

Automated Image Annotation for Semantic Indexing and Retrieval of Medical Images

Krishna A N

Associate Professor

Department of Computer Science and Engineering
S J B Institute of Technology, Bangalore, India

B G Prasad, Ph. D

Professor and Head

Department of Computer Science and Engineering
B N M Institute of Technology, Bangalore, India

ABSTRACT

Medical image retrieval to search for clinically relevant and visually similar images depicting suspicious lesions have been attracting research interest. Content-based image retrieval (CBIR) is an important alternate and complement to traditional text-based retrieval using keywords. We have implemented CBIR system based on effective use of texture information within the images obtained by statistical cooccurrence matrix method. Also, the method is improved by bridging the semantic gap between low-level visual features and the high-level semantic concepts using automated image annotations. In this paper, we have proposed a classification-based multi-class multi-label semantic model and the corresponding learning procedure to address the problem of automatic image annotation using J48 decision tree classifier and show its application to medical image retrieval. Hash structure is used to index images. Euclidian distance measure is used for similarity measurement. Both the methods are compared using precision and recall measures. Semantic indexing is shown to outperform CBIR for MR-T2 axial brain images.

General Terms:

Annotation, Retrieval

Keywords:

Cooccurrence matrix, Decision tree classifier, Semantic indexing.ifx

1. INTRODUCTION

The term image retrieval means finding similar images (may not be exact) from a large image archive with the help of some key attributes associated with the images or features inherently contained in the images. In the medical domain, the ultimate goal of image retrieval is to provide diagnostic support to physicians or radiologists by displaying relevant past cases, along with proven pathologies as ground truth. However, medical image retrieval can also be used as a training tool for medical students and residents in education, follow-up studies for detecting the growth of tumors, and for research purposes [1].

A large amount of research has been carried out on image retrieval (IR) in the last two decades. In general, the research efforts on IR can be divided into two types: text-based retrieval and CBIR. In the traditional text-based retrieval, images are manually annotated by humans and then indexing and retrieval is performed based on the annotated textual descriptions. Annotating a large amount of images manually is time-consuming, tedious and expensive and is often subjective, context-sensitive and incomplete. In CBIR, images are indexed and retrieved based on automatically extracted visual content features such as color, texture and shape. However, the research

effort has shown that there is a significant gap between the low-level visual features and the semantic concepts of images interpreted by the humans.

Though many sophisticated descriptors have been designed to describe color, texture and shape features, these alone do not adequately model the image semantics and have many limitations when applied to broad image categories. Extensive experiments on CBIR systems shows that low-level contents often fail to describe the high-level semantic concepts in users mind. Therefore, the performance of CBIR is still far from users expectations. In order to improve the retrieval accuracy of CBIR systems, research focus has been shifted from designing sophisticated low-level visual feature descriptors to narrowing down the semantic gap between the low-level visual features and the richness of human semantics. Ying Liu [2] survey identified five major categories of the state-of-the-art techniques to narrow down the semantic gap: (1) using object ontology to define high-level semantics; (2) using machine learning (either supervised or unsupervised) methods to extract (or learn) associations between low-level features and high-level concepts; (3) using relevance feedback to learn users intention; (4) generating semantic model to support image retrieval with high-level semantics; (5) fusing the evidences from HTML text and the visual content of images for WWW image retrieval.

The typical method of bridging the semantic gap is through the automated image annotation (AIA), which is a process of automatically assigning meta-data in the form of textual description or keywords to a digital image. Images are then indexed and retrieved similar to text-based retrieval. The main idea of AIA techniques is to learn semantic concept models automatically from a large number of training image samples, and use them to label new images. This method of AIA can be regarded as a type of multi-class image classification with a very large number of classes - as large as the vocabulary size. Therefore, AIA can be considered as a multi-class object recognition problem which is an extremely challenging task and still remains an open problem in computer vision. Because an image is an unstructured array of pixels, the first step in semantic understanding is to extract efficient and effective visual features from these pixels. Appropriate feature representation significantly improves the performance of the semantic learning techniques. Once images are represented with low-level features using either global methods or region methods, high level semantics can be learned from image samples. Assuming semantically labeled image samples are collected and represented with low-level features, a machine learning algorithm can then be trained using the feature-to-semantic label matching. Once trained, the algorithm can then be used to annotate new image samples.

2. RELATED WORK

A variety of learning methods have been applied for automatically annotating images, including co-occurrence model [3], machine translation model [4], latent space approaches [5], graphic models [6], classification approaches [7] - [9] and relevance language models [10] [11]. Mori proposed the co-occurrence model [3] in which the co-occurrence counts between words and image features are collected and used to predicate annotated words for images. Duygulu et al. [4] described images using a vocabulary of blobs. First, regions are created using a segmentation algorithm like normalized cuts. For each region, features are computed and then blobs are generated by clustering the image features for these regions across images. Each image is generated by using a certain number of these blobs. The Translation Model applies one of the classical statistical machine translation models to translate from the set of blobs forming an image to the set of keywords of an image. It views image annotation as a process of translation from visual language to texts and collects the co-occurrence information by the estimation of translation probabilities. Another way of capturing co-occurrence information is to introduce latent variables to link image features with words.

Standard latent semantic analysis (LSA) and probabilistic latent semantic analysis (PLSA) are applied to automatic image annotation [5]. Blei [6] extended the Latent Dirichlet Allocation (LDA) model and proposed a correlation LDA model which relates words and images. This model assumes that a Dirichlet distribution can be used to generate a mixture of latent factors. This mixture of latent factors is then used to generate words and regions. Expectation-Maximization is used to estimate this model. The classification approaches for automatic image annotation treat each annotated word as an independent class and create a different image classification model for every word. Work such as linguistic indexing of pictures [9], image annotation using SVM [8] and Bayes point machine [7] belong to this category. Recently, relevance language models [10] - [12] have been successfully applied to automatic image annotation. The essential idea is to first find annotated images that are similar to a test image and then use the words shared by the annotations of the similar images to annotate the test image.

Wei Li et al. [13] proposed a Maximum Entropy Model-based approach to the task of automatic image annotation. In the phase of training, a basic visual vocabulary consisting of blob-tokens to describe the image content is generated at first; then the statistical relationship is modeled between the blob-tokens and keywords by a Maximum Entropy Model constructed from the training set of labeled images. In the phase of annotation, for an unlabeled image, the most likely associated keywords are predicted in terms of the blob-token set extracted from the given image. Munirathnam et al. [14] proposes methods to use a hierarchy defined on the annotated words derived from a text ontology to improve automatic image annotation and retrieval. In particular, the hierarchy is used to generate visual vocabulary for representing images and as a framework for the proposed hierarchical classification approach for automatic image annotation.

In this paper, we propose a classification-based multi-class multi-label semantic model and the corresponding learning procedure to address the problem of automatic image annotation and show its application to medical image retrieval. Classification-based methods treat each keyword (concept) as a separate class, build the classifier model for each concept and employ trained classifiers to annotate an input image based on classification results. The common machine learning tools include Support Vector Machines (SVM) [15]-[19], Artificial Neural Networks (ANN) [20]-[28] and many data mining

techniques [29]-[32], which have recently proven to be good classifiers for medical images. Dengsheng Zhang et al. [33] give a detailed review of the above machine learning tools.

3. J48 DECISION TREE CLASSIFIER

A decision tree (DT) is a multi-stage decision making or classification tool. Depending on the number of decisions made at each internal node of the tree, a DT can be called a binary or an n-ary tree. Different from other classification models whose input-output relationships are difficult to describe, the input-output relationship in a DT can be expressed using human understandable rules, e.g., if-then-rules. A DT is trained using a set of labelled training samples, each represented by a number of attributes. During training, a DT is built by recursively dividing the training samples into non-overlapping sets and every time the samples are divided, the attribute used for the division is discarded. The procedure continues until all samples of a group belong to the same class or the tree reaches its maximum depth, when no attribute remains to separate them. This recursive partitioning gives rise to a tree-like structure. Due to the partitioning nature of the algorithm, DTs are very efficient even with large volumes of data. The tree has two types of nodes: internal and leaf node. Each internal node is associated with a decision governed by a certain attribute which divides the training samples most effectively. Each leaf node represents the outcome (class) of the majority samples that follow the path from the root of the tree to the corresponding leaf. The leaf nodes can be expressed with unique if-then-else rules. To label a new sample, the tree is traversed from the root node to a leaf node using the attribute value of the new sample. The decision of the sample is the outcome of the leaf node where the sample reaches. The most important feature of DT classifier is its capability to break down a complex decision making process into a collection of simpler decisions, thus providing a solution, which is often easier to interpret [34].

Several DT algorithms exist in the literature, including ID3, C4.5 and CART. These DTs differ by the type of attributes, the attribute selection criteria, the outcome, etc. ID3 is the simplest DT algorithm that works only with discretized attributes. On the other hand, C4.5 and CART can work with both discretized and continuous attributes. In ID3 and C4.5, the number of children of an internal node is equal to the number of attribute values of the selected attribute at that node. On the contrary, CART always splits an internal node into two children. Therefore, CART usually generates a bigger tree and takes more time than ID3 or C4.5. While ID3 and C4.5 are used only for classification, CART is used for both classification and numerical prediction. A number of data mining techniques are implemented in WEKA software, which contains tools for data pre-processing, classification, regression, clustering, association rules and visualization. WEKA written in Java runs on almost any platform and is available on the web at www.cs.waikato.ac.nz/ml/weka [35] [36]. We have used *J48 decision tree* classifier of WEKA for the classification.

J48 is a java implementation of C4.5 algorithm, which does not require discretization of numeric attributes in contrast to ID3. C4.5 builds the tree from a set of data items using the best attribute to test in order to divide the data item into subsets and then it uses the same procedure on each subset recursively. The main problem is deciding the attribute, that best partitions the data into various classes. ID3 and C4.5 uses the information theoretic approach to solve this problem. Information theory uses the concept of entropy, which measures the impurity of data items. The value of entropy is low when the class distribution is uneven and high when the class distribution is more. Information gain is a measure on the utility of each attribute in classifying the data items. It is measured using the entropy value. Information

gain measures the decrease of the weighted average impurity (entropy) of the attributes compared with the impurity of the complete set of data items. Therefore, the attributes with the largest information gain are considered as the most useful for classifying the data items. A series of improvements to ID3 including methods to deal with continuous attributes, missing values, over fitting of data, noisy data, computation efficiency and generating rules from trees are incorporated to make it a more widely accepted system for decision tree induction which is called C4.5 [37].

4. MULTI-CLASS MULTI-LABEL ANNOTATION MODEL

In this section, we describe the multi-class, multi-label framework for automatic image annotation. The framework includes three components: 1) feature representation; 2) building the classifier models for the individual concept (or vocabulary) using labelled training set of images and 3) propagating the labels of the training images to the query images.

4.1 Feature Representation

Images are represented either by global features, block-based local and spatial properties, or region-based local features. Inaccurate segmentation/blocking may make the feature representation imprecise and thereby decreasing annotation accuracy. We use global features for image representation: *entropy*, *energy* and *contrast* and perform classification-based automatic image annotation using the global texture features extracted from statistical cooccurrence matrix. Texture information is specified by a set of gray-tone spatial-dependence matrices that 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.

Suppose an image to be analyzed is rectangular and has N_x resolution cells in the horizontal direction and N_y resolution cells in the vertical direction. Suppose that the gray tone appearing in each resolution cell is quantized to N_g levels. Let $L_x = 1, 2, \dots, N_x$ be the horizontal spatial domain, $L_y = 1, 2, \dots, N_y$ be the vertical spatial domain, and $G = 1, 2, \dots, N_g$ be the set of N_g quantized gray tones. The set $L_y \times L_x$ is the set of resolution cells of the image ordered by their row-column designations. The image I can be represented as a function which assigns some gray tone in G to each resolution cell in $L_y \times L_x$; $I: L_y \times L_x \rightarrow G$.

Four closely related measures from which the texture features we have used are derived using angular nearest-neighbor gray-tone spatial-dependence matrices: $P(i, j, d, 0^\circ)$, $P(i, j, d, 45^\circ)$, $P(i, j, d, 90^\circ)$ and $P(i, j, d, 135^\circ)$. We assume that the texture-context information in an image I is contained in the "overall" or "average" spatial relationship which the gray tones in image I have to one another. More specifically, this texture-context information has been adequately specified by a matrix of relative frequencies P_{ij} with which two neighboring resolution cells, one with gray tone i and the other with gray tone j separated by distance d occur on the image. Such matrices of gray-tone spatial-dependence frequencies are a function of the angular relationship between the neighboring resolution cells as well as a function of the distance between them. Formally, for angles quantized to 45° intervals, the unnormalized frequencies are defined by (1)-(4).

$$P(i, j, d, 0^\circ) = \#\{(k, l), (m, n) \in (L_y \times L_x) \times (L_y \times L_x) \\ |k - m| = 0, |l - n| = d, I(k, l) = i, I(m, n) = j\} \quad (1)$$

$$P(i, j, d, 45^\circ) = \#\{(k, l), (m, n) \in (L_y \times L_x) \times (L_y \times L_x) \\ (k - m = d, l - n = -d) \text{ or } (k - m = -d, \\ l - n = d), I(k, l) = i, I(m, n) = j\} \quad (2)$$

$$P(i, j, d, 90^\circ) = \#\{(k, l), (m, n) \in (L_y \times L_x) \times (L_y \times L_x) \\ |k - m| = d, |l - n| = 0, I(k, l) = i, I(m, n) = j\} \quad (3)$$

$$P(i, j, d, 135^\circ) = \#\{(k, l), (m, n) \in (L_y \times L_x) \times (L_y \times L_x) \\ (k - m = d, l - n = d) \text{ or } (k - m = -d, \\ l - n = -d), I(k, l) = i, I(m, n) = j\} \quad (4)$$

where # denotes the number of elements in the set.

Note: These matrices are symmetric; $P(i, j; d, a) = P(j, i; d, a)$.

We compute four closely related measures $P(i, j, d, \theta)$ quantized to 45° intervals with $d = 1$ from which all our three texture features are derived. Out of the equations which define a total set of 14 measures of textural features [38], we have used the three most distinguishing parameters to describe the texture of an image as depicted by (5)-(7).

$$Energy = \sum_i \sum_j P(i, j)^2 \quad (5)$$

$$Entropy = - \sum_i \sum_j P(i, j) \log(P(i, j)) \quad (6)$$

$$Contrast = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1, |i-j|=n}^{N_g} P(i, j) \right\} \quad (7)$$

4.2 Building the Classifier Models for Annotation

In traditional classification problems, class labels are assumed to be mutually exclusive and each instance to be classified belongs to only one class. However, in the context of image annotation or semantic image classification, it is natural that one image belongs to multiple classes since image semantics is represented by both multiple semantic entities contained in the image and the relationships between them, causing the actual classes to overlap in the feature space [39].

A portion of the hierarchical classification of MR-T2 axial brain images has been used for our experiment as shown in Figure 1. We have considered a pre-defined textual set that contain a total of 15 keywords, each representing an individual concept in Figure 1. The *J48 decision tree* generated from WEKA tool has been used to obtain the classifier model for each concept. A large set of labelled images is required for training the classifier. There are 172 images with known concept which are used for training the classifier and 150 images that are used for testing. These images are of 256x256 GIF format obtained from brain atlas of Harvard University. For each image in the database, we extract the texture features as mentioned in the section 4.1. For each concept, the above mentioned texture feature attributes together with the associated label is transformed into a format that is suitable for applying the machine learning algorithm. With the help of WEKA tool, decision tree is obtained using only those attributes that are best able to differentiate the pattern by calculating the information gain. The objective is to get a decision tree that minimizes the error rate, with the highest number of correctly classified instances. The outcome of the classification for the individual concept is measured using precision and recall as shown in Table 1.

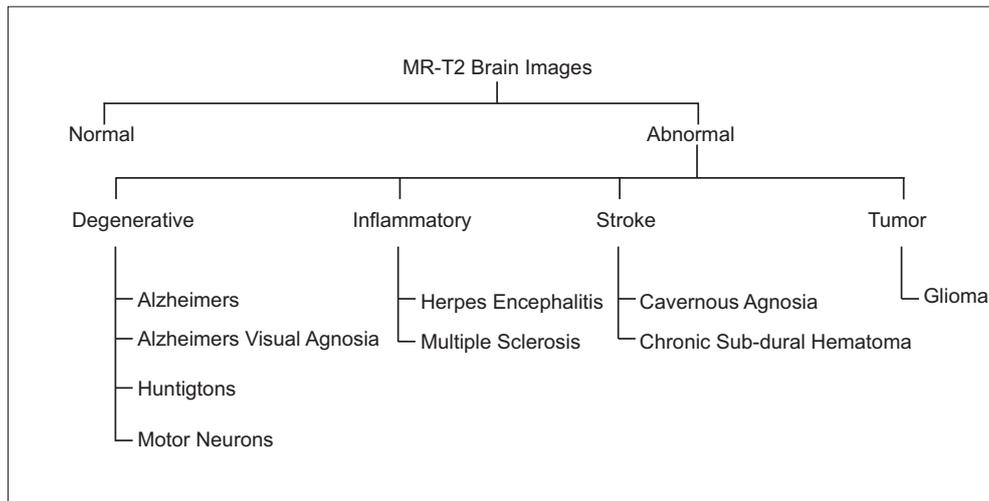


Fig. 1. A portion of the hierarchical classification of MR-T2 axial brain images

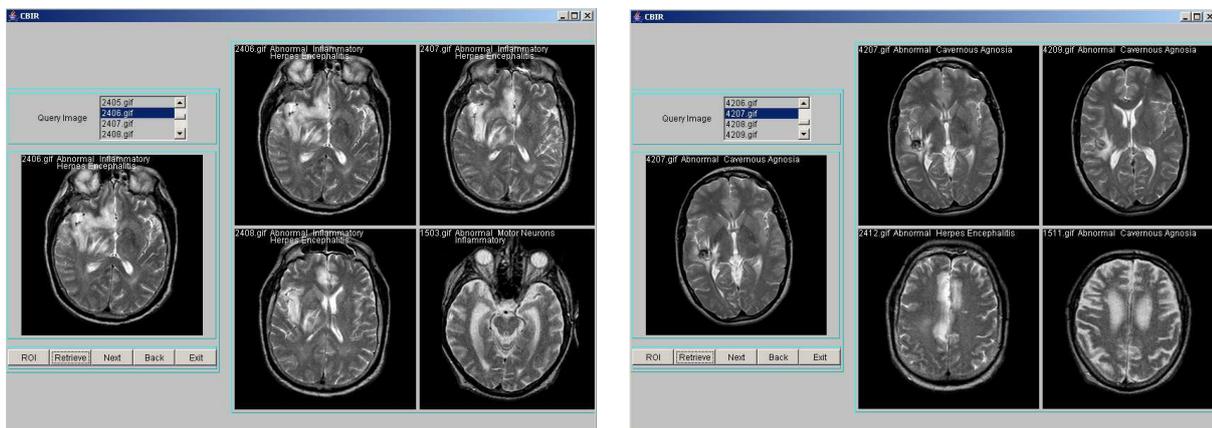


Fig. 2. A Few representative snapshots of CBR

Table 1. Results of J48 decision tree classifier

Class Name	Precision	Recall
Normal	0.935	0.967
Abnormal	0.993	0.986
Degenerative	0.902	0.917
Alzheimers	1.0	0.895
Visual Agnosia	0.95	1.0
Huntigtons	0.909	1.0
Motor Neurons	0.571	0.667
Inflammatory	0.688	0.786
Herpes Encephalitis	0.9	0.75
Multiple Sclerosis	0.385	0.313
Stroke	0.833	0.571
Cavernous Agnosia	0.4	0.286
Sub-dural Hematoma	0.905	0.905
Tumor	0.824	0.737
Glioma	0.824	0.737
Average	0.801	0.767

Table 2. J48 Decision Tree for the concept Degenerative

Contrast \leq 37078240
Entropy \leq 428289.375: No (105.0/4.0)
Entropy $>$ 428289.375: Yes (10.0)
Contrast $>$ 37078240
Entropy \leq 304816.5625
Entropy \leq 290763.34375
Entropy \leq 197924.359375: No (2.0)
Entropy $>$ 197924.359375: Yes (27.0)
Entropy $>$ 290763.34375: No (9.0)
Entropy $>$ 304816.5625: Yes (19.0)

a series of two-class classifiers for individual concept. In our annotation procedure, each word is generated independently by the decision tree obtained for each concept from texture feature attributes. Word correlations are not taken into consideration. Table 4 shows the result of automatic image annotation using our classification-based multi-class and multi-label image annotation model.

4.3 Multi-Label Annotation

The decision tree obtained for each concept is converted into if-then-else rules to assign the label to each image. For example, the decision tree of the concept *degenerative* shown in Table 2 is converted into if-then else rules shown in Table 3. To generate annotations, each image passes through

5. INDEXING AND RETRIEVAL

An important aspect of image retrieval system is the speed and efficiency with which a query object can be compared to the database objects. Indexing mechanisms are used to improve upon pair wise comparisons of the query to all database objects

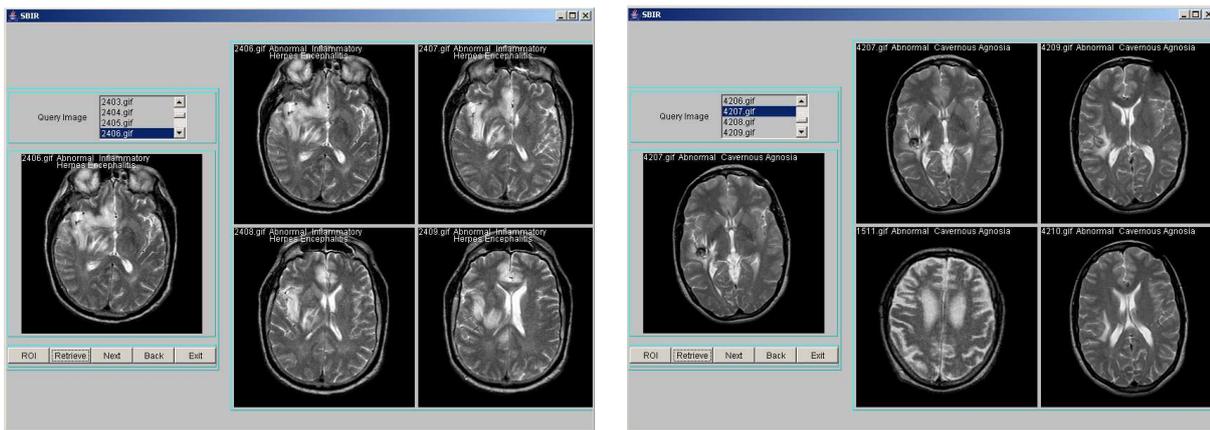


Fig. 3. A Few representative snapshots of SBIR

Table 4. Automatic image annotation results

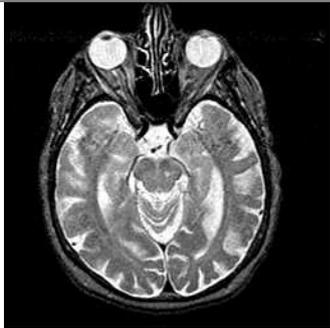
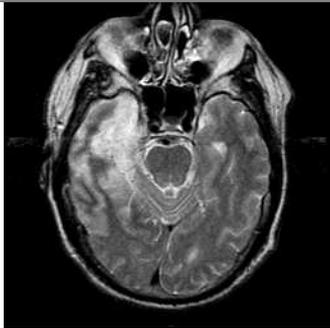
Images			
Actual Annotations	Abnormal Degenerative Alzheimers	Abnormal Inflammatory Herpes Encephalitis	Abnormal Stroke Subdural Hematoma
Automatic Annotations	Abnormal Degenerative Alzheimers	Abnormal Herpes Encephalitis	Abnormal Stroke Subdural Hematoma

Table 3. if-then-else rules for the decision tree in Table 2

```

if(cn ≤ 37078240) {
    if(ent > 428289.375) {
        System.out.print("Degenerative");
    }
}
else {
    if(ent ≤ 304816.5625) {
        if(ent ≤ 290763.34375) {
            if(ent ≤ 197924.359375) {
                System.out.print("Degenerative");
            }
        }
    }
    else {
        System.out.print("Degenerative");
    }
}

```

allowing retrievals to filter out large irrelevant portions of the database using distances to reference keys. A Hash structure is used to index all images. A combined index is adopted to point to all similar images in terms of the texture features. At query time, only those images that are in the same hash bucket as those of the queried image are compared for similarity, thus reducing the search space and time. The two indexing techniques implemented are: i) indexing based on image content and ii) indexing based on semantic keywords generated.

Table 7. Retrieval results of percentage recall rate for top 40 retrievals

Class Name	CBIR	SBIR
Normal	85	100
Alzheimers	84.21	100
Visual Agnosia	47.36	100
Huntigtons	70	100
Motor Neurons	58.33	83.33
Herpes Encephalitis	41.66	75
Multiple Sclerosis	37.5	37.5
Cavernous Agnosia	42.85	71.42
Sub-dural Hematoma	35	100
Glioma	63.15	63.15
Average	56.5	83.04

5.1 Content-Based Indexing and Retrieval (CBIR)

For each image in the database, the texture features: *entropy*, *energy* and *contrast* extracted are quantized to an integer values between 0 to 9. The combined index of these features is: $100 * [entropy] + 10 * [energy] + [contrast]$, where [] represents quantization. Each combined index stores feature data along with the image object. For a query image, the above mentioned texture features have to be computed, quantized and the combined index derived. Only those images that are stored at the combined index matching those of the query index, are extracted as resultant target images for a given query image. These resultant images are sorted using Euclidian distance measure in the decreasing order of similarity against the query image and displayed four images

Table 5. Retrieval results of percentage precision of CBIR for top 20 retrievals

Classes	Number of retrievals									
	2	4	6	8	10	12	14	16	18	20
Normal	100	100	100	100	100	100	100	100	100	100
Alzheimers	100	100	100	100	100	100	100	100	100	100
Visual Agnosia	100	100	100	100	80	66.66	71.42	75	66.66	70
Huntigtons	100	100	100	100	100	100	100	87.5	77.77	70
Motor Neurons	100	100	100	100	100	100	100	87.5	77.7	70
Herpes Encephalitis	100	100	100	75	60	66.66	57.14	50	55.5	50
Multiple Sclerosis	100	50	66.66	75	80	66.66	57.14	62.5	55.5	50
Cavernous Agnosia	100	100	66.66	50	60	66.66	71.42	75	66.66	60
Sub-dural Hematoma	100	100	100	100	100	100	100	87.5	77.7	70
Glioma	100	100	100	100	100	100	100	100	100	90
Average	100	95	93.33	90	88	86.66	85.71	82.5	77.77	73

Table 6. Retrieval results of percentage precision of SBIR for top 20 retrievals

Classes	Number of retrievals									
	2	4	6	8	10	12	14	16	18	20
Normal	100	100	100	100	100	100	100	100	100	100
Alzheimers	100	100	100	100	100	100	100	100	100	100
Visual Agnosia	100	100	100	100	100	100	100	100	100	100
Huntigtons	100	100	100	100	100	100	100	100	100	100
Motor Neurons	100	100	100	100	100	100	100	100	100	100
Herpes Encephalitis	100	100	100	100	100	100	100	100	100	90
Multiple Sclerosis	100	100	100	100	100	100	85.71	75	66.66	60
Cavernous Agnosia	100	100	66.66	75	80	83.33	71.42	75	77.77	80
Sub-dural Hematoma	100	100	100	100	100	100	100	100	100	100
Glioma	100	100	100	100	100	100	100	100	100	100
Average	100	100	96.66	97.5	98	98.33	95.71	95	94.44	93

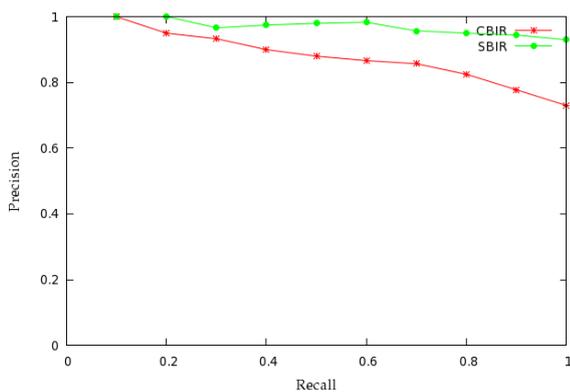


Fig. 4. Recall precision graph for top 20 retrievals

at a time along with the annotations using JAVA-AWT based GUI. A few representative snapshots of purely content-based indexing are shown in Figure 2.

5.2 Semantic-Based Indexing and Retrieval (SBIR)

For each image in the database, we extract the texture features: *entropy*, *energy* and *contrast* and annotate using our annotation model and then index them using the semantic keywords generated at the time of automatic image annotation. Though the length of the annotation is not restricted, indexing is done using first three keywords. With a pre-defined textual set that contains a total of 15 keywords (numbered 0 to 14), the combined index of the first three semantic keywords is $256 * [keyword1] + 16 * [keyword2] + [keyword3]$, where [] represents quantization. At each combined index, the texture feature data and the annotated keywords along with the image object is stored. For a query image, we extract the texture features: *entropy*, *energy* and

contrast and annotate using our annotation model and derive the combined index. Only those images that are stored at the combined index matching those of the query index, are extracted as resultant target images for a given query image. These extracted images are sorted in the decreasing order of similarity against the query image using Euclidian distance measure and displayed using JAVA-AWT based GUI, four images at a time. The snapshots of few representative retrievals of purely semantic-based indexing are shown in Figure 3.

6. PERFORMANCE ANALYSIS

The retrieval performance is measured using precision and recall. Precision measures the retrieval accuracy and is defined as the ratio between the number of relevant images retrieved and the total number of images retrieved. Recall measures the ability of retrieving all relevant images in the database. It is defined as the ratio between the number of relevant images retrieved and the total number of relevant images in the database. We have experimented with a sizable database of 300 images of 10 classes, each of 30 images. The experimentation was carried out to check the efficiency of the two indexing schemes described in section 5, as well as to study the performance of our semantic based indexing and retrieval results.

Each of the 300 images were used as query image and performance evaluated. Precision results are computed from the number of similar images (i.e., images belonging to the same class) in the top 20 retrieved images. Table 5 and 6 shows results depicting the *Precision rates* for 10 different classes tabulated for retrieval from top 2 to top 20 retrieved images. Each row presents average precision value for all query images on that class. Each column in the table presents precision rates computed using the specified number of retrievals. The results for SBIR is shown to be better when compared to CBIR, which leads to almost 25 percent increase in precision rates. Table 7

depicts the *Recall rates* for the same 10 different classes in the database. Here also, each of the 300 images were used as a query image and the number of matches in the top 40 retrieved images was counted and is shown to drastically increase recall rates by almost 50 percent. The precision recall graph for plotting the average precision retrieval rates for top 20 retrievals of the two indexing schemes is shown in Figure 4. In both the cases, precision decreases monotonically as a function of recall. We also see that performance of SBIR is better compared to CBIR.

7. CONCLUSIONS

We have implemented two methods of indexing namely: i) content-based indexing using statistical cooccurrence-based texture features and ii) semantic-based indexing using our automated annotation model. The performance of both the systems have been measured using standard recall versus precision graphs. Results obtained with the semantic-based indexing and retrieval are vastly improved giving retrieval results upto 96.86 percent on the average when compared to the content-based indexing and retrieval.

8. REFERENCES

- [1] Md. Mahmudur Rahman, Tongyuan Wang and Bipin C. Desai, Medical Image Retrieval and Registration: Towards Computer Assisted Diagnostic Approach, Proc. of the IDEAS Workshop on Medical Information Systems: The Digital Hospital (IDEAS-DH04).
- [2] Ying Liu, Dengsheng Zhang, Guojun Lu and Wei-Ying Ma, A survey of content-based image retrieval with high-level semantics, *Pattern Recognition* 40 (2007) 262-282.
- [3] Y Mori, H Takahashi and R Oka, Image-to-word transformation based on dividing and vector quantizing images with words, In MISRM99 First Intl. Workshop on Multimedia Intelligent Storage and Retrieval Management, 1999.
- [4] Duygulu P et al., Object Recognition as Machine Translation: Learning a Lexicon for a Fixed Image Vocabulary. in proceedings of 7th European Conference on Computer Vision, 2002.
- [5] Monay F and D Gatica-Perez, On Image Auto-Annotation with Latent Space Models, in proc. of ACM International Conference on Multimedia, 2003.
- [6] Blei D and M Jordan, Modeling Annotated Data, in proceedings of 26th International Conference on Research and Development in Information Retrieval (SIGIR), 2003.
- [7] Chang E et al., Cbsa: Content-Based Soft Annotation for Multimodal Image Retrieval Using Bayes Point Machines, *CirSysVideo*, 13(1): pp. 26-38, 2003.
- [8] Cusano C, G Ciocca and R Schettini, Image Annotation Using SVM, in proceedings of Internet Imaging IV, Vol. SPIE 5304, 2004.
- [9] Li J and J Z Wang, Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(19): pp. 1075-1088.
- [10] Jeon J, V Lavrenko and R Manmatha, Automatic Image Annotation and Retrieval Using Cross-Media Relevance Models, in Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval, 2003.
- [11] Lavrenko V, R Manmatha and J Jeon, A Model for Learning the Semantics of Pictures, in Proceedings of Advance in Neutral Information Processing, 2003.
- [12] Rong Jin Joyce Y and Chai Luo Si, Effective Automatic Image Annotation Via A Coherent Language Model and Active Learning, MM04, October 10-16, 2004, New York, USA, ACM 1-58113-000-0/00/0004.
- [13] Wei Li and Maosong Sun, Automatic Image Annotation Using Maximum Entropy Model, IJCNLP 2005, LNAI 3651, pp. 34-45, 2005 Springer-Verlag Berlin Heidelberg.
- [14] Munirathnam S, Joshua V, Mitchell B and Dan M, Exploiting Ontologies for Automatic Image Annotation, SIGIR05, August 15-19, 2005, Salvador, Brazil, ACM 1-59593-034-5/05/0008.
- [15] Gexiang Zhang, Weidong Jin and Laizhao Hu, Radar Emitter Signal Recognition Based on Support Vector Machines, 2004 8th International Conference on Control, Automation, Robotics and Vision Kunming, China, 6-9th December 2004 IEEE.
- [16] Madhubanti Maitra, Amitava Chatterjee and Fumitoshi Matsuno, A Novel Scheme for Feature Extraction and Classification of Magnetic Resonance Brain Images Based on Slantlet Transform and Support Vector Machine, SICE Annual Conference, August 20-22 2008, The University Electro-Communications, Japan, PR0001/08/0000-1130 2008 SICE.
- [17] Ahmed Kharrat et al., Automated Classification of Magnetic Resonance Brain Images Using Wavelet Genetic Algorithm and Support Vector Machine, Proc. 9th IEEE Int. Conf. on Cognitive Informatics (ICCI10) 978-1-4244-8040-1/10 2010 IEEE .
- [18] Bai Xing-li, Qian Xu, Medical Image Classification based on Fuzzy Support Vector Machines, 2008 International Conference on Intelligent Computation Technology and Automation, 978-0-7695-3357-5 2008 IEEE.
- [19] Bai Xingli, Tian Zhengjun, Medical Images Classification Based on Least Square Support Vector Machines, 978-1-4244-4507-3 2009 IEEE.
- [20] Wilburn E. Reddick et al., Automated Segmentation and Classification of Multispectral Magnetic Resonance Images of Brain Using Artificial Neural Networks, *IEEE Transactions on Medical Imaging*, Vol. 16, No. 6, December 1997.
- [21] Jan Larsen, Lars Nonboe Andersen, Mads Hintx-Madsen and Lars Kai Hansen, Design of Robust Neural Network Classifiers, 0-78034428-6/198, 1998 IEEE.
- [22] Qiang Ye, Paul W. Munro, Improving a Neural Network Classifier Ensemble with Multi-Task Learning, 2006 International Joint Conference on Neural Networks Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada July 16-21, 2006.
- [23] J Anitha, C Kezi Selva Vijila and D Jude Hemanth, An Enhanced Counter Propagation Neural Network for Abnormal Retinal Image Classification, 978-1-4244-5612-3/09 2009 IEEE.
- [24] Jiang Yun, Li Zhanhuai, Wang Yong and Zhang Longbo, A Better Classifier Based on Rough Set and Neural Network for Medical Images, Sixth IEEE International Conference on Data Mining - Workshops (ICDMW'06) 0-7695-2702-7 2006 IEEE.
- [25] Bo Pang, David Zhang, Naimin Li and Kuanquan Wang, Computerized Tongue Diagnosis Based on Bayesian Networks, *IEEE Transactions on Biomedical Engineering*, No 10, Vol. 51, October 2004.
- [26] Boaz Lerner and Roy Malka, Learning Bayesian Networks for Cytogenetic Image Classification, 18th Int. Conf. on Pattern Recognition (ICPR'06) 0-7695-2521-0/06 2006 IEEE.
- [27] Rong Chen, Edward H Herskovits, A Bayesian Network Classifier with Inverse Tree Structure for Voxelwise Magnetic Resonance Image Analysis, KDD05, August

- 2124 2005, Chicago, Illinois, USA, Copyright 2005 ACM 1-59593-135-X/05/0008.
- [28] Chunyi Lin, Lihong Ma, Jianyu Chen, Semantic Modeling for Multi-level Medical Image Semantics Retrieval, Third International Workshop on Advanced Computational Intelligence August 25-27 2010, Suzhou, Jiangsu, China.
- [29] Maria-Luiza Antonie, Osmar R. Zaane and Alexandru Coman, Application of Data Mining Techniques for Medical Image Classification, Proceedings of the Second International Workshop on Multimedia Data Mining (MDM/KDD2001), in conjunction with ACM SIGKDD conference, San Francisco, USA, August 26, 2001.
- [30] Sumeet Dua, Vineet Jain and Hilary W. Thompson, Patient Classification Using Association Mining of Clinical Images, 978-1-4244-2003-2/08 2008 IEEE.
- [31] P Rajendran and M Madheswaran, Pruned Associative Classification Technique For The Medical Image Diagnosis System, 2009 Second International Conference on Machine Vision, 978-0-7695-3944-7/10 2010 IEEE.
- [32] Lior Rokach and Oded Maimon, Top-Down Induction of Decision Trees Classifiers A Survey, IEEE Transactions on Systems, Man, And Cybernetics Part C: Applications and Reviews, Vol. 35, No. 4, November 2005.
- [33] Dengsheng Zhang, Md.MonirulIslam and GuojunLu, A review on automatic image annotation techniques, Pattern Recognition 45(2012) pp. 346-362.
- [34] Julie M. David and Kannan Balakrishnan, Significance of Classification Techniques In Prediction Of Learning Disabilities, International Journal of Artificial Intelligence & Applications (IJAAIA), Vol.1, No.4, October 2010.
- [35] Anil Rajput et al., Approaches of Classification to Policy of Analysis of Medical Data, IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.11, November 2009.
- [36] K P Soman, Shyam Diwakar and V Ajay, Insight into Data Mining: Theory and Practice, PHI Learning Private Limited, New Delhi, 2009.
- [37] G Meera Gandhi and S K Srivatsa, Adaptive Machine Learning Algorithm (AMLA) Using J48 Classifier for an NIDS Environment, Advances in Computational Sciences and Technology, ISSN 0973-6107 Vol. 3, No. 3 (2010), pp. 291304 Research India Publications.
- [38] R Haralick, K Shanmugam, and I Dinstein, Textural features for image classification, IEEE Trans. Systems on Man and Cybernetics, 3(6):610-621, 1973.
- [39] Wei Li and Maosong Sun, Multi-modal Multi-label Semantic Indexing of Images using Unlabeled Data, Int. Conf. on Advanced Language Processing and Video Technology, 978-0-7695-3273-8/08 2008 IEEE.