International Journal of Information & Computation Technology.
ISSN 0974-2239 Volume 9, Number 1 (2019), pp. 11-20

http://www. irphouse.com

GLCM and BoF Based Texture Representations for MRI Image Classification

P Dhivya[1], S Vasuki [2]

[1]*Department of Electronics & Communication Engineering, Velammal College of Engineering & Technology, Madurai, India.*

[2]*Department of Electronics & Communication Engineering, Velammal College of Engineering & Technology, Madurai, India.*

Abstract

Image representation is one of the major aspects of automatic classification algorithms. In this paper, different texture based feature extraction techniques have been utilized to represent medical MR images. They are categorized into two groups; (i) low-level image representation such as Gray Level Co-occurrence Matrix(GLCM) (ii) local patch-based image representation such as Bag of Features (BoF). These features have been exploited in different algorithms for automatic classification of medical MR images. Their classification performances obtained were analyzed with regard to the image representation techniques used. These experiments were evaluated on Oasis database consisting of 1260 medical MR images with 116 classes. Experimental results showed the classification performance obtained by exploiting BoF outperformed the other algorithms with respect to the image representation techniques used.

Keywords: Magnetic Resonance Images, Gray Level Co-occurrence Matrix, Bag of Features, Classification.

1. INTRODUCTION

Medical images are increasing at an alarming rate. This increasing number of images affects the interpreting capacity of radiologists. In order to reduce the burden of radiologists, automatic classification of medical images based on modality is the need of the hour. The important factor in automatic medical image classification based on modality is the texture feature used for classification purpose, because nice treatment on these subtleties can lead to good results.

Medical image possess a vast amount of texture feature relevant to clinical practice. For example, MR images of tissues are not capable of providing microscopic information that can be assessed visually. However, histological alterations present in some illnesses may bring about texture changes in MR images that are amenable to quantification through texture analysis [4].

Texture analysis plays an important role in assessing the spatial organization of different tissues and organs, overcoming the limits of the classical global measures [5]. Therefore, texture analysis has been widely explored in radiotherapic context, especially for the characterization of tumour in the planning phase and for the prediction of response to treatment.

In spite of several decades of development, most texture features have not been capable of performing at a level sufficient for real-world textures and are computationally too complex to meet the real-time requirements of many computer vision applications. The inherent difficulty in obtaining powerful texture representations lies in balancing two competing goals: high quality representation and high efficiency.

Bag of Features (local patch based image representation) is the approach that meets these goals by generating dictionary of features. It represents images using histograms of quantized appearances of local patches [10-13]. In recent years, many studies exploited this feature in various image classification domains including the medical domain. With increasing size of medical MRI archives, it is important to have simplistic, discrete representations and simple matching measures to preserve computational efficiency.

There is also an increase of digital information in medical domain where medical images of different modalities i.e., X-rays, CT scans, MRI scans, etc., are produced everyday in massive numbers. It is believed that the quality of such medical system can be improved by a successful classification of images so that the irrelevant images can be filtered out.

Automatic image classification is mapping images into pre-defined classes and it involves some basic principles such as representation where visual feature of the image are extracted and generalization which is training and evaluating the classifier. The first and most vital component of any classification system is image representation. It is categorized into two main approaches, (i) low-level image representation and (ii) patch based image representation.

In this paper, both low level image representation and local patch based image representation techniques are incorporated in different experiments for the task of automatic classification of medical MR images. The rest of the paper is organized as follows. Section 2 gives an overview of Bag of Features. Section 3 presents the proposed approach in detail. Experimental results and discussion are reported and analyzed in section 4. Finally, the overall conclusions of this study are presented in section 5.

2. BAG OF FEATURES (BoF)

The goal of texture representation or texture feature extraction is to transform the input texture image into a feature vector that describes the properties of a texture, facilitating subsequent tasks such as texture classification [16]. Since texture is a spatial phenomenon, texture representation cannot be based on a single pixel, and generally requires the analysis of patterns over local pixel neighbourhoods.

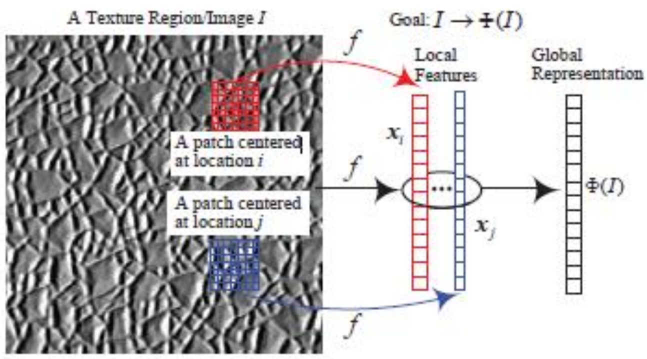

Fig.1 Schematic representation of texture representation process

Therefore, a texture image is first transformed to a pool of local features, which are then aggregated into a global representation for an entire image or region. Since the properties of texture are usually translationally invariant, most texture representations are based on an order-less aggregation of local texture features, such as a sum or max operation. Texture images can be statistically represented as histograms over a texton dictionary, referred to as Bag of Features (BoF) [19].

2.1 BoF Pipeline

The BoF pipeline is sketched in Fig. 2, consisting of the following basic steps.

2.1.1 Local Patch Extraction. For a given image, a pool of N image patches is extracted over a sparse set of points of interest over a fixed grid or densely at each pixel position.

2.1.2. Local Patch Representation. Given the extracted N patches, local texture descriptors are applied to obtain a set or pool of texture features of D dimension. We denote the local features of N patches in an image as $\{x_i\}_{i=1}^{N}, x_i \in \mathbb{R}^D$. Ideally, local descriptors should be distinctive and at the same time robust to a variety of possible image transformations, such as scale, rotation, blur, illumination, and viewpoint changes. High quality local texture descriptors play a critical role in the BoF pipeline.

2.1.3. Codebook Generation. The objective of this step is to generate a codebook (i.e., a texton dictionary) with K codewords $\{w_i\}_{i=1}^{K}, w_i \in \mathbb{R}^D$ based on training data. The codewords may be learned (e.g., by k-means).The size and nature of the codebook affects the representation followed and thus the discrimination power. The key here is how to generate a compact and discriminative codebook so as to enable accurate and efficient classification.

2.1.4. Feature Encoding. Given the generated codebook and the extracted local texture features $\{x_i\}$ from an image, the role of feature encoding is to represent each local feature x_i with the codebook, usually by mapping each x_i to one or a number of codewords, resulting a feature coding vector v_i (e.g. $v_i \in \mathbb{R}^K$). Of all the steps in the BoF pipeline, feature encoding is a core component which links local representation and feature pooling, greatly influencing the texture classification in terms of both accuracy and speed.

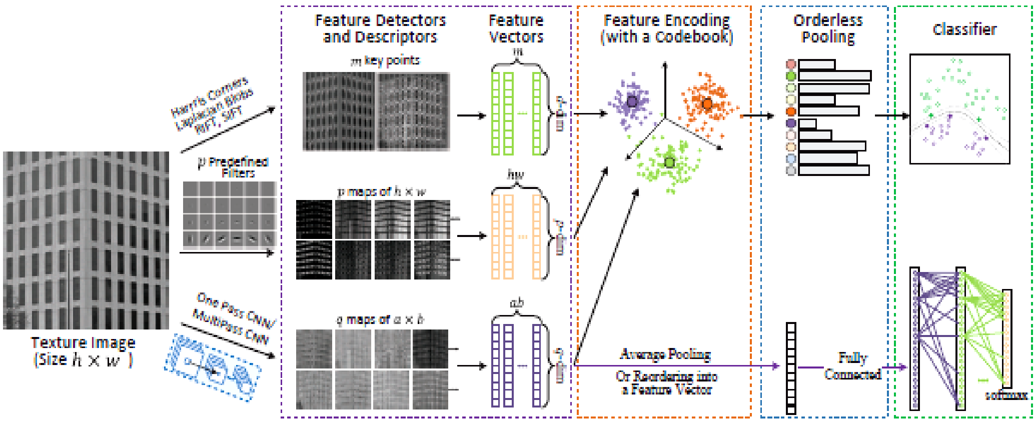

Fig. 2 General pipeline of BoF model

2.1.5. Feature Pooling. A global feature representation *y* is produced by using a feature pooling strategy to aggregate the coded feature vectors $\{v_i\}$. Classical pooling methods include average pooling, max pooling, and Spatial Pyramid Pooling (SPM).

2.1.6. Feature Classification. The global feature is used as the basis for classification, for which many approaches are possible: Nearest Neighbor Classifier (NNC), Support Vector Machines (SVM), neural networks, and random forests. SVM is one of the most widely used classifiers for the BoF based representation.

3. PROPOSED METHODOLOGY

The classification process consists of two steps, i.e. training phase and testing phase. In the training phase, the selected features are extracted from all the training images, and the classifier is trained on the extracted features to create a classification model. This model is then used to classify the test images into the predefined categories in the testing phase. Fig. 3 illustrates the training phase of the proposed classification framework.

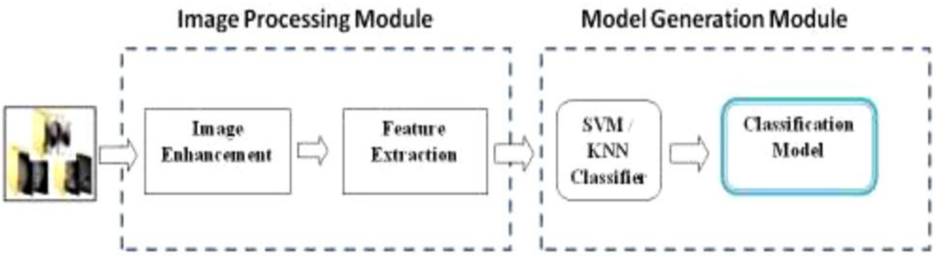

Fig.3 General Classification Framework (Training Phase)

3.1. Image Processing Module

Training phase consist of two modules; Image Processing and Model Generation. The image processing module composed of image enhancement and feature extraction as shown in Fig. 3.

3.1.1. Image Enhancement

Histogram equalization, one of the image enhancement techniques is applied to improve the quality of the image such as increasing the contrast of the image. This contrast adjustment provides better gray intensities distribution on the histogram. The method is useful in images with backgrounds and foregrounds that are both bright or both dark. In particular, the method can lead to better views of bone structure in MR images

3.1.2 Feature Extraction

The second component of image processing module is feature extraction. It plays an important role in the performance of any image classification because it can produce significant impact on the results of classification. Numerous low-level features such as colour, texture, shape are described in existing literature review. However, as MR images are gray level images and do not contain any colour information.

3.1.2.1 Texture

Texture contains important information regarding underlying structural arrangement of the surfaces in an image [21,22]. Gray Level Co-occurrence Matrices (GLCM) the commonly used feature extraction techniques, is employed for texture analysis.

GLCM is one of the well-known texture extraction techniques which measures second order texture characteristics. The GLCM of an N × N image, containing pixels with gray levels 0, 1, 2….G-1 is a matrix C (i, j), where each element of the matrix represents the probability of joint occurrence of intensity levels i and j at a certain distance and an angle θ. In this work, the four occurrence matrixes of GLCM are obtained from four different directions (θ ∈ {0°, 90°, 45°, and 135°}) at global level.

3.1.2.4. Bag of Features

The process of BoF started with detecting local interest point. Local interest point detectors have the task of extracting specific points and areas from images which are invariant to some geometric and photometric transformations. One of the popular approaches for the detection of local interest point is Difference of Gaussians (DoG) which is used in this experiment. DoG detector proposed by Lowe [25] has been built to be invariant to translation, scale, rotation, and illumination changes and samples images at different locations and scales. Next, distinctive feature that characterizes a set of keypoints for an image is extracted. Scale Invariant Feature Transform (SIFT) proposed by Lowe [26] is used to describe the grayscale image region around each keypoint in a scale and orientation invariant fashion .Each detected region is represented with the SIFT descriptor with the most common parameter configuration: 8 orientations and 4 × 4 blocks, resulting in a descriptor of 128 dimensions. Next step in implementation of bag of visual words is the codebook construction where the 128-dimensional local image features have to be quantized into discrete visual words. This task is performed using clustering or vector quantization algorithm. This step usually uses k-means clustering method, and use cluster center as visual vocabulary term. Upon identification of cluster centers, each image is represented as histograms of these cluster centers by simply counting the frequency of the words appear in an image. To accomplish this task, each feature vector in an image is assigned to a cluster center using nearest neighbor with a Euclidean metric.

3.2. Model Generation Module

As illustrated in Fig. 3, the next module after image processing is model generation. Note that the training set as well as the label of every images have been identified. Upon extraction of visual features from the entire training set, the extracted features as well as the label of every image in the dataset are fed into classifier to construct the classification model. Based on empirical results and several classification applications in same domain, Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) have shown a better classification performance as compared with other classification techniques.

SVM is very attractive for image classification as its aim to find the best hyperplane separating relevant and irrelevant vectors maximizing the size of margin. This optimum hyperplane has the maximum margin towards the sample objects, that is, the greater the margin, the less the possibility that any feature vector will be misclassified.

KNN is the most straightforward and simplest classifier in machine learning techniques. The KNN classification is based on majority vote of k-nearest neighbour classes. Classification is achieved by identifying the nearest neighbours to query example and using those neighbours to determine the class of query. In this work, k=9 is used, that means that the algorithm will take majority vote of its 9 nearest neighbours.

4. EXPERIMENTAL RESULTS

In this section, set of experiments were conducted to evaluate the classification performance obtained with respect to various image representation techniques. The database used in this research was OASIS dataset (URL: http:// www.oasis-brains.org/); this dataset contains 1260 MR images from 116 categories which differ from each other either on account of image modality, examined region, body orientation and biological system examined. 20% of 1260 training images were taken as test images to ensure that each class has representation in testing data and the remaining 80% are taken as training images.

4.1. Experiment 1

Fig. 4 depicts the classification accuracy of the first experiment with two different classifiers; SVM and KNN. The total classification accuracy obtained with SVM and KNN is 95.38% and 85.95% respectively. The result from Fig. 5 demonstrate that from 116 classes, the accuracy rate of 16 classes were 100% using SVM, and 9 classes obtained 100% accuracy using KNN. From the figure, it can be seen that 47 classes have classification accuracy of 0 %. This is because that all these classes have less than 15 training images and this would affect on their classification result as well as intra-class and inter-class similarity in images.

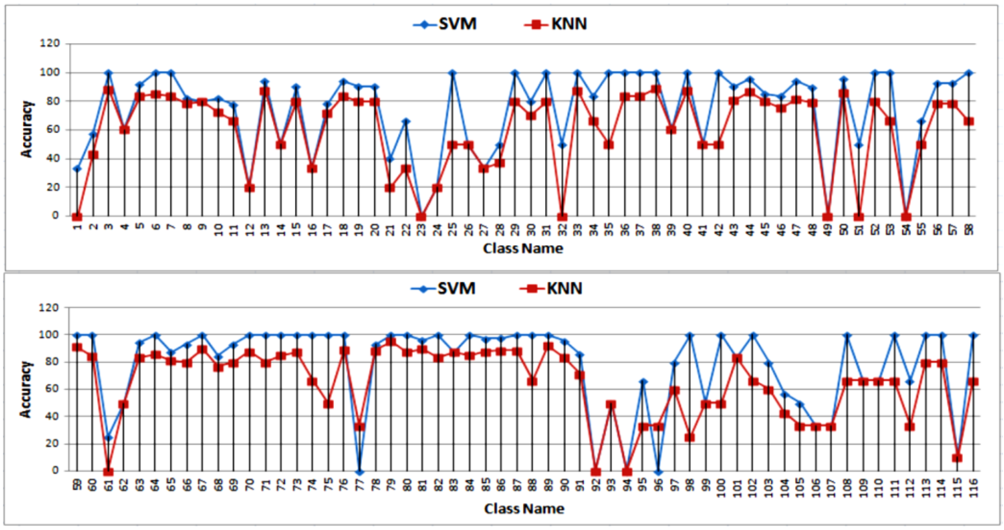

Fig. 4 Classification result for first experiment using SVM and KNN

4.2 Experiment 2

In this experiment, BoF are extracted from the training images and fed into classifier for constructing the classification model. The total classification rate obtained from the model generated by SVM and KNN classifier are 100% and 96% respectively. Fig. 5 shows the accuracy rate for each individual class obtained by this experiment.

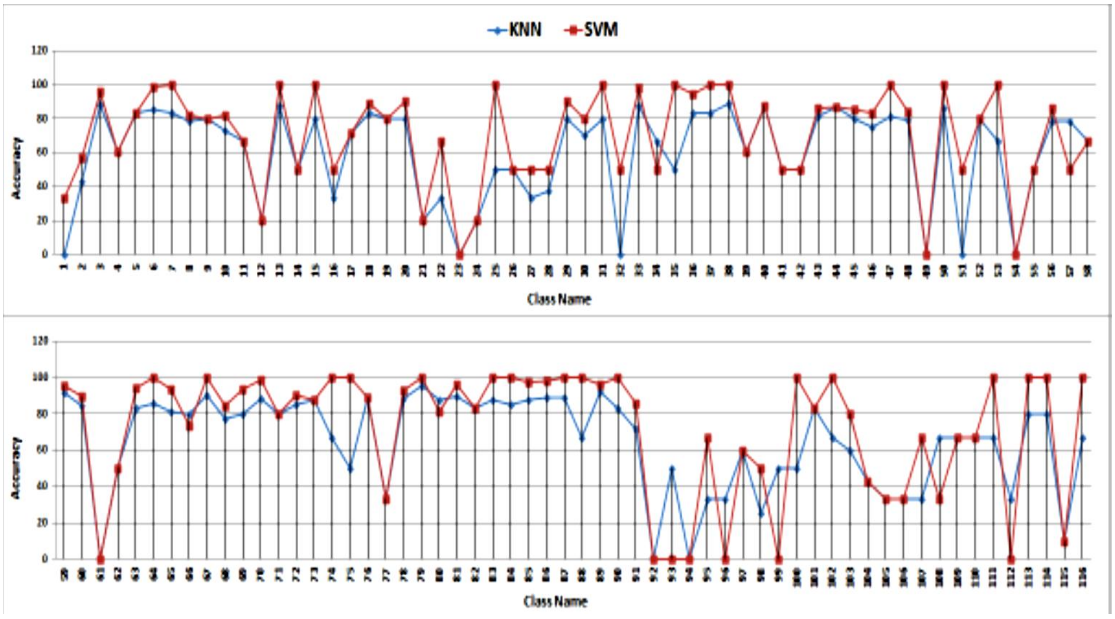

Fig. 5 Classification result for second experiment using SVM and KNN

5. CONCLUSION

This work presents different methods for automatic MR image classification with respect to the image representation techniques used. Image representation is categorized into two groups such as low-level image representation and patch-based image representation. In this paper, GLCM is applied as low-level image representation. BoF is used for local patch-based image representation. These features have been employed in different experiments for automatic classification of medical MR images. The evaluation for these experiments was conducted on OASIS medical image database. The experimental results indicates those classification methods constructed from BoF outperformed the other model generated from other representation technique.

REFERENCES

[1] W. Khaliq et al., “Comparison of a PACS workstation with laser hard copies for detecting scaphoid fractures in the emergency department”. Journal of Digital Imaging,. Vol. 23, No. 1, 2008, pp.100–103

[2] K. Seong-Hoon et al., “X-ray image classification using Random Forests with

Local Binary Patterns", International Conference on the Machine Learning and Cybernetics, 2010

[3] T. GuangJian et al., "Automatic medical image categorization and annotation using LBP and MPEG-7 edge histograms", International Conference on Information Technology and Applications in Biomedicine, 2008

[4] C. M. Zhy et al., "Segmentation of Ultrasound Image Based on Texture Feature and Graph Cut", International Conference on the Computer Science and Software Engineering, 2008.

[5] A. S. M. Sohail et al.,"Retrieval and classification of ultrasound images of ovarian cysts combining texture features and histogram moments". IEEE International Symposium on the Biomedical Imaging: From Nano to Macro, 2010.

[6] T. Deselaers et al., "Deformations, patches, and discriminative models for automatic annotation of medical radiographs", Pattern Recognition Letters, Vol. 29, No. 15, 2008, pp.2003-2010.

[7] I. Dimitrovski et al., "Hierarchical annotation of medical images", Pattern Recognition, Vol. 44, No. 10–11, 2011, pp.2436-49

[8] A. Mueen et al. , "Automatic Multilevel Medical Image Annotation and retrieval", Journal of Digital Imaging, Vol. 21, No. 3, 2008, pp. 290-295.

[9] T. Tommasi et al., "Discriminative cue integration for medical image annotation", Pattern Recognition Letter, Vol. 29, No. 15, 2008, pp. 1996-2002.

[10] U. Avni et al., "X-ray Categorization and Retrieval on the Organ and Pathology Level, Using Patch-Based Visual Words", IEEE Transaction on Medical Imaging , Vol. 30, No. 3, 2011, pp. 733-746.

[11] U. Avni et al., "Chest x-ray characterization: from organ identification to pathology categorization", International Conference on Multimedia Information Retrieval, Philadelphia, 2010.

[12] T. Deselaers et al., "Sparse Patch-Histograms for Object Classification in Cluttered Images", In: DAGM 2006, Pattern Recognition, 27th DAGM Symposium, Lecture Notes in Computer Science, Vol. 4174, berlin, Germany, 2006, pp. 201-211.

[13] K. Kesorn et al., "An Enhanced Bag-of-Visual Word Vector Space Model to Represent Visual Content in Athletics Images", IEEE Transaction on Multimedia, Vol. 14, No. 1, 2012, pp. 211-222.

[14] W. Zhou et al., "Principal Visual Word Discovery for Automatic License Plate Detection", IEEE Transaction on Image Processing., Vol. 21, No. 9, 2012, pp. 4269-4279.

[15] JM Aman et al., "Content-based image retrieval on CT colonography using rotation and scale invariant features and bag-of-words model", IEEE

International Symposium on Imaging: From Nano to Macro., 2010, pp. 1357-1360.

[16] F. Yang et al., “Visual tracking via bag of features”, IET Image Processing, Vol. 6, No. 2, 2012, pp. 115-128 .

[17] L. Sui et al., “Research on pornographic images recognition method based on visual words in a compressed domain”, IET Image Processing, Vol. 6, No. 1, 2012, pp. 87-93.

[18] S. Lazebnik et al., “Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories”. IEEE Conference on Computer Vision and Pattern Recognition, 2006, pp. 2169-2178.

[19] F. Jie et al., “Bag-of-Visual-Words Based on Clonal Selection Algorithm for SAR Image Classification”, IEEE Geoscience and Remote Sensing Letters, Vol. 8, No. 4, 2011, pp. 691-695.

[20] W. Jingyan et al., “Bag-of-Features Based Medical Image Retrieval via Multiple Assignment and Visual Words Weighting”, IEEE Transaction on Medical Imaging, Vol. 30, No. 11, 2011, pp. 1996-2011.

[21] R.M. Haralick et al., “Textural features for image classification”, IEEE Transaction on System, Vol. 3, 1973, pp.610-621.

[22] T. Ojala et al., “A Comparative Study of Texture Measures with Classification Based on feature Distributions”, Pattern Recognition, Vol. 29, 1996, pp.51-59.

[23] J. Canny, “A computational approach to edge detection”. IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, 1986, pp. 679-698.

[24] D. Keysers et al., “Deformation Models for Image Recognition”, IEEE Transaction Pattern Analysis and Machine Intelligence, Vol. 29, No. 8, 2007, 1422-1435.

[25] D. Lowe, “Object recognition from local scale-invariant features”, International Conference on Computer Vision, 1999, pp. 1150–1157.

[26] D. Lowe, “Distinctive image features from scale invariant key points”, International Conference on Computer Vision , Vol. 60, No. 2, 2004, pp. 91-110.

[27] H. Müller et al., “Overview of the ImageCLEFmed 2007 medical retrieval and annotation tasks”, In: Working Notes of the 2007 CLEF Workshop, 2007.