Comparative Evaluation of PCA-based Feature Transformation Techniques in Classification of Hepatic Focal Lesions with Ultrasound Images

A Dissertation submitted in fulfillment of the requirements for the Degree of

MASTER OF ENGINEERING

in

Electronic Instrumentation & Control Engineering

Submitted by

HIMANSHU PAUL
Regd. No.: 801351007

Under the Guidance of

Dr. DEEPTI MITTAL
Assistant Professor, EIED

2016

Electrical and Instrumentation Engineering Department
Thapar University, Patiala
(Declared as Deemed-to-be-University u/s 3 of the UGC Act., 1956)
Post Bag No. 32, Patiala – 147004
Punjab (India)

DECLARATION

I hereby certify that the work which is presented in dissertation entitled, "**Comparative Evaluation of PCA-based Feature Transformation Techniques in Classification of Hepatic Focal Lesions with Ultrasound Images**", in partial fulfillment of the requirements for the award of the degree of Master of Engineering in Electronic Instrumentation and Control, submitted to Electrical & Instrumentation Engineering Department of Thapar University, Patiala is as authentic record of my own work carried under the supervision of Dr. Deepti Mittal. It refers others researcher's work which are duly listed in the reference section. The matter contained in this dissertation has not been submitted, neither in part nor in full to any other degree to any other university or institute except as reported in text and references.

Place: Patiala

Date: 14/7/2016

(HIMANSHU PAUL)

(801451010)

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge and belief.

(DEEPTI MITTAL)

Assistant Professor

Electrical & Instrumentation Engineering Department

Thapar University, Patiala

Countersigned By:

(Dr. RAVINDER AGGARWAL)

Professor & Head

Electrical & Instrumentation Engineering Department

Thapar University, Patiala

(Dr. S.S. BHATIA)

Sr. Professor & Dean

Academic Affairs

Thapar University, Patiala

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

PCA – Principal Component Analysis

FPCA– Fast Principal Component Analysis

SPCA– Sparse Principal Component Analysis

KPCA– Kernel Principal Component Analysis

NFRPCA– Non-linear Fuzzy Robust Principal Component Analysis

US– Ultrasonic Imaging

HCC– Hepatocellular carcinoma

HEM– Hemangioma

MET – Metastases

PCs– Principal Components

GLCM – Grey Level Co-occurrence Matrix

GLRLM– Grey Level Run Length Matrix

FOS– First Order Statistics

SGLDM– Spatial Grey Level Dependence Matrix

TEM– Texture Energy Measures

Se- Sensitivity

Sp– Specificity

Ac- Accuracy

SVM- Support Vector Machine Classifier

ERR- Equal Error Rate

ANF –Average Number Of Features

Abstract

Dimensionality reduction is the most imperative stage to extract relevant information from the feature set used in a specific classification problem. Principal component analysis (PCA) is a widely accepted and frequently used dimensionality reduction technique that converts the set of correlated variables into set of uncorrelated variables. Various PCA-based techniques are available in the literature with different motives, but it is still unclear which one will perform well in the problem of detection and classification of focal hepatic lesions. Therefore, there is a need to perform a comparative evaluation to select the best PCA technique for dimensionality reduction of feature set in this specific problem. Consequently, in the present work a comparative evaluation has been performed with four PCA-based techniques, named as Fast PCA, Sparse PCA, Kernel PCA and Nonlinear fuzzy robust PCA. Among them, frequently applicable selected PCA techniques are (i) linear transformation techniques, viz., Fast PCA and Sparse PCA (ii) nonlinear transformation techniques, viz., Kernel PCA and Nonlinear fuzzy robust PCA. These PCA techniques have been applied on the 208 texture features to find out the best possible diagnostically important principal components in order to classify the five liver tissue categories [6]. Subsequently, these principal components are used to train the multi support vector machine classifier. The experimental results reveal that Sparse PCA outperforms the other PCA-based techniques showing the overall classification accuracy of 94%. Results also reveal that kernel PCA along with polynomial kernel outperforms the Fast PCA as it captures the high-order information from the feature space which is not possible by applying Fast PCA.

KEYWORDS: Principal component analysis, Fast principal component analysis, Sparse principal component analysis, Kernel principal component analysis, Nonlinear robust fuzzy principal component analysis.

Chapter 1

Introduction

1.1 Overview

Hepatic focal lesions refer to small area of the damaged liver tissues. Several types of lesions can build up in the liver and may be benign or malignant (cancer). Both lesion types are potentially disabling and life threatening. In fact, hepatic liver lesions are the third leading cause of cancer related deaths [1]. Liver cancer occurs mostly in the age of 45 years or more and is rare in children and young adults. Early and accurate diagnosis of hepatic focal lesions is the solution for implementing successful therapy. One of the most preferable imaging modality for the initial detection of hepatic focal lesions is B-mode (US) ultrasound imaging due to its low cost, non-invasive detection and real time imaging.

1.2. Need of dimensionality reduction

The diagnosis of focal liver lesions is a difficult task due to the large variation, overlapping and complexity in lesion description in terms of shape, size, texture and intensities with B-mode ultrasound images. Furthermore, different radiologists have different interpretations about atypical lesion cases on B-mode ultrasound images. They characterize the hepatic focal lesions along with normal liver tissues on the basis of their experience by visually assessing their texture patterns on ultrasound images. Texture of any tissue types can be broadly categorized as homogeneous or heterogeneous [6-8]. The diagnosis of focal liver lesions becomes a perplexing task for radiologist in the routine practice. Several research works have been done to avoid this discrepancy in the direction of designing the computer-aided diagnosis system. Classification of liver tissues in computer-aided diagnostic solutions is performed by extracting the textural features using spectral, statistical, and spatial filtering based feature extraction methods [2-8]. The information contents in the extracted features are highly interconnected due to which it results in the curse of dimensionality. Curse of dimensionality increases the computational complexity and classification error. Consequently the reduction in dimensionality is a desirable step that should be carried out as a pre-eminent stage before the supervised classification of focal liver lesions. This pre-eminent stage assists the further classification process by removing the redundant and irrelevant features which degrade the classification performance. Dimensionality

reduction of feature space is performed either using feature selection or feature transformation methods. Feature selection is the practice of reducing the feature set by selecting the relevant set of features from the original features without any transformation while feature transformation transforms the high-dimensional feature set into low dimensional feature set using certain linear or non-linear transformations. These transformed features keep as much information from the data as possible. Feature selection improves the classification performance by selecting the optimal set of features from the feature space. Selection of the optimal features becomes a difficult task in case of high-dimensional feature space. Feature transformation technique eradicates this limitation as it improves the classification performance by transforming the existing features into compact set of features. New transformed features retained most of the information in it. Feature transformation technique exhibit minimum reconstruction error between transformed feature set and existing original feature set.

1.3. Aims and Objectives

Curse of dimensionality and the risk of over fitting is the major issue while dealing with the high-dimensional feature space. It becomes difficult to operate on such a high-dimensional feature space as it is coupled with the high complexity. It becomes difficult to train the machine learning techniques due to such a high complexity of the system. Moreover, high-dimensional feature space consists of the irrelevant number of features which immensely increase the computational cost. To mitigate the problem of computational complexity and system complexity feature transformation becomes the desirable pre-processing step to build a compact model that can be learned in a reasonable amount of time and become efficient to use. Another objective for performing the feature transformation technique is that in order to approximate the function between input and output, it is reasonable and important to ignore such irrelevant features with little effect on the output, so as to keep the size of approximate model small.

1.4. Thesis overview:

Rest of the thesis is well thought-out and systematized as follow: chapter 2, represents the background theory regarding the hepatic focal lesions. Chapter 3, presents the literature review related to hepatic focal lesions and literature review related to PCA based various feature transformation techniques. Chapter 4, presents the methodology which is divided into two fragments. First fragment consists of the supporting structure which forms the ground base for

the second fragment and second fragment consists of proposed work consists of four PCA based feature transformation techniques. Chapter 5, provides a concise outline of multi SVM classifier followed by the description of performance parameters. In chapter 6, experimental setup has presented along with experimental results are discussed in the form of comparative analysis of two linear and two non-linear feature transformation techniques. Finally the thesis is winded up in the form of conclusion in chapter 7.

Chapter 2

Background Theory

2.1. Hepatic Focal Lesions Description

Benign or malignant (cancer) are the two types of focal lesions that can be build up in the liver. Liver tissues can be classified into two types, Normal and focal diseased. Focal diseased tissues are further classified into two types, fluid lesion and solid lesion. Cyst is the most commonly occur fluid lesion and in case of solid lesion it may be categorized into benign or malignant (cancer) lesion. Hemangioma (HEM) is the commonly occurring benign liver lesion which exhibits homogeneous texture patterns with hyperechoic appearance [6]. Malignant liver tumor can be broadly categorized as primary malignant lesion and secondary malignant lesion. Hepatocellular carcinoma (HCC) is commonly occurring primary malignant liver lesion and Texture pattern of HCC is heterogeneous with hypoechoic appearance and has hyper vascularity. Figure 2.1 depicts the sonographic appearance of HCC.

Metastases (MET) is the commonly occurring secondary malignant liver lesion which basically originates from other cancer effected organs. The texture pattern of MET is inhomogeneous with variable echo patterns including hypoechoic, isoechoic, hyperechoic appearance and has irregular vascularization. Figure 2.2, depicts the sonographic appearance of MET. In most of the cases it becomes difficult to distinguish between HCC and MET due to the overlapping of their sonographic appearances [6].

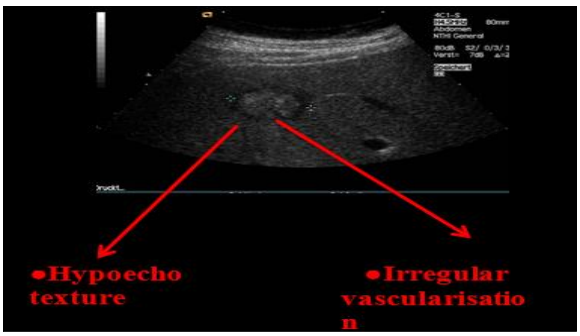

Fig 2.1: Sonographic appearance of MET

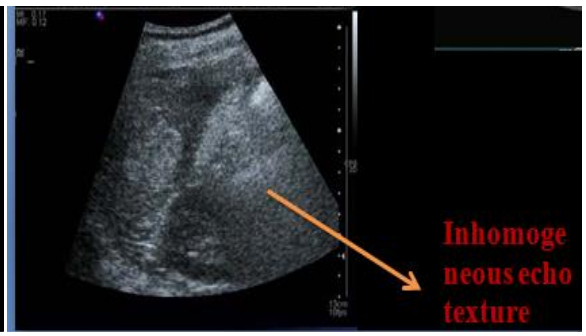

Fig 2.2: Sonographic appearance of HCC

For the detection of these hepatic focal lesions, B-mode (US) ultrasound imaging is preferred as the best imaging modality.

2.2. Ultrasound imaging modality

B-mode US imaging is real time imaging with low cost and non-invasive nature. This is the most significant imaging modality for the suspecting liver lesions. This imaging modality plays a vital role in scrutinizing the problem of HCC and portal hypertension.

2.2.1 Basic principle: Pulse echo approach is the basic principle behind it. Ultrasonic probe plays a vital role in its principle. This probe consists of the various piezo-electric crystals having analogous properties which convert an electronic beam into ultrasonic beam. This converted beam travels through various focal lesions out of which few of the waves gets reflected back. Acoustic impedances between two of the focal lesions become the main cause for the reflection. In reception mode these waves are captured by same ultrasonic probe where it gets converted back to the electronic beam. After the reception mode signal processing mode occurs where signal processing is performed over these electronic signals. Afterwards, these echoes are summed up and form an ultrasonic image.

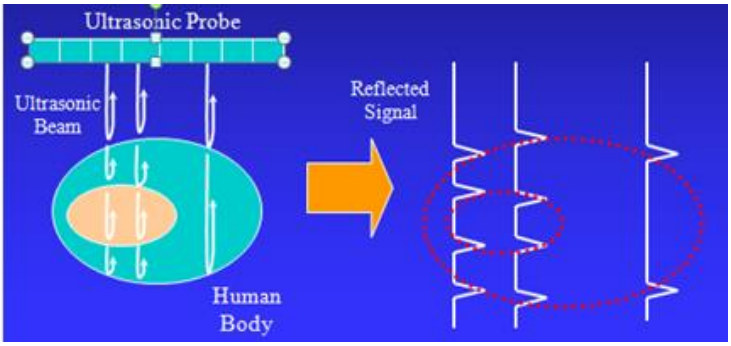

Fig 2.3: Basic principle of ultrasonic imaging

2.2.2. Underneath procedure of ultrasonic image formation: First step after interaction with focal lesions is reflection which is basically due to the acoustic impedances between two of the focal lesions.

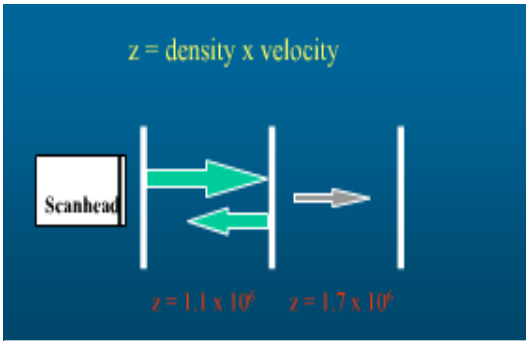

Fig: 2.4 Reflection

After reflection some of the waves gets scattered due to the high degree of roughness within the surface of tissues. Waves which are not scattered go deeper inside the tissues and get reflected from its deeper structure.

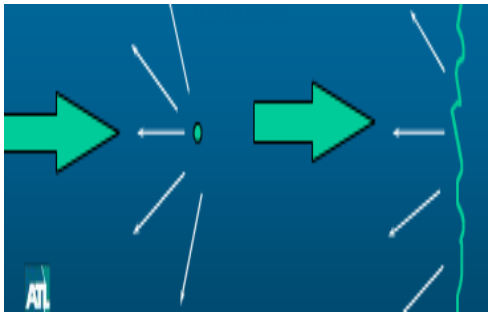

Fig: 2.5 Scattering

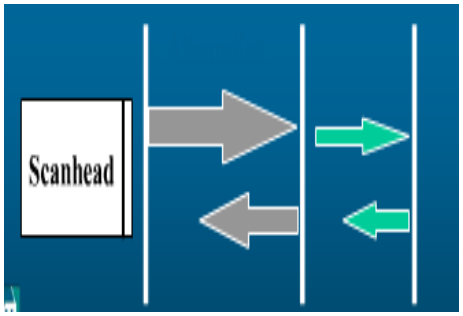

Fig: 2.6 Transmission

After transmission some of the waves which go so deep inside the tissue structure get attenuated. This attenuation occurs due to the fact that as the waves go deeper and deeper, more and more it becomes weaker and as a result it lose its strength and gets attenuated.

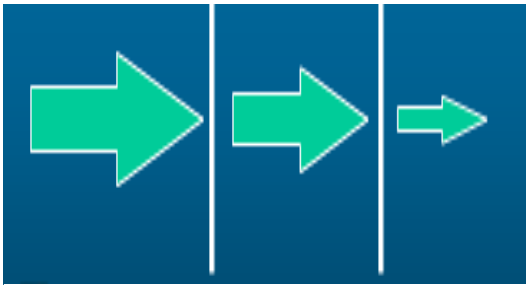

Fig: 2.7 Attenuation

Chapter 3

Literature Review

3.1 Literature review for hepatic focal lesions classification

Till date various research works have been done in the area of computer-aided diagnosis of focal liver lesions and there are very few researchers who have used feature transformation / selection method in this specific field. Table 3.1 depicts the brief detail of their related work.

Sujana *et al* used statistical features based on grey level run length matrix (GLRLM) and first order statistics (FOS) along with Linear discriminate analysis and neural network classifier for the characterization of HEM, HCC and MET [2]. Best classification results were obtained with neural network classifier in comparison with linear discriminate analysis.

Yoshida *et al.* used wavelet-packet-based texture features along with artificial neural network classifier for the characterization of HEM, HCC and MET [3]. The CAD systems designed in the work of Sujana *et al* and Yoshida *et al.* suffer from the limitation that they do not use any feature dimensionality reduction or transformation method for the removal of redundancy from the feature set to improve classification accuracy.

Balasubramanian *et al.* used PCA in their work for the classification of focal liver lesions [4]. They showed in their work that PCA helps in the reduction of computational time. They also mentioned that manually selected features along with neural classifier (NN) provide consistent and proficient classification with higher accuracy in comparison to PCA.

Poonguzhali *et al.* used PCA in their work to extract the best possible features from the feature set consisting of grey length co-occurrence matrix (GLCM), Law's and edge frequency based texture features [5]. They achieved a classification accuracy of 70% with applying the combination of GLCM and Law's features with K-means classifier for the classification of four liver tissue categories – Normal, Cyst, Benign and Malignant.

Mittal *et al.* first make the enhancement of US images and they made use of the box-plot studies for feature selection in order to remove highly noisy features and selected the 208 texture based features from the feature set consisting of 256 features on the basis of sensitivity analysis [6].They achieved the accuracy of 86.4% along with two step neural network classifier for characterization of focal liver lesions such as Cyst, HCC, MET, HEM and NOR. Since the use of box-plot studies help in the reduction of certain noisy features from the overall extracted features

but it is not helpful in order to remove the curse of dimensionality as there still retains certain redundancy in the data.

Virmani *et al.* used 104 texture features and its ratio features and applied PCA to these 208 texture features [7]. They achieved the accuracy of 87.2% along with SVM classifier for the classification of Cyst, HCC, HEM, MET and NOR.

Sakr *et al.* worked on enhanced images and they make use of haralick descriptors, histogram and statistical features for the characterization of Cyst, HEM, HCC and NOR [9]. They achieved the highest classification accuracy of 96.11% with multi-SVM classifier in comparison with KNN classifier.

Mittal *et al.* in their other work has performed the comparative analysis among one vs one multiclass SVM and tree structured based four binary SVMs [11]. In their work they showed an overall classification accuracy of 86.9% using tree structured based four binary SVMs and 93.1% using one vs one multiclass SVM by selecting 208 texture based features on the basis of sensitivity analysis using box-plot.

Mittal *et al.* and Sakr *et al.* both worked on enhanced images. The only difference is that Mittal *et al.* achieved the overall classification accuracy of 93.1% among five classes where as Sakr *et al.* achieved the accuracy of 96.11% among four classes.

Hwang *et al.* used PCA in order to select the most optimal features from the feature set consisting of GLCM, FOS, Law's and echogenicity [12]. They performed the binary classification among Cyst, HEM and malignant lesions using artificial neural network and achieved the classification accuracy of 99.7% among cyst vs HEM, 98.72% among cyst vs malignant and 96.13% among HEM vs malignant.

In most of the previous works, dimensionality of the feature space is reduced using PCA. The classical PCA is a linear transformation technique which performs the orthogonal transformation to convert the correlated feature set into set of linearly uncorrelated feature set. This linearly uncorrelated feature set is called *principal components* along which variance in the feature set is maximal. The general limitations of the above feature transformation method are that this linear PCA transformation technique is (i) based on eigen value decomposition method, where derived eigen values increase the computational complexity. (ii) unable to identify the diagnostically important features which were extracted originally from various feature extraction

methods. To overcome these constraints researchers have refined various versions of the PCA algorithm in the literature which is described in the section 3.2

Table 3.1: Literature review of Hepatic Focal Lesions classification

Authors year	**Liver image class**	**Enhancement method**	**Feature transformation method**	**Classifier used**	**Performance measure (%)**
1996 [2]	●HEM ●Malignant ●MET	Not used	Not used	●Linear discriminate analysis ●Neural Network	●Ac=79.6 ●Ac=100
2003 [3]	●HEM ●HCC ●MET	Not used	Not used	●Artificial neural network	●Se=90 ●Sp=80
2007 [4]	●NOR ●Malignant ●Cyst ●Benign	Not used	Principal component analysis	●K-means classifier ● Neural network	● Normal(Ac)=70 ● Cyst (Ac)=87.5 ●Benign(Ac)=77.5 ●Malignant(Ac)=88 ● Normal(Ac)=75 ● Cyst (Ac)=93.5 ● Benign(Ac)=80 ●Malignant(Ac)=90
2008 [5]	●NOR ●Malignant ●Cyst ●HEM	Not used	Principal component analysis	●K-means classifier	● GLCM (Ac)=67.5 ● Laws (Ac)=60 ●Autocorrelation(Ac)=51.25 ● Edge frequency (Ac)=48.75 ● GLCM + Laws (Ac)=70
2011 [6]	●Cyst ●HCC ●HEM ●MET ●NOR	Modified anisotropic diffusion method	Not used	●Two step neural network	●Ac=86.4
2013 [7]	●Cyst ●HCC ●HEM ●MET ●NOR	Not used	Principal component analysis	●Support vector machine	Ac=87.2
2014 [10]	●Cyst ●HCC ●HEM ●NOR	Not used	Not used	●Multi SVM ●KNN	●Ac=96.11 ●Ac=93.3
2015 [11]	●Cyst ●HEM ●Malignant	Not used	Principal component analysis	●Artificial Neural Network	● Cyst vs HEM (Ac)=99.7 ● Cyst vs Malignant (Ac)=98.72 ● Hem vs Malignant(Ac)=96.13

*se-sensitivity, sp-specificity, Ac-accuracy

3.2 Literature review of feature transformation techniques

3.2.1. Literature review based on linear feature transformation techniques

Table 3.2 describes the brief detail of its study.

K. Shrivastava *et al.* used classic PCA in their work along with SVM classifier for psoriasis risk stratification and image classification [13].They achieved the classification accuracy of 100% using the combination of high order spectra, textural and color features.

Attia *et al.* used classic PCA along with neural network classifier to reduce the set of statistical features and multi-scale wavelet-based features for the classification of ultrasound kidney diseases [14]. They achieved the classification accuracy of 95% using statistical features and accuracy of 97% using wavelet features among five image classes: normal, cyst, stone, tumor and failure by squeezing the information in first 15 principal components (PCs).

Ponomaryov *et al.* used classic PCA to select the most informative features along with SVM classifier for the diagnosis of breast cancer lesions [15]. They achieve the accuracy of 94.4% with high computational cost.

Bagnasco et al. also used classic PCA in their work for the detection of infections caused by mycophilic fungi on dried porcni mushrooms [16]. Characterization was performed on limited number of samples by identifying the pixel values on PC1 and PC2 scores.

Khedher et al. used classic PCA and partial least squares feature transformation methods along with SVM in their CAD system in order to extract the most informative features for the characterization of MRI images for the diagnosis of Alzheimer's disease [17].They performed the binary classification among three classes: healthy patients, Alzheimer disease patients and mild cognitive patients. They achieved the accuracy of 87.5% among healthy and alzheimer disease patients, accuracy of 73.2% among healthy and mild cognitive patients and accuracy of 76.7% among mild cognitive and Alzheimer disease patients.

Jing et al. used classic PCA along with SVM classifier for the classification of faults in industrial process [18]. They achieved the accuracy of 85% among 21 faults.

Neogi *et al.* used principal component analysis along with multi layer perceptron for the classification of fatty and normal liver [19]. In their work PCA helps in the reduction of computational complexity.

Liu *et al.* used classic PCA along with KNN classifier in order to select the efficient features for the detection of liver cancer [20]. They achieved the accuracy of 85% among liver cancer and normal liver .

Yu *et al.* also used classic PCA in their work to extract the informative features for the classification of non-stationary EEG signals using [21]. They achieved the overall classification accuracy of 71.7% along with SVM classifier.

Ilakkiya *et al.* also used classic PCA in their work for the classification of liver cancer tissues [22]. In their work they only perform the binary classification of liver tissues and achieved the overall classification accuracy of 95.8% along with the fuzzy neural network.

Inan *et al.* used classic PCA along with Apriori algorithm for selecting the optimal features together with artificial neural network for the detection of breast cancer [23]. In their method use of Apriori algorithm along with PCA achieved the overall classification accuracy of 95.8%.

Zhou et al. used classic PCA for face recognition problem [24]. They achieved the overall classification accuracy of 81.67% along with LDA and SVM classifier with limited number of samples.

Meng *et al.* used sparse PCA to enhance the properties of the classical PCA technique by pulling out principal components of given data with non-zero sparse loadings [25]. Sparse PCA in their method alleviates the problem of outliers and noise.

Lopez *et al.* used classic PCA a for the diagnosis of Alzheimer's disease [26].They achieved the overall classification accuracy of 96.7% using(PCA+LDA) along with SVM classifier which is 7.2% higher than PCA along with NN classifier.

Calisir *et al.* used classic PCA along with least square SVM classifier for the diagnosis of hepatitis diseases [27]. They achieve the accuracy of 96.12% by squeezing the information in first 10 PCs.

Thomaz *et al.* used classic PCA in face recognition problem [28]. In their work they achieved the average recognition rate of 72% using LDA and SVM classifier without enhancing any image artifacts.

Alok Sharma *et al.* used Fast PCA using fixed point algorithm [29]. It finds the dominant eigenvectors with less computational cost than the method based on eigen value decomposition.

Luan *et al.* used robust PCA for the recognition of face images with varying illumination and occlusion [30]. Robust PCA recover the each component that exhibits more discriminating information which is more beneficial for face identification.

Nguyen *et al.* used bidirectional PCA for pedestrian detection based on image reconstruction [31]. Bidirectional PCA preserves the structural shape of objects and is computationally efficient. Comparative analysis has been carried out between BDPCA and PCA. Bidirectional PCA yields better result in comparison with the classical PCA. Although their algorithm has one limitation that it exhibits more processing time.

3.2.2. Literature review based on non- linear feature transformation techniques

Du *et al.* used graph kernel PCA to extract the informative features for the classification of Attention Deficit Hyperactivity Disorder patients [32]. They achieved the overall classification accuracy of 94.9% along with SVM classifier.

Shao *et al.* used PCA and kernel PCA along with SVM classifier for classification of faults in vibration analysis [33]. In their work kernel PCA yields better result by yielding overall classification accuracy of 95% in comparison with the PCA.

Luukka *et al.* used non-linear fuzzy robust PCA (NFRPCA) along with similarity classifier in the characterization of medical data sets [34]. They achieved the overall classification accuracy of 72.2% using liver disorder dataset, 88.9% using hepatitis dataset and 97% using dermatology dataset. Obtained results in their work shows that NFRPCA performs better when there is an existence of highly overlapped non-linear features.

Wong *et al.* used Kernel PCA to extract the fixed-length feature vector from the original variable-size MLC template to represent the fingerprint template [35].

Licciardi *et al.* used the non-linear PCA for detection of targets and classification of synthetic Aperture Radar images [36]. The received signal is considered as a mixture of different basic scattering signals. Non-linear PCA separate these scattering signals in order to exploit these images in a more interpretable way.

Quan Wang *et al.* used Kernel PCA in face recognition [37]. In their work gaussian kernel PCA outperforms the classic PCA in revealing more complicated structures of data.

Table 3.2: Literature review of various feature transformation techniques

Literature review based on linear feature transformation techniques

	Authors year	Feature transformation method	Classifier used	Research area	Performance measures (%)	Pros	Cons
[13]	2016	PCA	●SVM	Classification of psoriasis skin images	●Ac=100 ●Se=100 ●sp=100	●Selects the most dominant features to obtain the optimized performance	●Color feature set and texture feature set yield less reliability index
[15]	2015	PCA	●Neural network	●Classification of US kidney diseases	●Ac=95	● All information is squeezed into first 15 PCs	●Methodology is applied on smaller dataset
[16]	2015	PCA	●SVM	●Classification of breast cancer lesions	●Ac=94.4	● Select the most informative features	● High computational cost
[17]	2015	PCA	●Classification using supervised rule	●Detect infections on porcini mushrooms	●Pixels on PC1<10 ●Pixels on PC2<7.5	●Provide valuable information for detection ●Lead to the development of non-destructive monitoring system.	●Analyzed limited number of samples.
[18]	2015	●PCA ●PLS	●SVM	●Detection of Alzheimer's disease	●Ac=88.49 ●Se=85.11 ●Sp=91.27	● PLS based method has less computational time	● Unable to achieve the good classification accuracy using PCA+SVM
[19]	2015	●PCA	●SVM	●Fault classification in industrial process	●Ac=85	●Achieve good accuracy with less computational efforts	●Unable to achieve real time online classification
[20]	2014	●PCA	●Multilayer perceptron	●Classification of fatty and normal liver	●Se=.78 ●Sp=.85	● Decreases the computational complexity	● Unable to achieve the good classification accuracy using PCA
[21]	2014	●PCA	●KNN	●Classification of liver cancer	●Ac=86.29	●Improve the classification of liver cancer	● low time efficiency
[22]	2014	●PCA	●SVM	●Classification of EEG signals	●ANF=35.3 ●OCA= 71.7	●Maintains the classification performance while decreasing the features effectively.	●Little loss in the classification accuracy.
[24]	2013	●PCA	●Fuzzy neural network	●Detection of liver cancer	●Ac=95.8	●Effective time performance ●Effective classification performance	● Only binary classification problem is considered
[25]	2013	●PCA ●AP	●ANN	●Detection of breast cancer	●Ac=98.29	● Feasible system for classification	● Increased system complexity
[26]	2013	●PCA	●LDA ●SVM	●Face recognition	●ARR=97.4 ●ARR=97	●efficiently handles SSS problems	●Little loss of discriminatory information
[27]	2012	Sparse PCA		Face recognition	●Ac=81.67	●Less sensitive to noise and outliers	●Attains approximate solution
[28]	2011	●PCA ●LDA	●Neural network ●SVM	Detection of Alzheimer's disease	●Ac=96.7 ●Ac=89.52	●Compresses the information in two features	●Unable to achieve the good results using NN+PCA
[29]	2011	●PCA	●LSSVM	●Hepatitis diagnosis system	●Ac=96.12	●Compresses the information in ten features	●Increased computational complexity with increase in the value of c and σ of SVM
[30]	2010	●PCA	●LDA	● Face recognition	●PCs=82 ●ARR=0.72	●Higher recognition rate using less linear features.	●Characterization is done without enhancing the image artifacts.
			●SVM		●ARR=.94 ●PCs=190	●Allow robust reconstruction of the data	

Literature review based on non- linear feature transformation techniques

[31]	2016	●Graph kernel PCA	●SVM	●Classification of ADHD patients	●Ac=94.91 ●Se=93.22 ●Sp=96.94	●Discover the discriminative subnetworks ●Discover the discriminative brain regions	●One kind of feature and modality has used

[32]	2014	●PCA ●Kernel PCA	●SVM	●Vibration Analysis	●Ac=.95003	●Reduced kernel principal components have good effect of classification	● Unable to get better results using PCA
[33]	2011	NFRPCA	Similarity classifier	Classification of medical data sets	●Ac=72.2 (liver disorder) ●Ac=88.9 (Hepatitis) ●Ac=97.0 (Dermatology)	Performs better where non-linear combinations of highly overlapped features exist.	Increased computational time in case of some datasets
[34]	2009	●Kernel PCA	●SVM ●LDA	●Detection of Alzheimer's disease	●Ac=92.31	●Extract appropriate features for Alzheimer's disease detection	●Computationally expensive

*se-sensitivity, sp-specificity, Ac-accuracy, ERR=Equal error rate, PCs=principal components, ARR=Average recognition rate, PCA=Principal component analysis, PLS= Partial least square, AP= Apriori algorithm, LDA= Linear discriminant analysis, SVM= Support vector machine, ANN= Artificial neural network, ANF= Average number of features , LSSVM= Least square support vector machine.

Fast PCA is the numerical remedy provided in the literature to eradicate the first limitation described in section 3.1. and Sparse PCA is the proficient method provided in the literature to overcome the second limitation described in section 3.1 as it plays a vital role in the selection of the most significant features that captures the maximum variance in the feature space. Both these Sparse PCA and Fast PCA are the linear transformation techniques. However, it has been mentioned in many literature that these linear transformation techniques are not very sensitive to the complex datasets [38]. They cannot capture the discriminative information present in the features extracted from high-order statistics. To overcome this limitation various non-linear feature transformation techniques have mentioned in the literature such as kernel PCA, Non-linear fuzzy robust PCA, Graph kernel PCA etc which is depicted in Table 3.2. The problem of classification of focal hepatic lesions has been dealt by extracting various statistical based texture features. These statistical features include first order statistics, second order statistics and high-order statistics. In order to capture the discrimination capability related to higher order statistics, non-linear transformation techniques seems to be more suitable then linear transformation techniques. Kernel PCA can handle the wide variety of non-linearity as it has the ability to perform the transformation using different kernels. Moreover, our problem involves the classification of cancerous tissues. There is a huge amount of overlapping among the features in case of primary and secondary tumors. In most of the literature it has mentioned that non-linear fuzzy robust PCA use certain fuzzy rules to handle such type of cases. Hence, kernel PCA and Non-linear fuzzy robust PCA are used for the classification of hepatic focal lesions. Finally, a comparative analysis is also performed in the present work between linear and non-linear feature transformation techniques.

CHAPTER 4

MATERIALS AND METHODS

4.1. Data description

For the present work,111 B-mode US images consisting of 16 normal liver images,15 HCC,18 HEM,17 Cyst and 45 MET are taken into consideration out of which 24 images (2 Cyst, 3 HCC, 8 HEM, 11 MET) consist of atypical cases and 87 images (15 Cyst, 12 HCC, 10 HEM, 34 MET) consist of typical cases.

4.2. Methods

The subsequent methodology consists of the two parts (i) Supporting structure (ii) proposed methodology which is illustrated in figure 4.1.

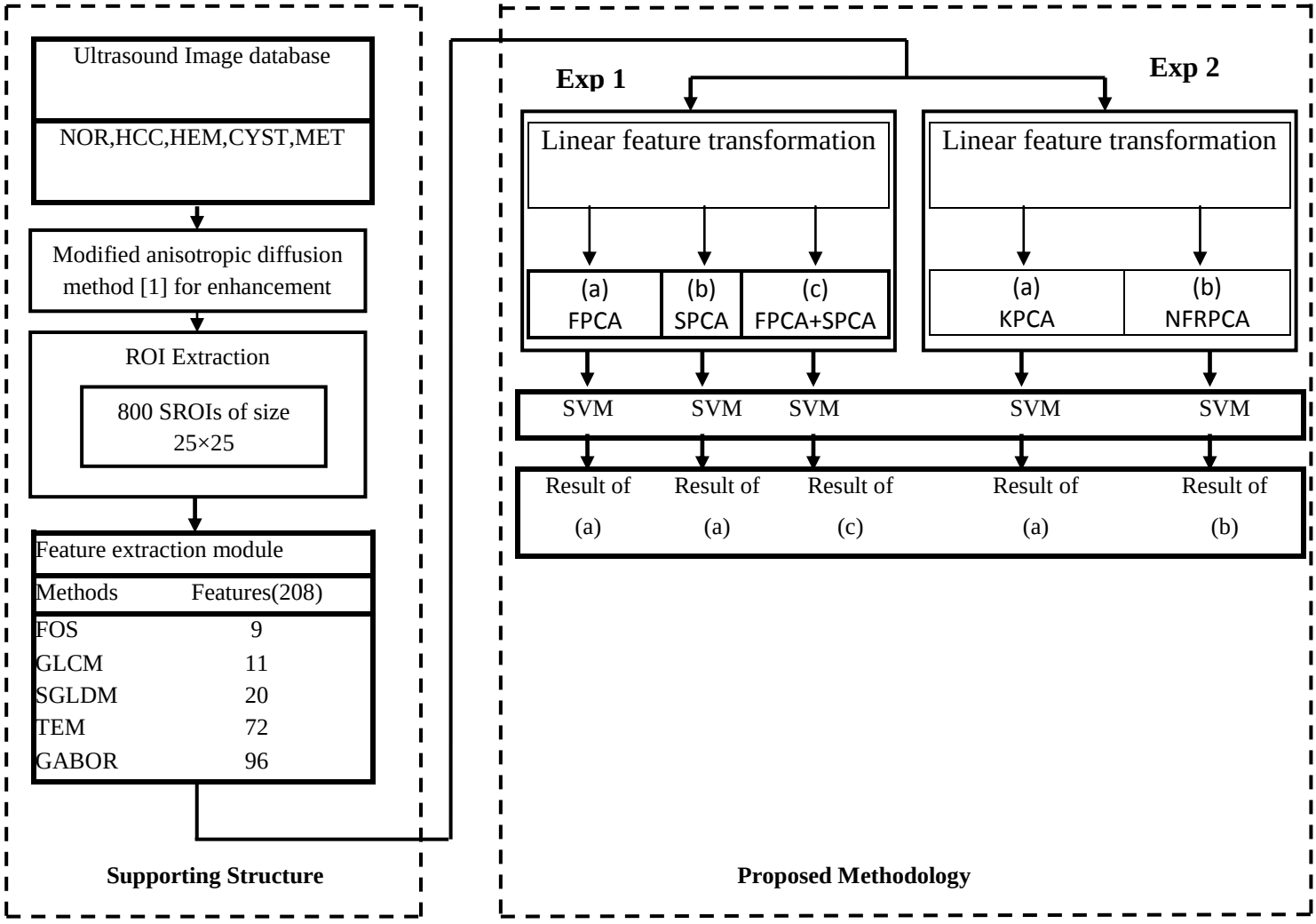

Fig 4.1: Flow chart of proposed methodology

4.2.1. Supporting structure

(i) Image enhancement

Low contrast and intrinsic speckle formation are the foremost curbs of acquired US images that inflicts the hepatic focal classification. To exterminate these curbs US images are enhanced by modified anisotropic diffusion method under the supervision of radiologists. Consequent enhancement strategy simultaneously perks up the visualization of acquired US images and intrinsic speckle noise and lead to the better hepatic focal classification.

(ii) Segmentation of region of interest

Marked portion of the samples for the identification of hepatic lesions is termed as region-of-interest (ROI). Segmentation of these marked ROIs into small number of non-overlapping ROIs is termed as segmented region-of-interest (SROI). Selection of SROI radically inflicts the classification of hepatic focal lesions. In the present work total of 800 non –overlapping SROIs have been segmented from enhanced images [6]. Out of these 800 SROIs 250 SROIs are used for training purpose and remaining 500 are taken for testing purpose. Distribution of these SROIs are given in Table 4.1.Sujana *et al*.[2] and Poonguzhali *et al*.[5] have considered SROIs of size 10×10 pixels for hepatic focal lesion classification. Yoshida *et al*.[3] has used SROI of size 60×60 pixels. Virmani *et al*.[7] has used SROI of size 32×32 pixels. Under the supervision of radiologist SROI of size 25×25 pixels is considered reasonable in terms of computational complexity and for computing consistent guesstimates for texture parameters.

Class	Training set	Validation set	Test set	Total
CYST	50	10	6	66
HCC	50	10	167	227
HEM	50	10	30	90
MET	50	10	125	185
NOR	50	10	172	232

Table 4.1: Distribution of SROI samples

(iii) *Feature extraction module*

Echogenicity and visual echotexture are two decisive factors for the detection of hepatic focal lesions. Contrast enhancement strengthens the features allied to echogenicity. Features allied to non-visual and visual echotexture are extracted using recurrently used analysis based on echotexture. Variety of textural features are computed based on first order statistics, second order statistics as well as higher order statistics which are depicted as (i) First order statistics (FOS) (ii) Grey level co-occurrence matrix (GLCM) (iii) Spatial grey level dependence matrix (GLRLM) (iv) Texture energy measures (TEM) (v) Gabor wavelet (GWT) [6]. Brief outline of the textural features cited above is presented in Table [4.1-4.5].

Above cited methods are elected on the basis of literature appraisal exclusively related to the hepatic focal lesions with US images. Different researchers have used different combinations of feature extraction methods. Balasubramanian *et al*.[4] has extracted total 88 textural features including GLCM, GLRLM, Law's and GWT. Poonguzhali *et al.*[5] have used GLCM, SGLDM, GWT and TEM for the classification of hepatic focal lesions. Virmani *et al.*[7] has extracted total 104 textural features including GLCM, SGLDM, GWT, TEM and spectral features. Hwang *et al*.[12] has extracted total 42 textural features FOS, GLCM, Law's and echogenicity. Mittal *et al*.[6] has extracted total 256 textural features from enhanced US images out of which 208 textural features are selected on the basis of sensitivity analysis using box-plot studies [6].

These 208 textural features become the foremost part of the supportive structure which provides a ethical support to the proposed methodology. This selected feature set may exhibit certain kind of redundancy. Therefore a dimensionality reduction methodology is designed to perform comparative evaluation to find out best PCA based transformation technique in the problem of hepatic focal lesions.

Table 4.2: Features extracted using First order statistics

Defination and Formula notations	Attributes	Name of the features	Features defination	Mathematical formulation
Estimate the texture features by computing the normalized grey level histogram of SROI image	1) Based on central tendency	Mean	Represents the average grey level value of SROI image	$\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} I(i,j)$
● I(i,j)= US image intensity ●I , j = Space variables ●M,N= Integers		Median	Measures the centre of intensities value in a SROI	$\mathbf{I(I,j)_{(n+1)/2}}$ for odd n $\mathbf{I(I,j)_{n/2} + I(I,j)_{(n+1)/2}}$ for even
●k= Number of occurrence of grey level ●L= Number of intensity values		Mode	Informs most frequently occurred intensity value in an image	$\mathbf{Max[\sum_0^{L-1} k]}$
● I_{max}= Largest grey level value ● I_{min}= Smallest grey level value	2) Based on diversity	Range	Measures the difference between largest and smallest grey level values.	$\mathbf{I_{max} - I_{min}}$
● $\bar{x}$ = Mean value		Variance	Shows the dispersion of normalized histogram about mean value	$\frac{1}{MN-1}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(I(i,j)-\bar{x}\right)^2$
		Standard deviation	Describe spread of grey level around the mean	$\sqrt{\mathrm{Var}_{\mathrm{FOS}}{}^2}$
●σ = Standard deviation	3)Based on shape	Skewness	Measures the symmetry and asymmetry of the grey level values around the mean	$\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\frac{I(i,j)-\bar{x}}{\sigma}\right)^3$
		●Kurtosis	Degree of peakedness of a distribution	$\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\frac{I(i,j)-\bar{x}}{\sigma}\right)^4$
●h(k) = Histogram	3)Based on histogram	●Entropy	Measures the information content in bits per pixel	$-\sum_{k=0}^{L-1} h(k)\,\ln\left[h(k)\right]$

Table 4.3: Calculated features using SGLDM

Defination and Formula notations	Attributes	Name of the features	Features definition	Mathematical formulation
●Second order statistical method. ● Calculate the textural features by considering spatial relationship of pixels in SROI image	1)Based on uniformity	Energy	Measures the uniformity of the texture.	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\{p(i,j)\}^2$
● N_g = Grey levels ● $p(i,j)$ =Probability density function	2)Based on inertia	Contrast	Measures the pixel intensity variations between pixel and its neighbor	$\sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g}\sum_{\substack{j=1 \\ \lvert i-j\rvert=n}}^{N_g} p(i,j) \right\}$
●n= Difference between intensity values i and j	3)Based on grey level linear dependence	Correlation	Measures the abrupt transitions in the image	$\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} \frac{(i-\mu_x)(j-\mu_y)\,p(i,j)}{\sigma_x\sigma_y}$
● μ_x, μ_y = Mean of rows and columns ● σ_x, σ_y = Standard deviation of rows and columns	4)Based on inverse difference moment	Homogeneity	Measures the uniformity	$\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} \frac{1}{1+(i-j)^2}\,p(i,j)$
	5)Based on second-order statistics	Entropy	Quantifies the level of randomness in the region of an image.	$\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} p(i,j)\log(p(i,j))$
μ is mean of μ_x and μ_y.	6)Based on second moment	Variance	Measures the heterogeneity	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-\mu)^2\,p(i,j)$
	7)Based on second moment	Maximum probability	Determine the most predominant pixel pair	$\max\,(p(i,j))$
	8)Based on second moment	Cluster shade	Measures the skewness of SGLDM matrix	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-\mu_x+j-\mu_y)^3\,p(i,j)$
	9)Based on grey-level histogram kurtosis	Cluster prominence	Measures the symmetry and asymmetry of the grey level values around the mean	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-\mu_x+j-\mu_y)^4\,p(i,j)$
	10)Based on spatial frequencies	Autocorrelation	Evaluates the linear spatial relationships between textural elements.	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i\,j)\,p(i,j)$
	11)Based on variability	Dissimilarity	Measures the local variability	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\lvert i-j\rvert\,p(i,j)$

*SGLDM= Spatial grey level dependence matrix, IMC= Information measure of correlation, IDN= Inverse difference normalized, IDMN= Inverse difference moment normalized.

Table 4.4: Features calculated using GLRLM

Defination and Formula notations	Name of the features	Features definition	Mathematical formulation
Method for extracting high-order statistical features by considering set of consecutive, pixels having the same grey level • n_r = Total number of runs in the image • j = Run length • $p(I, j)$= Probability density function	Short run emphasis	Measures short runs distribution	$\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{P(i,j)}{j^2}$
	Long run emphasis	Measures the distribution of long runs	$\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}P(i,j)*j^2$
	Low grey level run emphasis	Measures low grey level values distribution	$\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{P(i,j)}{i^2}$
	High grey level run emphasis	Measures the distribution of high grey level values	$\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}P(i,j)*i^2$
	Short run low grey-level emphasis	Measures short runs joint distribution and low grey level values.	$\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{P(i,j)}{i^2*j^2}$
	Short run high grey-level emphasis	Measures the joint distribution of short runs and high grey level values.	$\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{P(i,j)*i^2}{j^2}$
	Long run low grey-level emphasis	Measures the joint distribution of long runs and low grey level values.	$\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{P(i,j)}{i^2*j^2}$
	Long run high grey-level emphasis	Measures the joint distribution of long runs and high grey level values	$\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}P(i,j)*i^2*j^2$
	Grey-level non-uniformity	Measure the similarity of grey-level values throughout the image.	$\frac{1}{n_r}\sum_{i=1}^{M}\left(\sum_{j=1}^{N}P(i,j)\right)^2$
	Run-length non-uniformity	Measure the similarity of length of runs right through the image.	$\frac{1}{n_r}\sum_{j=1}^{N}\left(\sum_{i=1}^{M}P(i,j)\right)^2$
	Run-percentage	Measures distribution of runs and homogeneity of an image in a particular direction.	$\frac{n_r}{P(i,j)*j}$

*GLRLM= Grey level run length matrix

Table 4.5: Features extracted using TEM

Defination	Basic filters used	Convolution process	Filters	Steps involved in	Results

Feature extraction method	Defination in convolution	Mathematical formulation of Gabor filter	obtained after 3×3 convolution	Formula notations procedure	Mathematical formulation of filter parameters obtained after each step
Method of computing the textural features using local 2-D filter masks.	Gaussian , L3 =[1 2 1]	L3 * L3 = [1 4 6 4 1]	Level detector L5	Application of 2-D filter masks on US images	Generates 59 images having N rows and M columns
	Edge detector, E3 = [-1 0 1]	L3 * E3 = [-1 -2 0 2 1]	Edge detector E5	Performing windowing operation $\sum_{i=-7}^{i=7}\sum_{j=-7}^{j=7}\left\|I\left(x+i,y+j\right)\right\|$	Generates 59 TEM images.
	Laplacian, S3 =[-1 2 -1]	L3 * S3 = [-1 0 2 0 -1]	Spot detector S5	Combine similar features	Generates the set of 36 rotational invariant images
		E3 * S3 = [-1 2 0 -2 1]	Wave detector W5	Calculation of the statistical features from the rotational invariant texture energy images	Generates the set of 72 textural features.
		S3 * S3 = [1 -4 6 -4 1]	Ripple detector R5		

Table 4.6: Features calculated using Gabor wavelet

Gabor wavelet	●Method used for analyzing texture using multiscale texture analysis approach ●Formed from two components -Complex sinusoidal carrier -Gaussian envelope	$g(\tilde{x},\tilde{y}) = \overbrace{\left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left\{-\frac{1}{2}\left(\frac{\tilde{x}^2}{\sigma_x^2}+\frac{\tilde{y}^2}{\sigma_y^2}\right)\right\}}^{Gaussian\ envelop}.\ \overbrace{\exp i(2\pi u_0\tilde{x})}^{Complex\ sinusoidal}$	●U_0=Central frequency ●σ_x *and* σ_y= Shape factor of Gaussian envelope. ●$\hat{x}$ and $\hat{y}$ are axes. ● σ_u *and* σ_v are standard deviations	●$a = \left(U_h / U_l\right)^{-\frac{1}{P-1}}$ ●$\sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}}$

●**Steps involved in Gabor implementation**	●**Results obtained after each step**
1) Convolution of image with Gabor filter ● $I_k(x, y) = I(x, y) * g_k(\tilde{x}, \tilde{y})$ K= Number of Gabor filters	● Gabor outputs are generated corresponding to each Gabor filters
2) Implementation with 24 filters having 6 orientations (θ = 0, $\pi/6,\ \pi/3,\ \pi/2,\ 2\pi/3,\ 5\pi/6$) and four different scales (P = 0, 1, 2, and 3)	● Generation of 24 Gabor wavelets
3) Take $U_h = 0.4$ and $U_l = 0.05$ for implementing the Gabor wavelet	● Results in the Generation of total (5×4×6 =) 120 texture features

4.2.2. Proposed Methodology

Proposed methodology consists of the feature transformation techniques which basically maps the high-dimensional feature vectors into low dimensional feature vectors. Two linear and two non-linear feature transformation techniques are considered in this section as describe below:

(i) Linear feature transformation techniques: Linear feature transformation techniques extracts the new features by performing linear mapping on the original feature space by assuming that most of the features lie on or near the linear space of high-dimensional subspace.

In this section, two linear feature transformation techniques are discussed i.e Fast PCA and Sparse PCA along with the introduction of classic PCA.

Classic PCA: The basic of PCA is to extract the new features by transforming the original features. This transformation is performed linearly in classic PCA approach by capturing maximal variance of a feature set within a single variable which is termed as first principal component. Another highest variance that cannot be capture in first principal component lies in the second principal component and so on. The classic PCA approach is an eigen analysis that uses an orthogonal transformation to convert the possibly correlated feature set into a set of linearly uncorrelated feature set. An obvious advantage of this feature transformation technique is that it reduces the memory complexity. Conclusively form the literature, the main properties of classic PCA are that (i) it looks for the variance across the features as much as possible, and (ii) it retains the characteristics of original feature vectors after reconstructed as new set of feature vectors.

Figure 2 represents the steps involved in classic PCA. Let us assume we have a set of feature vector **X**, which is given by **X={**$x_1, x_2, x_3,$ **..., ** x_n **}** where 'n' is the number of feature vectors in this set. These feature vectors are calculated on the set A of SROI samples, which is expressed as $A= \{a_1, a_2, a_3,\ldots, a_d\}$, where '$d$' is the number of SROI samples. A term x_{mi} is also introduced which is the average calculated by taking the mean across i_{th} feature vector. This mean is subtracted from each of the data dimensions in order to calculate the central matrix X_C that is represented as:

$$X_C=\begin{bmatrix} x_{11}-x_{m1} & \cdots & x_{1n}-x_{mn} \\ \vdots & \ddots & \vdots \\ x_{d1}-x_{m1} & \cdots & x_{dn}-x_{mn} \end{bmatrix} \quad (1)$$

This d × n central matrix consists of features having zero mean. The n × n covariance matrix 'S' is calculated from this central matrix X_C, and expressed as:

$$S=\begin{bmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{bmatrix} \quad (2)$$

The components of S, denoted by S_{ij}, represents the covariance between the two random feature vectors given by x_i and x_j. Variance captured by these feature vectors represents their spread

across the zero mean. If these two feature vectors are uncorrelated to each other than their corresponding covariance will be zero S_{ij}= S_{ji}=0. Number of non-zero eigen values, represented as 'v' are calculated from this covariance matrix by satisfying the following characteristic equation:

$$\det(S - vI_n) = 0 \tag{3}$$

Where 'det' is the determinant of $(S - vI_n)$ and 'I' is n × n identity matrix given by:

$$I = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \tag{4}$$

After solving the characteristic equation total n non-zero eigen values, i.e, $v_1, v_2, v_3, \ldots, v_n$, are calculated which are present diagonally in the matrix D such as:

$$D = \begin{bmatrix} v_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & v_n \end{bmatrix} \tag{5}$$

All these 'n' calculated eigen values are sorted in the decreasing order of its value, out of which 'k' largest eigen values $(v_1, v_2, v_3, \ldots, v_k)$ are chosen. After selecting 'k' largest eigen values, their corresponding eigenvectors $\boldsymbol{w_k}$ are calculated using the equation:

$$\det(S - vI_n)\boldsymbol{w_k} = 0 \tag{6}$$

Let W_K is the n × k transformation matrix where k ≤ n which is expressed as:

$$W_K = \begin{bmatrix} w_{11} & \cdots & w_{1k} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nk} \end{bmatrix} \tag{7}$$

Each row of W_K represents the eigenvectors corresponding to the eigen values. The first row is the eigenvector analogous to the premier eigen value that is termed as first principal component whereas the second row represents eigenvector analogous to the next highest eigen value and termed as second principal component and so on. Variance captured by first principal component is maximum as compared to second one and continues likewise.

Original feature vector set **X** calculated on training SROI samples is now projected on these dominant eigenvector set W_K and the projection represents certain linear combinations which is expressed in matrix Z_{Train} as given below:

$$Z_{Train.} = \begin{bmatrix} x_{1i} - x_{mi} & \cdots & x_{1n} - x_{mn} \\ \vdots & \ddots & \vdots \\ x_{di} - x_{mi} & \cdots & x_{dn} - x_{mn} \end{bmatrix} \begin{bmatrix} w_{11} & \cdots & w_{1K} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nK} \end{bmatrix} \quad (8)$$

This projection minimizes the mean square error between the original feature set and transformed feature set and results into new feature values corresponding to each training SROI samples in matrix Z_{Train}. Afterwards, the original feature set of the test data is also reduced into new feature set by projecting testing data samples in the direction of principle components as obtained by the above procedure.

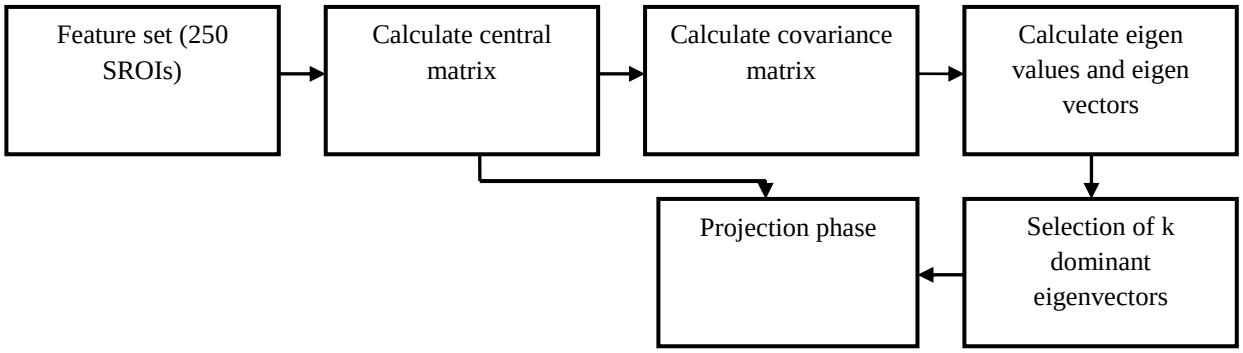

Fig 4.2: Flow chart representing steps of PCA

Fast PCA: Fast PCA, a modified version of classic PCA, is also based on eigen value decomposition method. The difference is eigen values are selected is this method on the basis of dimensionality parameters. Dimensionality is controlled by using two parameters (i) FE (ii) LE. FE refers to the index of the first eigen value to be kept and LE refers to the index of the last eigen value to be kept. Dimensionality parameters are chosen by following the given criteria:

These dimensional parameters should preserve 99% of the total information present in the feature set .i.e. LE = k should preserve 99% of the total information present in the feature set. This preserved energy is calculated using the given formula:

Retained information = (sum of selected eigen values/ sum of all eigen values) * 100

After meeting this criteria, k dominant eigen values represented as $(v_1, v_2, v_3, \ldots, v_k)$ are selected and all other eigen values are discarded. These selected eigen values are present diagonally in the matrix D such as:

$$D = \begin{bmatrix} v_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & v_k \end{bmatrix} \quad (9)$$

After calculating the k dominant eigen values their corresponding eigenvectors are calculated.

Sparse PCA: Sparse PCA is a modified version of classical PCA in a way that it adds sparsity constraint to the calculated eigenvectors. As in classical PCA it becomes difficult to interpret the original features because all of the principal components derived using this method are non-zero. In case of sparse PCA, one can interpret the subset of original features which are important for the classification purpose and these selected features capture most of the variance present in the feature space. In order to calculate the eigenvectors first central matrix is calculated which is explained in classic PCA and this d × n central matrix is given as:

$$X_C = \begin{bmatrix} x_{11} - x_{m1} & \cdots & x_{1n} - x_{mn} \\ \vdots & \ddots & \vdots \\ x_{d1} - x_{m1} & \cdots & x_{dn} - x_{mn} \end{bmatrix} \quad (10)$$

Calculated covariance matrix from the given central matrix is given as:

$$S = \begin{bmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{bmatrix} \quad (11)$$

After calculating the covariance matrix eigenvalues and eigenvectors are calculated using power method which is enlightened below:

Let us assume that $v_1, v_2, v_3 \ldots\ldots\ldots v_n$ are the eigenvalues of n×n covariance matrix S. v_1 is called the dominant eigenvalue of S if:

$$|v_1| > |v_i|, \quad i = 2, \ldots\ldots, n \quad (12)$$

The eigenvectors corresponding to v_1 is called the dominant eigenvectors $\boldsymbol{w_K}$ of v_1. Let us initialize the random vector $\boldsymbol{w_0}$ which is approximate to one of the dominating eigenvectors of S. Approximate eigenvectors $\boldsymbol{w_K}$ are then calculated as follow:

$$\boldsymbol{w_1} = S\, \boldsymbol{w_0} \quad (13)$$

$$\boldsymbol{w_2} = S\,(S\, \boldsymbol{w_1}) = S^2\, \boldsymbol{w_0} \quad (14)$$

$$\boldsymbol{w_3} = S\,(S\, \boldsymbol{w_2}) = S^2\, \boldsymbol{w_0} \quad (15)$$

⋮

$$\boldsymbol{w_K} = S\,(S^{k-1} w0) = S^{k}\,\boldsymbol{w_0} \quad (16)$$

After finding the dominant eigenvectors dominant eigen values v_K is calculated using **Rayleigh quotient given as:**

$$v_K = \frac{S\,\boldsymbol{w_K}\,.\,\boldsymbol{w_K}}{\boldsymbol{w_K}\,.\boldsymbol{w_K}}. \quad (17)$$

In order to impose the sparsity, combination of L_1 norm and L_2 norm [57] are used and yield the eigenvectors as:

$$\boldsymbol{w}^{K+1} = (1-\gamma)||\boldsymbol{w_K}||_2 + \gamma\,||\,\boldsymbol{w_K}\,||_1 \quad (18)$$

Equation (17) to (19) is repeated until it meets certain convergence criteria as follow:

$$\frac{v^{K+1} - v^{K}}{v^{K}} < \emptyset, \quad \emptyset = \frac{|v^{K+1}|}{v_K} \quad (19)$$

Where γ is sparsity controlling factor $\gamma \in [0,1]$. $(1-\gamma)||\boldsymbol{w_K}||_2$ factor in equation (18) minimizes the reconstruction error. L_1- norm penalty adds the non-zero loadings to the sparse matrix which is controlled by the sparsity controlling factor γ. In case of classical PCA $\gamma = 0$ whereas in case of sparse PCA $\gamma = 1$ yields the features having maximal variance. Figure 4 presents the steps involved in Sparse PCA.

(ii) Non- Linear feature transformation techniques: non-linear feature transformation techniques extracts the new features by performing non-linear mapping on the original feature space by assuming that most of the features lie on or near the non-linear space of high-dimensional feature space. In this section, two non-linear feature transformation techniques are discussed: (1) Kernel PCA (2) Non-linear fuzzy robust PCA.

Kernel PCA: Kernel PCA is non-linear expansion of classical PCA. KPCA performs the non-linear mapping using the similarity kernel function φ^0 which maps the original n-dimensional features into high dimensional feature space, P, by creating non-linear combinations of original features. The non-linear mapping function is defined as:

$$\boldsymbol{\varphi^0}: \mathbf{R^n} \rightarrow \mathbf{P} \quad (20)$$

$$\mathbf{X} \rightarrow \boldsymbol{\varphi}^{\mathbf{0}}(\mathbf{X}) \quad (21)$$

Then the classical PCA is performed in this P space. Figure 5 presents the steps involved in Kernel PCA. In order to perform the non-linear mapping three types of kernel functions are used which are depicted below:

● *Linear*:

$$k(x_i, x_j) = x_i{}^T x_j \quad (22)$$

● *Polynomial*:

$$k(x_i, x_j) = (x_i{}^T x_j + c)^d \quad (23)$$

●*Radial basis kernel*:

$$k(x_i, x_j) = \exp^{(-||x_i - x_j)^2/\sigma^2} \quad (24)$$

where k is n × n kernel matrix. Each element of this kernel matrix is the inner product of two feature vectors x_i and x_j in high dimensional feature space P which is given as:

$$k(x_i, x_j) = \varphi^0(x_i).\ \varphi^0(x_j). \quad (25)$$

Non-linear Fuzzy Robust PCA: This particular PCA is based on some fuzzy rules which basically remove the highly overlapped non-redundant features in more realistic way.

Non-linear Fuzzy Robust PCA algorithm:

Step 1: Initialize the statistical initial parameters : (i) learning coefficient $a_0 \in (0,1)$ (ii) soft threshold value τ which is a small positive number (iii) fuzzy variable U_1 which plays a vital role in removing the redundant features (iv) set the parametric value c.

Step 2: Set the iteration count r=1 and iteration bond R.

Step 3: Calculate the learning rate $a_T = a_0 \left(1 - \frac{r}{R}\right)$. (26)

Step 4: Repeat step 3 until r < R.

Step 5: Set counter i = 1and $\rho = 0$ parameter required to update the soft threshold value.

Step 6: initialize the random eigenvector $\boldsymbol{w}$ to certain non-zero vector.

Step 7: Initialize $\boldsymbol{w}^{old} = \boldsymbol{w}$

Step 8: project $\boldsymbol{w}^{old}$ on each feature vector such that:

$$Y = (\boldsymbol{w}^{old})^{T} * x_i \quad (27)$$

Step 9: Update the eigenvectors using the formula given as:

$$\boldsymbol{w}^{new} = \boldsymbol{w}^{old} + a_T \, \partial(x_i) \, (x_i \; E_0(x_i)^T \; \boldsymbol{w}^{old} \; G + E_0(x_i) \; f(Y)) \quad (28)$$

where f(Y) is a non linear function that depends on function G as

$$G = \frac{d}{dY}(f(Y) \quad (29)$$

$$f = \tanh(c * Y) \quad (30)$$

$\partial(x_i)$ is the parametric function which is controlled using fuzzy variable U_1, soft threshold τ and energy function $E_0(x_i)$ which basically measures the reconstruction error and given as:

$$\partial(x_i) = \left[\frac{1}{1 + \left[\frac{E_0(x_i)}{\tau} \right]^{1/U_1 - 1}} \right]^{U_1} \quad (31)$$

Energy function $E_0(x_i)$ is given as:

$$E_0(x_i) = x_i \; - \; w^{old} \; f(Y) \quad (32)$$

Step 10: Update soft threshold and parameter required to update the soft threshold value.

$$\rho = \rho + E_0(x_i) \quad (33)$$

$$\tau = \rho / K \quad (34)$$

Step 10: Repeat step 8- 10 until i = number of eigenvectors required (K) such that K< n (number of feature vectors).

Chapter 5

Performance Estimation

5.1 Performance parameters

Since the problem involves the detection of hepatic focal lesions, two performance parameters Sensitivity and specificity are used in order to measure the performance. These parameters are described below:

5.1.1 Sensitivity

Sensitivity is the statistical measure by which hepatic lesions can be detected correctly. Its mathematical representation is depicted below:

Sensitivity= TP/TP+FN ×100%

Where, TP represents the proportion of hepatic lesions which are correctly classified by the test and FN represents the number of cancerous tissues which are falsely detected as normal liver tissues.

5.1.2 Specificity

Specificity is the statistical measure by which normal liver tissues can be detected correctly. Its mathematical representation is depicted below:

Specificity= TN/TN+FP ×100%

Where, TN represents the proportion of normal liver tissues which are correctly classified by the test and FP represents the number of normal liver tissues which are falsely detected as hepatic lesions.

5.1.3 Performance parameters of Estimation matrix

The matrix used to estimate the performance of a classifier in terms of individual classification accuracy and overall classification accuracy is called confusion matrix. Particular confusion matrices are shown in Table 5 and Table 7 in context of characterization of focal hepatic lesions.

Individual classification accuracy= $\frac{\textbf{\textit{Correctly classified focal lesions of individual class}}}{\textbf{\textit{Total focal lesions of individual class}}}$

Overall classification accuracy= $\frac{\textbf{\textit{Sum of Correctly classified focal lesions of all individual classes}}}{\textbf{\textit{Sum of total focal lesions of all individual classes}}}$

5.2 Classification Module

5.2.1 *Multiclass SVM*

There exist a problem, where we have to deal with more than two classes such as in detection of focal hepatic lesions we have five classes HCC,MET,HEM,NOR,CYST. So, in such cases we use one-against-all methods for SVM multiclass classification. This method builds the K SVM models where K denotes the number of classes.

Chapter 6

Results and Discussions

6.1 Experimental design

In this section two set of experiments are executed for the classification of hepatic focal lesions. In section 1 linear feature transformation techniques are performed whereas in section 2 non linear feature transformation techniques are performed.

6.1.1 Experiment 1

This experiment is performed to compare the most efficient and frequently used linear transformation techniques. In this experiment Fast PCA and Sparse PCA are the selected linear transformation techniques. By considering the advantages of these two techniques one more experiment is performed in this section in which both of these transformation techniques are blend together.

6.1.2 Experiment 2

Second experiment is performed to compare the most efficient and frequently used non-linear transformation techniques. In this experiment Kernel PCA and Non-linear fuzzy robust PCA are the selected non- linear transformation techniques.

6.2 Results and discussions

Results acquired from these two experiments are illustrated in subsequent subsections:

6.2.1 Experiment 1

Experiment 1 is performed in three sections.

(a) Fast PCA

The use of Fast PCA for reducing the high dimensional feature space is based on fixed-point algorithm. First we examine the number of principal components (PCs) required in order to conserve the 99% of the total information present in the feature set. The energy for fixed number of PCs is computed as sum of the successive eigen values until the required PCs by the sum of all eigen values. So the number of PCs is selected by scrutinizing that how the energy varies by

varying the number of eigen values. Finally, we fix our investigation that first 20 PCs preserve the 98% of the total energy present in the feature space. Further, we analyze that how the classification accuracy gets effected on varying the number of principal components. Figure 6.1 depicts the effect of principal components on the classification accuracy. First we analyze the classification accuracy using first PC, then with second PC, and so on. Best classification accuracy of 89% is accomplished by squeezing the information in first 20 PCs that preserves the 98% of the total energy of the feature space also. On adding the 21th PC classification accuracy decreased below 89%. Finally, we set aside the number of PCs to 20 for the classification of hepatic focal lesions.

Further, Table 6.1 depicts that, that there is a very good detection of cyst with the accuracy of 100%. Further, there is a very good detection of the solid malignant tumors with the accuracy of 91.6% for HCC and 98.4 % for MET. Results obtained for the detection of solid benign tumors and normal liver tissues are satisfactorily with the accuracy of 76.6% and 81.9%.

(b) Sparse PCA

With Sparse PCA, we first investigate the track of most important PCs on the basis of captured cumulative variance. Next, sparse PCs are evaluated using equation (19) with a wide range of values for the sparsity controlling parameter γ. An appropriate value of γ is obtained on varying the value of γ in the range of 1 to 2 and the most optimum results are obtained by taking the value of γ as 1. After optimizing the value of γ, a wide range of values for the cardinality controlling parameter C is determined which basically determines the number of non-zero components in a each particular PC. An appropriate value of C is obtained by varying its value in the range of 1 to 5 and the most optimum value is achieved by inspecting the captured cumulative variance by each PCs. It has observed that number of non-zero components in each sparse PC exhibit more than 90% variance by taking the value of C as 2. Figure A depicts the effect of C on cumulative variance. Further, we analyze that how the classification accuracy gets effected on varying the number of PCs. Figure 6.1 depicts the effect of Sparse PCs on the classification accuracy. First we analyze the classification accuracy using first sparse PC, then with second sparse PC, and so on. Best classification accuracy of 94% is achieved by squeezing the information in first 10 PCs. On adding the 11th PC classification accuracy decreased below

94%. Finally, we set aside the number of sparse PCs as 10 for the classification of hepatic focal lesions.

Performance of SPCA can be analyzed from Table 6.1, which depicts that 100 % detection occurs in case of Cyst. In case of solid malignant tumors, secondary cancer is detected perfectly with the accuracy of 100% in comparison with primary cancer who's only 144 SROIs are detected correctly out of 167 with the accuracy of 86.2%. There is a good improvement in the detection of solid benign tumors and normal liver tissues with the accuracy of 93.3% by detecting 28 correctly SROIs out of 30 and 97% by detecting 167 correctly SROIs out of 172.

(c) Fusion of *Fast PCA and Sparse PCA:*

Using Fast PCA we observed that maximum accuracy is achieved by squeezing all the information in first 20 principal components. But after using Fast PCA we failed to gather the information that which features contribute maximum in the characterization of focal hepatic lesions. Hence to mitigate this problem and to decrease the time complexity Fast PCA and Sparse PCA have used successively in this section of experiment 1. First the best 20 eigenvectors are selected on the basis of calculated total energy criteria. Further, sparsity is added to these PCs by taking the optimum value of γ as 1. Next a suitable value of C is taken as 4 here on the basis of which number of non-zero components in each sparse PC exhibit more than 90% variance. Figure B depicts the effect of C on cumulative variance. Further, classification accuracy is examined on varying the number of sparse PCs. Figure 6.1 depicts the effect of Sparse PCs on the classification accuracy. First we analyze the classification accuracy using first sparse PC, then with second sparse PC, and so on. Best classification accuracy of 91% is achieved by squeezing the information in first 5 PCs. On adding the 6th PC classification accuracy decreased below 91%. Finally, we fixed our investigation by taking first 5 PCs into account.

It can be seen from Table 6.1, that by fusing these two techniques there is a very good improvement in the detection of primary cancer tissues by detecting 161 SROIs correctly out of 167. Results obtained for the detection of solid benign tumor and normal liver tissues are satisfactorily with the accuracy of 80% and 81.9%. In case of secondary cancer 123 SROIs are detected correctly out of 125 by exhibiting the accuracy of 98.4%.

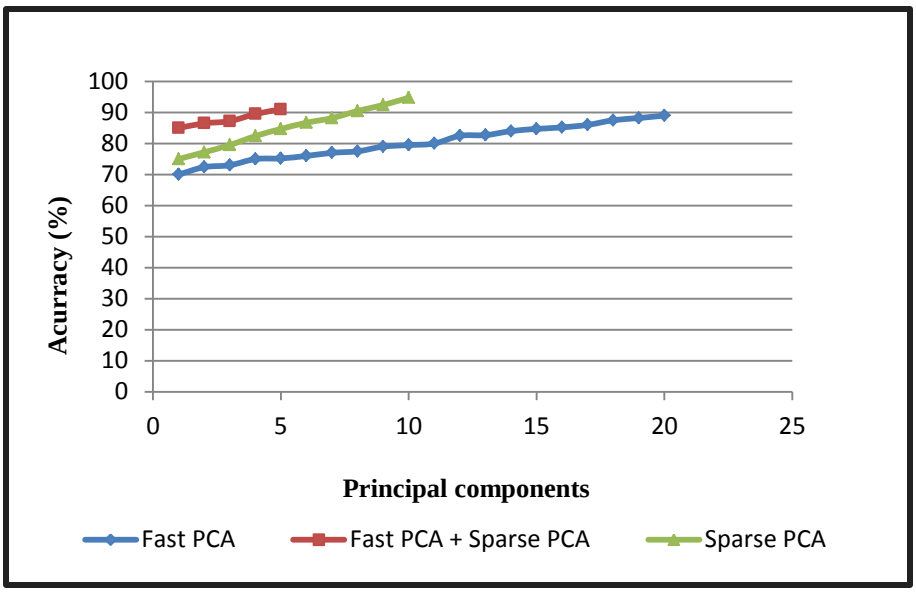

Fig.6.1. Effects of PCs on overall classification accuracy

6.2.1.1 *Comparative performance evaluation*

Table 6.1 shows the combined test results of these linear feature transformation techniques in form of confusion matrix along with SVM classifier. Results reveals that individual classification accuracy for CYST vs all is 100%, is almost equal for all types of feature transformation techniques. For the detection of NOR liver tissues, Sparse PCA performs better with 167 correctly classified SROIs out of 172 thus yielding the classification accuracy of 97.09% which is 15.19% higher than Fast PCA and (Fast PCA+ Sparse PCA). For the detection of primary cancerous tissues i.e. HCC, (Fast PCA+ Sparse PCA) yields very good results by exhibiting the classification accuracy of 96.4% with 161 correctly classified SROIs out of 167 thus yielding the 4.8% higher accuracy then Fast PCA and 10.2% higher accuracy then Sparse PCA. Secondary cancerous tissues i.e. MET are detected preeminently using sparse PCA as 28 SROIs are detected correctly out of 30 SROIs by exhibiting the classification accuracy of 93.3% which is 16.7% higher accuracy then Fast PCA and 13.3% higher accuracy then (Fast PCA+ Sparse PCA).

Table 6.1: Comparative analysis of linear feature transformation methods in terms of

Feature transformation method	Testing set	Class code					Individual classification accuracy	Overall classification accuracy
		CYST	HCC	HEM	MET	NOR		
	CYST	6	0	0	0	0	100	89.2
FPCA	HCC	4	153	4	4	2	91.6	
	HEM	0	2	23	5	0	76.6	
	MET	0	1	1	123	0	98.4	
	NOR	9	3	8	11	141	81.9	
SPCA	**CYST**	**6**	**0**	**0**	**0**	**0**	**100**	**94**
	HCC	**0**	**144**	**0**	**23**	**0**	**86.2**	
	HEM	**0**	**0**	**28**	**2**	**0**	**93.3**	
	MET	**0**	**0**	**0**	**125**	**0**	**100**	
	NOR	**2**	**3**	**0**	**0**	**167**	**97.09**	
FPCA+SPCA	CYST	6	0	0	0	0	100	91
	HCC	3	161	1	2	0	96.4	
	HEM	0	0	24	6	0	80	
	MET	0	1	1	123	0	98.4	
	NOR	8	11	3	9	141	81.9	

*FPCA=Fast Principal component analysis, SPCA= Sparse principal component analysis

Figure 6.2 depicts their comparative evaluation in terms of overall classification accuracy. It can be clearly seen from the figure that Sparse PCA outperforms the another two techniques by yielding the overall classification accuracy of 94% which has 4.8% higher accuracy than Fast PCA and 3 % higher accuracy then (Fast PCA + Sparse PCA).

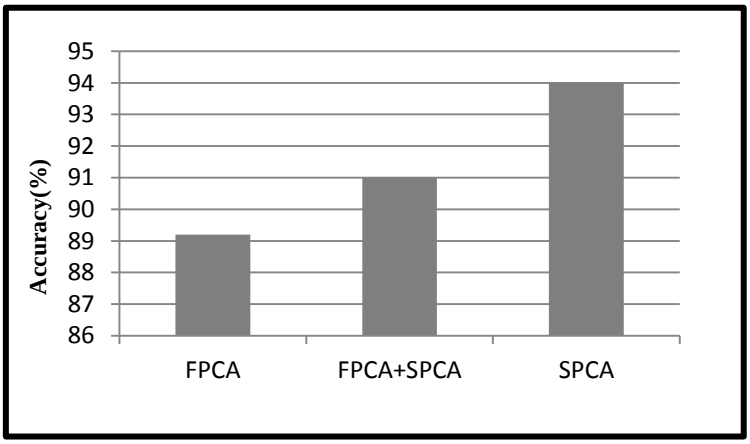

Fig: 6.2 Comparative analyses of linear transformation techniques

The comparative evaluation of these two techniques is also presented in Table 6.2, in terms of two medically imperative performance parameters, i.e. sensitivity and specificity. Table 6.2,

clearly demonstrates that there is a very good detection of the abnormal lesions with the sensitivity of 100% using Sparse PCA and (Fast PCA+ Sparse PCA) which is 0.7% higher than the detection using Fast PCA. Further, 97% of the normal liver tissues are detected correctly using sparse PCA and (Fast PCA + Sparse PCA) which shows a very high deviation of 15.1% in comparison with Fast PCA and (Fast PCA + Sparse PCA). Among fluid lesions and solid lesions there is 100 % detection using all the three techniques. Among primary and secondary cancer tissues (Fast PCA + Sparse PCA) outperforms the other two techniques with the sensitivity of 98.7% which is 1.3% higher than the Fast PCA and 12.5% higher than Sparse PCA. Among solid malignant and solid benign tumors Sparse PCA outperforms the other two with the sensitivity of 100% which is 1.8% higher than the Fast PCA and 07% higher than (Fast PCA + Sparse PCA).

Table 6.2, also depicts that Fast PCA exhibits less computational time of 7.54 sec in comparison with Sparse PCA which exhibit the computational time of 9.01 sec. Results also reveals that fusion of the two techniques further reduce the time by 1.12% in comparison with Sparse PCA.

Table 6.2: Comparative evaluation in terms of imperative performance parameters

Feature transformation technique	Computational time	Class code	Sensitivity (%)	Specificity
FPCA	7.54	Abnormal/Normal	99.3	81.9
SPCA	9.01		100	97
FPCA+SPCA	7.89		100	81.9
FPCA	7.54	Fluid Lesion/Solid Lesion	100	98.6
SPCA	9.01		100	100
FPCA+SPCA	7.89		100	99
FPCA	7.54	Primary Cancer/Secondary Cancer	97.4	99.1
SPCA	9.01		86.2	100
FPCA+SPCA	7.89		98.7	99.1
FPCA	7.54	Solid Malignant/Solid Benign	98.2	76
SPCA	9.01		100	93.3
FPCA+SPCA	7.89		99.3	80

*FPCA=Fast PCA, SPCA= Sparse PCA

6.2.2. Experiment 2

Experiment 2 is performed in two sections.

(a) Kernel PCA

This section is engrossed in the selection of best kernel function on the basis of which best PCs are extracted in order to characterize the hepatic focal lesions. Appropriate selection of kernel function implicit the performance of this transformation technique. In this section, best kernel mapping function is selected among two different kernel functions i.e. polynomial kernel and Gaussian kernel. For Gaussian kernel, 'σ' is the most imperative parameter which is defined by eq. (17). This parameter basically separates the two classes in most appropriate manner. Using Gaussian kernel mapping is performed by varying the value of σ in the range of 0 to 4 and optimal results are attained by taking σ=.031. Similarly in case of polynomial kernel 'd' is the most significant parameter defined by eq. (16). Using polynomial kernel mapping is performed by varying the value of d in the range of 1 to 5 and optimal results are attained by taking d = 1. Further, the effect of PCs is analyzed on the overall classification accuracy by varying its number. First we kept the track of PCs at 1 and observed the accuracy, then set the track of PCs at 2, 3, 4 and so on. Figure 6.4, depicts the effect of PCs on the overall classification accuracy. Finally we winded up our search by setting the track of PCs at 10.

It can be seen from Table 6.3, polynomial kernel PCA yields a very good results for the detection of MET and NOR liver tissues by exhibiting the accuracy of 100% and 98.4%. Results obtained for the detection of Cyst, HCC and HEM are satisfactorily as only 3 SROIs are detected correctly out of 6 in case of Cyst, 148 SROIs are detected correctly out of 167 in case of HCC and 22 SROIs are detected correctly out of 30 in case of HEM. Using Gaussian kernel PCA there is an improvement in the detection of Cyst which exhibits 100% accuracy by detecting all 6 SROIs correctly. Results obtained for the detection of solid malignant tumors are satisfactorily with the accuracy of 82.6% and 88.8%. Solid benign tumors and NOR liver tissues are detected very poorly with the accuracy of 70% and 37.7% only.

Comparative evaluation of polynomial and Gaussian kernel: Table 6.3 depicts the comparative evaluation between these two kernel functions in terms of overall classification accuracy. Results reveal that polynomial kernel yields the best results by exhibiting the overall classification

accuracy of 91.2% by squeezing the information in first 10 PCs. The yielded accuracy is 22.2% higher than the accuracy of Gaussian kernel which is only 69%.

(b) Non-linear fuzzy robust PCA:

In this section, first we examine the value of fuzzy variable U_1 which basically controls the energy function in order to minimize the reconstruction error. Value of U_1 is varied in the range of 1 to 4 and minimum reconstruction error is achieved by taking U_1 at 1.2. Figure 6.3, depicts the reconstruction error by varying this fuzzy variable.

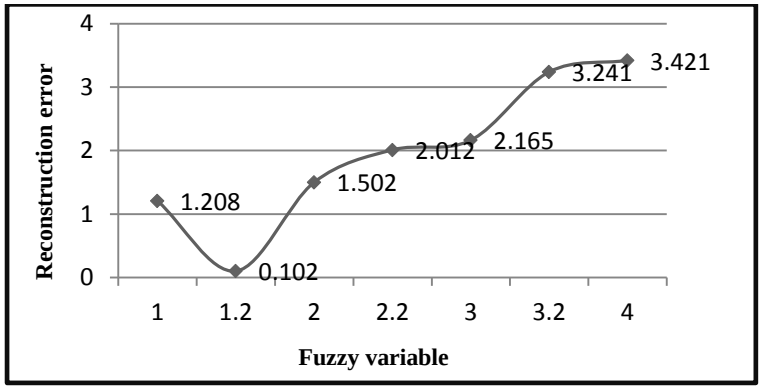

Fig: 6.3 Effect of fuzzy variable on reconstruction error

Next, we analyze the overall classification accuracy by taking different number of PCs. First we kept the track of PCs at 1 and observed the accuracy, then set the track of PCs at 2, 3, 4 and so on. Figure 6.4, depicts the effect of PCs on the overall classification accuracy. Finally we winded up our investigation by setting the track of PCs at 10.

It can be observed from the figure 6.4 that in case of polynomial kernel PCA overall classification accuracy of 91.2% is achieved by squeezing the information in first 10 PCs. In case of Gaussian kernel PCA overall classification accuracy is only 69% which is achieved by squeezing the information in first 10 PCs after that it goes on decreasing after further addition of the PCs. In case of NFRPCA overall classification accuracy of 76% is achieved by constricting the information in first 10 PCs. After the addition of further PCs accuracy goes on decreasing.

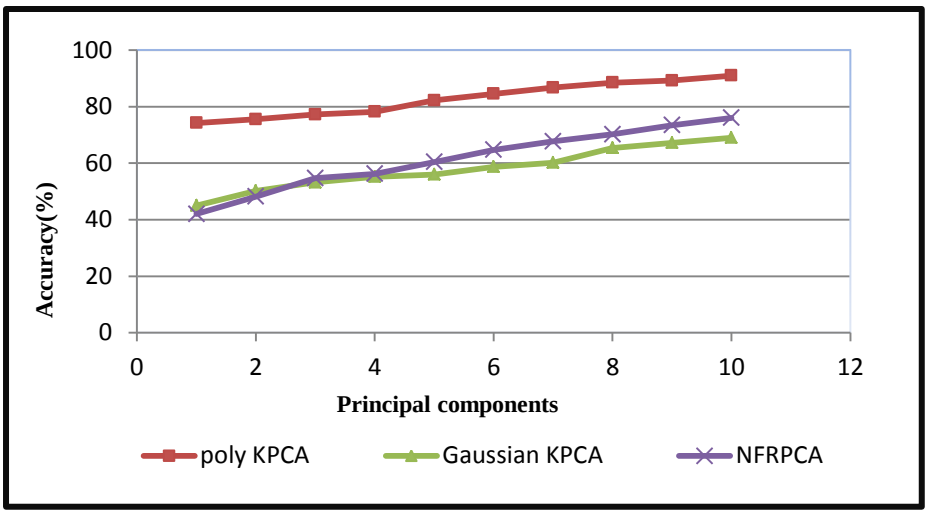

Fig: 6.4 Effect of PCs on overall classification accuracy

Table 7.3, reveals that there is 100% detection in case of Cyst. There is a vast improvement in the detection of solid malignant tumors with the accuracy of 99.4% for HCC and 97.6% for MET. Results obtained for the detection of solid benign tumors are satisfactorily with the accuracy of 70%. The NOR are liver tissues detected very poorly with the accuracy of only 37.7%.

6.2.2.1. *Comparative performance evaluation*

Table 6.3 depicts the combined test results of these two non-linear feature transformation techniques in terms of overall classification accuracy. We have observed that polynomial kernel PCA has outperformed the Gaussian kernel PCA, so in this section comparative analysis has done between polynomial kernel PCA and non-linear fuzzy robust PCA. In case of Cyst NRFPCA yield the highest classification accuracy of 100% which is 50% higher than polynomial kernel PCA. For the detection of primary cancerous tissue i.e. HCC, NRFPCA yields very good results by exhibiting the classification accuracy of 99.4% with 166 correctly classified SROIs out of 167 thus yielding the 10.8% higher accuracy then polynomial kernel PCA. For the detection of secondary cancerous tissue i.e. MET, polynomial kernel PCA yields terrific results by exhibiting the classification accuracy of 100% which is 2.4% higher than the NRFPCA. In case of HEM, results obtained by both the methods are satisfactory as only 22 SROIs are detected correctly out of 30 SROIs in case of polynomial kernel PCA by yielding the classification accuracy of 73.3% which is 3.3% higher than NRFPCA. In case of NOR liver

tissues polynomial kernel PCA yields a very good results by exhibiting the classification accuracy of 91.8% with the detection of 158 SROIs out of 172 SROIs in comparison with NRFPCA which yields the classification accuracy of only 37.7% by detecting only 65 SROIs out of 172 SROIs.

Table 6.3: Comparative analysis of non-linear feature transformation methods in terms of confusion matrix

Feature transformation method	**Testing set**	**Class code**					**Individual classification accuracy**	**Overall classification accuracy**
		CYST	**HCC**	**HEM**	**MET**	**NOR**		
Poly KPCA	**CYST**	**3**	**3**	**0**	**0**	**0**	**50**	**91.2**
	HCC	**0**	**148**	**10**	**8**	**1**	**88.6**	
	HEM	**0**	**0**	**22**	**8**	**0**	**73.3**	
	MET	**0**	**0**	**0**	**125**	**0**	**100**	
	NOR	**0**	**13**	**1**	**0**	**158**	**91.8**	
Gaussian KPCA	CYST	6	0	0	0	0	100	69
	HCC	9	138	2	12	6	82.6	
	HEM	1	3	14	12	0	46.6	
	MET	8	1	5	111	0	88.8	
	NOR	2	17	0	77	76	44.1	
NRFPCA	CYST	6	0	0	0	0	100	76
	HCC	0	166	1	0	0	99.4	
	HEM	1	0	21	8	0	70	
	MET	0	0	1	122	2	97.6	
	NOR	1	105	1	0	65	37.7	

*KPCA=Kernel Principal component analysis, Poly=polynomial, NRFPCA=Non-linear fuzzy robust Principal component analysis

Figure 6.5 depicts the comparative analysis of all these techniques in terms of overall classification accuracy.

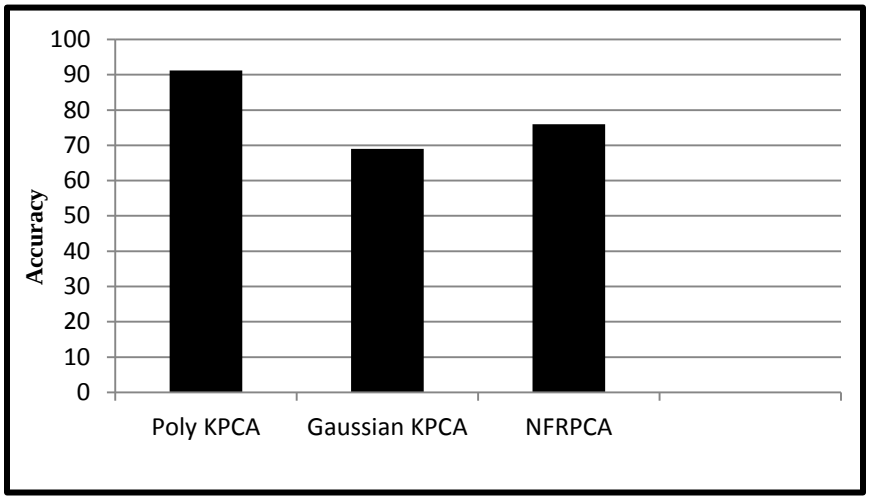

Fig: 6.5 Comparative analyses of non-linear feature transformation techniques

The comparative evaluation of these two techniques is also presented in Table 6.4, in terms of two medically imperative performance parameters, i.e. sensitivity and specificity. Table 6.4, clearly demonstrates that there is a very good detection of the abnormal lesions with the sensitivity of 99.6% using polynomial kernel PCA which is 0.3% higher than the detection using NFRPCA. Further, 91.8% of the normal liver tissues are detected correctly using polynomial kernel PCA which shows a very high deviation of 54.1% in comparison with NFRPCA. Among fluid lesions and solid lesions NFRPCA outperforms the polynomial kernel PCA with the sensitivity of 100% which is 50% higher than the polynomial kernel PCA. Among primary and secondary cancer tissues NFRPCA outperforms the polynomial kernel PCA with the sensitivity of 100% which is 5.2% higher than the polynomial kernel PCA. Among solid malignant and solid benign tumors NFRPCA outperforms the polynomial kernel PCA with the sensitivity of 99.3% which is 2.8% higher than the polynomial kernel PCA. Table 6.4, reveal that polynomial kernel PCA exhibits less computational time of 15.9 sec in comparison with other two.

Table 6.4: Comparative evaluation in terms of imperative performance parameters.

Feature transformation technique	Computational time	Class code	Sensitivity (%)	Specificity (%)
Poly KPCA	15.9	Abnormal/Normal	99.6	91.8
Gaussian KPCA	20.0		97.8	44.1
NFRPCA	135.6		99.3	37.7
Poly KPCA	15.9	Fluid Lesion/Solid Lesion	50	100
Gaussian KPCA	20.0		100	93.5
NFRPCA	135.6		100	99.6
Poly KPCA	15.9	Primary Cancer/Secondary Cancer	94.8	100
Gaussian KPCA	20.0		92	99.1
NFRPCA	135.6		100	100
Poly KPCA	15.9	Solid Malignant/Solid Benign	96.5	73.3
Gaussian KPCA	20.0		97.3	48.2
NFRPCA	135.6		99.3	72.4

*KPCA= Kernel PCA, NFRPCA= Non-linear fuzzy robust PCA

Chapter 7

Conclusion

This work presents the comparative performance analysis of linear and non-feature transformation techniques. The CAD system plays a vital role for radiologists in today's world, therefore it is essential to make this system efficient and more accurate. Use of feature transformation techniques make it more efficient and useful. High-dimensional features degrade the performance of classifier. Hence, all these feature transformation methods reduce the dimensionality of feature vectors in order to overcome the problem of over-fitting and curse of dimensionality. Obtained results escort to the conclusion that sparse PCA outperforms all other feature transformation techniques by exhibiting the overall classification accuracy of 94%. Results also reveal that Sparse PCA show 100 % detection among normal and focal liver tissues, fluid and solid lesions, solid benign and solid malignant tumors. Also, fusion of two techniques Fast PCA and Sparse PCA show 100 % detection among normal and focal liver tissues, fluid and solid lesions and sensitivity of 99.3% indicates a successful discrimination among primary and secondary cancer tissues. Further, NFRPCA performs very well in distinguishing primary cancer from secondary cancer, solid benign from solid malignant tumors and fluid lesions from solid lesions. Sensitivity of 99.6% using polynomial kernel PCA shows that there is very good detection of focal liver tissues from that of normal liver tissues. Results also reveal that Use of Fast PCA makes the system very efficient in case of time complexity.

Future work needs to focus on the ways of reducing this computational time complexity.