that is suitable for applying machine learning algorithms. The performance comparison of texture features and the classifiers is tabulated using precision and recall measures shown in Table 1 and 2 respectively.

Table 1. Precision

Class	WJ48	TJ48	HJ48	WRF	TRF	HRF
Inflammatory	0.81	0.85	0.9	0.94	1.0	1.0
Stroke	0.85	0.76	0.9	1.0	1.0	1.0
Tumor	0.83	0.93	0.9	0.86	0.95	1.0

Table 2. Recall

Class	WJ48	TJ48	HJ48	WRF	TRF	HRF
Inflammatory	0.85	0.85	0.9	0.85	0.95	1.0
Stroke	0.9	0.95	0.9	1.0	1.0	1.0
Tumor	0.75	0.7	0.9	0.95	1.0	1.0

8 Conclusions

In this paper, we have described the classification techniques for CT scan brain images. Three texture features and two classifiers are compared for classification of CT scan brain images from a large database of images. The texture features compared are Haralick, Tamura and Wold harmonic peaks which are implemented using JAVA. We have used data mining classifiers J48 and Random Forest of WEKA Data Mining tool for our experiment. Both the features and the classifiers are compared based on Precision and Recall measures as shown in Tables 1 and 2 respectively. Haralick features combined with Random Forest classifier is found to give the best results for classification of CT-scan brain images.

References

1. John, S., Ioannis, V., et al.: Computer Aided Diagnosis based on Medical Image Processing and Artificial Intelligence Methods. Nuclear Instruments and Methods in Physics Research A 569, 591–595 (2006)
2. Zheng, B.: Computer-Aided Diagnosis in Mammography using Content-Based Image Retrieval Approaches: Current Status and Future Perspectives. Algorithms, 828–849 (2009)
3. Antonie, M.-L., Zaane, O.R., Coman, A.: Application of Data Mining Techniques for Medical Image Classification. In: Proc. of the 2nd Int. Workshop on Multimedia Data Mining, San Francisco, USA (August 26, 2001)
4. Dua, S., Jain, V., Thompson, H.W.: Patient Classification using Association Mining of Clinical Images, 978-1-4244-2003-2/08 2008 IEEE
5. Haralick, R., Shanmugam, K., Dinstein, I.: Textural Features for Image Classification. IEEE Trans. on Systems, Man and Cybernetics 3(6), 610–621 (1973)
6. Tamura, H., Mori, S., Yamawaki, T.: Textural Features Corresponding to Visual Perception. IEEE Trans. on Systems, Man and Cybernetics (June 1978)
7. Liu, F., Picard, R.W.: Periodicity, Directionality and Randomness: Wold Features for Image Modeling and Retrieval. IEEE Trans. on Pattern Analysis and Machine Intelligence 18(7) (July 1996)