PERFORMANCE EVALUATION OF CONTENT BASED IMAGE RETRIEVAL FOR MEDICAL IMAGES

SASI KUMAR. M
Research Scholar, Department of CSE, Sathyabama University
Chennai, INDIA
msksasi@yahoo.com

Dr. Y.S. KUMARASWAMY
Sr. Professor & HOD, Department of MCA, Dayananda Sagar College of Engineering
Bangalore, INDIA
yskldswamy@yahoo.co.in

Abstract:

Content-based image retrieval (CBIR) technology benefits not only large image collections management, but also helps clinical care, biomedical research, and education. Digital images are found in X-Rays, MRI, CT which are used for diagnosing and planning treatment schedules. Thus, visual information management is challenging as the data quantity available is huge. Currently, available medical databases utilization is limited image retrieval issues. Archived digital medical images retrieval is always challenging and this is being researched more as images are of great importance in patient diagnosis, therapy, medical reference, and medical training. In this paper, an image matching scheme using Discrete Sine Transform for relevant feature extraction is presented. The efficiency of different algorithm for classifying the features to retrieve medical images is investigated.

Keywords: Content based Image Retrieval (CBIR), Medical images, Discrete Sine transform (DST), Boosting, J48.

1. Introduction

Content-based image retrieval (CBIR) is a highly complicated computer vision research area. Availability of steadily increasing visual and multimedia data and internet development highlight the need for the creation of thematic access systems offering more than simple text-based queries/requests based on matching exact database fields. Many programs/tools were developed to formulate/execute visual or audio content queries and help browse multimedia repositories. But no breakthrough was achieved regarding large varied databases with different documents having varied characteristics. Answers to questions regarding speed, semantic descriptors or objective image interpretations still remain unanswered.

Database images extract features after automatic pre-processing; generating feature vectors in content based image retrieval (CBIR) systems. Feature vectors are stored in feature databases and images classified. The query image is also similarly pre-processed for features extraction. Based on such similarity, specific database images are retrieved. Image retrieval plays a big role in handling visual information in medical applications. [1]. Image retrieval system depends on a multi-dimensional feature vector through use of extracted image information, computing similarity measures and correct database image identification with lowest distance metrics as regards the query image. Transform methods are used in image processing as many coefficients can be ignored to reduce feature vector size.

Low level features are used for Image Retrieval [2, 3], as all models combine low-level features to define a distance metric quantifying similarities between image models. A shortcoming of this procedure is low-level image features not always capturing an image similarly human perception. To state, semantic image content is tough for feature extraction with low level image features alone, this being called the semantic gap problem [4]. Medical Image Retrieval systems are not like regular image retrieval systems. For one, retrieval is associated with local pathological conditions where retrieval based on global signatures will be useless for medical databases. Converting images from spatial to frequency domain is currently used image retrieval procedure available in literature [5, 6].

Medical Image Retrieval systems are different from regular image retrieval systems in many ways. For one, the retrieval takes place with regard to pathological conditions that are local wherein retrieval based on global signatures would be senseless when used for medical databases.

Image retrieval plays a basic role in handling much visual information in medical applications [1]. Image retrieval system performance depends on the multi-dimensional feature vector formed through use of information extracted from images, computing of the similarity measures and accurate database image identification with lowest distance metrics as regards query images. Transform methods are used in image processing as many coefficients are ignored to reduce feature vector size.

Images, especially digital images, are produced in large amounts in the medical field for diagnosis and therapy. Geneva university hospital's radiology department produced more than 12,000 images a day. Cardiology currently produces the second highest producer of digital images specially videos of cardiac catheterization (2000 images for each case). Images have many uses in health care and biomedical research, but despite widespread use, little is known about how users search for and/or manage them. Two analyses reveal that image use is usually connected to the user's "role," like clinician, educator, and researcher. It is important to understand user needs and also provide systems to meet those needs because image collections/search interfaces proliferate on the internet and closed networks.

Archived digital medical images retrieval is being researched more as images are of great importance in clinical diagnosis. In this paper, an image matching scheme using Discrete Sine Transform for relevant feature extraction is presented. The efficiency of different algorithm for classifying the features to retrieve medical images is investigated. This paper is organized as follows: Section 2 reviews some of the related works available in the literature. Section 3 describes feature extraction and the classifier methodology, Section IV explains the experiment and results obtained. Section V concludes the paper.

2. Related Works

Ramamurthy [7] presented a medical images retrieval approach from large medical databases, requiring pre-processing, feature extraction, classification, retrieval and indexing steps to develop an efficient system. In this work, image segmentation was carried out for pre-processing, while basic shape feature was extracted using canny edge detection algorithm for feature extraction, while for classification, K-means classification algorithm was used. For image retrieval, Euclidian distance method values were calculated between query and database images. This work aims to provide a medical image retrieval system for medical diagnosis.

Quellec et al [8] proposed a CBIR method for diagnosis in medical fields. In this, images are indexed in a generically, without extracting domain-specific features: a signature is built into each image from wavelet transform. These signatures characterize wavelet coefficient distribution in each decomposition subband. A distance measure compares two image signatures and retrieves most similar images from the database when a physician submits a query image. To retrieve relevant images from a medical database, signatures and distance measure should be related to medical image interpretation. Subsequently the system requires much freedom to tune it to any pathology with image modality being introduced. The scheme proposed using a custom decomposition scheme to adapt the wavelet basis with lifting scheme framework. Weights are introduced between subbands. All parameters are tuned by an optimization procedure, using database medical image grading to define performance measures. System assessment is through two medical image databases: one for diabetic retinopathy follow up and another for mammography screening, as well as a general purpose database. Results are promising: a mean precision of 56.50%, 70.91% and 96.10% are achieved for the three databases, when the system returned five images.

Baranidharan, et al [9] addressed medical images retrieval problem from a multi-varied database. An algorithm based on energy information was proposed for gaining Hilbert Transform for medical images classification based on imaging modalities and body parts. Neural networks were used for image classification. This paper exploited image's spatial information to decide classification result and proposed a novel medical image classification method 2D-I Neural Networks using Fuzzy Logic for data pre-processing. The proposed Neural Network (NN) algorithm is a modified Elman network consisting of a hidden layer with Tanh activation function. Results show that classification accuracy improves when compared to standard Multi-Layer Perceptron Neural Network.

In content-based image retrieval, a prerequisite for successful retrieval is extracting many distinguishing features which describe important image content characteristics. Parameters underlying image segmentation and feature extraction should be set to ensure successful retrieval. Han et al [10] presented a parameter tuning method using simulated annealing to dynamically adjust important parameter values used in customized image processing algorithms to improve retrieval performance for high resolution CT lung images in computer-aided diagnosis. Notable improvement using $F\beta$ measure among five modules is from 0.56 to 0.81, which is a 44.64% increase (p=0.022). This method improves retrieval performance in many applications in medical imaging informatics.

Rajkumar et al., [11] presented a 2 step medical image retrieval framework for similar image retrieval from various features. An image subset was chosen through a wavelet filtering process and the image decomposed into 6 levels using wavelet transforms with extracted energies. Euclidean distance matched similar query and database images with dimensions being reduced through PCA use. Finally, calculated eigen vectors and similarity measures applied ensured efficient medical image retrieval resulting in improved retrieval accuracy due to reduced search space efficiency. Experiments with 200 medical images proved the proposed method's accuracy regarding precision and recall rate.

3. Methodology

3.1. Discrete Sine Transform (DST)

Each image's feature vector was extracted using Discrete Sine Transform (DST). Pixels one length away from each other was being chosen. The pseudo DST algorithm is given Figure 1:

1. Compute Image size MxN
2. For each alternate value 'i' in array M and array size less than M or M+1
3. For each alternate value 'j' in array N and array size less than N or N+1
4. Compute DST (array [xi, yj])
5. Store computed value in one dimensional array
6. Repeat from step 1 till all images are computed

Fig.1. Pseudo code for Discrete Sine Transform (DST)

DST is similar to discrete Fourier transform (DFT) the difference being use of only real numbers. Discrete sine transform is represented by

$$S_k = p_K \sum_{n=0}^{N-1} x_n \sin \frac{\pi(n+\frac{1}{2})(k+1)}{N} \quad k = 0,1,2,.........., N-1$$

$$p_k = \sqrt{\frac{2-\delta_{k,0}}{N}}$$

where x_n - original vector on N real numbers.

δ - Kronecker delta.

DST operates on real data with odd symmetry and hence the output data are shifted by half a sample [12]. The inverse of Discrete sine transform is given as

$$S_k^{III} = p_K \sum_{n=0}^{N-1} x_n q_n \sin \frac{\pi(n+1)(k+\frac{1}{2})}{N} \quad k = 0,1,2,.........., N-1$$

$$p = \sqrt{\frac{2}{N}}$$

$$q_n = \sqrt{\frac{1}{1+\delta_{n,0}}}$$

DST is preferred to Fast Fourier transform because of its simplicity and lowered time to compute medical image coefficients.

3.2. Classifiers

3.2.1. Naïve Bayes Classifier

The Naïve Bayes classifier works on a simple, intuitive concept. It is also seen that Naïve Bayes outperforms many comparatively complex algorithms, making use of variables in data samples, by observing them individually and independently [13, 14].

Naïve Bayes is a probability model based popular supervised learning classifier. Bayes rule for supervised learning is represented for an unknown target function f: X →Y, or equivalently P(Y|X). Assuming Y is a boolean-valued random variable, and X is a vector containing 'n' boolean attributes. That is X =(X_1, X_2, . . ., X_n) , where X_i is the Boolean random variable denoting the i^{th} attribute of X. Applying Bayes rule, P(Y = yi|X) can be represented as

$$P\left(Y = y_i \middle| X = x_k\right) = \frac{P\left(X = x_k \middle| Y = y_i\right) P\left(Y = y_i\right)}{\sum_j P\left(X = x_k \middle| Y = y_j\right) P\left(Y = y_j\right)}$$

where y_m denotes the m^{th} possible value for Y, x_k denotes the k^{th} possible vector value for X, and where denominator summation is over all legal values of the random variable Y. A method to learn P(YjX) is using training data to estimate P(X|Y) and P(Y) which are then used together with Bayes rule to determine $P(Y|X = x_k)$ for any new instance x_k.

3.2.2. Bagging

Breiman [15] introduced bagging including bootstrap and aggregation methods to improve unstable classification methods accuracy. For bagging, X bootstrap datasets, with x random selected examples, generated, with replacement from Y, decision tree is built using X samples. The predicted new sample class is obtained by majority vote. New examples are checked against X decision trees and results are noted. Though, simple single decision tree interpretation is lost, Bagging improves classification rule accuracy. This study uses bagging with J48 and decision stump.

For enhancing a given learning algorithm's accuracy, "Boosting" is resorted to. Boosting is a machine learning method which finds and combines rough rules to produce accurate classification. The boosting algorithm finds rules repeatedly through use of various training set subsets from the base learning algorithm. AdaBoost [16] is a common boosting technique.

3.3.3. J48

J48 is slightly modified C4.5 in WEKA. The C4.5 algorithm generates a classification-decision tree for a specific data-set through recursive data partitioning. Decision tree is grown using Depth-first strategy. The algorithm considers all tests that split a data set and selects a test that leads to best information gain. For every discrete attribute, one test with outcomes as many as distinct attribute values is considered. Binary tests involving every distinct values of the attribute are considered for each continuous attribute. To gather all binary tests entropy gain efficiently training data set of the node being considered is sorted for continuous attribute values with entropy gains of binary cut based on each distinct values being calculated in one sorted data scan, this process being repeated for every continuous attributes [17].

The J48 Decision tree classifier follows a simple algorithm. To classify a new item, a decision tree based on the attribute values of the available training data should first be created. When it encounters a set of items (training set) it identifies the discriminating attribute of various instances clearly. Among this feature's values, if there is a value which has no ambiguity, for which data instances within its category have similar value for target variable, then that branch is terminated the obtained target value being assigned to that branch.

3.3.4. Multilayer perceptron (MLP)

Multilayer perceptron (MLP) is a popular supervised learning network consisting of a input layer, one or more hidden layer and an output layer. Intra layer connections are formed by connecting every node from a layer to next layer's neurons. During training, each connection's scalar weight is adjusted. The outputs nodes produce output nodes. Feature vector x is input at input layer with output representing a discriminator between its class and other classes. Training examples, in training, are fed and the predicted outputs computed. The output and target output are compared and measured error is propagated back through network and weights adjusted [18, 19].

The training set of size m is represented as $T_M = \{(x_1, y_1), ..., (x_m, y_m)\}$ where $x_i \in R^a$ are the input vectors of dimension *a* and $y_i \in R^b$ are output vectors of dimension *b* and R represents a real numbers set. Let f_w represent the function with weight *w* for the neural network. Supervised learning adjusts weights so that:

$$f_w(x_i) = y; \forall (x_i, y_i) \in T_M$$

3.3.5. Support vector machine (SVM)

Support vector machine (SVM) is a linear machine constructing a hyperplane as a decision surface [20]. And it is based on structural risk minimization method; the error rate is bounded by sum of training-error rate and a term depending on Vapnik-Chervonenkis (VC) dimensions. SVM provides good generalization on pattern classification. SVM algorithm performance is based on inner-product kernel between a "support vector" x_i and input vector drawn vector *x*.

SVM uses mapping to the larger space to compute cross products with variables in original space lightening computational load. In larger space, cross products are defined using a kernel function *K*(*x*,*y*) which is selected to suit the problem domain. Cross products with a vector in space if constant is used define hyper planes [21].

Hyperplane defining vectors are linear combinations with parameters α_i of feature vectors which occur in a data base. After hyperplane selection, feature plane points x are defined by:

If $K(x,y)$ becomes small when y grows further from x, closeness degree is given by the sum measures of closeness of test point x to corresponding data base point x_i.

The above method measures closeness of each test point to data points originating from data sets awaiting discrimination. As points set mapped can be quite convoluted, complex discrimination happens between sets which are not convex in original space.

SVMs have performed well as a learning algorithm in the past and generally perform well on a various classification problems. They also allow rapid classification from trained models and can handle very high-dimensional input vectors.

The implementation's error function is given in the following equation

$$\tfrac{1}{2} w^T w + C \sum_{i-1}^{N} \xi_i + C \sum_{i-1}^{N} \xi^{\bullet} i$$

Which can be minimized to

$$w^T \phi(x_i) + b - y_i \le \varepsilon + \xi_i^{\bullet}$$

$$y_i - w^T \phi(x_i) - b_i \le \varepsilon + \xi_i$$

$$\xi_i, \xi_i^{\bullet} \ge 0, i = 1, \ldots, N$$

where C is capacity constant, w vector of coefficients, b a constant and ξ_i parameters for handling non separable data (inputs). The index i labels N training cases.

4. Results and Discussion

A program was developed in LabVIEW to handle multiple input images, outputting co-efficient of DST as a comma separated values. Various MRI scan images and noisy images from the dataset were used for evaluation. Sample dataset images are seen in Figure 2. Many were rotated by 90 degrees to simulate real time database search. The experiment used 4 medical image types with different noise degrees. 57 MRI scan images were the inputs and classified by boosting with J48 and decision stump. Around 60% data was classified as training set with the remaining being a test set.

The classification accuracy achieved by various classifiers are tabulated in Table 1 and shown in Figure 3.

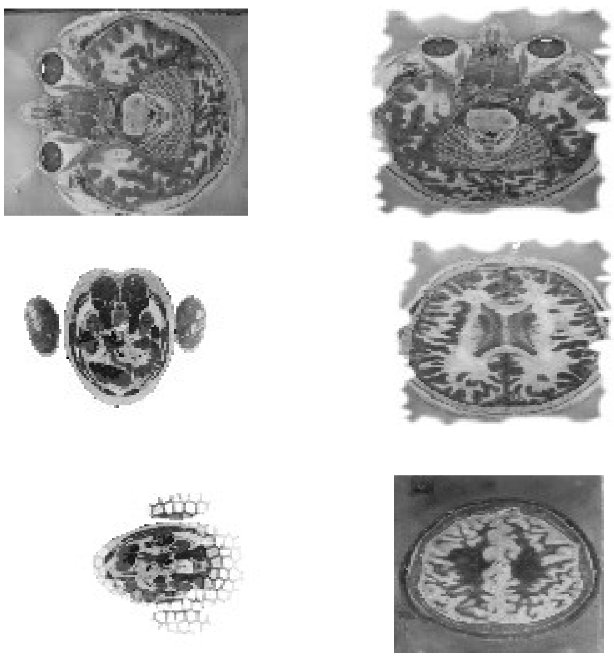

Fig. 2. Sample images used in the medical retrieval system

Table 1: Classification Accuracy achieved by various Classifiers

Classifier Used	Classification Accuracy %
Naïve Bayes	86.96
Bagging with J48	80.7
Support vector machine	89.65
Multilayer perceptron	**93.1**

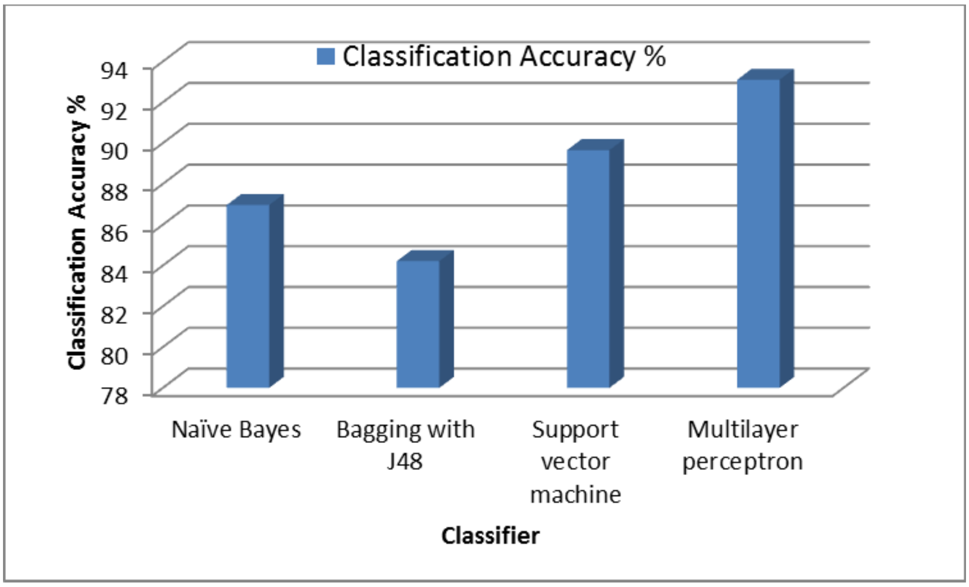

Fig. 3. Classification Accuracy achieved

It is evident from the graphs that the Multilayer perceptron achieve better classification accuracy than bagging with J48 and SVM. Further investigations are required to evaluate soft computing techniques in effort to improve the classification accuracy.

5. Conclusion

Image classification is an important step in image retrieval as it saves time while searching for images in a huge volume of database. Classification is identification of the image's different regions by which system retrieval efficiency is improved. This paper investigates classification accuracy for different classifiers. It extracted features with Discrete Sine Transform (DST) and extracted features were trained and classified with Naïve Bayes, bagging with j48, SVM and MLP. Results demonstrate that the Multilayer perceptron achieves better classification accuracy than naïve bayes, bagging with J48 and SVM.

References

[1] H. Muller, N. Michoux, D. Bandon, and A. Geissbuhler. A review of contentbased image retrieval systems in medical applications-clinical benefits and future directions. In International Journal of Medical Informatics, volume 73, pages 1–23, 2004.
[2] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors", IEEE Trans.Circuits S
[3] F. Jing, M. Li, H. J. Zhang, and B. Zhang, "An efficient and effective region-based image retrieval framework", IEEE Trans. Image Processing, vol. 13, no. 5, pp. 699–709, May 2004.
[4] Y. Rui, T. S. Huang, and S. F. Chang, "Image retrieval: Current techniques, promising directions, and open issues", J. Vis. Commun. Image Represen., vol. 10, pp 39–62, Mar. 1999.
[5] P. Kelly, T. Cannon, and D. Hush. Query by image example: The CANDID approach. In Storage and Retrieval for Image and Video Databases III, pages 238–248. SPIE Vol. 2420, 1995.
[6] H. B. Kekre, Dhirendra Mishra,"Digital Image Search & Retrieval using FFT Sectors of Color Images" published in International Journal of Computer Science and Engineering (IJCSE) Vol. 02,No.02,2010,pp.368-372.
[7] Ramamurthy, B., & Chandran, K. R. (2011). Cbmir: Shape-Based Image Retrieval Using Canny Edge Detection And K-Means Clustering Algorithms For Medical Images. International Journal of Engineering Science and Technology (IJEST) Mar.
[8] Quellec, G., Lamard, M., Cazuguel, G., Cochener, B., & Roux, C. (2010). Wavelet optimization for content-based image retrieval in medical databases. Medical image analysis, 14(2), 227-241.
[9] Baranidharan, T., & Ghosh, D. K. (2012). A Two Dimensional Image Classification Neural Network for Medical Images. European Journal of Scientific Research, 74(2), 286-291.
[10] Han, J. G., & Shyu, C. R. (2010). Improving Retrieval Performance in Medical Image Databases Using Simulated Annealing. In AMIA Annual Symposium Proceedings (Vol. 2010, p. 276). American Medical Informatics Association.
[11] K. Rajakumar, "An Integrated Approach for Medical Image Retrieval Using PCA and Energy Efficient Wavelet Transform", European Journal of Scientific Research ISSN 1450-216X Vol.51 No.3 (2011), pp.340-348.

[12] Yip, P., & Rao, K. (1980). A fast computational algorithm for the discrete sine transform. Communications, IEEE Transactions on, 28(2), 304-307.
[13] Moore, A. W. (2004). Naïve Bayes Classifiers. Statistical Data Mining Tutorials.
[14] Ridgeway, G., Madigan, D., Richardson, T., & O'Kane, J. (1998, August). Interpretable boosted naive Bayes classification. In Proceedings of the fourth international conference on knowledge discovery and data mining (pp. 101-104). AAAI Press.
[15] Breiman, L. (1996). Bagging predictors. Machine learning, 24(2), 123-140.
[16] Rätsch, G., Onoda, T., & Müller, K. R. (2001). Soft margins for AdaBoost. Machine learning, 42(3), 287-320.
[17] Loh, W. Y. (2009). Improving the precision of classification trees. The Annals of Applied Statistics, 3(4), 1710-1737.
[18] Leisch, F., Jain, L. C., & Hornik, K. (1998). Cross-validation with active pattern selection for neural-network classifiers. Neural Networks, IEEE Transactions on, 9(1), 35-41.
[19] Widrow, B., & Lehr, M. A. (1990). 30 years of adaptive neural networks: Perceptron, madaline, and backpropagation. Proceedings of the IEEE, 78(9), 1415-1442.
[20] Vandewalle, J. (1999). Least squares support vector machine classifiers. Neural processing letters, 9(3), 293-300.
[21] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.