Medical Image Categorization and Retrieval System in Radiology Using Bag of Visual Words Framework

S.Mahalakshmi and G.Raghuraman
Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam. Chennai, Tamil Nadu, India
mahalkshm92@gmail.com, raghurameng@ssn.edu.in

Abstract **- In This work we present an efficient image categorization and retrieval system applied to Image Clef 2009 medical image retrieval task. In this task we have presented methodology is based on local patch representation of the image content and a bag-of-features approach for defining image categories, with a kernel based SVM classifier. Two main tasks are addressed: First organ identification task; second the detection and identification of pathologies, i.e. shifting from the organ level to pathology level analysis. We used a large generic archive of 12,000 radiographs (IRMA) to tune the system parameters. We demonstrate automated organ detection on the IRMA collection as well as the generalization to a new data collection. We submitted one run, using support-vector-machines trained on the visual word histograms in multiple scales.**

We proposed system was helped to find discriminating orientation and body regions in X-ray images also organ-level discrimination we show an application to pathology level categorization of chest X-ray data. Results indicate detection of pathology at a sensitivity of 88.4% and a specificity of 81%. This is first step towards similarity-based medical image categorization that has a major clinical importance in computer-assisted diagnostics. It can identify suspicious pathological X-rays and alert the referring clinicians to potential emergencies. Overall it is hoped that the development of such systems will contribute to the improvement of safety and quality of medical services

Keywords **- Bag of visual words, Computer-aided diagnosis, Chest radiography, Medical image retrieval, Image categorization, Image retrieval, Image patches**

I. Introduction

Content based image retrieval refers to the ability to retrieve images on the basis of the image content. In recent decades, researchers have been on developing Content Based Image Retrieval (CBIR) systems to index and retrieve medical images. One of the reasons behind this research area is that using text alone to retrieve images might not work correctly. Throughout the world, the rapid growth of computerized Medical Imaging using Picture Archiving and Communication systems (PACS) in hospitals has generated a critical need for efficient and powerful search engines. In recent years the Growing workload on radiologists the need for computerized assisted diagnosis systems which could help the radiologist in prioritization and the diagnosis of findings [1]. Automated image categorization and retrieval system could easily support such needs once algorithmic solutions are found for diagnostic-level categorization, even on such an elementary level as healthy vs. pathology In this work CBIR system for Medical image retrieval, we need to deal with x-ray images and a technique for retrieving images on the basis of automatically-derived features. In medicine physicians and researchers are interested in being able to retrieve medical images based on low level features. Research on medical image retrieval using CBIR systems is not limited to text searches only but also extends to strategies that combine visual features of images with the text retrieval techniques. Some groups participating in the Cross Language Evaluation Forum (CLEF) are now using this approach, and the preliminary results from this combination of techniques seem to perform better with higher accuracy when compared to studies using only one technique. Thus this approach has the potential to improve the accuracy of CBIR systems. This would make these systems more helpful for radiologists in medical settings, researches in medical analysis, and medical students as well as teachers in academic healthcare environments. This CBIR system has all the basis capabilities needed for medical images and text retrieval thus the goal of this research is to boost the system performance by introducing some semantic types that are important to the retrieval of queries two different, parallel developments pursued are to be pursued in the IRMA project. Automatic classification using global image descriptions of X-rays with regard to the imaging modality, taking direction, body region-researcher and system function, Determination of diagnostically relevant local features to classified and registered images. The goal of the medical task is to retrieve relevant images based on an image query.

II. Related Works

As an important complementary search approach, content-based image retrieval (CBIR) has been one of the most active research areas in the field of computer vision over the last decade. In the medical field, CBIR also draws extensive attention [2]. Traditional global features include color features, texture features, and shape features. Recently, along with the rapid progress in the application of local descriptors in pattern recognition, computer vision, and image retrieval, the bag-of-features based methods derived from local features like key points or image patches have demonstrated promising performance on object classification and image retrieval tasks [1], [2]. Unlike text retrieval, image retrieval should create visual words first. Usually, k-means is adopted to cluster centers of features which are extracted from all images. These cluster centers are then used as a vocabulary for all the images to obtain word vector representations.

Avni et al. proposed an X-ray image categorization and retrieval method using patch-based visual word representations [1], while Zhi et al. developed a medical image retrieval method using scale invariant feature transform (SIFT) features [2]. Caicedo et al. conducted a comparison of different representations obtained from the bag-of-features approaches to classify histopathology images, including both the image patches and SIFT local features [4]. All the above methods build the histograms for image representation by assigning the local image feature descriptors to the single nearest visual word in the vocabulary, which is called nearest neighbour (NN) assignment in this paper. However, one inherent component of the transitional NN model is the assignment of the discrete visual words to the continuous image features, which shows a clear mismatch of this hard assignment with the nature of continuous features [5]. By explicitly modeling the ambiguity of visual word assignment, Van Gemert et al. improved the classification performance compared to the hard assignment of the traditional codebook model based bag-of-features methods [5]. However, the assignment is based on the usage of Gaussian kernel, which is very sensitive to the smoothing parameter .Jegou et al. increased the classification performance by using multiple assignment of descriptors to visual words at the cost of reduced efficiency [6]. The disadvantage of this method is

That it treats the entire candidate nearest neighboring visual words equally without considering the neighbourhood structure of the descriptors and the visual words. In [7], Yang et al. developed an extension of the spatial pyramid matching (SPM) method by generalizing the NN assignment based vector quantization to sparse coding (SC) followed by the multi scale spatial max pooling, and proposed a linear SPM kernel based on the SIFT sparse codes. They argued that the NN assignment may be too restrictive, giving rise to a coarse reconstruction of the local feature space. They relaxed the constraint by putting a -norm regularization on cluster membership indicators, which enforced cluster membership indicators to have a small number of nonzero elements. However, this method assumes that a local feature is reconstructed by all the visual words in the vocabulary, which causes complex computations. In this paper, we present a novel multiple assignment method by assuming a local descriptor can be linearly reconstructed by its neighbouring visual words. We will demonstrate that the local reconstruction assignment performs better than the global reconstruction methods such the SC assignment, on medical image retrieval tasks.

The GMM-KL Gaussian mixture modelling framework is used for matching and categorizing X-ray images by body regions. GMM-KL framework is a localized statistical framework for medical image retrieval. Image representation and matching framework for image categorization in medical image archives uses this framework. GMM-KL framework is a localized statistical framework for medical image retrieval. Image representation and matching framework for image categorization in medical image archives uses this framework. We are currently developing more efficient approximations for KL in order to enable such large archive processing [3].

III. Research Methodology

This work we present a patch based classification and retrieval system that is based on the bag of Visual Words (Bow) paradigm. This approach is recently introduced concept that has been successfully applied to scenery image classification tasks (see e.g. [8, 9, and 10]). . The use of Bow techniques for large scale radiograph archive categorization can be found in the Image CLEF competition. This approaches using patch-based, bag-of-visual-words concepts are gradually emerging in medical tasks.

The Bow model is based on the idea that it is possible to transform the image into a set of visual words and to represent the image (and objects within the image) using the statistics of appearance of each word as feature vectors. In our system the visual words are image patches (small sub images) that are clustered to form a dictionary consisting of a small set of representative patches. We utilize the Bow approach while implementing modifications which are relevant for medical images. A main advantage of this approach is avoiding the need for explicit object detection features. Previous methods are based on explicitly predefined features that are locally extracted from the image (e.g. gradient orientation, edge, line length and orientation). The proposed approach, which is based on medium size image patches, avoids the need for explicitly specified medical object features. Instead, the features are implicitly found as part of the unsupervised learning step composed of building the visual dictionary.

The patch-based image representations and "bag-of-features" classification techniques have been proposed for general object recognition tasks. In these approaches, a shift is made from the pixel entity to a patch – a small window centered on the pixel. In its most simplified form, raw pixel values (intensities) within the window are used as the components of the feature vector. It is possible to take the patch information as a collection of pixel values, or to shift the representation to a different set of features based on the pixels, such as SIFT features, and reduce the dimensionality of the representation via dimensionality reduction techniques, such as principle component analysis (PCA).

A very large set of patches are extracted from an image. Each small patch shows a localized "glimpse" at the image content; the collection of thousands and more such patches, randomly selected, have the capability to identify the entire image content (similar to a puzzle being formed from its pieces). A dictionary of words is learned over a large Collection of patches, extracted from a large set of images. Once a global dictionary is learned, each image is represented as a collection of words (also known as a "bag of words", or "bag of features"), using an indexed histogram over the defined words. The matching between images, or between an image and an image class, can then be defined as a distance measure between the representative histograms. In categorizing an image as belonging to a certain image class, well-known classifiers, such as the k- nearest neighbor and support-vector machines (SVM) .are used

The basic flow of this project is as follows:

Collect a lot of features from patch image .Use k-means to cluster those features into a visual vocabulary.

1. In the learning phase, we construct a visual vocabulary using a clustering algorithm usually; k-means is used to cluster centers of features which are extracted from all images in the database. These cluster centres are then used as a vocabulary (codebook) with visual words for all images to get word vector representations.

2. For each of the training images build a histogram of the word frequency (assigning each feature found in the training image to the nearest word in the vocabulary).Feed these histograms to an SVM. Build a histogram for test images and classify them with the SVM based on trained set.

3. First organ identification task done for given test or train image then pathology level analysis will be done. for each image we add label it will mention that image is healthy or pathology image

Bag of visual words

This model is simple, visual “words” (or SIFTS in our case) can be associated with particular images. Given a dictionary of vocabulary “words”, we learn the distribution of these words across images from each image. To do this we use an SVM that creates a model of “words”. We will use the SVM later to classify new test images. We will find the histogram of the vocabulary “words” for a test image and then try to match it to one of the image distributions with the SVM. The visual words model describes an image using a set of visual words called visual vocabulary. The vocabulary is obtained by clustering local features Extracted from images where each resulting cluster is a visual word. In this model, an image is finally represented by a histogram, where each bin corresponds to a visual word and the associated weight represents its importance in the image. Thereby, the construction of the histogram requires three steps: A) extracting visual features, B) building a visual vocabulary and C) indexing images.

The bag of words model workflow

1. Build Vocabulary (Dictionary)List for a Collection of Images
1. From each image, collect a lot features
2. Perform k-means to cluster these means into a visual vocabulary
2. Build a SVM Model for each Image
1. For each training image, build a histogram of the visual word frequencies
2. For each image, pump the corresponding training image histograms into an SVM to create a representative model
3. Classify a Test Image
1. Find the histogram of the visual vocabulary word frequencies
2. Test that histogram against the SVM models from each scene
4. Assign the image to the dataset that produced the best results

The dictionary building process extracts patches of a fixed size of 9x9 pixels with a grid of 6 pixels spacing. Patches are normalized to have 0 mean and 1 variance. We then compute a covariance matrix of a set of roughly 2,000,000 patches, and apply PCA to find its eigenvectors. Patch center coordinates are added to the feature set, in order to include information about the visual words layout. Running k-means algorithm on this set produces 1000 dictionary visual words The dictionary building process is repeated in 3 image scales: full resolution, 1/2 scale and 1/8 scale. The resulting dictionary is a concatenation of the 3 dictionaries from the 3 scales In the image representation step patches are extracted from each image using a dense grid around every pixel. An image is represented as a word histogram over the multi-scale dictionary.

Image classification is performed on the word histograms by an SVM classifier. Multi-class classification is implemented using one-vs-one heuristic .In the training step each IRMA code is treated as a separate label, without using the hierarchical nature of the code.

Dataset used

In this project we used Image CLEF (Cross Language Evaluation Forum) 2009 medical retrieval database used database consists of scanned X-Ray Images of different parts of human body Images. The goal of the CLEF medical retrieval task is to advance the performance of multimedia objects retrieval in the medical domain combining techniques from Information Retrieval and Content Based Image Retrieval (CBIR). A database contains 12729 fully classified radiographs for both training and test set. It was taken randomly from medical routine, is made available and can be used to train a classification system. It is taken randomly from medical routine, is made available and can be used to train a classification system. Images are labelled according to classification label sets. Images in the archive are labelled according to the IRMA coding system, with each category described by four axes: 1) A technical axis that describes the image modality; 2) a directional axis that defines body orientation, 3) an anatomical axis that describes the body region examined, and 4) a biological axis that describes the biological system being examined. The axes have a hierarchical description. To implement system we will take 1000 train and test images from Image CLEF 2009 database.

Using this train and test dataset we will perform image retrieval process.

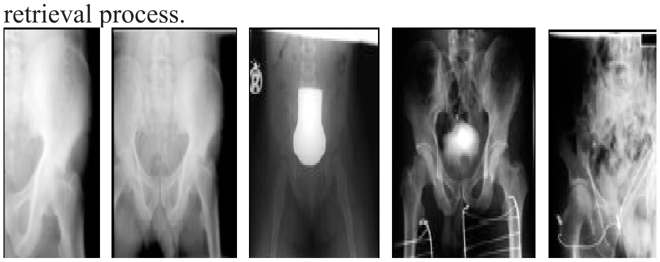

Fig. 1. Example of x-ray images belonging to the IRMA database with high intra-category Variability. All images share the same IRMA code 1121-120-800-700.

IV. Proposed Works

The basic steps of the bag-of-features image retrieval framework

A. Patch Extraction

In this process given input image is divided in to small patches. Feature detection approach is used here. We choose patch size 9*9.Patches are extracted every pixel in the image. Each patch shows a localized view of the image content. We chosen needs to be larger than pixels, since to capture the edges and corners present in the image. Common feature detection approaches are regular sampling grid, a random selection of points, or the selection of points with high information content using salient point detectors. We used all the information present in the image, by sampling rectangular patches of fixed size 9*9 around every pixel in the image. Extracted patches are normalized by subtracting its mean gray level, and divide the image by its own standard deviation. This will increase the brightness and provides local contrast enhancement within a patch. Whose patch images having single intensity value that images are abounded. Patches are normalized to have a zero mean and unit variance. Output of this process is a patch image collection.

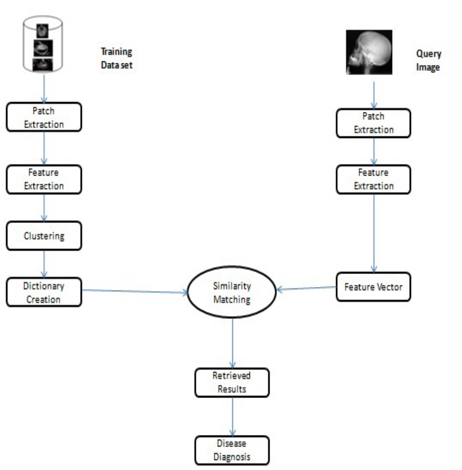

Fig.2 Bag-of-Feature Based Medical Image Retrieval Framework

B. Feature Extraction

In this process Normalized patch image is given input to this process. Patch image are projected into feature space using feature descriptors. Feature extraction was done using SIFT descriptors. It will detect and describe local features present in the image. Patch image Edge feature are extracted using the canny edge detector. Corner feature also extracted from a patch image. The dimensionality reduction in the feature space using the a principal component analysis procedure (PCA).It will reduce both computational complexity and the level of noise .Feature values are obtained for each image is combined to form a feature vector then addition of spatial co-ordinates to the feature vector .This will introduce spatial information into the image representation

C. Quantization

Feature collection given input to this process. Features are quantized then vector represented patches into visual words to generation of dictionary a visual word considered as a representative of set of similar patches. To perform K-means clustering over the vectors of the initial patch collection using K-means algorithm, and then cluster them into K groups in the feature space. The resultant cluster centres serve as a vocabulary of K visual words. In this process we generated dictionary of 1000 visual words .Due to the we added spatial coordinates as part of the feature space, the visual words have a localization component in them, which is reacted as a spatial spread of the words in the image plane.

D. From An Input Image To Representative Histogram

The most important step in CBIR is efficient image representation. Goal of the image representation step is to move from a 2d image to a vector of numbers Should preserve information of the image content enough to classify it correctly at the same time not be sensitive to object placement, artifacts and image quality In this histogram representation of given (training or testing) image over unique distribution of generated dictionary of visual Words. In our implementation, patches are extracted from every pixel in the image. The patches are projected into the selected feature space, and translated (quantized) to indices by looking up the most similar feature-vector in the generated dictionary. The dictionary lookup process is accelerated by comparing a new patch only to dictionary words at a certain radius from it. The dictionary generation process and the shift from a given image to its representative histogram. In this histogram representation we add spatial features. It will preserve both the image local content and spatial layout present in the image.

E. Image Classification

Image classification is based on a ground truth of manually categorized images. Classification was done using a SVM classifier with the histogram intersection kernel. Images are classified according to IRMA categorization of images.

Histogram intersection has no free kernel parameters, which makes it convenient for fast parameter Evaluation. The two other kernels have a free trade off parameter, and require careful optimization. In order to classify multiple categories

We use a binary classifier, where N (N - 1)/2 binary classifiers are trained for all pairs of categories present in the dataset. Whenever an unknown image is classified with a binary classifier it casts one vote for its preferred class, and the final result is the class with the most votes. Since each binary classifier runs independently, parallelization of both training and testing phases of the SVM is straight forward. It is implemented as a parallel enhancement of the LIBSVM library

F. Image Retrieval

Similarity measurement process to measure the distance between the given query image and the target image, the distance is where runs on the bins. Euclidean distance measurement method used to calculate distance between two images. The image retrieval result is not a single image but a list of images ranked by their similarities with the query image.

G. Performance Evaluation

Performance measures can be used to determine the degree to which the system reflects the notations of similarity desired by the user. Our system will give more relevant images than the existing method. The performance of the content based image retrieval system is evaluated after the execution of N number of queries using the measures like Precision and Recall.

Precision is the fraction of retrieved images that are relevant to the search.

$$\text{Precision} = \frac{\text{Number of Relevant images Retrieved}}{\text{Number of Retrieved Images}}$$

Recall is the fraction of relevant images that are retrieved.

$$\text{Recall} = \frac{\text{Number of Relevant images Retrieved}}{\text{Number of Relevant Images}}$$

This image retrieval system achieved overall classification rate 89.1%. The total running time for the whole system, training and classification, was approximately 40 minutes on the full resolution images, and 3 minutes on the 1/4 scaled down images. The retrieval system is also computationally efficient, with an average retrieval time of less than 400 ms per query

Chest X-Ray Pathology Detection And Classification: Sheba Archive

This medical image retrieval system that has demonstrated very strong classification rates while also providing efficiency in the retrieval process. The system has been applied to several large radiograph archives. We have recently applied it within the Image Clef competition [11], and demonstrated strong results. The topic of retrieval becomes of value on the clinical front, once the content involves a diagnostic-level categorization, such as healthy vs pathology. In a collaborative effort with Sheba medical center, a large academic medical facility, we address this concept in the identification and categorization of x-ray lung disease.

In this section we shift from organ-level analysis to a pathology-level analysis. We applied our system to chest X-rays obtained in the emergency room of Sheba Medical Center. We used 98 frontal chest images in DICOM format from the hospital PACS, taken during routine examinations.

X-ray interpretations, made by two radiologists, served as the reference gold standard. The radiologists examined all of the images independently; they then discussed and reached a consensus regarding the label of every image. For each image and pathology type, a positive or negative label was assigned: 38 of the images were diagnosed as normal, 55 images had at least one pathology and the other five images were labelled as inconclusive. Fig. 5 shows a set of healthy (a)–(c) and pathological images (d)–(m). Pathology data include 24 images with enlarged heart shadow [three examples shown in Fig. 5(d)–(f)], 19 images with enlarged mediastinum, Fig. 5(g)–(i), 17 images with right pleural effusion and 21 images with left pleural effusion, Fig. 5(j)–(l). Some patients had multiple pathologies. For example, Fig. 5(m) exhibits all Pathologies. We treated the multiple pathology detection as a

Set of binary classification tasks, where in each task we tried to detect an individual pathology.

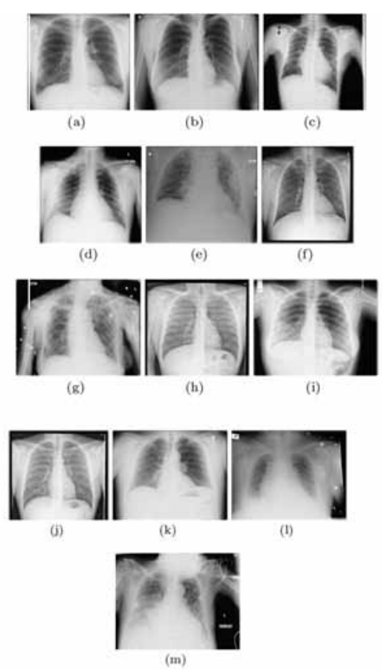

Fig.3 Frontal chest x-ray images, Sheba medical-center: (a-c) Healthy; (d-f) Enlarged heart; (g-i) Lung infiltrate; (j-l) Left or right effusion; (m) Multiple

pathologies: enlarged heart, lung infiltrate, left and right effusion.

We started by resizing the original high-resolution DICOM images to a maximal image dimension of 1024 pixels, and maintained the aspect ratio. We then followed the Feature extraction step to extract features, build a Visual dictionary, and represent an image as a histogram of visual words in multiple scales. We then detected each of the four pathologies using a binary SVM classifier, with a histogram intersection kernel. In addition to individual Pathology detection, we trained a classifier to distinguish between healthy images vs. a non-healthy image (with any Kind of pathology). This type of classifier can be useful for initial screening of suspicious images, in order to prioritize the radiologist's work.

In the task of individual-pathology detection, performance depended on the pathology type: it was fairly accurate in detecting enlarged hearts, with a sensitivity of 75:56% and specificity of 83:46%, and slightly less accurate in detecting lung infiltrates and effusions, which are more subtle findings. Frequently, research focuses on lung nodules. In this work we Looked at other areas beyond pulmonary nodules that could Benefit from computer-aided detection and diagnosis (CAD) in chest radiography. These include interstitial infiltrates, Right and left pleural effusion and cases of enlarged heart.

V. Comparison With Existing Work

In previous research work Image CLEF contest has text-based and image-based retrieval. Contest mainly based on the IRMA project X-ray library image which consists of medical radiographs. Images are classified by medical experts according to the imaging modality, the examined region, the image orientation with respect to the body and the biological system under evaluation. The Image CLEF provide for the analysis of feature spaces global versus local, similarity measures as well as classification schemes. Classification of images is also problem in existing research. A global image representation can miss-detect the pathology and result in a misclassification of the image. Due to incorrect classification of medical image it leads to incorrect identification of disease in a pathology result. Image classifier accuracy was less and

Also Image retrieval speed was less. This three key issues are addressed Accuracy of results, Memory usage, Query time using bag of visual words approaching this system.

In this proposed system we evaluated Two key characteristics are the representation of the patch as normalized raw values versus SIFT, and the use of spatial features as part of the representation space. We find that the using the raw (pixel) data, with minimal processing, it will gives good results, as long as a large amount of data are used. In this system we increase the dictionary size up to 1000 visual words. We show the percentage of correct classification averaged over 10 runs. In each run we used randomly chosen 10 667 images for training and the rest 2000 images were used in the test process. . Our system was tuned to achieve high accuracy in general medical image classification and retrieval. Low memory space was needed due to quantization of features. Our system will give more relevant images than the existing method. In this work we improved overall performance of the Medical image retrieval system. This system was tuned to achieve high accuracy, with an average of over 95% correct Classification of medical images .We show initial capabilities in image categorization into healthy vs pathology, along with discrimination into one of the pathology states. These capabilities can be generalized to larger data collections as well as additional pathology families. This system is also computationally efficient, with an average retrieval time of less than 400 ms per query

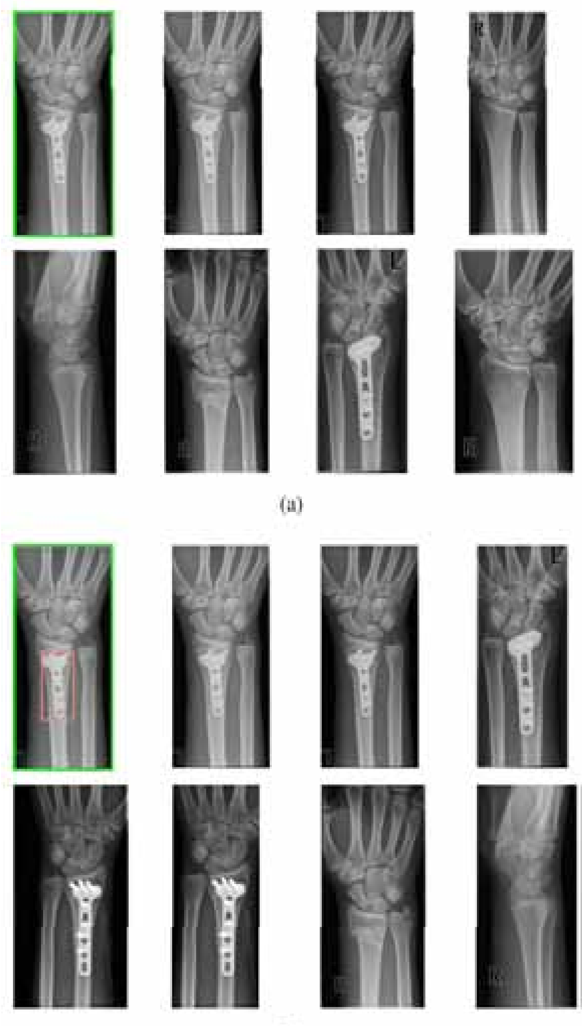

Fig. 4: A aample ROI query and retreival (a) Full image query (b) Query with selected ROI

VI. Experiments and Results

In this section we evaluate the proposed system for automated organ detection task and image retrieval task using Bag of visual words paradigm. We first investigate the sensitivity to various parameters that define the system. We then show classification and retrieval experiments on large radiograph archives. We focus on three components of the system: finding the optimal set of local features, finding the optimal dictionary size, and optimizing the classifier parameters we use large radio graphics images to tune the image retrieval system. The optimal parameters set was used in all four classification tasks.

We optimized the system parameters using several cross validation Experiments. In the following experiments, 10,667 images were used for training and 2000 randomly drawn images were used for testing and verification. In addition of spatial co ordinates to the patch as additional features it will improve the classification accuracy. The patch variance normalization step improves the classification rate as well: with no normalization, the average classification rate is 88:19, while with normalization it climbs to 90:9. Using SIFT features with the SVM classifier increased significantly the feature extraction time, and achieved an average of 85:4% classification accuracy; well below the classification rate of the raw patch based classifiers The three feature sets tested: raw patches, normalized patches and the 128 dimensional SIFT descriptors, were reduced in dimension using PCA. Classification was done using an SVM classifier with the histogram intersection kernel.

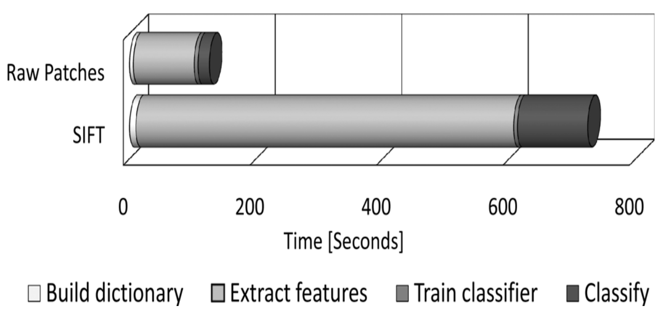

Fig.5 Running time using SIFT descriptors and normalized raw patches.

The advantage of using normalized raw patches over the SIFT descriptors is even more significant when considering the computational cost of the process. Using raw patches, the feature extraction step was significantly faster than with SIFT descriptors

We used the SVM classifier with three possible kernels: theHistogram intersection, the Radial Basis Function and the chi-square kernels. We used the optimal features and dictionary size consistently across all experiments. The system uses a set of densely extracted normalized raw patch features, with seven PCA components, spatial features with weight 6, and 1000 visual words. For classification we used the SVM algorithm with a kernel.

For image comparison, distance measures between the representative histograms were used. Retrieved Images were ranked by the distance between the targets Histogram and the histogram of the query image. When there were multiple query images, we used the minimal distance between the target and the query set.

A. ROI-Based Retrieval Results

A sample ROI query and retrieval are shown in Fig. 7(a). The query image (top left) is a left arm with a metal fixation device. Retrieved images are returned by order of similarity from left to right, top to bottom. Given query image is part of the database, the first returned image is the query image itself. Except for the first image, images 2 and 5 have similar images. In Fig. 7(b) the user selects the metal fixation device as a region of interest. The difference of visual words in the ROI is multiplied by inside the ROI, and outside the ROI. The selection of an ROI in this case retrieved images with a fixation in the top five returned images. This exemplifies how a simple weighting of the distance function can be used to locate an interesting object in the database. Since the coordinates are part of the features, the similarity distance is not invariant to large translations of the ROI. This method is therefore limited to locating similar objects in the near vicinity of the ROI.

VII. Conclusion

We presented a medical image categorization and retrieval system for large medical databases, based on compact bag-of-features image Representation. The system achieves comparatively good results in the Image Clef 2009 medical image retrieval task challenge, while maintaining efficient computation times. We provided a comprehensive overview of the methodology and its application to Image CLEF and in clinical settings. Statistical analysis of the results is shown on both the CLEF dataset, on the organ-level, and the Sheba chest X-ray dataset, on the pathology level. Retrieval is discussed in both domains, with initial discussion into ROI-based retrieval. Future work involves extending the system to handle larger collection of chest images and pathology types.

In the future we also would like to evaluate our methodologies in the more complex Image CLEF 2009 Medical Image Retrieval task database in order to test if the conceptualization of the image Content smaller concepts pays off in the case of unbalanced training/test examples distribution. Future databases involving more image modalities would be interesting to work with as well. We believe that the system presents a general approach that can assist radiologists and provide a preliminary computerized prioritization tool when a radiologist is not available. Future work involves extending the system to handle larger collection of chest images and pathology types.

References

[1] U. Avni et al., “X-ray image categorization and retrieval using patch-based visual words representation,” in Proc. 2009 *IEEE Int. Symp. Biomed. Imag.: from Nano to Macro,* 2009, vol. 1 and 2, pp.350–353

[2] L. J. Zhi et al., “Medical image retrieval using SIFT feature,” in Proc.2009 2nd *Int. Congress Image Signal Process.,* 2009, vol. 1–9, pp.2252–2255.

[3] H. Greenspan and A. T. Pinhas, “Medical image categorization and retrieval for pacs using the gmm-kl framework,” *IEEE Trans. Inf. Technol. Biomed.,* vol. 11, no. 2, pp. 190–202, Mar. 2007

[4] J. C. Caicedo, A. Cruz, and F. A. Gonzalez, “Histopathology image classification using bag of features and kernel functions. Artificial intelligence in medicine,” in *Proc. 12th Conf. Artif. Intell. Med.* (AIME 2009), 2009, p. 126

[5] J. C. van Gemert et al., “Visual word ambiguity,” IEEE Trans. Pattern Ana. Mach. Intell., vol. 32, no. 7, pp. 1271–1283, Jul. 2010.

[6]] H. Jegou, M. Douze, and C. Schmidt, “Improving bag-of-features for large scale image search,” *Int. J. Compute. Vis.,* vol. 87, no. 3, pp. 316–336, 2010.

[7] J. C. Yang et al., “Linear spatial pyramid matching using sparse coding for image classification,” in *Proc. IEEE Conf. Compute. Vis. Pattern Recognition.,* 2009, pp. 1794–1801.

[8] L. Feiffer and P. Persona. A Bayesian hierarchical model for learning natural scene categories. In CVPR 05, volume 2, pages 524–531, 2005

[9] E. Nowak, F. Juries and B. Trigs. Sampling strategies for bag-of-features image classification. In ECCV, pages 490–503. Springer, 2006.

[10] M. Varna and A. Zisserman. Texture classification: are filter banks necessary? In CVPR 03, volume 2, pages 691–8, 2003.

[11] H. Muller et al. Overview of the image clef med 2008 medical image retrieval task. In CLEF working notes. http://www.clef-compaign.org/2008/working notes/CLEF2008 WN-Contents.html.