Vol. 7(26), Oct. 2017, PP. 3644-3652

Contents list available at IJMEC

International Journal of Mechatronics, Electrical and Computer Technology (IJMEC)

Journal Homepage: www.aeuso.org

PISSN: 2411-6173 - EISSN: 2305-0543

Medical Image Retrieval Based On Ensemble Clustering

Farzad Golnoori, Negar Abbasi, Simindokht Jahangard, Farsad Zamani Boroujeni*

Department of Computer, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran

****Corresponding Author's E-mail:*** *farzadgolnoori@yahoo.com*

Abstract

With the pervasive availability and application of medical imaging equipment, every day, thousands of digital medical images must be stored in databases of hospitals and other medical facilities. Considering the sheer number of medical images taken every day, organization, classification, and retrieval of these images has become very challenging. One of the typical methods of image retrieval is the use of bag of visual words. In this approach, one of the most common techniques of constructing visual words is the use of K-means clustering. However, performance of K-means clustering largely depends on its initial cluster centers, and given the randomness of cluster center initialization in K-means, those BoW- based methods that utilize basic K-means often show a substandard performance. This paper presents a method to improve the accuracy of content-based medical image retrieval by optimizing the initial cluster centers of K-means by an ensemble clustering approach.In this method, after extracting the local features by SIFT descriptor, K-means clustering with random initialization will be run several times to yield several base clustering, and then a consensus algorithm selects, from among these initial results, the best clustering and more specifically the best visual words in terms of clustering criteria. This technique was tested on a group of standard x-ray images and the results showed that the use of ensemble clustering with MCLA serving as consensus algorithm leads to visual words that are more efficient and accurate than those created with randomly initialized centers.

Keywords: *content-based medical image retrieval,bag of words, SIFT descriptor,ensemble clustering.*

1. Introduction

The increasingly prevalent use of large medical image databases and their utility for medical data management, computer assisted diagnosis, research, and medical education and training necessitate the use of efficient content-based medical image search and retrieval systems. Today's digital medical images have numerous clinical applications and contain valuable information that can help healthcare professionals and other users with different levels of subject knowledge in universities and medical organizations to meet their information needs. As a result, image retrieval systems can play a vital role in effective use and dissemination of data in the academy and other organizations [1]. With the astonishing progress in information technology and Internet in the recent years, storage and retrieval of information and particularly images has become one of the most active lines of work in regard to multimedia systems [2]. These developments along with growing size of image databases and subjective

Article History:

DOI: IJMEC/10.225152
Received Date: Jun. 19, 2017
Accepted Date: Aug. 12, 2017
Available Online: Oct. 01, 2017

nature of tagging have made the text indexing increasingly dysfunctional. The efforts to resolve this problem have led to emergence and development of content-based image retrieval (CBIR) systems [3].

Typically, image retrieval systems operate in two stages. In the first stage, visual features of all images stored in the database will be automatically extracted and they will be indexed accordingly. In the second stage, system analyzes a user-provided query image to capture its low-level or visual features and searches for the image whose features are most similar to that query [4].

In response to deficiencies of text-based retrieval, many researchers have shown growing interest to content-based retrieval and developed several models for CBIR systems. The notable deficiencies of text-based retrieval include the time-consuming and costly process of manually annotating tags and keywords to images and subjective nature of tag and keyword assigned or selected by humans [5].

Unlike text-based retrieval, CBIR methods operate based on visual features such as color, texture, edge, shape, position and spatial relations of objects or areas within the image [6]. In recent years, CBIR systems have been developed for applications such as digital libraries, facial and fingerprint recognition, online shopping, trademark search, online search, and online publication, but only a limited number of these systems have been designed and implemented specifically for medical applications. Furthermore, most of the existing medical CBIR systems have been designed for images taken from a specific body member using a specific imaging system and are inapplicable to other images[7,8,9]. Thus, CBIR systems developed for general medical applications are very few in numbers[10,11,12].

In the last decade, the majority of large-scale image retrieval systems have been based on a concept called Bag of Features (BoF), which is mostly inspired by the concept of Bag of Words (BoW) in text mining. In BOF, each image is described as a set of order-less local features. BoW-based image retrieval consists of four main steps: 1- Feature extraction. 2- Construction of visual words: in the learning phase, clustering algorithms will be used to create a visual vocabulary, where the obtained cluster centers will serve as visual words. 3- Image representation and quantization: featured extracted from each image will be assigned to the constructed visual words, and images will be turned into a bag of words. 4- Retrieval: the histogram of query image will be compared with histogram of database images and similar images will be shown in output [14]. As mentioned, in the second step of BoW approach, a clustering algorithm must be used to construct visual words. The clustering algorithm most commonly used to obtain the cluster centers and construct the visual words is the K-means. K-means clustering consists of two phases: 1-initializing k random cluster centers, 2- assigning the features to the nearest center.

The most common metric for measuring distance from cluster centers is the Euclidean distance. Once features are assigned to cluster centers, the mean of all features assigned to each particular cluster will be set as the new center of that cluster. This process continuities until a new center is created. The quality of final cluster centers and constructed visual words largely depend on the choice of initial cluster centers. In the K-means clustering, initial centers are created randomly, so visual words may be described incorrectly, leading to lower retrieval precision. Also, because of this randomness, different runs of the algorithm with the same inputs may produce different outputs [15,16]. In this paper, ensemble clustering and specifically MCLA and HGPA [22] are used to produce optimal visual words and thereby improve clustering performance. The remainder of this paper is organized as follows: In section 2, the principles of ensemble clustering and related algorithms are explained. In section 3, the previous works on this subject are reviewed. In section 4, the proposed solution for medical image retrieval based on intelligent construction of visual words is described. In section 5, the results of tests conducted on the proposed method are presented. And finally, Section 6 concludes the paper**.**

2. Ensemble Clustering

Let X = { $X_1, X_2,\dots X_N$} be a set of N data points and let Π ={π_1, π_2, …, π_M}be a set of M base clustering results, which is referred to as a cluster ensemble. Each base clustering result (called an ensemble member) returns a set of clusters π_i={ C_1^i , C_2^i,…, $C_{k_i}^i$} such that $\cup_{j=1}^{k_i} C_j^i = X$ where k_i is the number of clusters in the i-th clustering. For each x ϵ X, C(x) denotes the cluster label to which the data point x belongs. In the i-th clustering, C(x) = j if x ϵ C_j^i . The problem is to find a new partition π^* of a data set X that summarizes the information from the cluster ensemble Π. The general framework of cluster ensembles is shown in Figure 2. Accordingly, multiple input clusterings, known as ensemble members or base clusterings, are intelligently aggregated to form a final data partition. There are two main stages of: (i) generating the cluster ensemble, and (ii) producing the final partition, which is normally referred to as a consensus function.

Base clusters are often produced by rerunning a clustering algorithm such as k-means (with random initialization), using random k-values in an unspecific interval with a set of unique parameters, using diverse subsets of data, or by the use of several different clustering algorithms [22].Having the ensemble, final partition can be obtained using a wide range of consensus functions. Consensus methods are generally divided into three main categories [22].

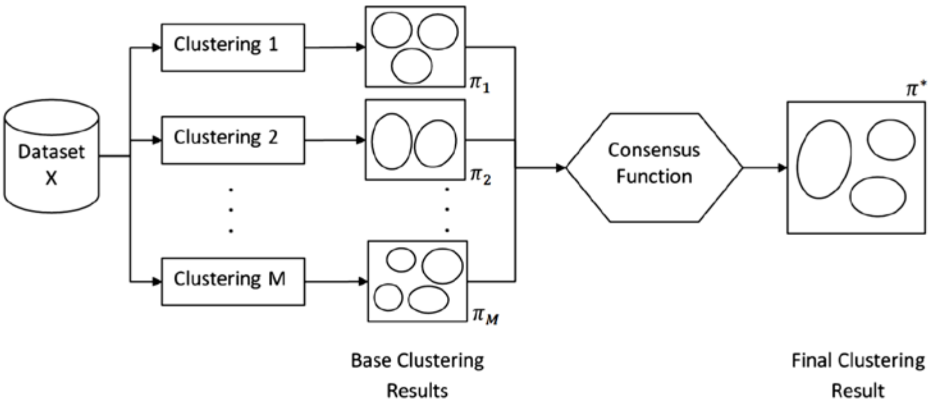

Figure 1: The Basic Process of Cluster Ensembles.

2.1 Pairwise similarity algorithm

In general, this method is based on pairwise similarity between data points. More specifically, for the data set X={$X_1, X_2, \dots ,X_N$}, this method first produces the set of clustering results (cluster ensemble) $\Pi = \{\pi_1 , \pi_2 , \dots ,\pi_M\}$ by obtaining M base clustering for the data set X, and then creates, for every member of the ensemble, a N×N similarity matrix denoted by S_m,m=1,2,...m, where each entry represents the relationship between two data points and equals 1 if both points are assigned to the same cluster, and is zero otherwise.

2.2 Graph-based methods

In this method, ensemble clustering problem is solved with the help of graph representations. Well-known examples of this approach include CSPA, HGPA, MCLA, HBGF. Cluster-based Similarity Partitioning Algorithm (CSPA) first creates a similarity graph where vertices represent data points and edges' weights represent the similarity scores obtained from the CO matrix. Then, a partitioning algorithm called METIS is used to partition the similarity graph into K clusters. Hyper Graph

Partitioning Algorithm (HGPA) first creates a hyper-graph, where vertices represent data points and hyper edges with equal weights represent clusters in the ensemble. HMETIS technique is then used to partition the graph into K equal parts. MCLA first constructs a graph representing the relation of cluster in the ensemble, where each vertex represents a cluster in the ensemble and each edge's weight between two vertices is obtained using the binary Jaccard measure. It then uses METIS technique to partition the graph into K clusters.

2.3 Feature-based approach

In this method, ensemble clustering problem is turned into a categorical clustering problem. More specifically, each base clustering is given a label, which is used as a new data point feature. Then, this feature is used to create the final solution.

3. Related Work

Caicedo et al. (2010) developed a multi-feature retrieval scheme, where texture feature were combined and DCM and SIFT descriptors were used to construct feature histogram. They used SVM classifier for image retrieval based on probability of proximity of input image to the cluster. Their results showed that SIFT is less effective in classification than DCT [17]. Wang et al. (2011) presented a BoF-based medical image retrieval scheme, where each image was represented by a set of local features obtained by image segmenting and key points was analyzed by SIFT descriptor. Assuming that local features can be reconstructed by their neighboring visual words, they used a multiple assignment along with a visual word weighting scheme to improve the clustering and retrieval efficiency. In this work, the sub-similarity histogram of visual words was used as the measure of similarity [14].

In the CBIR method proposed by Shakila and Ravindran (2013), the visual content of images is analyzed by BoF-based approach. In this approach, visual features are handled in two main procedures: 1- Finding the dependencies between visual patterns and high-level features, or in other words, finding the most relevant features and clustering with the lowest redundancy 2- automatic image annotation. This method was tested on histology images and was found to have 80% accuracy when utilized alongside neural networks[18]. Vanegas et al. (2014) presented a CBIR method composed of unsupervised feature learning (UFL) and BoF approaches based on classical descriptors such as SURF, SIFT, and DCT. They stated that UFL can be utilized to automatically learn the visual invariance properties of color, scale and rotation. Results of their tests showed that their approach has a better retrieval efficiency than conventional descriptors [19]. Most of the above image retrieval methods are based on BoF approach with k-means clustering used for constructing visual words. But given the randomness of cluster center initialization in k-means clustering, the constructed words may be suboptimal, and this may undermine the image retrieval precision. Thus, in this study, cluster centers are initialized by optimal ensemble clustering to avoid this issue.

4. Proposed Method

Flowchart of the proposed image retrieval method is shown in Figure 2 and steps of this method are described in the following.

International Journal of Mechatronics, Electrical and Computer Technology (IJMEC)
Universal Scientific Organization, www.aeuso.org
PISSN: 2411-6173, EISSN: 2305-0543

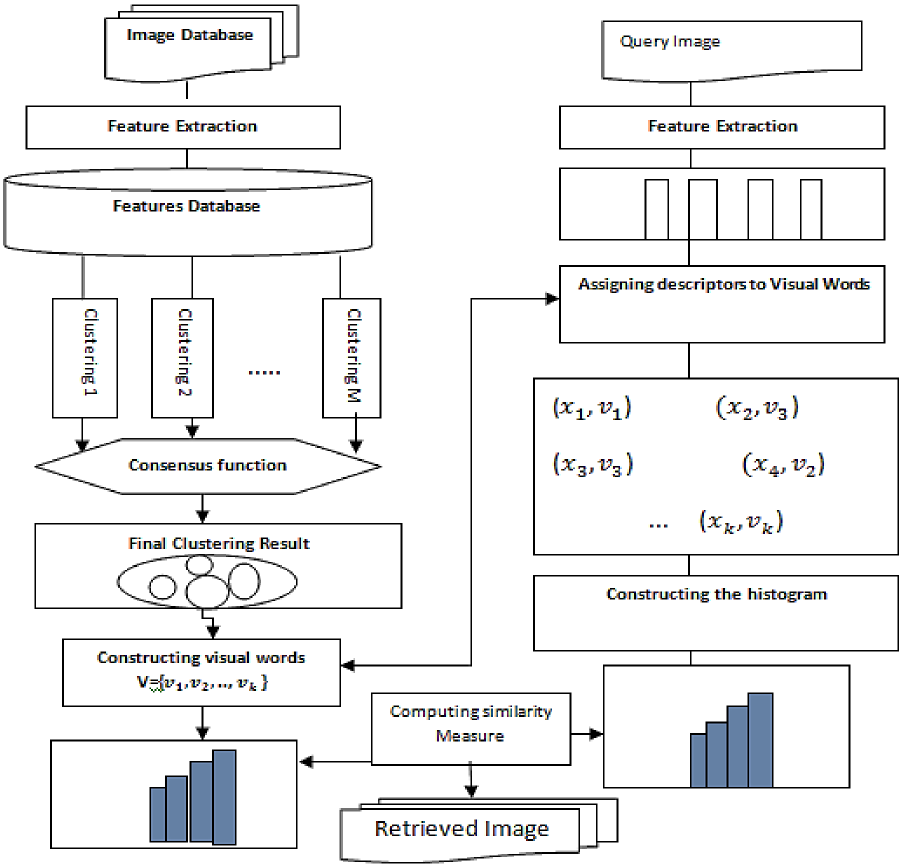

Figure 2: Proposed Method.

4.1 Feature extraction

Features are extracted by representing the input image as a set of small segments, and then capturing the local features and key points from these segments by the use of SIFT descriptor [21].

4.2 Construction of visual words

To construct visual words, K-means clustering (with random initialization) is used to obtain several base clusterings, which are then subjected to the ensemble function (here, HGPA and MCLA) to determine the final clustering. Next, centers are calculated by averaging the data points assigned to each cluster and resulting centers are considered as the centers of visual words.

4.3 Assigning features to visual words

In this step, each local feature x_l of the image X is assigned to the nearest visual word V_k in the vocabulary. Thus, image X is represented by the local features $\{(x_l, a_l)\}$ where a_l is calculated as follows:

$$a_l=\text{argmin}_K (D(V_K, X_l)) \quad (1)$$

In this relationship, $D(V_K, X_l)$ is the distance between code word V_k and feature X_l. After normalization of occurrences of visual word with the L1 norm, frequency vector of image X is created in the form of

International Journal of Mechatronics, Electrical and Computer Technology (IJMEC)
Universal Scientific Organization, www.aeuso.org
PISSN: 2411-6173, EISSN: 2305-0543

$HX = [h_1^x, h_2^x, \ldots, h_k^x]$ [1]. In this study, feature assignment is carried out by nearest neighbor approach[14].

4.4 Similarity evaluation

The similarity between histogram of query image and histograms of database images is often measured by criteria such as JSD and HIK. In this study, JSD, which has a good track record as image retrieval criteria, is used for this purpose[14].

5. Results

All steps of the proposed method were coded in MATLAB software, and then all tests were carried out on a laptop (Asus X550L) with 6Gb RAM and Corei7@3.1 GHz CPU. The proposed image retrieval system was evaluated by a set of x-ray images taken from different parts of the body from IRMA dataset . The Image Retrieval in Medical Applications (IRMA) database is a collection of 14,410 x-ray images that have been randomly collected from daily routine work at the Department of Diagnostic Radiology of the RWTH Aachen University. The downscaled images were collected from different ages, genders, view positions, and pathologies [23]. Each image in the dataset has an IRMA code. According to these codes, 193 classes are defined. In this study, 1409 of these images including six classes of spine, chest, breast, limbs, abdomen and skull were used for testing. In the course of test, SIFT feature extraction method was used to capture 140900 features from images, and each key point was represented by a 128 dimensional vector. Then, K-means clustering algorithm was run 10 times with random initialization and 376 clusters($k=\sqrt{140900}$). After obtaining the 10 base clusterings, the outputs of this step were turned into a 10 by 140900 matrix and then fed to the consensus function of MCLA and HGPA. The final clusterings, each containing 9 clusters, were then obtained from these algorithms. Given the high computation time of CSPA, this algorithm was not investigated. Having the optimal clustering, the mean value of members of each cluster (visual words) was considered as the new cluster center, and the resulting centers were used to create the histogram of every image. Features were assigned to visual words by the NN method. System was evaluated by the precision measure, which is known as a reliable criterion for this application. For this purpose, images were randomly selected from classes and precision of the system in retrieval of 5, 10, 15, 20, 30, 40, 50 and 100 images was measured with MCLA and HGPA used as ensemble algorithm and also without these algorithms (based directly on randomly initialized centers of K-means).

$$\text{Precision} = \frac{\text{Number of relevant Images retrieved}}{\text{Total number of Images retrieved}}$$

Table 1: Retrieval Precisions Using MCLA Algorithm For Ensemble Clustering .

5	10	15	20	30	40	50	100	class
1	1	1	1	1	0.87	0.74	0.41	spine
1	0.8	0.53	0.5	0.43	0.4	0.32	0.21	breast
0.8	0.9	0.86	0.9	0.93	0.9	0.92	0.86	chest
1	0.9	0.73	0.7	0.5	0.5	0.48	0.37	limbs
1	1	1	0.85	0.83	0.8	0.78	0.79	abdomen
0.8	0.9	0.86	0.85	0.9	0.9	0.86	0.73	skull

Table 2: Retrieval Precisions Using HGPA Algorithm For Ensemble Clustering.

5	10	15	20	30	40	50	100	class
0.8	0.5	0.33	0.3	0.26	0.25	0.22	0.17	spine
0.4	0.3	0.2	0.15	0.13	0.125	0.1	0.1	breast
0.4	0.2	0.13	0.15	0.16	0.125	0.18	0.18	chest
0.4	0.4	0.33	0.4	0.3	0.3	0.3	0.25	limbs
0.8	0.9	0.86	0.85	0.73	0.7	0.66	0.66	abdomen
1	1	1	1	0.9	0.9	0.86	0.68	skull

Table 3: Retrieval Precisions Using K-means And Without Ensemble Clustring.

5	10	15	20	30	40	50	100	class
1	1	0.93	0.75	0.56	0.45	0.36	0.18	spine
0.8	0.5	0.33	0.3	0.26	0.25	0.22	0.17	breast
0.6	0.4	0.53	0.5	0.53	0.5	0.46	0.46	chest
0.8	0.7	0.66	0.6	0.5	0.45	0.44	0.34	limbs
1	0.9	0.93	0.9	0.86	0.9	0.9	0,80	abdomen
0.8	0.9	0.86	0.65	0.56	0.55	0.5	0.34	skull

As the results show, for all six classes, image retrieval precision achieved with MCLA used as consensus algorithm is better than the precision achieved with HGPA and without ensemble clustering (with random centers). The results also show that using random centers is more efficient than the use of HGPA as consensus algorithm. It was also found that with MCLA, images with easiest retrieval are those taken from the chest, abdomen and skull, and images with most difficult retrieval are those taken from the chest. Images of the breast were also the most difficult to retrieve when using the HGPA or random centers.

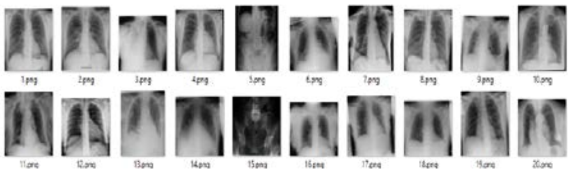

Figure 3: Top 20 Retrieved Chest Images Using MCLA

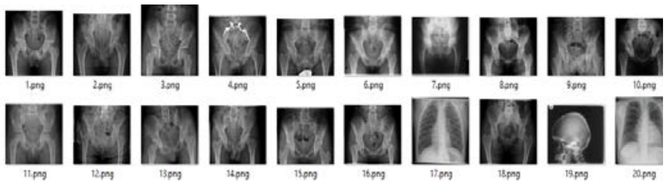

Figure 4: Top 20 Retrieved Abdomen Images Using MCLA

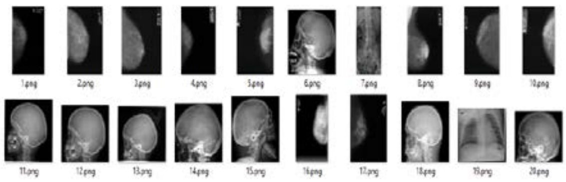

Figure 5: Top 20 Retrieved Breast Images Using MCLA

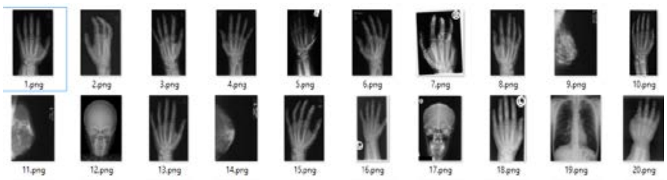

Figure 6: Top 20 Retrieved Limbs Images Using MCLA

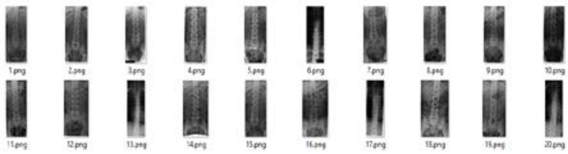

Figure 8: Top 20 Retrieved Spine Images Using MCLA

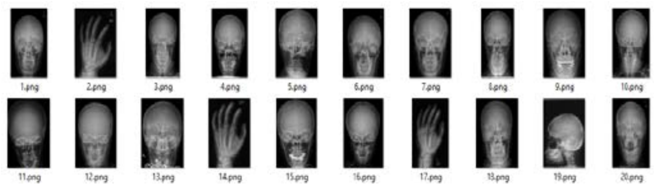

Figure 8: Top 20 Retrieved Skull Images Using MCLA

Conclusion

Image retrieval based on the bag of visual words is one of the most widely used image retrieval methods. In this method, visual words are constructed by the popular k-means clustering, where cluster centers are initialized randomly. Thus, different runs of k-means clustering on the same inputs may produce different outputs, and this may undermine the retrieval precision. This study examined the possibility of medical image retrieval by optimizing the initial cluster centers of k-means by the use of an ensemble clustering approach. As the results showed, precision achieved with MCLA used as ensemble algorithm was better than the precision achieved with randomly initialized cluster centers.

References

[1] A. Sedghi, S; Sanderson, M; & Clough, P. (2008). A study on the relevance criteria for medical images. Pattern Recognition Letters, 29, 2046-2057.

[2] B. Goodrum, A. A. (2000). Image information retrieval: An overview of current research. Informing Science, 3, 63-66.

[3] C. Long, F; Zhang, H; & Feng, D.D. (2003). Fundamentals of content-based image retrieval.Multimedia Information Retrieval and Management, 4.

[4] D. Oliveira, M. C; Cirne, W; & De Azevwdo Marques, P. M. (2007). Towards applying content-based image retrieval in the clinical routine. Future Generation Computer Systems, 23, 466-474.

[5] E. Smeulders, A. W. M; Worring, M; Santini, S; Gupta,A; & Jain, R. (2000). Content-based image retrieval at the end of the early years. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22, 1349-1380.

[6] F. Liu, Y; Zhang, D; Lu, G; & Ma, W. Y. (2007). A survey of content-based image retrieval with high-level semantics. Pattern Recognition, 40,262-282..

[7] G. Greenspan, H; & Pinhas, A.T. (2007). Medical image categorization and retrieval for PACS using the GMM-KL framework. Ieee Transactions on Information Technology in Biomedicine, 11, 190-202.

[8] H. Rahman, M. M; Desai, B. C; Bhattacharya, P.(2008). Medical image retrieval with probabilistic multi-class support vector machine classifiers and adaptive similarity fusion. Comput Med Imaging Graph, 32, 95-108.

[9] I. Yao, J; Zhang, Z. M; Antani, S; Long, R; & Thoma, G. (2008). Automatic medical image annotation and retrieval. Neurocomputing, 71, 2012-2022.

[10] J. Chu, W.W; Su, C.C; Cardenas, A.F; & Taira, R. K.(1998). Knowledge-based image retrieval with spatial and temporal constructs. Knowledge and Data Engineering, IEEE Transactions on, 10, 872-888.

[11] K. Lehmann, T. M; Guld, M. O; Thies, C; Plodowski, B; Keysers, D; Ott, B; & Schubert, H. (2004).IRMA-content-based image retrieval in medical applications. Medinfo, 11, 842-846.

[12] L. Antani, S; Lee, D; Long, L.R; & Thoma, G.R.(2004). Evaluation of shape similarity measurement methods for spine X-ray images. Journal of Visual Communication and Image Representation, 15, 285-302.

[13] M. Yu, J., Qin, Z., Wan, T., & Zhang, X. (2013). Feature integration analysis of bag-of-features model for image retrieval. *Neurocomputing*, *120*, 355-364.

[14] N. Wang, X., Yang, M., Cour, T., Zhu, S., Yu, K., & Han, T. X. (2011, November). Contextual weighting for vocabulary tree based image retrieval. In *Computer Vision (ICCV), 2011 IEEE International Conference on* (pp. 209-216). IEEE..

[15] O. Erisoglu, M., N. Calis, and S. Sakallioglu, A new algorithm for initial cluster centers in k-means algorithm. Pattern Recognition Letters, 2011. 32(14): p. 1701-1705.

[16] P. Jain, A.K., Data clustering: 50 years beyond K-means. Pattern recognition letters, 2010. 31(8): p. 651-666.

[17] Q. Caicedo, J. C. and E. Izquierdo (2010). "Combining Low-level Features for Improved Classification and Retrieval of Histology Images." T. Mass-Data Analysis of Images and Signals 2(1): 68-82.

[18] R. Ravindran, U. and T. Shakila (2013). "Content based image retrieval for histology image collection using visual pattern mining." International Journal of Scientific & Engineering Research 4(4).

[19] S. Vanegas, J., et al. (2014). Unsupervised feature learning for content-based histopathology image retrieval. Content-Based Multimedia Indexing (CBMI), 2014 12th International Workshop on, IEEE.

[20] T. Kato, T. (1992, April). Database architecture for content-based image retrieval. In *SPIE/IS&T 1992 symposium on electronic imaging: science and technology* (pp. 112-123). International Society for Optics and Photonics.

[21] V. Lowe, D. G. (2004). "Distinctive image features from scale-invariant keypoints." International Journal of Computer Vision 60(2): 91-110.

[22] W. Strehl, A., & Ghosh, J. (2002). Cluster ensembles---a knowledge reuse framework for combining multiple partitions. *Journal of machine learning research*, *3*(Dec), 583-617.

[23] X. Henning Muller, Paul Clough, Thomas Deselaers, and Barbara Caputo, ¨ ImageCLEF: Experimental Evaluation in Visual Information Retrieval, 2010.