Minor Project Report

on

Content-based medical image retrieval using Dictionary Learning

Submitted by

Aparna PL 14IT132

Rajat N 14IT134

S. Balakoti 14IT238

under the guidance of

Dr. Sowmya Kamath

Department of Information Technology, NITK Surathkal

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

In

INFORMATION TECHNOLOGY

At

Department of Information Technology

National Institute of Technology Karnataka, Surathkal.

CERTIFICATE

This is to certify that the project entitled "Content-based medical image retrieval using dictionary learning" is a bonafide work carried out under my guidance for the course IT399, by Aparna PL, Rajat N and S. Balakoti, students of VI Sem. B.Tech (IT) at the Department of Information Technology, National Institute of Technology Karnataka, Surathkal, during the academic year 2016-17, in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Information Technology, at NITK Surathkal.

Place: Mangalore

(Signature of the Guide)

Date: 01-05-2017

Dr. Sowmya Kamath

Name and Address of Organization: NITK, Surathkal

Department of Information Technology
National Institute of Technology Karnataka, Surathkal

End Semester Evaluation, May 2017

Name of Student	**Register Number**	**Signature with Date**
Aparna PL	14IT132	
Rajat N	14IT134	
S Balakoti	14IT238	

Place: Surathkal

Date: 01-05-2017

Abstract

This project is based on retrieving similar images to a query image from a large database. Sparse representation of images by dictionaries is the core concept used. Dictionary learning algorithms (k-SVD) and matching algorithms (Orthogonal Matching Pursuit) are implemented for updating dictionaries and matching images respectively. The main features of the method are that it requires no training data and works well on the medical databases which are not restricted to specific context. This Content based Image Retrieval concepts are used to get efficient and satisfying results compared to the traditional methods of image retrieval

Contents

List of Figures

1 Introduction

The problem of searching for similar images in a large collection of images and displaying them by their rank(based on some similarity criteria) based on the actual visual content(and not on metadata associated with it) of the image is called Content Based Image Retrieval (CBIR). The classical approach to this problem is based on text associated with the image. This is called Text Based Image Retrieval(TBIR) which is based on selecting similar images based on annotations associated with the image. It has various practical limitations. One of them is that these annotations have to be provided manually by a human and as the size of the repository of images increases, it becomes highly tedious and time-consuming. It becomes increasingly difficult to adequately represent image content through text. To deal with the shortcomings of text based image retrieval, CBIR approaches have been proposed. As the number of hospitals purchasing picture archiving and communication systems (PACS) is increasing day by day, the medical images are increasingly acquired, transferred and stored digitally due to space and resource constraints to store them physically. The size of collection of medical images kept in hospitals has exploded in the recent past due to ever-growing reliance on the latest medical diagnostic tools such as histopathology(study of tissues), computerized tomography(cross-section of human body obtained through X-rays) and radiology(X-rays). Medical image retrieval is very important in the area of medical image repositories to ease the process of decision making in the clinical scenario. It has the ability to perform retrieval of images of similar characteristics (same disease and modality) and nature. The medical images consisting of various modalities are a crucial source of functional and anatomical information for the medical research, detection of diseases in patients and for educational purpose. In any typical CBMIR system(Content based medical image retrieval), subtle variations between medical images can't be claimed to be unrelated. Thus, a Content Based Medical Image Retrieval (CBMIR) system with rotation-wise, scale-wise, illumination-wise(or any other transformation) invariant is highly valued.

In the recent past, sparse coding i.e representing signals with sparse vectors received a lot of attention from the image and signal processing communities. It basically involves representing an image(or any signal at hand) as a linear combination of some atoms in a dictionary. Sparse coding is a robust technique for the purpose of processing of data in non-traditional ways efficiently. It is because any image(or signal) of interest can be sparsely represented(sparse-coded) with the help of some dictionary. Identification of this dictionary may be achieved by observing the properties/ characteristics of images at hand. Of late, dictionaries learnt from the signals at hand were observed to have many interesting applications. Many interesting dictionary learning methods i.e identifying the dictionary based on the characteristics of the signals, like Method of Optimal Directions (MOD) and K-Singular Value Decomposition were developed to code each member of the repository of signals with some sparse vector. The growing domain of compressed sensing is capable of making use of sparsity present in medical images. The work carried out in this project is in the direction of proposing a CBMIR technique that makes use of concepts of sparsity.

1.1 Motivation

Some of the shortcomings associated with the present CBMIR techniques are 1) in majority of the cases, practitioners have to search through a huge number of medical images for detecting similar images, which is a time consuming process. 2) Most of the present techniques for browsing similar medical images do it through text based retrieval techniques (TBIR). The TBIR system, unlike CBIR, suffers from many shortcomings like the need for annotating images manually. Hence, the existing techniques/tools for browsing medical images through a large repository and retrieval techniques consume large amount of time and suffer in terms of accuracy. As mentioned earlier, the reliance on latest medical diagnostic tools such as radiology accelerated an increase in the quantity of medical images preserved in clinics. Digital image retrieval is pivotal in the upcoming field of medical image repositories/databases for the process of decision making in clinical scenario as it has the ability to retrieve images of similar nature and characteristics.

2 Literature review

In this section, various concepts useful for the proposed approach are described and other existing methods for medical image retrieval are reviewed.

2.1 Dictionary Learning

Sparse dictionary learning is an unsupervised learning method whose aim is to find a sparse representation of the input data i.e a vector where most of the elements are zero and only few are non-zero in the form of a linear combination of basic elements. These elements compose a dictionary and are themselves called atoms. Atoms in the dictionary i.e the columns of the dictionary may or may not be orthogonal, and they may be an overcomplete spanning set. Hence, this permits the representation of signal to be of higher dimensions than the dimensions of the signals being observed. These two properties give rise to having redundant atoms in the dictionary that permit more than one representations of the same signal. They also produce an improvement in the flexibility of representation of a signal through sparse coding. One of the key concepts in the process of dictionary learning is that the dictionary representing the database of images has to be deduced from the input data at hand. The need to represent a signal with as few components as possible has been evident in the past few years. This led to the emergence of the idea of sparse dictionary learning. The general practice before this approach was to use predefined dictionaries. But, in many cases a dictionary which is trained to fit only the input data can improve the sparsity to a great extent, which has several applications such as data decomposition, compression and analysis in the fields of audio and video processing, and in the area of image classification and denoising.

2.2 Sparse coding

The aim of sparse coding is to find the best sparse representation(which minimizes the error) of a signal at hand in terms of linear combinations of basis elements i.e the atoms of the dictionary. Given a dictionary D, and an input signal x, sparse coding methods give a sparse vector α as output which is the best approximate to the signal at hand.

2.3 Existing methods

Chu et al. [1] proposed a medical image retrieval techniques which uses the idea of knowledge-based retrieval for magnetic resonance imaging (MRI) images and computed tomography (CT). The segmentation of lesions of brain was performed automatically with the help of a semantic model represented through knowledge. In the Bag of Words (BOW) [2] approach, the image was partitioned into patches. These patches were sampled with keypoint detectors such as SIFT and each keypoint had a description which is a vector of some size(128 in case of SIFT). The keypoints were the most "interesting points" in an image and have the most information about an image. The descriptors obtained through a descriptor detector were used to classify CT images of liver. In [3], Adrien et. al. proposed a texture based analysis and retrieval of CT images of lungs through the use of Riesz wavelets. Guimond et al. [4] used the concept of volume of interest (VOI) which is user-selected for the retrieval of pathological MRI images of the human brain. Another technique based on wavelet optimization technique was proposed for CBIR in medical repository by Quellec et al. [5]. Feature selection and feature extraction algorithm based on Linear discriminate analysis(LDA) for segmentation and classification of 1-D radar signals and 2-D document and texture images with the help of wavelet packet was proposed by Etem and Chellappa[6]. Li. et. al[7] used group sparse representation with dictionary learning for denoising medical images of various categories and fusing them. In [8], Chen et al. used a clustering processing invariant to scale and in-plane rotation for providing invariance to rotation with the use of dictionaries. This techniques provides both Radon-based rotation and invariance to scale while clustering as applied to content based image retrieval. Fei et al. [9] proposed an approach to denoise CT images with the use of sparse representation using a global dictionary. This method not only improved the quality of CT images of the abdomen by using a denoising method based on dictionary learning, but also the training time also greatly reduced. In this approach, classes of images associated to different departments such as pathology, dermatology were considered separately.

2.4 Outcome of literature review

Most of the existing methods for medical image retrieval exhibit the following characteristics-
1. They require training phase before image retrieval for user input query is performed.

2. Image retrieval is restricted to a specific context and not applicable for all types of medical images.

2.5 Problem Statement

To group large databases of medical images quickly and produce relevant images ranked according to their relevance to the user-input query using sparse representation of images and performing dictionary learning for obtaining clusters of similar images. The images on which query is performed need not be restricted to a specific context.

2.6 Research objectives

The aim of the project is to make the searching of similar images in a large image repository quicker and efficient. The dictionary learning approach used in the proposed model not only produces results quickly but also can work well for a medical image database containing images of many different classes.

3 Methodology and Framework

3.1 System Architecture

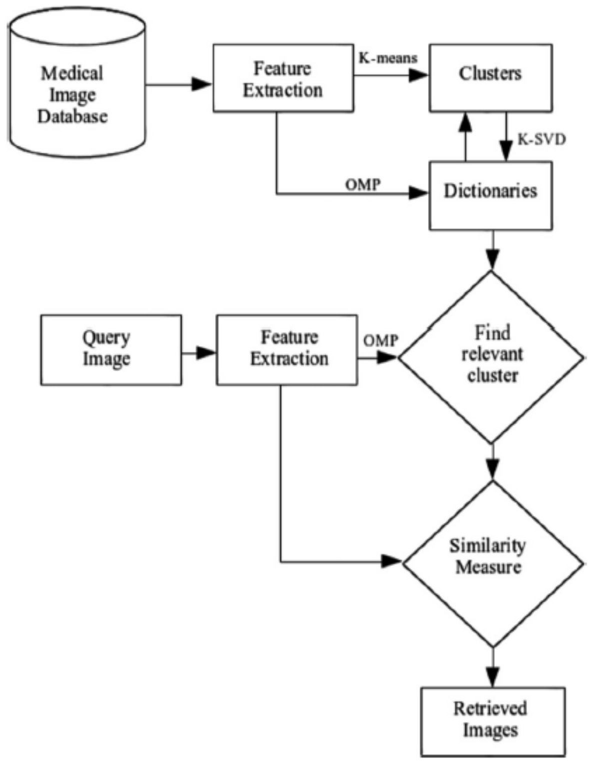


Figure 1: Proposed Method

The proposed model receives a database of medical images and a user query as input. The output is the set of all similar images ranked according to their relevance to the user query. Figure 1 shows the flow of the proposed approach. The IRMA medical image dataset containing 12000 images of 193 categories is used. The implementation language chosen is Python. For feature extraction and basic image I/O, methods of Scikit-image library are used. Clusters are initially formed using K-Means algorithm. The images are re-assigned to each cluster using OMP matching algorithm and the dictionary for each cluster are learnt using K-SVD. This step is repeated till clusters converge.

3.2 Algorithms and Techniques

3.2.1 Feature extraction

To depict the visual content of medical images, two types of feature extraction methods are considered. An image is divided into concentric circular regions each of equal area i.e each concentric circle has same number of pixels in the first method of feature extraction. 17 such concentric circles are considered. Mean and variance is computed for each concentric circle and they become elements of the feature vector. The size of this feature vector is 17X2=34. Figure 2 shows a visual representation of this feature extraction method. The mean and variance equations are given by:

$$y_{j,2k-1} = \frac{1}{P_k} \sum_{l \epsilon J_k} I_l y_{j,2k} = \sum_{l \epsilon J_k} I_l^2 - y_{j,2k-1}^2 \qquad (1)$$

for j=1,2,..M where M is total number of images and k=1,2,...,L where L is the number of total concentric circles and J_k(k=1,2,..L) for the pixels that fall in kth concentric circle. I_l is $l^t h$ pixel value.

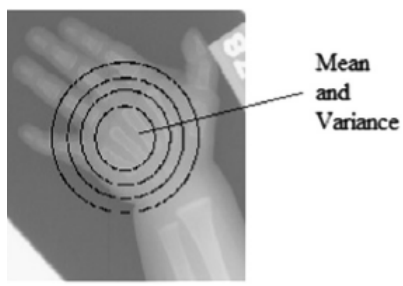


Figure 2: Feature extraction method 1

A second feature extraction method is considered where an image under consideration is partitioned into four blocks(i.e four rectangular regions) which results in four subimages. Another sub-image is considered at the center whose block size is same as the block size of the other four sub-images. This subimage is used to capture the information available at the central region of a medical image. Each sub-image obtained in the previous steps is divided into concentric circular regions of area consisting of same number of pixels. From each concentric circle, variance and mean of pixel intensity values are calculated using Equation 1. This feature extraction method is more appropriate for medical image repositories because the information obtained from the central area of a medical image is more rich and useful. Figure 3 shows a visual representation of this feature extraction method.

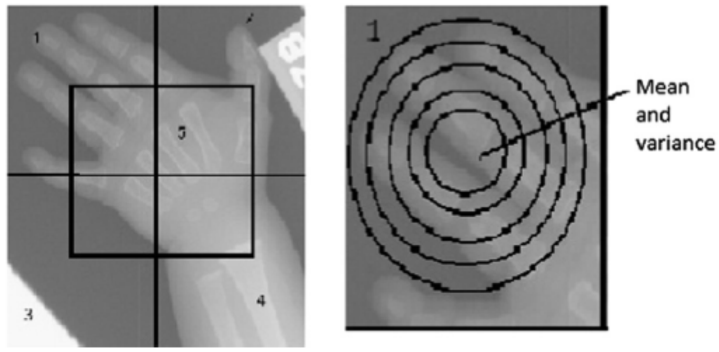


Figure 3: Feature extraction method 2

3.2.2 K-Means Clustering

After feature extraction the images are grouped into initial clusters based on the features extracted in the previous step. This is the first step in the process where naïve groups of similar images are formed and these clusters then form the basis for formation of initial dictionaries.

Given an initial set of k means (where k is the number of clusters), $m_1^{(1)}, .., m_k^{(1)}$, the algorithm progresses by switching between two main steps:

Assignment step: Assign each tuple to the cluster which gives the least difference between the mean of the cluster and the sample(tuple). The difference between the cluster and the mean is obtained through Eucledian distance.

$$S_i(t) = x_p : \|x_p - m_i^{(t)}\|^2 \leq \|x_p - m_j^{(t)}\|^2, \forall j, 1 \leq j \leq k, \tag{2}$$

where each x_p(sample) is assigned to one $S^{(t)}$.

Update step: Calculate the updated means to be the centroids of the samples in the new clusters obtained from last step.

$$m_i^{(t+1)} = \frac{1}{|S_i(t)|} \sum_{x_j \epsilon S_i^{(t)}} x_j \tag{3}$$

The algorithm is said to have converged when the assignments of samples to clusters no longer change and the process is then stopped.

3.2.3 K-SVD algorithm

Given a set of vectors $v_i, i = 1, .., n$, the K-SVD based DL(Dictionary Learning) method finds the dictionary D by solving the following optimization problem,

$$(D, \phi) = \underset{\hat{D},\hat{\phi}}{\arg\min} \|V - \hat{D}\hat{\phi}\|_F^2, \text{subject to} \|\gamma_i\|_0 \leq T_0 \forall i \tag{4}$$

where γ_i represents i^{th} column of ϕ, V is the matrix whose columns are v_i, and T_0 is the sparsity parameter. The columns of ϕ represent the sparse solutions of $v_i, i = 1, .., n$ in terms of dictionary D. Here, $\|A_F\|$ denotes the Frobenius norm which is defined as $\|A_F\|$ $= \sqrt{\sum_{ij} A_{ij}^2}$ and $|v|_0$ stands for the number of non-zero components in v. The K-SVD algorithm alternates between sparse coding and dictionary update steps.

3.2.4 Orthogonal Matching Pursuit

Matching pursuit (MP) is a sparse approximation algorithm which finds the sparse coded approximate signal for a given signal with an over-complete dictionary D. The algorithm for Orthogonal Matching Pursuit is shown below. For this algorithm, the input is a dictionary D and a signal. The output is a sparse representation of the signal.

Algorithm Matching Pursuit
Input: Signal: $f(t)$, dictionary D.
Output: List of coefficients $(a_n)_{n=1}^N$ and indices for corresponding atoms $(\gamma_n)_{n=1}^N$.
Initialization:
  $R_1 \leftarrow f(t)$;
  $n \leftarrow 1$;
Repeat:
  Find $g_{\gamma_n} \in D$ with maximum inner product $|\langle R_n, g_{\gamma_n} \rangle|$;
  $a_n \leftarrow \langle R_n, g_{\gamma_n} \rangle / \|g_{\gamma_n}\|^2$;
  $R_{n+1} \leftarrow R_n - a_n g_{\gamma_n}$;
  $n \leftarrow n + 1$;
Until stop condition (for example: $\|R_n\| < \text{threshold}$)
return

3.3 Detailed design methodology

3.3.1 Initial cluster formation

Features are extracted using both feature extraction methods listed earlier and initial clusters are formed by K-Means algorithm.

3.3.2 Dictionary Update

From the initial clusters: $C_1, C_2, \ldots, C_N$ we obtain the dictionaries $D_1, D_2, \ldots, D_N$ by using the K-SVD approach.

3.3.3 Cluster Assignment

In this step, updating of the clusters is performed from the dictionaries learnt from the initial clusters obtained with K-Means. The idea behind obtaining clusters is done based on the concept that any two images belonging to the same cluster have decomposition in terms of similar dictionary atoms. Let D be the concatenation of the dictionaries obtained after performing K-SVD on clusters obtained with naive Bayes. The proposed approach works by obtaining the most sparse representation of y_j in an appropriate dictionary D_i from:

$$\begin{aligned} \alpha^j &= \arg\min_{\omega} \|y_j - D\omega\|_2^2, \text{subject to} \|\omega\|_0 \leq T_0 \\ \hat{i} &= \arg\min_{i} \|y_j - D\delta_i(\alpha^j)\|_2^2, j = 1, 2, ..M \end{aligned} \tag{5}$$

where d_i is the characteristic function that picks the coefficients from the sparse vector obtained. Then y_j is assigned to C_i which is associated with the i^{th} dictionary. The step 1 of this process finds T_0 number of atoms from D that sparsely describe y_j, whereas the step 2 finds the number of non-zero elements from a particular dictionary i.e the concentration of atoms from a dictionary. Untill no significant change is observed in the clusters, the cluster assignment and dictionary update steps are repeated. Given a query image x_q, a cluster that is the closest to the query image is found by recognizing the appropriate dictionary that admits representation to x_q from the optimization problem in the following equation:

$$\beta = \arg\min_{\omega} \|x_q - D\omega\|_2^2, \text{subject to} \|\omega\|_0 \leq T_0 \tag{6}$$

The cluster C_i is considered to be the most relevant to the query image if it satisfies

$$\|x_q - D\delta_i(\beta)\|_2 < \|x_q - D\delta_j(\beta)\|_2 \forall j \neq \hat{i} \tag{7}$$

In case of a tie, that is when 2 dictionaries give the same minimum error,

$$\|x_q - D\delta_{\hat{i}}(\beta)\|_2 = \|x_q - D\delta_{\hat{l}}(\beta)\|_2 < \|x_q - D\delta_j(\beta)\|_2 \forall j \neq \hat{i}, \hat{l} \tag{8}$$

we search for relevance of X_q in C_i and C_l. After identifying the most relevant cluster to the query image, the relevant images within the cluster are identified with the help of a similarity metric. To determine similarity between two images based on the selected features, an appropriate similarity metric needs to be chosen.

Thus, the detailed steps of the proposed approach are:

Step 1. Extract features from the images of the medical image repository using feature extraction method-II.

Step 2. To obtain initial clusters, apply K-Means clustering over the samples features of samples obtained in the previous step.

Step 3. Obtain a dictionary for each cluster obtained in previous step using K-SVD method.

Step 4. For each dictionary, generate a new cluster by assigning the samples to a cluster that are sparsely represented by it as described in Equation 5

Step 5. Continue doing step 3 and step 4 till there is no significant change in the clusters i.e. the clusters have converged.

Step 6. For the user-input query image x_q, search for relevance in C_i, where D_i provides sparsest representation to x_q and display ranked images to the user based on certain similarity criteria such as Eucledian distance.

4 Implementation and Work Done

4.1 Experimental Framework

The simulations were performed on a Linux system with 8 GB RAM, with CPU speed 2.60Hz and x64-based processor. The language of implementation used was Python. Following Python libraries were used:

- Scikit-image: Image processing methods
- Scikit-learn: K-Means, OMP
- Numpy: Matrix manipulation

4.2 Step by step implementation

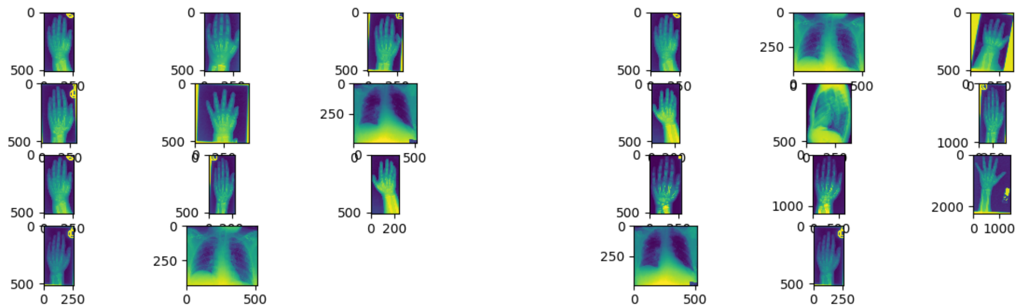
Dataset used: Subset of IRMA dataset of 300 images of 4 categories was used for experimental purpose. The following steps were implemented.

1. Feature extraction was performed using 'circle' method of Scikit-image and further processing it to obtain mean and variance of each concentric circle and obtaining feature vector of each image.
2. K-Means method of Scikit-image was used to form initial clusters.
3. K-SVD implementation of Python was applied on to clusters obtained in step 2 to obtain dictionary for each cluster.
4. OMP was used to obtain the sparse representation of each image and the dictionary giving the least error with Eucledian distance is used for its assignment to a cluster.
5. Repeat steps 3 and 4 till there is no significant change in the clusters.

4.3 Results and discussion

The unsupervised clustering algorithm based on dictionary learning algorithm was applied on the IRMA medical images dataset[10]. The IRMA medical image dataset is a dataset of consisting 12000 images from 193 categories. The Eucledian distance(ED) measurement and Cosine similarity(CS) were used for retrieval. Dictionary Learning(DL) and K-Means methods were used for clustering

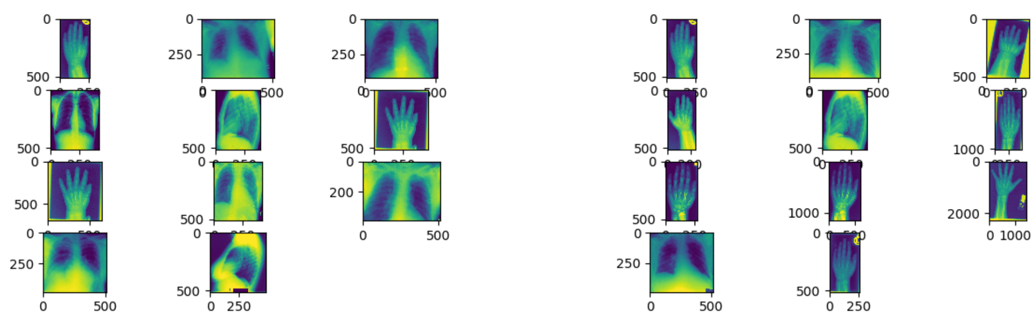
The retrieval performance of Eucledian distance and Cosine Similarity with clusters obtained using K-Means and Dictionary Learning approach for hand image as input are shown in Figure 4 and Figure 5 respectively. The precision obtained for K-means and DL was 0.6 and 0.7 respectively with EC and 0.2 and 0.6 respectively with CS.



(a) With DL, precision=0.7

(b) With K-Means, precision=0.6

Figure 4: Results of K-Means and DL clustering for hand as input image with ED

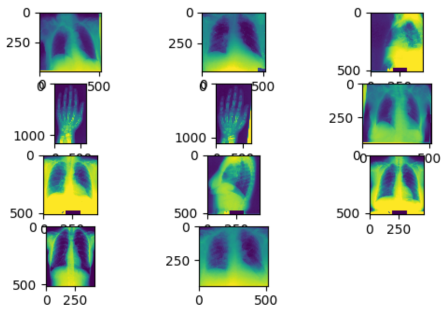


(a) With DL, precision=0.2

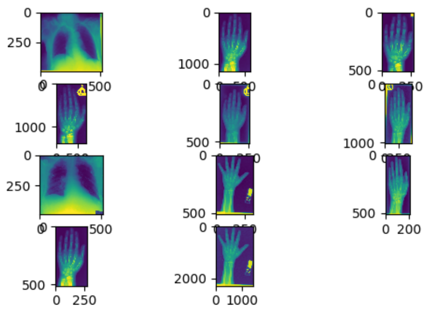
(b) With K-Means, precision=0.6

Figure 5: Results of K-Means and DL clustering for hand as input image with CS

Results obtained for lung image as input are shown in Figure 6 and Figure 7.

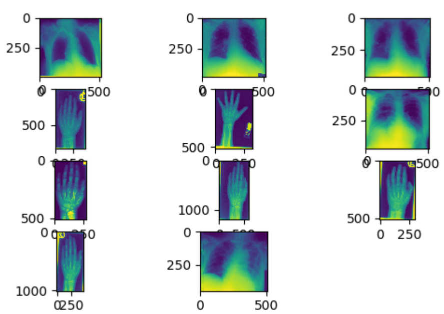


(a) With DL, precision=0.5

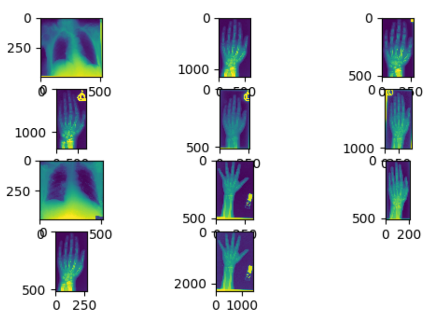


(b) With K-Means, precision=0.1

Figure 6: Results of K-Means and DL clustering for lung as input image with ED



(a) With DL, precision=0.2



(b) With K-Means, precision=0.4

Figure 7: Results of K-Means and DL clustering for lung as input image with CS

	CS	ED
DL	0.282	0.572
K-Means	0.237	0.237

Table 1: Mean average precision for 4 queries

	CS	ED
DL	0.33	0.87
K-Means	0.47	0.61

Table 2: Mean Reciprocal Rank for 4 queries

The mean average precision details and Mean Reciprocal Rank(MRR) are summarized in Table 1 and Table 2 respectively.

4.4 Analysis of results

ED was found to perform better than CS. This is because the during assignment of images to clusters, Eucledian distance was used for error measurement. So, during retrieval, when ED was used to rank images, it performed better than CS. DL gave better results than K-Means because clusters obtained after K-Means are naive clusters and DL used these clusters as foundation and improved them by performing assign and update steps repeatedly.

4.5 Challenges Faced

- Feature collection for images found to be time consuming. Takes around 60s on an average per image
- Clusters were found to be not converging in step 5. A threshold was chosen and if the difference between number of items per cluster was found to be less than the threshold, iterations were stopped.

4.6 Individual Contribution

Balakoti: Feature Extraction

Rajat: OMP

Aparna: OMP and K-SVD

Results and analysis: Aparna, Rajat, Balakoti

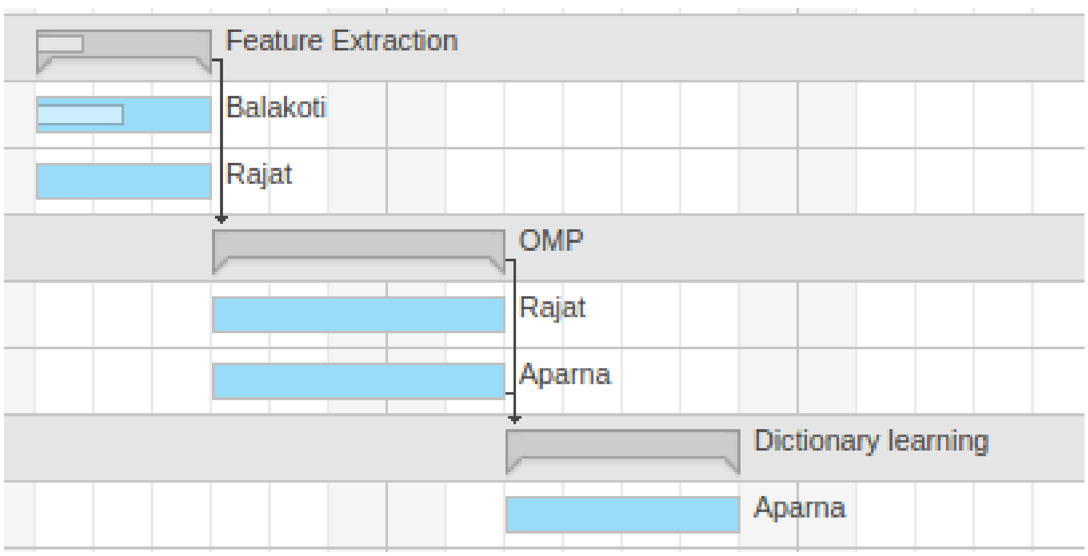


Figure 8: Gantt Chart for individual contribution

4.7 Innovative Work

- Compared results of distance based measure(ED) with similarity based measure(CS). ED was found to perform better than CS.
- Clustering was performed on a subset of dataset with higher number of images per cluster to check if it improved results. I did not improve results.
- Dataset was pre-processed first with both attribute-wise and row-wise normalization which gave better initial clusters(using K-means).

5 Conclusion and Future Work

The dictionary learning and K-means clustering approaches were compared on IRMA medical image database. Two distance measures namely- Eucledian distance and Cosine Similarity were used for ranking images. Dictionary Learning was found to outperform K-Means clustering for both Eucledian Distance and Cosine Similarity as can be seen from Table 1. The future work may involve automatically setting the number of components in the dictionary learning process. Possibility of including relevance feedback could be explored. Other matching pursuit algorithms like CoSaMP could be applied instead of OMP.

References

[1] W.W.Chu, C.C.Hsu, A.F.Cardenas, R.K.Taira, Knowledge based image retrieval with spatial and temporal constructs, IEEE Trans.Knowl. Data Eng.10 (December (6)) (1998)872–888.

[2] W.Yang, Z.Lu,Yu. M,M.Huang, Q.Feng, W.Chen, Content-based retrieval of focal liver lesions using Bag-of-Visual-Words representations of single-and multiphase contrastenhancedCTimages,J.Digit.Imaging25(2011)708–719.

[3] Depeursinge Adrien, Foncubierta Rodriguez Antonio, Van de Ville Dimitri, Muller Henning, Multiscale lung texture signature learning using the Riesz transform, Med. Image Comput. Comput.-Assist. Interv.-MICCAI(2004) 517–524.

[4] A.Guimond, G.Subsol, Automatic MRI database exploration and applications, Int. J.Pattern Recognit.Artif.Intell.11(1997)1345–1365.

[5] G. Quellec, M.Lamard, G.Cazuguel, B.Cochener, C.Roux, Wavelet optimization for content-based image retrieval in medical databases, J. Med. Image Anal. 14(2010)227–241.

[6] K. Etemand, R.Chellapa, Separability based multiscale basis selection and feature extraction for signal and image classification, IEEE Trans. Image Process. 7(October(10))(1998)1453–146

[7] S. Li, H. Yin, L. Fang, Group sparse representation with dictionary learning for medical image denoising and fusion, IEEE Trans. Biomed. Eng. 59 (December (12)) (2012).

[8] Y.C.Chen C.S.Sastry,V.M.Patel, P.J.Phillips, R.Chellappa, In-plane rotation and scale invariant clustering using dictionaries, IEEE Trans.ImageProcess.22 (June (6))(2013)2166–2180.

[9] F.Yu, Y.Chen, L.Luo, CT image denoising based on sparse representation using global dictionary in: Proceedings of IEEE Conference on Complex Medical Engineering, 2013,pp.408–411.

[10] T.M Deserno,"12000 images belonging to 193 categories", https://www.irma-project.org, 2008