Computer-Assisted Algorithms for Ultrasound Imaging Systems

R Bharath

A Dissertation Submitted to
Indian Institute of Technology Hyderabad
In Partial Fulfillment of the Requirements for
The Degree of Doctor of Philosophy

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Department of Electrical Engineering

July 2018

Declaration

I declare that this written submission represents my ideas in my own words, and where ideas or words of others 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.
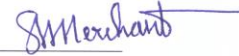
(Signature)

(R Bharath)

(Roll No. EE13P0007)

Approval Sheet

This thesis entitled "Computer-assisted Algorithms for Ultrasound Imaging Systems" by R Bharath is approved for the degree of Doctor of Philosophy from IIT Hyderabad.
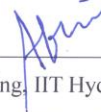
Professor Devendra Jalihal, Department of Electrical Engineering, IIT Madras

Examiner 1

Professor S. N. Merchant, Department of Electrical Engineering, IIT Bombay

Examiner 2

Dr. Abhinav Kumar, Department of Electrical Engineering, IIT Hyderabad

Internal Examiner

Dr. P Rajalakshmi, Department of Electrical Engineering, IIT Hyderabad

Adviser/Guide

Dr. Bheemarjuna Reddy Tamma, Department of Computer Science and Engineering, IIT Hyderabad

Chairman

Acknowledgements

Dedication

I dedicate this dissertation to my parents

Abstract

Ultrasound imaging works on the principle of transmitting ultrasound waves into the body and reconstructs the images of internal organs based on the strength of the echoes. Ultrasound imaging is considered to be safer, economical and can image the organs in real-time, which makes it widely used diagnostic imaging modality in health-care. Ultrasound imaging covers the broad spectrum of medical diagnostics; these include diagnosis of kidney, liver, pancreas, fetal monitoring, etc. Currently, the diagnosis through ultrasound scanning is clinic-centered, and the patients who are in need of ultrasound scanning has to visit the hospitals for getting the diagnosis. The services of an ultrasound system are constrained to hospitals and did not translate to its potential in remote health-care and point-of-care diagnostics due to its high form factor, shortage of sonographers, low signal to noise ratio, high diagnostic subjectivity, etc. In this thesis, we address these issues with an objective of making ultrasound imaging more reliable to use in point-of-care and remote health-care applications. To achieve the goal, we propose (i) computer-assisted algorithms to improve diagnostic accuracy and assist semi-skilled persons in scanning, (ii) speckle suppression algorithms to improve the diagnostic quality of ultrasound image, (iii) a reliable telesonography framework to address the shortage of sonographers, and (iv) a programmable portable ultrasound scanner to operate in point-of-care and remote health-care applications.

The objective of developing computer-assisted algorithms is to improve the diagnostic accuracy, nullify the bias caused due to the skill of a sonographer and to assist the semi-skilled technicians in diagnosis. Ultrasound scanning covers the broad diagnostic spectrum and involves high subjectivity in scanning, and hence computer-assisted algorithms need to be designed and engineered for each organ specifically. Here we focus on to develop computer-assisted algorithm for the typical problem in ultrasound diagnosis. The nonalcoholic fatty liver disease (NAFLD) is the leading cause of dysfunction of the liver and diagnosis of NAFLD through ultrasound imaging involves high subjectivity due to the minute variations observed in the characteristics of liver images across all grades of fatty liver. The quantification of fatty liver will be analyzed based on the characteristics of parenchymal texture, periportal veins, and diaphragm. The texture characteristics like morphology, echogenicity and homogeneity vary with the proportion of the fat in the liver, while the visibility of the diaphragm and periportal veins diminishes with the proportion of fat. Here, we propose two computer-assisted algorithms for NAFLD, one for characterization of ultrasonic fatty liver texture, while other for detection of diagnostically significant regions. As the texture properties vary with the concentration of fat, we formulate grading of fatty liver as a texture discrimination problem.

The scattering coefficients which give the stable and invariant representation for images are used as features for classifying the texture of a liver. To assist semi-skilled persons in scanning and to reduce the sonographers time in fatty liver diagnosis, we propose an algorithm for detection of the diaphragm, periportal veins, and texture of liver parenchyma. The signal to noise ratio of ultrasound images is very low due to the presence of speckle noise. Speckle delineates the boundaries of the organs, masks fine information about tissues, increase the variance between pixels thus reducing the diagnostic quality in ultrasound images. To enhance diagnostic quality, we propose a sparsity-based despeckling algorithm to effectively suppress the speckle noise. To address the shortage of sonographers in remote health-care, we propose a reliable web real-time communication (WebRTC) based streaming application. With the proposed application, the semi-skilled person can transmit the scanned ultrasound video in real-time to the expert-end for getting the diagnosis, and also it helps the semi-skilled person to get the real-time assistance from the sonographers in scanning the patients. Finally, to operate an ultrasound scanner in remote and point-of-care diagnostics, we propose a single system-on-chip based portable ultrasound scanning system. The proposed contributions on computer-assisted algorithms, image enhancement algorithms, WebRTC based telesonography and portable ultrasound scanner will leverage the potentiality of ultrasound diagnosis to use more reliably in point-of-care and remote diagnostics.

Contents

List of Figures

List of Tables

Chapter 1

Introduction to Ultrasound Imaging

1.1 Introduction

Medical ultrasound scanning is a diagnostic imaging modality which works on the principle of transmitting ultrasound waves into the body and reconstructs the images of internal organs based on the strength of the echoes. Depending on the type of application, ultrasound scanners are typically operated in 1-18 MHz frequency range. The resolution of the scanner is proportional to the frequency at which the device is operated, which is subjected to the depth of the organ scanned. Ultrasound scanners are operated at a higher frequency range of 7-18 MHz to image superficial structures like thyroid, breast, muscles, etc., since the high frequency gives better resolution of the tissues. The high-frequency ultrasound waves suffer from poor penetration due to high attenuation. Hence organs located at higher depths such as kidney, liver, pancreas, spleen, etc., are imaged with a low frequency ranging from 1-6 MHz [3].

Ultrasound scanning is the widely used diagnostic imaging modality in health-care due to following advantages [4]:

- Uses non-ionizing sound waves for sensing, hence considered to be the safest

imaging technology.

- Inexpensive compared to other imaging technologies like Magnetic Resonance Imaging (MRI), Computed Tomography (CT), etc.

- Have the capability to image the organs and moving structures like blood flow, cardiac motion, fetus, etc., in real-time.

- Covers a wide spectrum of medical diagnostics, which includes kidney, liver, pancreas, fetus monitoring, gynecology, cardiology, etc.

In addition to the above advantages, the ultrasound scanners add a lot of value in emergency health-care. For example, in situations like severe accidents, internal bleeding from the organs is the most serious threat to survival. Neither the injuries nor the bleeding may be apparently seen. Under the circumstances, the ultrasound scanning is very useful as a preliminary scan to pinpoint the location, cause and severity of hemorrhage, thus helping the doctors to take timely actions which can save the lives of patients [5]. Currently, the diagnosis through ultrasound scanning is hospital-centered, and the patients who are in need of ultrasound scanning has to visit the hospitals for getting the diagnosis. Conventional ultrasound scanners used in hospitals shown in Fig. 1.1 are placed on a cart positioned at a fixed location in hospitals. Ultrasound scanning is performed by the medical professionals trained in sonography known as sonologists.

1.2   Challenges

Despite being safer, economical and real time-imaging capability, ultrasound scanners have not been used to its fullest extent in health-care. Some of the bottlenecks encountered in maximum utilization of this technology are:

- **Form factor:** The conventional ultrasound scanners are of high form factor, and hence it becomes difficult to carry the ultrasound scanners to the patient's

Figure 1.1: Conventional ultrasound scanners (Siemens Acuson SC2000), used in current practice.

bedside to offer point-of-care (PoC) diagnostics.

- **Shortage of sonographers:** There is a severe shortage in the number of sonographers available to meet the current demand globally [6–8].

- **Low diagnostic accuracy:** The ultrasound images are contained with low signal to noise ratio, poorly defined organ boundaries, low contrast, etc., resulting in low diagnostic accuracy.

- **Prenatal gender determination act:** The ultrasound scanners have been used for gender determination of fetus, from where parents are opting for gender-selective abortions. This lead to a decline in the male-female sex ratio which is associated with lot more social issues. To prevent this, Government of India brought an amendment called Prenatal Gender Determination Act in the year 1994, which says that ultrasound scanning should only be used inside the hospital premises, and the sonographers should not reveal the gender of the fetus. These constrain the use of ultrasound scanners outside the hospital premises which limits the applications of ultrasound diagnostics to operate in PoC and remote health-care applications [8, 9].

1.3 Contributions

Considering the potentiality and limitations of ultrasound scanning in health-care, in this thesis, we focus on developing a reliable ultrasound technology customized for PoC and remote health-care diagnostics. The objective is realizable only if the ultrasound scanner is compact and comes with assistive technologies such that even a semi-skilled person is able to scan the patients in remote settings. As shown in Fig. 1.2, the contributions we made in this thesis to achieve the objective of taking ultrasound scanning from clinic centered diagnostics to PoC diagnostics are:

- **Computer-assisted diagnostics:**

  Ultrasound scanning covers a broad diagnostic spectrum and involves high subjectivity in scanning, and hence developing universal computer-assisted algorithms for complete ultrasound imaging is infeasible. The computer-assisted diagnostics need to be designed and engineered for each organ specifically. Here we focus on the development of a computer-assisted diagnostics for the representative problem in ultrasound diagnosis. Quantification of fat in the liver through ultrasound imaging is challenging, involves high subjectivity and carries a high clinical significance [10]. In this thesis, we propose a novel computer-assisted diagnostic algorithm for quantifying the fat in the liver. The characteristics of texture vary with the density of fat present in the liver hence we formulated quantification of fat in the liver as a texture discrimination problem. To characterize the texture of liver, we propose a scattering coefficient feature based classification. The propose algorithm can nullify the bias caused due to the skill of a sonographer and also assist the sonographers to take the decisions with high confidence.

- **Detection of diagnostically significant regions:**

  In ultrasound images, the diagnostic information is present only in particular regions called as the region of interest (RoI). The accuracy of diagnosis in ultra-

4
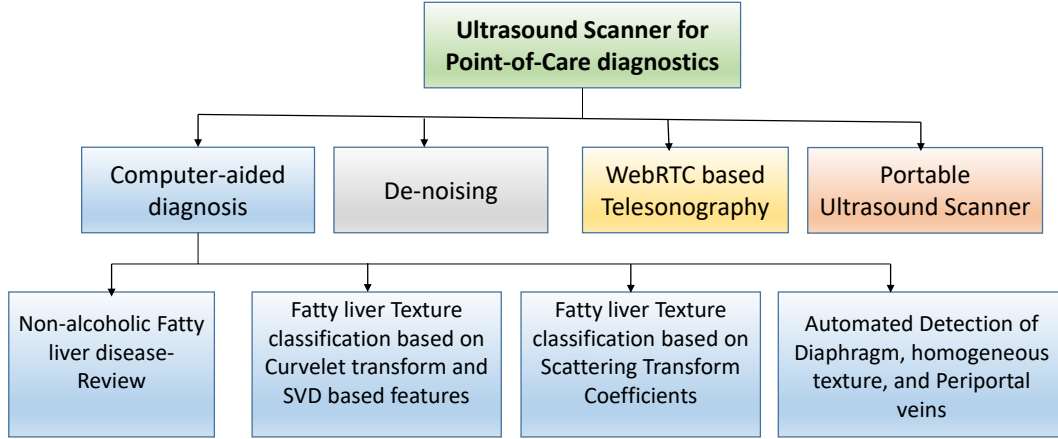
Figure 1.2: Contributions made in the thesis for making ultrasound scanning more reliable to operate at PoC and remote diagnostics.

sound imaging depends on detection and analysis of these regions. Automatic detection of these regions in ultrasound images will benefit in many ways: (i) assist the semi-skilled person to know whether the scanned image has the representative information useful for diagnosis, (ii) in the development of automated diagnostic algorithms, and (iii) to validate the presence of representative regions in images before transmitting it to the cloud or the expert-end for analysis in telesonography applications and (iv) to diagnose more patients in less time as it reduces the time taken by the sonographer in detecting the RoI. In this thesis, we propose algorithms for detecting diagnostically significant regions correspond to nonalcoholic fatty liver disease. Diaphragm, periportal veins and homogeneous texture of liver parenchyma are the three RoI's analyzed for diagnosing the nonalcoholic fatty liver disease. The shape and size of each RoI is organ specific hence we propose a specific algorithm for detecting diaphragm, periportal veins and homogeneous texture of liver parenchyma. The diaphragm is detected based on active contour segmentation, Viola Jones algorithm and GIST descrip-

tor, where periportal veins are detected using Viola Jones and GIST descriptor and homogeneous texture is detected based on histogram features and connected components algorithm.

- **Image enhancement algorithms:**

  The ultrasound images are affected by speckle noise which is multiplicative in nature. The speckle noise appears like dense granules and small worm-like structures conveying no significant information in the ultrasound image. Speckle noise occurs when the wavelength of sound is greater than the dimension of scanning tissue. Under this condition, the tissue will scatter the ultrasound wave that impinged on it. The scattered waves will randomly interfere with each other at the receiver side forming speckle noise. If the interference patterns add constructively it will result in bright spots or if the interference patterns add destructively it will result in black spots in an image resulting in granularity and small worm-like structures in an ultrasound image. Speckle noise masks the fine information about the tissues, delineates the boundaries of organ, thus reducing the diagnostic information present in the image. In this thesis, we propose a speckle suppression algorithm based on sparsity to suppress the speckle noise in ultrasound images.

- **Telesonography:**

  In general, telesonography addresses the lack of sonographers by allowing non-expertise or semi-skilled persons to scan the patients in remote health-care and transmit the scanned data to the expert-end for diagnosis. In conventional telesonography, the ultrasound video is lossy encoded with H.264/H.265 and transmitted to the expert end via wired or wireless communication technologies [11, 12]. These kind of techniques are high bandwidth demanding and did not adapt to the network conditions resulting in underutilization of the network resources and poor quality of service, making it unreliable for real-time telesonog-

raphy applications. In this thesis, we propose a web real-time communication (WebRTC) framework for reliable telesonography, where the end-users can connect with each other and share ultrasound video enabling real-time telediagnosis. The developed application is a server-based service and end-users does not need to have any installed software or plugins in their computing platforms to connect the other-end for establishing the connection and streaming the ultrasound video, thus enabling reliable real-time telediagnosis.

- **Portable ultrasound scanner:**
  As the conventional ultrasound scanners cannot be used in point-of-care diagnostics due to its high form factor, we propose a programmable portable ultrasound hardware framework based on a single system-on-chip. The system is designed for 16-channel and comes with adequate external interfaces to integrate with devices like finger-print scanner, the global positioning system (GPS), communication modules, etc., to ensure that the ultrasound scanning can be used securely in remote health-care environments.

1.4 Database Acquisition

The entire database used for developing CAD algorithms in this thesis is acquired from Asian Institute of Gastroenterology, Hyderabad, Telangana, India. The liver ultrasound images are acquired using 3.5 MHz curved array transducer with clinical Seimens S1000 ultrasound scanning system. The ultrasound images are captured to a depth of 15 cm by adjusting focal zone to the center of the organ to obtain minimal beam diffraction. Due to low inter and intra diagnosing agreement of fatty liver, we considered the opinion of three sonographers independently in labeling each image. All the three sonographers have more than 10 years of experience in ultrasonography. The images with 100% agreement between the sonographers are only considered for the analysis. The liver database of 657 labeled ultrasound images corresponding

to NAFLD is acquired in the process. These includes 203 Normal, 173 Grade I, 157 Grade II and 124 Grade III ultrasound images. Apart from liver images, we also collected a ultrasound kidney database of 600 images, 10 liver and 5 spleen ultrasound videos.

1.5   Thesis Overview

The thesis consists of four parts, in first part, we present problems associated in developing computer-aided diagnostics in ultrasound scanning, in particular we focus on developing algorithms for quantifying the fat content in the liver and detection of diagnostically significant regions corresponding to fatty liver. In the second part, we present speckle suppression algorithm to improve the diagnostic quality in ultrasound images. In the third part, we present the WebRTC framework for real-time telesonography. In the fourth part, we present the hardware framework for ultrasound scanning system designed for portability and PoC diagnostics.

The thesis is organized in the following way:

In Chapter **2**, we present the significance and challenges involved in grading the nonalcoholic fatty liver through ultrasound imaging. We discuss the methodologies employed in the literature regarding the quantification of fat in the liver and also we briefly mention about contribution towards the development of computer-assisted algorithms in quantifying the fat in the liver.

In Chapters **3** & **4**, we present the computer-assisted algorithms for quantification of fat in ultrasonic images, here we propose two novel algorithms one based on curvelet transform, and other based on scattering coefficients. The scattering coefficient features proved very effective in representing the texture of fatty liver which replicated in achieving good classification accuracy. The performance of the proposed algorithms are compared with widely used texture features.

In chapter **5**, we present an algorithm for the detection of diagnostically significant

regions useful for diagnosing the fatty liver. The algorithm is developed to detect the regions like the diaphragm, periportal veins, and texture of liver parenchyma, which are representative for grading the fatty liver. In Chapter **6**, we present a sparsity-based de-speckling algorithm for effective suppression of speckle noise in ultrasound images. In Chapter **7**, we present the WebRTC framework for real-time telesonography. In Chapter **8**, we present the hardware prototype of the system-on-chip (SoC) based portable ultrasound scanner designed for point-of-care applications. We summarize the entire thesis and potential future works in Chapter **9**.

Chapter 2

Nonalcoholic Fatty Liver Disease

2.1   Introduction

Accumulation of excess fat in liver cells is an abnormal condition of a liver known as a nonalcoholic fatty liver disease (NAFLD). It is etiologically associated with the hepatic manifestation of metabolic syndrome, specifically insulin resistance. Also, NAFLD is associated with obesity, type 2 diabetes, hyperlipidemia, side effects of certain medications, cardiovascular diseases, etc. NAFLD is one of the leading causes of liver dysfunction and is rapidly growing health problem in the world. It is expected that up to 30% of general population in the developed countries have the condition of NAFLD [13]. NAFLD is observed in 80-90% of obese patients, 30-50% of diabetic patients and 90% of hyperlipidemia patients [14]. If the underlying problem associated with NAFLD is not detected and treated, then NAFLD can progress into chronic liver diseases. From recent investigations, it is found that 50% of patients with NAFLD have led to fibrosis, 15% of patients with NAFLD have led to liver cirrhosis, while 3% of patients with NAFLD have led to liver failure resulting in liver transplantation [15]. Early and accurate detection of fatty liver becomes crucial to prevent liver suffer from chronic diseases.

Table 2.1: Grading and condition of NAFLD based on the percentage of fat present in the liver.

NAFLD	Degree of steotosis	Grading	Condition
Simple Steotosis	<5%	Grade 0	Normal
NASH	5-33%	Grade I	Mild
	33 -66%	Grade II	Moderate
	>66%	Grade III	Severe
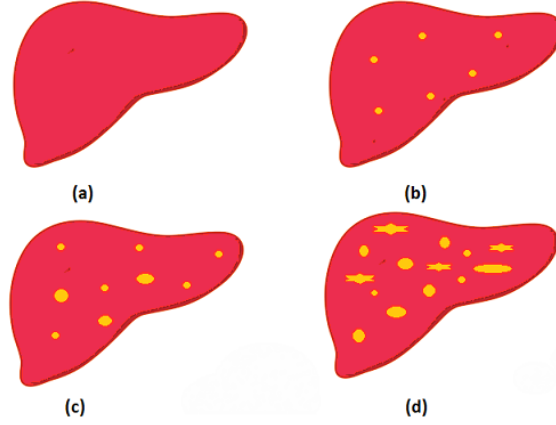
Figure 2.1: Graphical representation of NAFLD, yellow patches represents the fat or triglycerides (a) Normal, (b) Grade I, (c) Grade II, and (d) Grade III fatty liver.

The severity of NAFLD is characterized by the density of fatty granules accumulated in the tissues of a liver [16]. The visual representation regarding the presence of fatty granules corresponding to different grades of the nonalcoholic fatty liver is shown in Fig. 2.1 [1]. Brunt et al. categorized NAFLD into simple steatosis (Normal or Grade 0) and nonalcoholic steatohepatitis (NASH). NASH is further categorized into Grade I, Grade II and Grade III, as shown in Table. 2.1 [16]. If the concentration of fat in the liver is less than 5%, then the liver is considered as Grade 0 which is treated as a normal condition. If the concentration of fat in the liver is in between 5–33%, then the liver is considered to be in Grade I condition. Higher concentration of fat levels in liver such as 33–66% is considered as Grade II, and greater than 66% is considered as Grade III respectively. In general, Grade 0 and Grade I does not have an impact on the functionality of the liver and does not need medication.

Grade II and Grade III conditions impact the functionality of the liver and patients need medical attention to prevent the liver progressing into chronic diseases.

Fatty liver diagnosis is done through invasive and noninvasive procedures; invasive methods include biopsies, blood tests, etc., while noninvasive method includes imaging techniques like ultrasound scanning, Magnetic Resonance Imaging (MRI) and Computed Tomography (CT). Invasive procedures are painful, and it is associated with complications like infections, bleeding, bile leakage, etc. Hence, doctors recommend for noninvasive imaging procedures. Ultrasound scanning is widely used imaging modality for diagnosing the fatty liver since it offers real-time imaging, safer and economical compared to MRI and CT. Unlike MRI and CT (where images are captured automatically by systems without manual intervention), ultrasound scanning is performed by humans resulting in high subjectivity. Strauss et al. [10] reported that there is a mean interobserver and intraobserver agreement of 72% and 76% respectively in detecting the normal livers from the fatty livers, while quantifying the severity of fat there is a mean interobserver and intraobserver agreement of 47-59% and 59-64% respectively. Hence there is a need for developing computer-aided diagnostics (CAD) algorithms to nullify the bias caused due to the subjectivity in scanning and also to accurately quantify the fat in the liver.

Developing of CAD algorithms is challenging in ultrasound imaging due to the complexity involved from both technology and a human perspective. From technology perspective, ultrasound images contain a low signal to noise ratio, poorly defined organ boundaries, low contrast, obscuration of organs due to acoustic shadows, etc. From the human perspective, ultrasound scanning and diagnosis is performed by the sonographer, and hence the diagnosis will be biased to the skill of a sonographer. Also, the characteristics of the scanned organ vary with the anatomy of a patient, which includes age, height, gender, body mass index, position, and position of the transducer used to scan the organs, etc. CAD algorithms are developed based on RF data or completely formed ultrasound images. Current clinical ultrasound machines

do not have provision to acquire the raw echo data directly, and requires an expensive research interface [17] for acquiring the data. Depending on the depth of scan and analog to digital converter (ADC) resolution, the generated RF data for a single scan line is of very high dimension (for example, scanning to a depth of 10 cm, with ADC operating at 40 MHz will result in a total of 5200 samples per scan line) which makes it difficult to develop the machine learning algorithms. Hence researchers are mainly focused on developing CAD algorithms based on the images. The images needed for developing the CAD algorithms can be acquired from clinical ultrasound machines without any need for having external research interfaces. In addition to this, no additional effort is needed from the sonographers to capture these images, since it is inherent in the process of diagnosing the patients.

The liver ultrasound images corresponding to different grades of fatty liver is shown in Fig. 2.2, we can infer that there are minute variations in the characteristics of the image across different grades of fatty liver. The quantification of fat in the liver through ultrasound scanning is done by perceiving information in three regions of the liver such as a diaphragm, periportal veins, and texture of liver parenchyma. The characteristics of these regions change accordingly with the severity of fat present in the liver. The diaphragm, periportal veins, and texture of liver parenchyma of ultrasonic images is shown in Fig. 2.3. The visual characteristics of the RoI correspond to different grades of fatty liver is discussed in Table. 2.2. Based on the combined characteristics of these regions, sonographers quantify the severity of fat present in the liver.

2.2 Literature Survey

In literature, most of the research is focused on classifying normal liver with other liver diseases such as fatty, cirrhosis, hepatocellular carcinoma (HCC), fibrosis, etc [18–23]. While classification, authors have considered all grades (Grade I, Grade II and Grade
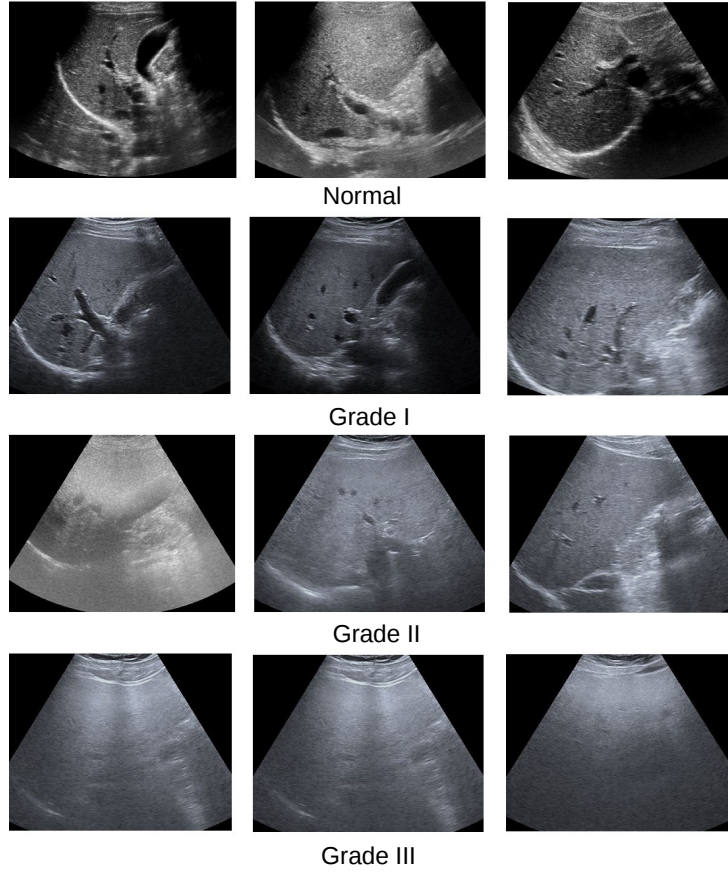
Figure 2.2: Ultrasound images correspond to different grades of nonalcoholic fatty liver. Images in each row belong to single category. We can observe minute differences in the characteristics of RoI across all grades of fatty liver images.

III) of fatty liver as one class, and further discrimination within the fatty liver is not addressed extensively. Accurate quantification of fat in the liver carries paramount importance in liver diagnosis. For example in liver transplantation, even a Grade I fatty liver donor can increase the potentiality of liver failure in the recipient and also there is a high probability that the donor will get diseased [24,25]. The patients who undergo liver resections with Grade III fatty liver are expected to suffer from post-operative complications [26]. Detecting the severity of fatty liver is of high importance such that the patients can take appropriate precautions to avoid the complications associated with the NAFLD.
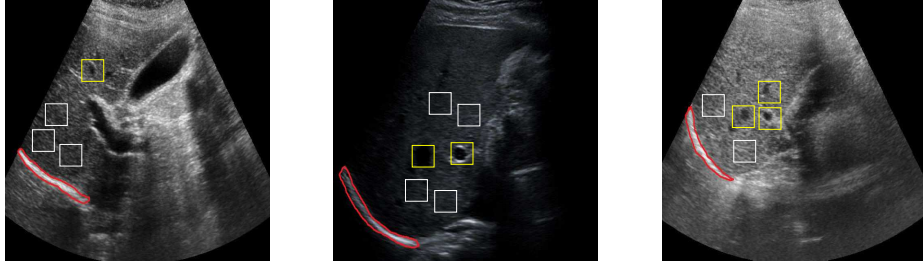
Figure 2.3: Ultrasonic liver images. Diaphragm region is enclosed with red contour, regions enclosed with yellow boxes indicates periportal veins and regions enclosed in white boxes represents homogeneous texture.

Table 2.2: Characteristics of diaphragm, periportal veins and texture of liver parenchyma correspond to different grades of fatty liver [1], [2].

Category	Characteristics
Normal	Visible echogenicity with visible periportal and diaphragm. The texture of liver parenchyma appears coarser and rugged.
Grade I	Increased hepatic echogenicity with visible periportal and diaphragmatic echogenicity. The texture appears less coarser and smooth.
Grade II	Increased hepatic echogenicity with imperceptible periportal echogenicity with partial obscuration of diaphragm. The texture appears more smooth and finer.
Grade III	Increased hepatic echogenicity with imperceptible periportal echogenicity and obscuration of diaphragm. The texture appears diffused and more finer.

Some of the image diagnostic methodologies appeared in literature in quantification of fat in the liver are:

Lupsor et al. [27] quantified the grades based on the attenuation coefficient (AC) and gray level co-occurrence matrix (GLCM) entropy features and concluded that the AC performs better compared to GLCM entropy features. The AC is computed by considering the pixel values along the vertical line (depth) of the image, while the GLCM entropy feature is computed from a rectangular region cropped from the homogeneous texture parenchyma of a liver. The authors computed the AC from RoI corresponding to the vertical straight line almost spanning the entire ultrasound image. Due to this, even a small change in one of the gain knobs of time gain compensation will change the statistics of the attenuation coefficients which will have an impact on the final classification accuracy.

Semra et al. [28] quantified the fatty content based on the gray relational grade (GRG) feature computed between liver and kidney parenchyma, here authors have

considered the database having both the liver and kidney organs present in the same image. The authors used multiple RoI's corresponding to liver and kidney parenchyma and used gray relational grade as a feature in doing classification. In this methodology, a cross-labeling of RoI's will have a significant impact on the classification accuracy.

Dan Mihai et al. [29], considered two RoIs from the liver for feature extraction, these features include: minimum attenuation (MIA) and maximum attenuation (MAA), maximum value for region (MAV), minimum value for region (MIV), median for liver (ML) parenchyma and median for kidney (MK) parenchyma. The features are extracted from the region of interest cropped from the liver and kidney parenchyma, and a dichotomy structure is employed for classification. Similar to [28], the authors considered the database having both liver and kidney organs present in the same image. The authors considered RoI covering the image region all over the liver parenchyma to compute the attenuation values. It is difficult to get a RoI of the large size without hepatic and portal veins, and also there is a need to validate the algorithm on a larger database

In [30], Cristian et al. used AC, backscattering coefficient (BS) and fit error (FE) as features to classify the different grades of the fatty liver. The features are extracted along the three vertical lines of the liver parenchyma.The authors consider RoI as three vertical lines spanning from top to bottom of the image. Extracting the features from three vertical lines is highly subjective to time gain compensation knobs, and getting the three RoIs without hepatic and portal veins is highly constrained.

In [1], Yin-Yin Liao et al. extracted multiple features from the radio frequency (RF) and liver ultrasound image, these include texture features, signal to noise ratio (SNR) and slope of the center frequency downshift (SCFD). Texture features include auto-correlation (AUC), sum average (SA) and sum variance (SV). The texture features are extracted from the RoI cropped from the liver parenchyma, while the AC and BS features are extracted from the RF data, the classification is done using a multinominal logistic model (MLM). Authors have dealt only with three classifica-

Table 2.3: Methodologies proposed in the literature for grading the fatty liver. Notations:- N: Normal, GI: Grade I, GII: Grade II, GIII: Grade III, the notation (N, GI) is considered as one class.

Authors	Features	Classifier	Database	Performance measure
M Lupsor et al. [27]	AC	SA	N: 24 NASH: 96	AUROC: Nvs GI: 0.951, N, GI vs GII, GIII: 0.879, N, GI, GII vs GIII: 0.859
Semra Icer et al. [28]	GRG	SA	N:45 GI: 30, GII: 55, GIII: 10	AUROC: N vs GI : 0.975, GI vs GII: 0.958, GII vs GIII: 0.949
Dan Mihai et al. [29]	MIA, MAA, MAV, MIV, ML, MK	RF	N:10, GI: 70, GII: 33, GIII: 7.	Accuracy: 91.7%
Cristian Vicas et al. [30]	AC, BS, FE	SVM	N: 25, GI: 32; GII: 37, GIII: 17.	AUROC: N vs GI, GII, GIII: 0.84, N, GI vs GII, GIII: 0.73, N, GI, GII vs GIII: 0.66
Yin-Yin Liao et al. [1]	AUC, SA, SV, SNR, SCFD, AC, BS	MLM	N: 151, Mild NAFLD: 127, Severe NAFLD: 106.	AUROC: N vs mild NAFLD: 0.73, N vs Severe NAFLD:0.81

tions namely: Normal, mild NAFLD and severe NAFLD. Classification of moderate NAFLD is not discussed.

The brief summary regarding the methodologies proposed in the literature in quantifying the fatty liver is reported in Table. 2.3. Few authors reported performance in terms of area under receiver operating characteristic (AUROC), while others analyzed performance based on the accuracy. The methodologies employed in the literature are diversified as the authors considered liver ultrasound images with spleen, liver ultrasound images with the kidney to quantify the fat in the liver, while Yin-Yin Liao et al. [1] considered multiple features corresponding to RF data and texture features to quantify the fatty content in the liver. Hence, we cannot have one common agreement to judge the performance of the proposed algorithms. In all the methodologies employed in the literature, authors have extracted features from more than one RoI, which includes pixels along wave propagation and homogeneous texture regions to classify the fatty liver. The methodologies employed in the literature are moderately successful in classifying the fatty content in the liver.

In this thesis, we propose a computer-aided diagnosis algorithm for quantifying

the fat in the liver based on the texture properties of liver parenchyma. Unlike the methodologies employed in the literature, we have considered only texture of liver parenchyma as RoI to quantify fat in the liver which makes it less subjective to the scanning parameters of the ultrasound machine and needs minimum intervention to crop the RoI from an image

We made following contributions to the development of a computer-assisted diagnostic algorithm for quantifying the fat in the liver.

- We propose a novel feature extraction scheme based on curvelet transform and SVD based features for characterizing the texture of liver.

- We propose a scattering network-based classification algorithm for quantifying the fat in the liver.

The performance of the proposed algorithm is compared with popularly used texture features. The database for developing the algorithms are acquired by collaboratively working with the sonographers in the scanning process.

2.3 Summary

In this chapter, we presented the importance of accurately quantifying the fat in liver and role played by ultrasound imaging in fatty liver diagnosis. In general, Doctors prefer ultrasound imaging over MRI and CT due to its safer, economical and real-time imaging capability. Quantifying the fat in the liver through ultrasound imaging is considered to be highly subjective due to the minute variations observed in the characteristics of fatty liver. It is inferred that there is very low inter and intra-agreement between the sonographers in grading the fatty liver, resulting in diagnostic uncertainty. Computer-assisted algorithms are needed to nullify the subjectivity caused due to the skill of the sonographers. We also discussed about different methodologies proposed in the literature in quantifying the fat in the liver and their limitations and also briefly discussed about the contributions of the thesis towards the development

of computer-aided diagnostics. In the next chapter, we propose a novel computer-assisted algorithm to characterize the texture of NAFLD based on curvelet transform and SVD based features.

Chapter 3

NAFL Texture Classification based on Curvelet transform and SVD

3.1 Introduction

The texture properties of liver parenchyma in ultrasound image varies with the severity of fat present in the liver [18, 31, 32]. Sonographers quantize the fatty content of the liver based on texture structural and perceptual properties of a liver; these include texture morphology, echogenicity, and degree of diffusion. The texture corresponds to different grades of NAFL is shown in Fig. 3.1. Visually, a minute difference is observed in textures correspond to different grades of fatty liver. The texture of normal liver appears coarser and rugged, and it becomes finer and finer as the concentration of fat increases.

As the properties of liver parenchymal texture is associated with the concentration of fat, we formulated the grading of fatty liver as a texture discrimination problem.

In literature, features like gray-level co-occurrence matrix GLCM [20, 22, 33–35], gray-level run-length matrices (GLRLM) [22, 35, 36], Gabor [35, 37], GIST [35], Laws texture measures [34], multi-resolution analysis [35, 38, 39] have been widely used in characterizing the ultrasonic texture of Normal, Fibrosis, HCC and cirrhosis diseases.
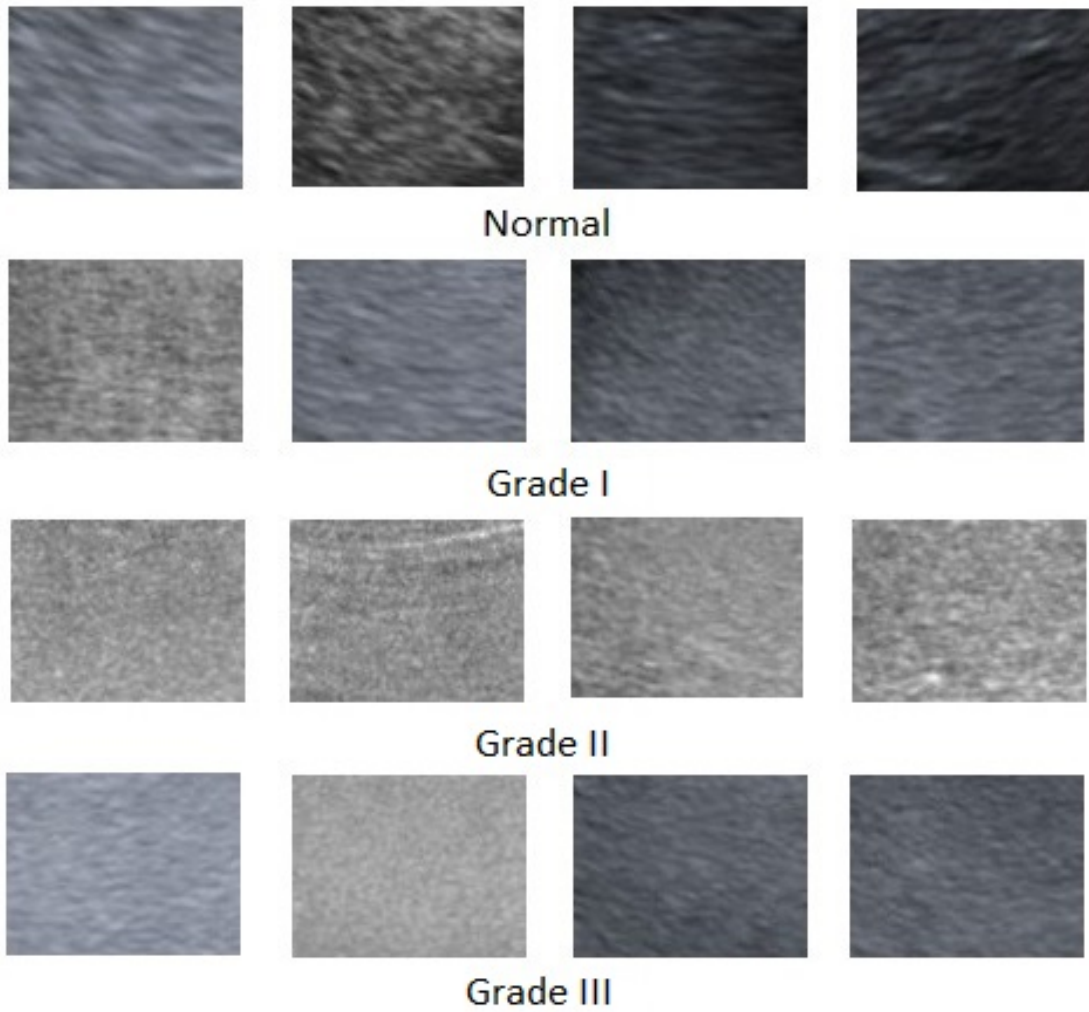
Figure 3.1: Textures corresponding to different grades of NAFL where images in each row belong to single class.

These features are moderately successful in classifying the texture of a liver, hence more than one set of features have been used for achieving higher classification accuracies [35, 40–44]. In this thesis, we propose a novel feature extraction scheme based on curvelet transform and singular value decomposition technique to effectively represent the texture of liver for classification. The effectiveness of the propose features is evaluated with the K-nearest neighbor and cubic kernel SVM classifier.

3.2 Curvelet Transform and SVD based Characterization of Ultrasonic Liver Texture

The texture of ultrasonic fatty liver varies with concentration of fat in the liver. For Normal liver texture appears very rugged, irregular and coarser in nature. The coarseness and irregularities of the texture reduces and becomes smoother and smoother as the concentration of fat increases. We also inferred that the texture of normal liver consist scars like structures that appears as curves. The texture properties like irregularities, coarseness and smoothness of the texture can be captured by extracting the gradient information of the texture. Higher the irregularity and coarseness better is the gradient magnitude. Curvelet transform has the ability to capture the gradient information with respect to different scales, orientations and space. In addition, it has the property of representing the finer curves in the texture due to its anisotropic scaling property. Hence, it prompted us to apply Curvelet transform on the image to capture the gradient properties and irregularities in the texture. Curvelet transform decomposes the image into sub-bands extracting the gradient information localized to scale, space and orientation. The total number of curvelet coefficients collectively present in all the sub-bands is of high dimension, which becomes difficult to model or to train a classifier. To get compact and efficient representation for each decomposed image, SVD technique is applied. The mean of the singular values is computed for each sub-band and considered as a feature representation for the texture [45]. Cubic kernel SVM and K- Nearest Neighbor (KNN) classifiers are used to analyze the performance of the propose feature extraction scheme in quantifying the fat in the liver.

The block diagram representation of the propose algorithm used for quantifying the texture of fatty liver is shown in Fig. 3.2. Initially, the curvelet transform is applied to the image texture decomposing the image into different subbands. SVD is applied to each subband for computing the singular values. The mean of singular
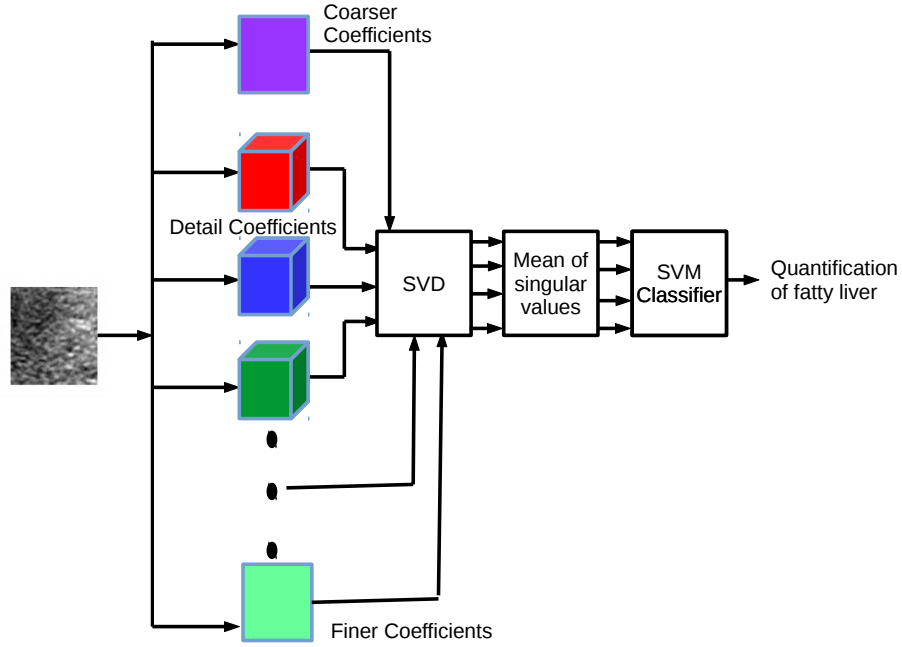
Figure 3.2: Block diagram representation of the propose method for quantifying the texture of a fatty liver.

values is then used as a feature to train the cubic SVM classifier.

3.2.1 Curvelet Transform

Curvelet transform was proposed by E. Candes and D. Donoho [46] to overcome the drawbacks of the conventional wavelet transform. The conventional wavelets lack directional representations, which lead to directional wavelets such as Gabor wavelets, curvelets, etc. Gabor wavelets can capture the direction information isotropically but lack direction sensitivity which is addressed with the curvelet transform.

One of the crucial tasks in image classification is to extract the representative features from an image. The features may be lines, edges, curves, textures, etc. The features are characterized with respect to scale, location, direction, geometry, etc., which motivated researchers to use scale-space filtering and multiresolution transforms for feature extraction. Curvelets are multiresolution transforms localized in scale, space, and direction, and gives superior performance in representing the texture, edges and

curves. Curvelets combined with other methodologies gave good results in various image processing applications like denoising [47], image representations [48], image enhancements [49], etc. Recently, curvelet transforms have been widely applied in medical image processing for developing automated diagnostic algorithms. In [50], Nayak et al. used curvelet transform for feature extraction to classify normal and pathological brain MR images. In [51], retinal blood vessels are effectively detected with high accuracy using the curvelet transform. In [52], curvelet transform with entropy features is used for automatic classification of normal and abnormal liver ultrasound images.

The objective of applying curvelet transform here is to enhance the finest curves present in the texture image with respect to different scales and orientations. The high directional sensitivity of the curvelets are obtained with the wedge functions, which makes it represent the curves more efficiently than the traditional wavelets. Curvelets give optimal sparse representation for the objects with C^2 singularities. The wavelet approximation $\tilde{f}$ for smooth object f with C^2 singularity or discontinuity using best m term wavelet thresholding can be obtained with $||f - \tilde{f}||_2^2 \approx m^{-1}$, while curvelet approximation $\tilde{f}_m^c$ will give $||f - \tilde{f}_m^c||_2^2 \approx Cm^{-2}(logm)^3$, resulting in a small asymptotic error compared to any other representations. Curvelets are multiscale transforms with strong direction characteristics, and the elements are highly anisotropic at fine scales with support following the parabolic scaling $width \approx length^2$.

To briefly explain the curvelet transform, we introduce the following notations, x is a spatial variable in $\mathbb{R}^2$, w is a frequency variable, r and θ represents the polar coordinates in the frequency domain. To construct the system of curvelet functions for each scale j, we define a window function U_j in Fourier domain as

$$U_j(r, \theta) = 2^{\frac{-3j}{4}} W(2^{-j}r) V\left(\frac{(2^{\lfloor \frac{j}{2} \rfloor} \theta)}{2\pi}\right), \tag{3.1}$$

where $\lfloor \frac{j}{2} \rfloor$ is the floor of $\frac{j}{2}$, $W(r)$ and $V(t)$ are the real valued, smooth and nonnegative windows supported on $r \in (1/2, 2)$ and $t \in [-1, 1]$ respectively. The support of U_j will be a polar wedge defined over the support of W and V. The $U_j(w)$ is equivalent to the Fourier transform of mother curvelet $\psi_j(x)$ defined in Eq. 3.2. The system of curvelets at scale 2^{-j} is acquired by translating and rotating the ψ_j, where the sequence of equispaced rotation angles is given by $\theta_l = 2\pi \cdot 2^{-\lfloor \frac{j}{2} \rfloor} \cdot l$, where $l = 0, 1, 2....$ such that $0 \leq \theta_l < 2\pi$. With the sequence of translation parameter $k = (k_1, k_2) \in Z^2$, the curvelet function at scale 2^{-j}, orientation θ_l and position $x_k^{(j,l)} = R_{\theta_l}^{-1}(k_1 2^{-j}, k_2 2^{-\frac{j}{2}})$ is defined as

$$\psi_{j,l,k}(x) = \psi_j(R_{\theta_l}(x - x_k^{(j,l)})), \tag{3.2}$$

where R_θ represent rotations of θ radians The curvelet coefficients c of an element $f \in \mathbb{R}^2$ is obtained as

$$c(j, l, k) := \; < f, \psi_{j,l,k} > = \int_\mathbb{R} f(x) \overline{\psi_{j,l,k}(x)} dx. \tag{3.3}$$

The digital curvelet transforms are always computed in the frequency domain. The curvelet coefficients in frequency domain is computed as

$$c(j, l, k) := \frac{1}{(2\pi)^2} \int \hat{f}(w) \overline{\hat{\psi}_{j,l,k}(w)} dw \tag{3.4}$$
$$= \frac{1}{(2\pi)^2} \int \hat{f}(w) U_j(R_{\theta_l} w) e^{i<x_k(j,l),w>} dw.$$

Implementation of curvelet transform known as the first generation of curvelets is very complicated. Candes and Donoho proposed a simpler and fast second generation transform called fast discrete curvelet transform (FDCT) [53]. FDCT is implemented in two versions namely wrapping of specially selected Fourier samples and unequally spaced fast Fourier transforms (USFFT), both having same computational complexity. In this work, we used USFFT based curvelet transform. Considering an $n \times n$

image in Cartesian arrays of the form $f[t_1, t_2]$, $0 \leq t_1, t_2 < n$, the curvelet transform of an image using USFFT is obtained in the following way.

1. The Fourier samples of the input arrays are obtained by applying a 2D fast Fourier transform (FFT) as

$$\hat{f}[n_1, n_2] = \sum_{t_1, t_2 = 0}^{n-1} f[t_1, t_2] e^{-i2\pi(n_1 t_1 + n_2 t_2)/n}; -n/2 \leq n_1, n_2 < n/2. \qquad (3.5)$$

2. For each pair of scale j and angle l, Fourier samples $\hat{f}[n_1, n_2 - n_1 tan\theta_l]$ are obtained from interpolating $\hat{f}[n_1, n_2]$ for $n_1, n_2 \in P_j$, where

    $P_j = \{(n_1, n_2) : n_{10} \leq n_1 < n_{10} + L_{1,j}; n_{20} \leq n_2 < n_{20} + L_{2,j}\}$,

    where, $L_{1,j}$ and $L_{2,j}$ are the length and width of a rectangle, (n_{10}, n_{20}) are the pixel index corresponding to bottom of the rectangle.

3. The interpolated samples are then multiplied with a frequency window $\tilde{U}_j$ to obtain

$$\tilde{f}_{j,l}[n_1, n_2] = \hat{f}[n_1, n_2 - n_1 tan\theta_l]\tilde{U}_j[n_1, n_2]. \qquad (3.6)$$

4. Curvelet coefficients are obtained by applying the inverse 2D FFT.

$$C^D(j, l, k) = \sum_{n_1, n_2 \in P_J} \tilde{f}_{j,l}[n_1, n_2] e^{i2\pi(k_1 n_1/L_{1,j} + k_2 n_2/L_{2,j})}. \qquad (3.7)$$

The curvelet coefficients corresponding to the texture of liver parenchyma is shown in Fig. 3.3, the difference in finer details of the texture corresponding to different grades is better visualized in the curvelet coefficients which is not obvious in the original texture.

3.2.2 SVD

In general, the classification of images via wavelet models are built by modeling the wavelet coefficients [54–56]. The number of wavelet coefficients in each sub-image is
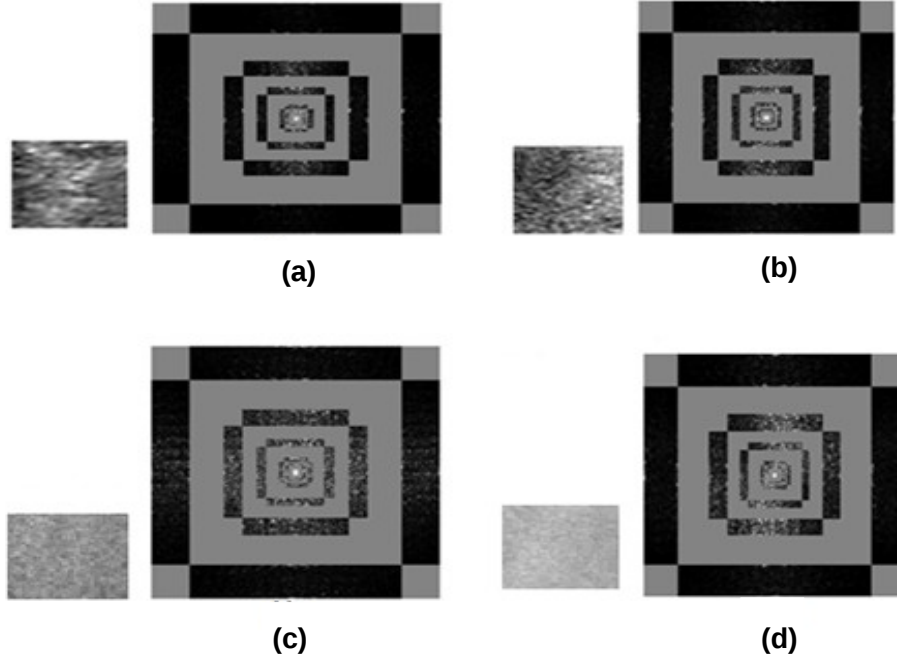
Figure 3.3: The curvelet coefficients of different grades of fatty liver texture. (a) Normal, (b) Grade I, (c) Grade II, (d) Grade III of size 128×128. We can observe discrimination in the curvelet coefficients corresponding to different grades of fatty liver, which is not obvious in the texture image.

of very high dimension, and hence it is complex to model the data. To reduce the dimension of the wavelet coefficients, features such as energy, mean, standard deviation, fractal dimensions, etc., have been exclusively used for classifying the texture. These features are moderately successful in characterizing the texture of a liver. Hence it prompted researchers to use more than one set of features to represent the texture of a liver [40]. To quantify the curvelet coefficients in each subband and to get the good representation, SVD is applied on each sub-image. SVD is widely used in image classification algorithms [45] image de-noising [57], dimensionality reduction [58], solving system of linear equations [59], etc. If C_i is the curvelet transformation coefficient matrix then SVD of C_i, which is of the size M × N is obtained as

$$C_i = P_i \sum_i Q_i^T, \tag{3.8}$$

here, P_i is a M × N orthogonal matrix with Eigenvectors as columns of $C_i C_i^T$, Q_i is a N×N orthogonal matrix whose Eigenvectors are columns of $C_i^T C_i$, and $\sum_i$ is a N×N diagonal matrix with singular values $\sigma_1, \sigma_2 \ldots \ldots \sigma_n$ arranged in decreasing order $\sigma_1 \geq \sigma_2 \geq \sigma_3 \cdots \geq \sigma_n \geq 0$. The singular values are computed as the squareroot of the Eigenvalues of a matrix $C_i C_i^T$ or $C_i^T C_i$. The mean of the singular values $\sum_i$ are computed and considered as a feature for each decomposed image.

3.2.3   Classifiers

The algorithm is evaluated using two supervised algorithms support vector machine (SVM) and KNN classifier. The brief introduction regarding the classifiers is given below.

SVM

SVM learns a model from the training features that separate the different classes. SVM is a binary classifier, hence to classify more than two classes One-vs-One approach is used. SVM works only with linearly separable data, to work with linearly non-separable data, SVM is operated with kernel operators. In this study, we used SVM with a cubic kernel of order three for classification [60]. The inbuilt SVM function available in the MATLAB 2017a version is used in the experiment

KNN

The KNN classify the incoming feature by computing K nearest neighbors with the training features. The nearest neighbors between the feature vectors are computed based on the Euclidean norm. The features in the training example which have less distance with the feature that has to be classified is considered as the nearest neighbor.
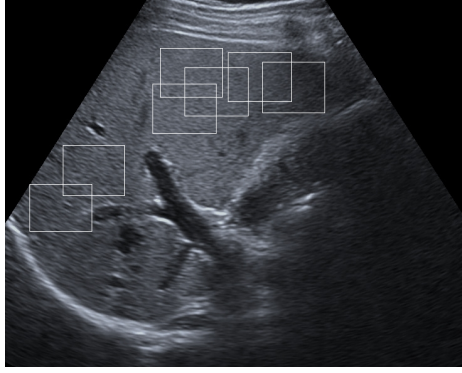
Figure 3.4: Liver ultrasound image. Rectangular boxes in the image represents the texture used for classification.

The classification is done based on a majority voting rule. In this study, we analyzed the performance by fixing $K=5$.

3.3 Database for Analysis

The database used in the analysis consists of 196 Normal, 173 Grade I, 157 Grade II and 124 Grade III images. The texture in all the images is cropped in the homogeneous regions of liver parenchyma as shown in the Fig. 3.4. The ultrasound images are acquired using 3.5 MHz curved array transducer with clinical Seimens S1000 ultrasound scanning system. The ultrasound images are captured to a depth of 15 cm by adjusting the focal zone to the center of the organ for minimal beam diffraction. Care is taken that the texture is free from blood vessels, acoustic shadows, hepatic and portal veins. Each sample cropped from the image is of size 78×100. Multiple samples are cropped from the same image ensuring overlapping samples does not contain more than half of the pixels in common, which is similar to the methodology adapted in [39], [40], [61]. The analysis is carried on a database consisting of 1000 texture patches of a liver, where each category consists of 250 images.

Table 3.1: Accuracy of the propose algorithm with respect to the features computed for wedges of each scale.

		Accuracy			
Scale	Feature size	Real		Complex	
		KNN	SVM	KNN	SVM
1	1	64.7	25.8	64.3	27.9
2	8	53.8	54.4	52.8	51.7
3	16	63.4	62.6	67.7	67.6
4	16	76.4	75.0	79.1	79.4
5	32	89.9	88.8	90.6	90.9
6	32	88.4	88.7	89.7	88.7
7	1	48.9	25.9	50.6	25.5

Table 3.2: Accuracy of the propose algorithm

Image size	Scale	Feature size	Accuracy			
			Real		Complex	
			KNN	SVM	KNN	SVM
128×128	7	106	93.9	95.0	93.7	95.7

3.4   Results

The proposed algorithm is evaluated with ten-fold cross-validation scheme. The performance of the algorithm is measured using accuracy as a metric, which is termed as the ratio of correctly classified images to the total number of images tested. To apply the curvelet transform, the image is resized to 128×128 as to get more number of features. The individual accuracy of the features correspond to each scale in quantifying the fatty liver is shown in Table. 3.1. The accuracy of the features correspond to wedges of first three scales are very low, while wedges correspond to scales 4, 5 and 6 performed moderately, and at the scale 7, the classification accuracy is very poor. The maximum classification accuracy is given by the features computed at the scale 5 by complex curvelet transform. The proposed features computed over complex curvelet transform gave maximum classification accuracy of 95.7% for cubic SVM classifier.
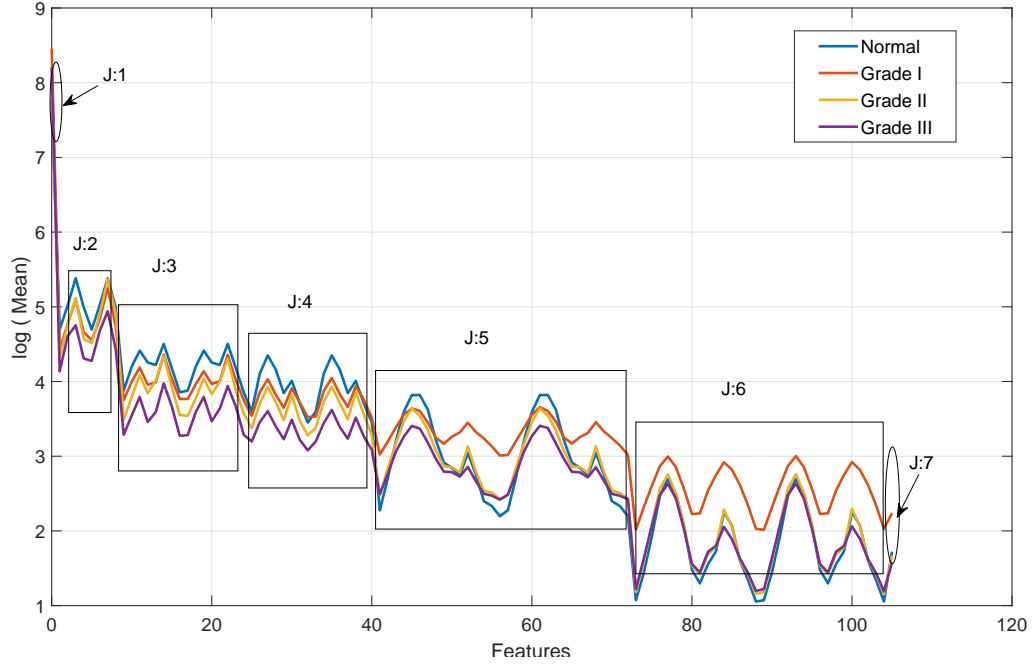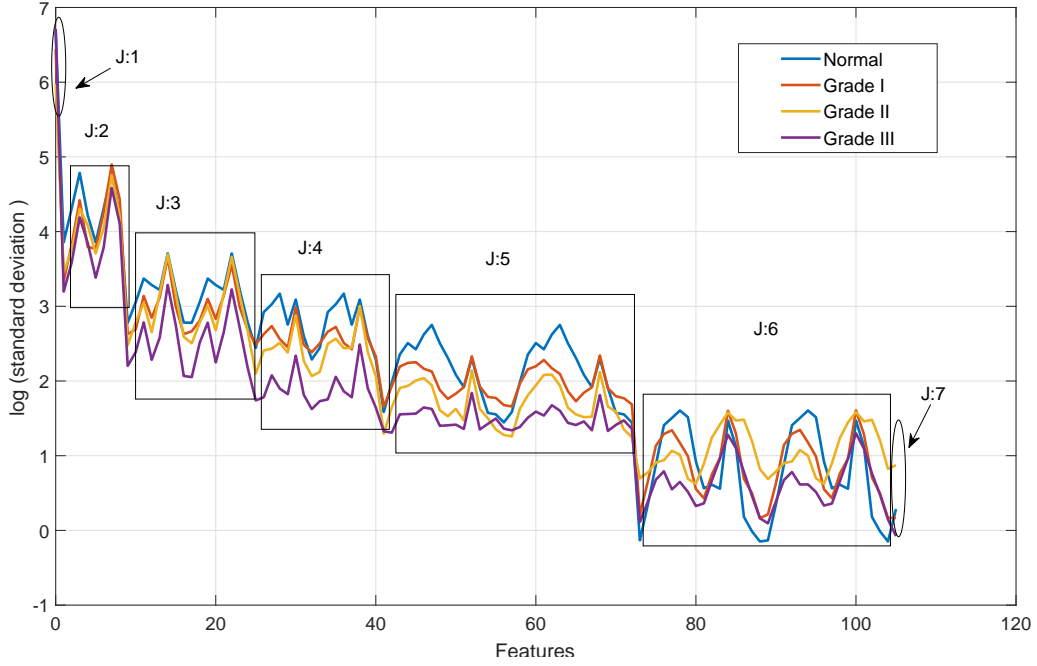
Figure 3.5: (a) *Log* mean and (b) *Log* standard deviation of the features of Normal, Grade I, Grade II and Grade III fatty liver is computed over 250 images of each class. *J* in the figure correspond to the features in each sub-band. Log of mean and log of standard deviation of features are considered for better visual representation. At scales J=2,3 and 4, the log mean of the Normal texture is found higher, followed by Grade I, II and III respectively, whereas the log mean of Normal, Grade II and Grade III found to be equal at J=5 and 6 respectively.

Table 3.3: Confusion matrix of the propose algorithm.

	Predicted Class			
True Class	Normal	Grade I	Grade II	Grade III
Normal (250)	242	7	0	1
Grade I (250)	7	228	5	10
Grade II (250)	0	3	247	0
Grade III (250)	0	7	3	240

3.4.1 Statistical Analysis of the Features

The Fig. 3.5, refer the log of mean and log of standard deviation for the proposed features computed over 250 images for each category. The features correspond to sub-bands at J=2, J=3 and J=4 we can infer a clear distinction in the mean of the sub-bands across all grades of fatty liver. The features of the normal liver had high mean followed by Grade I, Grade II and Grade III respectively. At scale J=5 and J=6 Grade II possessed high mean compared to Normal, Grade II and Grade III categories. High standard deviation is observed for all the curvelet transform based features.

3.4.2 Confusion Matrix

The confusion matrix for the proposed algorithm based on curvelet transform and SVD based features is shown in Table. 3.3. Normal, Grade I, Grade II and Grade III images classified with an accuracy of 96.8%, 91.2%, 98.8% and 96% respectively for complex curvelet coefficients.

From confusion matrix, we infer that seven images of the Normal liver are misclassified as Grade I fatty liver, one image as Grade III and a clear distinction is observed between Normal and Grade II classes. While classifying Grade I, twenty-two images are misclassified, out of which seven images are classified as Normal, five images as Grade II and ten images as Grade III respectively. Since Grade I lies between Normal and Grade II categories, misclassifications of Grade I to Normal and Grade II can be justifiable, but surprisingly ten images of Grade I are misclassified as Grade III.

In classifying the Grade II images, three images are misclassified as Grade I, and a clear distinction is observed between Grade II and Normal, Grade II and Grade III classes. Ten images of Grade III fatty liver are misclassified, out of which seven images are misclassified as Grade I and three images as Grade II respectively. It is worthy to note that only seven images of Grade I are misclassified as Normal, and no image of Grade II and, Grade III are misclassified as Normal which is crucial in medical diagnostics. The misclassification between the textures have mainly occurred between Grade I and Grade III classes. Considering all the three Grades I, II and III as positive cases (images with disease) and Normal images as negative cases, the propose algorithm resulted with a sensitivity of 95.3% (715 images out of 750 fatty liver images correctly classified as diseased) and specificity of 96.8% (242 images out of 250 normal liver images classified correctly as normal) respectively. The computations are performed on an Intel Core i7 processor with 16 GB RAM running with 2.8 GHz clock cycle using MATLAB. The proposed algorithm took 0.3 seconds to classify the texture image.

3.5   Summary

In this chapter, we formulated the grading of fatty liver as a texture discrimination problem and classified the texture by representing with a novel feature extraction scheme based on curvelet transform and SVD based features. From the experimental analysis, we can infer that classification accuracy is increased with more detail coefficients. The amount of detail information extracted from the curvelet transform is the constraint to the size of the image and hence we can not get further gradient information useful for classification. To overcome the limitation, in the next chapter, we propose a scattering network-based features for fatty liver texture classification. Scattering network provides as many number of features extracting different levels of information from the texture image useful for classification.

Chapter 4

NAFL Texture Classification based on Scattering Coefficients

4.1 Introduction

In this chapter, we propose an algorithm for fatty liver texture classification based on scattering coefficients (SC) features [62]. SC give stable invariant image representations for deformations, and preserves high-frequency information which is useful for classification.

4.1.1 Scattering Coefficients

The SC are invariant to global translations and Lipschitz continuous to local deformations, thus providing good features for classification [62]. SC are computed via invariant scattering convolution network (ISCN) architecture [63] as shown in Fig. 4.1. ISCN builds the invariant representations by progressively cascading wavelet transform with modulus and averaging operators. Since the high-frequency components are the main source of instability, the ISCN successively maps high-frequency wavelet coefficients to lower frequencies and averages the lower frequencies to get translational invariant representation. Deformation invariance to the features is obtained
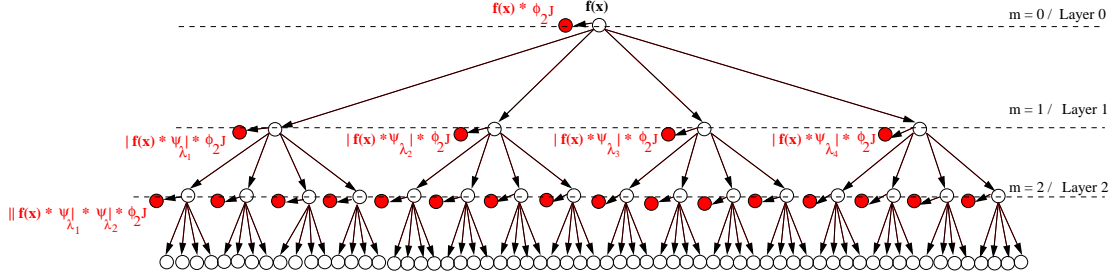
Figure 4.1: ISCN architecture. The red bubbles correspond to the SC matrix. * indicates the convolution operator and m defines the depth of the network.

by the wavelet transform since wavelets are localized waveforms which are stable to deformations. In ISCN architecture (shown in Fig. 4.1), the red bubble indicates the SC while the white bubble indicates the propagator signal $|.|$ on which low pass filtering generates the SC. Mathematically, SC are obtained by progressively cascading wavelet transform with modulus and averaging operators. In the zeroth layer of the network, the image is convolved with a low pass scaling function ϕ defined as

$$Layer\ 0,\ S_x(0) : f(x) \star \phi_{2^J}, \tag{4.1}$$

where $f(x)$ represents an image, $\phi(x)_{2^J} = 2^{-2J}\phi(2^{-J}x)$ represent Gaussian low pass filter, J denotes the largest scale space variable and $\star$ denotes the convolution operation.

In the first layer, modulus of the complex wavelet transform of an image is convolved with a low pass filter.

$$Layer\ 1,\ S_x(\lambda_1)\ :\ \mid f(x) \star \psi_{\lambda_1} \mid \star\ \phi_{2^J}, \tag{4.2}$$

where $\psi_{\lambda_1}(x) = 2^{-2j_1}\psi(2^{-j_1}r_\theta x)$ represents all the rotated and dilated versions of the band pass wavelets with $\lambda_1 = (2^{j_1}, \theta)$, $0 \leq j_1 < J$, j_1, is a scale space variable, and r_θ represent group of rotations computed as θ: $2\pi l/L$, where $0\leq l < $ L. The SC $S_x(\lambda_1)$ extracted at Eqn. 4.2 is equivalent to SIFT descriptors [64]. The $f \star \psi_\lambda$ computes gradient of an image at different directions and scales capturing high

frequency. Since high frequency is the source of variability, modulus operator applied on $f(x) \star \psi$ is interpreted as nonlinear demodulator that pushes wavelet coefficient energy towards lower frequencies, which is captured by ϕ_{2^J}. Additionally, modulus operator avoids the wavelet coefficients leading to zero while averaging which result in loss of information. The ψ is a complex wavelet which is of the form

$$\psi = \psi_a + i\psi_b, \tag{4.3}$$

where ψ_a and ψ_b denotes real and imaginary parts of wavelet. The modulus operator on complex wavelet is computed by

$$\mid \psi \mid = (\psi_a^2 + \psi_b^2)^{1/2}. \tag{4.4}$$

The translation invariance is obtained by averaging the modulus of complex wavelet coefficients. The averaging operation i.e., convolution of ϕ_{2^J} with $\mid f(x) \star \psi \mid$ will result in a shift invariance. Here, Morlet wavelet is chosen for ψ, defined as

$$\psi(x) = e^{-|x|^2/(2\sigma^2)}\alpha(e^{i.x.\zeta} - \beta). \tag{4.5}$$

The Morlet wavelet is obtained by multiplying the Gaussian window with a complex exponential. Here β is adjusted such that the area under the wavelet becomes zero. In simulations, the parameters σ, ζ are set to 0.8 and $3\pi/4$ respectively.

The SC of the second layer is obtained as

$$Layer\ 2,\ S_x(\lambda_1, \lambda_2)\ :\mid f(x) \star \psi_{\lambda_1} \mid \star\ \psi_{\lambda_2} \mid \star\ \phi_{2^J}, \tag{4.6}$$

where ψ_{λ_1} corresponds to all scales and rotations, while ψ_{λ_2} corresponds only to the scales $j_2 < j_1$, since the second order interference coefficients are negligible for $j_2 \geq j_1$ [62]. The coefficients obtained in the second layer extract co-occurrence information of the image at two scales 2^{j_1} and 2^{j_2} corresponding to two different

36

orientations, which is interpreted as a interaction coefficients, hence called as SC [65].

Similarly, the network is extended to deeper layers in the following way.

$$Layer\ m,\ S_x(\lambda_1 \ldots \lambda_m) : || \cdots | f(x) \star \psi_{\lambda_1} |\ \star \psi_{\lambda_2} | \ldots \tag{4.7}$$
$$\cdots | \star \psi_{\lambda_m} | \star \phi_{2^J}.$$

The high frequencies lost in averaging with ϕ are retained by convolving the signal with wavelet coefficients at higher scales, suppressing the phase, and averaging the signal with ϕ_{2^J}. Every time the image convolved with ψ or ϕ, the resultant image is down sampled by a corresponding filter scale 2^j to reduce the computations. The SC matrix obtained at each node is of dimension $[M/2^{J-1}, N/2^{J-1}]$, where M, N are the dimensions of an image. For the ultrasonic image of size $[78, 100]$ and for length $J = 4$, the SC matrix will be of dimension $[10, 13]$ with each coefficient locally invariant to a width of eight pixels. The ISCN network will result in 1, 32, 384 SC matrices from layer 0, 1 and 2 respectively, constituting a total of 417 SC matrices.

As the depth of the network increases, more invariant and discriminative features for classification can be obtained [63]. To visualize the discriminative nature of SC corresponding to the texture of different grades of fatty liver, the disk covering the entire frequency support of image is displayed as sectors as shown in Fig. 4.2 [63]. If $\psi(w)$ is the Fourier transform centered at frequency η, then $\psi(w)_{2^{-j}r}$ has a support centered at $2^{-j}r$ with bandwidth proportional to 2^{-j}. Each sector in the disk corresponds to local Fourier transform energy of an image over support of $\psi_\lambda(w)$. From Fig. 4.2, we can infer that SC give discriminative features for different grades of fatty liver both in the second and third layer.

The SC obtained after the first layer is embedded with information about the SC from where it is propagated from the previous layers. The efficacy of the features extracted from ISCN in classification lies in extracting the representations which are not present in the previous layer. To decorrelate the redundant information present
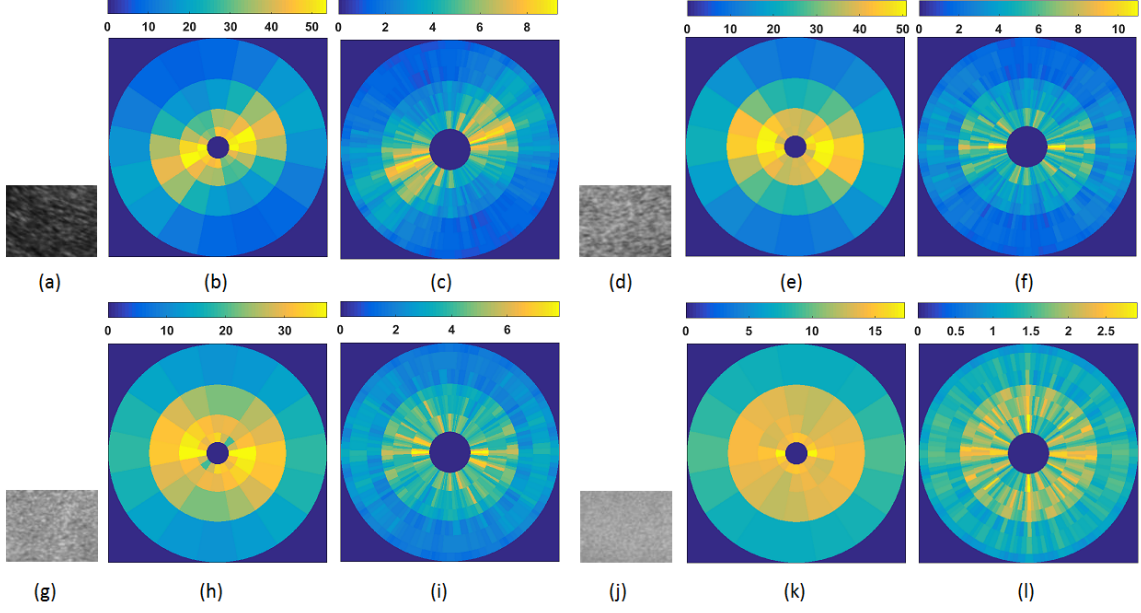
Figure 4.2: Perceptual difference observed in SC (4 scales and 8 orientations) of the fatty liver. (a), (d), (g), (j). Texture pattern of Normal, Grade I, Grade II, Grade III fatty liver respectively. (b), (e), (h), (k). First layer SC of Normal, Grade I, Grade II, Grade III respectively. (c), (f), (i), (l). Second layer SC of Normal, Grade I, Grade II, Grade III respectively.

in deeper layers, the SC are normalized by dividing children node $S_x(\lambda_1, \lambda_2)$ with corresponding parent node $S_x(\lambda_1)$, known as transfer SC [66].

4.1.2 Compressed Transfer SC Features

For image size [78, 100], we will obtain a 417 SC matrices (for $j=4$, $L=8$ and $m=2$). The dimension of the SC is approximately seven times greater than image size which makes it difficult to model or to train the classifier. In [67], authors have summed SC along rows or columns of SC matrix obtaining a vector representation for each SC matrix, here much more compact representation is obtained by summing all the SC present in each matrix S_q as given below

$$feature(q) = \sum_{m=1}^{M}\sum_{n=1}^{N} S_q(m, n), \tag{4.8}$$

here q indicates the index of the feature, m and n represent the index of the transfer SC in S_q. The S_q denotes the SC matrix extracted from all the networks layers.

The normalized SC along with the SC corresponding to zero-th and first layer are collectively used as features for classification.

The computational complexity of computing SC is in the order of $O((r_\theta/3)^m \, T \, log \, T)$, where $T = M \times N$ denotes the number of pixels present in an image.

The efficiency of SC features is evaluated with EKNN classifier. In EKNN classifier, instead of using a single classifier, multiple KNN classifiers with subspace of features per classifier have been used in classification. The optimal number of classifiers, the number of features per classifier and number of nearest neighbors used for each classifier are determined by cross-validation. The unseen feature is classified by combining the outputs of all the classifiers and decision is made on a majority voting rule.

4.2 Results

4.2.1 Optimal Scale and Orientation of a ISCN

To determine optimal ISCN architecture for achieving maximum classification accuracy, the propose algorithm is tested with various configurations of scales and orientations for a network depth $m = 2$. The accuracy of SC features with respect to different scales and orientations is shown in Table. 4.1. The results are presented for transfer SC features with EKNN classifier using ten-fold cross-validation scheme [30]. For EKNN classifier, number of learners is fixed to 30, subspace dimension of features for each classifier is taken half of its feature size. The classification accuracy increased with increase in scales and orientations. Maximum accuracy of 98.2% is achieved at four scales and eight orientations. Further increase in scales and orientations did not improve the classification accuracy. The maximum accuracy of 98.2% is further increased to 98.8% by optimal selection of parameters in the EKNN classifier.

Table 4.1: Accuracy(%) of the SC features for different scales and orientations with EKNN classifier, depth of the network $m = 2$ and size of the image 78×100. Features from all the layers are concatenated as a single feature vector and used in classification. For EKNN classifier, number of learners is fixed to 30, subspace dimension of features is taken half of feature size.

Scales (j)	Orientations (L)							
	1	2	3	4	5	6	7	8
2	74.4	88.1	90.1	91.4	90.9	92.2	93.3	94.8
3	83.2	91.0	93.0	95.5	94.9	96.5	96.9	97.4
4	88.3	92.9	94.6	95.2	96.2	96.9	98.0	98.2
5	87.6	93.3	95.4	96.1	96.5	97.6	96.9	97.8
6	91.5	94.6	95.5	95.9	96.2	96.1	95.6	96.2

4.2.2 Parameter Selection of EKNN classifier.

The optimal parameters for EKNN classifier are computed hierarchically. For example, SC features for 4 scales and 8 orientations with network depth 2, optimal parameters are computed in the following way. Initially, the optimal nearest neighbors are computed using the 10 fold cross validation as shown in Fig. 4.3a. For $K=1$, 2 and 5 nearest neighbors, the cross-validation error is low, and it is 3.9%, 4.2%, and 4.8% respectively, and error tends to increase with an increase in nearest neighbors. Since $K = 1$ is susceptible to noise, we preferred $K=5$ over $K = 1$ despite its low error loss. The Fig. 4.3b shows the 10 fold classification error with respect to the number of randomly selected features. KNN gave minimum classification error 1.2% when the number of randomly selected features is 93. The error rate is high if the number of randomly selected features are less and also tend to increase slightly if the number of randomly selected features are more than 93. The error rate is high when the number of features are less since less number of features does not provide enough information or clues for classification. As the number of features increases, information about data increases hence the error rate decreases and achieves minimum at one stage. Further increase in features will create a confusion in the data as the features will contradict each other resulting in increase in the error rate. By setting $K=5$ and the number of randomly selected features to 93, the optimal learners are computed as shown in Fig. 4.3c. For 63 learners, the EKNN resulted in the lowest error rate of

1.2%. However the same low error rate can be obtained for 65, 85 and more than 90 learners, we chose 63 learners in this case as the computational complexity increases with an increase in the number of learners.

4.2.3 Statistical Analysis of SC Features

The correlation between the SC features and different grades of fatty liver is analyzed by computing the mean and standard deviation of the SC of 250 images of each class. The mean and standard deviation of the SC with respect for different grades of fatty liver is shown in Fig. 4.4 and Fig. 4.5. It is observed that the mean of SC features is reduced with the increase in the severity of fat in the liver. This is justified as the SC quantifies the gradient information correspond to different scales and orientations. The smoothness of the ultrasonic texture varies proportionally with the severity of the fat [18], representing low gradient information and results in the low magnitude of SC. From the first layer statistics, we can infer a clear distinction is observed between the SC features across all grades of fatty liver. Similar to first layer SC features, the mean of second layer SC features reduced with an increase in the severity of fat, except the mean of Grade II fatty liver is found to be lower than the Grade III fatty liver.

4.2.4 Confusion Matrix

The confusion matrix of SC features with EKNN classifier is shown in Table. 4.2. The proposed algorithm classified Normal image texture with an accuracy of 98.8%, Grade I with an accuracy of 96.8 %, Grade II with an accuracy of 100%, and Grade III with an accuracy of 99.6% respectively. We can infer that only three images of Normal texture get misclassified, out of which two images are classified as Grade I, one image as Grade II, and a clear distinction is observed between Normal and Grade III images. While classifying the Grade I, eight Grade I images get misclassified, out of which two images are classified as Normal, five images as Grade II and one image
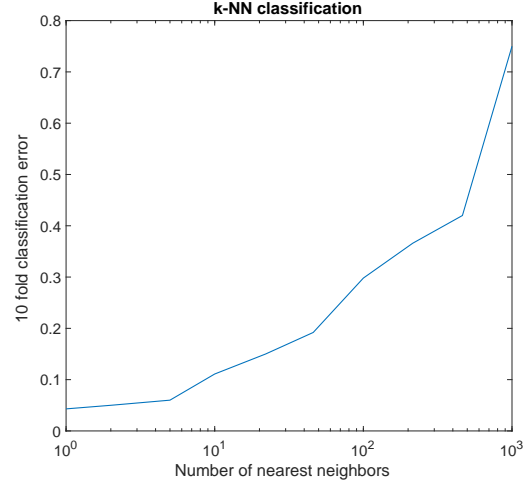
(a)
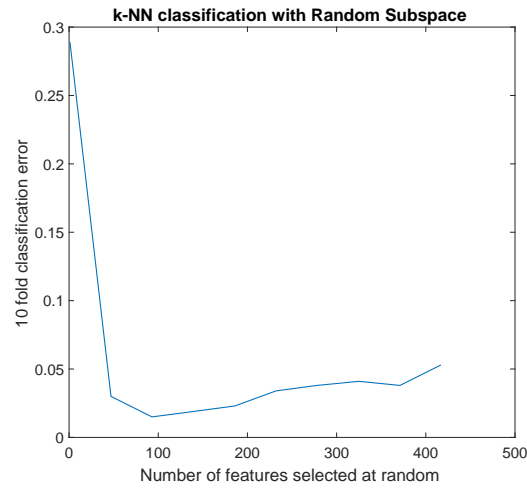
(b)
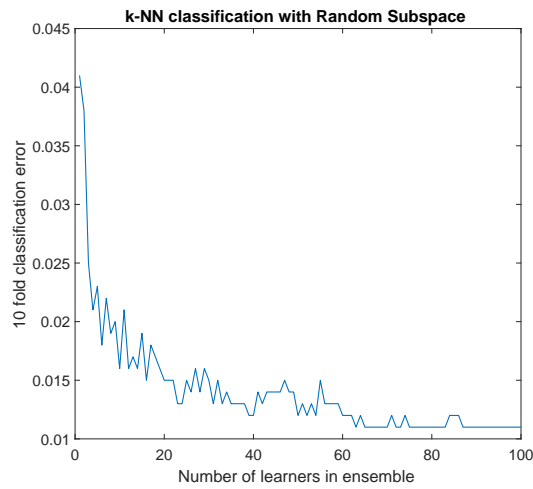
(c)

Figure 4.3: Optimal parameter selection of EKNN classifier for SC features (a). Performance of KNN with respect to K, (b) Performance of KNN classifier with respect to the random selection of the number of features, (c) Performance of KNN classifier with respect to the number of learners.
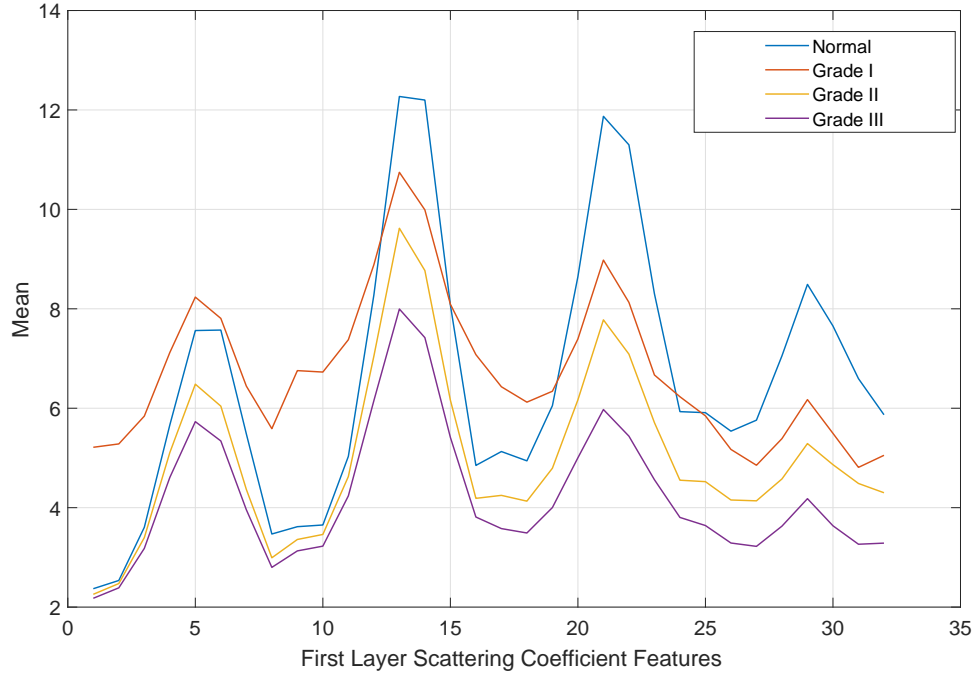
(a)
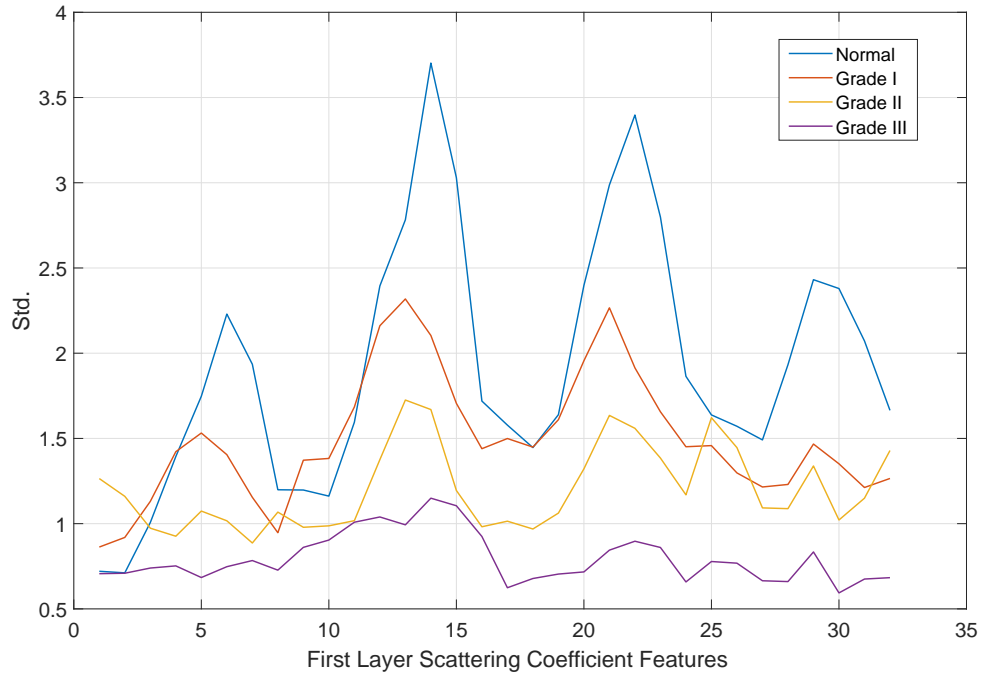
(b)

Figure 4.4: Correlation of first layer SC features with the diseases. (a) Mean and (b) standard deviation (Std.) of the first layer SC features for Normal, Grade I, Grade II and Grade III fatty liver computed over 250 images of each class.
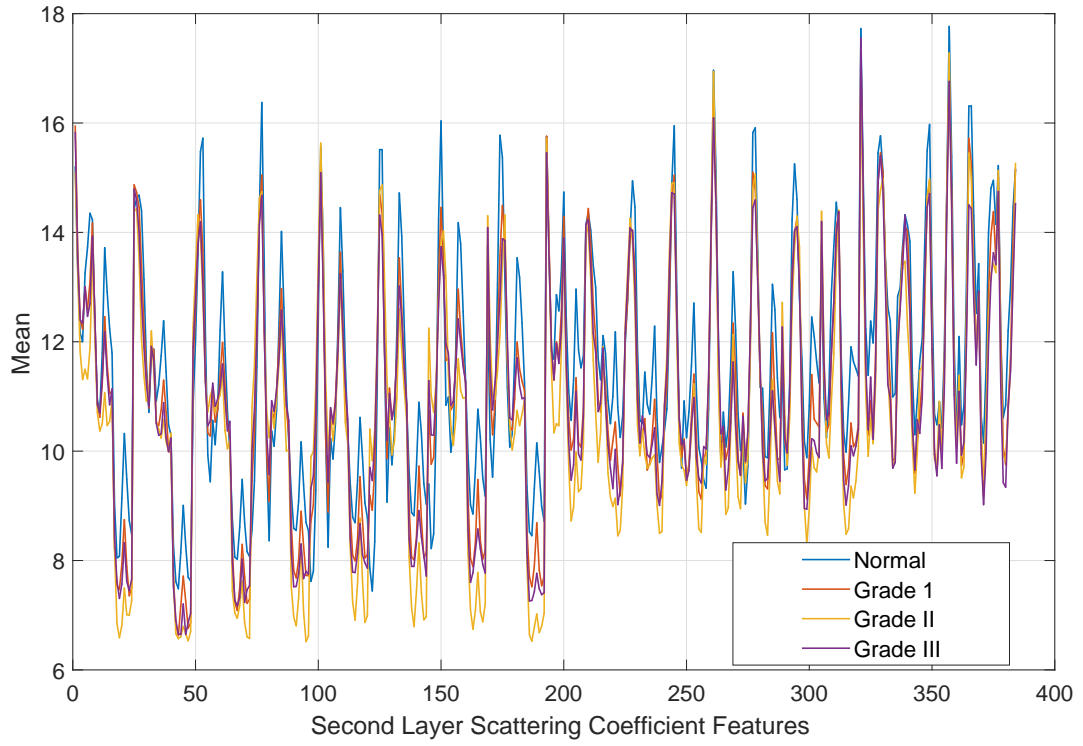
(a)
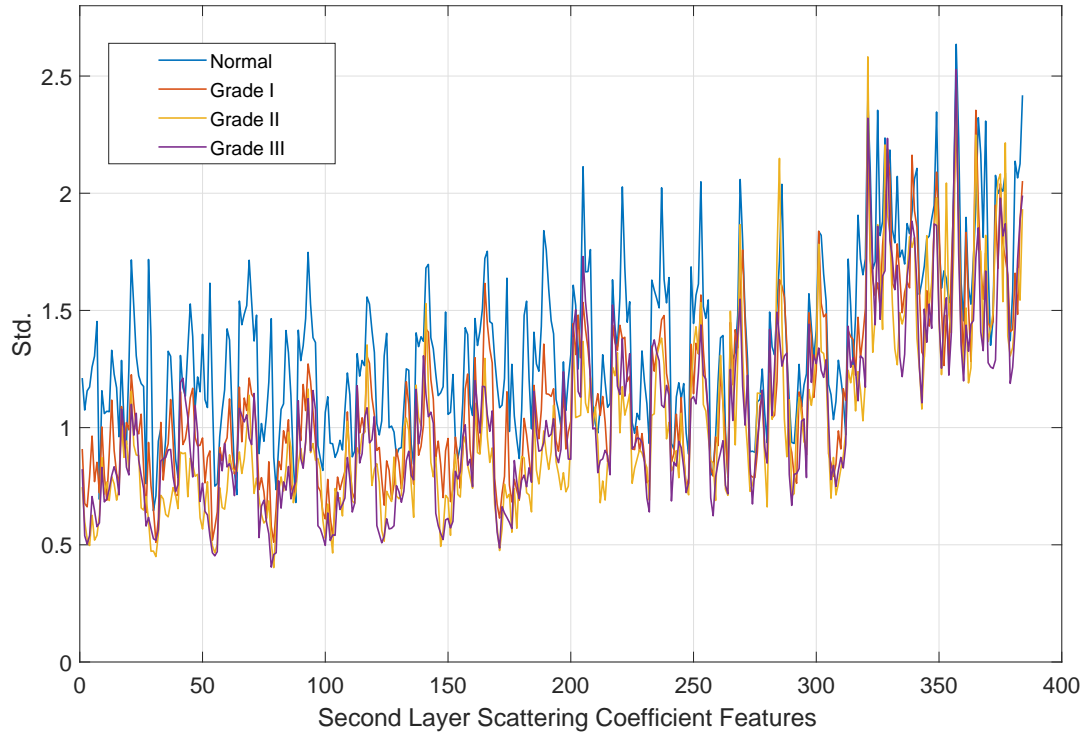
(b)

Figure 4.5: Correlation of second layer SC features with the diseases (a) Mean and (b) standard deviation of the second layer SC features for Normal, Grade I, Grade II and Grade III fatty liver computed over 250 images of each class.

44

Table 4.2: Confusion matrix for the SC features extracted for ISCN for network depth $m = 2$ with EKNN classifier.

True Class	Predicted Class			
	Normal	Grade I	Grade II	Grade III
Normal (250)	247	2	1	0
Grade I (250)	2	242	5	1
Grade II (250)	0	0	250	0
Grade III (250)	0	0	1	249

as Grade III. Since the characteristics of Grade I liver texture lies in between the Normal and Grade II the misclassification's of Grade I images to Normal and Grade II are justifiable. Grade II images classified with an accuracy of 100% and a clear distinction is observed between Grade II and Normal, Grade I and Grade III classes. For Grade III, only one image is misclassified as Grade II, and a clear distinction is observed between Grade III and Normal, Grade I classes.

Considering Grade I, II and III as positive and Normal images as negative classes, the proposed algorithm gave a sensitivity of 98.8% and a specificity of 98.8% . It is also worthy to note that no images of Grade II and Grade III are classified as Normal and Grade I, which is crucial in medical diagnostics since the patient needs medical attention under the circumstances.

4.3 Popularly used Texture Features

The performance of SC features in classification is compared with widely used ultrasonic texture features. The brief introduction regarding these features are explained below:

4.3.1 GLCM Features

GLCM features capture the spatial relationship between the pixels present in an image. The spatial relationship between the pixels is characterized regarding how often two pixels with intensities i and j occur in the specific direction and distance [68]. To extract the texture features, four GLCM's corresponding to directions $0^0, 45^0, 90^0$

and 135^0, with distance between two pixels as one unit is computed [40]. From each GLCM, 13 features were extracted, these include angular second moment, contrast, correlation, variance, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, and two features regarding information measures of correlation. For four GLCM's, a total of 52 features were computed.

4.3.2 GLRLM Features

The GLRLM features capture the texture information by computing the run-length of a pixel with specific gray value occurring in a specific direction [69]. Eleven GLRLM features corresponding to each direction $0^0, 45^0, 90^0, 135^0$ are computed, constituting a total of 44 features. The GLRLM features include short run emphasis, long run emphasis, gray-level nonuniformity, run length nonuniformity, run percentage, low gray-level run emphasis, high gray level run emphasis, short-run low gray level-emphasis, short-run high gray-level emphasis, long run low gray-level emphasis and long run high gray-level emphasis [70].

4.3.3 Laws Texture Features

In this approach, each image is convolved with a set of nine 5×5 masks to capture the local variation in the texture. The convolution masks are generated from the following vectors: $L_5 = [1\,4\,6\,4\,1]$; $E_5 = [-1\,-2\,0\,2\,1]$; $S_5 = [-1\,0\,2\,0\,-1]$; $R_5 = [1\,-4\,6\,-4\,1]$. The vector L_5 detect the center weighted local average, E_5 detect the edges, S_5 detect the spots, R_5 detect the ripples. The 2D convolution masks are obtained by computing the outer product of the vectors such as $L_5 E_5, L_5 R_5, E_5 S_5, S_5 S_5, R_5 R_5, L_5 S_5, E_5 E_5, E_5 R_5$ and $S_5 R_5$. For example, the mask $S_5 R_5$ is computed as $[-1\,0\,2\,0\,-1] \times [1\,-4\,6\,-4\,1]^T$. After applying these convolution masks on the image, the energy is computed over the coefficients of the convolved image resulting in nine features for single image [34].

4.3.4 GIST Features

GIST captures the gradient information with respect to different scales and orientations for different parts of the image giving a rough description of a surface [71]. GIST features are computed in the following way. Initially, the image is convolved with 32 Gabor filters generated by 4 scales and 8 orientations resulting in 32 feature maps. Each feature map is divided into 16 regions which are obtained by partitioning the image into a 4×4 grid, and coefficients in each region are averaged. All the averaged values corresponding to 32 feature maps are concatenated resulting in a total of 512 features. These features are used to represent the texture of liver parenchyma [72].

4.3.5 Multiresolution Features

Multiresolution feature extracts the texture information by decomposing the image into sub images with various resolutions. Multiresolution framework proposed in [39], [40] have been used for comparison. The images are decomposed using M-band wavelet and Gabor filter bank. From each decomposed sub-image, energy is computed from each sub-band, and considered as a feature for classification. Using M-band wavelet, the image is decomposed into 45 sub images, resulting in 45 M-band wavelet energy (Wav-Energy) features. Gabor filter bank with five radial frequencies ($\sqrt{2}/2^5, \sqrt{2}/2^4, \sqrt{2}/2^3, \sqrt{2}/2^2 and \sqrt{2}/2^1$) and six orientations ($0^0, 30^0, 60^0, 90^0, 120^0 \ and \ 150^0$) have been used for obtaining a total of 30 sub images. Energy in each Gabor sub images (Gabor-Energy) is computed and considered as a feature for representing the texture.

4.3.6 Comparative Analysis

The accuracy of popularly used texture features with respect to KNN, cubic kernel SVM and EKNN classifiers in grading the texture of fatty liver is shown in Table. 4.3. The number of nearest neighbors for KNN classifier is set to five, while for cubic kernel SVM classifier, the inbuilt SVM function available in the MATLAB 2017a version is

Table 4.3: Performance of the popularly used texture features in grading the fatty liver. CT+ SVD refers to the features proposed in chapter. 3

Features	KNN	SVM	EKNN
GLCM	86.7	92.3	91.4
GLRLM	85.8	92.4	90.4
Laws	82.9	88.5	91.6
GIST	80.8	90.1	88.7
Wavelet_Energ	87.5	90.4	88.2
Gabor_Energ	87.4	92.8	91.1
CT+SVD	93.7	95.7	94.5
SC features	94.0	96.5	98.8

used in the experiments.

GLCM and GLRLM feature performed better with SVM classifier resulting with an accuracy of 92.3% and 92.4% respectively, while Laws texture measure resulted in better accuracy for EKNN classifier with an accuracy of 91.6%. GIST descriptor performed better with 90.1% accuracy for the SVM classifier. The energy features extracted from wavelet sub-band and Gabor coefficients performed better with an accuracy of 90.4% and 92.8% with SVM classifier. The curvelet transform and SVD based features performed better than the texture features proposed in the literature with an accuracy of 95.7% with SVM classifier, whereas the SC features resulted with a maximum accuracy of 98.8% with EKNN classifier performing better than all the texture features. The translational invariant representation, frequency preservation of the network along with its discriminative representation for the textures proved very effective for achieving higher classification accuracy from SC features. This is justified as the INSN preserves the frequency information and extracts the different levels of gradient information from the texture. For example, the first layer SC gave SIFT type descriptor and second layer SC gives complementary information for classification. The first layer SC gave the similar performance compared to the Gabor wavelets in classifying the texture, but when we use the second layer SC along with first layer SC, the classification is significantly improved. The significant difference in classification compared to conventional wavelets is obtained due to second layer SC features.

4.4 Summary

In literature, researchers are mainly focused on classifying normal liver with respect to the fatty liver and not much emphasis is laid on further quantification of fat in the liver. Accurate quantification of fat in the liver is essential to access the functionality of the liver and to prevent the complications associated with fatty liver diseases. In this study, we tried to quantify the fatty content in the liver based on the discrimination observed in the characteristics of the homogeneous texture of liver parenchyma. We used SC based features to represent the texture of a liver which proved very effective in classifying the texture with EKNN classifier. The proposed algorithm classified the textures correspond to different grades of fatty liver with an accuracy of 98.8%. The performance of the SC features is compared with widely used texture features in representing the texture of ultrasonic liver and the results show that the proposed SC features perform better than the texture features proposed in the literature.

The proposed algorithm is developed based on the analysis of homogeneous texture of liver parenchyma, and hence to apply the proposed algorithm on the complete images, the user has to mark region of interest to apply the algorithm. In general, there are instances where the texture of liver parenchyma is heterogeneous, under those circumstances, the proposed algorithm may result in erroneous result. In these circumstances, analyzing the characteristics of the diaphragm and periportal veins will help in accurately quantifying the fat in the liver. Detecting these regions will help in developing robust automated diagnostic algorithms. In the next chapter, we propose an algorithm for detection of diagnostically significant regions pertaining to fatty liver.

Chapter 5

Detection of Diagnostically Significant Regions of NAFL

5.1 Introduction

The severity of fat in ultrasonic liver images is quantified based on characteristics of three regions in the image namely diaphragm, periportal veins and texture of liver parenchyma. The characteristics of these regions vary with the severity of fat in the liver, and is subjected to low signal to noise ratio, low contrast, poorly defined organ boundaries, etc., hence locating these regions in ultrasound images is challenging task for the sonographers and the accuracy of the diagnosis is mainly influenced based on the perception of these region of interest (RoI). As the characteristics of each RoI differ from other RoIs, we propose a specific algorithm for detecting each RoI. The accuracy of detecting diaphragm is computed as the efficiency of algorithm in placing at-least one bounding box on the diaphragm of liver parenchyma, while the accuracy of detecting periportal veins is computed as the ratio of total number of periportal veins detected to the total number of periportal veins present. The accuracy of the algorithm in detecting the homogeneous region is seen as its ability to place at-least one RoI inside the liver parenchyma.

The block diagram representation of the propose algorithm in detecting each RoI is shown in Fig. 5.1. The framework consists of three independent algorithms each specifically designed for detecting RoI of a diaphragm, periportal veins and texture of liver parenchyma. The description regarding the detection of each RoI is discussed in the following sections.

5.2 Diaphragm Detection

The diaphragm is the primary muscle used in the process of inspiration, or inhalation. It is a dome-shaped sheet of muscle present into the lower ribs. The diaphragm is situated above the liver and appears as slanted 'U' shape with an upper part trimmed in the liver ultrasound image. The shape, size and length of a diaphragm depends on the anatomy of patient and position of the probe used to acquire the liver images. Since the diaphragm lies above the liver, it appears in a lower left portion of the ultrasound image when captured from a subxiphoid view [73]. The block diagram representation of the propose algorithm for detecting RoI of a diaphragm is shown in Fig. 5.2. Initially, Viola Jones (VJ) algorithm [74] is used to detect the RoI of a diaphragm. VJ algorithm which is primarily proposed for detecting faces in real-time also proved that the same framework is effective in detecting other parts like noses, eyes, upper body parts, cars, stop signs, etc. Recently VJ algorithm has been used in medical image analysis to detect organs such as carotid artery, kidney, pelvis and proximal femur of a hip joint, etc [75], [76]. The VJ algorithm works in three stages namely feature extraction, AdaBoost training and cascade of classifiers. The brief description regarding the VJ algorithm is discussed below.

5.2.1 Feature Extraction

Haar-like features are derived from the kernels [77], few of these kernels are shown in Fig. 5.3. The white region in kernel correspond to weight w_0=-1 and black region
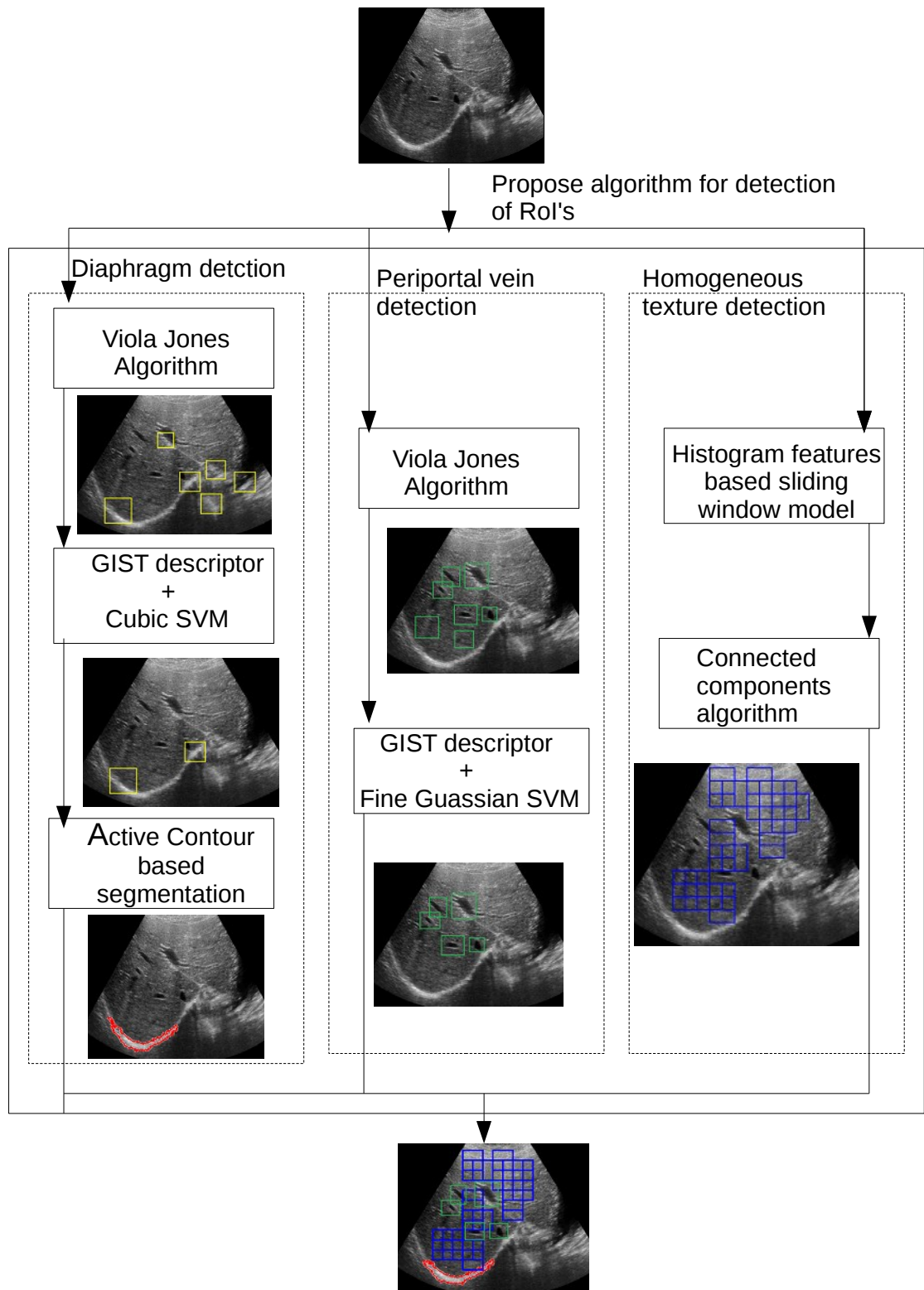
Figure 5.1: Block diagram representation of the propose algorithm for detection of RoI's in the liver ultrasound image.
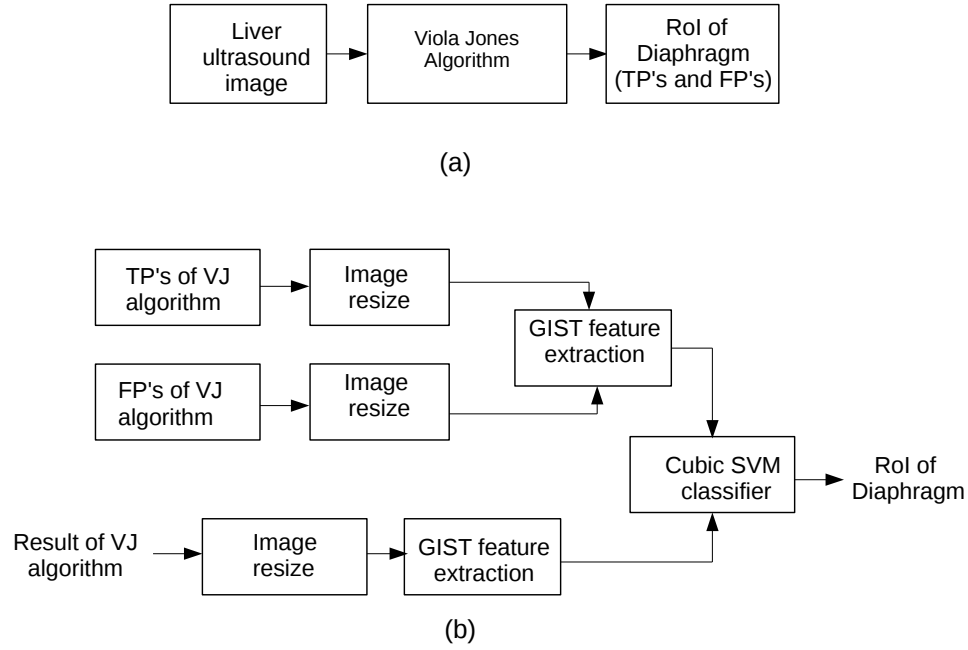
Figure 5.2: (a) Block diagram representation of the VJ algorithm used for detecting the RoI of a diaphragm and (b) Algorithm used to reduce the FP's resulted from VJ algorithm.

corresponds to $w_1 = +1$. The value of these features are then computed using the formula $f(x) = w_0 r_0 + w_1 r_1$, where $f(x)$ is the response of a given Haar-like feature to the input image x, w_0 is the weight of the black area r_0 and w_1 is the weight of the black area r_1. The number of pixels in areas r_0 and r_1 vary because the features are generated for various possible combinations and positions in a given window. These dimensions start from the single pixel and extend up to the size of given window.

The features generated using these kernels are independent of image content. The process of feature generation is explained as follows: considering kernel in Fig. 5.3(a), which is initially of two-pixel column width (one pixel white and one-pixel black), the feature value of $f(x)$ is computed. The kernel is shifted from the top left of the image by one pixel, and new feature value is calculated. Similarly, the kernel is then moved across the complete image until it reaches the right bottom of the image with

all features computed. Hence the features are evaluated hundreds of times as kernel moves across all the rows of the image and every time new feature value is updated in the feature list. Later the kernel is increased to the four-pixel width (two white pixels and two black pixels), and the process is repeated to get new feature values. The process is repeated for all the kernels, capturing all variations in size and position. Computing sum of pixels in a given area is a computationally expensive procedure. To address these an intermediate image representation called integral image approach is employed [78].
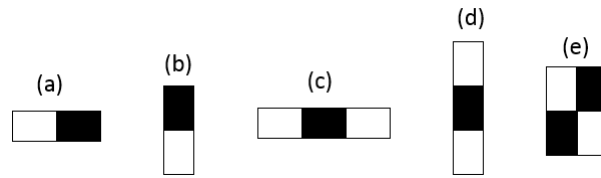
Figure 5.3: Kernels used to extract Haar like features from the images.

5.2.2 AdaBoost based Learning and Cascade Architecture

Evaluating a large number of features for detecting an object in an image is highly expensive, to address this an AdaBoost algorithm is used [79]. An AdaBoost learning algorithm is used to select the best features and to train the classifiers that use them. AdaBoost builds the strong classifier as a weighted linear combination of weak learners, where weak learners are chosen as decision stumps. To improve the detection rate, a cascade of classifiers architecture is used. Each cascade classifier is trained with a combination of weak learners. The complexity of detecting the RoI increases with an increase in the number of cascaded classifiers. The cascaded classifier architecture improves the detection rate by eliminating the non-RoI's in the earlier stages of the cascaded classifiers.

Initially, the VJ algorithm is trained with positive and negative training images. The negative and positive images used in training the VJ algorithm for detecting RoI of a diaphragm is shown in Fig. 5.4 and Fig. 5.5 respectively. The positive images are
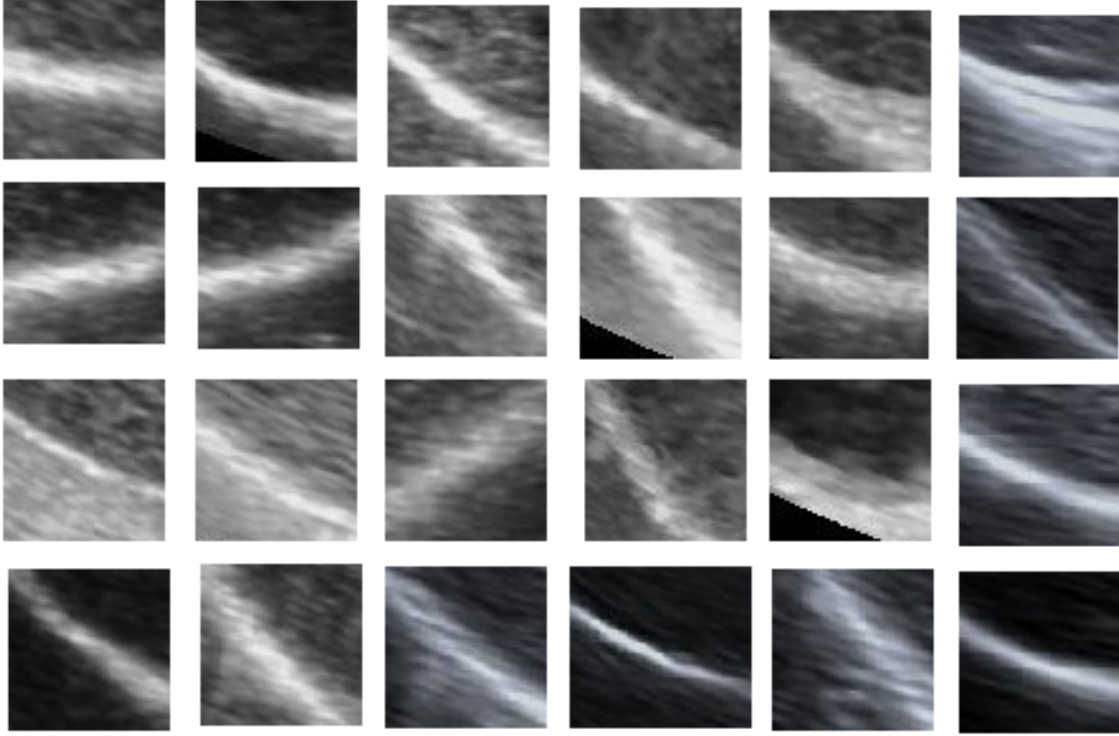
Figure 5.4: Images used as positive examples in training the VJ algorithm for detection of diaphragm.

extracted by cropping RoI of a diaphragm from overlapping regions of liver ultrasound images.

The size of the diaphragm used in training the VJ algorithm is varied in the range of 48×48 to 64×78 pixels. The entire region of the diaphragm is not considered for training due to the complex and varied structure of the diaphragm. While for the negative training images, care is taken that the images do not contain shapes appear as the diaphragm. We choose kidney ultrasound image as negative training examples since no part of the kidney image resemble like a diaphragm. The VJ algorithm is trained with 741 labeled diaphragms and 50 kidney images. 741 labeled diaphragms are obtained from 200 images by cropping the overlapping regions on the diaphragm. The RoI's detected by the VJ algorithm is shown in Fig. 5.6.

Although VJ algorithm results with high accuracy in detecting the diaphragm, it suffers from high false positives (FP's). While developing CAD algorithms, care should be taken that the algorithm will result in less number of FP's, since analyzing
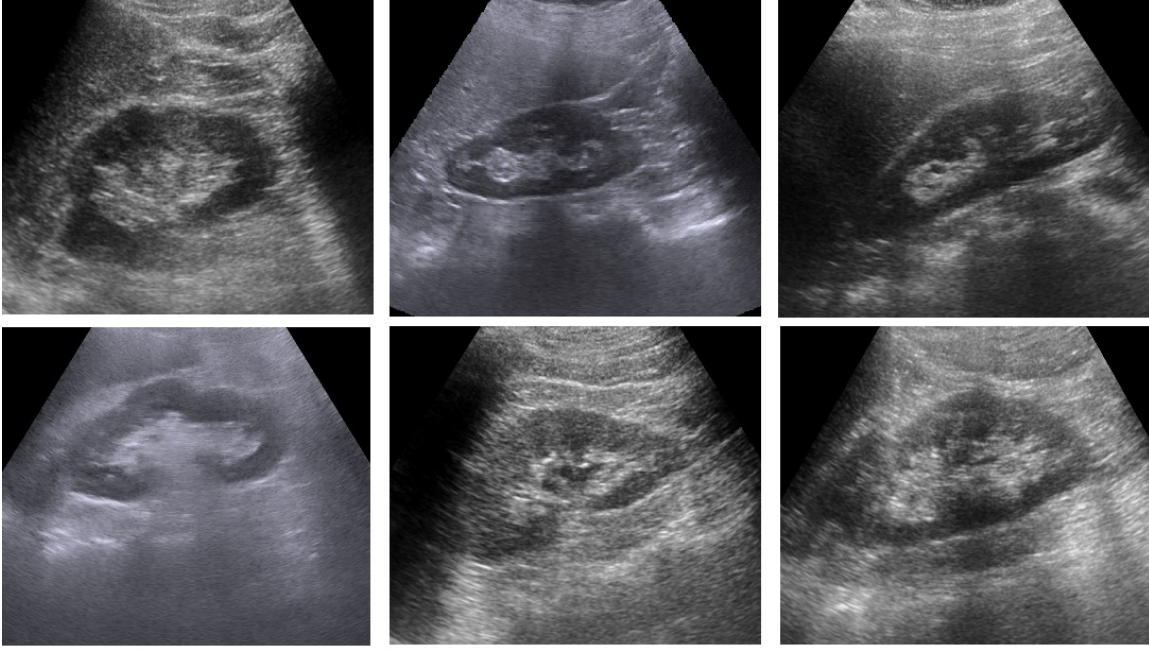
Figure 5.5: Images used as negative examples in training the VJ algorithm for detection of diaphragm. All these images belongs to the kidney.

the FP's for diagnosis will lead to the faulty diagnosis. The VJ algorithm mostly detected border of vein walls and nonhomogeneous texture regions as FP's. To eliminate FP's, the second stage of classification algorithm trained with true positives (TP's) and FP's resulted from VJ algorithm is developed. The TP's and FP's are obtained by testing the VJ algorithm on the training data. The second stage of the classification algorithm is developed based on the GIST descriptor and cubic SVM classifier. The features for classification framework is chosen based on the experimentation analysis. The classification framework is evaluated with popularly used texture features such as Histogram of Oriented Gradients (HOG) [80], histogram features and GIST descriptor [71] with SVM classifier, Out of all these features, GIST descriptor performed better in classifying TP'as and FP's.
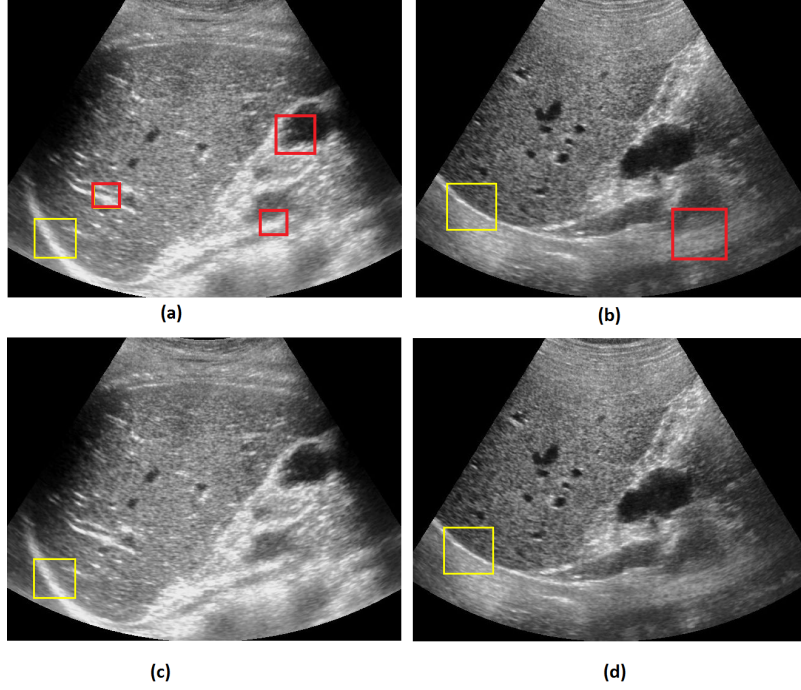
Figure 5.6: The red rectangular boxes indicates the FP's and the yellow rectangular boxes indicates the true positives. (a), (b) Result of the VJ algorithm in detecting the diaphragm. (c), (d) Elimination of the FP's after applying GIST based cubic kernel SVM classification.

5.2.3 GIST Descriptor based Cubic SVM Classifier

The block diagram representation regarding the classification between TP's and FP's is shown in Fig. 5.2(b). The TP's and FP's resulted from VJ algorithm is of different size. Hence in developing the algorithm, all the TP's and FP's images are resized to a fixed size of 64×64 pixels, this size is chosen based on the cross-validation. From each resized image, GIST descriptor is extracted. GIST descriptor gives the low dimensional representation of the scene by extracting the spatial envelope of an image [81]. GIST represent the features like naturalness, ruggedness, openness, roughness and expansion of a scene [82]. GIST descriptor from each image is extracted in the following way. Initially, the image is convolved with 32 Gabor filters corresponding to four scales and eight orientation resulting in 32 feature maps. Each feature map is divided into 4×4 grid resulting in 16 regions. The coefficients of each region is averaged resulting in 16 features for each feature map. The averaged values of all 32 feature maps will result in a total of $16 \times 32 = 512$ GIST descriptor. The extracted

features are then used to train the cubic SVM classifier. The classifier is trained with 250 TP's and 896 FP's images resulted from the VJ algorithm. With 5 fold cross-validation scheme, the algorithm resulted with an accuracy of 94.6% in classifying TP's and FP's.

After applying the GIST descriptor based cubic SVM classification, the FP's have been significantly reduced. But the output of this algorithm will not give the complete detection of the diaphragm. For complete detection of the diaphragm, active contour-based segmentation [83], active shape modeling (ASM) [84] algorithms can be used. ASM model fails due to high non-rigidness in the shape of the diaphragm. Hence we employed active contour-based segmentation for detecting the entire contour of a diaphragm.

5.2.4 Active Contour based Segmentation

Segmenting RoI in ultrasound image poses a lot of difficulties due to large variations observed in the diaphragm from image to image. The active contour model (ACM) [83] helps to obtain closed object contour as a segmentation result under the circumstances of noise and poorly defined boundaries. Active contour is a curve evolution algorithm which confirms to the image edges. In active contour, we initialize the curve and try to optimize it towards its lower energy. The energy of the curve is defined as

$$E(c) = E_{internal}(c) + E_{external}(c), \tag{5.1}$$

where $E_{internal}(c)$: depends on the shape of the curve given by

$$E_{internal}(c) = \int_0^1 \alpha ||c^l(s)||^2 ds + \beta ||c^{ll}(s)||^2 ds, \tag{5.2}$$

In Eqn. 5.2, the 1st term determines amount of stretch in the curve and 2nd term represents the smoothness of the curve. If the curve is relaxed it has low energy, suppose if it has stretched then the curve will have high energy. Similarly if the curve
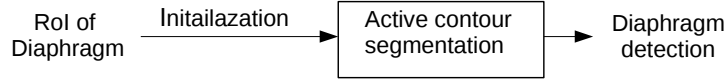
Figure 5.7: Block diagram representation for the active contour model based segmentation of a diaphragm.
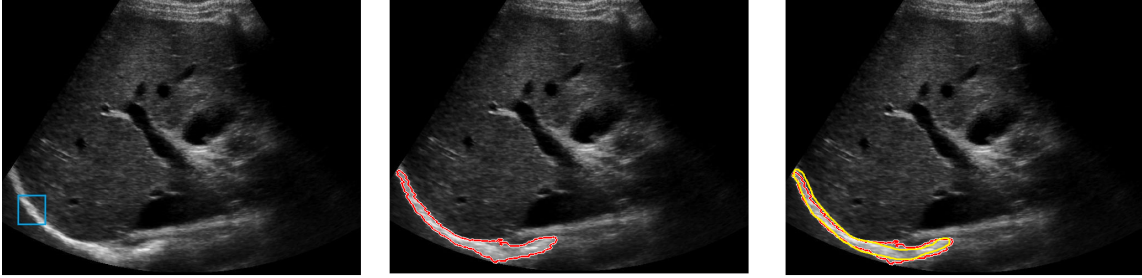
Figure 5.8: Blue box indicates manual initialization, red lines indicates the segmented diaphragm and yellow line indicates the manual segmentation. We can observe there is a strong resemblance between manual segmentation and active contour based segmentation.

is too curvy it has high energy and vice versa. Here α and β are the weights which controls stretch and smoothness of the curve. The $E_{external}(c)$ deals with the image intensity or edges of an image.

$$E_{external}(c) = \int_0^1 -||\nabla I(c(s))||^2 ds \qquad (5.3)$$

If there is no edge then $E_{external}$ becomes zero or if there is an strong edge then the $E_{external}$ will be lower due to negative of gradient of the image. Likewise the curve evolve towards the lower energy conforming to the edges of an image. The energy components in active contour are converted into forces. When the external and internal forces balance each other, we get the minimum energy. Contour displacement stops if minimum energy is reached or the specified number of iterations are over.

The block diagram representation for the active contour-based segmentation is shown in the Fig. 5.7. The result of active contour after manual initialization is shown in the Fig. 5.8. Here, we can see there is a close agreement between active contour segmentation and manual segmentation. Many variants of active contours like gradient vector flow [85], balloon model [86], diffusion snakes [87], active contours

with edges [83], geodesic active contours [88], etc., have been proposed in the literature for effective segmentation. Here, we employed active contours with edges algorithm for segmenting the diaphragm.

5.3 Periportal Vein Detection

The framework used to detect RoI for diaphragm also worked effectively in detecting the periportal veins. The segmentation algorithm is not employed for periportal vein detection since the entire vein is enclosed inside the RoI. Periportal veins are detected in two stages, in the initial stage, VJ algorithm is used to detect the periportal veins. The positive training images used to train the VJ algorithm is shown in Fig. 5.9. In general, more than one periportal vein vessels will appear in liver ultrasonic images. Hence all the visible periportal veins are considered in training the VJ algorithm. The size of cropped images employed in training the VJ algorithm is in the range of 15×25 to 32×42 pixels. The positive training images are extracted by cropping the periportal veins corresponding to all grades of fatty liver. We used the same negative training images (refer Fig. 5.5,) which we have used for training the VJ algorithm for detecting the diaphragm detection. Care is taken that kidney images used in negative training examples do not have cyst abnormalities since the cyst in kidney resembles as periportal vein in the liver ultrasound image. The VJ algorithm is trained with 829 labeled periportal veins and 50 kidney images. The VJ algorithm has resulted in both TP's and FP's as shown in Fig. 5.10. The regions where there is enough contrast in images are detected as FP's, since the periportal veins also provide similar contrast in their respective spatial locations. To eliminate the FP's, a second stage classifier algorithm with GIST descriptor and quadratic SVM classifier is employed. The second stage classifier algorithm is trained with TP's and FP's of the VJ algorithm which is developed for detecting the periportal veins. Before training the classifier, the sizes of all TP's and FP's are resized to a standard $32 \times$
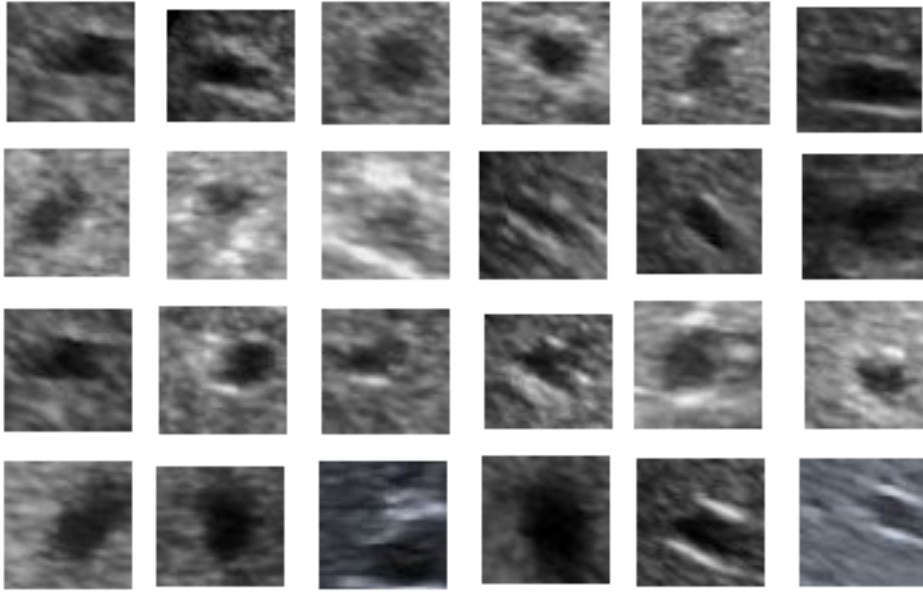
Figure 5.9: Images used in training the VJ algorithm for detecting the periportal veins.

32-pixel size, this size is chosen based on empirical analysis. The second stage of the classification algorithm is trained with 350 images of TP's and 200 images of FP's resulted from the VJ algorithm. With 5 fold cross-validation, the GIST descriptor based quadratic SVM classifier resulted in an accuracy of 93.7% in classifying TP's and FP's of periportal veins.

5.4 Homogeneous Texture Detection

The block diagram representation regarding the detection of RoI for homogeneous textures in liver parenchyma is shown in Fig. 5.11. To avoid homogeneous region lying outside the liver parenchyma, the organ is segmented manually before applying the algorithm. Since the texture does not have any shape specific information, the VJ algorithm performed poorly in detecting the homogeneous RoI in a liver. The liver mainly consists of periportal veins, the texture of liver parenchyma and diaphragm. Therefore, the RoI detection for homogeneous texture is framed as a binary classification problem, where one class belongs to RoI of homogeneous texture regions and
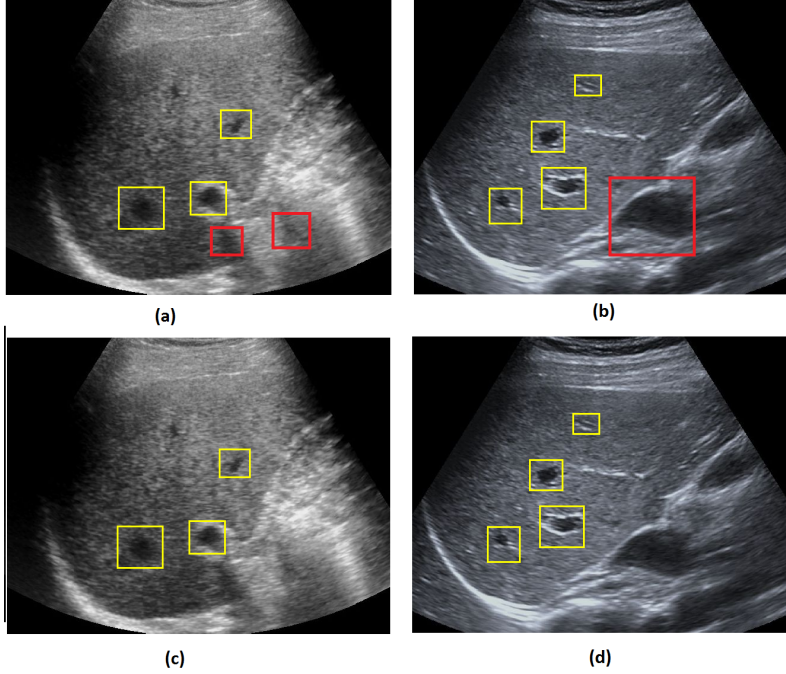
Figure 5.10: Automated periportal vein detection. Red boxes indicates the FP's and yellow boxes indicates TP's. (a), (b) Performance of the VJ algorithm in detecting the periportal veins. (c), (d) Performance of the algorithm after applying the GIST feature based quadratic SVM classifier.

another class belongs to RoI's of the diaphragm and periportal veins. The homogeneous texture is better represented with histogram features; these include mean, variance, skewness and kurtosis. The histogram appeared symmetry for homogeneous texture and nonsymmetry for nonhomogeneous texture regions giving discriminative representation across the homogeneous and nonhomogeneous RoI's. The homogeneous RoI's in liver parenchyma is detected using a sliding window based approach. A window of size 64×64 with the sliding length of 32 pixels is chosen for the analysis. The Gaussian SVM classifier is used to discriminate homogeneous and nonhomogeneous regions based on the histogram features. The homogeneous texture patches used in training the classifier is shown in Fig. 5.12. A total of 400 homogeneous textures and 550 diaphragm and periportal images extracted from all grades of fatty liver is used in training the Gaussian SVM classifier. The sliding window model resulted with 100% accuracy in detecting the homogeneous texture regions along with the FP's. The FP's are resulted due to periportal veins. Since the periportal veins

(a)

(b)
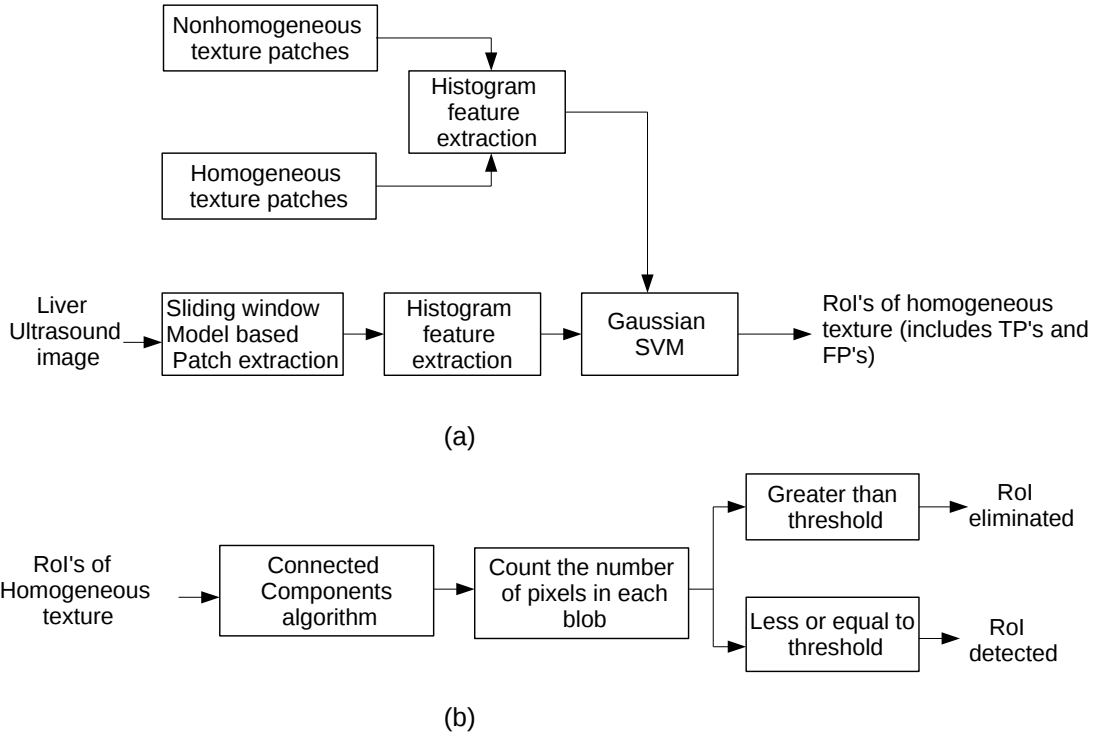
Figure 5.11: Algorithm for detecting RoI correspond to homogeneous texture of liver parenchyma.

appears as blobs, we eliminated the FP's by detecting the blobs in the homogeneous region. This is done by using a connected components algorithm [89]. The connected components algorithm is applied in the following way. Initially each detected homogeneous texture is binarized by setting a threshold to 20. The value 20 is chosen based on the empirical analysis. The pixels with intensity value less than 20 is set to '1', while the pixels with intensity greater than or equal to threshold 20 is set to '0'. Then the number of pixels in each blob is computed by fixing a connectivity to 6. It is observed that we can eliminate all FP's present in the liver parenchyma whose pixel count in each blob is more than 150. The optimal values for the connected components algorithm is chosen based on empirical analysis. The performance of the propose algorithm in detecting the homogeneous textures is shown in Fig. 5.13. The multiple RoI's are obtained for the homogeneous texture since the parenchyma covers wide area in the liver.
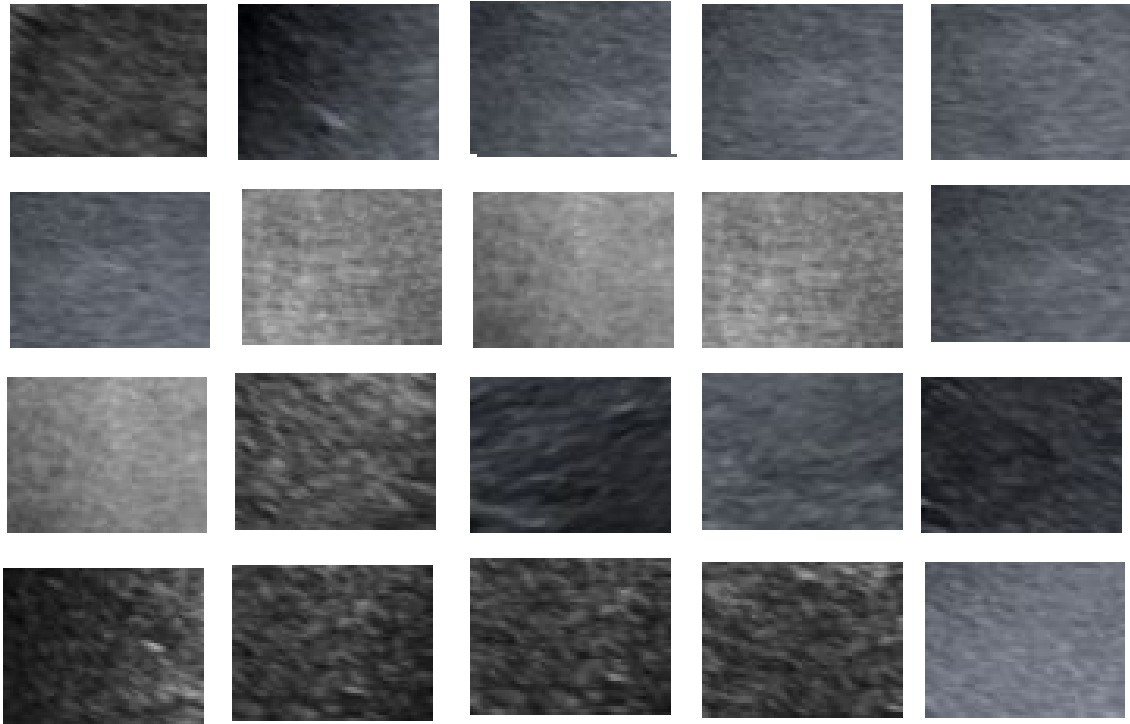
Figure 5.12: Homogeneous texture patches used to train the classifier for RoI detection of texture.
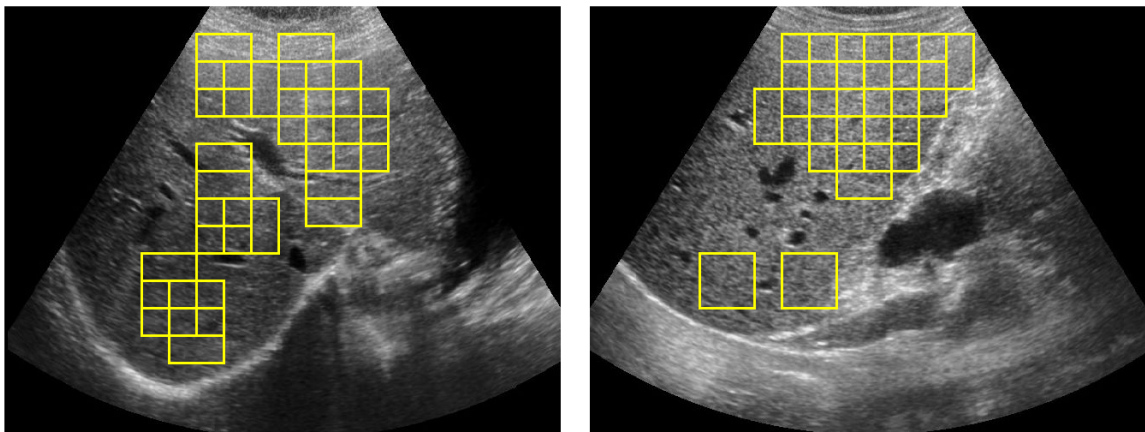
Figure 5.13: Performance of the propose algorithm in detecting RoI for homogeneous patches. Yellow boxes indicates the RoI of homogeneous texture patterns corresponding to liver parenchyma.

5.5 Database for Training and Testing

A total of 200 images were used in training process; these include 80 Normal, 45 Grade I, 40 Grade II and 35 Grade III images. The propose algorithm is tested with two sets of database, one consisting of 180 images which include 75 Normal, 60 Grade I and 45 Grade II images respectively, here Grade III images are not considered due to high obscuration of the RoI's. The second database consists 157 liver images. The images used in the first test case has not been used in the second test case. Performance of the proposed algorithm on the first database gives the overall accuracy of the algorithm in detecting the RoI's, while testing on the second database gives the performance of the algorithm on individual categories of different grades of fatty liver.

5.6 Results

The VJ algorithm is trained by setting following parameters: the number of cascade classifiers=5, merge threshold=80, window enlargement in each step is set to 1.1 with sliding window shifted by one pixel. The VJ algorithm detects the RoI using sliding window approach due to which multiple detections will occur for each RoI. The multiple detections are eliminated by merging all the overlapped detections by taking an average of the coordinates of detected RoI's. Decreasing the merge threshold increased the number of FP's along with TP's and vice-versa. The optimal merge threshold of 80 is chosen based on the empirical analysis. The VJ algorithm used 7, 9, 11, 13 and 19 features at each stage of cascaded classifiers respectively. The performance of the propose algorithm in detecting RoI for a diaphragm when VJ algorithm trained with different false alarm rates is shown in Table. 5.1. The algorithm is tested on 180 liver ultrasound images. The maximum classification accuracy of 97% is obtained at a false alarm rate of 0.5. Here the VJ algorithm detected the diaphragm with an accuracy of 98.2% with 450 FP's, after applying the GIST based cubic SVM classifier, FP's has been significantly reduced to 113.

Table 5.1: The accuracy of the propose algorithm for placing the RoI on a diaphragm.

False alarm rate	Performance of Viola Jones algorithm		Performance of the algorithm after applying GIST descriptor + cubic SVM classifier	
	Accuracy(%)	FP's	Accuracy(%)	FP's
0.1	91.0	160	85.0	14
0.2	93.5	194	87.4	21
0.3	94.7	260	89.7	44
0.4	96.0	390	96.0	78
0.5	98.2	450	97.0	113

Table 5.2: Accuracy of the propose algorithm in detecting the periportal veins

False alarm rate	Performance of Viola Jones algorithm		Performance of the algorithm after applying GIST descriptor + Quadratic SVM classifier	
	Accuracy (%)	False Positive	Accuracy (%)	False Positive
0.1	89.7	84	84.0	9
0.2	91.7	131	87.0	15
0.3	93.0	141	89.6	16
0.4	94.9	314	90.7	25
0.5	96.0	425	91.0	43

The performance of propose algorithm in detecting the periportal veins is shown in Table. 5.2. The VJ algorithm used 7, 8, 10, 11 and 14 features in each cascaded stage respectively. The optimal trade-off between TP's and FP's is obtained at a merge threshold of 50. The propose algorithm detected periportal veins with a maximum accuracy of 91% for the VJ algorithm when trained with a false alarm rate of 0.5. The VJ algorithm detected periportal veins with a maximum accuracy of 96% with 425 FP's. After applying GIST descriptor with quadratic SVM classifier, the number of FP's have significantly reduced to 43, which in turn also reduced the overall detection accuracy of periportal veins from 96% to 91%.

The histogram based Gaussian SVM classifier detected the RoI for homogeneous texture with 100% accuracy. The algorithm resulted in multiple RoI's for homogeneous textures along with FP's. The FP's are completely removed using the connected components algorithm.

Table 5.3: Accuracy of the propose algorithm in detecting the RoI's with respect to different grades of fatty liver.

Category	Database	Diaphragm (%)	Periportal veins (%)
Normal	48	97.5	93.7
Grade I	33	95.5	92.3
Grade II	45	93.2	90.8
Grade III	31	65.7	80.3

5.6.1 Performance of the Propose Algorithm with Respect to Different Grades of Fatty Liver.

The performance of the proposed algorithm in detecting the RoI with respect to different grades of fatty liver is presented in this section. The proposed algorithm is tested on a database consisted of 157 ultrasound images which include 48 normal, 33 Grade I, 45 Grade II and 31 Grade III images. For Normal, Grade I and Grade II images, the diaphragm and periportal veins are visible or partially visible to the normal eye. For 31 images of Grade III fatty liver, only for 15 images the diaphragm and periportal veins are partially visible, whereas for other images diaphragm and periportal veins are totally obscured and is not visible to the human eye. The accuracy of the proposed algorithm corresponding to different grades of the fatty liver is shown in Table. 5.3. The proposed algorithm detected diaphragm and periportal veins of the normal liver with an accuracy of 97.5% and 93.7% respectively. The accuracy of detecting diaphragm and periportal veins get reduced with an increase in the severity of the fat which resulted in obscuration of RoI. Even there is a high obscuration of RoI for Grade III fatty liver, the proposed algorithm detected diaphragm and periportal veins with an accuracy of 65.7% and 80.3% respectively. The classification accuracy is only computed for the images where diaphragms and periportal veins are perceivable to the human eye. For the images where diaphragm and periportal veins are not visible, the algorithm did not detect any RoI in the region with zero FP's. The proposed algorithm detected RoI for homogeneous patches with a 100% accuracy for all the liver ultrasound images. Since the organ boundaries are not visible in

Table 5.4: A brief summary regarding the performance of the propose algorithm in detecting each RoI.

RoI detection	Remarks
Diaphragm	The bounding boxes enclosing some part of the diaphragm is detected using the VJ and GIST descriptor based cubic SVM classifier. The detected RoI is then used to initialize the active contour algorithm to detect the entire diaphragm. The bounding box position of RoI enclosing the diaphragm did not have an effect in detecting the entire contour of diaphragm since the active contour algorithm tends to capture the strong edges. The accuracy of detecting diaphragm decreased with increase in severity of fat in the liver due to obscuration of diaphragm.
Periportal veins	The bounding boxes enclosing periportal veins is detected using the VJ and GIST descriptor based quadratic SVM classifier. All the detected RoI's are completely enclosed within the RoI. The accuracy of detecting periportal veins decreased with increase in severity of fat in the liver due to obscuration of periportal veins.
Homogeneous texture	The RoI's are detected using the combination of histogram features based Gaussian SVM classifier and connected components algorithm. The algorithm resulted in multiple overlapping RoI's, and the end user can select an appropriate number of RoI's depending on the application. The accuracy of detecting RoI for homogeneous texture did not effected with the severity of fat present in the liver.
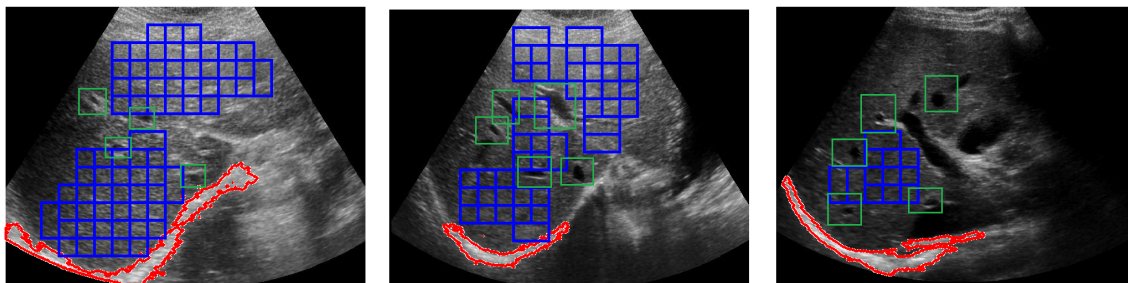
Figure 5.14: Performance of the propose algorithm in detecting the RoI's. The region enclosed with red color represents the diaphragm, green boxes indicate the detection of periportal veins and blue boxes indicates the RoI of a homogeneous texture.

Grade III fatty liver ultrasound images, the RoI's detected outside the liver image is not recognizable. The brief remarks regarding the performance of the propose algorithm in detecting each RoI is discussed in Table. 5.4. The result of the proposed algorithm in detecting the diaphragm, periportal veins and homogeneous texture of liver parenchyma is shown in Fig. 5.14.

5.6.2 Segmentation Accuracy

The accuracy of segmenting the diaphragm and periportal veins is measured using the Dice coefficient which is given by

$$D = \frac{2(A \cap B)}{|A| + |B|}, \tag{5.4}$$

A and B represents the segmentation resulted due to two different procedures respectively. $|.|$ represents cardinality. D lies in between 0 and 1, the higher the value better is the agreement between the segmentation procedures. For diaphragm detection, the mean D for manual initialized active contour segmentation to the manually segmented diaphragm is 0.60 is obtained, for manually initialized active contour to automated segmentation a mean D of 0.71 is obtained and for manual segmentation to the automated segmentation a mean D 0.59 is obtained. The mean D values are computed for 32 images, which consist of 10 Normal, 10 Grade I, 10 Grade II and 2 Grade III ultrasound images. A difference of 0.1 is observed in mean D between manually initialized active contour and automated segmentation procedure showing a close resemblance between manually initialized and automatically initialized active contour segmentation. While in the detection of periportal veins, the proposed algorithm resulted with a mean D of 0.71 when tested on 40 labeled periportal veins, which consist of 10 labeled periportal veins in each Normal, Grade I, Grade II and Grade III.

The experiments are performed on a Windows desktop computer with Intel Core i7 processor, 16 GB RAM running with 2.8 GHz clock using MATLAB 2017a version. For an image of size 500×650, the algorithm took approximately 2.42 s for detecting RoI of a diaphragm, 2.5 s for segmenting the diaphragm after initialization, 4.3 s for periportal detection and 2.3 s for homogeneous texture detection.

5.7   Summary

In this chapter, we proposed an algorithm for detection of diagnostically significant regions of fatty liver namely diaphragm, periportal veins and homogeneous texture in ultrasonic liver images. Since the characteristics of these regions vary with the fatty content of the liver, it poses a serious challenge in detecting these regions. The proposed algorithm detected RoI of diaphragm, periportal veins and homogeneous texture of liver with an accuracy of 97%, 91% and 100% respectively. The accuracy of the proposed algorithm in detecting the RoI's in fatty liver decreased with increase in the severity of fat in the liver, this is justifiable since the visibility of diaphragm and periportal veins reduces with the severity of the disease. In literature, computer-aided algorithms developed for diagnosing the fatty liver involves manual intervention to crop the RoI's, hence making the algorithms semi-automated. The proposed algorithm eliminates the need for manual intervention except for homogeneous texture detection, which can be used to develop accurate and automated diagnostic algorithms. The proposed algorithm benefits the sonographers to diagnose more number of patients by reducing the time needed to locate RoI's in the ultrasound image. The proposed algorithm can also be used to assist the semi-skilled persons to know whether the scanned image has representative information for diagnosis before transmitting it to the expert or cloud for analysis.

Chapter 6

Sparseland Model for Speckle Suppression

6.1 Introduction

Ultrasound images are inherently affected by a multiplicative noise called speckle. The speckle appears as small worms or snake like structures in the image revealing no significant information about the tissue structure and so it is considered as noise. Speckle is uncorrelated with respect to spatial, temporal and frequency of operation of ultrasound probe. The uncorrelated property of speckle is used to suppress the noise at RF acquisition stage using different compounding techniques [90]. In frequency compounding, the same imaging sector is scanned with multiple frequencies and the corresponding RF signals are averaged. Phase compounding involves averaging the RF data acquired by scanning the same position with multiple probe positions [91]. Compounding techniques are employed in the ultrasound machines at a cost of increased computations. Speckle suppression also done after complete formation of an ultrasound image by estimating the random behavior of speckles in the image.

Mathematically speckle can be modeled as sum of large number of complex phasors which results in constructive and destructive interference at the receiver side [92]. The

constructive interference leads to bright spots and destructive interference leads to dark spots similar to dense *salt and pepper* like noise in the image. The intensity of envelop detected RF data *J(m, n)* affected by speckle noise is given as

$$J(m, n) = (P(m, n) * I(m, n)) N_\times(m, n) \qquad (6.1)$$

where the multiplicative noise $N_\times(m, n)$ is sample wise independent of past, future samples and uncorrelated to the image pixel value *I(m, n)*, *P(m, n)* is the point spread function (PSF) of the ultrasound imaging system, (m, n) represents the spatial position of the tissue in the scan plane. The log transformation, which is used to compress the dynamic range of envelope detected data in ultrasound imaging system modify the multiplicative model into an additive model.

$$J(m, n)_{al} = (P(m, n) * I(m, n))_{al} + N_+(m, n) \qquad (6.2)$$

$N_+(n, m)$ is the additive noise term and suffix al represents the term after log compression. The behavior of $I(m, n)$ in fully formed image is modeled as summation of complex phasors in a random walk model which results bright pixels due to constructive interference and dark pixels due to destructive interference [92].

$$I(m, n) = \sum_1^p a_p(m, n) \, e^{j\varphi_p(m, n)} \qquad (6.3)$$

p is a positive integer which is generally considered very large, a_p and φ_p are amplitudes and phases of scattering echoes from tissues.

Smoothing the image is one common solution seen in literature to address speckle suppression. These techniques differs from one another based on smoothing criteria employed to determine the degree of smoothness. The operation of the filter on a particular pixel depends upon local statistics of the pixel surrounding around it.

These filters are biased to the size of the window used for finding local statistics and smoothing.

Sparse and redundant representations over learned dictionaries looks promising for image deblurring [93] and denoising applications [94]. In this thesis, we show how sparse and redundant representation over learned dictionaries of speckle affected ultrasound image effectively leads to speckle suppression. The propose algorithm is tested individually on coherent, diffused speckle pattern images and real time ultrasound B-mode (Liver and Kidney) images. Coherent speckles in the ultrasound image appears as dense *salt and pepper* like noise and diffusion speckle pattern appears as *small worm* like structures. The speckle suppression is performed on images, which is generated from Matlab using K-wave toolbox [95]. The performance of the algorithm is evaluated based on how efficiently it can smooth the image by preserving mean and reducing the standard deviation in homogeneous region. The performance is compared with widely used speckle reduction filters like Frost [96], Lee [97] and Speckle reduction anisotropic diffusion (SRAD) filters [98].

6.2 Representation of an Image Over Sparse and Redundant Over-Complete Dictionary.

6.2.1 Problem formulation

Sparse modeling is used in many image processing applications like denoising, inpainting, mosaicing etc. Speckle suppression in ultrasound images is mathematically formulated in the following way.

$$\min f(\underline{X}) = \frac{1}{2} \parallel \underline{X} - \underline{Y} \parallel_2^2 + G(\underline{X}) \tag{6.4}$$

$\underline{Y}$: Given measurements.

$\underline{X}$: Unknown to be recovered.

$\underline{Y}$ is the speckle image and we want to recover clean image $\underline{X}$. We do not want to recover the image that is not to far from the noisy image and that is the penalization that we have here i.e., $\| \underline{X} - \underline{Y} \|_2^2$. $\| \underline{X} - \underline{Y} \|_2^2$ gives the mean square error between the speckle image and restored image. In Eqn. 6.4, $\| \underline{X} - \underline{Y} \|_2^2$ also seen as variance of speckle noise. The image that minimizes Eqn. 6.4 is noisy image itself and so we have not done much. The second term in Eqn. 6.4 indicates the prior information regarding the image that has to be recovered. This is simply a Bayesian point of view, adapting the Maximum-A-Posterior (MAP) estimation and the basic idea is computing the $\underline{X}$ that minimizes $f(\underline{X})$. Eqn. 6.4 is seen as prior and likelihood estimation with probabilistic interpretation. In this scenario we choose a prior to be

$$G(\underline{X}) = \lambda \| \alpha \|_0 \qquad (6.5)$$

where $\| \alpha \|_0$ gives sparsity of the signal representing number of non-zero coefficients.

6.2.2 Sparseland Modeling of an Image.

To construct sparse modeling of an image, we will begin by constructing the sparse model for overlapping fixed image patches [94]. An image patch of size $\sqrt{m} \times \sqrt{m}$ pixels is arranged lexicographically as single column $x \in \Re^m$. To define a sparseland model for this column vector, we need to construct a dictionary of size $D \in \Re^{m \times k}$. If $k > m$ then the dictionary is said to be overcomplete and it is redundant. We simplify the model assuming the matrix D is fixed and it is known to us. Every image patch x in the image is represented sparsely over the dictionary by solving

$$\hat{\alpha} = arg \min_{\alpha} \ \| \alpha \|_0 \quad s.t \quad D\alpha \approx x \qquad (6.6)$$

s.t stands for such that.

we will get $\parallel \alpha \parallel_0 \ll m$. Mathematically speaking, the signal is represented by a linear combination of few column vectors of redundant dictionary matrix D. The constraint $D\alpha \approx x$ is equivalent to $\parallel D\alpha - X \parallel_2^2 \leq \epsilon$, ϵ is the magnitude of error allowed. We need to define the parameter Q such that $\parallel \alpha \parallel_0^0 \leq Q \ll m$. It states that it uses atmost Q columns of D for representing the image patch.

We have (ϵ, Q, D) with us and need to find the sparse representation $\hat{\alpha}$

$$\hat{\alpha} = arg \min_{\alpha} \parallel \alpha \parallel_0 \quad subject \quad to \quad \parallel D\alpha - Y \parallel_2^2 < T_h \qquad (6.7)$$

T_h is the error allowed and the recovered image is given by $D\hat{\alpha} = \hat{X}$. The optimization of Eqn. 6.7 is an NP hard problem and cannot be solved as it is. Eqn. 6.7 can be solved using relaxation and greedy algorithms by modifying the equation as

$$\hat{\alpha} = arg \min_{\alpha} \parallel D\alpha - Y \parallel_2^2 \quad + \quad \mu \parallel \alpha \parallel_1 \qquad (6.8)$$

Orthogonal matching pursuit (OMP) optimization algorithm is employed for obtaining the $\hat{\alpha}$ due to its efficiency and simplicity. .

6.2.3 Despeckling of Image from Local Patches.

Despeckling an ultrasound image Y of size $\sqrt{M} \times \sqrt{M}$ where $M \gg m$ is done by constructing a larger dictionary. The larger dictionary is obtained by just scaling the dictionary of image patch containing the basis of curvelet or contourlet transforms [99]. This is not possible if we use small and fixed dictionary $D \in \Re^{m \times k}$. We can solve this problem by another way by tiling the results of all patches in the image forming the complete despeckled image. Blocking artifacts is seen on the resulted image due to tiling of patches and is overcomed by constructing the dictionary for overlapping patches and averaging the results of the patches accordingly [100].

Considering every patch in the image belonging to sparseland model (ϵ, Q, D),

MAP estimator can be rewritten as

$$\{\alpha_{ij}, \bar{X}\} = arg \min_{\alpha_{ij}, X} \lambda \parallel \underline{X} - \underline{Y} \parallel_2^2 + \sum_{ij} \mu_{ij} \parallel \alpha_{ij} \parallel_0 + \sum_{ij} \parallel D\alpha_{ij} - B_{ij}\underline{X} \parallel_2^2 \quad (6.9)$$

The first term in the equation represents the error between the recovered image and noise image. B_{ij} is binary image used to extract the patch of an image of size $\sqrt{m} \times \sqrt{m}$. Eqn. 6.9 has two unknowns $\alpha_{ij}, \underline{X}$ and is solved by fixing $\underline{X} = \underline{Y}$ and seeking for $\hat{\alpha}_{ij}$. The optimum value of $\hat{\alpha}_{ij}$ for each image patch is obtained by solving

$$\hat{\alpha}_{ij} = arg \min_{\alpha} \mu_{ij} \parallel \alpha \parallel_0 + \parallel D\alpha - x_{ij} \parallel_2^2 \quad (6.10)$$

OMP algorithm is employed to obtain $\hat{\alpha}_{ij}$ for each image patch by picking one column at a time from a dictionary and stopping it when $\parallel D\alpha - x_{ij} \parallel_2^2 < T_h$. This is operated for every image patch of size $\sqrt{m} \times \sqrt{m}$ one at a time on sliding window model basis. Eqn. 6.9 is in quadratic form and its closed form solution is given by

$$\bar{X} = (\lambda I + \sum_{ij} B_{ij}^T B_{ij})^{-1} (\lambda Y + \sum_{ij} B_{ij}^T D\hat{\alpha}_{ij}) \quad (6.11)$$

Eqn. 6.11 is seen as averaging the result of sparseland model of shifted overlapping patches of a speckle image. We formulated all the above equations assuming the dictionary matrix D is given to us. Various dictionaries are proposed in the literature and we choose to have discrete cosine transform as dictionary due its uncorrelated basis structure which tends to have high sparsity. We have to update the parameters D and α_{ij} based on the image patches iteratively unless the required condition is met. D and α_{ij} can be updated using patches from a set of clean images or from the corrupted image itself. D is the common dictionary used to represent all the patches in the image. In this thesis, we updated D and α_{ij} using the patches from

the speckled image. The final generalized problem is formulated as

$$\bar{X} = arg \min_{\{\alpha_{ij}\}_{ij}, X, D} \lambda \parallel \underline{X} - \underline{Y} \parallel_2^2 + \sum_{ij} \parallel D\underline{\alpha_{ij}} - B_{ij}X \parallel_2^2 \quad s.t \quad \parallel \alpha_{ij} \parallel_0 < Q \quad (6.12)$$

Algorithm 1: Algorithm for finding Sparse representation of an image.

1. In Eqn. 6.12, we have to fix three terms X, D and α_{ij}. To find this, we are going to fix two terms and find the third term. Let us fix $\underline{X}=\underline{Y}$ and overcomplete dictionary D with DCT.

2. Repeat J times

   - **Sparse matrix update** using OMP algorithm, Compute sparse vector α_{ij} for each patch $B_{ij}X$

   $$\forall_{ij} \min_{\underline{\alpha}_{ij}} \parallel \alpha \parallel_0 \, s.t \sum_{ij} \parallel B_{ij}X - D\underline{\alpha} \parallel_2^2 \leq \epsilon$$

   - **Dictionary update stage**: update each atom a=1,2,...k in D by
   -Find the set of blocks in image that uses this atom $v_a = \{(i,j)|\alpha_{ij}(a) \neq 0\}$
   -Find the corresponding error for each index $(i,j) \in w_a$

   $$e_{ij}^a = B_{ij}X_{ij} - \sum_{n \neq a} d_n \alpha_{ij}(n) \quad (6.13)$$

   -Matrix E is formed with columns $\{e_{ij}^a\}_{(ij) \in w_a}$
   -E_a is factorized as $U\nabla V^T$ using SVD algorithm. Fix the updated dictionary column d_l be the first column of U. Coefficients of $\{\alpha_{ij}^a\}_{(ij) \in w_a}$ is updated by the entries of V multiplied by $\nabla(1,1)$. Iteration of step 2 is called K-SVD.

3. Now D and α_{ij} known, compute $\underline{X}$ by

   $$\bar{X} = (\lambda I + \sum_{ij} B_{ij}^T B_{ij})^{-1}(\lambda Y + \sum_{ij} B_{ij}^T D\hat{\alpha}_{ij}) \quad (6.14)$$

   which is a simple averaging of shifted patches.

(i,j) in Eqn. 6.14 corresponds to spatial location of image patch in the image. Representing the image with few number of columns naturally reduces the noise and averaging the patches leads to smoothness of the image. The algorithm for finding sparseland model for speckle suppression is shown in Algorithm.1.

The sparseland model algorithm for speckle suppression requires to initialize few parameters, they are listed below.

1. **Initial dictionary** : overcomplete DCT.

2. **Overlapping stepsize** Δ: Interval between neighboring blocks.

3. **Sigma**: used to determine the target error for sparse-coding each block.

4. **Training block T**: Number of blocks extracted for training.

5. **Dict D** :Dictionary size.

6. **Block size** $\sqrt{m} \times \sqrt{m}$: size of image patch to process.

7. **Iteration number I**: number of K-SVD training iterations to perform.

8. λ: Specifies the relative weight attributed to the noisy input signal in determining the output.

Table 6.1: K-wave parameters used to simulate the ultrasound image is shown in Fig. 6.3(a)

Transducer width	14.1593 mm (64 grid points)
Number of elements	64
Number of active elements	64
Element width	221.2389 um (1 grid points)
Sound speed	1540 m/s
Focus distance	30 mm
Elevation focus distance	30 mm

The order of complexity for each pixel in sparseland model is of the order $O(mkQJ)$, where m is patch size, k is number of columns in dictionary, Q is the sparsity of each column in coefficient matrix and J is number of stages used for updating the dictionary.

6.3 Results

The performance of the algorithm is evaluated based on four types of ultrasound images, namely simulated coherent and fully formed speckle, real liver and kidney ultrasound images. The experiment is conducted by varying the parameters listed in

Table 6.2: Performance of the propose speckle suppression algorithm in suppressing the speckle noise. Reg. 1 and Reg. 2 corresponds to the homogeneous region correspond to Fig. 3.2

		Mean	Std.	SNR
Noisy	Reg. 1	31.23	12.21	2.5
	Reg. 2	121.95	19.9	6.12
Frost	Reg. 1	31.10	5.89	5.3
	Reg. 2	121.29	10.29	11.78
Lee	Reg. 1	30.10	5.18	5.8
	Reg. 2	122.0	10.77	11.32
SRAD	Reg. 1	39.42	4.4	8.95
	Reg. 2	167	10.18	16.40
Proposed	Reg. 1	27.68	2.7	10.25
	Reg. 2	121.63	6.1	19.93
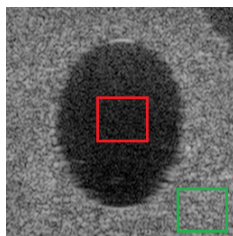
Figure 6.1: Red Box represents Region 1 (Reg. 1) and Green Box represents Region 2 (Reg. 2)

section II. The optimal suppression based on smoothing of homogeneous regions is obtained for the values $\sigma=4$, $I=20$, $\lambda=6$ by fixing parameters $T=40000$, dictionary size D-8 x 256, Block size- 8 x 8. The window size for computing local statistics of Lee and Frost is fixed at 5 x 5. The performance of the SRAD filter is fixed for 30 iterations. The performance of the filter in suppressing the speckle noise is analyzed based on conventional speckle signal to noise ratio (SNR) [101], which is computed as ratio of mean of the pixels to the standard deviation of pixels present in homogeneous region of an image. The Table. 6.2 shows the SNR computed for two homogeneous regions for the Fig. 6.1, we can infer that the propose speckle suppression filter has significantly enhanced the SNR compared to other filters. The Frost, Lee, SRAD filters able to preserve the mean in de-speckled image, while the propose filter able to reduce the standard deviation of the pixels resulting in high SNR compared to other filters.

The performance of Frost, Lee, SRAD and proposed method in suppressing the
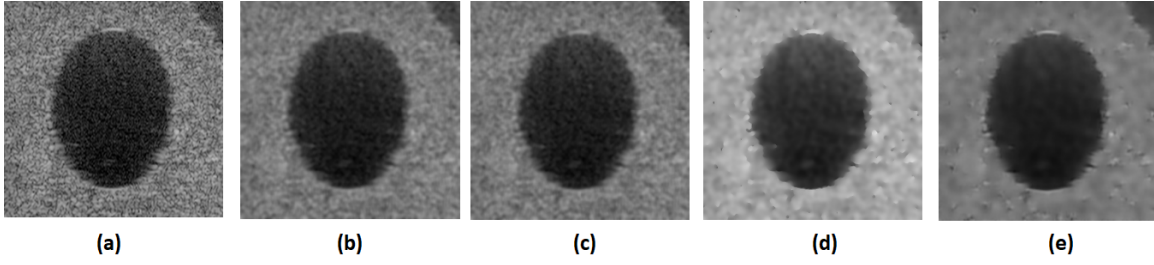
Figure 6.2: (a). Coherent Speckle image, De-speckling performance of (b) Frost, (c) Lee, (d) SRAD, (e) Proposed method.
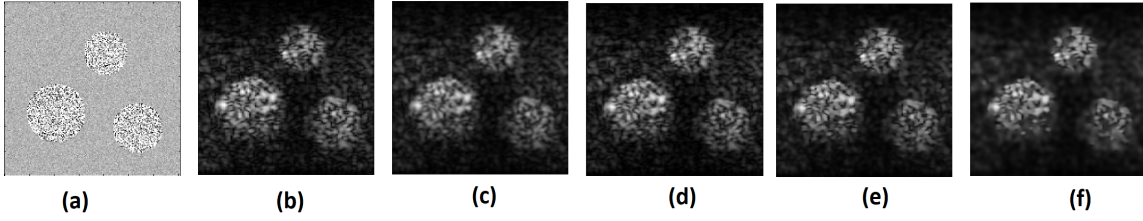
Figure 6.3: (a) Scattering phantom (b) Simulated ultrasound image of Scattering phantom. De-speckling performance of (c) Frost, (d) Lee, (e) SRAD, and (f) Proposed method.
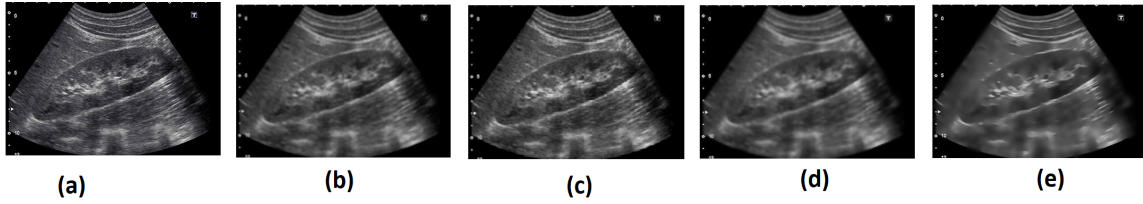
Figure 6.4: (a) Kidney ultrasound image. De-speckling performance of (b) Frost, (c) Lee, (d) SRAD, (e) Proposed method.
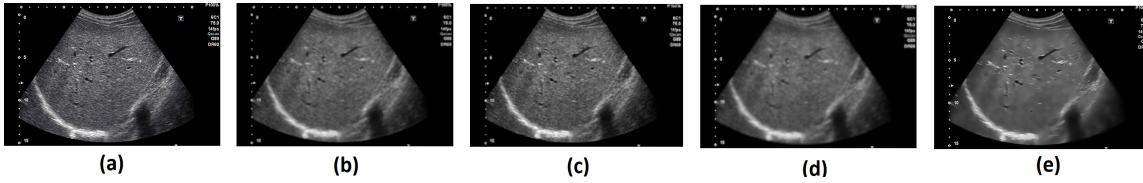
Figure 6.5: (a) Liver ultrasound image. De-speckling performance of (b) Frost, (c) Lee, (d) SRAD, (e) Proposed method.

speckle noise is shown in Fig. 6.2, we can infer from the despeckled images that the granular noise has been effectively suppressed from the proposed de-speckling filter. To know the performance of the filter on fully formed speckle, an ultrasound phantom image is generated in Matlab using K-Wave tool box. The K-wave parameter specifications used for generating an ultrasound image for the template Fig. 6.3(a) are shown in Table. 6.1. The small snake like structures in the image in Fig. 6.3(b) corresponds to speckle patterns.We can infer that the proposed speckle suppression filter effectively suppressed the small worm like structures in the despeckled image compared to other filters. Fig. 6.4 and Fig. 6.5 shows the performance of filters on kidney and liver images respectively. From visual analysis, we can infer that the diagnostic quality of ultrasound image is enhanced using sparseland model comparatively to Frost, Lee and SRAD filters. The algorithm is implemented in MATLAB 7.9 version on core i5 processor with 2.8 GHz clock speed. Sparseland model took 10.2 seconds for execution for an image of size 256×256.

6.4   Summary

Presence of speckle noise in ultrasound images is one of the reason for having low diagnostic accuracy for ultrasound scanning. To enhance the diagnostic quality in ultrasound images, we proposed an speckle suppression filter based on sparse representation of the image patch. The proposed algorithm proved very effective in suppressing the speckle noise thus enhancing the diagnostic quality in ultrasound images. With the proposed de-speckling filter, sonographers can infer more information from the despeckled ultrasound images which will help them to do accurate diagnosis.

Chapter 7

WebRTC Framework for Telesonography

7.1 Introduction

Ultrasound scanners does not translated to its potential in remote diagnostic health-care due to shortage of sonographers. Telesonograhy addresses the shortage of sonographers by allowing semi-skilled persons to scan the patients in remote areas and send it to the sonographer sitting elsewhere for getting the diagnosis. In conventional telesonography applications, the ultrasound video is encoded and transmitted to the expert end via wired or wireless communication technologies [11, 12], these kind of techniques are high bandwidth demanding and does not adapt to the network conditions resulting in under utilization of the network resources and poor quality of service. In telesonography, an inherent assumption is made that the scanned and transmitted ultrasound video sent to the expert or cloud is always representative enough for doing diagnosis which may not always hold good due to semi-skilled nature of the expertise. To have a reliable real-time telesonography, the semi-skilled person must able to connect distantly located expertise and transmit the scanned data in real-time for getting diagnosis.
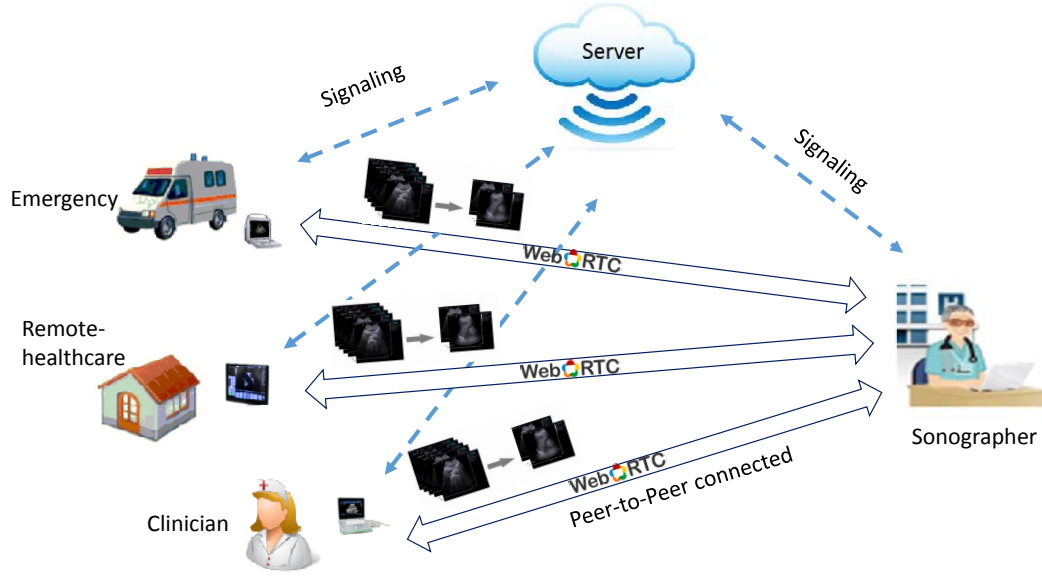
Figure 7.1: Architecture of WebRTC based telesonography.

Internet-based videophone services become more popular due to the widespread availability of the Internet connectivity, computing power, information and communication technologies. Video chat applications like Skype, Facebook, Google Hangouts, Facetime, WebRTC, etc., are some of the most popularly used video streaming applications. These video chat applications come with inbuilt rate control mechanisms that adapt to variable bandwidth, packet loss, delay, jitter, and packet error providing real-time better quality of service.

In this chapter, we propose a WebRTC (Web Real-Time Communication) based framework for real-time telesonography. WebRTC is a server-based service, where two end users can access the service to share multimedia data in a peer-to-peer fashion. The end users need to have only HTML5 supported browsers in their computing platforms for sharing ultrasound data making telesonography more reliable. The graphical representation of WebRTC based telesonography is shown in Fig. 7.1.

The quality of ultrasound video delivered to the end-user is analyzed by varying network settings at the transmitter side in a controlled manner. Liver ultrasound video is considered for the analysis. The performance of WebRTC is evaluated based

on the subjective quality analysis [102–104]. Subjective analysis is quantified by computing mean opinion scores (MOS) of the subjects for various network settings. We report for minimum network conditions needed for acceptable video quality for telediagnosis in particular to liver ultrasound video.

7.2   WebRTC

WebRTC [105] is a revolutionary technology providing browsers real-time communication capabilities with a set of application programming interfaces (APIs) and communication protocols. WebRTC is an open source project supported by Google, Mozilla, and Opera. The standardization of WebRTC is jointly developed by world wide web consortium (W3C) and the Internet engineering task force (IETF). The set of real-time communication protocols is standardized by IETF, while APIs are standardized by W3C.

Before WebRTC, to enable voice and video services within the browser, the user has to install that particular service real-time communication stack as a plugin (for example Skype, Google hangouts, etc.). The multimedia communication stack is built into the web browser internals with WebRTC. Hence the developers can make use of HTML5 APIs for developing multimedia services. WebRTC uses real-time congestion control algorithm proposed by Google [106]. WebRTC did not standardize the signaling protocol, and the users can choose any pre-existing signaling protocol like session initiation protocol, jingle, etc. WebRTC eliminates the need for installing third-party plugins like flash players, customized multimedia stacks for playing multimedia content making it widely adaptable in developing videophone services [107].

WebRTC provides a peer-to-peer communication by extending client-server semantics. In general, WebRTC uses Trapezoidal or Triangle architectural web server models. In WebRTC Trapezoid model, both browsers are running a web application downloaded from a different server, while in Triangle model both browsers are
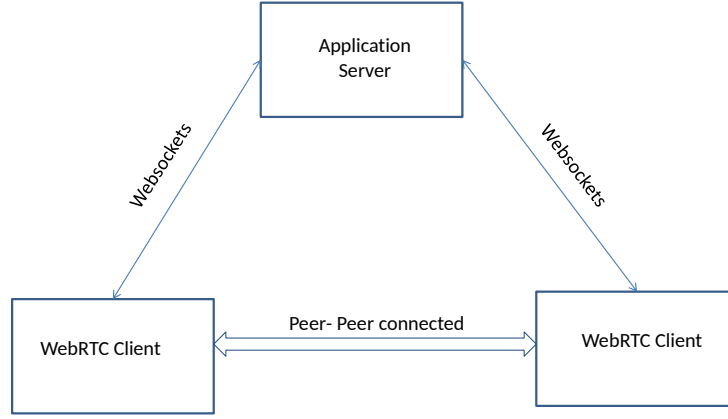
Figure 7.2: WebRTC network architecture.

running the same web application downloaded from the same webpage. Here, we employed the Triangle model web architecture as shown in Fig. 7.2. The web application is implemented on a dedicated server using nodejs [108]. To traverse firewalls and NATs, we run multiple TURN servers on dedicated machines. Websockets is used for signaling between browsers and server. WebRTC supported browsers Google Chrome and Firefox are used to establish a connection with a web server and to download the WebRTC javascript to configure the browser internals during call establishment. If the end points are behind NATs and firewalls, the media data flows through TURN servers else it flows directly between endpoints. The server application is run on an Intel core i7 processor with 16 GB RAM running on a 2.8 GHz clock cycle.

7.3 Experimental Setup

To study the performance of the WebRTC framework with respect to various network settings, we built a custom-made test bed. Two clients running a specific videophone service are connected to a Network Address Translation (NAT) router. Each host is connected wirelessly to the Internet through Wi-Fi router and consisted a private IP address. NEWT [109], a software-based network emulator tool is used to control the network parameters like bandwidth, packet loss, delay and packet error at the transmitter side. The ultrasound video is injected into videophone service using a

<div align="center">(a)   (b)   (c)</div>
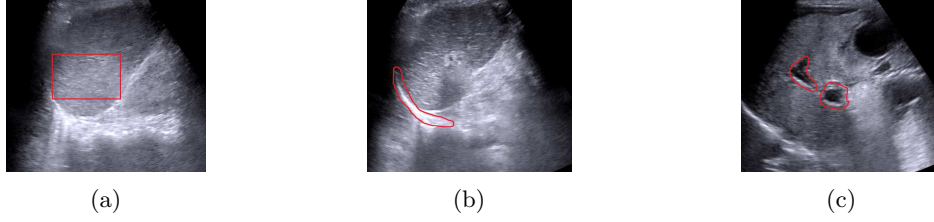
Figure 7.3: Region of interest (bounded with markers) seen by the sonographers for doing diagnosis. (a) Liver parenchyma accessed through texture and echogenecity, (b) Diaphragm visibility, (c) Artery and vein visibility.

virtual video camera tool Manycam [110]. The streaming videos are recorded using the CamStudio screen recording software [111].

7.3.1   Details of Ultrasound Video Considered for Study

In this study, the quality of delivered ultrasound video is analyzed for 169 different network settings with three subjects. Normal liver ultrasound video is considered for the subjective analysis. Liver diagnosis is characterized based on liver parenchyma along with visualization of the intrahepatic portal and vein vessel borders, and diaphragm. Even a slight diffusion in texture results in diagnosing the liver as higher grades [112]. The regions particularly seen in liver ultrasound image for diagnosis are, texture morphology, diaphragm, artery, and vein vessel visibility is shown in Fig. 7.3. The liver ultrasound video used in the study is captured using Siemens S1000 ultrasound scanner with a convex probe operating at 4 MHz frequency. The images are acquired to a depth of 15 cm with 15 frames per second for a duration of 10 s. The video is recorded in avi format with a 770×1024 resolution. The captured ultrasound video contains the images of normal liver showing parenchyma, diaphragm, hepatic and portal vein as shown in Fig. 7.4, these images are sampled at a frequency of one second for visualization.
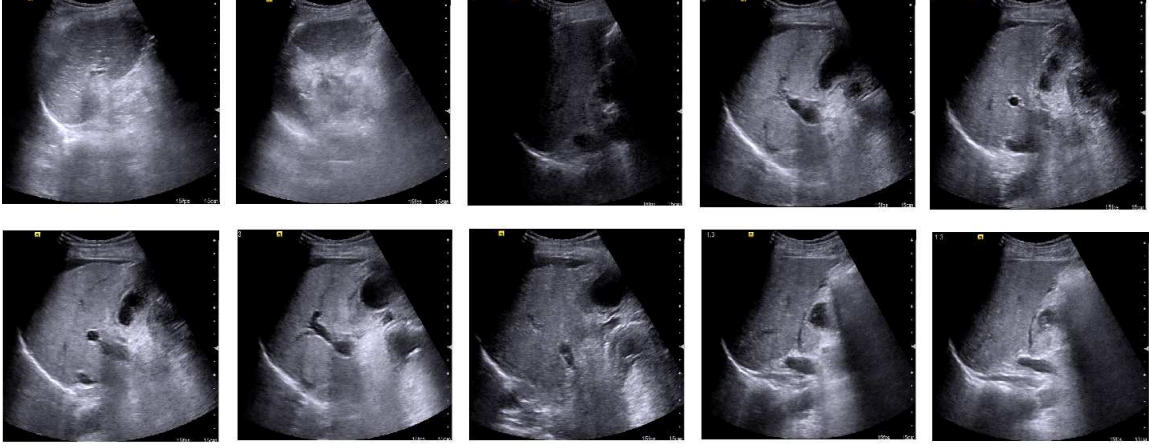
Figure 7.4: Images sampled from test ultrasound video at a sampling frequency of one second.

7.4 Subjective Quality Analysis

The subjective quality analysis is carried out with three experts who are trained in ultrasonography. They are presently working as sonographers in Hyderabad, India. For subjective analysis, we followed the procedure similar to Double Stimulus Continuous Quality Scale variant II methodology which is recommended in ITU-R BT.500-11 document [113, 114]. The streaming video is viewed on a laptop monitor with 1366 × 768 screen resolution. The laptop ran with Microsoft Windows 7 operating system with Intel Core i5 processor, 4 GB RAM, and 2.6 GHz clock cycle. Before taking the ratings, the subjects are made aware of the artifacts caused due to streaming applications. The subjects are asked to rate for a particular network configuration on a scale having categories namely *"Bad"*, *"Poor"*, *"Fair"*, *"Good"* and *"Excellent"*. These categories are translated into 1, 2, 3, 4 and 5 values respectively. The unimpaired video is considered to be of excellent quality. A rating of 5 is the highest value in the scale given to the ultrasound video, which is having the same clinical information to that of the unimpaired video. A rating of 4 is given to the video for which there is slight distortion in clinical information and still, the video is good enough for doing the diagnosis. A rating of 3 is given to the video for which there is a minor loss in the clinical information and still acceptable for doing the diagnosis. A rating below 3

is given for diagnostically unacceptable video. The ratings are given on a continuity scale such that the subjects do not undergo uncertainty in rating the video between two categories (integers). The results of a particular video instance are computed by presenting MOS. The MOS $\bar{\mu}_j$ for a particular network configuration is given by

$$\bar{\mu}_j = \frac{1}{N} \sum_i^N \mu_{ij},$$
(7.1)

where μ_{ij}: score of observer i for test configuration j, N: number of observers. Likewise the scores are computed for each network configuration.

7.5 Results

The subjective analysis is conducted with respect to variation in network parameters like delay, bandwidth, packet error at the transmitter side. The quality of video delivered at the receiver side is analyzed by setting Bandwidth (BW) at 2000, 1800, 1600, 1400, 1200, 1000, 800, 600, 500, 400, 300, 200, 100 kbps, Delays (DL) at 0, 50, 100, 200, 500, 1000 ms and Packet Error (PE) at 1%, 2%, 3%, 4%, 5%, 6%, 7% . For each setting in DL, and PE, the quality of ultrasound video is analyzed by varying the BW from 2 Mbps to 100 kbps. This results in an evaluation of 169 videos per videophone service. Since every user has to give a rating for 169 videos, the subjective analysis is carried out in multiple sessions. In each session, subjective evaluation is carried out for 25 to 30 instances based on their individual comfort. No repetition is done due to the evaluation of a large number of instances. The rejection of subjects is not considered since it consisted of only three subjects.

Fig. 7.5 present MOS for WebRTC application in delivering reliable ultrasound video for diagnosis, the MOS are plotted for bandwidth vs delays and bandwidth vs packet error. WebRTC delivered reliable ultrasound video for BW greater than 1400 kbps tolerating up to 500 ms delay and for BW greater than 1600 kbps tolerating up to 4% PE. The ultrasound video delivered through WebRTC framework was deteriorated
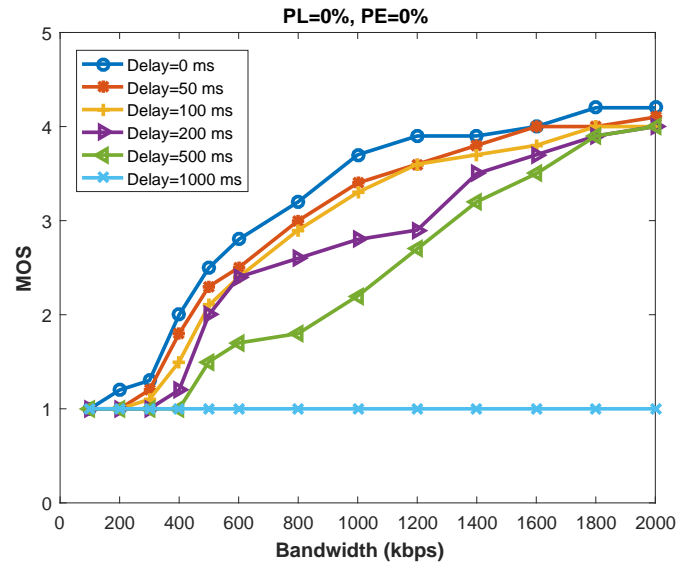
(a)
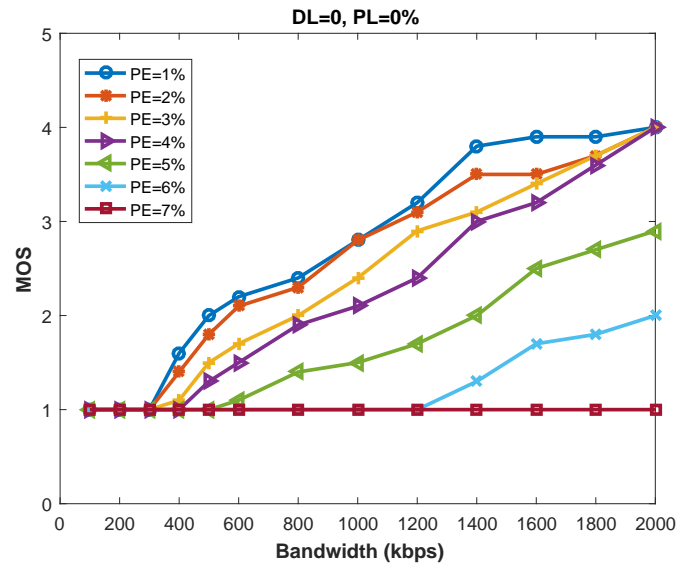
(b)

Figure 7.5: MOS for WebRTC for different network configurations

(a)	(b)	(c)

Figure 7.6: Frames captured at (a) DL:200 ms, BW: 800 kbps, PE=0 %. (b) DL:0 ms, BW: 2000 kbps, PE=5 % (c) DL:0 ms, BW: 800 kbps, PE=2 %. We can observe the image regions get smoothened due to the streaming artifacts loosing of texture characteristics in the image.

with spatial and temporal artifact's. Spatial artifact's like blocking and ringing are observed, and it is affected nonuniformly in the streamed ultrasound video. Temporal artifact's like frame freezing and frame dropping are observed in affecting the quality of experience of the user. Spatial artifacts deteriorated the video quality by morphing the edge regions and smoothing the texture making the image to loose diagnostic quality. Frame dropping and freezing had less impact on diagnosis especially in scenarios like liver ultrasound videos where fewer frame rates are enough for diagnosis. In our subjective study, the subjects are found to be more sensitive to spatial artifacts than temporal artifact's. For reference, the quality of liver images sampled at different network configurations of a streaming video is shown in Fig. 7.6

7.6 Summary

In this chapter, we propose a WebRTC based video streaming application in delivering the ultrasound video for real-time telesonography applications. We analyzed the performance of WebRTC under different network conditions in delivering the live ultrasound video. An inference is made that, we can deliver a reliable ultrasound video if the network bandwidth has BW greater than 1400 kbps tolerating up to 500 ms delay (with PE=0%) and for BW greater than 1600 kbps tolerating up to 4% PE (with DL=

0 ms). Recent advancements in IoT and multi-media communications significantly improved the standards in healthcare which introduced a new paradigm in patient monitoring and diagnostic healthcare. Since smartphone, tablet and laptop-based ultrasound scanners are coming with Internet connectivity, our study in evaluating the performance of WebRTC service will give future directions in using these services. In this study, the performance of WebRTC framework is evaluated by varying bandwidth for different delays and packet error. The performance of the WebRTC framework is not evaluated for combinations where packet loss, delay and packet error occurs simultaneously which is taken as future extension of this work.

Chapter 8

SoC based Portable Ultrasound Scanning System for PoC Applications

8.1 Introduction

The services of ultrasound scanning are centralized to hospitals and did not translate to its potential in PoC diagnostics due to its high form factor. Recent advancements in computing platforms like field programmable gate arrays (FPGAs), graphical processing units (GPUs), digital signal processors (DSPs), application specific integrated circuit (ASIC), etc., reduced the size of the ultrasound scanning system to the portable level [115–117]. PoC diagnostics enables fast and instantaneous diagnostics and has the potentiality to save lives especially in critical care situations like accidents, causalities, military, etc. [5], [118]. The desired characteristics of the portable ultrasound scanner such as low form factor, low cost, low power consumption can be obtained with ASIC's [5], [116]. But, the systems developed on ASIC is specifically meant for targeted consumer applications, hence in a continuously evolving field like ultrasound, further up-gradation of the system is not possible which acts as a serious limitation

on the system. Hence researchers are focused on programmable computing platforms for developing portable ultrasound scanning (PUS) systems [119]. Developing the portable ultrasound scanner (PUS) on programmable computing platforms allows for easy up-gradation and validation of the new algorithms. Motivated by this reason, researchers have used DSP processors for implementing the ultrasound signal processing algorithms [117, 119]. Implementing the entire ultrasound signal processing on a DSP processor poses a serious challenge in executing computationally intensive beamforming algorithms due to limited data rate capability. Due to high programming flexibility and support of high data rates, the FPGA's are highly reliable to perform complete ultrasound signal processing algorithms [117], to control and coordinate between different processing modules on FPGA, a separate ARM processor is used. This makes the entire system cumbersome and also results in high form factor due to the need for an additional general purpose processor. Some of the primary issues found in conventional and portable ultrasound scanning systems are:

1. Since the ultrasound scanners are meant for imaging, there is no provision for accessing the raw RF data from the scanner, which will be very much helpful in easy validation of the new algorithms and also in developing the computer-aided algorithms. At present, most of the computer-aided diagnosis (CAD) is based on the image based analytics [19, 24]. But it is proved that robust CAD algorithms can be developed by directly working on the RF echo data [120, 121], presently, there is a no provision in acquiring the RF data from the clinical ultrasound scanners. To acquire the raw RF data from the clinical ultrasound scanners, we need to use highly expensive interfaces [17], [122].

2. In programmable ultrasound scanners, not much focus has been laid in providing the debug interfaces, which will be needed while updating the PUS with new algorithms [115, 116].

In addressing these issues, we propose a programmable PUS that can not only

be used for PoC diagnosis, but also for acquiring the clinical RF data that can be used for research activities. The main contributions and implications of the propose programmable PUS are:

- We propose a programmable PUS using a single SoC, thus achieving miniaturization, and provided with adequate external interfaces for more reliable diagnosis.

- The proposed PUS is provided with JTAG and UART ports for efficient debugging of FPGA and the ARM processor respectively.

- Provision is provided to acquire the raw RF echo data, which can be used for developing highly accurate computer-aided algorithms, thus eliminating the need for having expensive research interfaces.

In this chapter, we present a programmable Zynq SoC (Xilinx Inc., San Jose, CA) based hardware architecture for implementing complete PUS system customized for PoC diagnostics. The Zynq SoC XC7Z045FBG676 is a hybrid processor that comes with both FPGA and ARM Cortex A9 processor. The propose system is developed for 16 channels and has the provision to extend for 32 channels. To make the ultrasound scanning more reliable to use in terms of connectivity, the system is provided with external interfaces like Ethernet, memory card slot, HDMI port, VGA port, two USB OTG ports and RS232. A Wi-Fi module can be interfaced with propose PUS to transfer the image data wirelessly to the mobile computing platforms.

8.2 A Single SoC based Portable Ultrasound Scanning System

The overall block diagram representation of the propose PUS system is shown in Fig. 8.1 and the corresponding hardware implementation at circuit level is shown in Fig. 8.2. Based on the functionality, the ultrasound signal processing is categorized into front-end, mid-end and back-end processing modules [3]. Dedicated hardware is
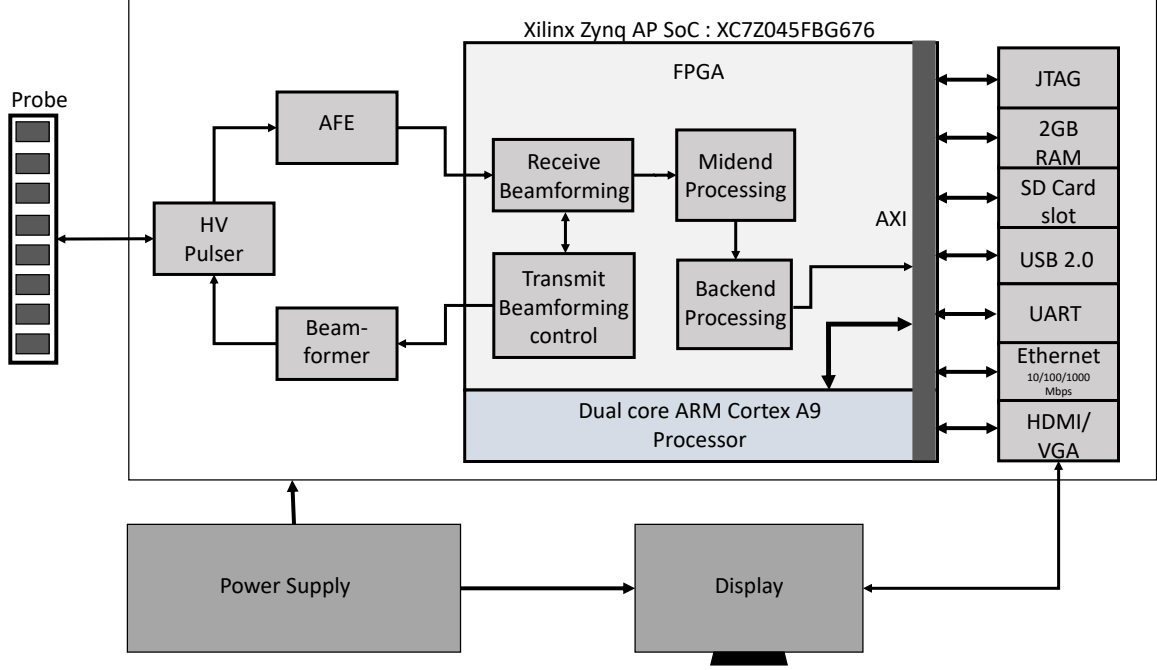
Figure 8.1: Block diagram representation of the proposed PUS representing the different signal processing modules present in the ultrasound scanner. AXI represent advanced extensible interface.

developed for front-end processing, while the mid-end and back-end processing algorithms are implemented on the SoC. The front-end processing deals with transducer excitation and data acquisition module. Dedicated hardware is developed to meet the high-speed real-time data processing requirements. Mid-end and back-end processing modules deal with signal processing algorithms required for image formation from the RF data, and it is implemented on the FPGA.

The board is provided with VGA and HDMI ports that are directly connected to the FPGA such that the monitors can be connected to display the scanned images. The onboard USB 2.0 can be used for interfacing with other computing platforms like smart phones, tablets, etc. for performing other ultrasound signal processing algorithms, and also can be used as a display device. All other external ports are interfaced with an on-chip ARM processor. The board is provided with a 2 GB RAM and a secure digital (SD) card slot for storing the scanned data and for running a Linux operating system for ARM processor at 667 MHz clock frequency. The onboard
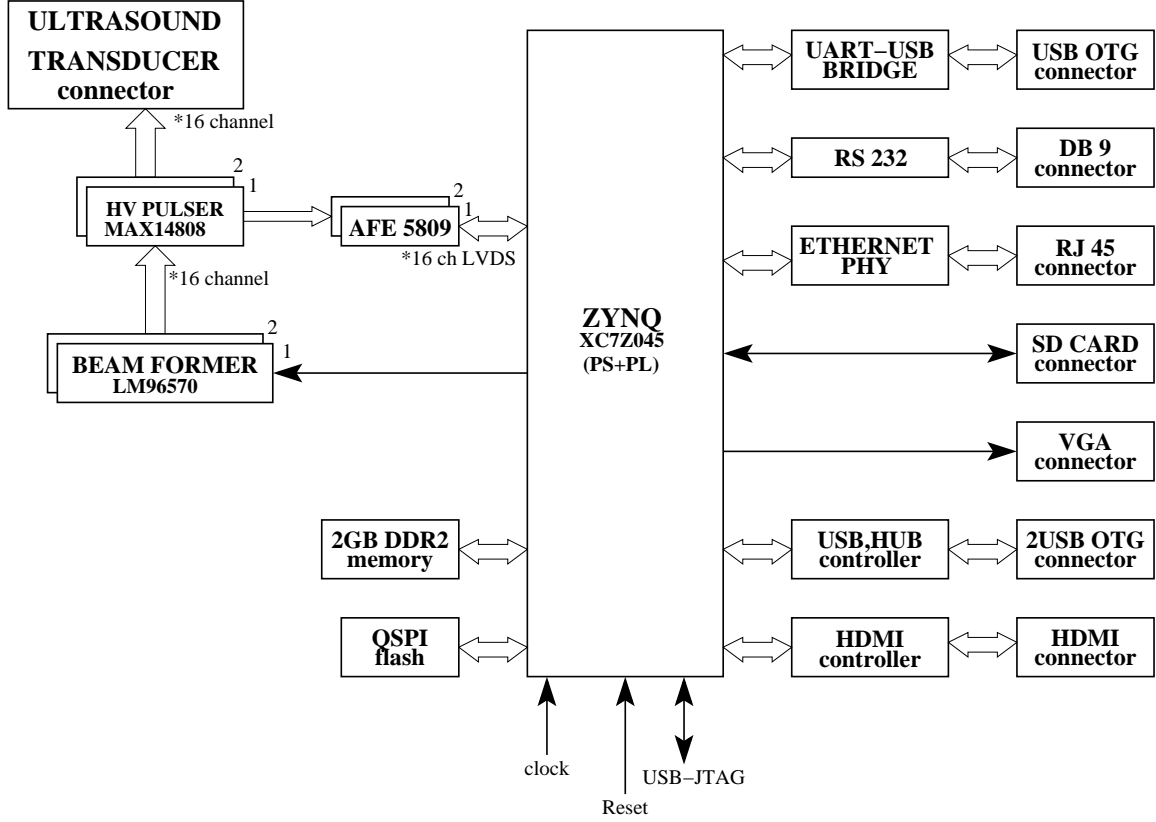
Figure 8.2: Circuit level implementation of the proposed PUS. Here SoC comes with a combination of processing system (PS) and programming logic (PL).

Ethernet port can be used to transfer the data with 10/100/1000 Mbps speed, which can be used to transfer or upload the scanned data to the server or cloud for central storage.

The front-end processing module consists of hardware circuitry needed for exciting the transducer elements, transmit beamforming and signal conditioning algorithms like low noise amplification and time gain compensation. Since these algorithms needed to be performed in real-time, a dedicated hardware module is employed to execute these algorithms.

A voltage of ± 80 V is needed for exciting the transducer elements, this voltage is generated by high voltage (HV) pulser MAX14808 IC (Maxim Integrated, SanJose, California, United States), each HV pulser can excite 8 channels at a time and hence two HV pulsers are employed to excite the 16 channels. For performing transmit

beamforming, LM 96570 IC (Texas Instruments, Dallas, USA) has been used, each IC can be programmed to generate the delay pattern for 8 channels. The RF data is acquired using AFE 5809 (Texas Instruments, Dallas, USA) via HV pulser. The AFE consist of high-speed analog to digital converter (ADC) capable of operating at a max of 65 MHz sampling frequency with 12 bit resolution, similar to beamformer and HV pulser, AFE is also capable of operating for max of 8 channels, and hence two IC's are employed to receive the signals from the 16 channels. The output of the AFE is transferred to the FPGA for RF data processing, which includes envelope detection, log compression and scan conversion algorithms.

8.3   Hardware Prototype

The prototype of the proposed PUS is shown in Fig. 8.3. The transducer consisted of 64-elements with 96 pin header. The PCB board is provided with a pin header to hold the transducer probe firmly while scanning. The same pin header can be used to connect both linear and curvilinear probe. The specifications of the developed PUS is shown in Table. 8.1. The monitor can be connected to the board via VGA or through HDMI port. The PUS is provided with all the necessary resources needed for complete clinical ultrasound scanning system. To obtain compact size for the PUS, the developed hardware circuitry for the front-end processing, power management circuitry, external interfaces and Zynq SoC are fabricated on an eight-layer printed circuit board.

8.4   Summary

In this chapter, we presented the programmable hardware framework for portable ultrasound scanning system based on a single SoC. The propose PUS is provided with external interfaces to connect other modules like global positioning system (GPS) for
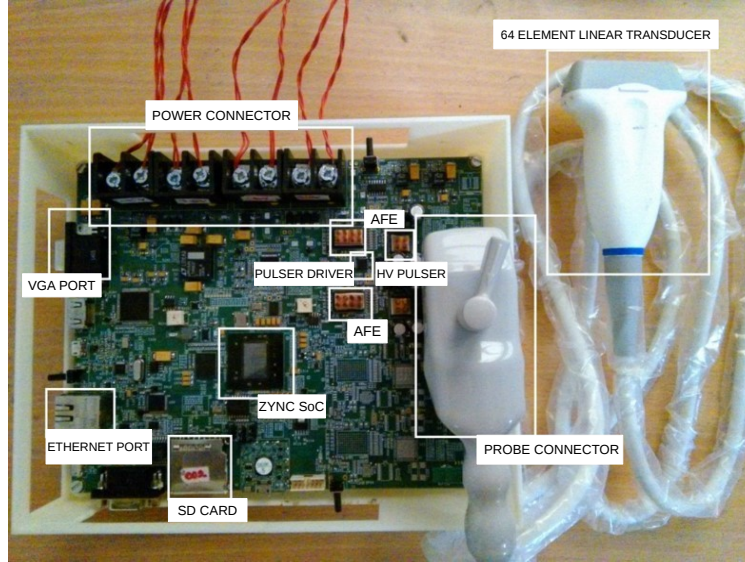
Figure 8.3: A Single SoC based programmable portable ultrasound scanner.

Table 8.1: Specifications of the propose programmable PUS.

Mode	Specifications
Size	235 mm x 205 mm
Weight	0.45 kg
Probe	5 MHz, 64 elements, linear and curvilinear array probe
Display	15 inch LCD, Resolution - 1366 x 768, Input : VGA
CPU	Dual core ARM cortex A9 processor
System memory	2 GB DDR2 (SDRAM), 256 MB (flash)
FPGA	XC7Z045FBG676
Processor	ARM cortex A9 processor
Peripherals	USB 2.0 host and client, UART, Ethernet, HDMI and VGA
Scanning mode	B-Mode

tracking, biometric authentication for security, Wi-Fi models for IoT compatibility, etc. which will be beneficial while operating in remote healthcare. The developed hardware framework can be used for testing of novel portable ultrasound signal processing algorithms and also provides the provision to access the raw RF data, which will be useful for developing computer-aided diagnostic algorithms.

Note: The hardware development of portable ultrasound scanner is result of collaborative work carried with M.Tech research assistants namely Punit Kumar, Chandra Sekhar, Vivek Akkala, Arun Kumar and project staff namely Suresh Puli and Harsha

Ponduri. It took a time duration of three years to develop the working prototype of the proposed portable ultrasound scanner.

Chapter 9

Summary and Future Scope

In this chapter, we summarize the work presented so far. Subsequently, we discuss broader impact of the contributions of this dissertation in ultrasound imaging, and the future scope of this work.

9.1 Summary

Despite being safer, economical and having real-time imaging capability, the ultrasound scanners have not been leveraged to its fullest extent in health-care. There are many reasons, among which shortage of sonographers, low diagnostic accuracy leave room for improvement. Against this backdrop, we attempted to provide engineering perspectives in ultrasound diagnostic imaging that involves image-based analysis and diagnostics, reliable telesonography development of portable ultrasound scanning system. To this end, acknowledging that ultrasound diagnosis being a vast discipline, highly subjective and comprehensive coverage is infeasible. We targeted on problems which pose a significant challenge in both clinically and also from engineering perspective. In particular, we focused on developing algorithms to assist semi-skilled persons to scan the patients in remote areas, these include (i) automated diagnosis (ii) development of portable ultrasound scanning system, and (iii) IoT enabled telesonography. In this endeavor, we come to know that, finding universal solutions

for whole gamuts of ultrasound imaging is unrealistic. Instead, we find that understanding specific problems and developing problem-specific solutions were found to be robust and clinically deployable. Accordingly, our perspective on ultrasound imaging is categorized into image analysis, telesonography and development of portable ultrasound scanning system.

The ultrasound datasets for developing automated diagnostics is not publicly available for research activities. Due to subjectivity involved in ultrasound scanning, developing the organized ultrasound data sets is highly challenging. In this regard, we developed ultrasound datasets for the nonalcoholic fatty liver disease. Further, by using these datasets, we proposed an algorithm for quantifying the fatty content in the liver. In addition, we also proposed an algorithm for detecting the diagnostically significant regions corresponding to fatty liver. Ultrasound images are affected by speckle noise which reduces the diagnostic quality, in this regard, we proposed speckle suppression algorithm for enhancing the image quality. To address the shortage of sonographers, we proposed a WebRTC based IoT enabled telesonography and conducted subjective video quality analysis to evaluate the performance of the proposed system. We developed a programmable portable ultrasound scanning system based on single SoC for research and point-of-care diagnostics.

9.2 Future Scope

As future scope of this thesis, we are focused on achieving following objectives:

9.2.1 Image Analytics

1. We will develop an automated algorithm based on multiple RoIs detected from chapter 5. We will investigate appropriate features to be used for representing each RoI, weights given for each RoI in classifying the fatty liver. Since the characteristics of fatty liver changes accordingly with the severity of fatty liver, we

will explore ordinal classification framework for automated grading of fatty liver. We also explore Grey theory to capture the uncertainty observed in ultrasound images.

2. As WebRTC framework proved reliable for real-time telesonography, we will develop and integrate automated organ validation algorithms to WebRTC to validate whether the scanned data has representative information useful for diagnosis before transmitting to the expert. This will reduce the effort and time spent by the sonographer to diagnose the patients, and also guides the semi-skilled person to scan appropriately to acquire the representative information from the patients.

Publications

Journals

1. R. Bharath, P. Rajalakshmi, MA Mateen. "Multi-modal framework for automatic detection of diagnostically important regions in nonalcoholic fatty liver ultrasonic images," *Elsevier Biocybernetics and Biomedical Engineering.* 2018 Apr 5.

2. R. Bharath R, PK Mishra, P. Rajalakshmi. "Automated quantification of ultrasonic fatty liver texture based on curvelet transform and SVD." *Elsevier Biocybernetics and Biomedical Engineering.* 2017 Dec 27.

3. Divya Krishna, V Akkala, R. Bharath, P Rajalakshmi, AM Mohammed, SN Merchant, UB Desai. "Computer Aided Abnormality Detection for Kidney on FPGA Based IoT Enabled Portable Ultrasound Imaging System." *Elsevier Innovation in Research and Biomedical Engineering*",37.4, pp:189-97, 2016.

4. R. Bharath., P. Kumar, C. Dusa, V Akkala, P Suresh, P. Harsha, D Krishna, P. Rajalakshmi, SN Merchant, MA Mateen, and U.B Desai, "FPGA-Based Portable Ultrasound Scanning System with Automatic Kidney Detection." *MDPI Journal of Imaging*, 1.1, pp.193-219, 2015.

5. R. Bharath, Punit Kumar, D. Santhosh Reddy, P. Rajalakshmi "A Single SoC based Programmable Portable Ultrasound Scanning System Hardware Framework for PoC and Clinical Research Activities" *Elsevier Biocybernetics and Biomedical Engineering*, **under minor revision**.

Conferences

1. R. Bharath, Punit Kumar, Santhosh Reddy and P. Rajalakshmi, "Programmable Ultrasound Front-End Processing Module for Research Activities", 40th Annual International Conference of the IEEE EMBC, Honolulu, HI, 2018. *Accepted*

2. R. Bharath and P. Rajalakshmi, "Deep scattering convolution network based features for ultrasonic fatty liver tissue characterization," 39th Annual International Conference of the IEEE EMBC, Seogwipo, 2017, pp. 1982-1985. doi: 10.1109/EMBC.2017.8037239

3. P. Vaish, R. Bharath and P. Rajalakshmi, "Smartphone based automatic organ validation in ultrasound video," 39th Annual International Conference of the IEEE EMBC, Seogwipo, 2017, pp. 4289-4292. doi: 10.1109/EMBC.2017.8037804

4. R. Bharath, P. Rajalakshmi and U. B. Desai, "Subjective liver ultrasound video quality assessment of internet based videophone services for real-time telesonography," IEEE 19th International Conference on, Healthcom, Dalian, 2017, pp. 1-6. doi: 10.1109/HealthCom.2017.8210802

5. R. Bharath and P. Rajalakshmi, "Non-local means kernel regression based de-speckling of B-mode ultrasound images," IEEE 18th International Conference on, Healthcom, Munich, 2016, pp. 1-6. doi: 10.1109/HealthCom.2016.7749482

6. P. Vaish, R. Bharath, P. Rajalakshmi and U. B. Desai, "Smartphone based automatic abnormality detection of kidney in ultrasound images," IEEE 18th International Conference on, Healthcom, Munich, 2016, pp. 1-6. doi: 10.1109/HealthCom.2016.7749492

7. R. Bharath, P. Vaish and P. Rajalakshmi, "Implementation of diagnostically driven compression algorithms via WebRTC for IoT enabled tele-sonography,"

2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, 2016, pp. 204-209. doi: 10.1109/IECBES.2016.7843443

8. R. Bharath, D. S. Reddy, P. Kumar and P. Rajalakshmi, "Novel architecture for wireless transducer based ultrasound imaging system," IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, 2016, pp. 432-436, 2016. doi: 10.1109/IECBES.2016.7843487

9. P. K. Mishra, R. Bharath, P. Rajalakshmi and U. B. Desai, "Compressive sensing ultrasound beamformed imaging in time and frequency domain," 17th International Conference on, HealthCom, Boston, MA, 2015, pp. 523-527. doi: 10.1109/HealthCom.2015.7454558

10. R. Bharath, D Chandrashekar, A. Vivek, Divya Krishna, Harsha Ponduri, P Rajalakshmi, U B Desai "Portable ultrasound scanner for remote diagnosis," 2015 17th International Conference on, HealthCom, Boston, MA, 2015, pp. 211-216.doi: 10.1109/HealthCom.2015.7454500

11. M. Srinivas, R. Bharath, P. Rajalakshmi and C. Krishna Mohan, "Sparseland model for speckle suppression of B-mode ultrasound images," 2015 Twenty First National Conference on Communications (NCC), Mumbai, 2015, pp. 1-6. doi: 10.1109/NCC.2015.7084842

12. K. D. Krishna, V. Akkala, R. Bharath, P. Rajalakshmi and A. M. Mohammed, "FPGA based preliminary CAD for kidney on IoT enabled portable ultrasound imaging system," IEEE 16th International Conference on Healthcom, Natal, 2014, pp. 257-261. doi: 10.1109/HealthCom.2014.7001851

13. A. Vivek, R. Bharath, P. Rajalakshmi, and Punit Kumar. "Compression techniques for IoT enabled handheld ultrasound imaging system." In Biomedical Engineering and Sciences (IECBES), 2014 IEEE Conference on, pp. 648-652. IEEE, 2014.

14. R. Bharath and Rajalakshmi, "Fast Region of Interest detection for fetal genital organs in B-mode ultrasound images," 5th ISSNIP-IEEE Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC), Salvador, 2014, pp. 1-5. doi: 10.1109/BRC.2014.6880981

References

[1] Y.-Y. Liao, K.-C. Yang, M.-J. Lee, K.-C. Huang, J.-D. Chen, and C.-K. Yeh. Multifeature analysis of an ultrasound quantitative diagnostic index for classifying nonalcoholic fatty liver disease. *Scientific reports* 6, (2016) 35,083.

[2] D. Singh, C. J. Das, and M. P. Baruah. Imaging of non alcoholic fatty liver disease: A road less travelled. *Indian journal of endocrinology and metabolism* 17, (2013) 990.

[3] P. R. Hoskins, K. Martin, and A. Thrush. Diagnostic ultrasound: physics and equipment. Cambridge University Press, 2010.

[4] T. L. Szabo. Diagnostic ultrasound imaging: inside out. Academic Press, 2004.

[5] S. P. Stawicki, J. M. Howard, J. P. Pryor, D. P. Bahner, M. L. Whitmill, and A. J. Dean. Portable ultrasonography in mass casualty incidents: The CAVEAT examination. *World journal of orthopedics* 1, (2010) 10.

[6] [Online]. Available: Statistics Division Ministry of Health and Family Welfare, Government of India 2017, https://data.gov.in/search/site?query=Radiographers .

[7] [Online]. Available: Sonographer Workforce Survey Analysis-Society of Radiographers, https://www.sor.org/printpdf/book/export/html/11716 .

[8] M. Craig. Sonographer Shortages: A Day Late and a Dollar Short? *Journal of Diagnostic Medical Sonography* 19, (2003) 261–271.

[9] S. Oak et al. Eagle's eye and lion's heart-medicolegal issues in ultrasonography practice. *Indian Journal of Radiology and Imaging* 12, (2002) 105.

[10] S. Strauss, E. Gavish, P. Gottlieb, and L. Katsnelson. Interobserver and intraobserver variability in the sonographic assessment of fatty liver. *American Journal of Roentgenology* 189, (2007) W320–W323.

[11] M. Razaak, M. G. Martini, and K. Savino. A study on quality assessment for medical ultrasound video compressed via HEVC. *IEEE Journal of biomedical and health informatics* 18, (2014) 1552–1559.

[12] A. Panayides, M. S. Pattichis, C. S. Pattichis, C. P. Loizou, M. Pantziaris, and A. Pitsillides. Atherosclerotic plaque ultrasound video encoding, wireless transmission, and quality assessment using H. 264. *IEEE Transactions on Information Technology in Biomedicine* 15, (2011) 387–397.

[13] Y. Takahashi and T. Fukusato. Histopathology of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. *World journal of gastroenterology: WJG* 20, (2014) 15,539.

[14] S. Bellentani, F. Scaglioni, M. Marino, and G. Bedogni. Epidemiology of nonalcoholic fatty liver disease. *Digestive diseases* 28, (2010) 155–161.

[15] G. Marchesini, M. Brizi, A. M. Morselli-Labate, G. Bianchi, E. Bugianesi, A. J. McCullough, G. Forlani, and N. Melchionda. Association of nonalcoholic fatty liver disease with insulin resistance. *The American journal of medicine* 107, (1999) 450–455.

[16] E. M. Brunt, C. G. Janney, A. M. Di Bisceglie, B. A. Neuschwander-Tetri, and B. R. Bacon. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *The American journal of gastroenterology* 94, (1999) 2467.

[17] T. Wilson, J. Zagzebski, T. Varghese, Q. Chen, and M. Rao. The ultrasonix 500RP: A commercial ultrasound research interface. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 53.

[18] U. Raeth, D. Schlaps, B. Limberg, I. Zuna, A. Lorenz, G. Van Kaick, W. J. Lorenz, and B. Kommerell. Diagnostic accuracy of computerized B-scan texture analysis and conventional ultrasonography in diffuse parenchymal and malignant liver disease. *Journal of Clinical Ultrasound* 13, (1985) 87–99.

[19] P. Bharti, D. Mittal, and R. Ananthasivan. Computer-aided Characterization and Diagnosis of Diffuse Liver Diseases Based on Ultrasound Imaging: A Review. *Ultrasonic imaging* 39, (2017) 33–61.

[20] Y. M. Kadah, A. A. Farag, J. M. Zurada, A. M. Badawi, and A.-B. Youssef. Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images. *IEEE transactions on Medical Imaging* 15, (1996) 466–478.

[21] U. R. Acharya, O. Faust, F. Molinari, S. V. Sree, S. P. Junnarkar, and V. Sudarshan. Ultrasound-based tissue characterization and classification of fatty liver disease: A screening and diagnostic paradigm. *Knowledge-Based Systems* 75, (2015) 66–77.

[22] S. Pavlopoulos, E. Kyriacou, D. Koutsouris, K. Blekas, A. Stafylopatis, and P. Zoumpoulis. Fuzzy neural network-based texture analysis of ultrasonic images. *IEEE Engineering in Medicine and Biology Magazine* 19, (2000) 39–47.

[23] U. R. Acharya, H. Fujita, V. K. Sudarshan, M. R. K. Mookiah, J. E. Koh, J. H. Tan, Y. Hagiwara, C. K. Chua, S. P. Junnarkar, A. Vijayananthan et al. An integrated index for identification of fatty liver disease using radon transform and discrete cosine transform features in ultrasound images. *Information Fusion* 31, (2016) 43–53.

[24] E. Goceri, Z. K. Shah, R. Layman, X. Jiang, and M. N. Gurcan. Quantification of liver fat: a comprehensive review. *Computers in biology and medicine* 71, (2016) 174–189.

[25] L. De Luca, R. Westbrook, and E. A. Tsochatzis. Metabolic and cardiovascular complications in the liver transplant recipient. *Annals of Gastroenterology: Quarterly Publication of the Hellenic Society of Gastroenterology* 28, (2015) 183.

[26] V. Khullar, A. Dolganiuc, and R. J. Firpi. Pre-and-post transplant considerations in patients with nonalcoholic fatty liver disease. *World journal of transplantation* 4, (2014) 81.

[27] M. Lupşor, R. Badea, C. Vicaş, S. Nedevschi, M. Grigorescu, C. Radu, H. Stefanescu, and D. Crisan. Non-invasive steatosis assessment in NASH through the computerized processing of ultrasound images: Attenuation versus textural parameters. In Automation Quality and Testing Robotics (AQTR), 2010 IEEE International Conference on, volume 2. IEEE, 2010 1–6.

[28] S. İçer, A. Coşkun, and T. İkizceli. Quantitative grading using grey relational analysis on ultrasonographic images of a fatty liver. *Journal of medical systems* 36, (2012) 2521–2528.

[29] D. M. Mihailescu, V. Gui, C. I. Toma, A. Popescu, and I. Sporea. Automatic evaluation of steatosis by ultrasound image analysis. In Electronics and Telecommunications (ISETC), 2012 10th International Symposium on. IEEE, 2012 311–314.

[30] C. Vicas, S. Nedevschi, M. Lupsor, and R. Badea. Automatic detection of liver capsule using Gabor filters. Applications in steatosis quantification. In Intelligent Computer Communication and Processing, 2009. ICCP 2009. IEEE 5th International Conference on. IEEE, 2009 133–140.

[31] R. Hernaez, M. Lazo, S. Bonekamp, I. Kamel, F. L. Brancati, E. Guallar, and J. M. Clark. Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: A meta-analysis. *Hepatology* 54, (2011) 1082–1090.

[32] R. Allan, K. Thoirs, and M. Phillips. Accuracy of ultrasound to identify chronic liver disease. *World Journal of Gastroenterology: WJG* 16, (2010) 3510.

[33] K. Ogawa, M. Fukushima, K. Kubota, and N. Hisa. Computer-aided diagnostic system for diffuse liver diseases with ultrasonography by neural networks. *IEEE transactions on Nuclear Science* 45, (1998) 3069–3074.

[34] C.-M. Wu, Y.-C. Chen, and K.-S. Hsieh. Texture features for classification of ultrasonic liver images. *IEEE Transactions on medical imaging* 11, (1992) 141–152.

[35] J. Santos, J. S. Silva, A. A. Santos, and P. Belo-Soares. Detection of pathologic liver using ultrasound images. *Biomedical Signal Processing and Control* 14, (2014) 248–255.

[36] A. Andrade, J. S. Silva, J. Santos, and P. Belo-Soares. Classifier approaches for liver steatosis using ultrasound images. *Procedia Technology* 5, (2012) 763–770.

[37] A. Ahmadian, A. Mostafa, M. Abolhassani, and Y. Salimpour. A texture classification method for diffused liver diseases using Gabor wavelets. In Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the. IEEE, 2006 1567–1570.

[38] W.-L. Lee. An ensemble-based data fusion approach for characterizing ultrasonic liver tissue. *Applied Soft Computing* 13, (2013) 3683–3692.

[39] W.-L. Lee, Y.-C. Chen, and K.-S. Hsieh. Ultrasonic liver tissues classification by fractal feature vector based on M-band wavelet transform. *IEEE Transactions on Medical Imaging* 22, (2003) 382–392.

[40] C.-C. Wu, W.-L. Lee, Y.-C. Chen, and K.-S. Hsieh. Evolution-based hierarchical feature fusion for ultrasonic liver tissue characterization. *IEEE journal of biomedical and health informatics* 17, (2013) 967–976.

[41] A. Alivar, H. Danyali, and M. S. Helfroush. Hierarchical classification of normal, fatty and heterogeneous liver diseases from ultrasound images using serial and parallel feature fusion. *Biocybernetics and Biomedical Engineering* 36, (2016) 697–707.

[42] M. Singh, S. Singh, and S. Gupta. An information fusion based method for liver classification using texture analysis of ultrasound images. *Information Fusion* 19, (2014) 91–96.

[43] R. T. Ribeiro, R. T. Marinho, and J. M. Sanches. Classification and staging of chronic liver disease from multimodal data. *IEEE transactions on biomedical engineering* 60, (2013) 1336–1344.

[44] R. T. Ribeiro, R. T. Marinho, and J. M. Sanches. An ultrasound-based computer-aided diagnosis tool for steatosis detection. *IEEE journal of biomedical and health informatics* 18, (2014) 1397–1403.

[45] S. Selvan and S. Ramakrishnan. SVD-based modeling for image texture classification using wavelet transformation. *IEEE Transactions on Image Processing* 16, (2007) 2688–2696.

[46] E. Candès, D. Donoho, and A. Curvelets. surprisingly effective nonadaptive representation for objects with edges. Curve and Surfaces, L. Schumaker et al 1999.

[47] J.-L. Starck, E. J. Candès, and D. L. Donoho. The curvelet transform for image denoising. *IEEE Transactions on image processing* 11, (2002) 670–684.

[48] J.-L. Starck, D. L. Donoho, and E. J. Candès. Astronomical image representation by the curvelet transform. *Astronomy & Astrophysics* 398, (2003) 785–800.

[49] J.-L. Starck, F. Murtagh, E. J. Candès, and D. L. Donoho. Gray and color image contrast enhancement by the curvelet transform. *IEEE Transactions on image processing* 12, (2003) 706–717.

[50] D. R. Nayak, R. Dash, B. Majhi, and V. Prasad. Automated pathological brain detection system: A fast discrete curvelet transform and probabilistic neural network based approach. *Expert Systems with Applications* 88, (2017) 152–164.

[51] M. S. Miri and A. Mahloojifar. Retinal image analysis using curvelet transform and multistructure elements morphology by reconstruction. *IEEE Transactions on Biomedical Engineering* 58, (2011) 1183–1192.

[52] U. R. Acharya, U. Raghavendra, H. Fujita, Y. Hagiwara, J. E. Koh, T. J. Hong, V. K. Sudarshan, A. Vijayananthan, C. H. Yeong, A. Gudigar et al. Automated characterization of fatty liver disease and cirrhosis using curvelet transform and entropy features extracted from ultrasound images. *Computers in biology and medicine* 79, (2016) 250–258.

[53] E. Candes, L. Demanet, D. Donoho, and L. Ying. Fast discrete curvelet transforms. *Multiscale Modeling & Simulation* 5, (2006) 861–899.

[54] N. Dasgupta and L. Carin. Texture analysis with variational hidden Markov trees. *IEEE transactions on signal processing* 54, (2006) 2353–2356.

[55] M. N. Do and M. Vetterli. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. *IEEE transactions on image processing* 11, (2002) 146–158.

[56] G. Van de Wouwer, P. Scheunders, and D. Van Dyck. Statistical texture characterization from discrete wavelet representations. *IEEE transactions on image processing* 8, (1999) 592–598.

[57] R. Kakarala and P. O. Ogunbona. Signal analysis using a multiresolution form of the singular value decomposition. *IEEE Transactions on Image processing* 10, (2001) 724–735.

[58] S. H. Jensen, P. C. Hansen, S. D. Hansen, and J. A. Sorensen. Reduction of broad-band noise in speech by truncated QSVD. *IEEE Transactions on Speech and Audio Processing* 3, (1995) 439–448.

[59] B. De Moor. The singular value decomposition and long and short spaces of noisy matrices. *IEEE transactions on signal processing* 41, (1993) 2826–2838.

[60] C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)* 2, (2011) 27.

[61] D. Meng, L. Zhang, G. Cao, W. Cao, G. Zhang, and B. Hu. Liver fibrosis classification based on transfer learning and FCNet for ultrasound images. *IEEE Access* 5, (2017) 5804–5810.

[62] S. Mallat. Group invariant scattering. *Communications on Pure and Applied Mathematics* 65, (2012) 1331–1398.

[63] J. Bruna and S. Mallat. Invariant scattering convolution networks. *IEEE transactions on pattern analysis and machine intelligence* 35, (2013) 1872–1886.

[64] E. Tola, V. Lepetit, and P. Fua. Daisy: An efficient dense descriptor applied to wide-baseline stereo. *IEEE transactions on pattern analysis and machine intelligence* 32, (2010) 815–830.

[65] J. Bruna and S. Mallat. Classification with invariant scattering representations. In IVMSP Workshop, 2011 IEEE 10th. IEEE, 2011 99–104.

[66] J. Bruna. Scattering representations for recognition. Ph.D. thesis, Ecole Polytechnique X 2013.

[67] R. Lan and Y. Zhou. Medical image retrieval via histogram of compressed scattering coefficients. *IEEE journal of biomedical and health informatics* 21, (2017) 1338–1346.

[68] R. M. Haralick, K. Shanmugam et al. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics* 610–621.

[69] M. M. Galloway. Texture analysis using grey level run lengths. *NASA STI/Recon Technical Report N* 75.

[70] X. Tang. Texture information in run-length matrices. *IEEE transactions on image processing* 7, (1998) 1602–1609.

[71] A. Oliva and A. Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision* 42, (2001) 145–175.

[72] U. R. Acharya, H. Fujita, S. Bhat, U. Raghavendra, A. Gudigar, F. Molinari, A. Vijayananthan, and K. H. Ng. Decision support system for fatty liver disease using GIST descriptors extracted from ultrasound images. *Information Fusion* 29, (2016) 32–39.

[73] A. Sarwal, F. O. Walker, and M. S. Cartwright. Neuromuscular ultrasound for evaluation of the diaphragm. *Muscle & nerve* 47, (2013) 319–329.

[74] P. Viola and M. J. Jones. Robust real-time face detection. *International journal of computer vision* 57, (2004) 137–154.

[75] C. Chu, J. Bai, L. Liu, X. Wu, and G. Zheng. Fully automatic segmentation of hip CT images via random forest regression-based Atlas selection and optimal graph search-based surface detection. In Asian Conference on Computer Vision. Springer, 2014 640–654.

[76] K. Říha, J. Mašek, R. Burget, R. Beneš, and E. Závodná. Novel method for localization of common carotid artery transverse section in ultrasound images

using modified Viola-Jones detector. *Ultrasound in medicine and biology* 39, (2013) 1887–1902.

[77] A. Mohan, C. Papageorgiou, and T. Poggio. Example-based object detection in images by components. *IEEE transactions on pattern analysis and machine intelligence* 23, (2001) 349–361.

[78] P. Simard, L. Bottou, P. Haffner, and Y. LeCun. Boxlets: a fast convolution algorithm for signal processing and neural networks. In Advances in Neural Information Processing Systems. 1999 571–577.

[79] Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences* 55, (1997) 119–139.

[80] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 1. IEEE, 2005 886–893.

[81] M. Douze, H. Jégou, H. Sandhawalia, L. Amsaleg, and C. Schmid. Evaluation of gist descriptors for web-scale image search. In Proceedings of the ACM International Conference on Image and Video Retrieval. ACM, 2009 19.

[82] A. Oliva. Gist of the scene. In Neurobiology of attention, 251–256. Elsevier, 2005.

[83] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on image processing* 10, (2001) 266–277.

[84] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam. Use of active shape models for locating structures in medical images. *Image and vision computing* 12, (1994) 355–365.

[85] C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on image processing* 7, (1998) 359–369.

[86] L. D. Cohen. On active contour models and balloons. *CVGIP: Image under-standing* 53, (1991) 211–218.

[87] D. Cremers, C. Schnorr, and J. Weickert. Diffusion-snakes: combining statistical shape knowledge and image information in a variational framework. In Variational and Level Set Methods in Computer Vision, 2001. Proceedings. IEEE Workshop on. IEEE, 2001 137–144.

[88] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International journal of computer vision* 22, (1997) 61–79.

[89] R. M. Haralick and L. G. Shapiro. Computer and robot vision. Addison-wesley, 1992.

[90] L. P. M Berson, A Roncin. Compound scanning with an electrically steered beam. *Ultrasonic imaging* 3, (1981) 303–308.

[91] M. Berson, A. Roncin, and L. Pourcelot. Compound scanning with an electrically steered beam. *Ultrasonic imaging* 3, (1981) 303–308.

[92] J. W. Goodman. Statistical properties of laser speckle patterns. In Laser speckle and related phenomena, 9–75. Springer, 1975.

[93] L. Ma, L. Moisan, J. Yu, and T. Zeng. A dictionary learning approach for Poisson image deblurring. *IEEE Transactions on medical imaging* 32, (2013) 1277–1289.

[94] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image processing* 15, (2006) 3736–3745.

[95] B. E. Treeby and B. T. Cox. k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields. *Journal of biomedical optics* 15, (2010) 021,314.

[96] V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. Holtzman. A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Transactions on pattern analysis and machine intelligence* 157–166.

[97] J.-S. Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE transactions on pattern analysis and machine intelligence* 165–168.

[98] Y. Yu and S. T. Acton. Speckle reducing anisotropic diffusion. *IEEE Transactions on image processing* 11, (2002) 1260–1270.

[99] B. Matalon, M. Elad, and M. Zibulevsky. Improved denoising of images using modelling of a redundant contourlet transform. In Wavelets XI, volume 5914. International Society for Optics and Photonics, 2005 59141Y.

[100] O. G. Guleryuz. Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part I: theory. *IEEE Transactions on image processing* 15, (2006) 539–554.

[101] S. W. S. J. M. S. Wagner, Robert F. and H. Lopez. Statistics of speckle in ultrasound B-scans. *Sonics and Ultrasonics* 3, (1983) 156–163.

[102] X. Zhang, Y. Xu, H. Hu, Y. Liu, Z. Guo, and Y. Wang. Profiling skype video calls: Rate control and video quality. In INFOCOM, 2012 Proceedings IEEE. IEEE, 2012 621–629.

[103] K. Yamagishi and T. Hayashi. Qrp08-1: Opinion model for estimating video quality of videophone services. In Global Telecommunications Conference, 2006. GLOBECOM'06. IEEE. IEEE, 2006 1–5.

[104] S. Winkler and R. Campos. Video quality evaluation for Internet streaming applications. In Human Vision and Electronic Imaging VIII, volume 5007. International Society for Optics and Photonics, 2003 104–116.

[105] [Online]. Available: http://engineering.purdue.edu/ mark/puthesis .

[106] H. Lundin, S. Holmer, and H. Alvestrand. A google congestion control algorithm for real-time communication on the world wide web. *IETF Informational Draft, 2012* .

[107] [Online]. Available:https://www.youtube.com/watch?v=HCE3S1E5UwY .

[108] [Online]. Available: https://nodejs.org/en/.

[109] Network Emulator for Windows Toolkit (NEWT),[Online]: Available:https://blog.mrpol.nl/2010/01/14/network-emulator-toolkit/ .

[110] [Online]. Available: https://manycam.com/ .

[111] [Online]. Available: Video streaming recording software:camstudio.org .

[112] H. Tchelepi, P. W. Ralls, R. Radin, and E. Grant. Sonography of diffuse liver disease. *Journal of Ultrasound in Medicine* 21, (2002) 1023–1032.

[113] R. I.-R. BT. Methodology for the subjective assessment of the quality of television pictures .

[114] R. B. ITU-R. 500-12 Methodology for the subjective assessment of the quality of television pictures, 2009. *Google Scholar* .

[115] J. Kang, C. Yoon, J. Lee, S.-B. Kye, Y. Lee, J. H. Chang, G.-D. Kim, Y. Yoo, and T.-k. Song. A system-on-chip solution for point-of-care ultrasound imaging systems: Architecture and ASIC implementation. *IEEE transactions on biomedical circuits and systems* 10, (2016) 412–423.

[116] H.-Y. Tang, D. Seo, U. Singhal, X. Li, M. M. Maharbiz, E. Alon, and B. E. Boser. Miniaturizing ultrasonic system for portable health care and fitness. *IEEE transactions on biomedical circuits and systems* 9, (2015) 767–776.

[117] G.-D. Kim, C. Yoon, S.-B. Kye, Y. Lee, J. Kang, Y. Yoo, and T.-K. Song. A single FPGA-based portable ultrasound imaging system for point-of-care appli-

cations. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 59.

[118] S. P. Stawicki and D. P. Bahner. Modern sonology and the bedside practitioner: evolution of ultrasound from curious novelty to essential clinical tool 2015.

[119] F. K. Schneider, A. Agarwal, Y. M. Yoo, T. Fukuoka, and Y. Kim. A fully programmable computing architecture for medical ultrasound machines. *IEEE Transactions on Information Technology in Biomedicine* 14, (2010) 538–540.

[120] S. Granchi, E. Vannacci, E. Biagi, and L. Masotti. Multidimensional spectral analysis of the ultrasonic radiofrequency signal for characterization of media. *Ultrasonics* 68, (2016) 89–101.

[121] M. Moradi, P. Abolmaesumi, P. A. Isotalo, D. R. Siemens, E. E. Sauerbrei, and P. Mousavi. Detection of prostate cancer from RF ultrasound echo signals using fractal analysis. In Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE. IEEE, 2006 2400–2403.

[122] P. Tortoli, L. Bassi, E. Boni, A. Dallai, F. Guidi, and S. Ricci. ULA-OP: An advanced open platform for ultrasound research. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 56, (2009) 2207–2216.