ISSN 1996-3343

Asian Journal of Applied Sciences

Knowledgia
SCIENTIFIC
A Place to Publish Outstanding Research

http://knowledgiascientific.com

Asian Journal of Applied Sciences 8 (1): 71-78, 2015
ISSN 1996-3343 / DOI: 10.3923/ajaps.2015.71.78

Medical Image Classification Using Multi-Vocabulary

Ahmed Mueen, Abdullah AL-Malaise AL-Ghamdi, Bassam Zafar and Mohammad Awedh
King Abdulaziz University, Saudi Arabia

Corresponding Author: Ahmed Mueen, King Abdulaziz University, Saudi Arabia

ABSTRACT

In this study, the bag-of-visual-word based medical image classification technique was investigated. A new approach for medical image classification was proposed by introducing multi steps image classification using three different visual vocabularies. First, image is classified into general category by constructing that level of vocabulary. In second step image is classified into middle level by building another type of vocabulary and in last step specific vocabulary is calculated to perform exact classification. The proposed algorithm was evaluated on IRMA 2005 database consisting of 9,000 medical x-ray images of 57 classes. The accuracy rates obtained from three vocabularies are 95, 92 and 90%.

Key words: Image classification, medical x-ray images, feature extraction, bag-of-words

INTRODUCTION

Over the last decade, huge amount of images are produced in the areas of education, entertainment, geographical information, remote sensing systems and medicine. Medical images in particular, are produced daily in large amount of various imaging modalities e.g., Computer Tomography (CT), Magnetic Resonance Imaging (MRI), X-rays, etc. Consequently, this gives us a challenging problem of developing a system that can provide effective and efficient management of these valuable resources and able us to quickly and accurately search and navigation through the enormous digital archives. Many automatic medical image clustering and classification technique were introduced by researchers (Goldberger *et al*., 2006, 2007; Greenspan and Pinhas, 2007; Mueen *et al*., 2007) in recent years. The main task for clustering and classification of images is extracting visual features of the image which is one of the most important aspects of such approaches (Muller *et al*., 2004). A visual feature includes color, texture and shape image characteristics. Shyu *et al*. (2002) used feature extraction based on physicians perceptual categories and claim better retrieval accuracy than traditional approaches. Mueen *et al*. (2007) introduced new medical image classification method by using multi-level features. Coelho and Ribeiro (2010) suggested an approach using global descriptors from MPEG7, GIST and compact composite descriptors for medical image retrieval. Many other published studies (Kim *et al*., 2010; Tian *et al*., 2008; Wang *et al*., 2011) used visual features for clustering and categorizing objects for medical image retrieval. A recent trend in Medical Content-Based Image Retrieval (M-CBIR) is to use state-of- the-art SIFT features (Lowe, 2004) and the bag-of-visual-words (Sivic and Zisserman, 2003) image representation for large scale image retrieval system. The bag-of-visual-words representation is inspired from text retrieval techniques (Squire *et al*., 2000) and now becomes a well know method in content-based image retrieval and object recognition. In this approach, SIFT descriptor (Lowe, 2004) is used to describe regions of interest within the image as feature which is both scale rotation invariant. A local descriptor is then quantized into a visual word depending on a size of visual vocabulary. Each image is represented as a histogram of visual words and similarity between images is calculated by matching their histograms. Impressive results have been achieved with this approach in recent years. Furthermore, various approaches have been proposed to make bag-of-visual-words technique more effective and efficient (Liu *et al.*, 2008; Quack *et al.*, 2007; Aly *et al.*, 2009). In the present method the image patches were used as visual words that are clustered to form a dictionary (Avni *et al.*, 2009). This approach is evading the need for explicit object detection features which has been successfully applied to scenery image classification tasks (Nowak *et al.*, 2006). The concept modeling was also considered for organizing concepts and knowledge representation which help medical image retrieval system to extract potentially relevant images from the image database. Most of existing medical image classification techniques do not consider the hierarchical relationship between different levels (Breen *et al.*, 2002). Therefore, in this study, multi-level classification of image using three different vocabularies of visual words are presented.

MATERIALS AND METHODS

Benchmark image database: In this study, the research database used is from IRMA x-ray library (Lehmann, 2013). It consists of 9,000 training set and 1,000 test set radiograph images. Training set is classified into 57 predefined classes. Figure 1 shows some images from different classes. The medical radiographs collected randomly from daily routine work at the RWTH University Hospital of Aachen, Germany.

Classification of these images is non-trivial task (Pinhas and Greenspan, 2004) due to the high intra class variability and inter-class similarity among classes, presence of clothes, jewels and medical instruments and imbalance of training samples, whereby some classes have a huge number of samples; for example class 12 has 2563 images and class 2 has only 32 images.

Bag-of-word based medical image classification framework: The present framework was presented for the task of medical image classification. The block diagram of the proposed framework is show in Fig. 2. The framework is composed of three phases: Feature extraction phase; vocabulary

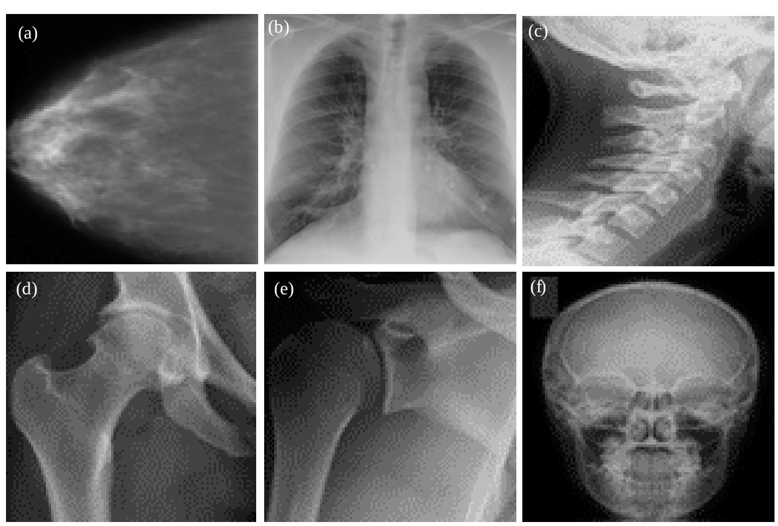

Fig. 1(a-f): Images from different classes

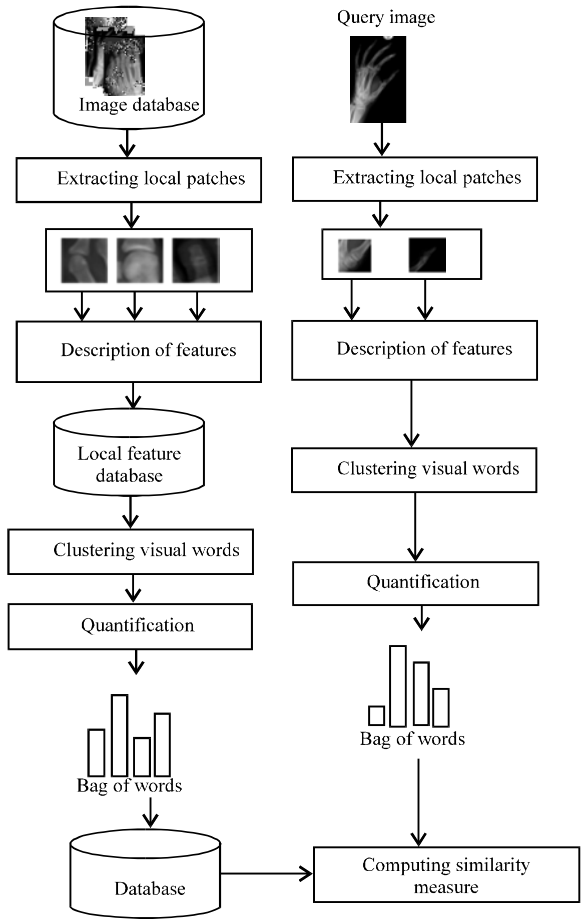

Fig. 2: Medical image classification framework

construction phase and classification phase. The patches are sampled around every pixel using fixed patch size. These patches are clustered to form visual vocabulary of visual words. Then, the patches are extracted from all the training images and mapped to the cluster centers to create each image a histogram of visual words. Histogram intersection is used as similarity measure with Support Vector Machine (SVM) classifier for classification:

- **Representation of images:** Appropriate feature representation significantly increases the performance of medical image classification and retrieval systems. Various approaches have been proposed to detect more discriminative features form image (Lowe, 2004) that are locally extracted from image. In this study, the local patches were extracted which have shown promising results for object classification and retrieval task (Avni *et al*., 2009). A patch around every pixel is extracted but patches along the border of an image are considered noise and are excluded. Let, X be a bag of features of an image and $\{x_l\}$, $l = 1,....$ L be a collection of local features extracted from X. The patch size of 9×9 pixels was used and Principal Component Analysis (PCA) was applied to reduce data dimensionality form 81 to 15. All the patches with single intensity value of black are eliminated:

- **Construction of vocabulary:** Next step in implementation of bag of visual words is the vocabulary construction based on a representative set of images. The main step in vocabulary building is clustering the patches. This task is performed using vector quantization process which usually done by k-means clustering algorithm. Let, V be a visual vocabulary then $V = \{v1, ..., v_K\}$ with K visual words for all images to get word vector representation. Once the cluster centers are recognized, each feature vector in an image is assigned to a cluster center using nearest neighbor method with a Euclidean metric and lastly each image is represented as histogram of these cluster centers by simply counting the occurrence of the words appear in an image
- **Classification:** The multi-class Support Vector Machine (SVM) classifier is used for classification. Bag-of-visual-words representation was extracted from training dataset. Then it was used as inputs to SVM classifier to build the model. The training dataset divided into two parts. The first part was 80% images used to construct the classification models and the remaining 20% of the training images were taken as for test images for evaluation purpose of the generated model

Hierarchical classification: IRMA medical image database is classified into 57 predefined classes and each class has been annotated with short description. Using this annotation, three level of hierarchy was constructed (Mueen *et al*., 2008). In this study, three visual vocabularies were created according to each level of hierarchy. First vocabulary contain descriptors from all the 57 classes, second vocabulary consist of 29 classes descriptors and third and last vocabulary from 9 classes according to hierarchy. Every image in training and test stage clusters into three different vocabularies. Therefore, the classification has three phases. First, the image is classified into one broad class that is one of the main 9 classes. In second phase, image is classified in one of 29 middle level classes. Third phase gives exact classification which is classification in one of the 57 classes.

Experimental evaluation: Each class in the database contains different number of sample images. Number of training images in some classes are very high and other classes have limited training images. This unbalanced number of training images may affect classification results because most probably few classes may dissolve into other classes at training stage. There are 57 different image classes within the archive, they are differing in either the examined region, the image orientation with respect to the body or the biological system under evaluation. The distribution of the images across the categories is non-uniform the most common class contains over 25% of the images in the database while few categories are represented by less than 0.1% of the images.

Comparison: In this experiment, Bag-of-Visual-Words technique is used. Therefore, the size of a vocabulary is an important factor for the classification performance. Different vocabulary size has been considered starting from 200 words followed by 300, 400, 500 and 600 words.

Two classifiers, SVM and k-NN are compared. The SVM is generally used for statistical learning and classification. Mostly, SVM classifier deals with binary classification problem but currently two multiple classification approaches, one-against-one and one-against-all are also in use. In our experiment, we have chosen one-against-all due to its speed. The second most commonly used classification method k-NN is used for further comparison. The measurement used to compare classification results is the average accuracy, that is:

$$\text{Average accuracy (Correctness rate)} = \frac{\text{No. of images classified correctly}}{\text{Size of test dataset}}$$

RESULTS AND DISCUSSION

In the present experiment, the input image was represented as collection of small patches. The image patch was essentially croped around every pixel, using a patch size of 9×9 pixels. The sample image patch was clustered to form visual vocabulary. The results shown in Fig. 3 refers, to the correctness rates of different vocabulary size. The vocabulary sizes we consider are 200, 300, 400, 500, 600. As illustrated in Fig. 3, increasing the vocabulary size improves the classification accuracy. Based on these experiments, the best performance was achieved with vocabulary size of 500 words. This result is strange as compare in natural image classification and retrieval.

Normally, in natural image database vocabulary size is between 1-10 k (Jegou *et al.*, 2010). However, in medical domain the present result shows that medical image classification needs less vocabulary size than what is normally used in natural images classification. This is due to the fact, that natural image contents are much more complex than medical image contents. Furthermore, the present result can be confirmed by other studies (Avni *et al.*, 2009). They obtained very good medical image retrieval results with a vocabulary size of 700 visual words only. Figure 3 also demonstrates that SVM classifier perform better than K-NN classifier with an average rate of 90% which is 4% more than the K-NN classifier. Specially in medical domain classification performance of SVM classifier outperform other classifiers which has been proved in many studies (Zare *et al.*, 2013; Avni *et al.*, 2011).

For hierarchal classification three experiments were performed under three vocabularies (one vocabulary at each level). The hierarchical approach improved the result about 5%. Concept hierarchy defines the semantic image concepts and their logical associations. The first level of semantic hierarchy covers more general perceptions and then they become more specific at lower levels.

The present experiment involves three level hierarchies with three vocabularies: each level contains one vocabulary. The first level of hierarchy consists of 9 main concepts which correspond to 9 main regions of human body. Second level contains 29 concepts. The last and most specific, level contains 57 concepts.

The accuracy of the first level with the vocabulary of 9 classes achieved was 95%; the middle level with 29 classes obtained an accuracy of 92%; the third level accuracy was 90%. Classification results of each class for all three levels are shown in Fig. 4a-c.

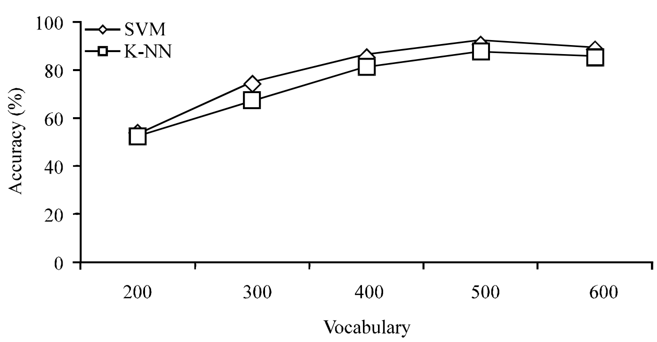

Fig. 3: Effect of vocabulary size, for SVM and K-NN classifers

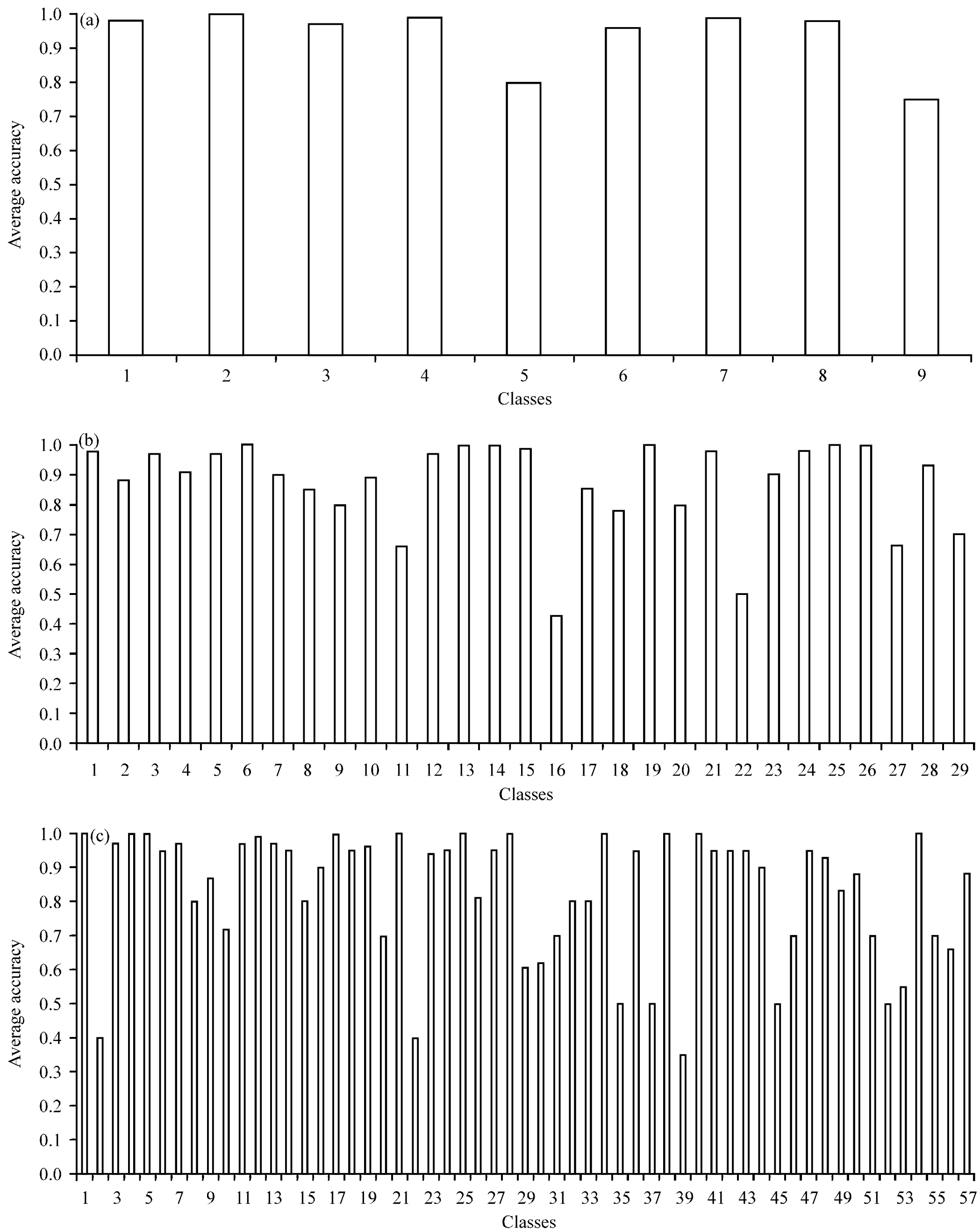

Fig. 4(a-c): Classification results of (a) 1st, (b) Middle and (c) Last level classes

In multi-vocabulary image classification, images are classified into most relevant image concept. The present approach uses concept hierarchy with the combination of visual vocabularies. For each level of hierarchy, one vocabulary was constructed to classify image in relative classes. Other than medical images text keywords used for object classes and image concept become the text keywords for annotation and classification (Gao *et al.*, 2006). In contrast the present approach uses concept hierarchy which represents hierarchical knowledge and organization and for each level of hierarchy one vocabulary was built to categorize image in relative classes.

CONCLUSION

In this study, a novel hierarchical-based medical image classification technique was presented. The classification was done by three classification steps with three separate visual vocabularies. Instead of absolute decision for a class, the proposed technique decides first about broad category which is the first level of hierarchy. In second step, new visual vocabulary was constructed with combination of different classes according to the middle level of hierarchy. Lastly, third vocabulary was calculated to specify the exact class of an image. In order to improve the discriminative power, bag-of-visual-word is used to extracted patch on each pixel. The experiment results show the strength of our method. The present study uses simple multi-level hierarchy; one future work is to build more interrelation between levels. In addition, the present study can be used in an online medical image classification and retrieval system.

ACKNOWLEDGMENT

The study was supported by King Abdullah City of Science and Technology under the grant No. A-C-11-0406. The image data used in this study is courtesy of TM Lehmann, Department of Medical Informatics, RWTH Aachen, Germany.

REFERENCES

Aly, M., P. Welinder, M. Munich and P. Perona, 2009. Towards automated large scale discovery of image families. Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition Workshops, June 20-25, 2009, Miami, FL., USA., pp: 9-16.

Avni, U., H. Greenspan, M. Sharon, E. Konen and J. Goldberger, 2009. X-ray image categorization and retrieval using patch-based visualwords representation. Proceedings of the International Symposium on Biomedical Imaging: From Nano to Macro, June 28-July 1, 2009, Boston, MA., USA., pp: 350-353.

Avni, U., H. Greenspan, E. Konen, M. Sharon and J. Goldberger, 2011. X-ray categorization and retrieval on the organ and pathology level, using patch-based visual words. IEEE Trans. Med. Imaging, 30: 733-746.

Breen, C., L. Khan and A. Ponnusamy, 2002. Image classification using neural networks and ontologies. Proceedings of the 13th International Workshop on Database and Expert Systems Applications, September 2-6, 2002, France, pp: 98-102.

Coelho, F. and C. Ribeiro, 2010. Evaluation of global descriptors for multimedia retrieval in medical applications. Proceedings of the Workshop on Database and Expert Systems Applications, August 30-September 3, 2010, Bilbao, Spain, pp: 127-131.

Gao, Y., J. Fan, X. Xue and R. Jain, 2006. Automatic image annotation by incorporating feature hierarchy and boosting to scale up SVM classifiers. Proceedings of the 14th Annual ACM International Conference on Multimedia, October 23-27, 2006, Santa Barbara, CA., USA., pp: 901-910.

Goldberger, J., S. Gordon and H. Greenspan, 2006. Unsupervised image-set clustering using an information theoretic framework. IEEE Trans. Image Process., 15: 449-458.

Goldberger, J., H. Greenspan and J. Dreyfuss, 2007. An optimal reduced representation of a mog with applicatios to medical image database classification. Proceedings of the Conference on Computer Vision and Pattern Recognition, June 17-22, 2007, Minneapolis, MN., USA., pp: 1-6.

Greenspan, H. and A.T. Pinhas, 2007. Medical image categorization and retrieval for PACS using the GMM-KL framework. IEEE Trans. Inform. Technol. Biomed., 11: 190-202.

Jegou, H., M. Douze and C. Schmid, 2010. Improving bag-of-features for large scale image search. Int. J. Comput. Vision, 87: 316-336.

Kim, S.H., J.H. Lee, B. Ko and J.Y. Nam, 2010. X-ray image classification using random forests with local binary patterns. Proceedings of the International Conference on Machine Learning and Cybernetics, Volume 6, July 11-14, 2010, Qingdao, China, pp: 3190-3194.

Lehmann, T., 2013. IRMA x-ray library. Institut de Recherche Mathematique Avancee (IRMA), Strasbourg, France.

Liu, D., G. Hua, P. Viola and T. Chen, 2008. Integrated feature selection and higher-order spatial feature extraction for object categorization. Proceedings of the Conference on Computer Vision and Pattern Recognition, June 23-28, 2008, Anchorage AK, USA., pp: 1-8.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision, 60: 91-110.

Mueen, A., M.S. Baba and R. Zainuddin, 2007. Multilevel feature extraction and X-ray image classification. J. Applied Sci., 7: 1224-1229.

Mueen, A., R. Zainuddin and M.S. Baba, 2008. Automatic multilevel medical image annotation and retrieval. J. Digit. Imaging, 21: 290-295.

Muller, H., N. Michoux, D. Bandon and A. Geissbuhler, 2004. A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. Int. J. Med. Inform., 73: 1-23.

Nowak, E., F. Jurie and B. Triggs, 2006. Sampling strategies for bag-of-features image classification. Proceedings of the 9th European Conference on Computer Vision, May 7-13, 2006, Graz, Austria, pp: 490-503.

Pinhas, A.T. and H.K. Greenspan, 2004. A continuous and probabilistic framework for medical image representation and categorization. Proceedings of the SPIE Conference on Medical Imaging, April 19, 2004, USA., pp: 230-238.

Quack, T., V. Ferrari, B. Leibe and L. Van Gool, 2007. Efficient mining of frequent and distinctive feature configurations. Proceedings of the 11th International Conference on Computer Vision, October 14-21, 2007, Rio De Janeiro, Brazil, pp: 1-8.

Shyu, C.R., C. Pavlopoulou, A.C. Kak, C.E. Brodley and L.S. Broderick, 2002. Using human perceptual categories for content-based retrieval from a medical image database. Comput. Vision Image Understanding, 88: 119-151.

Sivic, J. and A. Zisserman, 2003. Video Google: A text retrieval approach to object matching in videos. Proceedings of the 9th International Conference on Computer Vision, October 13-16, 2003, Nice, France, pp: 1470-1477.

Squire, D.M., W. Muller, H. Muller and T. Pun, 2000. Content-based query of image databases: Inspirations from text retrieval. Pattern Recogn. Lett., 21: 1193-1198.

Tian, G., H. Fu and D.D. Feng, 2008. Automatic medical image categorization and annotation using LBP and MPEG-7 edge histograms. Proceedings of the International Conference on Information Technology and Applications in Biomedicine, May 30-31, 2008, Shenzhen, China, pp: 51-53.

Wang, J., Y. Li, Y. Zhang, C. Wang, H. Xie, G. Chen and X. Gao, 2011. Bag-of-features based medical image retrieval via multiple assignment and visual words weighting. IEEE Trans. Med. Imaging, 30: 1996-2011.

Zare, M.R., A. Mueen and W.C. Seng, 2013. Automatic classification of medical X-ray images using a bag of visual words. IET Comput. Vision, 7: 105-114.