CLASSIFICATION OF MEDICAL DATA BASED ON SPARSE REPRESENTATION USING DICTIONARY LEARNING

A THESIS

submitted by

M. SRINIVAS

for the award of the degree

of

DOCTOR OF PHILOSOPHY

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY HYDERABAD

DECEMBER 2014

To my

Family & Friends

Declaration

I declare that this written submission represents my ideas in my own words, and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited, or from whom proper permission has not been taken when needed.

(Signature)

M. SRINIVAS

(– Student Name –)

CS10P002

(Roll No)

THESIS CERTIFICATE

This is to certify that the thesis entitled **Classification of Medical Data Based On Sparse Representation Using Dictionary Learning** submitted by **Mettu Srinivas** to the Indian Institute of Technology, Hyderabad for the award of the degree of Doctor of Philosophy is a bonafide record of research work carried out by him under my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

C. Krishna Mohan

Hyderabad - 502 205 | Dr. C. Krishna Mohan

Date: 06/05/2015 | Dept. of Computer Science and Engg.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. C. Krishna Mohan for providing me with the opportunity to do my research work under his guidance. His emphasis on steady and committed effort has motivated me during the course of the research work. I have immensely benefited from the excellent research environment that he has created and nurtured.

I extremely grateful to our director, Prof. U.B. Desai, for providing excellent computing facilities and such a nice atmosphere for doing my research work. I am extremely thankful to my doctoral committee members Dr. Bheemarjuna Reddy Tamma, Dr. Sri Rama Murty, Dr. C. S. Sastry and my doctoral committee chairman Dr. M. V. Pandurangarao. My special thanks to Dr. C. S. Sastry for his guidance and encouragement.

I convey my heartfelt thanks to Software Systems Research Lab (SSRL) team for their help in completing the project work successfully. I thank Smt. M. Lakshmi and Shri. Praveen Vemuri, for their affection, encouragement and support during my PhD.

My special thanks to Debaditya Roy, N. Pattabhi Ramaiah, A. Ravi Shankar, V. Nageswarao, Ramu Naidu and my classmates for their support, motivation and help at critical moments during my project work.

I would like to take this opportunity to thank my beloved parents Shri. M. Odelu and Smt. M. Laxmi for making me what I am today.

I would like to thank my brother Dr. M. Ravinder, my sisters Krishna Veni, Rama Devi and Srilatha, my brothers-in-law Sadanandam, Bhaskar and Kumara Swamy, my nephews Vamshi Krishna, Goutham and Babu and my niece's Mounica, Srinidhi and Sathvika for their affection and support.

I thank all my batch mates A. Gopi, J. Krishna, K. Ravi Kumar, Satya Narayana, Aravind Kumar, Vandana, Sweetha, Phani Shankar, Chandan, Rafi, Ragavendra, Naresh, Pravanjan, Harsha, Kiran, Suresh, Laxman, Nagababu and Sandeep for making my stay at IIT enjoyable. I am grateful to all the faculty and non-teaching staff with in the department as well as outside the department for all the help I have received during my stay at IIT Hyderabad.

I wish to thank Dr. K.P Supreethi, Anitha, Pramod, Santosh Kumar, Thejaswini, K. Santosh Kumar, Radhika, Jenni, Shiva, Ganesh, Nagendar, Venkat Reddy, Keerthi Deep, Joseph, Durga, Sreenu, Prasanth, Anil, Shiva Sairam, Sandeep R.B, Mukesh, Satya, Rama Raju, Thomas, Goutham Pilla, Krishna Veni, Poorna, Prakash and Subbu for extending help at different times.

Finally, I thank everyone who helped me directly or indirectly during my stay at IIT Hyderabad.

M. Srinivas

ABSTRACT

Keywords: *Content based medical image retrieval; classification; sparse representation; dictionary learning; clustering; modality; multi-level classification; support vector machines; on-line dictionary learning; K-SVD; OMP; ℓ_1-lasso; multi-scale dictionary learning; adaptive dictionary learning.*

Due to the increase in the sources of image acquisition and storage capacity, the search for relevant information in large medical image databases has become more challenging. Classification of medical data into different categories is an important task, and enables efficient cataloging and retrieval with large image collections. The medical image classification systems available today classify medical images based on modality, body part, disease or orientation. Recent work in this direction seek to use the semantics of medical data to achieve better classification. However, representation of semantics is a challenging task and sparse representation has been explored in this thesis for this task.

In this thesis, we explore new methods for grouping of medical data into different classes based on sparse representation and dictionary learning. The sparsity seeking dictionary learning approaches typically exploit the framework of under-determined setting and hence work on some implicit assumptions on the database. The methods proposed here vastly reduce the search time and improve accuracy of retrieved images. In application, however, one often encounters databases which are not so big that the sparsity promoting under-determined framework cannot be efficiently deployed.

An algorithm for classification of medical images based on edge information extracted from various body parts using ℓ_1-lasso sparse representation and on-line dictionary learning (ODL) is proposed. Edge information is extracted from an image by dividing the image into patches and each patch into concentric circular regions to pro-

vide discriminative information useful for classification of medical images. The ability of on-line dictionary learning to achieve sparse representation of an image is exploited to develop dictionaries for each class using edge-based features.

A single classifier may not be suitable for classification of various kinds of medical image datasets. Most of the medical datasets have the problem of data imbalance i.e. unequally distributed training samples among all the classes, which gives rise to poor classification performance with any of the standard single classifier. We aim to address the problem of data imbalance of medical data using multi-level classification approach. A multi-level classifier combines correctly classified examples in the first level with the training data and supplies them as input to the next level classifier. So, if there is any imbalance in the data, it can be alleviated by this approach. For the first stage of classification, on-line dictionary learning (ODL) is used. Support vector machine (SVM) is used for the second level of classification and together with on-line dictionary learning forms the multi-level classification approach.

Another problem in medical imaging is the classification of medical images captured by acquisition source (i.e modalities). Capturing images using different modalities suffers from significant contrast variation among the images of the same organ or body part. Due to this large variation, existing image classification and retrieval algorithms do not perform well for different modality images. We propose to address this issue by using multi-scale wavelet representation and dictionary learning. Wavelet features extracted from an image provide discriminative information useful for classification of medical images. Multi-scale wavelets are employed to compensate for the varying scale of intensity in the images captured by the aforementioned sources. Cardiovascular diseases (CVD) are a leading cause of unnecessary hospital admissions. Hence, automated detection of abnormal heartbeats captured by electronic cardiogram (ECG) signals is vital. We employ an approach to classify abnormal heartbeat patterns from standard heartbeat patterns using adaptive dictionary learning on a standard ECG database.

We propose a method for clustering of medical image datasets using sparse representation and dictionary learning. The basic idea is to group similar images into clusters that are sparsely represented by the dictionaries and simultaneously learn dictionaries from the clusters using K-SVD. The mean and variance over concentric circular regions in the image are calculated and used as features for providing a rotation invariant image retrieval scheme.

In summary, this thesis opens up the area of sparse representation and dictionary learning to a lot of medical applications particularly in classification and retrieval. The main idea of this work is to explore the applicability of sparsity and dictionaries on various medical datasets like IRMA (X-ray), ICBM (MRI, DTI, MRA, FMRA), MIT-BIH (ECG) and UCI (PIMA, SPECTF, WBC, Heart SATALOGS). We have shown that sparse representation with any of the dictionary learning algorithms like K-SVD and on-line dictionary learning (ODL) is quite suitable for a myriad of classification, clustering and retrieval tasks on different medical datasets.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ED - Euclidean Distance

MRI - Magnetic Resonance Imaging

CT - Computed Tomography

PET - Positron Emission Tomography

MRI Magnetic Resonance Imaging

ODL - On-line Dictionary Learning

MOD - Method of Optimal Directions

WSQ - Wavelet Scalar Quantization

PCA - Principal Component Analysis

SVD - Singular Value Decomposition

LDA - Linear Discriminate Analysis

k-NN k- Nearest Neighbour

ICA - Independent Component Analysis

TBIR - Text Based Image Retrieval

CBIR - Content Based Image Retrieval

BC - Bayes Classifier

NN - Neural Network

IRMA - Information Retrieval in Medical Applications

MLP - Multi-layer perceptron

SOM - Self-organizing Map

DTI - Diffusion Tensor Imaging

MRA - Magnetic Resonance Angiography

FMRI - Functional Magnetic Resonance Imaging

ICBM - International Consortium For Brain Mapping

OMP - Orthogonal Matching Pursuit

ECG - Electronic Cardiogram

CVD - Cardiovascular Diseases

CHAPTER 1

INTRODUCTION TO CONTENT BASED IMAGE CLASSIFICATION AND RETRIEVAL

In the last few years, thousands of millions of images have become available on the Internet. The increase of these image collections is compelling people in various professions, for example, medicine, architecture, geography, design, computer aided design, advertising and publishing to use them in various applications. Meanwhile, the study of image classification and retrieval, which is concerned with efficiently accessing similar type of images from large image collections, has become a more interesting and challenging task. Nevertheless, one cannot utilize the information in these image collections unless they are organized for efficient search and retrieval of data. Image classification and retrieval is all about techniques for storing and retrieving images both efficiently and effectively.

Previously, searching and retrieving similar images an image database was based on human annotation, i.e. each image in a database is given some keywords to denote the semantic meaning of the image. Thus, classification and retrieving images was based on the keywords of images. This type of image retrieval is called as text based image retrieval (TBIR) [1]. Now, many search engines that claim to do image retrieval perform text based image retrieval like Google, QBIC and AltaVista. These search engines search the text around the image, such as captions, file names, and paragraphs located close to the image to search for relevant items to the query. This text based image retrieval method has many limitations, namely, as the size of image collection gets increasingly large manually annotating each image becomes very difficult. Annotating an image based on human perception is very subjective. Different people may assign different annotations to images with similar visual contents. The problem of searching for similar images in a large image repository based on content is called Content Based Image Retrieval (CBIR) [2]. The term content in CBIR refers to colors [3], shapes [3–5], textures [3, 4], or any other information that can be possibly obtained from the image itself. Indexing remarkably affects the speed of data access besides supporting the accuracy for retrieval process and thus, is a significant factor in cataloging image database systems. Content based image indexing tends to facilitate automatic identification and abstraction of the visual content of an image. CBIR has the potential to greatly enhance the functionality of Picture Archiving and Communication Systems (PACS).

In the early 1990s Content Based Image Retrieval (CBIR) was proposed to overcome the limitations of text based image retrieval. There are many differences between content based image retrieval systems and classic information retrieval systems. The major differences are that in CBIR systems images are indexed using features extracted from the content itself and the objective of CBIR systems is to retrieve similar images to the query rather than exact matches. The similarity in most CBIR systems is quantified and the database entries are ranked based on their similarity to the query image. Similar images are retrieved as the result of a query. Different users may be interested in different parts of the same image. So, similarity based retrieval is more flexible than exact matching, and gives better performance in case of queries such as finding the images similar to the given image. The capability of present CBIR systems has been limited by their use of only primitive features like, color, shape, texture, spatial relationships among objects and these features can be used in most CBIR applications.

In Section 1.1, we briefly describe various tasks involved in content based image classification and retrieval. In Section 1.2, we discuss certain issues related to medical image classification and retrieval that are addressed in this thesis. Section 1.3 outlines the organization of the thesis.

1.1 TASKS INVOLVED IN MEDICAL IMAGE CLASSIFICATION AND RETRIEVAL

The objective of content based image retrieval is to develop techniques to automatically extract and retrieve relevant similar images from the large database. In conventional content based image retrieval systems, the query image is given to the CBIR system where the CBIR system will retrieve images from raw (unstructured) image database related to query image. Content based image retrieval involves three major tasks as shown in Fig. 1.1.

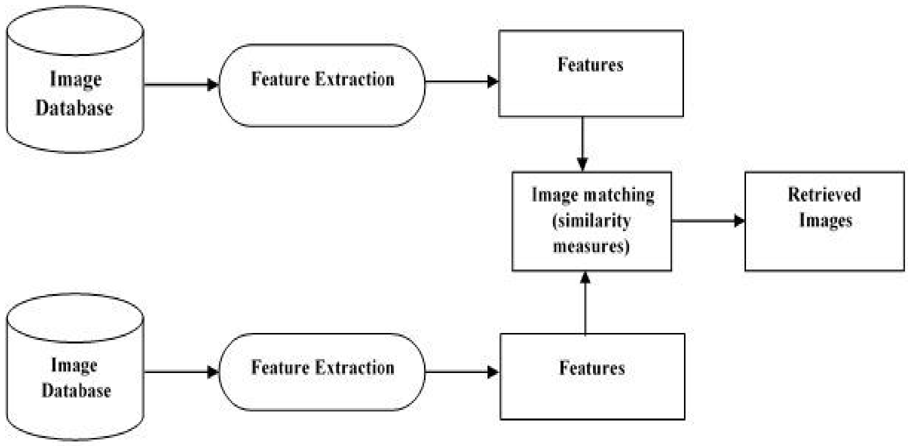

Fig. 1.1: Process diagram for CBIR.

The major functions of a CBIR are as follows:

- Analyze the contents of the source information and represent the contents of the analyzed sources in a way that will be suitable for matching user queries. This step is normally time consuming since it has to process all the source information (images) in the database.
- Analyze user queries and represent them in a form that will be suitable for matching with the source database, which is similar to the source images in the database.

- Define an approach to match the search queries with information in the stored database. Retrieve the images relevant to the query image.

1.1.1 Feature extraction

Feature extraction technique is the process of describing the image by considering parameters known as features (color, edge, texture etc.) from a given image. A feature is defined as a descriptive parameter that is extracted from an image [6]. The effectiveness of medical image classification and retrieval mainly depends on the effectiveness of features used for the representation of the content. An important issue is the choice of suitable features for a given task. Effective image retrieval can be achieved by collaboratively using color, edge density, boolean edge density, texture and histogram bins. These features are discussed in this section.

1.1.1.1 *Color*

Color has proven to be the most importent feature and almost all methods used color information. Although most of the images are in the RGB (Red, Green, Blue) color space, this space is only rarely used for indexing and querying as it does not relate well to the human color perception. It only works well for images taken under exactly the identical conditions each time. Other spaces such as HSV (Hue, Saturation, and Value) or the CIE Lab and Luv spaces are much better with respect to human perception and are used more commonly. This means that differences in the color space are close to the differences between colors that humans perceive. There are different types of color spaces available which are appropriate for different purposes. Some of the color spaces that we often come across are RGB, HSV, CIE Lab and Luv [3]. Color feature can be comprised of histogram bins or average, standard deviation or variance in an opted color space.

1.1.1.2 *Texture*

Texture [7] is another important property of images. Texture features [8] of images refer to the visual patterns that have properties of homogeneity that do not result from the presence of only a single color or intensity. Image texture content provides information of image properties such as smoothness, coarseness, and regularity which is useful in a CBMIR system. Basically, texture representation methods can be classified into two categories: structural and statistical. Structural methods including morphological operator and adjacency graph, describe texture by identifying structural primitives and their placement rules. They tend to be most effective when applied to textures that are very regular. Statistical methods, including Fourier power spectra, co-occurrence matrices, shift-invariant principal component analysis (SPCA), tamura feature, Markov random field, fractal model and multi-resolution filtering techniques such as Gabor [9] and wavelet transform, characterize texture by the statistical distribution of the image intensity.

1.1.1.3 *Shape Retrieval*

Shape features [8] of objects or regions have been used in many content-based image retrieval systems. Compared with color and texture, shape features are normally described after images have been segmented into regions. Since, accurate and robust image segmentation is onerous to achieve. The use of shape features for image classification and retrieval has been restricted to some applications where objects or regions are readily available. The methods for shape description can be classified into boundary or region- based methods. A good shape representation feature for an object should be invariant to translation, rotation and scaling. More information is given in [8].

1.1.1.4 *Semantics*

Current CBIR systems retrieve similar images from a collection on the basis of the low level features of images, such as shape, color and texture. Nevertheless, some systems attempt to finding similar images that are semantically relative to a given query. Semantically similar is meant in the sense of human visual similarity perception (or called high level in CBIR).

1.1.1.5 *Edge Information*

Another choice for characterizing an image is its edge information. The advantage of this feature is that it is sufficiently invariant to illumination changes. Its main disadvantage is computational cost, noise sensitivity, and when not post-processed, high dimensionality.

1.1.2 Indexing for retrieval and browsing

Effective indexing and fast searching of images on basis of visual features pose a significant issue in content based image retrieval. Commonly, a tree structure is utilized to store image information since it has high dimensional metric space. R-tree [10], R*-tree [11], VP-tree structure [12] and Hybrid Tree [13] are some of the widely used tree structures. A majority of these multi-dimensional indexing methods perform significantly well for dimensions up to 20. A variant of R-tree employed in the indexing of spatial information is known as R*-tree. Both point and spatial data are supported at the same instant by an R*-tree but they are more complex compared to R-trees. Norbert Beckmann, Hans-Peter Kriegel, Ralf Schneider and Bernhard Seeger put forth the concept of R*-tree in 1990. Though R*-tree displays significant improvements over the R-tree variants its reinsertion method poses a considerable overhead. Database systems organizing both multidimensional points and spatial data can benefit from the R*-trees. Reduction of the area, margin and overlap of the directory rectangles are the basis for an R*-tree. The R*-tree utilizes an algorithm analogous to that of the R-trees for query and delete operations. The primary difference lies in the insert algorithm. To be precise the mode of selection of which branch to insert the new node into and the methodology for splitting a full node in an R*-tree differs from that of the R-tree [11]. The Indexing tree structure can perform efficiently in dictionary operations. But, it can not be used for finding similarity among the images in the database. So, indexing tree structure was limited to structure the image database so that efficient retrieval is possible.

1.2 ISSUES ADDRESSED IN THIS THESIS

The previous section briefly described the various issues involved in content based image retrieval. The objective of this research work is to develop methods for classification of medical data. The motivation for this objective stems from the need to organize large collection of medical images, for efficient classification and retrieval. The problem of image classification is addressed in the context of medical images. Classification of medical images into various genres or categories is an important task and most of the medical image classification systems available today classify medical images based on modality, body part, disease or orientation. In this thesis, we address two issues which are important for efficient classification and retrieval of medical data, namely, representation and classification. Representations can entangle and hide more or less the different explanatory factors of variations in the data. The objective of image classification is to categorize a given image into one of the predefined categories. Medical image classification is an important task in content based medical image retrieval. With the help of classification, the accuracy and retrieval speed of relevant images in content based image retrieval vastly improves. Classification of medical images based on various body parts using on-line dictionary learning (ODL) and ℓ_1-lasso sparse representation on edge-based features is performed since different body parts are distinctly characterized by edge information.

Another important issue is the classification of imbalanced data. Most of the medical datasets pose data imbalance problems which give poor classification performance with single classifiers. In this thesis, we propose a method to address the problem of data imbalance in medical images using multi-level classification approach. A multi-level classifier combines correctly classified examples in the first level with the training data and supplies them as input to the next level classifier.

Another issue in medical imaging is classification of medical images captured by different sensors. Capturing images using different modalities suffers from significant contrast variation between the images of the same organ or body part. Due to this large variation, existing image classification and retrieval algorithms do not perform well for different modality images. In this chapter, we propose a new classification technique, namely, sparse representation based multi-scale dictionary learning to classify the different type of modality images. Wavelet features extracted from an image provide discrimination useful for classification of medical images obtained from different sensors. Another important application area which is explored is automated detection of abnormal heartbeats captured by electronic cardiogram (ECG) signals. We employ an approach to classify abnormal heartbeat patterns from standard heartbeat patterns using adaptive dictionary learning on a standard ECG database.

Content based medical image retrieval (CBMIR) is the process of extracting relevant images to a query image, based on content rather than annotation. The key issues with CBMIR are the choice of features for representation of images, similarity/distance metric and a generic algorithm for retrieval of rotation invariant based similar images. The proposed CBMIR approach concentrates on retrieving rotation invariant resultant images and improving the accuracy of retrieved images with the help of clustering. A given image is partitioned into concentric circular regions of equal area and the mean and variance of each such area are considered as features for rotation invariant representation of images. These features are then used for the proposed dictionary learning based clustering and sparsest representation based classification algorithms.

1.3 ORGANIZATION OF THE THESIS

An overview of the existing approaches to image classification and retrieval is presented in Chapter 2. Some research issues are identified in both these tasks which are addressed in this thesis. In Chapter 3, content based medial image classification is performed using on-line dictionary learning (ODL) and ℓ_1-lasso sparse representation on edge-based features. In Chapter 4, the problem of imbalanced data problem in medical image classification is addressed by using multi-level classification approach. A new method for classification of medical images based on modality is proposed in Chapter 5, using the framework of multi-scale dictionary learning algorithm. In Chapter 6, the problem of automated detection of abnormal heartbeats captured by electronic cardiogram (ECG) signals is addressed. An adaptive dictionary learning based classification technique is used to classify the normal and abnormal heartbeats from ECG medical database. In Chapter 7, dictionary learning based clustering method is proposed for content based medical image retrieval. An approach to group similar images into clusters that are sparsely represented by the dictionaries and simultaneously learn dictionaries from the clusters using K-SVD method is proposed. A query image is matched with the existing dictionaries to identify the dictionary with the sparest representation using OMP algorithm. Then, images in the cluster associated with this dictionary are compared using a similarity measure to retrieve images similar to the query image. Considering mean and variance over concentric circular regions as features facilitates rotation invariance based image retrieval. Chapter 8 summarizes the research work carried out as part of this thesis, highlights the contributions of the work and discusses directions for future work.

CHAPTER 2

OVERVIEW OF APPROACHES FOR CONTENT BASED MEDICAL IMAGE CLASSIFICATION

This chapter reviews some of the existing approaches to content based medical image classification and retrieval. The problem of content based image retrieval is briefly described in Section 2.1. The two important components of algorithms for image classification, namely, features for representation of images, and similarity/distance metric, are discussed in terms of the commonly made choices for these components. The existing algorithms for content based image retrieval are then reviewed. In Section 2.3, the existing approaches to image classification are reviewed, with particular focus on the classification of medical images. Some research issues arising out of the review of existing methods are identified, which are addressed in this thesis.

2.1 EXISTING METHODS FOR CONTENT BASED MEDICAL IMAGE CLASSIFICATION AND RETRIEVAL

There are hundreds of millions of images available on the Internet. Nevertheless, one cannot utilize the information in these image collections unless they are organized for efficient search and retrieval of data. Therefore, the need of an efficient method to retrieve digital images is recognized by the public. There are two approaches to image classification, namely, text based approach and content based approach. The former solution is a more traditional approach which indexes images by using keywords. The keyword indexing of digital images is useful but requires a considerable level of effort and often limited for describing image content. The alternate approach, the content based image retrieval indexes images by using the low level features of the digital images and the searching depends on features being automatically extracted from the image. Content based image retrieval [2], is the term used to describe the process of retrieving images from a database on the basis of the internal features of images. In CBIR, digital images are indexed [11] by summarizing their visual contents through automatically extracted features such as texture, color, and shape. There exist different ways to express the query. The query can be defined by submitting one or more example images, providing a rough sketch of the desired item or by providing textual description of the object. CBIR retrieves stored digital images from a collection by comparing features extracted from the images. The most common features used are mathematical measures of color, texture or shape [2]. The CBIR system identities those stored images whose feature values match those of the query most closely and displays these found images to the user. In the following section, some of the frequently used types of features used for image retrieval will be described. The first step in content based medical image retrieval is to select an appropriate feature set for the image database. The selection of the feature set should be done in such a way that it should approximate images which are semantically similar to be as close to each other as possible in the feature space. The next step is to prepare a query image for retrieval i.e. extract features from the query image. Finally, an appropriate similarity measure is employed for retrieving the most similar images from the database. A block diagram of traditional content based image retrieval is shown in Fig. 2.1.

A common approach to feature extraction is to segment the images into regions [14] based on a certain similarity criterion. Regions from the segmentation result can then be used in region based queries for CBIR. This enables the user to include only the relevant regions when formulating a query. Chu et al. [15] described a knowledge based image retrieval of computed tomography (CT) and magnetic resonance imaging (MRI) images. In this approach, the brain lesions were automatically segmented and represented to form a knowledge based semantic model. Cai et al. [16] proposed a CBIR system for functional dynamic positron emission tomography (PET) images of

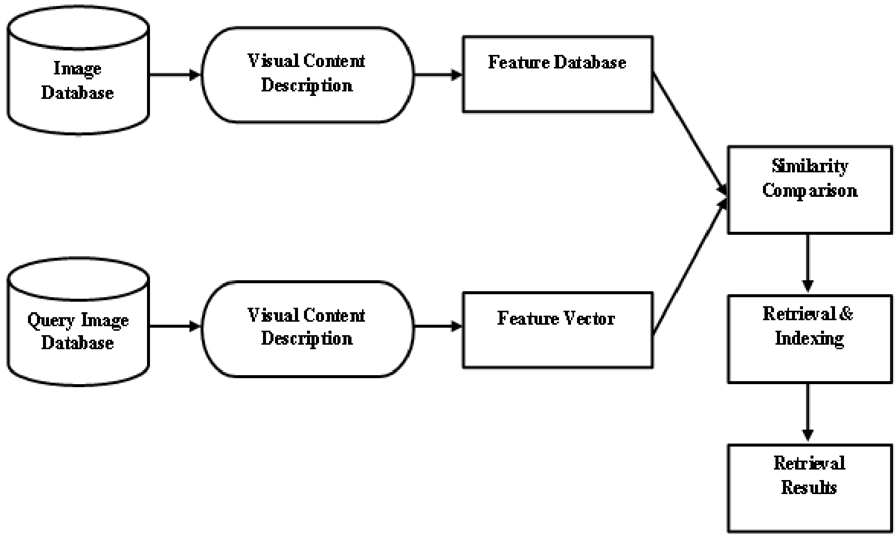

Fig. 2.1: Traditional content based image classification and retrieval system.

the human brain, where segmented clusters of tissue time activity curves from the temporal domain were used in the computation of similarity measure for retrieval. In [17], the delineations of the regions of interest were manually performed on the key frame from the stack of high resolution CT images and were used as features to represent the entire image. Some CBIR systems use segmentation to represent the regions, such as the ones used for retrieval of tumor shape and the shape of regions in spine X-ray images.

Guimond et al. [18] introduced user-selected volume of interest (VOI) for the retrieval of pathological brain MRI images. In [19], group sparse representation with dictionary learning for medical image denoising and fusion was used. Wavelet optimization techniques for content based image retrieval in medical database were described in G. Quellec et al [20]. Linear discriminate analysis (LDA) based selection and feature extraction algorithm for classification and segmentation of one dimensional radar signals and two-dimensional texture and document images using wavelet packet was proposed by Etemand and Chellappa [21]. Recently, similar algorithms for simultaneous sparse signal representation and discrimination were proposed [22]- [23]. In [24], Yi. Chen et al. proposed in-plane rotation and scale invariant clustering using dictionaries. This approach provides Radon-based rotation and scale invariant clustering as applied to content based image retrieval on Smithsonian isolated leaf, Kimia shape and Brodatz texture datasets. Fei et al. [25] described a CT image denoising based on sparse representation using global dictionary. This approach improves low dose CT abdomen image quality through a dictionary learning based denoising method and accelerates the training time at the same time. Some of the existing medical CBIR systems as follows:

ASSERT : This system was developed in the school of electrical and computer engineering at Purdue University [17]. It is designed specifically for high resolution computed tomography images of the lung, since it uses some perceptual features specific to those images. It also includes gray-level features, such as the gray-level mean and standard deviation, texture features such as contrast, entropy and homogeneity and shape features such as the area. The feature vectors are indexed using the multi-hash method described in [26]. In Fig. 2.2, shows the some of the query related images with the ASSERT tool.

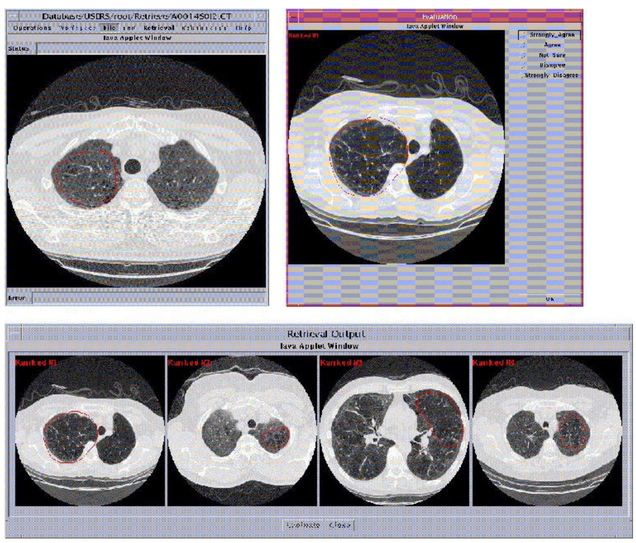

Fig. 2.2: Some of the retrieved images with ASSERT tool.

IRMA : The IRMA system [27] was developed on the Aachen University of Technology. It is focused on the querying of medical images using manually defined prototypes in a first stage and features are extracted from frequency, texture and structure analysis in regions segmented in a multi-scale blob-representation (blob tree). Those features are then indexed using a cluster-based approach.

The strategies adopted in the field of medicine often are (a) to use of more complex gray-level features (e.g. increase the number of gray level bins in the histogram), (b) to limit searches by creating prototypes for several well defined categories and (c) to use features that are specific of those images. MedGIFT uses the first strategy, IRMA uses the second and ASSERT uses the third. In Fig. 2.3, shows the some of the query related images with the IRMA tool.

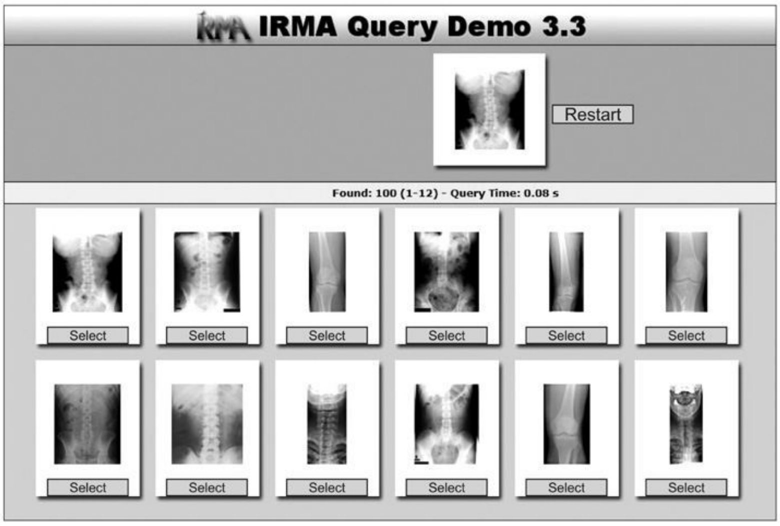

Fig. 2.3: Some of the retrieved images with IRMA tool.

2.2 COMPONENTS OF CONTENT BASED IMAGE CLASSIFICATION AND RETRIEVAL METHODS

An important component of content based image retrieval algorithms is the set of features extracted from a image or from a region of the image. Another component is the similarity measure that is used to detect the presence of a image retrieval. We present below the different choices that can be made for each component, along with their advantages and disadvantages. A content based image retrieval algorithm can then be designed by suitably choosing each component.

2.2.1 Features used for representation of an image

Feature extraction is the process of describing the image by considering parameters known as features (color, edge, texture etc) from a given image. A feature is defined as a descriptive parameter that is extracted from an image [6]. The effectiveness of image retrieval depends on the effectiveness of features/attributes used for the representation of the content. An important issue is the choice of suitable features for a given task. Effective image retrieval can be achieved by collaboratively using color, edge density, boolean edge density, texture and histogram bins. These features are discussed in this section.

2.2.1.1 Extraction of gray-level features

Color has been the most effective and most widely used feature in CBIR [28, 29]. In medical CBIR, the color of each pixel is restricted to a gray levels intensity, which is already available, so it is quite straightforward to extract meaningful gray-level features. The objective is to transform the local gray-level information of each pixel into a global gray-level distribution of the full image, where visually similar images have similar representations.

Gray-level histogram

The most popular method of extracting gray level features of an image is to construct its histogram [29, 30]. A histogram is a statistical description that captures the gray levels distribution of an image. To construct an histogram, we discretized the intensity of the gray levels into a set of bins, and count the number of pixels whose intensity is in that bins range [31, 32]. In CBIR, the histogram is discretized into 256 bins, where the first bin has the number of black pixels (absence of color) and the last bin has the number of white pixels. Mainly, histograms suffer from two problems that limit their reliability. Perceptually, similar colors problem [30] is due to the very small difference between intensity values in neighboring bins. Sometimes, almost identical intensities are not assigned to the same bin, but to a neighbor. This means that the difference between the histograms of perceptually similar images (such as two images taken with different light conditions) can be quite big . An even bigger problem is the absence of any spatial information [33]. We can shuffle all the pixels in an image but the histogram remains untouched and therefore, the images are considered equal.

Partition-based histograms

Partition-based histograms incorporate spatial information by splitting the picture into $k \times k$ partitions, each one with its own histogram to store the local area gray-level information [34]. The spatial information emerges because only the corresponding pairs of local histograms are compared.

Color coherence vectors

Due to the absence of any spatial information in a histogram, an image with a large area of a given gray-level can be considered similar to another image that has many small areas of the same gray level scattered. To solve this problem, Pass et al. [33] proposed the Color Coherence Vectors (CCV) method. We start by identifying all the similar gray-level regions (connected components) in the image and count the number of pixels they have. If the number of pixels in a connected component is bigger than a given threshold, then they are classified as coherent. Otherwise, they are classified as incoherent. Not only a CCV has all the information present in histograms (i.e. we can convert a CCV into a histogram simply by adding the coherent and incoherent pixels for each pair), but it also measures if a gray-level is in a large area or scattered. A big number of coherent pixels are able to distinguish images with big similar gray levels areas from images with small scattered areas, even if both histograms are equal. Unfortunately, we still miss important spatial information. We do not know how many regions are present, how big they are or their location. Another potential problem is the definition of the threshold used to classify the pixels coherence. Too low and even pixels in small regions will be coherent, too high and there will only be coherent pixels in large regions.

2.2.1.2 Extraction of texture features

Texture [35] is another important property of images. Texture features [36] of images refer to the visual patterns that have properties of homogeneity that do not result from the presence of only a single color or intensity. Image texture content provides information of image properties such as smoothness, coarseness, and regularity which is useful in a CBIR system. Basically, texture representation methods can be classified into two categories: structural and statistical. Structural methods including morphological operator and adjacency graph, describe texture by identifying structural primitives and their placement rules. They tend to be most effective when applied to textures that are very regular. Statistical methods, including Fourier power spectra, co-occurrence matrices, shift-invariant principal component analysis (SPCA), tamura feature, world decomposition, markov random field, fractal model and multi-resolution filtering techniques such as Gabor [37] and wavelet transform, characterize texture by the statistical distribution of the image intensity. Three classical approaches have been developed to describe textures, namely, structural, spectral and statistical. The structural approach assumes that the elements of a texture (textels) are placed under some rules. The spectral approach converts the image to the frequency domain to obtain features from its power spectrum. The statistical approach uses the statistical distribution of the pixels gray-level intensity to identify features. More recently, other methods were proposed inspired by human visual system (HVS) using multichannel filtering at different spatial frequencies and orientations.

2.2.1.3 Extraction of shape features

Shape features [36] of objects or regions have been used in many content-based image classification systems. Compared with color and texture features, shape features are usually described after images have been segmented into regions or objects. Since robust and accurate image segmentation is difficult to achieve, the use of shape features for image retrieval has been limited to special applications where objects or regions are readily available. The methods for shape description can be classified into boundary or region-based and contour-based methods. A good shape representation feature for an object should be invariant to translation, rotation and scaling.

Region based features

Methods that extract region based features take into account all the pixels within the shape. Each shape is mapped onto a fixed sized grid or circle to achieve scale, rotation and translation invariance. This normalized shape is viewed as a probability density of a two-dimensional variable, from which orthogonal moments that describe some global properties of the shape can be computed [38, 39]. However, they are unable to capture its local properties, thus failing to achieve partial occlusion invariance.

Contour based features

Methods that extract contour based features are more popular, since they extract both global and local features using only the shape boundary coordinates $(x(t), (y(t))$, $t = 0$, 1, ..., L-1 where, L is a fixed number of samples (data points). All shapes are sampled into these data points so that (a) each shape signature (i.e. the representation containing the features) has the same size, to facilitate the comparison between shapes and (b) to smooth the shape, reducing unwanted details, and increasing the computational efficiency [38]. Some of counter based features are complex coordinates, centroid distance, curvature and cumulative angular function.

2.2.2 Measure of similarity

Similarity measurement [3] is one of the key point in content based image classification and retrieval. An important step in most clustering is to select a distance measure, which will determine how the similarity of two elements is calculated. In CBIR, images are represented as features in the database. Once the features are extracted from the indexed images, the retrieval becomes the measurement of similarity between the features. Commonly used similarity measures are :

- The Euclidean distance (also called distance as the crow flies or 2-norm distance). A review of cluster analysis in health psychology research found that the most common distance measure in published studies in that research area is the Euclidean distance or the squared Euclidean distance.
- The Manhattan distance (one-norm)
- The maximum norm (infinity norm)
- The Mahalanobis distance corrects data for different scales and correlations in the variables
- The angle between two vectors can be used as a distance measure when clustering high dimensional data.
- The Hamming distance measures the minimum number of substitutions required to change one member into another.

Euclidean distance [32] is the most common metric for measuring the distance between two vectors, and is given by the square root of the sum of the squares of the differences between vector components.

2.3 EXISTING METHODS FOR MEDICAL IMAGE CLASSIFICATION

Efficiently searching and retrieving of data in the large image collections poses significant technical challenges as the characteristics of the medical images differ from other general purpose images. Some methods have been explored in recent years to automatically classify medical image collections into multiple semantic categories for efficient retrieval [40]. For example, in [41], the automatic categorization of 6231 radiological images into 81 classes is achieved by utilizing a combination of low level global texture features with low resolution scaled images and a K-nearest neighbor (KNN) classifier. Although, these approaches demonstrate promising results for medical image classification and retrieval, classification and searching similar images in a large database is still a challenge.

An X-ray image categorization and retrieval method using patch-based visual word representations is proposed in [42]. The feature extraction process is based on local patch representation of the image content and a bag-of-features approach for defining image categories, with a kernel based support vector machine (SVM) classifier. The method is especially effective in discriminating orientation and body regions in X-ray images, and in medical visual retrieval. In [43], a descriptor is proposed which combines local features with global shape features. The descriptor combines edge of whole image with edge density of sub-images and it is known as the edge density histogram descriptor (EDHD). The image retrieval and classification is then done based on euclidean distance and with the help of support vector machines.

A learning based classification framework based on local binary pattern (LBP) feature is proposed in [44]. Local binary pattern is extracted from each image in database with the help of an LBP operator which labels image pixels by thresholding neighborhood of each pixel with the center value and considers the result as a binary number, which is then classified using a maximum margin SVM. Moreover, a merging technique is applied on the overlapped classes. These overlapped classes are detected in merging scheme with the help of measures such as correctness rate of each class, similarity of imaging body organ and misclassification ratio. In [45], a least square support vector machines is used for breast cancer classification. Least square SVM (LSSVM) simplifies the required computation by solving a linear equation set. This equation set embodies all available information about the learning process. The most important difference between SVM and LSSVM is that LSSVM uses a set of linear equations for training, while SVM uses a quadratic optimization problem which greatly reduces the computational cost. Wavelet optimization techniques for content based image retrieval in medical database are described in [20].

In, [46] shape and texture features are extracted from breast MRI images and genetic algorithm is applied to select the best feature to be used for classification process. To improve classification performance three different classifiers, namely, multi-layer perceptron (MLP), generalized regression neural network (GRNN) and support vector machine (SVM) are combined to from a multi-classifier system. Bartosz et al. [47] introduced under sampling balanced ensemble method to solve the imbalance problem. The construction of multiple independent classifiers is typically a non-trivial problem. In, [48] a cost-sensitive ensemble classification algorithm is proposed. The data imbalance problem is addressed by employing cost-sensitive decision trees as base classifiers which are trained on random feature subspaces to ensure diversity, and an evolutionary algorithm for simultaneous classifier selection and fusion. Marco Vannucci et al. [49] described a binary classification method named Labeled SOM Classification Unbalanced Sets (LASCUS) that can be applied to uneven datasets and sensitive problems such as malfunction detection. LASCUS method is based on the use of a self-organizing map (SOM) and fuzzy inference system (FIS). The SOM creates a set of clusters to be associated either to frequent or unfrequented situations while the FIS determines such association on the basis of data distribution.

Modality classification and its use in text based image retrieval in medical databases is proposed in [50]. Visual descriptors and text features are used for classifying the medical images. Medical image classification is then done with the help of support vector machines classifier. In [51], explore different type of medical image modality and retrieval strategies. Bags of visual words and fisher vectors representations are integrated to perform medical modality classification. Wavelet optimization techniques for content based image retrieval in medical database are described in [20].

2.4 ISSUES ADDRESSED IN MEDICAL IMAGE CLASSIFICATION

This thesis is mainly focused on the issues related to the efficient classification and retrieval of medical images. Image classification and retrieval which is concerned with efficiently accessing similar type of images from large image collections, has become more interesting and more challenging as the medical datasets have grown over the years. The existing medical image search and retrieval techniques are not very efficient in terms of time and accuracy of search result because most of the existing tools for searching medical images use text based image retrieval techniques.

Text based image classification suffers from some serious limitations, namely, when the size of image collection gets increasingly large, annotating each image manually is very difficult. Also, different people may give different annotations to images with similar visual content. Improving the classification accuracy and reducing the retrieval time are important issues in medical images.

In most medical imaging systems, the same body part is captured from different orientations and magnification by the same sensor. Devise a rotation and scale invariant classification and retrieval system is a real challenge. Medical images are captured by different sensors (modalities). Images captured from various modalities suffer from significant contrast variation between the images of the same organ or body part. Due to this large variation, existing image classification and retrieval algorithms do not perform well for different modality images.

In addition, most of the medical datasets pose data imbalance problem i.e. unequally distributed training samples among all the classes, which gives rise to poor classification performance with standard single classifiers. Finally, one of the most important problems in medical CBIR is to find images with similar anatomical regions and diseases which can greatly reduce the effort exerted by physicians to manually analyse and annotate the disease region.

2.5 SUMMARY

This chapter reviewed some of the existing approaches to content based medical image classification and retrieval. Various steps involved in content based image classification and retrieval system and the related work is briefly described. Also, the existing approaches for all the components of content based image retrieval system are reviewed. Some research issues arising out of the review of existing methods are addressed in this thesis.

CHAPTER 3

CLASSIFICATION OF MEDICAL IMAGES USING EDGE-BASED FEATURES AND DICTIONARY LEARNING

In this chapter, an approach for classification of medical images using edge-based features is proposed. We demonstrate that the edge information extracted from an image by dividing the image into patches and each patch into concentric circular regions provide discriminative information useful for classification of medical images by considering 18 categories of radiological medical images, namely, skull, hand, breast, cranium, hip, cervical spin, pelvis, radiocarpaljoint , elbow etc. The ability of on-line dictionary learning (ODL) to achieve sparse representation of an image is exploited to develop dictionaries for each class using edge-based feature. A low rate of misclassification error for these test images validates the effectiveness of edge-based features and on-line dictionary learning models for classification of medical images.

Digital image retrieval techniques are becoming increasingly important in the field of medical image databases. The increasing dependence on modern medical diagnostic techniques like radiology, histopathology and computerized tomography has led to an explosion in the number of medical images stored in hospitals. Images of various body parts and modalities are becoming an important source of anatomical and functional information for the diagnosis of diseases, medical research and education [52]. However, one cannot utilize the information in these image collections unless they are organized for efficient search and retrieval of data. Efficiently searching and retrieving of data in these large image collections poses significant technical challenges as the characteristics of the medical images differ from other general purpose images. Some methods have been explored in recent years to automatically classify medical image collections into multiple semantic categories for effective retrieval [40]. For example, in [41], the automatic categorization of 6231 radiological images into 81 classes is achieved by utilizing a combination of low level global texture features with low resolution scaled images and a K-nearest neighbor (KNN) classifier. Although these approaches demonstrate promising results for medical image classification and retrieval, classification and search of similar images in a large database is still a challenge due to the enormity of the search space. Searching similar images in a large image repository on the basis of their visual content is called Content Based Image Retrieval (CBIR) [53]. The traditional text based image classification and retrieval (TBIR) approach has many practical limitations like the images in the collection have to be annotated manually which becomes very difficult as the size of the image collection increases and time consuming. Another important limitation of TBIC and TBIR is inadequacy in representing the image content [54]. Content based image classification and retrieval approaches are proposed to overcome the limitations of text based image classification and retrieval. Digital image retrieval techniques are crucial in the emerging field of medical image databases for clinical decision making process.

Medical image classification is an important task in content based medical image retrieval (CBMIR). Automatic medical image classification is a technique for assigning a medical image to an appropriate class among a number of medical image classes. In medical image classification, several methods have been proposed in the literature [55]- [56]. One approach to content based medical image retrieval is proposed in [55], in which medical images are classified based on body orientation, biological system, anatomical region and image modality. The performance of the classification is evaluated on IRMA database and the best classification result is achieved by using distorted tangent distance in a kernel density classifier. The CBMIR system can achieve better performance by filtering out the images of irrelevant classes from the medical database through classification. This significantly reduces the search space and time for retrieving similar type of images. So, image classification is indeed an important stage in a CBMIR system.

The major limitations associated with existing text based image classification and retrieval techniques are: 1) It is time consuming as the physicians have to search through a large number of images for identifying similar images. 2) Most of the existing tools for searching medical images use text based image classification and retrieval techniques. These text based image classification suffer from several limitations [51] and the most important one is the need for manual annotation. Thus, the existing medical image search and retrieval techniques are not very efficient in terms of retrieval time and accuracy of search results.

In this chapter, we address the some issues in text based image classification and retrieval. The content based image classification techniques serve as an alternative to text based image classification. Moreover, CBMIR overcomes the need for manual annotation and human perception. Also, finding similar images in large volumes of medical image databases is a difficult task. Classification of medical images enables the efficient retrieval of relevant images from the large database and reduces the search space and time.

Selection of features for adequately representing the class specific information is an important process in medical image classification. An X-ray image categorization and retrieval method using patch-based visual word representations is proposed in [42]. The feature extraction process 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. The method is especially effective in discriminating orientation and body regions in X-ray images, and in medical visual retrieval. In [43], a descriptor is proposed which combines local features with global shape features. The descriptor combines edge of whole image with edge density of sub-images and is known as the edge density histogram descriptor (EDHD). The image retrieval and classification is then done based on euclidean distance and with the help of support vector machines. A learning based classification framework based on local binary pattern(LBP) feature is proposed in [44]. Local binary pattern is extracted from each image in database with the help of an LBP operator which labels image pixels by thresholding neighborhood of each pixel with the center value and considers the results as a binary number, which is then classified using a maximum margin SVM. Moreover, a merging technique is applied on the overlapped classes. These overlapped classes are detected in merging scheme with the help of measures such as correctness rate of each class, similarity of imaging body organ and misclassification ratio. In [45], a least square support vector machines is used for breast cancer classification. Least square SVM (LSSVM) simplifies the required computation by solving a linear equation set. This equation set embodies all available information about the learning process. The most important difference between SVM and LSSVM is that LSSVM uses a set of linear equations for training while SVM uses a quadratic optimization problem which greatly reduces the computational cost. The extracted feature database is constructed by merging some already existing features in the original database with some new visual content features that are extracted from the medical images using image processing techniques. Wavelet optimization techniques for content based image retrieval in medical database are described in [20].

In most cases, medical images can easily be classified based on edge information. In this chapter, we propose a novel feature extraction method using the edge information. Medical images of different body parts contains different edge information. An edge image is divided into patches and each patch into concentric circular regions. Mean and variance of pixel intensity values in each region is computed. Mean and variance are global measurements and these are more suitable with deterministic methods. In this method, different orientations of same shape images are combined into a single class, in order to achieve better classification. The reason for combining multiple classes to solve a given classification problem is due to the fact that in medical applications, numerous classes of any given medical image database have considerable overlap. Hence, a single class with limited features cannot classify images correctly [54].

Sparse representation has received a lot of attention from the research in signal and image processing. Sparse coding involves the representation of an image as a linear combination of some atoms in a dictionary [57]. Sparse representation is a powerful tool for efficiently representing data. This is mainly due to the fact that signals and images of interest tend to enjoy the property of being sparse in some dictionary. These dictionaries are often learned directly from the training data. Several algorithms like on-line dictionary learning (ODL) [58], K-SVD [59] and method of optimal directions (MOD) [60] have been developed to process training data. Sparse representation is used to match the input query image with the appropriate class. Linear discriminant analysis (LDA) based selection and feature extraction algorithm for classification using wavelet packet has been proposed by Etemand and Chellappa [21]. Recently, similar algorithms for simultaneous sparse signal representation and discrimination have also been proposed [22], [61]. In [62], a method for simultaneously learning a set of dictionaries that optimally represent each cluster is proposed. This approach was later extended by adding a block incoherence term in their optimization problem to improve the accuracy of sparse coding.

In this chapter, we propose an approach for classification of medical images on image retrieval in medical applications (IRMA) database [63] using on-line dictionary learning approach. Learned dictionaries are used to represent datasets in sparse model of IRMA medical images. Dictionaries are designed to represent each class. For a given N number of classes, we design N dictionaries to represent the classes. Each image associated with a dictionary provides the best sparsest representation. For every image in the given set of images $\{y_i\}_{i=1}^{n}$, ODL is used to seek the dictionary D that has the sparsest representation for the image. We define $l(\hat{D}, \hat{\Phi})$ as the optimal value of the l_1-lasso sparse coding problem [64]. This is accomplished by solving the following optimization problem:

$$l(\hat{D}, \hat{\Phi}) = \arg\min_{D,\Phi} \frac{\mathbf{1}}{N} \sum_{i=1}^{N} \frac{\mathbf{1}}{\mathbf{2}} \| Y_i - D\Phi_i\|_{\mathbf{2}}^{\mathbf{2}} \text{ subject to } \|\Phi_i\|_{\mathbf{1}} \leq \lambda, \tag{3.1}$$

where Y is the matrix whose columns are y_i, and λ is the sparsity parameter. D denotes the learned dictionary, Φ represents the sparse representation vectors, N denotes the number of classes, and Y represents the training database. The ODL algorithm alternates between sparse coding and dictionary update steps. Several efficient pursuit algorithms have been proposed in the literature for sparse coding [60], [65]. The simplest one is the l_1-lasso algorithm [64]. The main advantage with ODL algorithm is its computational speed as it uses l_1-lasso algorithm for sparse representation. In sparse coding step, dictionary D is fixed and representation vectors Φ_i are identified for each example y_i. Then, the dictionary is updated atom by atom in an efficient way.

The rest of the chapter is organized as follows. Section 3.1 presents the proposed method. Experiments of content based medical image classification application are described in detail in section 3.2. Finally, we draw the conclusions in section 3.3.

3.1 MEDICAL IMAGE CLASSIFICATION USING DICTIONARY LEARNING

The present work provides a method for medical image classification using dictionary learning. There are many advantages to this approach. Firstly, the edge and patch based feature extraction method proposed to classify the data. Secondly, the entire dataset is represented with the help of fixed small size of dictionary which greatly reduces computational time. Moreover, the classification performance improves because of the uniform dictionary size irrespective of number of training images.

The proposed CBMIR framework is shown in Fig. 3.1. First, the features are extracted from the images of the each training dataset. A dictionary is generated for each class using the on-line dictionary learning (ODL) algorithm. Then, given test data is compared with the existing dictionaries to identify the dictionary with the sparest representation using l_1 lasso algorithm. Finally, test data is assigned to the class associated with the sparsest dictionary. Fig. 3.2 (a) shows some of sample IRMA medical images.

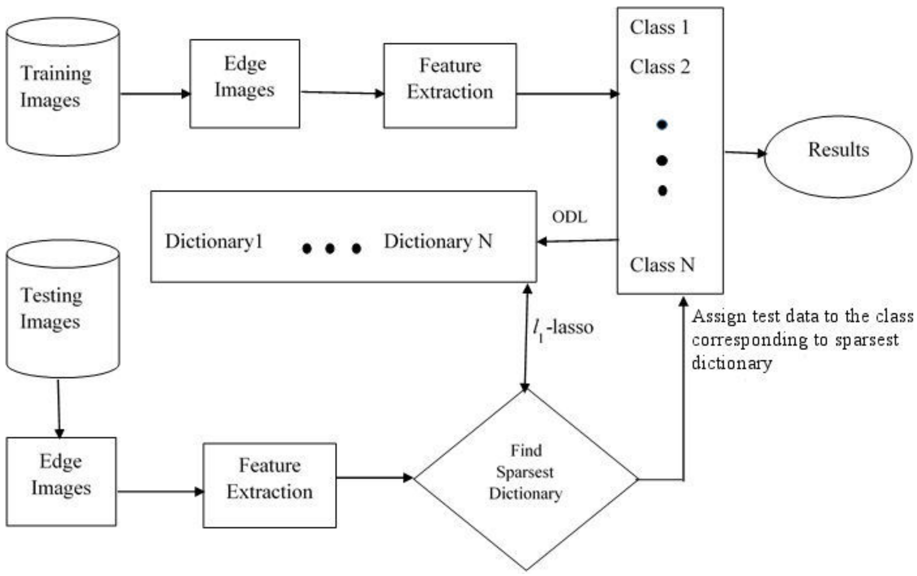

Fig. 3.1: Block diagram of the proposed medical image classification.

3.1.1 Feature extraction

The performance of a CBIR system depends on how well the extracted features capture the semantics of an image. Generally, content based medical image classification and retrieval techniques use fundamental visual features like image colour, shape and texture yielding vector with thousands of features. However, use of multiple features should give better classification accuracy. We consider two types of feature extraction methods to represent the content of medical images. In the first method, edge based feature extraction is used to extract edge information of the medical images. Since, medical images of different body parts contains different shapes and different edge information, medical images can be easily be classified based on the edge features. Canny edge [66] detection algorithm is used for finding the edges of the images as shown in Fig. 3.2 (b). This feature extraction method is more suitable for medical image databases. In the second method, patch based feature extraction method is used on edge images. An edge image is divided into equal size of patches as shown in Fig. 3.2 (c). Each patch of the image is partitioned into concentric circular regions of equal area as shown in Fig. 3.2 (d). The mean and variance of pixel intensity in each

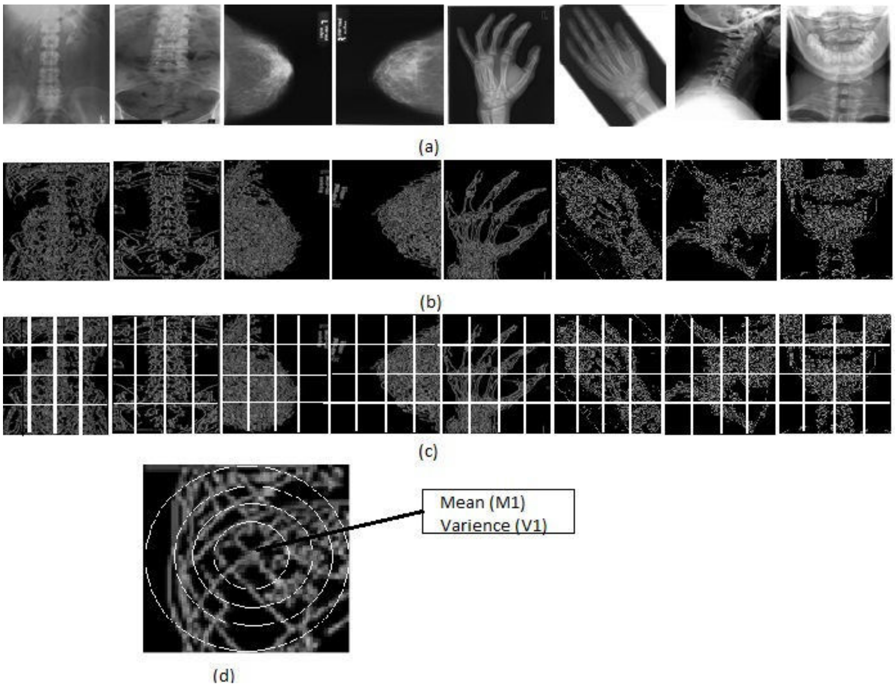

Fig. 3.2: (a) Samples of IRMA medical images. (b) Edge images of samples in (a). (c) Images are divided into equal size of patches. (d) A patch is divided into concentric circular regions.

circular region becomes a component of the feature vector using equations (2) and (3), where P is the number of pixels in each region, m is the mean of pixels intensity values and S is the variance of pixels intensity values in each region. This approach accomplishes the best representation of the contents of an image.

$$m = \frac{1}{P}\sum_{k=1}^{P}(y_k) \tag{3.2}$$

$$S = \sum_{k=1}^{P}(y_k - m)(y_k - m)^t, \tag{3.3}$$

The procedure for feature extraction is as follows:

1. Extract edge information from medical images.

2. Divide each edge image into 16 equally sized (50×50) patches.

3. Partition each patch of the every image into 4 concentric circular regions, such that each circular region has the same number of pixels as the other regions.
4. Calculate mean and variance of each circular region and use them as components of the feature vector. The size of the feature vector for each image is 128×1 (16 (patches) $\times$ 4 (regions) $\times$ 2 (features - mean and variance)).

3.1.2 Proposed method

In this proposed method, we introduce a sparsity based medical image classification by representing the test data as a sparse linear combination of training data from a dictionary. In this work, class $C = [C_{\mathbf{1}}, \ldots, C_N]$ consists of training samples collected directly from the image of interest. In the proposed sparsity model, images belonging to the same class are assumed to lie approximately in a low dimensional subspace. Given N training classes, the $p^{\mathbf{th}}$ class has K_p training images $\{y_i^N\}$ i=1,..., K_p. Let b be an image belonging to the $p^{\mathbf{th}}$ class, then it is represented as a linear combination of these training samples:

$$b = D^p \Phi^p, \tag{3.4}$$

where D^p is $m \times K^p$ a dictionary whose columns are the training samples in the $p^{\mathbf{th}}$ class and Φ^p is a sparse vector. The proposed method is summarized in algorithm 1.

Proposed method consists of two steps:

1) Dictionary Construction: Construct the dictionary for each class of training images using on-line dictionary learning algorithm [58]. Then, the dictionaries $D = [D_{\mathbf{1}}, \ldots, D_N]$ are computed using the equation:

$$(\hat{D}_i, \hat{\Phi}_i) = \arg\min_{D_i, \Phi_i} \frac{\mathbf{1}}{N} \sum_{i=1}^{N} \frac{\mathbf{1}}{\mathbf{2}} \| C_i - D_i \Phi_i \|_{\mathbf{2}}^{\mathbf{2}} + \lambda \| \Phi_i \|_{\mathbf{1}},$$

satisfying $C_i = \hat{D}_i \hat{\Phi}_i$, i$= 1, 2, \ldots, N$.

2) Classification: In the classification process, the sparse vector Φ for given test image is found in the test dataset $B = [b_{\mathbf{1}}, \ldots, b_l]$. Using the dictionaries of training samples

$D = [D_{\mathbf{1}}, \ldots, D_N]$, the sparse representation Φ satisfying $D\Phi{=}B$ is obtained by solving the following optimization problem:

$$\begin{aligned} \Phi^j &= \arg\min_{\Phi} \tfrac{1}{2}\| b_j - D\Phi_j\|_{\mathbf{2}}^{\mathbf{2}} \quad \text{subject to } |\Phi_j\|_{\mathbf{1}} \leq T_{\mathbf{1}}, \textbf{ and} \\ \hat{i} &= \arg\min_i \| b_j - D\delta_i(\Phi^j)\|_{\mathbf{2}}^{\mathbf{2}} \;\; j = 1, \cdots, t, \end{aligned} \tag{3.5}$$

where δ_i is a characteristic function that selects the coefficients. A test image b_j is assigned to class C_i if the i^{th} dictionary that is associated with C_i class gives maximum sparsity for b_j among all the dictionaries while considering l_1 - distance. This procedure is summarized in algorithm 2.

Algorithm 1 : Dictionary Construction for each training class dataset using on-line dictionary learning algorithm (ODL)

Input: Training class dataset $N \epsilon \mathrm{R}^{m \times n}$(C_1, ..., C_N), and $T \epsilon R$ (regularization parameter)

Output: Construct N Dictionaries $D \epsilon \mathrm{R}^{m \times k} = [d_1, ..., d_N]$ (k$\ll n$).

Dictionary construction:

Step 1. For i=1 to N do

Step 2. Construct dictionary D_i for each training class C_i using on-line dictionary learning algorithm (ODL).

$$(\hat{D}_i, \hat{\Phi}_i) = \arg\min_{D_i, \Phi_i} \frac{\mathbf{1}}{N} \sum_{i=1}^{N} \frac{\mathbf{1}}{\mathbf{2}} \| C_i - D_i \Phi_i\|_{\mathbf{2}}^{\mathbf{2}} + \lambda \| \Phi_i\|_{\mathbf{1}}$$

satisfying $C_i = \hat{D}_i \hat{\Phi}_i$, i= $1, 2, \ldots, N$.

Step 3. End for

Step 4. Return D_i

Algorithm 2 : Classification based on sparse Representation

Input: A dictionary matrix $D \epsilon \mathrm{R}^{m \times k}=[d_1, \ldots, d_N]$ for N classes, and a test data $B \epsilon R^{m \times z}$.

Step 1. Normalize the columns of D to have unit l_2- norm

Step 2. Solve the l_1- norm minimization problem by:

$$\begin{aligned} \Phi^j &= \arg\min_{\Phi} \frac{1}{2}\| b_j - D\Phi_j \|_{\mathbf{2}}^{\mathbf{2}} \quad \text{subject to} \| \Phi_j \|_{\mathbf{1}} \leq T_{\mathbf{1}}, \\ \hat{i} &= \arg\min_i \| b_j - D\delta_i(\Phi^j) \|_{\mathbf{2}}^{\mathbf{2}} \;\; j = \mathbf{1}, \cdots, t \end{aligned} \tag{3.6}$$

Step 3. Assign test data b_j to class C_i

3.2 EXPERIMENTAL RESULTS

Experiments are carried out on IRMA medical database, in which each image is of size 200 × 200 pixels. Majority of medical images are generally gray-scale images such as X-ray, CT, etc. Fig. 3.2 (a) shows some of the sample ImageCLEF images of IRMA database. For classification of medical images, 5400 sample images of skull, breast, chest, hand etc., spanning 44 different classes with different orientations are used and these classes are described in Table 3.1. The main problem in classifying medical radiological images is high inter class overlap and intra class variability in some of the classes [54]. To address this problem, different merging techniques are used in literature [54]. In our proposed work, a merging technique is devised where different orientations of the same shaped image are merged into a single class (i.e. number of classes are reduces from 44 to 18) as shown in Table 3.2. Moreover, the proposed method works for images with different orientations. Each class consists of 300 training and 50 testing images, and experiments are run through 5-fold cross validation. The results obtained from these experiments are presented in Table 3.3.

The performance of the proposed method is compared with other classification techniques and given in Table 3.3. The proposed method gives best classification performance of 98.5% as compared to other image classification techniques such as linear discriminant analysis (LDA), kernel SVM, neural network (NN), K-Nearest neighbor

(KNN) and Bayes classifier (BC). The classification performance of different classifiers are shown in Fig. 3.3 in terms of a confusion matrix.

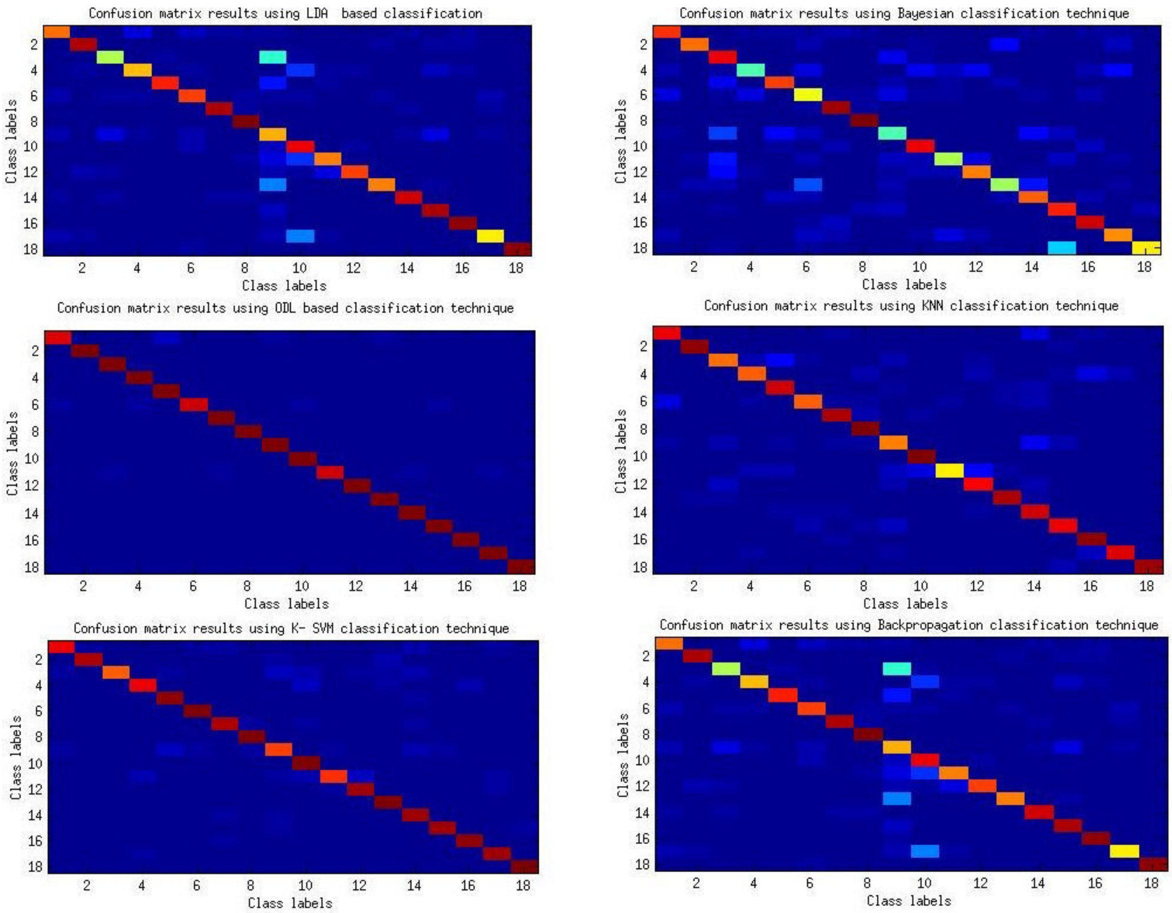

Fig. 3.3: Confusion matrix using (a) LDA classification (b) Bayesian classification (c) ODL classification (d) KNN classification (e) K-SVM classification (f) NN classification

Linear discriminant analysis classifier and Bayes classifier give the classification performance of 77% and 74%, respectively. Neural network classifier is tested with different number of hidden layers. Among these, neural network classifier gives the classification performance of 82%. KNN gives best performance of 88.1% with K=5. When K value increases, the KNN classification performance decreases. The performance of KNN with different K values are shown in Fig. 3.4.

Kernel SVM gives highest performance of 94% using polynomial kernel function.

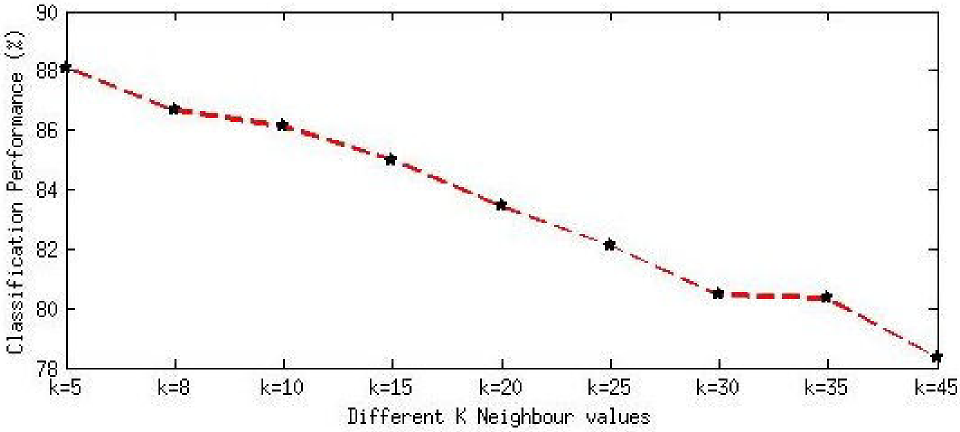

Fig. 3.4: Performance of KNN classifier using different K values.

Further, kernel SVM is explored with different types of kernels, namely, linear, polynomial, RBF, and sigmoid. The best classification results among all classes with various kernels is shown in Fig. 3.5. Also, the performance of all kernels on each individual class is shown in Fig. 3.6. From the experimental results, it is observed that the feature vector selected from multiple features and on-line dictionary based classifiers gives the best performance among all the other classifier methods.

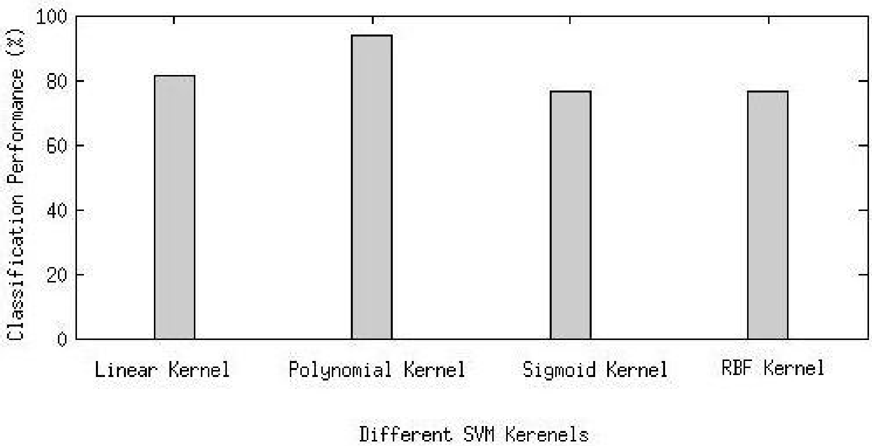

Fig. 3.5: Classification performance of different types of SVM kernels.

Over the years, various methods have been done by taking different number of

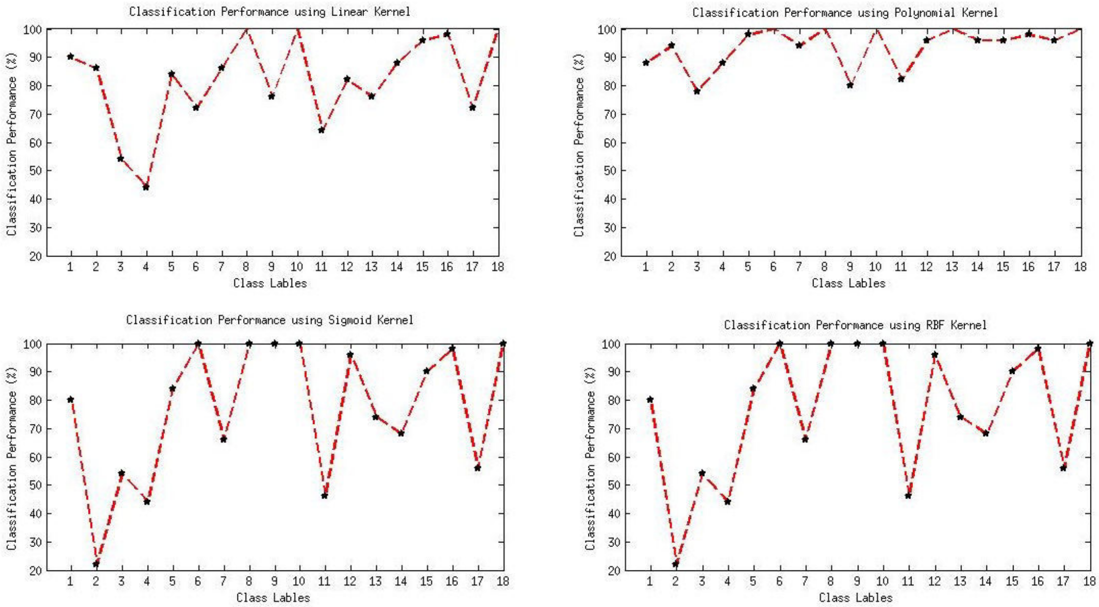

Fig. 3.6: Classification performance of each class using (a) linear kernel SVM. (b) polynomial kernel SVM. (c) sigmoid kernel SVM. (d) RBF kernel SVM.

images from the IRMA medical database. In [55], best classification error rate of 8.0% was achieved for a set of 1617 images from IRMA database. Database consisting of 9100 medical x-ray images of 40 classes are considered in [54]. It provides accuracy rate of 90.83% on 25 merged classes in the first level. Next, if correct classes were considered within the best three matches, then the performance increases to 97.9%. In [41], medical images are classified into 80 classes describing the image direction and modality. In this, 6231 training images are used for classification of medical images and 85.5% correctness is obtained. In [67], for a database consisting of 5000 medical images of 20 classes, classification accuracy of 81.96% is achieved. In [56], an evaluation on a dataset of 1500 images of IRMA database achieved a classification rate of 97.5% in a 17-class classification problem. Fesharaki et al. [68] used the IRMA database for medical image classification. Database includes 4937 X-ray images belonging to 28 different classes. Classes are separated based on the angle of photography and the anatomical area and an accuracy rate of 82.87% was achieved.

3.3 SUMMARY AND CONCLUSIONS

In this work, we have presented an approach for classification of X-ray images using edge-based features and have leveraged the ability of dictionary learning to achieve sparse representation of an image in order to develop dictionaries for each class. Also, a comparative study with other classifiers like kernel SVM, NN, LDA, KNN and Bayes classifier was conducted. The X-ray images database containing 18 categories, namely, skull, hand, breast, cranium, hip, cervical spine, pelvis, radio-carpal joint, elbow etc. were used for training and testing the models. The experimental results indicate that the edge-based features can provide better discrimination among the classes when used in conjunction with on-line dictionary learning. Preliminary computational results are promising and have the potential for practical applications in image classification. The proposed method has achieved best performance of 98.5% which is significantly better than the existing classifiers on the same images.

Table 3.1: X-ray image classes: anatomical, direction. [6](A=Coronal, B=Axial, C=Other orientation D=Sagittal and E=Rotated)

Class	Anatomic	Direction	Class	Anatomic	Direction
1.	Abdomen gastrointestinal	A	23.	Pelvis	C
2.	Abdomen uropoietic	A	24.	Foot	A
3.	Left Breast	B	25.	Radiocarpaljoint	A
4.	Left Breast	C	26.	Radio carpal joint	D
5.	Right Breast	B	27.	Knee	A
6.	Right Breast	C	28.	Knee	D
7.	Hand	A	29.	Knee	B
8.	Hand	C	30.	Elbow	A
9.	Hand	E	31.	Elbow	D
10.	Neck	A	32.	Upperleg	A
11.	Neck	D	33.	Lowerleg	A
12.	Neck	C	34.	Chest bones	A
13.	Cranium	A	35.	Facial cranium	C
14.	Cranium	D	36.	Weber ankle	C
15.	Cranium	C	37.	Weber ankle	A
16.	Hip	A	38.	Shoulder	A
17.	Thoracic spine	D	39.	Shoulder	C
18.	Spinal card	D	40.	Fibrous dysplasia	A
19.	Cervical Spin	D	41.	Fibrous dysplasia	C
20.	Chest	A	42.	Fibrous dysplasia	E
21.	Chest	D	43.	Anklet joint	A
22.	Pelvis	A	44.	Anklet joint	D

Table 3.2: Merged classes of same images with different orientations.

Class	Anatomic numbers	Class	Anatomic numbers
C1	1, 2	C10	25, 26
C2	3, 4, 5, 6	C11	27, 28, 29, 30, 31
C3	7, 8, 9	C12	32, 33
C4	10, 11, 12	C13	34
C5	13, 14, 15	C14	35
C6	16, 17, 18, 19	C15	36, 37
C7	20, 21	C16	38, 39
C8	22, 23	C17	40, 41, 42
C9	24	C18	43, 44

Table 3.3: Comparison of classification performance (%) using different classifiers.

Classes \ Classifiers	NN	K-SVM	BC	ODL	KNN	LDA
C1	76	88	82	90	88	86
C2	94	94	76	100	98	28
C3	54	78	88	100	76	60
C4	68	88	44	100	78	44
C5	84	98	80	100	92	96
C6	80	100	60	92	78	100
C7	94	94	96	100	94	90
C8	100	100	100	100	100	100
C9	70	80	44	100	74	44
C10	88	100	88	100	100	100
C11	74	82	54	92	64	48
C12	80	96	74	100	86	96
C13	74	100	52	100	94	74
C14	92	96	78	100	92	70
C15	94	96	84	100	88	90
C16	98	98	92	100	98	98
C17	64	96	72	100	90	56
C18	98	100	64	100	96	100
Average	82	94	74	**98.5**	88.1	77

CHAPTER 4

CATEGORIZATION OF MEDICAL DATA USING A GENERIC MULTI-LEVEL CLASSIFICATION APPROACH

Classification of medical data is one of the most challenging pattern recognition problem. As stated in literature, a single classifier is unable to solve all medical image classification problems due to high sensitivity to noise and other imperfections like data imbalance. So, several individual classifiers have been studied to solve the different types of classification problems arising in medical datasets but all have proven to be useful on some specific datasets. Hence, in this chapter, we propose a generic multi-level classification approach for categorization of medical data using sparsity based dictionary learning and support vector machine. The proposed technique demonstrates the following advantages: 1) It shows encouraging performance over all the datasets considered. 2) It addresses the problem of data imbalance. 3) It needs no fusion and ensemble of methods in multi-level classification. The results presented on the 5 standard UCI medical datasets demonstrate the efficacy of the proposed multi-level classification approach.

The increasing dependence on modern medical diagnostic techniques like radiology, histopathology and computerized tomography has led to an explosion in the number of medical images stored in hospitals. Several medical image classification systems are available today that categorize medical images based on modalities, body parts, disease and orientation. However, one cannot utilize the information in these image collections unless the data is organized for efficient search and retrieval. With a single classification technique, it may not be possible to solve all medical image classification problems due to imbalanced data problems. So, different classification techniques have to be employed to classify specific datasets. Combinations of multiple classification techniques have been found to give better classification performance than a single classifier. In this chapter, we use two different types of classification approaches to correctly classify various types of medical data. Here, support vector machine (SVM) and sparse representation based on-line dictionary learning (ODL) classification approaches are used to categorize the different types of medical data.

In [46], shape and texture features are extracted from breast MRI images and genetic algorithm is applied to select the best feature to be used for classification process. To improve classification performance three different classifiers, namely, multilayer perceptron (MLP), generalized regression neural network (GRNN), and support vector machine (SVM) are combined to from a multi-classifier system. Influence of different types of distance measures on the performance of a multiple classifier system consisting of one-class classifiers were described by Bartosz et al. [69]. One of the problems in medical image classification is that medical datasets are often imbalanced i.e. more samples of some classes compared to others. Bartosz et al. [47] introduced under sampling balanced ensemble method to solve the imbalance problem. The construction of multiple independent classifiers is typically a non-trivial problem. In [70], atlas-based segmentation and multiple classifiers methods are proposed to solve this problem. The application of performance based decision fusion methods to multi-classifier atlas-based segmentation method is evaluated. Each of 20 subjects is segmented using each of the remaining 19 as the atlas. The resulting 19 segmentations per subject are combined into a final segment. The classification methods proposed in literature often have difficulties with breast cancer datasets. The main reason being that training data is imbalanced with more benign cases recorded than malignant ones.

In [48], a cost-sensitive ensemble classification algorithm is proposed. The data imbalance problem is addressed by employing cost-sensitive decision trees as base classifiers which are trained on random feature subspaces to ensure diversity, and an evolutionary algorithm for simultaneous classifier selection and fusion. Yok-Yen Nguwi et al. [71] introduced an unsupervised self-organizing learning with support vector ranking for imbalanced datasets. This model uses support vector machines for selecting variables so that the problem of imbalanced data distribution can be relaxed. Then, the ranker features are clustered using emergent self-organizing map (ESOM) so as to provide clusters for unsupervised classification. Marco Vannucci et al. [49] described a binary classification method named labeled SOM classification unbalanced sets (LASCUS) that can be applied to uneven datasets and sensitive problems such as malfunction detection. LASCUS method is based on the use of a self-organizing map (SOM) and fuzzy inference system (FIS). The SOM creates a set of clusters to be associated either to frequent or unfrequented situations while the FIS determines such association on the basis of data distribution.

Single classification method is not suitable for classification of various medical image datasets as can be seen in literature [47, 48, 70]. For example, WBC dataset is best classified by KNN classifier and other datasets with KNN classifier gives less classification performance. An extensive literature review revealed the following problems:

1. A single classifier system caters to only a specific medical dataset and performs poorly on others as can be seen in literature [47]. Moreover, it is very susceptible to noise in the data and the performance degrades considerably when noise data is fed as input to any of the individual classifiers.

2. Most of the medical datasets pose data imbalance problems. The imbalanced datasets usually give poor classification performance with standard single classifiers [71]. A multi-level classifier combines correctly classified examples in the first level with the training data and supplies them as input to the next level classifier. So, if there is any data imbalance regarding less number of training samples then it can be alleviated by this method.

3. Main problem with multi-classifier system is how to select classifiers to form an ensemble, and how to fuse the individual decisions of the base classifiers into a single

decision [48].

The problems stated above could be addressed by using multi-level classification approach which does not require ensemble or fusion methods for combining multiple classifiers. Combining the training data along with correctly classified test samples could address the problem of data imbalance. The lack of training data for a given class is compensated by the test samples incorporated in the training data after correct classification.

One of the dictionary learning algorithms, namely, on-line dictionary learning is used in conjunction with support vector machines in the proposed method. Support vector machine is a robust method that has been widely used for classification in various pattern recognition applications. This method was first proposed for classification and regression tasks by Vapnik [72]. We demonstrate the efficacy of our approach on various UCI datasets [73] meant for medical applications. Initially, sparsity based dictionary learning algorithm is applied to classify medical data. Next, correctly classified test data and training data are merged into a single training dataset and given as input to the SVM classifier.

The rest of the chapter is organized as follows. Section 4.1 gives a brief account on dictionary learning. Section 4.2 presents the proposed multi-level classification based on dictionary learning and support vector machine. Experiments on different medical applications are discussed in section 4.3. Finally, in section 4.4, we present the conclusions.

4.1 SPARSE REPRESENTATION AND DICTIONARY LEARNING

Sparse representation has received a lot of attention from the research in signal and image processing. Sparse coding involves the representation of an image as a linear combination of some atoms in a dictionary [57]. It is a powerful tool for efficiently processing data in non-traditional ways. This is mainly due to the fact that closely related images tend to enjoy the property of being sparse in some dictionary. These dictionaries are often learned directly from training data. Several algorithms like on-line dictionary learning (ODL) [58], K-SVD [59] and method of optimal directions (MOD) [60] were developed to process this data. A sparsity measure is used to match the input query image with the appropriate class.

Learned dictionaries give sparse models to represent various datasets in UCI medical data corpus. For a given number of classes N, we design an equal number of dictionaries to represent the classes. Each image is associated with a dictionary that provides the sparsest representation. For every image in the given set of images $\{y_i\}_{i=1}^{n}$, on-line dictionary learning (ODL) is used to seek a dictionary D that has the sparsest representation for the image. We define $l(\hat{D}, \hat{\Phi})$ as the optimal value of the l_1 sparse coding problem [64]. This is accomplished by solving the following optimization problem.

$$l(\hat{D}, \hat{\Phi}) = \arg\min_{D,\Phi} \frac{\mathbf{1}}{N} \sum_{i=1}^{N} \frac{\mathbf{1}}{\mathbf{2}} \| Y_i - D\Phi_i \|_{\mathbf{2}}^{\mathbf{2}} \ \textbf{subject to} \ \| \Phi_i \|_{\mathbf{1}} \leq \lambda, \tag{4.1}$$

where Y is the matrix whose columns are y_i and λ is the sparsity parameter. D denotes the learned dictionary, Φ represents the sparse representation vectors, N denotes the number of classes and Y represents training database. The ODL algorithm alternates between sparse coding and dictionary update steps. In the sparse coding step, the dictionary D is fixed and the representation vectors Φ_i are identified for each example y_i. Several efficient pursuit algorithms [60, 65] have been proposed in the literature for sparse coding. The simplest one is the l_1-lasso algorithm [64]. In the next step, the dictionary is updated atom by atom.

4.2 MULTI-LEVEL CLASSIFICATION APPROACH TO MEDICAL DATA

In this section, we explain the multi-level classification scheme to improve the performance on imbalanced medical datasets. The motivation for present work is to overcome some problems involved in single and multiple classifier systems related to medical database classification problems which are stated in the previous section.

The proposed multi-level classification scheme for medical datasets is depicted in Fig. 4.1. In the training phase, dictionaries are developed based on sparsity of training feature vectors for each class using on-line dictionary learning and all the dictionaries are combined to form a single dictionary. During testing, the sparsity of a test data is computed with the dictionaries of each class using the l_1-lasso distance. The class which exhibits maximum sparsity is then assigned as the class for that test data. Then, correctly classified results are merged with original training dataset to form a new training dataset. The updated training data and the original test data sets are given as input to support vector machine classifier to classify medical database. The three different phases of the proposed classification system, namely, feature extraction, sparse coding based on dictionary learning, and SVM are described in the following subsections.

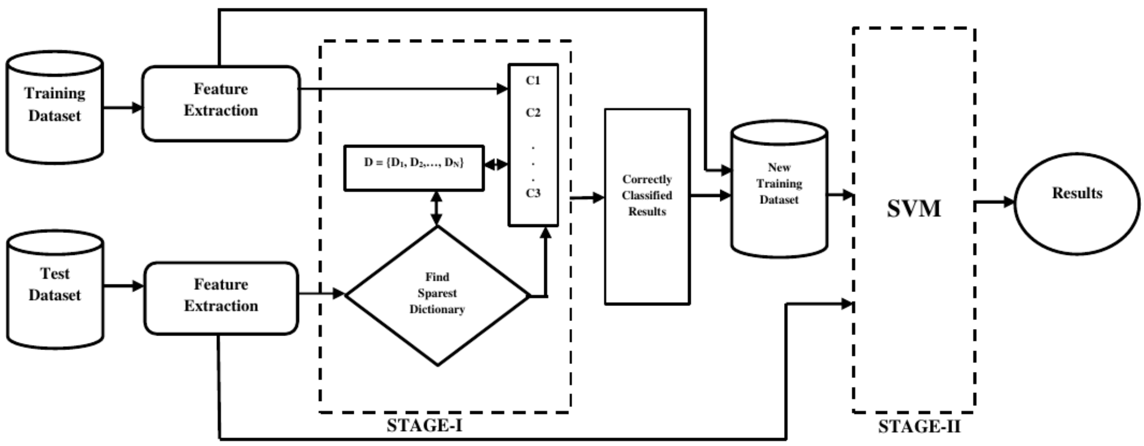

Fig. 4.1: Block diagram of the multi-level classification framework using on-line dictionary learning and support vector machine.

4.2.1 Feature extraction

In this work, five different types of medical image datasets are used, namely, SPECTF (Heart), Heart-Statlog, Wisconsin Breast Cancer Diagnostic (WBCD), Pima Indians

Diabetes (PIMA) and Wisconsin Breast Cancer (WBC) all from the UCI repository. Different medical datasets contain different types of feature values. A brief description of the features extracted from various datasets are presented below:

Dataset 1: Wisconsin Breast Cancer Diagnostic (WBC):

This data set contains 30 continuous features, computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. They describe characteristics of the cell nuclei present in the image, such as the perimeter, the area, the symmetry, and the number of concave portions of the contour.

Dataset 2: Pima Indians Diabetes:

The Pima Indians Diabetes data set contains 8 features. The features include age, number of times pregnant, diastolic blood pressure and body mass index, among others.

Dataset 3: SPECTF (Heart):

The SPECTF data set contains 44 continuous feature patterns which was created for each patient.

Dataset 4: Wisconsin Breast Cancer (WBC):

The Breast data set contains 9 features. The features include clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli and mitoses.

Dataset 5: Heart-Statlog :

The Breast data set contains 13 features. The features include age, sex, chest pain type, resting blood pressure, serum cholesterol, fasting blood sugar, resting electro cardio graphic results, maximum heart rate achieved, exercise induced angina, old peak, the slope of the peak exercise ST segment, number of major vessels.

Now the features extracted from the above datasets are given as input to form a sparse dictionary using on-line dictionary learning. The following subsection describes about dictionary construction and sparsity based classification approach.

4.2.2 On-line dictionary learning and sparsity based classification

In the proposed method, we introduce, at the first level, a sparsity based medical image classification algorithm by representing the test data as a sparse linear combination of training data from a dictionary. On-line dictionary learning is a data-driven approach which provides the best possible sparse representation for the image thereby improving the accuracy of classification. Class $C = [C_{\mathbf{1}}, \ldots, C_N]$ consists of training samples collected directly from the image of interest. The images related to the same classes are assumed to approximately lie in a low dimensional subspace. For a given N classes, the $p^{\mathbf{th}}$ class has K_p training images $\{y_i^N\}$ i=1,..., K_p. Let b an image belongs to the $p^{\mathbf{th}}$ class, then it is represented by a linear combination of these training samples.

$$b = D^p \Phi^p, \tag{4.2}$$

where D_p is a $m \times K_p$ dictionary whose columns are the training samples in the $p^{\mathbf{th}}$ class. And, Φ_p is a sparse vector. On-line dictionary learning and sparsity based classification method mainly consists of two steps:

- *Dictionary Construction:* Construct the dictionary for each class of training images using on-line dictionary learning algorithm [58]. Then, the dictionaries $D = [D_{\mathbf{1}}, \ldots, D_N]$ are computed using the equation.

$$(\hat{D}_i, \hat{\Phi}_i) = \arg\min_{D_i, \Phi_i} \frac{\mathbf{1}}{N} \sum_{i=1}^{N} \frac{\mathbf{1}}{\mathbf{2}} \| C_i - D_i \Phi_i \|_{\mathbf{2}}^{\mathbf{2}} + \lambda \| \Phi_i \|_{\mathbf{1}}$$

  satisfying $C_i = \hat{D}_i \hat{\Phi}_i, \quad i = \mathbf{1}, \mathbf{2}, \ldots, N.$

- *Classification:* In the classification process, find the sparse vector Φ for given test image in the test dataset $B = [b_{\mathbf{1}}, \ldots, b_l]$. The dictionary of training samples $D = [D_{\mathbf{1}}, \ldots, D_N]$, the sparse representation vector Φ satisfying $D\Phi$=B is obtained by solving the following optimization problem:

$$
\begin{aligned}
\Phi^j &= \arg\min_{\Phi} \tfrac{1}{2}\|b_j - D\Phi_j\|_{\mathbf{2}}^{\mathbf{2}} \quad \text{subject to } \|\Phi_j\|_{\mathbf{1}} \leq T_{\mathbf{1}}, \text{ and} \\
\hat{i} &= \arg\min_i \|b_j - D\delta_i(\Phi^j)\|_{\mathbf{2}}^{\mathbf{2}} \;\; j = \mathbf{1}, \cdots, t \;,
\end{aligned}
\tag{4.3}
$$

where δ_i is a characteristic function that selects the coefficients. A test image b_j is assigned to class C_i if the i^{th} dictionary that is associated with C_i class gives maximum sparsity for b_j among all the dictionaries while considering l_1 - distance. The following subsection describes the second level classification approach using SVM.

4.2.3 Multi-level classification approach

In a single-level classification system, the classes having more examples achieve better representation than the ones having fewer examples in order to achieve generalization. This leads to lower classification performance for imbalanced datasets. In multi-classifier approach, choosing a suitable method for combination of results from various classifiers involved requires exhaustive testing. Fusion and ensemble techniques are widely used as combination methods [74], but they are ill defined and dataset specific. Multi-level classification eliminates this whole process of choosing combinations and delivers better performance on all the datasets considered here. The new training dataset is formed based on the first level classification results obtained using dictionary learning. After first level classification, correctly classified results are merged with initial training dataset members to form a new updated training dataset. In the second level of classification, support vector machine classifier is used to categorize test data based on new updated training dataset. The training set in multi-level classification is augmented by correctly classified examples, the dependence on noisy training data as well as the bias towards highly populated classes is significantly reduced. This multi-level classification approach is more suitable for imbalanced medical datasets. The proposed method is different from adaboosting method. Adaboosting method uses the weighted voting technique and a weight assigned to a classifier depends on its error on the training set.

4.3 EXPERIMENTAL RESULTS AND DISCUSSION

In our experiments, we have used five different types of medical datasets, namely, SPECTF (Heart), Heart-stalog, Wisconsin Breast Cancer Diagnostic (WBCD), Pima Indians Diabetes (PIMA) and Wisconsin Breast Cancer (WBC) selected from UCI database. Different medical data contain different types of objects and feature values shown in Table 4.1.

Table 4.1: Datasets used in experiments.

Dataset name	# of objects	# of attributes	# of classes
WBC	699	9	2
WBCD	569	32	2
SPECTF	267	44	2
PIMA	768	8	2
Heart-Statlog	270	12	2

In our experiments, we have used two different types of breast cancer datasets. Both datasets represent binary classification problems (i.e benign and malignant), and both are highly imbalanced datasets. Wisconsin Breast Cancer original (WBC) is a well known and publicly available breast cancer dataset made available by the University of Wisconsin hospitals [75]. In total, there are 699 samples of which 241 are malignant and 458 are benign. Another breast cancer dataset is Wisconsin Breast Cancer Diagnostic (WBCD) data. WBCD data consists of 569 instances with 32 binary attributes. Single Proton Emission Computed Tomography (SPECTF) heart data set is composed as normal and abnormal classes. It consists of 267 instances with 44 attributes. There are 40 samples of each class in the training datasets and test datasets composed of 172 normal samples and 15 abnormal samples. PIMA dataset contains the data from all female patients of at least 21 years old. The database consists of 768 instances, each with 8 attributes. Heart-Statlog dataset is composed as absence and presence classes. It consists of 270 instances with 12 attributes.

The performance of the proposed system is evaluated by measuring classification accuracy, sensitivity and specificity. Sensitivity and specificity are statistical measures for the performance of a binary classification test. The sensitivity and specificity are calculated from true positive (TP), false negative (FN), false positive (FP), and true negative (TN).

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FN + FP)} \ , \tag{4.4}$$

$$Sensitivity = \frac{(TP)}{(TP + FP)} \ , \ and \tag{4.5}$$

$$Specificity = \frac{(TN)}{(TN + FP)}. \tag{4.6}$$

The proposed method gives better classification performance than existing single and multiple classifier methods applied on five different type of UCI medical datasets. Experiments on the WBCD, WBC, PIMA and Heart-Statlog datasets were run through 5 -fold cross validation. The experimental results obtained from these experiments are presented in Table 4.2 - 4.4.

Table 4.2 shows the performance evaluation obtained on WBCD data. In this table, F-BP, F-kNN, F-SVM, F-Bayes and multi-agent classifiers give good classification results because they use multi-classification technique. However, determining the right ensemble for fusion is a difficult task. This problem is alleviated in our implementation which utilizes multi-stage classification and hence, the true performance of both classifiers is explored giving rise to 99.1% accuracy. Table 4.3 depicts the results of the proposed method and various classifiers on Wisconsin Breast Cancer original (WBC) dataset. It can be noted that the proposed method gives highest performance of 98 % compared with the existing single and multiple classification methods.

Table 4.4 represents the performance of proposed method with existing methods on Heart-StatLog dataset. The performance of the proposed method gives 88%, which is the best classification accuracy when compared to other single and multiple

Table 4.2: Performance comparison of multi-level classification with state-of-the-art approaches on Wisconsin Breast Cancer Diagnostic dataset.

Author	**Method/Classifier**	**Accuracy (%)**
Balasubramanian, V. et al. 2009 [76]	Random Sampling	92.8
	Margin-based SVM	83.6
	Query by Committee	80
	Ho-WechslersInitial QBT	46.8
	GQBT	28
Fangqing Peng et al. 2009 [77]	Multi-agent	96.58
	F-Bayes	96.32
	F-BP	96.11
	F-KNN	96.26
	F-SVM	96.04
Jing Wei et al. 2013 [78]	k2	94.03
	SDBNS	95.59
	ECFBN	95.76
Proposed	**ODL+SVM**	**99.1**

classification methods including the state-of-the-art methods. Table 4.5 enlists the various classification schemes applied on the Pima indians diabetes dataset including our proposed method. The PIMA database consists of 768 instances, each with eight attributes. A total of 268 patients were diagnosed as having diabetes and 500 patients are healthy persons without diabetes. Performance of proposed method on this dataset depicts the classification accuracy close to state-of-the-art.

Table 4.6 represents the performance of proposed method with existing methods on SPECTF (Heart) dataset. The performance of the proposed method gives 97.8%, which is the best classification accuracy when compared to other single and multiple classification methods including the state-of-the-art approaches. In Table 4.7, the

Table 4.3: Performance comparison of multi-level classification with state-of-the-art approaches on Wisconsin Breast Cancer original (WBC).

Author	**Method/Classifier**	**Accuracy(%)**
Myungraee Cha et al. 2014 [79]	Support vector data description	94.8
	Density weighted SVDD	96.2
	Liu et al. (2013) [80]	96.6
Yuwono, M. et al. 2012 [81]	Multi-agent	96.8
	RCE	96
	RCE+	96.08
	Swarm RCE+	95.89
Duch, W. et al. 2012	K2MLP	97
Yuanyuan Guo. et al. 2012 [82]	1-NN	92.46
	LLGC	65.52
	SVM	96
	TSVM	97
Ramos-Pollan, R. 2010 [83]	Grid based	95.8
Sheng-Yi Jiang et al. 2009 [84]	C4.5	96.09
	RIPPER	95.99
	Naive-Bayes	97.32
Proposed	**ODL+SVM**	**98**

performance of various single classifiers on different medical datasets is presented. It can be noted that on PIMA dataset QDA gives best performance of 83.3%. However, in all other datasets, the proposed classification scheme out performs all others, making it reliable for use over a variety of medical applications.

Figs. 4.2 and 4.3 show the sensitivity and specificity measures of different type of classifiers on various UCI medical datasets. Proposed method gives best sensitivity and specificity results among all classifiers.

Table 4.4: Comparison of performance of classification with state-of-the-art approaches on Heart-StatLog dataset.

Author	**Method/Classifier**	**Accuracy(%)**
Christoph F. Eick et al. 2004 [85]	Nearest representative	83.8
	Wilson	80.4
	1-NN	76.7
	C4.5	78.2
Rodda, S et al. 2007 [86]	Associative Classifier	82.81
Sheng-Yi Jiang et al. 2009 [84]	C4.5	81.48
	RIPPER	82.33
	Naive-Bayes	84.33
Kemal Polat et al. 2009 [87]	Combining of RBF kernel F-score feature selection and LS-SVM classifier	83
Yuanyuan Guo. et al. 2010 [88]	1-NN	53.26
	LLGC	70.4
	SVM	57.8
	TSVM	83.93
Koji Ouchi et al. 2011 [89]	Logistic Regression with a ridge estimator	83.7
	Naive Bayes	83.7
Wodzisaw Duch et al. 2012 [82]	LVQ	85.07
Proposed	**ODL+SVM**	**88**

4.4 SUMMARY AND CONCLUSIONS

In this chapter, a multi-level classification approach using on-line dictionary learning and SVM classification methods for UCI medical data classification is proposed. In

Table 4.5: Performance comparison of multi-level classification with state-of-the-art approaches on Pima Indians Diabetes dataset.

Author	**Method/Classifier**	**Accuracy(%)**
Yuwono, M 2012 [81]	RCE	65.64
	RCE+	65.64
	Swarm RCE+	65.6
Duch, W. et al. 2012 [82]	Naive Bayes	75.3
	SVML	77.08
Ouchi, K. et al. 2011 [89]	logistic regression with a ridge estimator	77.21
Yuanyuan Guo. et al. 2010 [88]	1-NN	64.84
	LLGC	65.1
	SVM	70
	TSVM	71
Sheng-Yi Jiang et al. 2009 [84]	C4.5	77.73
	RIPPER	77.3
	Naive Bayes	77.28
Chen, S.-C et al. 2006 [90]	SA+BPN	82.16
Zhongwei Li et al. 2006 [91]	Cascade Structure	79.89
Christoph F. Eick et al. 2004 [85]	Nearest representative	74.5
	Wilson	73.4
	1-NN	69
	C4.5	74.5
Proposed	**ODL+SVM**	**82**

all the datasets barring one (PIMA), the performance of the proposed multi-level classification scheme is significantly better than the single classifiers. On-line Dictionary learning being a data-driven approach provides the better possible sparse representa-

Table 4.6: Performance comparison of multi-level classification with state-of-the-art approaches on SPECTF (Heart) dataset.

Author	**Method/Classifier**	**Accuracy(%)**
Myungraee Cha et al. 2013 [79]	support vector data description	82.7
	Density Weighted SVDD	95.4
	Liu et al. (2013) [80]	90
Jing Wei et al. 2013 [78]	k2	94.03
	SDBNS	95.59
	ECFBN	95.76
Kumar, R. et al. 2013 [92]	mc-MKC Matrix Completion - Multiple Kernel Completion	79.9
	mc-SVM Matrix Completion -SVM	79.1
Duch, W et al. 2012 [82]	SVMG	80.18
Cui Li-lin et al. 2010 [93]	TCM-IKN N	90
Tian, D. et al. 2007 [94]	C-GAME+Johnson+c4.5	84.4
	RMEP+Johnson+c4.5	41
	C4.5	81.7
Proposed	**ODL+SVM**	**97.8**

tion for the images thereby improving the accuracy of classification. Also, multi-level classification scheme works better than other multiple classifier schemes which suffer from the problem of ensemble selection. Thus, this method proves to be an all-round strategy for medical image classification.

Table 4.7: Comparison of performance (in %) using individual classifiers on different medical datasets.

Method \ Dataset	WBCD	WBC	Heart-StatLog	PIMA	SPECTF
KNN	94.6	96.5	72.5	68.5	70.5
Neural Network	89.3	86.1	82.3	83.3	73.26
Naive Bayes	92	97.2	71.2	75.9	81.8
LDA	89.3	91.6	74.5	81.4	58.2
QDA	92.1	90.2	74.5	**83.3**	53.4
SVM	96.2	85.4	76	81.3	73.4
ODL	96.5	96.5	79	81.4	94.1
(ODL+SVM)	**99.1**	**98**	**88**	82	**97.8**

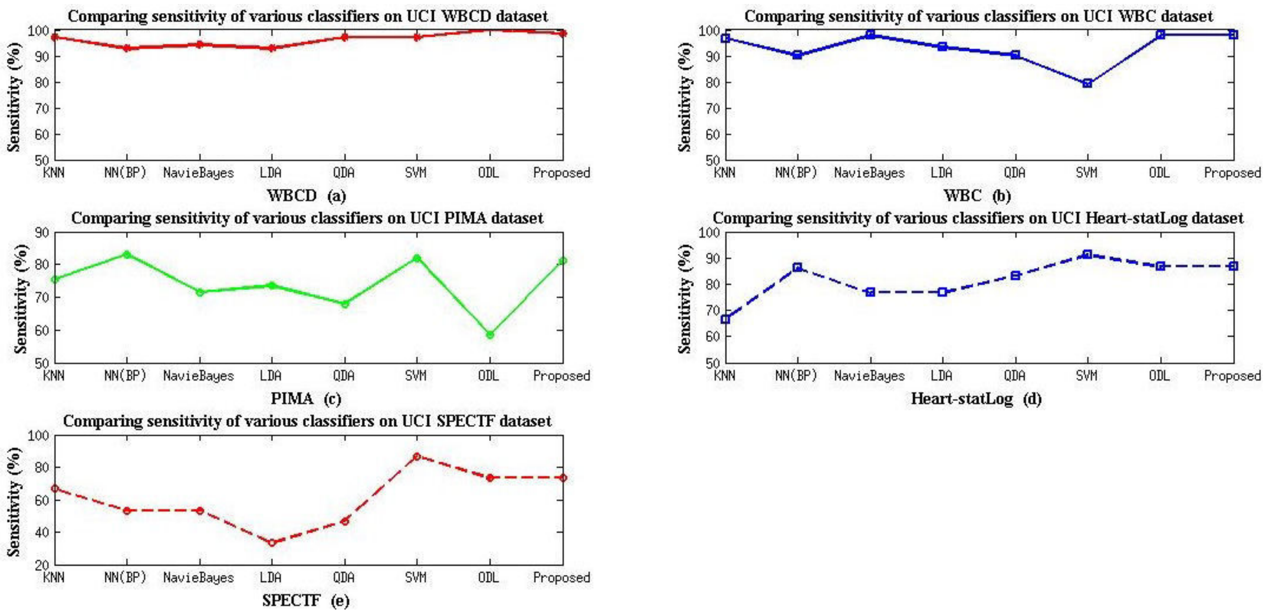

Fig. 4.2: Sensitivity measure of proposed method (ODL+SVM) on various UCI medical datasets.

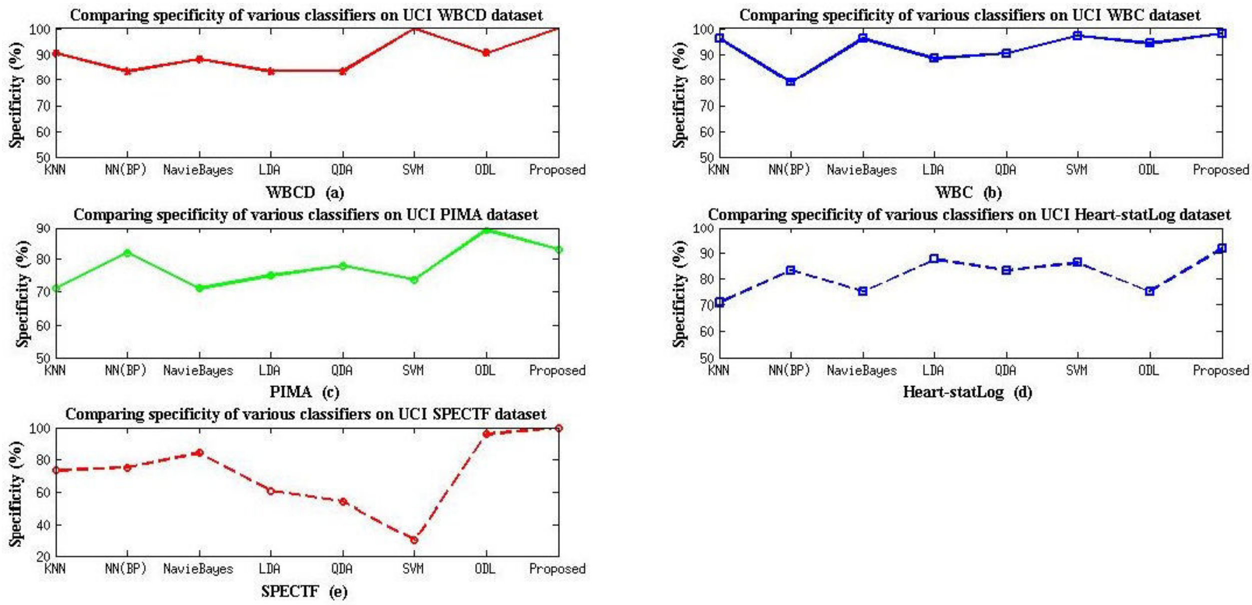

Fig. 4.3: Specificity measure of proposed method (ODL+SVM) on various UCI medical datasets.

CHAPTER 5

CLASSIFICATION OF MEDICAL IMAGES CAPTURED BY DIFFERENT SENSORS BASED ON MULTI-SCALE WAVELET REPRESENTATION USING DICTIONARY LEARNING

In this chapter, we propose a method for classification of medical images captured by different sensors (modalities) based on multi-scale wavelet representation using dictionary learning. Wavelet features extracted from an image provide discriminative information useful for classification of medical images, namely, diffusion tensor imaging (DTI), magnetic resonance imaging (MRI), magnetic resonance angiography (MRA) and functional magnetic resonance imaging (FRMI). The ability of on-line dictionary learning (ODL) to achieve sparse representation of an image is exploited to develop dictionaries for each class using multi-scale wavelet features. The experimental analysis performed on a set of images from the ICBM medical database demonstrates efficacy of the proposed method.

Modern medical diagnostic techniques like radiology, histopathology and computerized tomography generate a lot of medical images that need to be indexed, archived and stored for future use. The medical image classification systems available today classify medical images based on modality, body part, disease or orientation. The enormous amount of medical images with a wide variety of image modalities such as diffusion tensor imaging (DTI), magnetic resonance imaging (MRI), magnetic resonance angiography (MRA) and functional magnetic resonance imaging (FRMI) are available on medical databases. Effectively and efficiently searching and retrieving of medical image data in these different modality image collections poses significant technical challenges as the characteristics of the medical images differ from other general purpose images. Traditional text based image retrieval (TBIR) cannot handle these problems because of its many practical limitations [51]. One of these problems is that images in the collection have to be annotated manually which becomes very difficult as the size of the image collection increases and time consuming. Another important limitation of TBIR is inadequacy in representing the image content [54]. Content based image retrieval approaches were proposed by [53] to overcome the limitations of text based image retrieval. Content Based Image Retrieval (CBIR) gives a way of searching similar images in a large image repository on the basis of their visual content. When applied for medical images, CBIR can retrieve images of similar nature (like same modality and disease) and characteristics and this process is known as Content Based Medical Image Retrieval (CBMIR).

Medical image classification is an important task in CBMIR. Automatic medical image classification is a technique for assigning a medical image to an appropriate class among a number of medical image classes. In medical image classification, several methods and algorithms have been presented in the literature [55]- [56]. One approach to content based medical image retrieval is proposed in [55], in which medical images are classified based on body orientation, biological system, anatomical region and image modality. The performance of the classification is evaluated on IRMA database and the best classification result is achieved by using distorted tangent distance in a kernel density classifier. The CBMIR system can achieve better performance by filtering out the images of irrelevant classes from the medical database because it reduces the search space and time for retrieving similar type of images. This establishes the importance of image classification in a CBMIR system. In literature, it has been suggested that modality is one of the most important filters that can limit the search and retrieval time [95].

Content based medical image classification (CBMIC) overcomes the need for manual annotation and human perception. Also, finding similar images in large volumes of medical image databases is a difficult task. Modality based classification of medical images enables the efficient retrieval of relevant images from the large database and reduces the search space and time. Multimodality during capturing images suffers from significant contrast variation between the images of the same scene. Due to this large variation, existing image classification and retrieval algorithms do not perform well for different types of modality images.

Selection of features for adequately representing the class specific information is an important process in medical image classification. The classification performance mostly depends on the extracted features. Commonly, there exists a semantic gap between the content of an image and its visual features. Thus, decreasing the semantic gap through extracting more effective features has still remained as a challenging topic in content based image classification and retrieval task. Wavelet features were used to overcome the semantic gap between low level and high level features [96]. Features extracted from sub-bands in a multi-resolution space are useful for extracting some high level features. And, capturing images of various modalities suffers from significant contrast variation between the images of the same organ or body part. Due to this large variation, existing image classification and retrieval algorithms do not perform well for different modality images. In this chapter, we propose a new classification technique, namely, sparse representation based multi-scale dictionary learning to classify the different type of modality images. Multi-scale image representation can handle the semantic gap between low and high level features and intensity variations of the different modality images.

An X-ray image categorization and retrieval method using patch-based visual word representations is proposed in [97]. The feature extraction process is based on local patch representation of the image content and a bag-of-features approach for defining image categories. These features are then applied to a kernel SVM for classification. The method is especially effective in discriminating orientation and body regions in X-ray images, and in medical visual retrieval. Modality classification and its use in text based image retrieval in medical databases is proposed in [50]. Visual descriptors and text features are used for classifying the medical images. Medical image classification is then done with the help of support vector machines classifier. In [51], different types of medical image modality classification and retrieval strategies are explored. Bags of visual words and Fisher vectors representations are integrated to perform medical image modality classification. Quellec et al. [20] proposed a CBIR system where each image is represented by its wavelet transform. The distribution of wavelet coefficients in each sub band defines a signature. The signature thus obtained is compared to the signature of the query image using a distance measure based on pathology and image modality. The similarity is also weighted between sub-bands and the procedure to obtain weight is guided by an optimization procedure.

5.1 FEATURE EXTRACTION

The performance of a content based image classification system depends on the representation of an image as a feature vector. Generally, content based image classification techniques use fundamental visual features like image's color, shape and texture yielding vectors with thousands of features. But, using these features directly, one cannot classify images easily. In the proposed method, multi-scale wavelet packet decomposition based feature extraction method is used. Wavelet features were used to overcome the semantic gap between low level and high level features [96]. Wavelet packet decomposition can be implemented by progressively applying two channel filter banks. At every stage each filter bank comprises of a low-pass (L) and a high-pass (H) filter and whose sampling frequency is half of that of the previous stage. As a consequence, the original image can be decomposed into four sub-images, namely, both horizontal and vertical directions have low-frequencies (LL), the horizontal direction has low frequencies and the vertical one has high-frequencies (LH), the horizontal direction has high frequencies and the vertical one has low frequencies (HL) and both horizontal and vertical directions have high-frequencies (HH) sub-images. Next, construct a gradient vector for each sub-image. Similar approach applied for the entire training and testing database images to form a feature vector. The procedure for feature extraction is as follows:

1. Apply the wavelet packet decomposition on an original image to obtain the LL, LH, HL and HH sub-images.
2. Construct a gradient vector for each sub-image.
3. Repeat step (1) and (2) for all original training and testing images to form a gradient feature vector
4. Combine the similar sub-bands (e.g. LL) from all the images of each class to form a feature vector matrix. This will generate four feature vector matrices for the four sub-bands for each class.

The following subsection describes introduction about sparse representation.

5.1.1 Sparse representation

Sparse representation has received a lot of attention from the research in signal and image processing. Sparse coding involves the representation of an image as a linear combination of some atoms in a dictionary [57]. These dictionaries are often learned directly from the wavelet coefficients of training data. Several algorithms like on-line dictionary learning (ODL) [58], K-SVD [59] and method of optimal directions (MOD) [60] have been developed to process training data. Sparse representation is used to match the input query image with the appropriate class. Linear discriminant analysis (LDA) based selection and feature extraction algorithm for classification using wavelet packet has been proposed by Etemand and Chellappa [21]. Recently, similar algorithms for simultaneous sparse signal representation and discrimination have also been proposed [22], [61]. In [62], a method for simultaneously learning a set of dictionaries that optimally represent each cluster is proposed. This approach was later extended by adding a block incoherence term in their optimization problem to improve the accuracy of sparse coding. Multi-scale dictionary learning is proposed in [98]. It combines the advantages of generic multi-scale representations with the K-SVD dictionary learning method.

In this chapter, we propose a modality based classification method for International Consortium for Brain Mapping (ICBM) database using wavelet based on-line dictionary learning approach. Learned dictionaries are used to represent datasets in sparse model of ICBM medical images. Dictionaries are designed to represent each class. For a given N number of classes, we design N dictionaries to represent the classes. Each image associated with a dictionary provides the best sparsest representation. For every image in the given set of images $\{y_i\}_{i=1}^{n}$, ODL is used to seek the dictionary D that has the sparsest representation for the image. We define $l(\hat{D}, \hat{\Phi})$ as the optimal value of the l_1-lasso sparse coding problem [64]. This is accomplished by solving the following optimization problem:

$$l(\hat{D}, \hat{\Phi}) = \arg\min_{D,\Phi} \frac{\mathbf{1}}{N} \sum_{i=1}^{N} \frac{\mathbf{1}}{\mathbf{2}} \| Y_i - D\Phi_i\|_{\mathbf{2}}^{\mathbf{2}} \text{ subject to } \|\Phi_i\|_{\mathbf{1}} \leq \lambda, \tag{5.1}$$

where Y is the matrix whose columns are y_i and λ is the sparsity parameter. D denotes the learned dictionary, Φ represents the sparse representation vectors, N denotes the number of classes and Y represents the training database. The ODL algorithm alternates between sparse coding and dictionary update steps. Several efficient pursuit algorithms have been proposed in the literature for sparse coding [60], [65]. The simplest one is the l_1-lasso algorithm [64]. Main advantage with ODL algorithm is its computational speed as it uses l_1-lasso algorithm for sparse representation. In sparse coding step, dictionary D is fixed and representation vectors Φ_i are identified for each example y_i. Then, the dictionary is updated atom by atom in an efficient way.

The rest of the chapter is organized as follows. Section 5.2 presents the proposed method. Experiments of modality based medical image classification application using sparse representation are discussed in detail in section 5.3. Finally, we draw the conclusions in section 5.4.

5.2 MEDICAL IMAGE CLASSIFICATION USING SPARSE REPRESENTATION AND ON-LINE DICTIONARY LEARNING (ODL) ALGORITHM

The present work provides a method for medical image classification using the framework of multi-scale dictionary learning. There are many advantages to this approach. Firstly, the feature extracted from sub-bands in a multi-resolution space are useful for extracting some high level features. With the help of high level features to overcome the semantic gap. Secondly, the entire dataset is represented with the help of fixed small size of dictionary which greatly reduces computational time.

The following subsection describes sparse representation based classification method.

5.2.1 Sparsity based medical image classification

In this proposed method, we introduce a sparsity based medical image classification by representing the test data as a sparse linear combination of training data from a dictionary. In this chapter, each class $C_i = [c_{ib1}, \ldots, c_{ib4}]$ (each class contains 4 sub-bands feature vector matrices i.e. for class $C_1 = [c_{1b1}, c_{1b2}, c_{1b3}, c_{1b4}]$) consists of all classes training samples collected directly from the wavelet coefficients of same sub-bands. In the proposed sparsity model, images belonging to the same class are assumed to lie approximately in a low dimensional subspace. Given N training classes, the $p^{\mathbf{th}}$ class has K_p training images $\{y_i^N\}$ i=1,. . . , K_p. Let r be an image belonging to the $p^{\mathbf{th}}$ class, then it is represented as a linear combination of these training samples:

$$r = D^p \Phi^p, \tag{5.2}$$

where $D^{\mathbf{p}}$ is $m \times K^p$ a dictionary whose columns are the training samples in the $p^{\mathbf{th}}$ class and $\Phi^{\mathbf{p}}$ is a sparse vector.

Proposed method consists of two steps:

1) Dictionary Construction: In the wavelet packet decomposition, domain contains a collection of coefficient images or sub-bands. The different wavelet coefficients capture data at different scales and orientations. As such it makes sense that separate dictionaries be used to represent these images. Construct the dictionary for each sub-band of class (D_{ib}), where i is the number of classes (i.e. i=1,..,4) and b is the number of sub-bands in each class (i.e. b=1,...,4) using on-line dictionary learning algorithm [58]. Then, the dictionaries for all training class on same sub-band is $D_b = [D_{\mathbf{1}b}, \ldots, D_{\mathbf{4}b}]$ (if b=1, then D_{4b} means fourth class and first sub-band dictionary) and computed using the equation:

$$(\hat{D}_i, \hat{\Phi}_i) = \arg\min_{D_i, \Phi_i} \frac{\mathbf{1}}{N} \sum_{b=1}^{\mathbf{4}} \sum_{i=1}^{N} \frac{\mathbf{1}}{\mathbf{2}} \| C_{ib} - D_{ib} \Phi_{ib} \|_{\mathbf{2}}^{\mathbf{2}} + \lambda \| \Phi_{ib} \|_{\mathbf{1}},$$

satisfying $C_i = \hat{D}_i \hat{\Phi}_i, \quad i = \mathbf{1}, \mathbf{2}, \ldots, N.$

2) Classification: In this classification process, the sparse vector Φ for given test image is found in the test dataset $Z = [z_{\mathbf{1}}, \ldots, z_l]$. The dictionaries of training samples of each class on same sub-band are given by $D_b = [D_{\mathbf{1}b}, \ldots, D_{\mathbf{4}b}]$. The sparse representation Φ satisfying $D_b \Phi$=Z is obtained by solving the following optimization problem:

$$\begin{array}{lcl} \Phi^l & = & \arg\min_{\Phi} \sum_{b=1}^{4} \frac{1}{2} \| z_{lb} - D_b \Phi_{bl} \|_{\mathbf{2}}^{\mathbf{2}} \quad \text{subject to } \| \Phi_{\mathbf{l}} \|_{\mathbf{1}} \leq T_{\mathbf{1}}, \\ \textit{and} \ \ \hat{i} & = & \arg\min_i \| z_l - D\delta_i(\Phi^l) \|_{\mathbf{2}}^{\mathbf{2}} \ \ l = \mathbf{1}, \cdots, t, \end{array} \tag{5.3}$$

where δ_i is a characteristic function that selects the coefficients. Then z_l is assigned to C_i associated with the i^{th} dictionary. It means, finding the sparsest dictionary for a given test data using l_1 -lasso algorithm. Then, test data is assigned to the class associated with this sparsest dictionary.

In the classification phase, each sub-image acquired from the test image is matched with the trained dictionaries of only that sub-image. The class which yields maximum sparsity is chosen as the class for that sub-band. Once all the sub-images are evaluated, the class which agrees with the majority of the sub-bands is chosen as the category for the test image.

5.3 EXPERIMENTAL RESULTS

In this section, we show the effectiveness of the proposed modality based medical image classification method using multi-scale dictionary learning and sparse representation. Data used in the preparation of this work were obtained from the international consortium for brain mapping (ICBM) database (www.loni.usc.edu/ICBM). The ICBM project (principal investigator John Mazziotta, M.D., University of California, Los Angeles) is supported by the national institute of biomedical imaging and bioEngineering. ICBM is the result of efforts of co-investigators from UCLA, montreal neurologic institute, university of texas at san antonio, and the institute of medicine, juelich/heinrich heine university - germany.

Experiments are carried out on ICBM medical database, in which each image is of size 200×200 pixels. Majority of medical images are generally gray scale images such as X-ray, FMRI, MRI etc. The main problem in classifying medical radiological images is high inter class overlap and intra class variability in some of the classes [54]. For tackling this problem, wavelet packet decomposition based feature extraction method is used to overcome semantic gap between low level features and high level features. Moreover, the proposed method works for images with various sensors. ICBM database consisting of a four different type of image modalities such as diffusion tensor imaging (DTI), magnetic resonance imaging (MRI), magnetic resonance angiography (MRA) and functional magnetic resonance imaging (FRMI). Entire database of images are divided into 70% training and 30% testing for each class and experiments are run through 5-fold cross validation. Each class consists of 5587 training and 1482 testing images. Proposed method tested with various wavelet families, namely, Harr, Daubechies, Coiflets, Symlets, Discrete Meyer, and Biorthogonal. The experimental results are presented in Table 5.1. The proposed method was tested with dictionaries size of 60, 80 and 100. Generally, accuracy improves for larger sized dictionaries. However, after a certain point, increase in dictionary size does not yield better classification accuracy. The dictionary size at this point of time gives the best possible sparse representation of the given feature descriptor. In our case, recognition rate of 91.6% was obtained for dictionary size of 80. The confusion matrices for SVM, KNN, Bayesian and the proposed classification method on the ICBM dataset are shown in Figs. 5.1, 5.2, 5.3 and 5.4, respectively.

Table 5.1: Classification accuracy (%) of multi-scale dictionary learning method using wavelet decomposition based features and different dictionary sizes.

Wavelet Families	**60 Dict**	**80 Dict**	**100 Dict**
Daubechies(db4)	86.3	86.9	85.6
Daubechies(db10)	85.6	85.9	84.4
Harr(db2)	90.3	**91.6**	90.7
Discrete Meyer	86.7	86.2	86.6
Coiflets	87.2	87	86.8
Symlets2	90.2	89.5	87
Biorthogonal	87	87.8	87

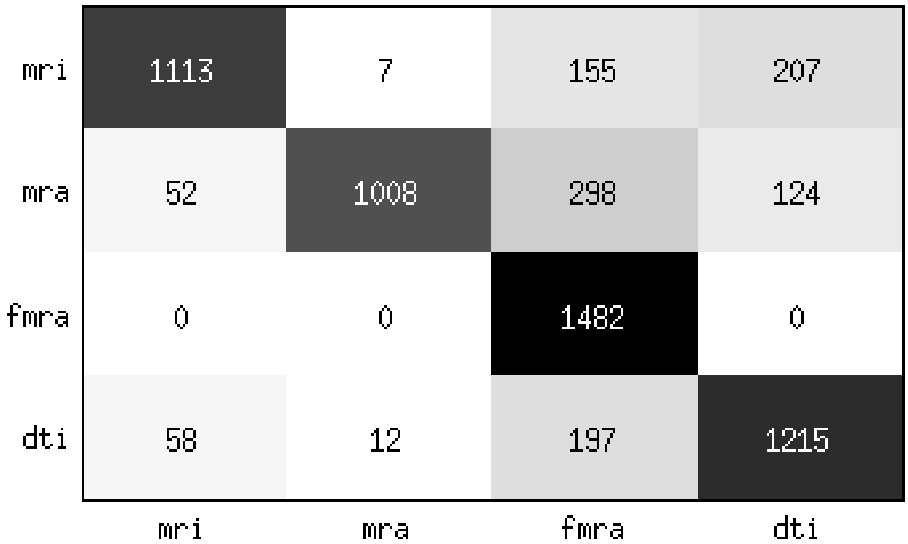

Fig. 5.1: Confusion matrix of medical modality image classification using SVM with haar wavelet feature.

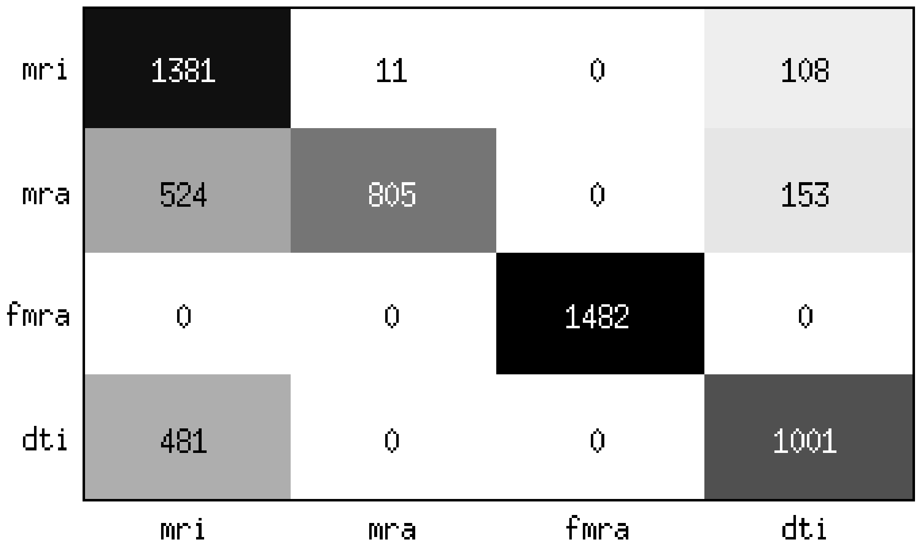

Fig. 5.2: Confusion matrix of medical modality image classification using neural network method with haar wavelet feature.

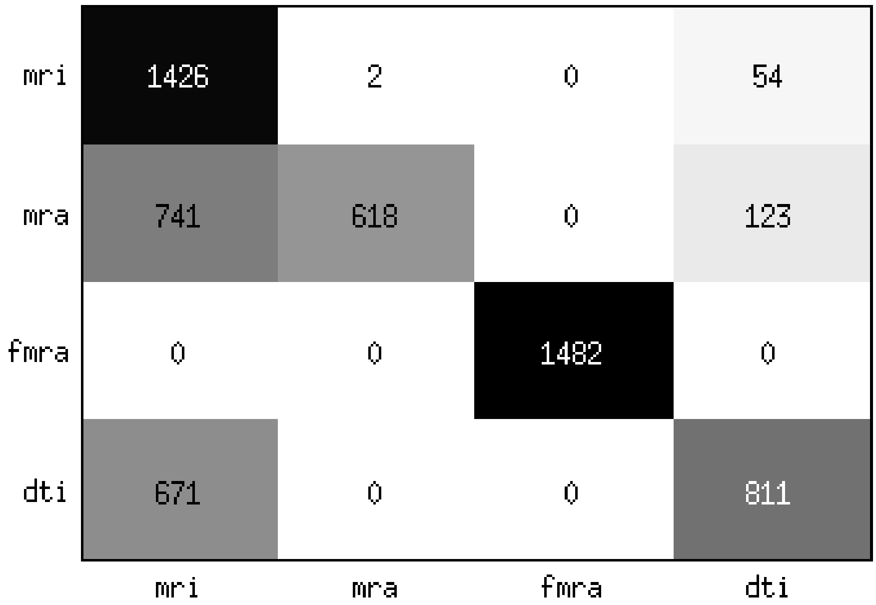

Fig. 5.3: Confusion matrix of medical modality image classification Bayesian classification with haar wavelet feature.

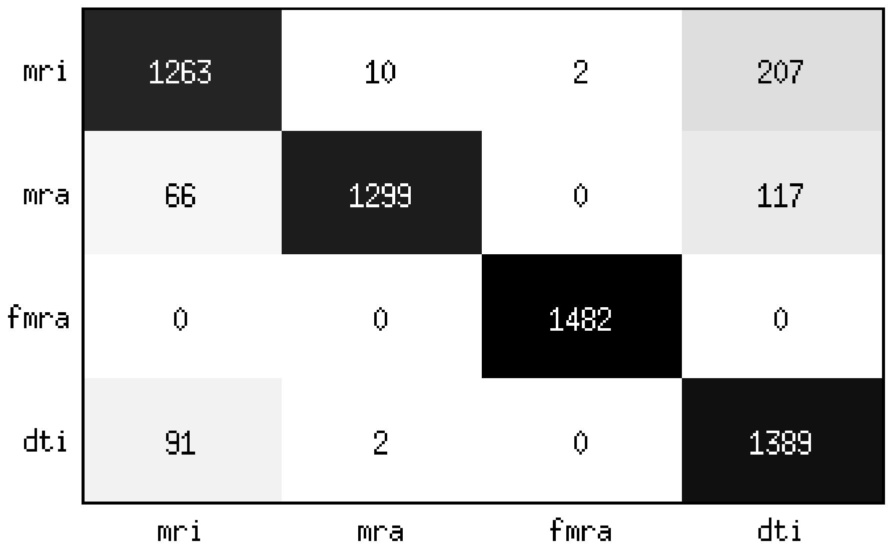

Fig. 5.4: Confusion matrix of medical modality image classification using multi-scale dictionary learning.

The proposed method gives classification performance of 91.6% which is better than other image classification techniques such as SVM, neural network, and Bayes classifier. The classification performance of different classifiers are shown in Table 5.2.

Table 5.2: Classification accuracy (%) of the multi-scale dictionary learning method with different classifiers on ICBM dataset.

Classifier	**Accuracy (%)**
SVM	81.2
Neural Network(BP)	78.3
Bayesian	73.1
Proposed	**91.6**

Wavelet packet decomposition generates gradient vectors individually for each of the four sub-bands. Although distinct, these gradient vectors by themselves do not have enough discriminative capabilities. Using different combinations of the gradient vectors may yield different discriminating characteristics [99].

Classification accuracy of different possible combinations of the gradient vectors extracted from the four sub-bands are presented in Table 5.3. It can be observed that LL sub band contains more information among the four sub-bands. The classification accuracy based on the gradient vectors extracted from the LL sub-band is 84.3%. The classification accuracy based on the gradient vectors extracted from the LH, HL, and HH sub-bands were 73.4, 70.2, and 73.8 %, respectively. To increase the classification accuracy, we can combine all sub-bands sparsity results. Various combination sequences were tried and best classification accuracy of 91.6% was achieved after combining the dictionaries from all the sub-bands. Testing images are classified based on majority of the all sub-bands sparsity results.

Table 5.3: Classification accuracy (%) of multi-scale dictionary learning method based on individual and all combination of the sub-bands obtained from wavelet decomposition.

Subband	**Accuracy (%)**
LL	84.3
LH	73.4
HL	70.2
HH	73.8
LL+LH+HL+HH	**91.6**

5.4 SUMMARY AND CONCLUSIONS

In this chapter, we proposed a method for classification of medical images captured by different sensors (modalities) based on multi-scale wavelet representation using dictionary learning. We have exploited the ability of ODL to achieve sparse representation of an image, to develop dictionaries for each class using wavelet features. Other classifiers, namely, SVM, NN and Bayes were also examined. The medical images database containing four different type of modality(sensors) images, namely, diffusion tensor imaging (DTI), magnetic resonance imaging (MRI), magnetic resonance angiography (MRA) and functional magnetic resonance imaging (FRMI) was used for training and testing the models. Experimental results indicate that the wavelet packet decomposition based features provide useful information for discriminating the classes. Preliminary computational results are promising and have the potential for practical image classification. The proposed method has achieved best performance of 91.6%. The experimental results suggest that the proposed method performs better when compared to other classification approaches.

CHAPTER 6

CLASSIFICATION OF HEARTBEAT USING ADAPTIVE LEARNING

Cardiovascular diseases (CVD) are a leading cause of unnecessary hospital admissions as well as fatalities placing an immense burden onthe healthcare industry. A process to provide timely intervention can reduce the morbidity rate as well as control rising costs. Patients with cardiovascular diseases require quick intervention. Towards that end, automated detection of abnormal heartbeats captured by electronic cardiogram (ECG) signals is vital. While cardiologists can identify different heartbeat morphologies quite accurately among different patients, the manual evaluation is tedious and time consuming. In this chapter, we propose new features from time and frequency domains and further more, feature normalization techniques to reduce inter-patient and intra-patient variations in heartbeat cycles. Our results using the adaptive learning based classifier emulate those reported in existing literature and in most cases deliver improved performance, while eliminating the need for labeling of signals by domain experts.

Modern medical diagnostic techniques like radiology, histopathology and computerized tomography generate a lot of medical images that need to be indexed, archived and stored for future use. The medical image classification systems available today classify medical images based on modality, body part, disease or orientation. Classification of heartbeats is a fundamentally challenging problem. Cardiovascular diseases (CVD) are a leading cause of fatality representing 30% of all global deaths [100]. In 2008, an estimated 17.3 million individuals died of cardiovascular diseases. Third world countries account for 80% of CVD related deaths. In 2010, CVD related illnesses cost the United States healthcare industry \$316.4 billion. A large number of admissions to hospitals are unnecessary and avoidable. Due to inadequate preventive measures, CVD related fatalities continue to rise. It is imperative that we find a solution that reduces these fatalities. One way is to identify high risk patients is using simple and inexpensive tools. An automated system that can identify potential risks of patients can aid optimizing the usage of medical resources. Such systems must be able to identify patterns in cardiovascular activity that can pose a threat to the patients. Further more, in rural areas, where access to healthcare facilities is poor, early detection systems can be potentially life saving and cost effective. Electrocardiogram (ECG) is a widely used device to monitor heart function irregularities. At present, an expert cardiologist analyzes ECG plots to detect abnormalities. However, such an analysis is done over short durations of an ECG signal. Since, certain kinds of heartbeat arrhythmias are time consuming to detect, the patient may require long term monitoring. Hu et al [101] and Chazal et al [102] proposed a set of time domain and ECG morphology features and evaluated the classification performance using Linear Discriminant Analysis. Both approaches require that in addition to the standard training set, a specified number of heartbeats of a new test patient is labeled by a domain expert and added to the training set, which may be difficult to obtain in practice. Wiens et al [103] proposed an active learning technique to reduce the number of labeled heartbeats required for a new test patient. Other approaches, Alvarado et. al. [104] focused on data compression without compromising on classification performance.

In this work, we build on existing techniques and propose a technique to detect two types of heartbeat arrhythmias, namely, ventricular ectopic beats (VEB) and supra ventricular ectopic beats (SVEB). We propose new features from time and frequency domains and further more, a data normalization technique to reduce inter-patient and intra-patient variations. Our results are comparable to those reported in existing literature and in most cases give improved performance. The chapter is organized as follows. Section 6.1 describes the sources of data, data sets, and features used. In Section 6.2, classification methodology is described. Section 6.3 describes the results and comparisons with existing methods.

6.1 DATA DESCRIPTION

Heartbeat patterns in an ECG signal is identified by a cardiac cycle consisting of P-QRS-T waveforms. The P-QRS-T waveforms consist of 5 successive deflections in amplitude, known as P, Q, R, S and T waves as shown in Fig. 6.1.

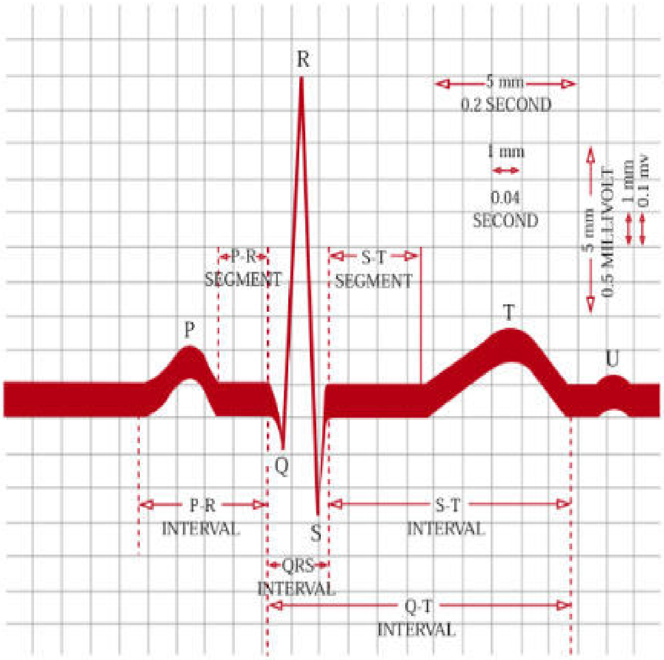

Fig. 6.1: Cardiac cycle of a typical heartbeat represented by the P-QRS-T wave form.

These patterns tend to vary within a patient recording resulting in intra-patient variations. In addition to intra-patient variations, these patterns exhibit inter-patient variations. This makes heartbeat classification a challenging problem. To effectively classify a heartbeat, a classifier must be able to take into account both inter-patient and intra-patient variations in ECG signal. Fig. 6.2 shows the inter-patient variation of heartbeat pattern for patient 119 and 106. We used MIT/Beth Israel Hospital (BIH) Arrhythmia Database available in PhysioBank archives [105]. The database includes 48 Electrocardiogram (ECG) recordings obtained from 47 subjects. Each ECG recording is sampled at 360 Hz for a duration of half hour. ECG recording is susceptible to noise such as power line interference and baseline wander. Before the feature extraction, the ECG signal is preprocessed to reduce the baseline wander and 60 Hz power line interference. To remove baseline wander, signal is passed through median filters of window sizes 200ms and 600ms. The first median filter removes P-waves and QRS complexes and second median filter removes the T-waves leaving behind the baseline wander. By subtracting the baseline wander from the original signal, we obtain the filtered signal. The power line interference is removed using a notch filter centered at 60Hz. The database has annotations for 20 different types of heartbeats, with each heartbeat annotated by an expert cardiologist. The annotation includes the location of the R-Peak and the corresponding heartbeat label. The R-Peak is the peak of QRS complex as seen in Fig. 6.2. The heartbeat label indicate the type of heartbeat.

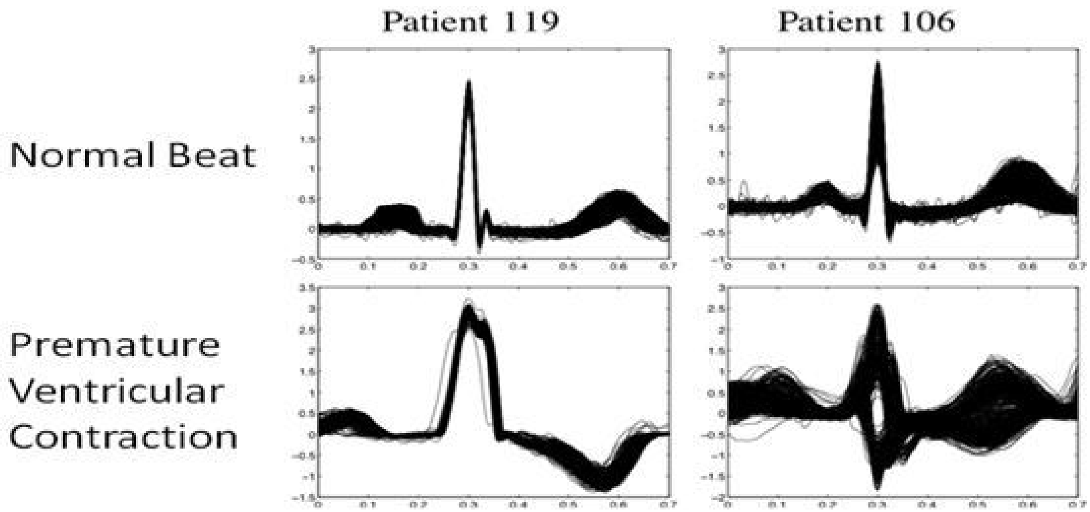

Fig. 6.2: Examples of heartbeat shapes from the MIT-BIH data set.

Each column represents a patient and each row the beats for that specific class. Variations can be seen in the beat morphology across patients as well for a single patient (Source Alvarado et. al. [106])

American association of medical instrumentation (AAMI) protocol define five classes of heartbeat. In accordance with the AAMI protocol, we grouped together the 20 types of heartbeats available in MIT-BIH arrhythmia database into five classes. They are normal and bundle branch block beats (N), supra-ventricular ectopic beats (SVEBs), ventricular ectopic beats (VEBs), fusion of normal and VEBs (F), and unknown beats (Q). Although there exist 5 classes, our problem is a binary classification problem. For the detection of SVEB, a heartbeat is classified as either SVEB or not SVEB (N, VEB, F and Q). Similarly, for the detection of VEB, the heartbeat is classified as either VEB or not VEB (N, SVEB, F and Q). The data was divided into two disjoint sets of patients DS1 and DS2, containing 22 patients each. In accordance with the AAMI protocol [107], four patients with paced beats were not considered for the study. The training dataset was derived from dataset DS1 and testing dataset was derived from dataset DS2. In other words, training set DS1 is used to train the global classifier, which is then tested on test set DS2 containing a new set of patients. Note that our approach do not require apriori knowledge of patient specific labeled beats from the test set, unlike certain other techniques [102], [103], [104] in existing literature. DS1 and DS2 comprise of the following recordings:

DS1 = {101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230};

DS2 = {100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234};

Paced beats = {102, 104, 107, 217}.

Note that paced beats are excluded from analysis.

6.1.1 Feature Extraction

The time domain features, ECG morphology features and frequency domain features are extracted from the ECG signal. Out of the 18 features extracted, 12 features are time domain features, 2 are ECG morphology and 3 are frequency domain features. The 18$^{\text{th}}$ feature is a flag, indicating 0 or 1. Time domain features include RR Interval features, QRS duration, QR duration, RS duration and T wave duration, energy of QRS complex, energy of QR segment, energy of RS segment and energy of T wave.

Energy of a signal is calculated as the sum of squares of magnitude of samples in that segment. The RR interval features include the pre-RR interval, post-RR interval, average RR interval and local average RR interval. Pre-RR interval is the time interval between the current R-peak and the preceding R-peak and post-RR interval is the time interval between the current R-peak and the next R-peak. Average RR interval is the average of all the RR intervals in a recording. Local average RR interval is calculated as the average of 10 RR intervals surrounding a heartbeat. QRS duration is the time interval between the QRS onset and the QRS offset. QR duration is the time interval between the QRS onset and the R-peak. RS duration is the time interval between the R-peak and the QRS offset.

The ECG morphology features consist of fixed interval morphology features from the QRS complex and the T wave of a heartbeat cycle. In order to form ECG morphology features, the ECG signal was down sampled to 120 Hz. Once down sampled, 2 samples to the left of R-peak, the sample value at R-peak and 2 samples to the right of R-peak were extracted. In order to extract the T wave features, 9 samples representing the T wave were extracted. Linear interpolation was applied to extract the T wave samples [108]. The frequency domain features include maximum Fourier coefficients at QRS complex, QR segment of QRS complex and RS segment of QRS complex. In addition to time domain features, ECG morphology features and frequency domain features, we also extracted the P wave flag, which is a binary flag representing the presence or absence of P wave associated with a beat. In total, we extracted 18 different types of features for lead A. The features were extracted for every heartbeat in the 30 minute recording of each patient. Feature selection involves the selection of the best subset of 18 features that maximize the classifier performance. We used three time domain (pre-RR interval, local average RR interval and energy of T wave), five ECG morphology (R peak, 2 samples to the left of R-peak at 120 Hz and 2 samples to the right of R-peak at 120 Hz) and two frequency domain (Max. Fourier coefficient of QR segment and max. Fourier coefficient of RS segment) as feature vector.

6.2 CLASSIFICATION

In this chapter, we develop an approach for classification of normal and abnormal heartbeat using adaptive learning. We designed the classifier for use in a clinical setting, where physicians have little time to label beats, let alone tune classifier parameters. Then, correctly classified results are merged with original training dataset to form a new training dataset. The updated training data and the original test data sets are again given as input to classifier to classify medical database. This process is repeated until results are converged. Adaptive learning based classification approach improves the classification accuracy when compared with single time classification approach. The proposed method is different from adaboosting method. Adaboosting method uses the weighted voting technique and a weight assigned to a classifier depends on its error on the training set.

6.3 EXPERIMENTAL RESULTS

A variety of metrics are used in the realm of classification. Adhering to common practice in heartbeat classification, we used the metrics listed below. The classification results are reported in terms of accuracy (Acc) and calculated from true positive (TP), false negative (FN), false positive (FP), and true negative (TN). Accuracy is defined as follows:

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FN + FP)}, \tag{6.1}$$

In our experiments, training data consists of 45833 normal heart beat samples, 942 SVEB samples, and 3785 VEB samples, and test data consists of 44228 normal heart beat samples, 1836 SVEB samples and 3219 VEB samples. Two different ways of experiments are conducted on this training and testing datasets. First one is to classify the normal, SVB, and VEB heart beats using various classifiers with and with out adaptive learning mechanism. Table 6.1 reports the classification results using single time classification approach. Classification performance is measured in terms of its accuracy. The results of single classification techniques such as linear discriminant analysis (LDA), QDA, dictionary learning (DL), neural network (NN), K-Nearest neighbor (KNN) and Bayes classifier (BC) are shown in Table 6.1. Columns in Table 6.1 represents the classifiers accuracy results.

Table 6.1: Comparison of classification performance (%) using individual classifiers without adaptive learning.

Classifiers / Classes	QDA	LDA	KNN	NN	DL
Normal	95.6	99.1	91.9	99.4	84.3
SVEB	91.6	77.8	48.2	87	56.4
VEB	92.8	85.8	69	91.9	78.2

Table 6.2 reports the classification performance using adaptive learning based classification approach. Proposed approach gives improved performance compared with the individual classifiers.

Table 6.2: Comparison of classification performance (%) using individual classifiers with adaptive learning.

Classifiers / Classes	QDA	LDA	KNN	NN	DL
Normal	97.3	99.2	99.6	99.6	86.4
SVEB	93.6	94.2	68.8	89.5	68.2
VEB	96.4	97.8	83.6	92.6	84.8

Second approach is to classify only SVEB and VEB heart beats using various classifiers with and without adaptive learning mechanism. Table 6.3 reports the classification results using single time classification approach. Table 6.4 reports the classification performance using adaptive learning based classification approach. Among these, proposed approach produces improved performance relative to sensitivity and positive predictive value.

Table 6.3: Comparison of classification performance (%) using individual classifiers without adaptive learning.

Classifiers / Classes	QDA	LDA	KNN	NN	DL
SVEB	93.2	84.5	78.1	94.7	78.4
VEB	94.6	98.3	95.2	97.4	88.2

Table 6.4: Comparison of classification performance (%) using individual classifiers with adaptive learning.

Classifiers / Classes	QDA	LDA	KNN	NN	DL
SVEB	95.5	93.6	96	97.4	89.2
VEB	97.6	97.3	97.1	98.6	91.3

6.4 SUMMARY AND CONCLUSIONS

In this chapter, we have shown that by distinguishing between inter-patient and intra-patient variations, classification performance can be improved significantly. We proposed a new set of features in the time domain and frequency domain, and demonstrated the significance of using pre-RR interval information for classification. Furthermore, the proposed method is fully automated and it eliminates the requirement for patient specific labeled data.

CHAPTER 7

CONTENT BASED MEDICAL IMAGE RETRIEVAL USING DICTIONARY LEARNING

In this chapter, a clustering method using dictionary learning is proposed to group X-ray medical images based on sparse representation for efficient search and retrieval from large database. An approach to group similar images into clusters that are sparsely represented by the dictionaries and simultaneously learn dictionaries from the clusters using K-SVD method is proposed. A query image is matched with the existing dictionaries to identify the dictionary with the sparest representation using orthogonal matching pursuit (OMP) algorithm. Then, images in the cluster associated with this dictionary are compared using a similarity measure to retrieve images similar to the query image. The performance of the proposed method is examined for IRMA test image database. The experimental results demonstrate the efficacy of the proposed method in retrieval of medical images.

There are billions of images available on the internet. Nevertheless, one cannot utilize the information in these image collections unless they are organized for efficient search and retrieval of data. The problem of searching for similar images in a large image repository based on the content is called content based image retrieval (CBIR) [53]. The traditional text based image classification and retrieval (TBIR) approach has many practical limitations like the images in the collection have to be annotated manually which becomes very difficult as the size of the image collection increases and time consuming [51]. Another important limitation of text based image classification (TBIC) and TBIR is inadequacy in representing the image content [54]. Content based image retrieval (CBIR) approaches are proposed to overcome the limitations of text based image retrieval. Digital image retrieval techniques are crucial in the emerging field of medical image databases for clinical decision making process. Digital image retrieval can be used to retrieve images of a similar nature (like same modality and disease) and characteristics. The increasing dependence on modern medical diagnostic techniques like radiology, histopathology and computerized tomography has led to an explosion in the number of medical images stored in hospitals. The images of various modalities are becoming an important source of anatomical and functional information for the diagnosis of diseases, medical research and education [52]. Existing medical CBIR systems also suffer from some serious limitations which are as follows: 1) In most cases, physicians have to browse through a large number of images for identifying similar images which is time consuming. 2) Most of the existing tools for searching medical images use text based image retrieval techniques. So, the existing medical image search and retrieval techniques are not very efficient in terms of search time and accuracy of results. Another important issue in medical CBIR is to find images with similar anatomical regions and diseases. For example, in case of brain tumor images, the tumor can be at any of the different stages and an image of the tumor in a state could be in any orientation. So, there is a need for rotation invariant medical image retrieval technique to find images (of different orientation) of a similar (same stage) tumor.

In this chapter, we address the issues mentioned above in the proposed method for content based medical image retrieval (CBMIR). The use of clustering enables retrieval of relevant images from the large database. We use a dictionary learning based clustering algorithm, namely K-SVD algorithm [59], to group the images in medical databases. This clustering technique improves the retrieval speed and search results. The selection of features for adequately representing the class specific information is an important process in CBIR. For facilitating this, an image is divided into four sub-images of equal size. In addition, we consider another sub-image which is centered on the image of interest and is of the same size as the other four sub-images because in most of the medical images the subject is in the center. Then each sub-image is partitioned into concentric circular regions around the center. The mean and variance of pixel intensities in each concentric circular region are considered as component of the feature vector.

Sparse representation received a lot of attention from the signal and image processing communities. Sparse coding involves the representation of an image as a linear combination of some atoms in a dictionary [57]. It is a powerful tool for efficiently processing data in nontraditional ways. This is mainly due to the fact that signals and images of interest is sparsest in some dictionary, which may be identified based on the properties of signals at hand. Of late, the dictionaries learnt from the data were found to be useful for several applications. Several interesting dictionary learning methods like K-SVD and method of optimal directions (MOD) [60] were developed to provide each member of database with sparse representation. The dictionary based methods rely on the premise that two signals belonging to the same cluster have decomposition in terms of similar atoms (columns) of a dictionary. Making use of this property, an input query is matched with the appropriate cluster.

In this chapter, we propose a content based medical image retrieval (CBMIR) algorithm using dictionary learning approach. We demonstrate the usefulness of our approach on image retrieval in medical applications (IRMA) database [63]. For a given M, the number of clusters, M dictionaries are designed to represent the clusters. Every image in the database is associated with a dictionary based on the sparsity criterion. Given a query image, we once again invoke the concept of sparsity to identify appropriate cluster, wherein we search for relevant images.

The rest of the chapter is organized as follows. Section 7.1 gives brief account of dictionary learning and the survey of related work. Section 7.2 presents the proposed content based medical image retrieval using dictionary learning method. Experiments of CBMIR application are discussed in detail in section 7.3. Finally, section 7.4 concludes this chapter.

7.1 DICTIONARY LEARNING

Given a set of vectors $\{v_i\}_{i=1}^{n}$, the K-SVD based dictionary learning method finds the dictionary D by solving the following optimization problem:

$$(\hat{D}, \hat{\Phi}) = \arg\min_{D,\Phi} \| V - D\Phi\|_{\mathbf{F}}^{\mathbf{2}} \text{ subject to } \|\gamma_i\|_{\mathbf{0}} \leq T_{\mathbf{0}} \forall i, \tag{7.1}$$

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. Φ represents sparse representation vector. Here, $\|A\|_F$ denotes the Frobenius norm which is defined as $\|A\|_F = \sqrt{\sum_{\mathbf{ij}} {A^2}_{ij}}$. The K-SVD algorithm alternates between sparse coding and dictionary update steps. Various efficient pursuit algorithms were proposed in the literature for sparse coding [60], [65]. The simplest one among all is the orthogonal matching pursuit (OMP) algorithm [65]. In sparse coding step, dictionary D is fixed and representation vectors γ_i are identified for each example y_i. Then, the dictionary is updated atom by atom in an efficient way.

7.2 CBMIR USING DICTIONARY LEARNING

In this section, we propose a method for clustering data using dictionary learning. The present work is inspired by the ideas embedded in [24] and differs from it as follows :

- The sparsity seeking dictionary learning approaches typically exploit the framework of under-determined setting and hence, work on some implicit assumptions on the database. In applications, nevertheless, one often encounters databases which are not so big that the sparsity-promoting under-determined framework could not efficiently be deployed. We come to this point in our simulation work.
- When not using labelled data (as is the case with present work), one may not have enough members in a cluster, which prevents the applicability of K-SVD.
- As Radon transform is $O(N^2 logN)$ procedure, the present approach avoids using it. This, of course, results in some computational savings.

The problems stated above could be addressed by down sampling the images or by projecting them to lower dimensional spaces. Instead, the present work extracts a small set of features that describe the images well for CBMIR.

7.2.1 Feature extraction

Two types of feature extraction methods are considered to represent the content of medical images. In the first feature extraction method, an image is partitioned into concentric circular regions of equal area for rotation invariant representation which is shown in Fig. 7.1.

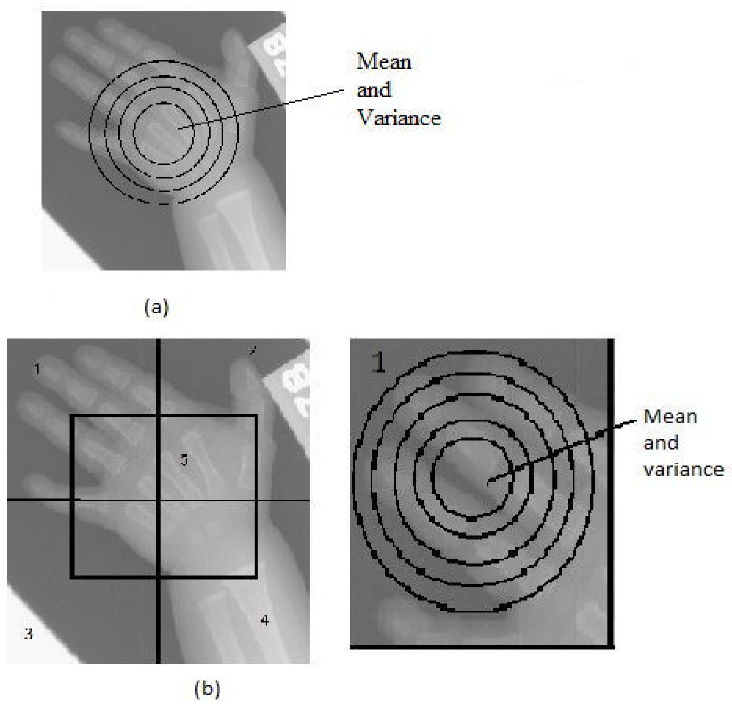

Fig. 7.1: Feature extraction.(a) Image is partitioned into concentric circular regions of equal area. (b) Image is divided into sub-images and partitioned into concentric circular regions of equal area.

The mean and variance of pixel intensity in a circular region become components of the feature vector which are defined as follows :

$$m = \frac{1}{P}\sum_{k=1}^{P}(y_k), \tag{7.2}$$

$$S = \sum_{k=1}^{P}(y_k - m)(y_k - m)^t, \tag{7.3}$$

where P is the number of pixels in each region, m is the mean of intensity of pixel values and S is the variance of intensity of pixel values in each region. This approach accomplishes the rotation invariant representation of the contents of an image.

In the second feature extraction method, an image is divided into four blocks resulting in four sub-images shown in Fig. 7.1(b). Also, another sub-image which is of same block size as other four sub-images is considered in order to capture the rich information available at the center of medical images. Each sub-image is partitioned into concentric circular regions of equal area from which the mean and variance of pixel intensity values are computed. This feature extraction method is more suitable for medical image databases because of the rich information of medical images available at the center of images.

7.2.2 Proposed Method

In this section, an approach to content based medical image retrieval technique using dictionary learning is proposed. The feature vector consisting of mean and variance of pixel intensity values are extracted from the images in the database. Initial clusters are formed by applying K-means clustering algorithm on the extracted features. Then, a dictionary is generated for each cluster using K-SVD method. A new cluster is created for each dictionary, by assigning the images that are sparsely represented by the dictionary using orthogonal matching pursuit (OMP) algorithm. The K-SVD algorithm is again used to generate the dictionaries for new clusters. The updated dictionaries are then used to generate the clusters using OMP algorithm and generated clusters are then used to update the dictionaries using K-SVD iteratively, until clusters converge. Given query image is matched with the existing dictionaries to identify the dictionary with the sparest representation using OMP algorithm. The images in the cluster associated with this dictionary are compared using a similarity measure to retrieve images similar to the query image. The entire process of proposed content based medical image retrieval is summarized in algorithm 3.

Let $\{y_j\}_{j=1}^{M}$ be the database of images represented as vectors. Suppose N is the number of clusters. Define $D = [D_{\mathbf{1}}, \ldots, D_N]$, as the concatenation of dictionaries corresponding to N clusters. Let C_i be the matrix containing images as columns corresponding to the i^{th} cluster. Then, the proposed method may be summarized as follows:

- *Cluster assignment*: The cluster assignment begins with arbitrary dictionaries $D = [D_{\mathbf{1}}, \ldots, D_N]$, where N is the number of clusters. Our proposed method considers obtaining the sparsest representation of y_j in an appropriate dictionary $D_{\hat{i}}$ from:

$$\begin{aligned} \alpha^j &= \arg\min_\omega \|y_j - D\omega\|_{\mathbf{2}}^{\mathbf{2}} \quad \text{subject to } \|\omega\|_{\mathbf{0}} \leq T_{\mathbf{0}}, \\ \hat{i} &= \arg\min_i \|y_j - D\delta_i(\alpha^j)\|_{\mathbf{2}}^{\mathbf{2}} \;\; j = \mathbf{1}, \cdots, M, \end{aligned} \tag{7.4}$$

  where δ_i is a characteristic function that selects the coefficients and ω is a sparsity matrix. Then, $y_{\mathbf{j}}$ is assigned to $C_{\hat{\mathbf{i}}}$ associated with the i^{th} dictionary.

- *Dictionary update*: From the initial clusters $C_{\mathbf{1}}, C_{\mathbf{2}}, \ldots, C_N$, the dictionaries $D_{\mathbf{i}}$ are updated by using the K-SVD approach described in Eq. (1). Then, the new dictionaries are computed as:

$$(\hat{D}_{\mathbf{i}}, \hat{\Phi}_i) = \arg\min_{D_i, \Phi_i} \|C_i - D_{\mathbf{i}}\Phi_i\|_F^{\mathbf{2}} \;\; \text{subject to } \; \|\gamma_i\|_{\mathbf{0}} \leq T_{\mathbf{0}} \; \forall i,$$

  satisfying $C_i = \hat{D}_i \hat{\Phi}_i, \quad i = \mathbf{1}, \mathbf{2}, \ldots, N.$

The cluster assignment and dictionary update steps are repeated till there is no significant change in the clusters C_i. The above mentioned clustering procedure can be summarized as the optimization problem :

$$\min_{\{D_i\},\{C_i\}} \sum_{j=1}^{M} \sum_{x \in C_i} \min_j \|y_j - D\delta_i(\alpha)\|_{\mathbf{2}}^{\mathbf{2}} + \gamma\|\alpha\|_{\mathbf{1}}, \tag{7.5}$$

where $\gamma > 0$. The above two step process of clustering and dictionary update is repeated till the convergence of clusters. Given a query image x_q, we find the cluster that is closest to the query image by identifying the corresponding dictionary admitting representation to x_q. After identifying the most relevant cluster, we find the relevant images with in the cluster using a similarity metric. The related images (search results) within the cluster are identified based on the distance criterion. To evaluate similarity between images based on the selected features, an appropriate similarity/dissimilarity metric needs to be chosen. A large class of similarity measures are used in the literature [109]. In this chapter, we use three type of similarity metrics, namely, Euclidean distance (ED), Mahalanobis distance (MD) and cross correlation (CC). The proposed algorithm is summarized as follows:

Algorithm 3 : Summary of the proposed CBMIR procedure

1. Extract features from the medical image database.
2. Apply K-means clustering algorithm on the extracted features to generate initial clusters.
3. Generate the dictionary for each cluster using K-SVD method.
4. Create new cluster for each dictionary, by assigning the images that are sparsely represented by it.
5. Repeat steps 3 and 4 till the clusters are converged.
6. For the query image q, search for relevant images in $C_{\hat{i}}$, where $D_{\hat{i}}$ provides sparsest representation to q.

7.3 EXPERIMENTAL RESULTS

The performance of the content based medical image retrieval task is measured in terms of recall $R = N_c/N_m$ and precision $P = N_c \ / \ (N_c + N_f)$ where N_m is the total number of actual (or similar) images, N_c is the number of images detected correctly, and N_f is the number of false alarms. A good performance requires both recall and precision to be high, that is, close to unity. Recall is the portion of total relevant images retrieved where as precision indicates the capability to retrieve relevant images. A compromise between recall and precision is obtained by using a measure combining both as, $F_1 = \frac{2\times(R\times P)}{R+P}$. Ideally, F_1 should be close to unity.

Given some of retrieved images, the average retrieval performance is defined as the average number of relevant images retrieved over all query images of a particular class. We compare the performance of proposed method with that of CBMIR obtained by K-means and fuzzy C-means clustering algorithms on the same image database. Experimental results are evaluated on proposed, K-means and fuzzy C-means clustering procedures using two different types of feature extraction methods on the same image database. The performance is measured on IRMA database and the results are shown in Tables 7.1 - 7.6.

7.3.1 Database Description and Results

Majority of medical images are generally gray scale images such as X-ray, CT etc. The ImageCLEF medical image database is made available by IRMA group from the University Hospital of Aachen Germany. The main goal of ImageCLEF is to create a standard environment for the evaluation and improvement of medical image retrieval from heterogeneous collections containing images as well as text. For the details on the database and the ImageCLEF benchmark evaluation for the medical annotation task one may refer to IRMA website [63].

In the IRMA database considered for CBMIR application, where each image is of size 120×120 pixels. For evaluating rotation invariant based CBIR, 2600 sample images of skull, breast, chest, hand etc. are selected. The database members when considered in matrix form as columns results in a matrix of size $(120)^2 \times 2600$. This matrix being tall and slim may not in general provide sparse representation to q. Consequently, to bring CBMIR problem into the rich theory of compressed sensing, which is based on the undetermined setting, one needs to generate feature vectors of database members. In the first feature extraction method (FE-I), each image is partitioned into 17 concentric circular regions, such that each circular region has the same number of pixels as the other region. The mean and variance of these circular regions are used to design the feature vector. So, the size of each feature vector is 34×1 (due to 17 means and 17 variances) for one image. In the second feature extraction method (FE-II), image is partitioned into five sub-images and each sub-image is partitioned into 4 concentric circular regions, such that each circular region has the same number of pixels as the other regions. The mean and variance of pixel intensity in a circular region become components of the feature vector and size of each feature vector is 40×1 (due to 4 means and 4 variances from each of 5 sub images). This procedure is applied to all the database members and 14 more images are used for testing. The performance of the proposed method is evaluated with three different cluster sizes of 3, 4 and 5 (N=3, N=4 and N=5) with dictionary size of 65 are shown in Tables 7.1, 7.2, 7.3, 7.4, 7.5, and 7.6. This size of dictionary was chosen to show the retrieval results because it can be seen from Table 7.9 that the precision and recall obtained for dictionary size of 65 is better than all the other dictionary sizes which were considered for evaluation. Experiments were also carried out with various values of the residual ϵ and it was found that $\epsilon = 0.005$ gave the best retrieval performance. In all the tables presented in this section, the value of ϵ has been considered as 0.005.

Table 7.1, 7.3 and 7.5 represent the average precision and recall for N being 3, 4 and 5 using the proposed method, fuzzy C-means and K-means clustering methods and using first feature extraction method. Table 7.2, 7.4 and 7.6 represent the average precision and recall for N being 3, 4 and 5 using the proposed method, fuzzy C-means and K-means clustering methods and using second feature extraction method. The performance in Table 7.1 was computed against the top 10 most accurately retrieved images for each test image using first feature extraction method and Euclidean distance as similarity measure. Through proposed method, for 3 clusters, the best performance of 92.1% precision and 79.6% recall was obtained. Similarly, the results for 4 clusters gave the best performance of 90.7% precision and 78.8% recall, and for 5 clusters the best performance of 87.8% precision and 74.7% recall was obtained. The best performance obtained using the fuzzy C-means clustering is 67.8% precision and 68.2% recall. In other cases, the performance of fuzzy C-means clustering algorithm was less accurate. The K-means clustering algorithm resulted in the performance of 62.1% precision and 32.6% recall. In other cases, the performance of K-means clustering algorithm was further less.

Table 7.1: Performance measure (%) of the proposed, fuzzy C-means and K-means clustering methods obtained with the first feature extraction method and the Euclidean distance as similarity measure.

Query	Proposed Method-I			Fuzzy C-Means-I			K-Means Clustering-I		
Images/Clusters	N=3	N=4	N=5	N=3	N=4	N=5	N=3	N=4	N=5
1.png	100	90	100	0	0	50	50	0	45
2.png	80	90	100	10	0	40	70	25	10
3.png	70	70	100	100	40	60	60	20	30
4.png	100	100	100	100	50	60	80	100	55
5.png	100	100	80	90	50	50	60	85	45
6.png	100	100	100	80	60	40	80	100	100
7.png	90	100	80	50	100	20	50	0	10
8.png	100	80	100	100	50	0	50	0	60
9.png	70	80	80	80	40	60	70	20	30
10.png	100	100	80	90	50	50	70	80	50
11.png	100	90	100	20	10	50	50	0	50
12.png	90	100	60	90	50	40	70	90	80
13.png	100	80	80	90	40	0	50	0	10
14.png	90	90	90	50	60	20	60	10	50
precision (%)	**92.1**	90.7	87.8	**67.8**	42.8	38.5	**62.1**	37.8	41.4
recall(%)	**79.6**	78.8	74.7	**68.2**	62	60.7	**32.6**	35.5	60.2

Table 7.2 shows the performance of evaluation obtained with the second feature extraction method and Euclidean distance as similarity measure. Through the second method of feature extraction, the best performance of 97.14% precision and 80.1% recall were obtained. The best performance using the fuzzy C-means clustering is 74.8% precision and 60% recall and K-means clustering is 62.6% precision and 48% recall. From our simulation results, it can be concluded that the 2^{nd} feature extraction method gives better performance than the 1^{st} method.

Table 7.3 and 7.4 represent the average precision and recall of proposed, fuzzy C-means and K-means clustering methods using first and second feature extraction methods, respectively, using cross correlation as similarity measure. It can be inferred from Table 7.3 and 7.4 that the proposed method using second feature extraction method (93.7% precision and 83.2% recall) gives better performance than the fuzzy C-means and K-means clustering algorithms.

Table 7.5 and 7.6 represent the average precision and recall of proposed, fuzzy C-means and K-means clustering methods using first and second feature extraction methods, respectively, by using Mahalanobis distance as similarity measure. From the results in Table 7.5 and 7.6, it can be concluded that the proposed method performs better (62.8% precision and 47.2% recall) than fuzzy C-means and K-means clustering methods.

Table 7.7 and 7.8 represent the average precision and recall of proposed methods with increasing and decreasing number of concentric circular regions for the first feature extraction method and Euclidean distance as similarity measure. The results obtained in the table 7.7 portray that decreasing the number of concentric regions (<17) for feature extraction yields less performance. This is because of the reduction in feature vector size and the creation of non-optimal dictionaries for clustering. Moreover, in table 7.8, it can be seen that increasing the number of concentric regions (>17) allows artifacts from X-ray images near the boundaries of the images to contribute to the feature vector, thereby reducing performance.

Table 7.9 depicts the results of the proposed method with various dictionary and

Table 7.2: Performance measure (%) of the proposed, fuzzy C-means and K-means clustering methods using second feature extraction method and Euclidean distance as similarity measure.

Query	Proposed Method-II			Fuzzy C-Means-II			K-Means Clustering-II		
Images/Clusters	N=3	N=4	N=5	N=3	N=4	N=5	N=3	N=4	N=5
1.png	100	100	100	70	60	60	30	30	0
2.png	100	100	90	80	70	60	40	10	20
3.png	90	70	100	80	70	50	57.1	50	50
4.png	100	80	100	90	50	40	50	40	50
5.png	100	100	90	90	70	20	70	40	100
6.png	90	90	80	60	40	80	80	20	70
7.png	100	100	100	50	70	70	100	100	50
8.png	100	100	80	25	0	0	80	40	0
9.png	90	88.8	90	85.7	83.3	87.5	100	90	43
10.png	100	90	90	90	90	80	90	60	32
11.png	100	77.7	60	83.3	50	16.6	100	60	0
12.png	100	100	90	77.7	80	80	80	40	50
13.png	90	90	40	83.3	60	100	0	40	30
14.png	100	100	100	83.3	20	0	0	40	20
precision(%)	**97.14**	91.8	86.4	**74.8**	54	48	**62.6**	45	43
recall(%)	**80.1**	83.2	76.9	**60**	58.2	68	**48**	38	32

cluster sizes. It can be noted that the dictionary size of 65 yields the best performance. This can be attributed to the fact that redundancy of information increases with increasing dictionary sizes due to fewer number of training images. Higher dictionary sizes can be accommodated by increasing the number of training images.

In Fig. 7.4, on every row, the first element represents the query image while the other represent those retrieved by the proposed method with Euclidean distance as

Table 7.3: Performance measure (%) of the proposed, fuzzy C-means and K-means clustering methods using first feature extraction method and cross correlation as similarity measure.

Query	Proposed Method-I			Fuzzy C-Means-I			K-Means Clustering-I		
Images/Clusters	N=3	N=4	N=5	N=3	N=4	N=5	N=3	N=4	N=5
1.png	90	60	50	20	30	60	68.4	0	60
2.png	70	70	50	20	30	50	47.3	0	44.4
3.png	100	90	80	100	80	80	65	35	0
4.png	90	60	80	100	50	90	33.3	55	50
5.png	70	80	90	100	50	50	33.3	50	50
6.png	100	100	100	80	60	60	91.6	25	55
7.png	90	100	100	100	100	30	100	25	0
8.png	40	30	20	50	30	0	37.5	0	0
9.png	90	80	80	90	70	70	60	30	0
10.png	80	90	80	100	50	50	40	50	50
11.png	80	60	50	20	30	50	63.2	0	70
12.png	100	80	60	70	50	50	80	30	50
13.png	60	40	20	40	30	10	32.4	10	0
14.png	50	50	25	50	20	0	30	0	20
precision(%)	**79.2**	70.7	68.5	**67.1**	48.5	46.4	**55.8**	22.1	32.1
recall(%)	**65.7**	65	68	**69.8**	75	71.4	**57.4**	47.4	47

similarity metric.

Fig. 7.3 shows the average precision and recall of the proposed, fuzzy C-means and K-means clustering methods using first and second feature extraction methods with three different similarity measures. Among these, the proposed method has better performance (97.1% precision and 80.1% recall) with the Euclidean distance based similarity measure as shown in Fig. 7.3.

Table 7.4: Performance measure (%) of the proposed, fuzzy C-means and K-means clustering methods using second feature extraction method and cross correlation as similarity measure.

Query	Proposed Method-II			Fuzzy C-Means-II			K-Means Clustering-II		
Images/Clusters	N=3	N=4	N=5	N=3	N=4	N=5	N=3	N=4	N=5
1.png	90	100	90	70	90	70	70	70	40
2.png	100	100	90	80	90	80	80	60	70
3.png	97	90	94.1	60	40	50	57.1	50	67
4.png	90	90	100	50	50	30	90	67	63
5.png	100	100	50	50	90	30	90	0	70
6.png	90	90	100	90	50	100	90	78	90
7.png	100	80	80	100	50	50	0	50	0
8.png	95	80	100	100	0	20	83.3	50	0
9.png	100	80	100	87.5	40	50	50	10	90
10.png	100	80	100	90	50	40	70	90	80
11.png	90	60	80	83.3	90	70	70	70	70
12.png	90	100	100	88.9	50	90	100	70	50
13.png	80	70	20	83.3	10	30	50	40	50
14.png	90	100	100	83.3	20	20	60	40	50
precision(%)	**93.7**	87.1	86	**79.6**	51.4	52.1	**68.6**	53.2	56.4
recall(%)	**83.2**	80.1	76.8	**69.6**	50.1	64.8	**71.9**	50.4	49.8

Fig. 7.4 shows comparison between retrieval time and feature vector size for different cluster sizes. This plot indicates that increasing the feature vector size contributes to an increase in retrieval time as expected.

Table 7.5: Performance measures (%) of the proposed, fuzzy C-means and K-Means clustering methods using first feature extraction method and Mahalanobis distance as similarity measure.

Query	Proposed Method-I			Fuzzy C-Means-I			K-Means Clustering -I		
Images/Clusters	N=3	N=4	N=5	N=3	N=4	N=5	N=3	N=4	N=5
1.png	60	70	100	0	0	50	31.5	0	35
2.png	40	30	60	0	0	33.3	26.5	0	05
3.png	40	50	50	60	20	28.5	30	10	05
4.png	30	40	60	70	50	50	18	50	45
5.png	30	70	30	70	50	50	18	55	45
6.png	50	70	50	20	20	0	16.6	05	30
7.png	90	70	50	50	80	0	52.6	03	0
8.png	100	90	90	70	30	0	75	0	30
9.png	50	50	50	60	20	30	30	10	0
10.png	30	40	50	70	60	50	20	50	30
11.png	60	60	90	0	10	60	30	0	30
12.png	50	70	40	40	30	0	20	30	40
13.png	90	80	80	30	20	10	70	20	40
14.png	80	90	70	40	30	10	40	0	45
precision(%)	57.1	60.7	**62.1**	**41.4**	30	26.5	**34.1**	19.2	27
recall(%)	56.4	54.6	**60**	**42.9**	49.3	32	**41.5**	30.7	39.5

7.4 SUMMARY

In this chapter, a novel dictionary learning based clustering method for content based medical image retrieval is proposed. Mean and variance of pixel intensity values are used as feature vector and K-SVD method is used to generate dictionaries for each cluster. The performance of the proposed method is evaluated using IRMA database. The first feature extraction (FE1) method aims at providing rotation invariant CBIR,

Table 7.6: Performance measure (%) of the proposed, fuzzy C-means and K-Means clustering method using second feature extraction method and Mahalanobis distance as similarity measure.

Query	Proposed Method-II			Fuzzy C-Means-II			K-Means Clustering-II		
Images/Clusters	N=3	N=4	N=5	N=3	N=4	N=5	N=3	N=4	N=5
1.png	40	70	70	0	30	30	30	30	0
2.png	30	60	60	50	20	20	20	10	30
3.png	100	40	80	29	43	38	20	19	35.2
4.png	40	30	50	38	20	33.3	33.3	17	62.2
5.png	40	60	30	50	20	0	33.3	20	62.2
6.png	50	60	70	50	40	40	60	20	59
7.png	80	83.3	70	71.4	40	75	100	0	0
8.png	90	90	50	88	20	0	67	0	0
9.png	10	20	80	14.2	16.6	50	10	20	0
10.png	30	20	50	10	10	0	30	20	50
11.png	60	30	30	50	30	30	30	30	30
12.png	30	30	60	33.3	20	40	70	20	20
13.png	50	50	90	66.6	0	0	30	10	20
14.png	50	80	90	66.6	16.6	0	30	10	60
precision(%)	50	53.8	**62.8**	**44**	23.3	25.4	**40.2**	16.1	30.6
recall(%)	49.3	49.4	**47.2**	**38.7**	29.1	27	**55.6**	24.7	31.6

while the second (FE2) method aims at taking into consideration the rich information available at the center. The experimental results show that FE2 method gives superior performance compared to FE1. The extensive experimental work is carried out with different cluster sizes, with different number of concentric circular regions, different column sizes for dictionaries, different similarity metrics and with different initial clustering algorithms. It is observed that when N is 3, number of concentric

Table 7.7: Performance measure (%) of the proposed method with decreasing feature vector size (No.of concentric circles is 7) using Euclidean distance, cross correlation and Mahalanobis distance as similarity measure.

Query	Euclidean distance			Cross correlation			Mahalanobis distance		
Images/Clusters	N=3	N=4	N=5	N=3	N=4	N=5	N=3	N=4	N=5
1.png	80	80	30	50	30	30	30	60	10
2.png	100	90	30	90	40	30	30	30	20
3.png	90	90	90	100	90	90	70	50	50
4.png	100	100	70	90	90	80	30	30	50
5.png	100	90	90	100	100	80	50	50	20
6.png	40	90	70	70	100	90	20	70	30
7.png	100	80	80	100	90	60	40	20	20
8.png	90	100	70	40	20	10	70	90	70
9.png	100	100	100	100	100	90	10	30	30
10.png	60	100	80	80	90	100	10	70	60
11.png	90	90	90	70	40	80	10	60	40
12.png	30	50	60	40	50	70	30	30	20
13.png	60	100	80	20	70	40	30	50	50
14.png	90	100	70	60	60	10	90	90	80
precision(%)	80.7	**90**	72.1	**72.1**	69.2	61.4	37.1	**52.1**	39.2
recall(%)	67.9	**80.4**	54.9	**66.3**	61.8	50.2	35	**53.4**	36.2

circles is 17, one achieves better F_1 performance of 87.5% with Euclidean distance as similarity metric. As medical images come with different scaling factors, our future work aims at addressing scale invariance as well in CBMIR.

Table 7.8: Performance measure (%) of the proposed method with increasing feature vector size (No.of concentric circles=23) using Euclidean distance, cross correlation and Mahalanobis distance as similarity measure.

Query	Euclidean distance			Cross correlation			Mahalanobis distance		
Images/Clusters	N=3	N=4	N=5	N=3	N=4	N=5	N=3	N=4	N=5
1.png	100	90	80	90	70	60	50	40	20
2.png	70	80	90	80	100	80	40	20	30
3.png	90	90	70	100	100	100	60	40	50
4.png	80	100	100	90	100	90	40	30	20
5.png	100	100	90	100	100	100	30	30	40
6.png	90	90	70	100	90	90	20	50	30
7.png	100	90	90	100	100	100	60	50	40
8.png	100	90	100	20	20	10	60	70	80
9.png	100	100	90	100	100	100	40	20	50
10.png	70	70	70	80	70	100	50	40	30
11.png	80	90	80	80	100	80	20	20	40
12.png	90	90	80	90	90	80	40	70	60
13.png	100	80	100	80	70	90	60	60	80
14.png	100	90	90	60	40	30	90	80	90
precision(%)	**91.4**	89.2	85.7	**83.5**	82	79.2	**47.8**	44.2	47.1
recall(%)	**71.6**	76.3	73.8	**68.9**	71.9	68.7	**47.4**	45.5	43.3

Table 7.9: Performance measure (%) of the proposed method with different dictionary sizes.

	Proposed method-I			Proposed method-II		
Column size of D_i/Clusters	N=3	N=4	N=5	N=3	N=4	N=5
60	89	82.3	82	93	91	86.4
65	**92.1**	90	**87.8**	**97.1**	**91.8**	**93**
70	86.2	**90.7**	82	91.2	89.3	89
75	88.1	82	80.2	92.3	90	90
80	86.4	80.4	86	93.2	88.6	91.3
85	88.3	78.4	84	95	89.1	88.6

Fig. 7.2: Some of the retrieved images, first column contains the query images and remaining columns correspond to the retrieved images.

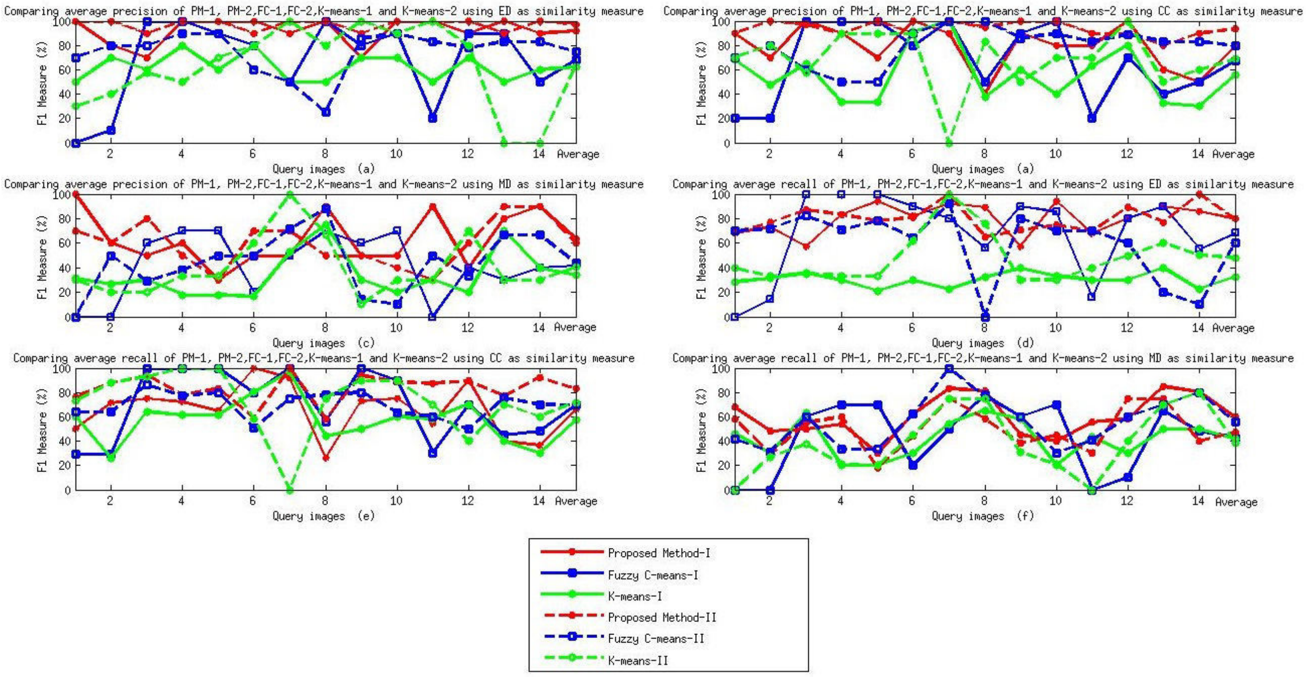

Fig. 7.3: Comparision of average precision and recall of proposed, fuzzy C-means and K-means clustering methods using first (I) and second (II) feature extraction methods with three different distance similarity measures. (a) Highest precision recorded $(\%)$ using Euclidean distance as similarity measure. (b) Highest precision recorded $(\%)$ using cross correlation as similarity measure. (c) Highest precision recorded $(\%)$ using Mahalanobis distance as the similarity measure.(d) Highest recall recorded$(\%)$ using Euclidean distance as similarity measure. (e) Highest recall recorded $(\%)$ using cross-correlation as the similarity measure. (f) Highest recall recorded $(\%)$ using Mahalanobis distance as similarity measure. Here, x-axis refers to different query images and the y-axis refers to F_1 performance.

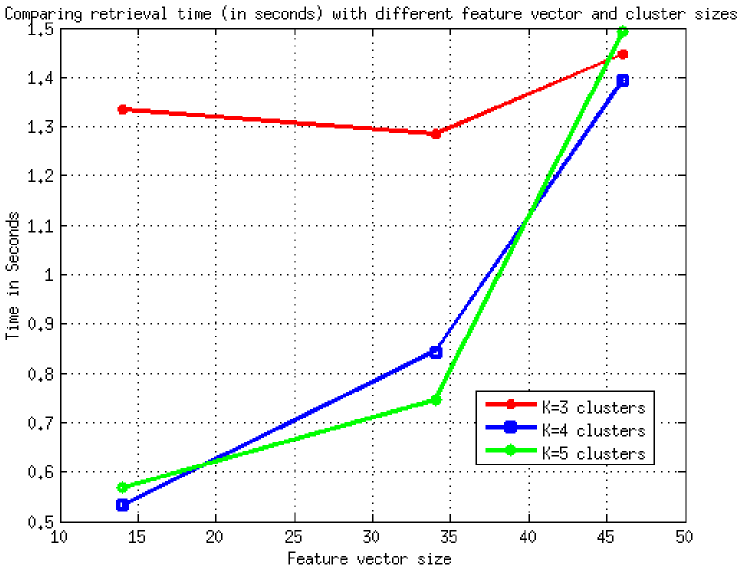

Fig. 7.4: Comparison between retrieval time and feature vector size for different cluster sizes.

CHAPTER 8

CONCLUSIONS

8.1 SUMMARY AND CONCLUSIONS

In this thesis, new approaches were proposed to address some issues in classification and retrieval of medical data. Image classification and retrieval which is concerned with effectively and efficiently accessing similar type of images from large image collections, has become more interesting and more challenging as the medical datasets have grown over the years. The medical data classification is an important task in the context of content based medical image retrieval and present some challenging problems. The key issues in medical data classification are representation of class specific information using suitable features, and developing methods to capture information present in the features. In this thesis, new methods for grouping of medical images into different classes based on sparsest representation and dictionary learning were proposed. The sparsity seeking dictionary learning approaches typically exploit the framework of under-determined setting and hence work on some implicit assumptions on the database. The methods proposed here vastly improve the speed and accuracy of retrieved images.

The existing medical image search and retrieval techniques are not very efficient in terms of time and accuracy of search results because most of the existing tools for searching medical images use text based image retrieval techniques. Text based image classification suffers from some serious limitations, namely, when the size of image collection gets increasingly large, manually annotating each image is very difficult. Also, different people may give different annotations to images with similar visual content. Improving the classification accuracy and reducing the retrieval time are important issues in medical image classification and retreival. Content based image classification and retrieval approaches were proposed to overcome the limitations of text based image classification and retrieval. Digital image retrieval techniques are crucial in the emerging field of medical image databases for clinical decision making process.

An algorithm for classification of medical images based on edge features extracted from various body parts using ℓ_1 -lasso sparse representation and on-line dictionary learning (ODL) was proposed. Edge information was extracted from an image by dividing the image into patches and each patch into concentric circular regions to provide discriminative information useful for classification of medical images. The ability of on-line dictionary learning to achieve sparse representation of an image was exploited to develop dictionaries for each class using edge-based features.

In most medical imaging systems, the same body part was captured from different orientations and magnification by the same sensor. Coming up with a rotation and invariant classification and retrieval system was a real challenge. The mean and variance over concentric circular regions in an image were calculated and used as features for providing a rotation invariant image retrieval scheme.

Medical images are captured by different sensors (modalities). Capturing images of various modalities suffers from significant contrast variation between the images of the same organ or body part. Due to this large variation, existing image classification and retrieval algorithms do not perform well for different modality images. Our proposal to address this issue was based on multi-scale wavelet representation using dictionary learning. Wavelet features extracted from an image provide discriminative information useful for classification of medical images. Multi-scale wavelets were employed to compensate for the varying scale of intensity in the images captured by the aforementioned sources.

In addition, most of the medical datasets pose the problem of data imbalance i.e. unequally distributed training samples among all the classes, which gives rise to poor classification performance results with standard single classifiers. The proposed method uses a multi-level classifier to combine correctly classified examples in the first level with the training data and supply them as input to the next level classifier. So, if there is any data imbalance i.e. less number of training samples, it can be alleviated by this method.

Adaptive dictionary learning based classification is used to classify normal and abnormal heartbeat patterns from an ECG database. A relevant application is automated detection of heart diseases based on abnormal heartbeat patterns.

The problem of the search for relevant information in large medical image databases in content based medical image retrieval was addressed. This problem deals with the retrieval of rotation invariant based similar images and improves the accuracy of similar retrieval images with the help of clustering technique. We also proposed a method for clustering of medical data based on sparse representation using dictionary learning. The basic idea is to group similar images into clusters that are sparsely represented by the dictionaries and simultaneously learn dictionaries from the clusters using K-SVD.

8.2 CONTRIBUTIONS OF THE WORK

The main contributions of this thesis are summarized as follows:

- Classification of X-ray images using on-line dictionary learning based on the sparse representation of edge-based features was proposed. This method was used to classify various body parts present in medical images. Edge-based features are used to classify the medical images since different body parts are distinctly characterized by edge information.
- Multi-level classification framework involving on-line dictionary learning and support vector machine for medical data classification has been proposed. A multi- level classifier that combines correctly classified examples in the first level with the training data and supplies them as input to the next level classifier has been devised. The ability of multi-level classification approach is more suitable for imbalanced medical datasets.

- Classification of medical images based on acquisition source (modality) represented by multi-scale wavelets using on-line dictionary learning has been proposed. Wavelet features extracted from an image provide discriminative information useful for classification of medical images. Multi-scale wavelets are employed to compensate for the varying scale of intensity in the images captured by the aforementioned sources.
- Adaptive dictionary learning based classification approach for detection of abnormality in ECG signals has been proposed.
- A new clustering method was proposed for content based medical image retrieval based on sparse representation and dictionary learning. The mean and variance over concentric circular regions in an image are calculated and used as features for providing a rotation invariant image retrieval scheme. The methods proposed here vastly improve accuracy of retrieved images and reduce the search time.

8.3 DIRECTIONS FOR FUTURE RESEARCH

- The proposed method for clustering of medical images for content based medical image retrieval assumes an under-determined setting i.e. the number of instances are much less than the number of attributes. So, improvements have to be made to make it work in an over-determined system as well.
- In the case of clustering, there is no guarantee that a cluster will have enough members and consequently, dictionary learning cannot be effectively applied. The main problem would be to reform the classification problem in under-determined setting.
- Medical images come with different transformations (such as scaling), future work aims at addressing the invariant CBMIR with respective to other transformations.

APPENDIX A

SPARSE REPRESENTATIONS

Suppose that there are K medical image classes, and each class has a set of N medical images. Let a d-dimensional feature vector be extracted from each medical image. Let A_k be a $d \times N$ matrix of feature vectors of the k^{th} class, where the column $a_{kn} = [a_{kn1}a_{kn2}...a_{knd}]^T$ denotes the d-dimensional feature vector of the n^{th} medical image belonging to the k^{th} class.

$$A_k = [a_{k1}a_{k2}...a_{kn}...a_{kN}] \in R^{d \times N} \tag{A.1}$$

An medical image dictionary A can be defined as follows:

$$A = [A_1A_2...A_k...A_K] \in R^{d \times KN} \tag{A.2}$$

where K represents some of the feature vectors from K different medical image classes. Let $y \in R^d$ be an observed feature vector extracted from a test medical image. The y can be be expressed as a linear weighted sum of columns of medical image dictionary A as

$$y = \sum_{k=1}^{K} \sum_{n=1}^{N} x_{kn} a_{kn} \tag{A.3}$$

where the scalar x_{kn} is the weight associated with the column a_{kn}. The above equation can also be written in the matrix form as

$$y = Ax \tag{A.4}$$

and the residual can be written as

$$r(y) = y - Ax \tag{A.5}$$

The observation vector y belongs to a particular class meaning that it is approximately comes in the linear span of the training vectors of that medical class. In other words, the coefficients of the weight vector x that does not belong to that particular medical image class are very close to zero and also x gives more sparsity with very few nonzero coefficients. The given system of linear equations in (A.4) is under-determined, since the size of the feature vector (d) is much greater than the number of feature vectors concatenated in the medical image dictionary. So it does not give unique solution, the sparsest solution can be obtained from the infinitely many solutions by solving the following optimization problem

$$min_x ||x||_0 \ subject \ to \ y = Ax \tag{A.6}$$

where $||x||_0$ is zero norm of weight vector x which mean the number of nonzero coefficients in weight vector x. There were many iterative algorithms proposed like matching pursuit (MP), and orthogonal matching pursuit (OMP) to address the above optimization problem. In the proposed medical image classification and retrieval methodology, OMP algorithm is chosen to calculate the approximate sparse weight vector x [65]. The main goal of the algorithm is to identify sparse weight vector x which gives a few nonzero coefficients. These coefficients will determine the few columns of A that participate in the representation of observation vector y. The algorithm chooses those columns in a greedy fashion. The following are the steps involved in OMP algorithm [65].

1. The sparse weight vector x is initialized with zero, ($x^0 = 0$). The initial residual is, $r^0(y) = yAx^0 = y$. The solution support is initialized with $S^0 = Support\{x^0\} = \phi$

2. Since the residual error depends on $||y||_2$, a fraction of $||y||_2$ can be used as error threshold, i.e., $\theta_0 = \lambda ||y||_2$ where $0 < \lambda < 1$. The value should not be very high

or very low. If the value is very high, it may not capture the iris class-specific characteristics. On the other hand, a low value of λ may spoil the sparsity of the weight vector x while minimizing the residual error.

3. The first iteration of the algorithm starts with $k = 1$.

4. The errors are computed for all columns of A using $\theta(c) = min_{z_c}||a_c z_c - r^{k-1}||_2^2$. Where c represents the column index and $z_c = a_c^T r^{k-1}/||a_c||_2^2$.

5. Among all the column errors, find a minimizer c_0 from $\theta(c)$ in such a way that the column should not be an element in previous solution support and $\theta(c_0) \leq (c)$. Update the solution support S^k by adding the minimizer c_0 to previous solution support S^{k-1}.

6. Based on the updated solution support S^k, compute the sparse weight vector x^k by solving the $min||y - Ax||_2^2$.

7. The residual is again computed for the current iteration using $r^k = bAx^k$.

8. If the l_2 norm for the updated residual is below the predefined error threshold θ_0. Then x^k becomes the solution. Otherwise, repeat the steps from 4, by incrementing k by 1.

A.1 LEAST ABSOLUTE SHRINKAGE AND SELECTION OPERATOR (LASSO) ALGORITHM

Another commonly used sparse representation of an algorithm called Least Absolute Shrinkage and Selection Operator (LASSO) and another termed Least Angle Regression (LARS). The LASSO is an L_1 regression technique introduced by Tibshirani (1996) and it is shrinkage and selection method for linear regression It minimizes the usual sum of squared errors, with a bound on the sum of the absolute values of the coefficients. It has connections to soft-thresholding of wavelet coefficients, forward stagewise regression, and boosting methods.

Given a matrix of signals $A = [a^1, ..., a^n] \in R^{m\times n}$ and a dictionary D in $R^{m\times p}$, depending on the input parameters, the algorithm returns a matrix of coefficients $X = [\alpha^1, ..., \alpha^n] \in R^{p\times n}$ such that for every column a of A, the corresponding column α of A is the solution of

$$\min_{\alpha\in\mathbf{R^{m\times n}}} \|x\|_1 \text{ s.t. } \|a - D\alpha\|_\mathbf{2}^\mathbf{2} \ \le \lambda, \tag{A.7}$$

A.2 ODL ALGORITHM

Assuming the training set composed of i.i.d. samples of a distribution $p(x)$, its inner loop draws one element X_t at a time, as in stochastic gradient descent, and alternates classical sparse coding steps fr computing the decomposition α_t of X_t over the dictionary D_{t-1} obtained at the previous iteration, with dictionary update steps where the new dictionary D_t is computed by minimizing over C the function

$$\hat{f}_\mathbf{t}(\mathbf{D}) \doteq \frac{\mathbf{1}}{t}\sum_{i=1}^{t}\frac{\mathbf{1}}{\mathbf{2}}\|X_i - D\alpha_i\|_\mathbf{2}^\mathbf{2} \ + \ \lambda\|\alpha_i\|_\mathbf{1}, \tag{A.8}$$

where the vectors α_i are computed. Algorithm is summarized in Algorithm 4.

Algorithm 4 : Online dictionary learning.

Input : $X \in R^m \sim p(X)$ (random variable and an algorithm to draw i.i.d samples of p), $\lambda \in R$ (regularization parameter), $D_0 \in R^{m \times k}$ (initial dictionary), T (number of iterations).

1. $A_0 \leftarrow 0, B_0 \leftarrow 0$ (rest the past information).
2. for $t = 1$ to T do
3. Draw X_t from $p(X)$.
4. Sparse coding: compute using LARS

$$\alpha_{\mathbf{t}} \doteq \arg\min_{\alpha \in \mathbf{R}^i} \frac{\mathbf{1}}{\mathbf{2}} \|X_t - D_{t-1}\alpha\|_{\mathbf{2}}^{\mathbf{2}} \; + \; \lambda \|\alpha_i\|_{\mathbf{1}}. \tag{A.9}$$

5. $A_t \leftarrow A_{t-1} + \alpha_t \alpha_t^T$.
6. $B_t \leftarrow B_{t-1} + X_t \alpha_t^T$.
7. Compute D_t using Algorithm 5, with D_{t-t} as warm restart, so that

$$(D_{\mathbf{t}}) \doteq \arg\min_{D \in C} \frac{\mathbf{1}}{t} \sum_{i=1}^{t} \frac{\mathbf{1}}{\mathbf{2}} \|X_i - D\alpha_i\|_{\mathbf{2}}^{\mathbf{2}} \; + \; \lambda \|\alpha_i\|_{\mathbf{1}}.$$

$$= \arg\min_{D \in C} \frac{1}{t} \left(\frac{1}{2} Tr(D^T D A_t) - Tr(D^T B_t)\right). \tag{A.10}$$

8. end for
9. Return D_T (learned dictionary).

Algorithm 5 : Dictionary Update.

Input : $D = [d_1, ..., d_k] \in R^{m \times k}$ (input dictionary),

$A = [a_1, ..., a_k] \in R^{k \times k} = \sum_{i=1}^{t} \alpha_i \alpha_i^T$,

$B = [b_1, ..., b_k] \in R^{m \times k} = \sum_{i=1}^{t} X_i \alpha_i^T$,

1. repeat
2. for $j = 1$ to k do
3. Update the j-th column to optimize for (A.10):

$$
\begin{aligned}
(u_j \leftarrow \frac{1}{A_{jj}}(b_j - Da_j) + d_j). \\
d_j \leftarrow \frac{1}{max(\|u_j\|_2, 1)} u_j.
\end{aligned}
\tag{A.11}
$$

4. end for
5. until convergence
6. Return D (updated dictionary).

APPENDIX B

SUPPORT VECTOR MACHINES

The support vector machine (SVM) is a linear machine pioneered by Vapnik [72]. The main idea of an SVM is to construct a hyperplane as the decision surface in such a way that the margin of separation between positive and negative examples is maximized. The notion that is central to the construction of the support vector learning algorithm is the innerproduct kernel between a support vector $\mathbf{x}_i$ and a vector $\mathbf{x}$ drawn from the input space. The support vectors constitute a small subset of the training data extracted by the support vector learning algorithm. The separation between the hyperplane and the closest data point is called the margin of separation, denoted by ρ. The goal of a support vector machine is to find a particular hyperplane for which the margin of separation ρ is maximized. Under this condition, the decision surface is referred to as the optimal hyperplane. Fig. B.1 illustrates the geometric construction of a hyperplane for two dimensional input space. The support vectors play a prominent role in the operation of this class of learning machines. In conceptual terms, the support vectors are those data points that lie closest to the decision surface, and therefore the most difficult to classify. They have a direct bearing on the optimum location of the decision surface.

The idea of an SVM is based on the following two mathematical operations [72]:

1. Nonlinear mapping of an input pattern vector onto a higher dimensional feature space that is hidden from both the input and output.

2. Construction of an optimal hyperplane for separating the patterns in the higher dimensional space obtained from operation 1.

Operation 1 is performed in accordance with Cover's theorem on the separability of patterns [72]. Consider an input space made up of nonlinearly separable patterns.

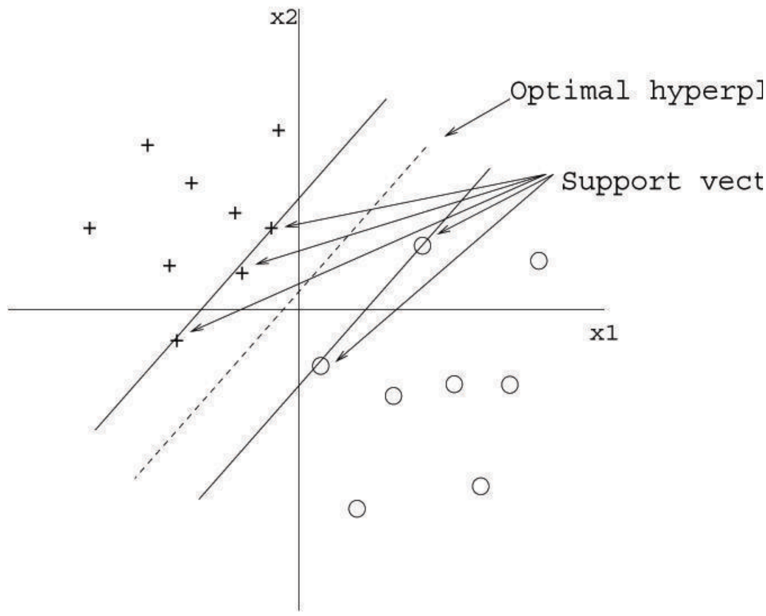

Fig. B.1: Illustration of the idea of support vectors and an optimal hyperplane for linearly separable patterns.

Cover's theorem states that such a multidimensional space may be transformed into a new feature space where the patterns are linearly separable with a high probability, provided the transformation is nonlinear, and the dimension of the feature space is high enough. These two conditions are embedded in operation 1. The separating hyperplane is defined as a linear function of the vectors drawn from the feature space. Construction of this hyperplane is performed in accordance with the principle of structural risk minimization that is rooted in Vapnik-Chervonenkis (VC) dimension theory [110]. By using an optimal separating hyperplane the VC dimension is minimized and generalization is achieved. The number of examples needed to learn a class of interest reliably is proportional to the VC dimension of that class. Thus, in order to have a less complex classification system, it is preferable to have those features which lead to lesser number of support vectors.

The optimal hyperplane is defined by:

$$\sum_{i=1}^{N_L} \alpha_i d_i K(\mathbf{x}, \mathbf{x}_i) = 0 \tag{B.12}$$

where $\{\alpha_i\}_{i=1}^{N_L}$ is the set of Lagrange multipliers, $\{d_i\}_{i=1}^{N_L}$ is the set of desired classes and $K(\mathbf{x}, \mathbf{x}_i)$ is the innerproduct kernel, and is defined by:

$$\begin{aligned} K(\mathbf{x}, \mathbf{x}_i) &= \varphi^T(\mathbf{x})\varphi(\mathbf{x}_i) \\ &= \sum_{j=0}^{m_1} \varphi_j(\mathbf{x})\varphi_j(\mathbf{x}_i), \qquad i = 1, 2, \ldots, N_L \end{aligned} \tag{B.13}$$

where $\mathbf{x}$ is a vector of dimension m drawn from the input space, and $\{\varphi_j(\mathbf{x})\}_{j=1}^{m_1}$ denotes a set of nonlinear transformations from the input space to the feature space. $\varphi_0(\mathbf{x}) = 1$, for all $\mathbf{x}$. m_1 is the dimension of the feature space. From (B.12) it is seen that the construction of the optimal hyperplane is based on the evaluation of an innerproduct kernel. The innerproduct kernel $K(\mathbf{x}, \mathbf{x}_i)$ is used to construct the optimal hyperplane in the feature space without having to consider the feature space itself in explicit form.

The design of a support vector machine involves finding an optimal hyperplane. In order to find an optimal hyperplane, it is necessary to find the optimal Lagrange multipliers which are obtained from the given training samples $\{(\mathbf{x}_i, d_i)\}_{i=1}^{N_L}$. Dimension of the feature space is determined by the number of support vectors extracted from the training data by the solution to the optimization problem (B.12).

APPENDIX C

LINEAR DISCRIMINANT ANALYSIS

Discriminant analysis method developed in 1936 by R.A. Fisher. And it is a multivariate classification method. In discriminant analysis, the mian objective is to predict class labels of individual observations based on a set of predictor variables.

The purpose of linear discriminant analysis (LDA) is to find the linear combinations of the predictor variables that gives the best possible separation between the groups of observations. Linear discriminant analysis is also known as "canonical discriminant analysis".

Given dataset there are N different groups, each assumed to have a multivariate normal distribution with mean vector($n = 1, ..., K$) and common covariance matrix . The actual mean vectors and covariance matrices are almost always unknown. With the help of maximum likelihood methods are used to estimate these parameters.

The basic method of LDA isto classify observations y_i to the group n, which minimize the within group variance i.e.,

$$n = argmin_n (y_i - \mu_n)^T \Sigma^{-1} (y_i - \mu_k) \tag{C.14}$$

Under multivariate normal assumptions, this is equivalent to finding the group that maximizes the likelihood of the observation. Generally, we can estimate prior probability using the proportion of the number of observations in each group to the total. For example, let $\pi_n = \frac{m_n}{m}$ be the proportion of group n such that $\pi_1 + ... + \pi_n$=1. Then, contrary to maximizing the likelihood value, the posterior probability is maximized. The observation relates to a particular group,

$$n = argmax_n[-\frac{1}{2}(y_i - \mu_n)^T \Sigma^{-1}(y_i - \mu_k) + log\pi_k] \tag{C.15}$$

Simplifying (C.15), the n LDA functions are

$$d_n(y) = y^T \Sigma^{-1} \mu_n - \frac{1}{2}\mu_n^T \Sigma^{-1} \mu_k + log\pi_k \tag{C.16}$$

When the assumption of common covariance matrix is not satisfied, an individual covariance matrix for each group is used.

In the binary case, two linear discriminantfunctions are built asfollows:

$$d_1(y) = y^T \Sigma^{-1} \mu_1 - \frac{1}{2}\mu_1^T \Sigma^{-1} \mu_1 + log\pi_1 \tag{C.17}$$

$$d_2(y) = y^T \Sigma^{-1} \mu_2 - \frac{1}{2}\mu_2^T \Sigma^{-1} \mu_2 + log\pi_2 \tag{C.18}$$

If $d_1(y) > d_2(y)$ the observation y will be assigned to first group, otherwise to second group. The two discriminantfunctions can also be combined i.e.,

$$d(y) = d_1(y) - d_2(y)$$

$$= y^T \Sigma^{-1}(\mu_1 - \mu_2) - \frac{1}{2}(\mu_1 + \mu_2)^T \Sigma^{-1}(\mu_1 - \mu_2) + log\frac{\pi_1}{\pi_2} \tag{C.19}$$

If $d(y) > 0$, the observation y will be assigned to group one, otherwise to group two. The last two parts in the equation (C.19) are constant given a data set. The discriminant function coefficients are $D = \Sigma^{-1}(\mu_1 - \mu_2)$. The coefficients reflect the joint contribution of the variablesto the function, thereby showing the influence of each variable in the presence of the others. The standardized coefficients $D^* = diag(\Sigma)D$ are computed by multiplying each coefficient by the standard deviation of the corresponding variables. When the variable scales differ substantially, the standardized coefficient vector provides better information about the relative contribution of each variable to the canonical discriminantfunction.

Suppose there are two groups of p predictor variables, which allow for construction of LDA functions using all predictors. A practical process is to choose significant variables using stepwise procedure, which usesthe Wilks Lambda statisticsto identify significant independent variables of the discriminant functions (Siotani et al. 1985, Rencher 1993). The Wilks Lambda criterion maximally discriminates between groups by maximizing the multivariate F ratio in the tests of differences between the group-means.

APPENDIX D

DATABASES

- Information Retrieval in Medical Applications (IRMA) database
- MIT/Beth Israel Hospital (BIH) database
- International Consortium for Brain Mapping (ICBM) database
- UCI repository

REFERENCES

[1] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik, "Blobworld: A system for region-based image indexing and retrieval," in *Visual Information and Information Systems*, pp. 509–517, Springer, 1999.

[2] R. C. Veltkamp and M. Tanase, "Content-based image retrieval systems: A survey," *rapport no UU-CS-2000-34*, 2000.

[3] F. Long, H. Zhang, and D. D. Feng, "Fundamentals of content-based image retrieval," in *Multimedia Information Retrieval and Management*, pp. 1–26, Springer, 2003.

[4] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 8, no. 6, pp. 460–473, 1978.

[5] R. C. Veltkamp and M. Hagedoorn, *State of the art in shape matching.* Springer, 2001.

[6] M. A. Smith and T. Chen, *Image and video indexing and retrieval.* Academic Press, 2000.

[7] P. Brodatz, *Textures: a photographic album for artists and designers.* Dover New York, 1966.

[8] C. W. Niblack, R. Barber, W. Equitz, M. D. Flickner, E. H. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin, "QBIC: project querying images by content, using color, texture, and shape," in *IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology*, pp. 173–187, International Society for Optics and Photonics, 1993.

[9] E. Loupias and N. Sebe, "Wavelet-based salient points: Applications to image retrieval using color and texture features," in *Advances in Visual Information Systems*, pp. 223–232, Springer, 2000.

[10] A. Guttman, *R-trees: A dynamic index structure for spatial searching*, vol. 14. ACM, 1984.

[11] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger, *The R*-tree: an efficient and robust access method for points and rectangles*, vol. 19. ACM, 1990.

[12] P. N. Yianilos, "Data structures and algorithms for nearest neighbor search in general metric spaces," in *Proceedings of the fourth annual ACM-SIAM Symposium on Discrete algorithms*, pp. 311–321, Society for Industrial and Applied Mathematics, 1993.

[13] K. Chakrabarti and S. Mehrotra, "The hybrid tree: An index structure for high dimensional feature spaces," in *Data Engineering, 1999. Proceedings., 15th IEEE International Conference on*, pp. 440–447, 1999.

[14] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Region-based image querying," in *Content-Based Access of Image and Video Libraries, 1997. Proceedings. IEEE Workshop on*, pp. 42–49, IEEE, 1997.

[15] W. W. Chu, C.-C. Hsu, A. F. Cárdenas, and R. K. Taira, "Knowledge-based image retrieval with spatial and temporal constructs," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 10, no. 6, pp. 872–888, 1998.

[16] W. Cai, D. Feng, and R. Fulton, "Content-based retrieval of dynamic pet functional images," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 4, no. 2, pp. 152–158, 2000.

[17] C.-R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. M. Aisen, and L. S. Broderick, "Assert: a physician-in-the-loop content-based retrieval system for hrct image databases," *Computer Vision and Image Understanding*, vol. 75, no. 1, pp. 111–132, 1999.

[18] A. Guimond, G. Subsol, and J.-P. Thirion, "Automatic mri database exploration and applications," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 11, no. 08, pp. 1345–1365, 1997.

[19] S. Li, H. Yin, and L. Fang, "Group-sparse representation with dictionary learning for medical image denoising and fusion," *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 12, pp. 3450–3459, 2012.

[20] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, and C. Roux, "Wavelet optimization for content-based image retrieval in medical databases," *Medical image analysis*, vol. 14, no. 2, pp. 227–241, 2010.

[21] K. Etemad and R. Chellappa, "Separability-based multiscale basis selection and feature extraction for signal and image classification," *Image Processing, IEEE Transactions on*, vol. 7, no. 10, pp. 1453–1465, 1998.

[22] F. Rodriguez and G. Sapiro, "Sparse representations for image classification: Learning discriminative and reconstructive non-parametric dictionaries," tech. rep., DTIC Document, 2008.

[23] M. Ranzato, F. J. Huang, Y.-L. Boureau, and Y. Lecun, "Unsupervised learning of invariant feature hierarchies with applications to object recognition," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pp. 1–8, IEEE, 2007.

[24] Y.-C. Chen, C. S. Sastry, V. M. Patel, P. J. Phillips, and R. Chellappa, "In-plane rotation and scale invariant clustering using dictionaries," *IEEE Transactions on Image Processing*, vol. 22, no. 6, pp. 2166–2180, 2013.

[25] F. Yu, Y. Chen, and L. Luo, "Ct image denoising based on sparse representation using global dictionary," in *Complex Medical Engineering (CME), 2013 ICME International Conference on*, pp. 408–411, IEEE, 2013.

[26] L. Grewe and A. C. Kak, "Interactive learning of a multiple-attribute hash table classifier for fast object recognition," *Computer Vision and Image Understanding*, vol. 61, no. 3, pp. 387–416, 1995.

[27] T. M. Lehmann, M. O. Güld, C. Thies, B. Plodowski, D. Keysers, B. Ott, and H. Schubert, "Irma-content-based image retrieval in medical applications," *Medinfo*, pp. 842–8, 2004.

[28] H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval systems in medical applicationsclinical benefits and future directions," *International journal of medical informatics*, vol. 73, no. 1, pp. 1–23, 2004.

[29] F.-H. Kong, "Image retrieval using both color and texture features," in *Machine Learning and Cybernetics, 2009 International Conference on*, vol. 4, pp. 2228–2232, IEEE, 2009.

[30] K. Konstantinidis, A. Gasteratos, and I. Andreadis, "Image retrieval based on fuzzy color histogram processing," *Optics Communications*, vol. 248, no. 4, pp. 375–386, 2005.

[31] M. J. Swain and D. H. Ballard, "Color indexing," *International journal of computer vision*, vol. 7, no. 1, pp. 11–32, 1991.

[32] J. R. Smith and S.-F. Chang, "Tools and techniques for color image retrieval.," in *Storage and Retrieval for Image and Video Databases (SPIE)*, vol. 2670, pp. 2–7, 1996.

[33] G. Pass, R. Zabih, and J. Miller, "Comparing images using color coherence vectors," in *Proceedings of the fourth ACM international conference on Multimedia*, pp. 65–73, ACM, 1997.

[34] K.-L. Tan, B. C. Ooi, and C. Y. Yee, "An evaluation of color-spatial retrieval techniques for large image databases," *Multimedia Tools and Applications*, vol. 14, no. 1, pp. 55–78, 2001.

[35] C. Böhm, S. Berchtold, and D. A. Keim, "Searching in high-dimensional spaces: Index structures for improving the performance of multimedia databases," *ACM Computing Surveys (CSUR)*, vol. 33, no. 3, pp. 322–373, 2001.

[36] R. Bayer and E. McCreight, *Organization and maintenance of large ordered indexes.* Springer, 2002.

[37] C. Faloutsos, *Searching multimedia databases by content*, vol. 3. Springer, 1996.

[38] D. Zhang and G. Lu, "Shape retrieval using fourier descriptors," in *In Proceedings of 2nd IEEE Pacific Rim Conference on Multimedia*, pp. 1–9, Citeseer, 2001.

[39] D. Zhang and G. Lu, "Content-based shape retrieval using different shape descriptors: A comparative study.," in *ICME*, pp. 1139–1142, 2001.

[40] A. Mojsilovic and J. Gomes, "Semantic based categorization, browsing and retrieval in medical image databases," in *Image Processing. IEEE International Conference on*, vol. 3, pp. 145–148, IEEE, 2002.

[41] T. M. Lehmann, M. O. Güld, T. Deselaers, D. Keysers, H. Schubert, K. Spitzer, H. Ney, and B. B. Wein, "Automatic categorization of medical images for content-based retrieval and data mining," *Computerized Medical Imaging and Graphics*, vol. 29, no. 2, pp. 143–155, 2005.

[42] U. Avni, H. Greenspan, M. Sharon, E. Konen, and J. Goldberger, "X-ray image categorization and retrieval using patch-based visualwords representation," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, pp. 350–353, IEEE, 2009.

[43] F. Li-dong and Z. Yi-fei, "Medical image retrieval and classification based on morphological shape feature," in *Intelligent Networks and Intelligent Systems (ICINIS), 2010 3rd International Conference on*, pp. 116–119, IEEE, 2010.

[44] M. R. Zare, M. Awedh, A. Mueen, and W. C. Seng, "Merging scheme-based classification of medical x-ray images," in *Computational Intelligence, Modelling and Simulation (CIMSiM), 2011 Third International Conference on*, pp. 253–258, IEEE, 2011.

[45] B. Xingli and T. Zhengjun, "Medical images classification based on least square support vector machines," in *Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on*, pp. 1–4, IEEE, 2009.

[46] F. Keyvanfard, M. A. Shoorehdeli, and M. Teshnehlab, "Feature selection and classification of breast MRI lesions based on multi classifier," in *Artificial Intelligence and Signal Processing (AISP), 2011 International Symposium on*, pp. 54–58, IEEE, 2011.

[47] B. Krawczyk and G. Schaefer, "Ensemble fusion methods for medical data classification," in *Neural Network Applications in Electrical Engineering (NEUREL), 2012 11th Symposium on*, pp. 143–146, IEEE, 2012.

[48] B. Krawczyk, G. Schaefer, and M. Wozniak, "A cost-sensitive ensemble classifier for breast cancer classification," in *Applied Computational Intelligence and Informatics (SACI), 2013 IEEE 8th International Symposium on*, pp. 427–430, IEEE, 2013.

[49] M. Vannucci and V. Colla, "Novel classification method for sensitive problems and uneven datasets based on neural networks and fuzzy logic," *Applied Soft Computing*, vol. 11, no. 2, pp. 2383–2390, 2011.

[50] P. Tirilly, K. Lu, X. Mu, T. Zhao, and Y. Cao, "On modality classification and its use in text-based image retrieval in medical databases," in *Content-Based Multimedia Indexing (CBMI), 2011 9th International Workshop on*, pp. 109–114, IEEE, 2011.

[51] G. Csurka, S. Clinchant, and G. Jacquet, "Medical image modality classification and retrieval," in *Content-Based Multimedia Indexing (CBMI), 2011 9th International Workshop on*, pp. 193–198, IEEE, 2011.

[52] H. D. Tagare, C. C. Jaffe, and J. Duncan, "Medical image databases a content-based retrieval approach," *Journal of the American Medical Informatics Association*, vol. 4, no. 3, pp. 184–198, 1997.

[53] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 12, pp. 1349–1380, 2000.

[54] H. Pourghassem and H. Ghassemian, "Content-based medical image classification using a new hierarchical merging scheme," *Computerized Medical Imaging and Graphics*, vol. 32, no. 8, pp. 651–661, 2008.

[55] D. Keysers, B. B. Wein, T. M. Lehmann, H. Ney, *et al.*, "Statistical framework for model-based image retrieval in medical applications," *Journal of Electronic Imaging*, vol. 12, no. 1, pp. 59–68, 2003.

[56] H. Greenspan and A. T. Pinhas, "Medical image categorization and retrieval for pacs using the gmm-kl framework," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 11, no. 2, pp. 190–202, 2007.

[57] I. Ramirez, P. Sprechmann, and G. Sapiro, "Classification and clustering via dictionary learning with structured incoherence and shared features," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 3501–3508, IEEE, 2010.

[58] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proceedings of the 26th Annual International Conference on Machine Learning*, pp. 689–696, ACM, 2009.

[59] M. Aharon, M. Elad, and A. Bruckstein, "-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4311–4322, 2006.

[60] K. Engan, S. O. Aase, and J. Hakon Husoy, "Method of optimal directions for frame design," in *Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on*, vol. 5, pp. 2443–2446, IEEE, 1999.

[61] K. Huang and S. Aviyente, "Sparse representation for signal classification," in *NIPS*, pp. 609–616, 2006.

[62] P. Sprechmann and G. Sapiro, "Dictionary learning and sparse coding for unsupervised clustering," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 2042–2045, IEEE, 2010.

[63] T. M. Lehmann, H. Schubert, D. Keysers, M. Kohnen, and B. B. Wein, "The irma code for unique classification of medical images," in *Medical Imaging 2003*, pp. 440–451, International Society for Optics and Photonics, 2003.

[64] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," *Advances in neural information processing systems*, vol. 19, pp. 801–808, 2007.

[65] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *Signal Processing, IEEE Transactions on*, vol. 41, no. 12, pp. 3397–3415, 1993.

[66] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 8, no. 6, pp. 679–698, 1986.

[67] M. M. Rahman, P. Bhattacharya, and B. C. Desai, "A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 11, no. 1, pp. 58–69, 2007.

[68] N. J. Fesharaki and H. Pourghassem, "Medical x-ray images classification based on shape features and bayesian rule," in *Computational Intelligence and Communication Networks (CICN), 2012 Fourth International Conference on*, pp. 369–373, IEEE, 2012.

[69] B. Krawczyk and M. Wozniak, "Experiments on distance measures for combining one-class classifiers," in *Computer Science and Information Systems (FedCSIS), 2012 Federated Conference on*, pp. 89–92, IEEE, 2012.

[70] T. Rohlfing, D. B. Russakoff, R. Brandt, R. Menzel, and C. Maurer Jr, "Performance-based multi-classifier decision fusion for atlas-based segmentation of biomedical images," in *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, pp. 404–407, IEEE, 2004.

[71] Y.-Y. Nguwi and S.-Y. Cho, "An unsupervised self-organizing learning with support vector ranking for imbalanced datasets," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8303–8312, 2010.

[72] V. Vapnik, *The nature of statistical learning theory.* springer, 2000.

[73] A. Asuncion and D. Newman, "Uci machine learning repository," 2007.

[74] A. M. Canuto, A. M. Santos, M. C. Abreu, V. M. Bezerra, F. M. Souza, and M. F. G. Junior, "Investigating the use of an agent-based multi-classifier system for classification tasks," in *Neural Information Processing*, pp. 854–859, Springer, 2004.

[75] W. H. Wolberg and O. L. Mangasarian, "Multisurface method of pattern separation for medical diagnosis applied to breast cytology.," *Proceedings of the national academy of sciences*, vol. 87, no. 23, pp. 9193–9196, 1990.

[76] V. Balasubramanian, S. Chakraborty, and S. Panchanathan, "Generalized query by transduction for online active learning," in *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pp. 1378–1385, IEEE, 2009.

[77] F. Peng, L. Li, W. Xu, W. Liu, J. Zhang, and G. Shao, "The identification of breast mass based on multi-agent interactive information fusion method," in *Bioinformatics and Biomedical Engineering, 2009. ICBBE 2009. 3rd International Conference on*, pp. 1–4, IEEE, 2009.

[78] J. Wei, H. Yu, and J. Wang, "The research of bayesian method from small sample of high-dimensional dataset in poison identification," in *Software Engineering and Service Science (ICSESS), 2013 4th IEEE International Conference on*, pp. 705–709, IEEE, 2013.

[79] M. Cha, J. S. Kim, and J.-G. Baek, "Density weighted support vector data description," *Expert Systems with Applications*, vol. 41, no. 7, pp. 3343–3350, 2014.

[80] B. Liu, Y. Xiao, L. Cao, Z. Hao, and F. Deng, "Svdd-based outlier detection on uncertain data," *Knowledge and information systems*, vol. 34, no. 3, pp. 597–618, 2013.

[81] M. Yuwono, S. W. Su, B. D. Moulton, and H. T. Nguyen, "Optimization strategies for rapid centroid estimation," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 6212–6215, IEEE, 2012.

[82] W. Duch, N. Jankowski, and T. Maszczyk, "Make it cheap: learning with o (nd) complexity," in *Neural Networks (IJCNN), The 2012 International Joint Conference on*, pp. 1–4, IEEE, 2012.

[83] R. Ramos-Pollán, J. M. Franco, J. Sevilla, M. A. Guevara-López, N. Gonzalez de Posada, J. Loureiro, and I. Ramos, "Grid infrastructures for developing mammography cad systems," in *Engineering in Medicine and Biology Society (EMBC), IEEE 2010 Annual International Conference*, pp. 3467–3470, IEEE, 2010.

[84] S.-y. Jiang, X. Li, Q. Zheng, and L.-x. Wang, "Approximate equal frequency discretization method," in *Intelligent Systems, 2009. GCIS'09. WRI Global Congress on*, vol. 3, pp. 514–518, IEEE, 2009.

[85] C. F. Eick, N. Zeidat, and R. Vilalta, "Using representative-based clustering for nearest neighbor dataset editing," in *Data Mining, 2004. ICDM'04. Fourth IEEE International Conference on*, pp. 375–378, IEEE, 2004.

[86] S. Rodda and M. Shashi, "An improved associative classifier," in *Conference on Computational Intelligence and Multimedia Applications, 2007. International Conference on*, vol. 2, pp. 286–290, IEEE, 2007.

[87] K. Polat and S. Güneş, "A new feature selection method on classification of medical datasets: Kernel f-score feature selection," *Expert Systems with Applications*, vol. 36, no. 7, pp. 10367–10373, 2009.

[88] Y. Guo, X. Niu, and H. Zhang, "An extensive empirical study on semi-supervised learning," in *Data Mining (ICDM), 2010 IEEE 10th International Conference on*, pp. 186–195, IEEE, 2010.

[89] K. Ouchi, A. Nakamura, and M. Kudo, "An efficient construction and application usefulness of rectangle greedy covers," *Pattern Recognition*, vol. 47, no. 3, pp. 1459–1468, 2014.

[90] S.-C. Chen, S.-W. Lin, T.-Y. Tseng, and H.-C. Lin, "Optimization of back-propagation network using simulated annealing approach," in *Systems, Man and Cybernetics, 2006. SMC'06. IEEE international conference on*, vol. 4, pp. 2819–2824, IEEE, 2006.

[91] Z. Li, "A support vector machine training algorithm based on cascade structure," in *Innovative Computing, Information and Control, 2006. ICICIC'06. First International Conference on*, vol. 3, pp. 440–443, IEEE, 2006.

[92] R. Kumar, T. Chen, M. Hardt, D. Beymer, K. Brannon, and T. Syeda-Mahmood, "Multiple kernel completion and its application to cardiac disease discrimination," in

Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on, pp. 764–767, IEEE, 2013.

[93] C. Li-lin, Z. Hai-chao, Z. Lin-ke, and L. Rui-peng, "Improved k nearest neighbors transductive confidence machine for pattern recognition," in *Computer Design and Applications (ICCDA), 2010 International Conference on*, vol. 3, pp. V3–172, IEEE, 2010.

[94] D. Tian, J. Keane, and X.-J. Zeng, "Core-generating approximate minimum entropy discretization for rough set feature selection: an experimental investigation," in *IEEE International Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007.*, pp. 1–6, IEEE, 2007.

[95] V. Gál, E. Kerre, and M. Nachtegael, "Multiple kernel learning based modality classification for medical images," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, pp. 76–83, IEEE, 2012.

[96] E. Rashedi, H. Nezamabadi-Pour, and S. Saryazdi, "A simultaneous feature adaptation and feature selection method for content-based image retrieval systems," *Knowledge-Based Systems*, vol. 39, pp. 85–94, 2013.

[97] U. Avni, H. Greenspan, E. Konen, M. Sharon, and J. Goldberger, "X-ray categorization and retrieval on the organ and pathology level, using patch-based visual words," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 3, pp. 733–746, 2011.

[98] B. Ophir, M. Lustig, and M. Elad, "Multi-scale dictionary learning using wavelets," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 1014–1024, 2011.

[99] P. Huang, S. Dai, and P. Lin, "Texture image retrieval and image segmentation using composite sub-band gradient vectors," *Journal of Visual Communication and Image Representation*, vol. 17, no. 5, pp. 947–957, 2006.

[100] R. Mark and G. Moody, "Mit-bih arrhythmia database directory," *Cambridge: Massachusetts Institute of Technology*, 1988.

[101] Y. H. Hu, S. Palreddy, and W. J. Tompkins, "A patient-adaptable ecg beat classifier using a mixture of experts approach," *Biomedical Engineering, IEEE Transactions on*, vol. 44, no. 9, pp. 891–900, 1997.

[102] P. de Chazal and R. B. Reilly, "A patient-adapting heartbeat classifier using ecg morphology and heartbeat interval features," *Biomedical Engineering, IEEE Transactions on*, vol. 53, no. 12, pp. 2535–2543, 2006.

[103] J. Wiens and J. V. Guttag, "Active learning applied to patient-adaptive heartbeat classification," in *Advances in neural information processing systems*, pp. 2442–2450, 2010.

[104] A. S. Alvarado, C. Lakshminarayan, and J. C. Principe, "Time-based compression and classification of heartbeats," *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 6, pp. 1641–1648, 2012.

[105] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–220, 2000.

[106] G. Doquire, G. De Lannoy, D. François, and M. Verleysen, "Feature selection for interpatient supervised heart beat classification," *Computational intelligence and neuroscience*, vol. 2011, pp. 1–7, 2011.

[107] R. Mark and R. Wallen, "Aami-recommended practice: Testing and reporting performance results of ventricular arrhythmia detection algorithms," *Association for the Advancement of Medical Instrumentation, Arrythmia Monitorring Subcommittee, AAMI ECAR*, 1987.

[108] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ecg morphology and heartbeat interval features," *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 7, pp. 1196–1206, 2004.

[109] S.-H. Cha, "Comprehensive survey on distance/similarity measures between probability density functions," *International journal of mathematical models and methods in applied sciences*, vol. 1, no. 2, pp. 300–307, 2007.

[110] S. Haykin and N. Network, "A comprehensive foundation," *Neural Networks*, vol. 2, no. 2004, 2004.

List of Publications

CONFERENCES

1. M. Srinivas and C. Krishna Mohan, "Multi-level classification : A generic classification method for medical data sets," Communicated to *IEEE Int. Conf. on International Joint Conference on Neural Networks (IJCNN 2015)*, Ireland, Jul. 2015.

2. M. Srinivas, Ramkrishna Bharath, Pachamuthu Rajalakshmi and C. Krishna Mohan, "Sparseland Model for Speckle Suppression of B-mode Ultrasound Images," Communicated to *IEEE National Conference on Communication (NCC-2015)*, IIT Mumbai, India, Feb. 2015.

3. Debaditya Roy, M. Srinivas, and C. Krishna Mohan, "Sparsifying Dense Features for Action Classification," Communicated to *IEEE Int. Conf. on Perception and Machine Intelligence* , Kolkata, India, Feb. 2015.

4. M. Srinivas and C. Krishna Mohan, "Classification of Medical Images Using Edge-Based Features and On-line Dictionary Learning," Communicated to *IEEE Int. Conf. on Perception and Machine Intelligence* , Kolkata, India, Feb. 2015.

5. M. Srinivas and C. Krishna Mohan, "Medical Images Modality Classification using Multi-scale Dictionary Learning,"in *Proc. 19th IEEE Int. Conf. on Digital Signal Processing (DSP)*, Hong Kong, Aug. 2014.

6. M. Srinivas, Debaditya Roy, and C. Krishna Mohan, "Learning Sparse Dictionaries for Music and Speech Classification," in *Proc. 19th IEEE Int. Conf. on*

Digital Signal Processing (DSP), Hong Kong, Aug. 2014.

7. Shyju Wilson, M. Srinivas and C. Krishna Mohan, "Dictionary based action video classification with action bank," in *Proc. 19th IEEE Int. Conf. on Digital Signal Processing (DSP)*, Hong Kong, Aug. 2014.

8. M. Srinivas, Debaditya Roy and C. Krishna Mohan, "Music Genre Classification using On-line Dictionary Learning," in *Proc. IEEE Int. Conf. on International Joint Conference on Neural Networks (IJCNN)*, Beijing, July. 2014.

9. M. Srinivas and C. Krishna Mohan, "Medical Image Indexing and Retrieval using Multi-Feature Extraction Method," in *Proc. IEEE Int. Conf. on Computational Intelligence and Information Technology (CIIT)*, Mumbai, Oct. 2013.

10. M. Srinivas and C. Krishna Mohan, "Efficient clustering approach using incremental and hierarchical clustering methods," in *Proc. IEEE Int. Conf. on International Joint Conference on Neural Networks (IJCNN)*, Barcelona, Jul. 2010.

JOURNALS

1. M. Srinivas, R. Ramu Naidu, C. S. Sastry and C. Krishna Mohan, "Content Based Medical Image Retrieval Using Dictionary Learning," Communicated to *Journal of Neurocomputing (Elsevier)*. Oct. 2014.

2. M. Srinivas, Tony Basil and C. Krishna Mohan, "Adaptive learning based heartbeat classification," Communicated to *Bio-Medical Materials and Engineering* . Oct. 2014.

CURRICULUM VITAE

1. **Name:** M. Srinivas

2. **Date of Birth:** 24^{th} April 1983

3. **Permanent Address:**
   S/o. M. Odelu
   D.No: 1-105/3, Hitex colony
   P.O. Mancherial - 504208
   Adilabad Dist., Andhra Pradesh, India

4. **Educational Qualifications:**
   - December 2014: Doctor of Philosophy in Computer Science and Engineering (Ph.D., Dept. of CSE, IIT Hyderabad, A.P., India)
   - May 2008: Master of Technology in Software Engineering (M.Tech., JNTU, Hyderabad, A.P., India)
   - May 2005: Bachelor of Technology in Computer Science and Information Technology (B.Tech., JNTU, Hyderabad, A.P., India)

5. **Employment History:**
   - Oct 2008 - Dec 2009: Lecturer in JNTU., Anantapur, A.P., India.

DOCTORAL COMMITTEE

1. **Chairperson:** Dr. M. V. Pandurangarao

2. **Guide:** Dr. C. Krishna Mohan

3. **Members:**

   - Dr. Bheemarjuna Reddy Tamma (Dept. of CSE)
   - Dr. C. S. Sastry (Dept. of Mathematics)
   - Dr. K. Sri Rama Murty (Dept. of EE)