COMPUTER AIDED DIAGNOSIS OF LIVER DISEASES USING ULTRASOUND IMAGES

A Thesis

Submitted in the fulfilment of the requirement for the award of the degree of

DOCTOR OF PHILOSOPHY

in

ELECTRICAL AND INSTRUMENTATION ENGINEERING

Submitted by

Puja Bharti

(Registration No.: 951204004)

Under the supervision of

Dr. Deepti Mittal

Associate Professor, EIED, TIET

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Electrical and Instrumentation Engineering Department
Thapar Institute of Engineering & Technology,
Patiala -147004, Punjab, India

Oct 2019

DECLARATION

I, **Puja Bharti** hereby certify that the work presented in this thesis report entitled "**Computer Aided Diagnosis of Liver Diseases Using Ultrasound Images**" in fulfilment of the requirements for the award of the degree of **Doctor of Philosophy** and submitted in Electrical and Instrumentation Engineering Department of Thapar Institute of Engineering & Technology, Patiala, is an authentic record of my own research work carried out during a period from July 2012 to Sep 2019 under the supervision of **Dr. Deepti Mittal,** Associate Professor, Electrical and Instrumentation Engineering Department of Thapar Institute of Engineering & Technology, Patiala.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Puja Bharti

(Puja Bharti)

Registration No. 951204004

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

(Dr. Deepti Mittal)

Associate Professor

Electrical and Instrumentation Engineering Department

Thapar Institute of Engineering and Technology (Deemed to be University)

Patiala, Punjab, India

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. Deepti Mittal, Associate Professor, Department of Electrical and Instrumentation Engineering, Thapar Institute of Engineering & Technology, Patiala, for introducing the present topic and for her inspiring guidance, constructive criticism, essential support, valuable suggestions and constant encouragement during the entire research work. Her dynamism and foresightedness helped me in a big way in completion of this work. I really feel privileged to have worked under her supervision.

I express my sincere thanks to the members of my doctoral research committee Dr. Sanjay Sharma, Professor, Department of Electronics and Communication Engineering, Thapar Institute of Engineering & Technology; Dr. Gagandeep Kaur, Associate Professor, Department of Electrical and Instrumentation Engineering, Thapar Institute of Engineering & Technology; and Dr. Jainy Sachdeva, Associate Professor, Department of Electrical and Instrumentation Engineering, Thapar Institute of Engineering & Technology for providing necessary facilities and valuable suggestions throughout this research work and in the submission of this manuscript.

I am very thankful to Dr. R.S. Kaler, Senior Professor and Head, Department of Electrical and Instrumentation Engineering, Thapar Institute of Engineering & Technology and Dr. Mandeep Singh, Professor and PhD Coordinator, Department of Electrical and Instrumentation Engineering, Thapar Institute of Engineering & Technology for their constructive discussions and valuable suggestions during the meetings of my research committee. I am also thankful to all other faculty members of the Department of Electrical and Instrumentation Engineering for their support.

I extend my warm thanks to all teaching and non-teaching staff, Department of Electrical and Instrumentation Engineering, Thapar Institute of Engineering & Technology for their day-to-day co-operation.

I would like to thank Dr. Rupa Ananthasivan, Consultant from the Department of Radiology, Manipal hospital, Bangalore for granting ethical clearance to the research proposal and for their clinical advice. I am also thankful to the clinicians of the department to provide all the necessary help in collecting and providing liver ultrasound images data.

I will certainly fail in my duty if I am not to thank my friends: Deepika, Heena, Sharad, Sikandar, Abhay and Suman for their support, care and affection. It is impossible for me to realize 'my world' without them.

I am indebted to my beloved parents and in-laws for all their support during the PhD work. I express my deepest appreciation for my brother Inderjeet, sisters Sanjeevni and Ruchi and uncle R.K.Garg whose blessings and wishes helped me to reach this far.

Finally, I owe deeply to my husband Yuvraj and daughter Gitisha for their unwavering patience, understanding, and encouragement, as the present work is finally concluded because of their support for which I shall ever remain indebted.

(Puja Bharti)

Abstract

Liver is one of the most important organs in the human body, because it plays a vital role in digestion and in the metabolization of protein, drugs, toxins etc. Diseases of the liver are becoming major cause of morbidity and mortality all over the world. Ultrasound is preferable imaging modality by medical practitioners for the screening of liver diseases. It is preferable because of its several advantages; it is a non-invasive, cost effective, non-ionizing, portable and real time imaging modality. Medical practitioners/radiologists diagnose liver diseases by the visual examination of ultrasound images. Visual examination is a subjective criterion and is highly dependent on the expertise of radiologists in the domain. This may lead to ambiguity in the diagnostic procedure. To improve the objectivity in the diagnostics of liver diseases, various approaches of designing computer-aided support system are explored in this thesis. In addition, the visualization of liver diseases in ultrasound images becomes a tedious task due to overlapping textural characteristics of liver diseases. Thus, to design the computer-aided diagnostic (CAD) system, in-depth textural analysis of ultrasound images is performed to quantify information related to liver diseases.

This research work is carried out with a database of 189 B-mode ultrasound images. The database was developed by collecting images from patients visiting Manipal Hospital, Bangalore, India during the period from March 2013 to August 2014. These patients were recommended for liver examination. In this duration of one year, ultrasound images were collected from 94 patients; among those 67 were male (age range: 21-70) and 27 were female (age range: 23-61). The database encompasses 48 images of normal liver and 141 images of abnormal liver. Images related to abnormal liver comprise of 50 images of chronic liver, 50 images of cirrhotic liver, and 41 images related to HCC evolved over cirrhosis.

Diagnostically relevant areas in liver images are regions-of-interest (ROIs). The ROIs were identified and marked by the radiologist in all the images. Regions-of-interest were segmented into maximum possible square shape regions to create a large dataset for research work. These

segmented regions are termed as segmented-regions-of-interest (SROIs). In this research work, three datasets are formed to optimize the design of CAD system. First dataset consists of 754 SROIs of size 32 x 32 pixels and is termed as original1 dataset. Second dataset consists of 400 SROIs of size 227 x 227 pixels and is termed as original2 dataset. Third dataset consists of ultrasound images on which enhancement was performed. This dataset is termed as enhanced dataset and has 754 SROIs of size 32 x 32 pixels.

Ultrasound images suffer from low contrast and inherent formation of speckle. This may affect visual evaluation of liver diseases by radiologist. Therefore, in this research work, an image enhancement method is designed to reduce speckle and improve contrast of ultrasound images without the loss of diagnostic information. The proposed method is based on the scaling with neutrosophic similarity score (NSS), where an image is represented in the neutrosophic domain through three membership subsets; degree of truth *(T)*, indeterminacy *(I)* and falseness *(F)*. The NSS measures the belonging degree of pixel to the texture and is calculated with multi-criteria: intensity, local mean intensity and edge detection. Then, NSS is utilized to extract the enhanced coefficient and this coefficient is applied to scale the input image. This scaling reflects contrast improvement and denoising effect on ultrasound images. The performance of proposed enhancement method is evaluated using both subjective and objective criteria. Results demonstrate that the proposed enhancement method enhances the contrast of ultrasound images and thus making the visualization of diagnostic information easier for the radiologist in comparison to original images.

In ultrasound images, echo pattern of liver represents the texture of liver tissues. Hidden information in texture can be extracted using various mathematical descriptors to quantify information. These mathematical descriptors are termed as features and play a crucial role in differentiating liver tissue types. The choice of features also has an important influence on the accuracy of learning algorithm and the time required for execution. In this research work, features are obtained from traditional hand-designed methods and deep learned methods. The features obtained via traditional hand-designed methods are termed as handcrafted features. On the other hand, the features obtained via deep learned methods are termed as deep learned features.

On the basis of extensive literature review, the traditional hand-designed methods chosen for this research work are (i) gray-level difference matrix (GLDM), (ii) gray-level co-occurrence matrix (GLCM) and (iii) ranklet transform (*ranklets*). GLDM method relies on the calculation of local derivatives between pair of gray-levels. These local derivatives can provide information about the roughness of liver surface. In literature, the derivative calculation of GLDM method is based on the absolute differences between two neighboring pixels. However, the use of pixel pair, only in calculation of local derivatives, may not be able to provide the best approximation of calculated features from GLDM. Hence in this research work, GLDM features are calculated with better approximation of derivatives by using greater than two pixels. This proposed extension of GLDM is designed by computing the absolute difference of intensities using three, five and seven pixels. A total set of 148 features (64 of GLDM, 48 of GLCM and 36 *ranklets*) were extracted from handcrafted methods. However, this feature space may contain redundant features and the relevance of features in discriminating among liver tissue types is not known in advance. Thus, a hybrid feature selection method is designed in this research work. This method selects the optimal set of features to reduce computational complexity and improve performance. Further, the features extracted from methods such as GLDM and its extension, GLCM and *ranklets* may contain complementary information. To take the benefit of this information, feature fusion schemes (i) serial feature combination, (ii) serial feature fusion and, (iii) hierarchical feature fusion are also implemented in this research.

The deep learned features extracted from deep neural network can be beneficial as they present the possibility of learning complex patterns from a large dataset. The hidden layers inside the deep neural network architecture automatically extract features without user intervention to give the best texture discriminating performances in a multi-class texture classification problem. In computer vision, convolutional neural networks (CNNs) are one of the most successful feature learning methods in performing various tasks. In this research work, two major approaches that employ CNN for feature extraction to classify liver diseases: (i) pre-trained CNN, and (ii) fine-tuning with transfer learning, are utilized**.** In the first approach, the convolutional base is kept in its original form and then its outputs are used to feed the classifier. Here the pre-trained AlexNet model is used and features are extracted from its different layers. In the second approach, the pre-trained AlexNet model is adapted to the current research problem by replacing the last fully connected layer (intended for 1000 classes) with a new fully connected layer for four classes. The initial filter weights of CNN are derived from ImageNet dataset and then fine-tuned (optimized) through back-propagation so that they better reflect the four liver classes.

The success of a computer-based system depends both on the features and classification method. An efficient set of textural features decides the correct detection of liver diseases and an appropriate classification method provides potential to produce accurate classification using these features. Thus, in this research an ensemble model is designed with different base classifiers: k-NN, SVM, and RF; and combiner: majority voting. Independent training of all the base classifiers is done and finally, output of classifiers is combined with majority voting. The proposed ensemble model can be regarded as a combination of different strategies that enhances the generalization ability of model, as each component in the model learn about some part of classification problem. The CAD system designed by using this proposed ensemble model and handcrafted features extracted from original1 dataset gave 96.6% classification accuracy. Sensitivity and specificity of normal/chronic/cirrhosis/HCC were 96.3%/95.5%/97.5%/96.9% and 99.2%/98.0%/98.2%/99.8% respectively. The CAD system designed with proposed ensemble model and deep learned features extracted from original2 dataset gave 68.4% and 75.4% classification accuracy with pre-trained CNN and fine-tuned CNN. Here, it was observed that CAD system designed with handcrafted features gave better performance for classifying liver tissue type as compared to deep learned features. Finally, a CAD system was designed with handcrafted features and proposed ensemble model with enhanced dataset gave 97.9% accuracy which show the high precision capability of the system in predicting the test data to its actual liver tissue class. The experiment results also indicate the high performance of this proposed CAD system with the overall average sensitivity of 97.4% revealing its ability to perform well in predicting the true positive cases in each liver tissue type correctly. Sensitivity and specificity of normal/chronic/cirrhosis/HCC were 97.9/96.5/98/97.5% and 99.2/99.2/98.5/99.8% respectively. It was observed that there was 1.3% increase in accuracy with the CAD system designed using enhanced dataset in comparison to the CAD system designed with original1 dataset.

Finally, it can be concluded that the CAD system designed to classify four liver tissue types comprises of (i) enhanced ultrasound images, (ii) hierarchal fused features set of handcrafted features and, (iii) ensemble classifier. The substantial performance of experiments on clinical database supports the strong candidature of the proposed computer-aided system to assist radiologist/clinicians in the diagnosis of liver diseases.

Contents

List of Abbreviations

CNN	convolutional neural network
CHL	chronic liver
CIR	Cirrhosis
NOR	Normal
HCC	hepatocellular carcinoma
SVM	support vector machine
RF	rotation forest
ANN	artificial neural network
k-NN	k-nearest neighbor
GLCM	gray-level co-occurrence matrix
GLDM	gray level difference matrix
HFS	hybrid feature selection
SFS	sequential forward selection
SBS	sequential backward selection
NSS	neutrosophic similarity score
TEM	texture energy measurement
AMBE	absolute mean brightness error
PSNR	peak-signal-to-noise-ratio
SSIM	structural similarity index
MS-SSIM	multi-scale structural similarity
UIQI	universal image quality index
GLRLS	gray-level run length statistics
CAD	computer-aided diagnosis
ROI	region-of-interest
SROIs	segmented-regions-of-interest
NS	neutrosophic set
NSS	neutrosophic similarity score
CFS	correlation feature selection
GA	genetic algorithm

PSO	particle swarm optimization
PPV	positive predictive value
NPV	negative predictive value
PCA	principal component analysis
CT	computed tomography
MRI	magnetic resonance imaging
SFS_{kNN}	SFS with kNN as classifier
SFS_{SVM}	SFS with SVM as classifier
TP	true positive
TN	true negative
FN	false negative
FP	false positive
$HFS_{k\text{-}NN}$	hybrid feature selection with kNN
HFS_{SVM}	hybrid feature selection with SVM
$f(i,j)$	original ultrasound image
$f_{NS}(i,j)$	neutrosophic image
c_{in}	intensity criteria
c_{mi}	local mean intensity criteria
c_e	edge detection criteria
$E(i,j)$	edge image using sobel operator
$f_{en}(i,j)$	enhanced image
$\bar{f}$	mean of original image
$\overline{f_{en}}$	Mean of enhanced image
σ_f	standard deviation of original image
$\sigma_{f_{en}}$	standard deviation of enhanced image
RC	ranklet coefficient
conv	Convolution
fc	Fully connected

List of Figures

List of Tables

Chapter 1

Introduction

1.1 OVERVIEW

Liver is one of the most important organs in the human body. It has a key role in the digestion and regulation of protein, carbohydrate, lipid and amino acid metabolism, and is also responsible for metabolization of environmental toxins and drugs [1]. In light of this, it is highly expected that diseases of the liver are substantial reason of illness and mortality. Liver diseases such as chronic liver, cirrhosis, and hepatocellular carcinoma (HCC), represent a failure in hepatic metabolic and synthesis processes. These diseases accounts for approximately two million deaths per year worldwide, one million deaths due to complications of cirrhosis and one million deaths due to viral hepatitis and HCC [2]. According to the estimates of World Health Organization (WHO) in India 259,749 or 2.95% of total deaths are due to these liver diseases [3]. The major causes of these diseases are infection with hepatitis virus A, B, C, D or E; obesity; metabolic disorder or over consumption of alcohol [4]. In the year 2018, 325 million people were infected with hepatitis B virus and hepatitis C virus globally [5]. Another survey presents that nearly 65% of people with chronic hepatitis B and 75% of those with chronic hepatitis C are unaware of their infection [6]. Usually, the affected people realize their infection only when they are persistently ill. Thus, the need for early detection of liver diseases is necessary, as it can facilitate timely treatment.

The early detection needs a reliable and accurate diagnostic procedure that allows doctors to detect the severity of liver disease. Although there are many diagnostic modalities, liver biopsy is the gold standard for diagnosing progression of liver disease [7]. However, biopsy is invasive and cannot be used repeatedly in follow-ups. It also associates pain and risk of bleeding. Therefore, it is important to consider noninvasive diagnostic modalities in differentiating of liver diseases.

1.2 LIVER IMAGING MODALITIES

The liver imaging modalities such as ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI) are commonly used diagnostic modalities by the expert radiologists to detect liver diseases [8]. These imaging modalities are non-invasive as they do not penetrate the skin physically. They play a vital role in detection of liver diseases by identifying the appearance of liver surface, size, contours etc. However, there is no perfect liver imaging modality for all hepatological applications and needs. Additionally, each imaging modality is limited by its mechanism and nature of image being formed. The selection of modality to be used for diagnosis is based on the expert radiologist, the availability of the equipment, the condition of the patient, and the experience of the radiologist. Moreover, most of the details related to abnormality should be detected so that the best treatment to the patient is ensured. Therefore, choosing the modality with the most benefits and least limitations is essential. With the brief introduction of MRI, CT and ultrasound modalities in the following paragraphs, some comparative statements are made to get the better understanding of the preferences among imaging modalities in order to detect liver diseases.

Magnetic resonance imaging form a high-resolution detailed cross-section images of the internal organs and structures [9]. It is usually used in the detection of joint and bone complications in addition to torn cartilage and ligaments. The MRI scanner looks like a long tube with a table in the middle, allowing the patient to slide in and out. The scanner produces magnetic field around the patient and pulses radio waves into the area of the body to be imaged [10]. MRI scan has several advantages i.e. non-invasive and painless procedure, gives a detailed image as compared to ultrasound and CT scan. However, some of the prominent limitations of MRI scan are (i) it is an expensive imaging technique, (ii) the scan time is longer as compared to CT and ultrasound i.e. 30-40 mins, (iii) its availability is limited due to its expensive machine and installation, and (iv) the machine produces a lot of noise.

A computed tomography scan, uses a more powerful and sophisticated X-ray to form a detailed, high-quality images of bones, organs, blood vessels and soft tissue [11]. The CT scan takes a 360º image of the internal structures. The CT scanner has a slab in the middle on which patient lie down, the scanner rotates around the patient and cross-section images of the body part are produced. CT scan has several advantages i.e. non-invasive and painless procedure and gives a detailed image. However, some of the prominent limitations of CT scan are (i) it is an expensive imaging technique than ultrasound, (ii) the scan time is longer as compared to ultrasound i.e. 10-30 mins, (iii) high doses of radiation are involved which may increase the probability of cancer, and (iv) its availability is limited due to its expensive machine and installation.

An ultrasound scan, uses high-frequency sound waves to form images of internal organs and tissues [12]. A transducer is used to take an ultrasound image. Transducer emits high-frequency sound waves and these waves disperse through the body. After sound waves are emitted into the body, the waves bounce off of the structures inside body and travel back to the transducer. The duration and patterns in which these waves come back to the transducer is used to form an ultrasound image. These waves are continuously moving and follow the transducer, producing real-time images. Ultrasound scan has several advantages i.e. non-invasive, painless, cost effective, shortest scan time compared to CT and MRI, portable and easy to install, no risk of radiations and gives a real-time image. The disadvantage of ultrasound scan is that the diagnostic accuracy is lower as compared to CT and MRI.

In summary, MRI, CT and ultrasound imaging modalities of liver have their own advantages and disadvantages. Ultrasound is an attractive imaging modality in general because of unique set of features. These features include (i) low cost that makes it feasible for a patient from low socio-economic status, (ii) widespread availability that makes its clinical relevance high worldwide, (iii) risk-free system that makes it the first preference among other imaging modalities for screening of liver abnormalities, and (iv) fast and portable that makes it convenient to use. In continuation, the ultrasound images are not difficult to study once the viewer is trained. As a result, ultrasound imaging is an extremely useful type of diagnostic imaging and the most widely employed imaging technique for screening purposes and detection of liver abnormalities. Therefore, ultrasound images are utilized in the present work to design computer-aided system to diagnose liver abnormalities.

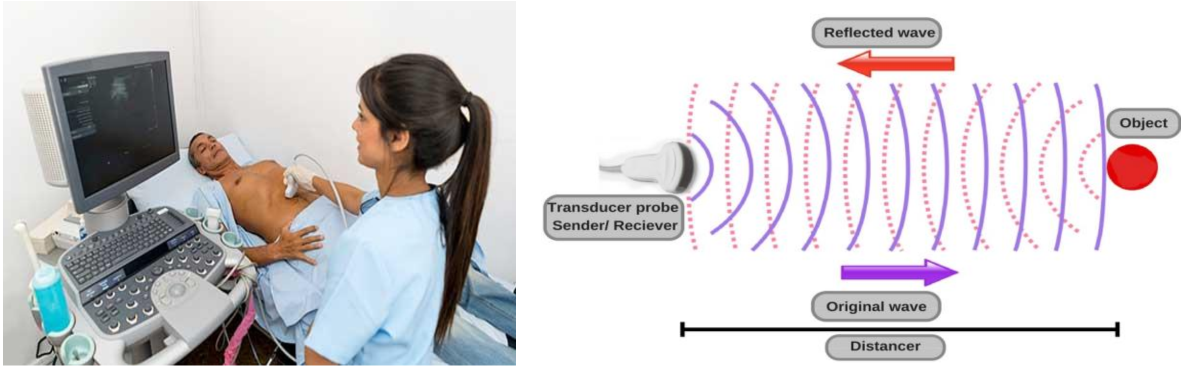
1.3 ULTRASOUND IMAGING

Ultrasound imaging modality create images of tissue by using uses high-frequency sound waves [13]. The ultrasound machine sends ultrasound pulse into the body and is able to convert the returned echoes into an image. These echoes contain spatial and contrast information. The ultrasound imaging is useful in imaging fluid filed spaces and soft tissues like the liver, heart, pancreas, kidneys, and blood vessels of the neck and abdomen [12]. The basis of its operation, as shown in figure 1.1, is the transmission of sound waves by transducer into the body. Then transducer receives the sound waves that bounce off and these sound waves are used to create an image.

Radiologist use these liver images to follow, diagnose, and treat liver abnormalities such as chronic liver, cirrhosis, liver cancer etc. To comprehend an ultrasound image of the liver, it is necessary to know how the ultrasound imaging system works and how reflected sound is converted into an image.

1.3.1 Ultrasound Image Formation

The ultrasound scanner comprises of two primary parts: (i) a transducer probe and (ii) a computer system [14].



(a) (b)

Figure 1.1 (a) Liver imaging with ultrasound, (b) Path of transmitted and reflected sound wave from transducer probe

In the beginning of ultrasound scan, a gel is applied on the selected region of the skin to be diagnosed. The application of gel is an important step to impede formation of small air bubbles, which could otherwise possibly form in-between the probe and skin and hinder the sound waves during transmission and reception. Arc-shaped sound beam is generated from the transducer probe which is focused on the selected area. The transducer probe is a hand-held device attached to the scanner by a cord. This probe contains one or more quartz crystals called piezoelectric crystals. These crystals vibrate when electric current is applied to them and produce sound waves with higher frequencies. The location of the tissue through which ultrasound beam will penetrate determine the type of transducer probe to be used in examination. The ultrasound frequencies used in diagnostics, range from 2 MHz to approximately 15 MHz. The transducer probes with high-frequency provide less penetration of the ultrasound waves through the tissue; however, ultrasound images of higher-resolution are produced. These probes are usually employed to visualize superficial structures, namely pulmonary pleura, muscle, tendons and vasculature. On the other hand, lower-frequency probes provide more penetration and employed to image deeper structures. However, the potential to image deeper structures results in low-resolution of ultrasound image. These probes are beneficial in examining internal organs such as the inferior vena cava, the abdominal aorta, gallbladder and pelvic. For instance, the convex array or curvilinear probe with frequency range of 2-5 MHz is largely utilized for examining deep structures in the abdominal region. The transmitted ultrasound waves from transducer probe pass through the thin layer of skin, but bounce off internal organs, fluids and tissues. These bounced or reflected waves are received by the probe, which are converted into electric signals and sent to computer. The computer uses this information and performs calculations, considering the speed of waves and the time taken by echo to reach the probe, to generate real-time images. During ultrasound scan, real-time images of the tissue structure are acquired in B-mode scanning. These gray-scale B-mode or brightness mode images comprise of bright dots which represents the ultrasound echoes and their brightness is regulated by the strength of echo. These gray-scale images are visualization of tissue structures and useful for diagnostic purposes. In the cases such as examining the abdominal regions, where ultrasound beam travels deeper, the returning echoes are attenuated and result in lower resolution. Then time gain compensation helps radiologists fine-tune the image gray-scale at specific depths. Finally, radiologist freeze and save the still images. These images can be printed or stored on hard disk of ultrasound machine.

1.3.2 Speckle

An anatomical display of an organ on ultrasound images aids the radiologist in assessing the health of the organ. The speckle is an unavoidable characteristic of ultrasound imaging that degrades the quality of ultrasound image [15]. Speckle presence is due to interference effects between overlapping echoes. The superposition of returning echoes coming with random phases and amplitudes tends to produce complex interference pattern, defined as speckle noise. It demonstrates little association to the macroscopic properties of tissue. This noise has the tendency to mask the important diagnostic details, which results in diverting the diagnosis. However, speckle noise is a random process, it is not without information. The statistics of speckle, largely rely on the microstructure of tissue, can be helpful for distinguishing tissue types.

1.4 LIVER ANATOMY AND LIVER DISEASES

The anatomical view of human liver is demonstrated in figure 1.2. Liver is the largest organ that weigh approximately 1.5kg and is located in the upper right quadrant of the abdomen [1]. It is reddish brown in color and is wedge-shaped organ. The two large sections of liver are left and right lobes. These left and right lobes are further segmented into eight parts. The liver is connected to common hepatic duct and two large blood vessels: the hepatic artery and the portal vein and also closely connected with the small intestine. The main functions of liver are (i) metabolism of drugs and environmental toxins, (ii) regulation of blood clotting by producing proteins, (iii) metabolism of bilirubin and production of bile, and (iv) regulation of protein, carbohydrate and lipid metabolism.

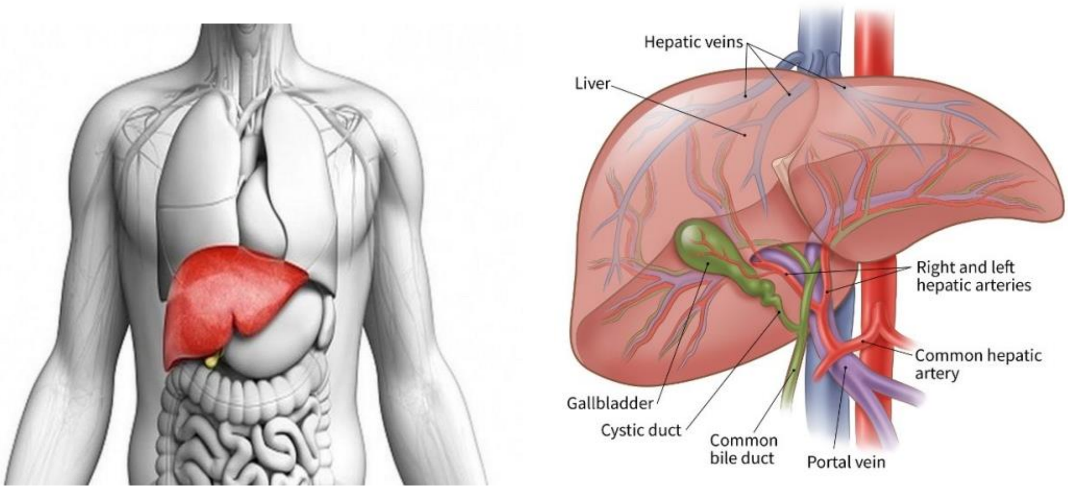


Figure 1.2 Photographic view of liver

The ultrasound evaluation of liver is performed by assessing various liver characteristics such as the liver size, contour, echogenicity, structure, portal vein velocity and spleen size [16]. The echogenicity is the ability to bounce an echo. The organs are characterized based on the amount of echoes they generate i.e. echogenicity and how equally distributed these echoes are i.e. homogeneity. The regions that reflect no echoes are anechoic and appear black in ultrasound. The regions that reflect less echoes as compared to their neighbouring structures are hypoechoic and appear as varying shades of darker grey. The regions that are reflect more echoes as compared to their neighboring structures are hyperechoic and appear as varying shades of lighter grey. The structure of liver is determined as either 'homogeneous' or 'course' echo pattern. Experienced radiologist visualizes these features in liver ultrasound images to characterize the normal liver and liver abnormalities.

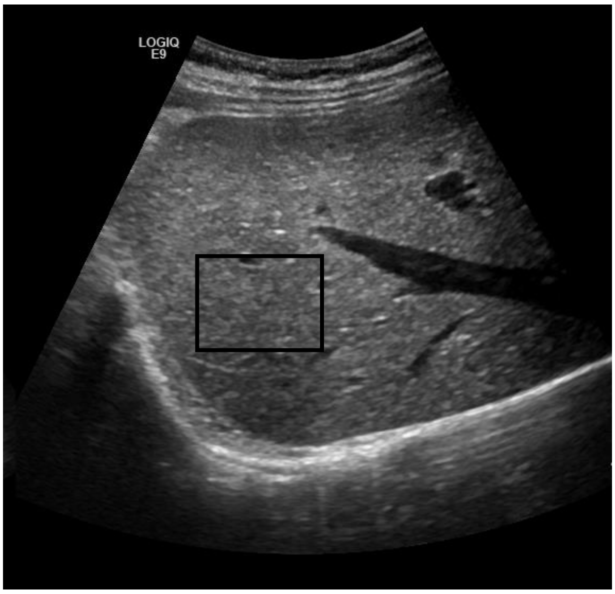


Figure 1.3 Appearance of normal liver on ultrasound image

1.4.1 Normal Liver

The characteristics of normal liver are (i) smooth liver surface with no lumps protruding or indentations, (ii) sharp liver borders, (iii) homogeneous echotexture of liver parenchyma, (iv) normal liver and spleen size and (v) normal portal vein diameter [17]. The appearance of normal liver on ultrasound is isoechoic and homogeneous as shown in figure 1.3. The marked rectangular area shows the normal liver texture.

The normal liver shows homogeneous echo reflectivity, composed of low and medium echoes. The hepatic and portal vein, the hepatic arteries and bile duct appear as anechoic structures. These ultrasound findings of liver can provide valuable information about normal liver.

1.4.2 Chronic Liver

Chronic liver refers to disease of the liver which lasts over a prolonged period of six months. It involves a process of progressive destruction and regeneration of the liver parenchyma. The prevalent causes of disease are alcohol abuse, viral hepatitis and metabolic disorders which leads to damage of liver cells. The chronic liver is major cause of illness and death in the developing countries. The accurate evaluation of the disease is crucial for timely treatment.

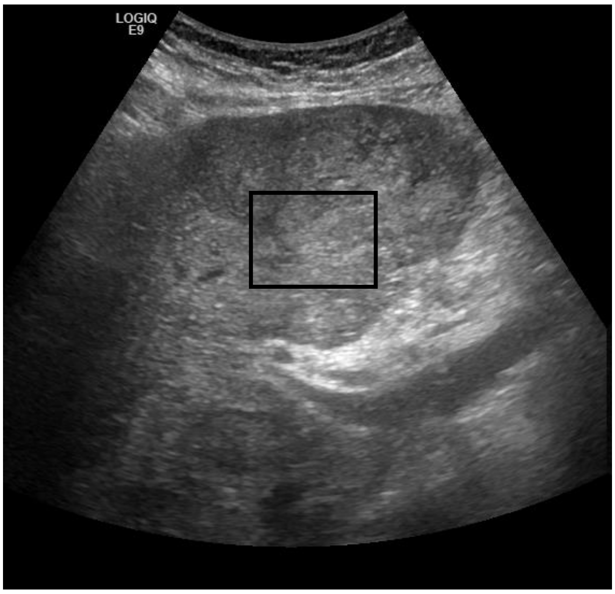


Figure 1.4 Appearance of chronic liver on ultrasound image

The characteristics of chronic liver are (i) coarse echotexture, (ii) irregular liver surface, (iii) blunt liver edge, (iv) enlarged liver (>15cm mid clavicular line) and spleen size (>13 cm), and (v) dilated portal vein diameter (>13mm). The appearance of chronic liver on ultrasound is heterogeneous as shown in figure 1.4. The marked rectangular area shows the chronic liver texture. The coarse echotexture is due to the loss of uniform smooth hepatic echotexture of liver due to chronic hepatitis.

1.4.3 Cirrhosis

Cirrhosis is a liver disease characterized by hepatic fibrosis and formation of nodules. These changes generally affect the whole liver. Cirrhosis evolves slowly over years to decades. This disease can happen at any phase of life, causes critical liver injury and is a primary reason of premature demise. This disease is usually caused by prolonged excessive alcohol consumption and chronic viral hepatitis. The prolonged biliary damage or obstruction, and persistent blockage of vein returning from the liver are also contributing reasons for cirrhosis occurrence. The patients with this disease may show symptoms of jaundice, ascites, variceal haemorrhage and abdominal pain, weight loss and anorexia.

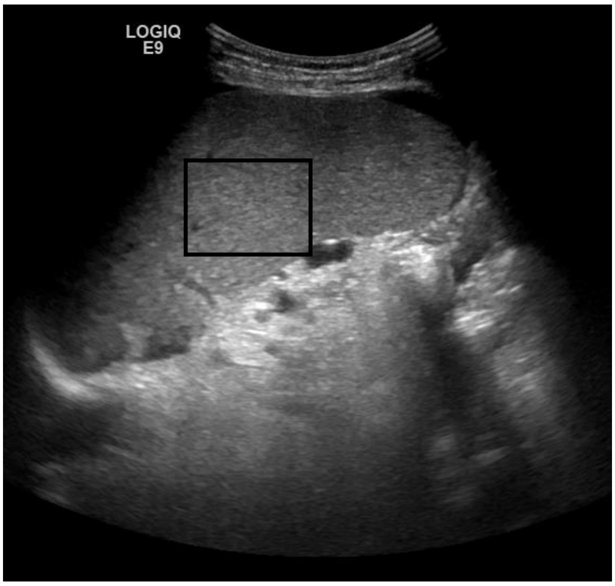


Figure 1.5 Appearance of cirrhosis liver on ultrasound image

The prime characteristics of liver cirrhosis are (i) increase in hepatic fibrosis, (ii) continuous and widespread death of hepatic or liver cells, and (iii) enlargement of liver resulting in loss of the normal liver architecture. The progressive hepatic cell destruction and increase in fibrosis lead to gradual reduction in liver size. The liver becomes irregular, hard and non-tender as the disease progresses. The clinical presentation of cirrhosis is highly inconsistent. In some cases, disease is entirely asymptomatic and may get diagnosed accidentally at ultrasound scanning. Other cases may show isolated enlargement of liver, splenomegaly or signs of portal hypertension. The appearance of cirrhosis liver on ultrasound image is highly coarse. In figure 1.5, the marked rectangular area shows the cirrhotic liver texture. The cirrhosis appear on liver with (i) highly coarse echotexture, (ii) nodular liver surface, (iii) round liver edge and (iv) shrunken liver size (<10cm in mid clavicular line) [18].

1.4.4 Hepatocellular Carcinoma Evolved Over Cirrhosis

Hepatocellular carcinoma is the sixth most common liver cancer world-wide [19]. According to WHO in 2018, 1.34 million deaths occurred due to viral hepatitis B and C globally [5]. At least 60% of liver cancer cases were due to untimely diagnosis and treatment of viral hepatitis B and C. The chronic hepatitis B infection is the foremost risk factor for the occurrence of HCC. Further, the possibility of HCC increases in patients having cirrhosis [4]. In 75–90% of patients having HCC also have cirrhosis and it is the major risk factor for HCC [19] . Many liver cancers detected are asymptomatic till they develop to an advanced stage. Macroscopically, it is usually seen that the HCC cancer appears in the form of a single or multiple nodule along with cirrhosis. Hepatic artery is the source from where cancer cells take blood and then grows by invasion into the portal vein and its radicals. Usually it is difficult to distinguish between the signs of a small tumor from those of cirrhosis. The marked area shown in figure 1.6 is the appearance of HCC evolved over cirrhotic liver on ultrasound.

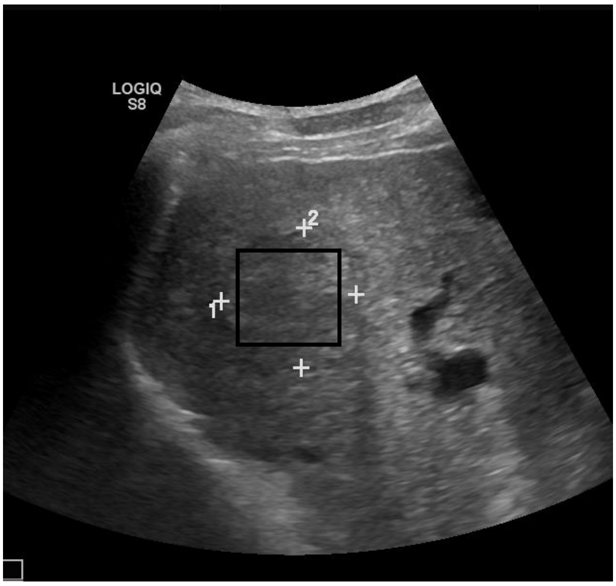


Figure 1.6 Appearance of HCC evolved over cirrhosis liver on ultrasound image

HCC evolved over cirrhosis may show (1) multifocal appearance initially, (2) increased or decreased echogenicity in relation to the adjacent liver parenchyma and (3) internal heterogeneity due to haemorrhage, necrosis or fat. For rest of the study, HCC evolved over cirrhosis is written as HCC.

1.5 NEED OF COMPUTER AIDED LIVER DISEASE DETECTION METHOD

Ultrasound imaging is one of the frequently and non-invasive technique for imaging liver. However, due to qualitative and subjective nature of acquiring the ultrasound images these images can be affected by imaging conditions such as illumination, machine settings. The visualization criteria for distinguishing various liver diseases is slightly confusing and highly based on the radiologist's experience. This results in ambiguity of the diagnostic procedure and reduces its objectivity and reproducibility. The visualization of liver diseases is a tedious task with ultrasound images because of their varying coarseness, contrast with respect to the background etc. Moreover, while working with a large number of images even the experienced radiologists can commit errors due to fatigue or boredom [20]. Thus, there is a need to develop a computer-aided detection (CAD) system that can be used by radiologists as a non-invasive diagnostic tool to support their observations. Liver disease diagnosis with the assistance of an efficient and reliable computer-aided system may also reduce the frequency of biopsy or further examinations.

1.6 NEED OF IMAGE ENHANCEMENT

It is crucial that all the characteristics of the abnormality present in an ultrasound image must be perceived clearly to obtain high diagnostic accuracy. The presence of noise, that occur during liver image acquisition, adversely effects the interpretation and analysis of liver ultrasound images. Speckle noise leads to low contrast and masks the diagnostically important features present in liver ultrasound image. This noise is introduced due to interference effect between overlapping echoes. The detection of liver abnormality by an experienced radiologist is done by visualizing the texture variations in a liver ultrasound image. Therefore, the texture of liver ultrasound images is highly desirable characteristics for accurate liver disease detection. Thus, there is a need to develop a methodology to enhance the visual quality of liver ultrasound images so that visualization of the liver texture improves and provide efficient detection of abnormality present in the image. In the present work, an enhancement method is designed to improve visualization of details available in liver ultrasound images in order to provide accurate detection of liver diseases.

1.7 NEED OF FEATURE EXTRACTION

The ultrasonic texture is a vital attribute in the radiological diagnosis. Texture variations in ultrasound images of liver play an important role in detection of liver abnormalities. The radiologists visualize ultrasonographic feature i.e. texture, in ultrasound image to differentiate among liver classes. These visual features provide guidelines for identifying suitable mathematical feature descriptors, which are utilized in the designing of CAD system. In an image analysis system, the textures are characterized by mathematical descriptors that help differentiate among textures. Further, texture analysis techniques have successfully been reported for differentiating liver classes. Thus, the extraction of appropriate set of mathematical features from ultrasound images is required to study the possibility of numerical separation among liver classes.

1.7.1 Handcrafted Feature

The CAD system for classification of liver diseases depend on extraction of local features from the images; consequently, substantial amount of CAD research work is concentrated on understanding, discovering and improving methods for extracting features [21], [22]. Number of features are manually designed, or 'handcrafted' by keeping application in mind. Designing of these features is usually a function of finding the right trade-off between accuracy and computational efficiency. The grey-level difference matrix (GLDM), for example, is frequently used to extract 'handcrafted' texture features for classification of liver diseases [23]. The GLDM method is based on local derivatives of grey levels and provide information about roughness of liver surface. Its observed that GLDM is computed with absolute difference between two neighboring pixels. However, use of only two pixels to calculate local derivatives may not provide the best approximation to calculate features, this may lead to incorrect characterization of liver classes. Thus, there is need to compute the local derivatives as accurately as possible. In addition to GLDM, the other frequently used feature extracted methods are grey-level co-occurrence matrices (GLCM) and ranklet transform (*ranklets*) for classification of liver classes [20], [24], [25].

Feature selection of handcrafted features:

The feature space, consists of all the extracted features, may contain features with less discriminating ability and regarded as redundant or irrelevant features. The presence of large number of features may result in learning model to overfit, resulting in degradation of model's performance [26]. Moreover, relative relevance of the feature is not known in advance. Additionally, there may be some noisy instances i.e., outliers which may not lie in their labeled class. Thus, selection of features is essential to reduce computational complexity and improve the classification accuracy by (i) handling noisy instances and (ii) removing redundant and irrelevant features. Feature selection targets to select a subset of relevant features from the feature space and this selected feature set generally results in improving learning performance and lowering computational time. Thus, in the present work, for feature selection, a hybrid feature selection (HFS) method is designed by combining filter (ReliefF) and wrapper (sequential forward selection [SFS]) methods.

1.7.2 Deep Learned Feature

Recently much of the research is going on medical diagnosis via deep neural networks. From literature it can be inferred that the problem of medical images classification can be solved by using deep learning models. These models enable the creation of complex networks and deep layers of these models act as feature generators. The deep learned features are directly learned from the input images. The CNN models have proven their success in classifying images of both nonmedical and medical domains [27]. Therefore, there is a need to explore the deep learned features and examine the performances obtained from them. Thus, in the present work deep learned features are extracted from pre-trained CNN model and fine-tuned CNN model.

1.8 NEED OF CLASSIFICATION SYSTEM

The aim of CAD system is to achieve a high classification accuracy for liver classes i.e. normal, chronic, cirrhosis and HCC and maintain a low false positive classification rate. Thus, classification is a key step to improve accuracy and reduce false positive classification rate.

In addition to extraction of representative features from ultrasound liver images, it is also important to have an efficient classifier. Although, a lot of research has been done in classification of liver diseases, but only some of that research has taken the advantage of combining different classifiers [28], [29]. The multiple individual classifiers whose individual predictions are fused through a combining strategy, usually a voting scheme, to classify new patterns constitute an ensemble classifier model. In literature it was observed that multi classifier systems have given better results than individual classifiers [30]. Therefore, there is a need of multi classifier system or ensemble classifier model to accurately classify liver tissue types. In the present work, a majority voting ensemble classifier is designed. It deploys three base classifiers: k-nearest neighbors (k-NN), support vector machine (SVM), and rotation forest (RF).

1.9 OBJECTIVES OF THE PRESENT WORK

In view of the needs as discussed in preceding sections, the proposed work aims to enhance image, extract and select features and classify liver classes such as normal, chronic, cirrhosis and HCC by using ultrasound images, with the following objectives:

i. Development of database of clinical ultrasound B-mode images of normal liver, chronic liver, cirrhosis and HCC.
ii. Study and establish the diagnostically important mathematical features that will differentiate among mentioned liver diseases (taking expert radiologist help).
iii. Assess and implement/modify enhancement algorithm for enhancement of ultrasound images.
iv. Extraction and selection of relevant set of features important for diagnostic purpose.
v. Architect a classifier to classify images into different stages of liver diseases.
vi. Verify and validate the results of enhancement algorithm with the interpretation of expert radiologist(s).

1.10 ORGANIZATION OF THE THESIS

This thesis contains eight chapters and two Appendices. Chapter 1 covers the introduction, need and specific research objectives. Chapter 2 covers the current status of CAD based on liver ultrasound images and the summary of limitations observed from literature survey. The improvements incorporated in the present work to overcome these limitations have been listed in concluding remarks. In chapter 3, the detailed methodology and the development of the liver image database of the present research work is presented. In chapter 4, the detailed procedure of the methods designed for ultrasound image enhancement has been presented. In chapter 5, the methods designed for feature extraction i.e. 'handcrafted features' and 'deep learned features' have been discussed in detail. In chapter 6, the method designed for the classification of liver diseases is described. In chapter 7 all the results and discussions are covered. Finally, Chapter 8 covers the conclusions drawn on this research work and its future scope. The handcrafted features extracted from original and enhanced images, and the sample enhanced images of each liver class, have been provided in Appendix-A and B.

1.11 CONCLUDING REMARKS

A general background about the liver diseases and their characteristic visual appearance on ultrasound images has been provided. The need of (i) computer-aided diagnosis system, (ii) image enhancement, (iii) extraction of features, and (v) classification system has been discussed. Finally, the objectives of the present study have been presented.

Literature Review

2.1 INTRODUCTION

Computer-aided diagnosis system assists radiologist/clinician by carrying out repetitive tasks in the diagnosis and detection of liver diseases. Several mathematical techniques, such as noise filtering methods, contrast enhancement methods, feature extraction, classification methods are being used to develop such systems for the quantitative analysis of liver ultrasound images. It is also possible to represent visual understanding of many radiologists in a single computational system. Quantification of liver diseases using computer-aided systems requires sequential application of various digital image processing steps on liver ultrasound images. Specifically, these image processing steps are image enhancement, segmentation of regions-of-interest, feature extraction and selection, and classification of liver tissue type. In this chapter, review of significant work done in image enhancement, feature extraction and selection and classification with liver ultrasound images is provided.

2.2 IMAGE ENHANCEMENT METHODS

Ultrasound images suffer from low contrast and inherent formation of speckle. Speckle is an unavoidable characteristic of ultrasound images [31]. The CAD system designed by utilizing ultrasound images may get impacted in performance due to speckle noise and low contrast. Thus, several researchers have attempted for the removal of speckle noise. The commonly used adaptive statistic filters based on multiplicative nature of speckle include Lee filter, Kaun filter, Wiener filter, enhanced frost filter and adaptive median filters [32]–[35]. Speckle noise is reduced by these adaptive filters by adjusting the size of filtering window. However, image resolution gets decreased inside the window, which in turn blurs the edges. These low-pass filtering methods are simple and reduces speckle noise in homogeneous area; however, these methods display limitations on preserving sharp textures of images or reducing speckle noise in area near to edges [32].

The anisotropic diffusion (AD) method was developed by Perona and Malik that was based on heat equation and method performs better in homogeneous regions with preserving the edge information in images which are corrupted by additive noise but does not perform well for the multiplicative speckle noise [36]. Yu et al. applied coefficient of variation to AD method and proposed SRAD method [33]. This method limits in retaining textural features due to its smoothing and eliminating nature. The other anisotropic diffusion filters methods are OSRAD, SUSAN_AD and DTD [37]–[39] . Most of these methods perform well in the suppression of speckle noise but result in over-smoothening i.e. important diagnostic details get lost. Mittal et al. modified SRAD by limiting speckle reduction and combining regularization method to preserve the fine-details of ultrasound liver images [40]. However, selection of the parameters is subjective and image enhancement is highly dependent on the value of parameters. Although nonlocal means filters such as PPB, OBNLM and Guo et al. have better speckle noise removing effect, the complexity of these filters is usually very high. Consequently, they cannot meet the real-time need of computer-aided system [41]–[43]. The methods based on wavelet, shearlet, curvelets and bandelets are also used for de-speckling [44]–[48]. These methods perform better on edge preservation; however, artifacts often occur.

The total variation regularization-based image restoration model was first presented by Rudin et al. [49]. This method performs well in preserving edges and image sharpness while removing noise. Aubert et al. gave a non-convex model used for reducing speckle noise by utilizing a maximum a posteriori (MAP) estimator [50]. Authors applied the fidelity term of Aubert model and proposed a convex model of multiplicative noise [51]. Koundal et al. proposed variation method based on Nakagami distribution that reduces the speckle noise in medical ultrasound images [52] . Their method preserved the edge information well however the fine-details of texture were lost in processing. Additionally, these variational methods usually display staircase effect due to the loss of fine-details of texture in ultrasound images.

These methods fail to preserve fine-details of texture because of the fuzziness caused by echo patterns in ultrasound images. It is observed that indeterminacy information in ultrasound images often gets overlooked in the traditional de-speckling methods. This problem is mostly solved by using fuzzy set [53]–[55]. However, the fuzzy set can only handle the membership degree and fails to deal with non-membership or indeterminacy degree of pixels.

Neutrosophic sets (NS) are generalization of fuzzy set and can deal with indeterminant information in images. Neutrosophic sets have been successfully applied in image de-noising applications [56], [57]. The method proposed by Guo et al. was based on NS approach wherein entropy in NS domain was measured to estimate the indetermination. Here γ-median filtering operation was applied to remove noise and decrease indeterminacy [58]. Faraji et al. used NS-based technique to remove noise and enhance facial features [59]. Qi et al. used pixel-wise adaptive neutrosophic filter based on indeterminacy to remove salt-and-pepper noise [60]. Guo et al. applied NS and directional α-mean operation for detection of edges [61]. The method removed the noise effect and detected the edges. Shahin et al. applied neutrosophic similarity score (NSS) scaling for enhancement of pathological images [62]. The NSS was computed under multi-criteria and then utilized to scale the input image. The similarity measurement used by Shahin to measure the similarity between two elements in NS under multi-criteria was given by Guo et al. [63]. In present work, a method is proposed for noise reduction and contrast enhancement in ultrasound liver images is based on NSS.

2.3 REGION-OF-INTEREST SEGMENTATION

The methods available for identification and segmentation of medical images are specific to the imaging modality and body part. Thus, while selecting a segmentation procedure all these aspects are considered. Limitations exist in each modality and so in the ultrasound imaging. Ultrasound images suffer from (i) low brightness contrast, (ii) gray level discontinuity (iii) strong speckle noise and (iv) attenuated artifacts. These factors affect the outcomes of automatic/semi-automatic segmentation with ultrasound images in comparison to images captured by other imaging modalities, e.g., CT or MRI [64]. There are several automatic/semi-automatic algorithms based on region growing approaches, clustering approaches, atlas-based approaches and model-based approaches, but their success rate is not much promising in the cases where we have to exclude many small hepatic vessels from the liver parenchyma [65]. Further, there exist limited studies on the segmentation of two-dimensional ultrasound images with diffuse liver diseases. The standard approach for accurate identification and segmentation of ROIs is considered to be the expert's manual delineation. However, manual segmentation is tedious and take lot of time. In comparison to automatic, semi-automatic and manual segmentation, the interactive segmentation put the user into a main role during the segmentation task. The user interactions, preferably minimal, are desired to correctly segment regions-of-interest [66]. Thus, in present work, the use of an interactive segmentation approach is preferred.

Square-shaped ROIs have been used by lot of researchers because most of the texture methods are based on matrices calculation which can be easily conducted on square matrices [67]. Thus, the square shape of ROI is selected for present work. These ROI are segmented into small regions such that their size should not include blood vessels and artifacts [68]. In order to have a good sampling distribution for estimating reliable statistics, many researchers have demonstrated that size of segmented-ROIs (SROIs) must be at least 800 pixels. Whereas in few research works, a size of 1,000 pixels was also recommended to estimate reliable statistics [67]. The size of SROIs such as 10×10, 25×25, 30×30, 32×32, 40×40, 50×50, 60×60 and 64×64 pixels have been used by different studies for classification of liver diseases [69]–[73]. For extracting 'handcrafted' features from ultrasound images, after discussion with the radiologist, the size of SROI 32×32 pixels was considered suitable for this research work to estimate reliable statistics and to extract maximum SROIs from the database. Whereas, for extracting 'deep learned' features with AlexNet model, SROI size of 227×227 was considered as defined in network [74].

2.4 FEATURE EXTRACTION

Feature extraction is key step in CAD system for characterization of liver tissue type. The echo patterns are represented as the texture of an image and can be exhibited by the texture metrics or mathematical descriptors. These mathematical descriptors are termed as features and play crucial role in differentiating the liver tissue type. The features can be extracted by using traditional hand-designed methods and deep learning methods. The review of feature extraction studies for classifying liver diseases is given below:

2.4.1 Handcrafted Features

Wu et al. used gray-level co-occurrence matrices (GLCM), fourier power spectrum, gray-level difference matrix (GLDM), Laws' texture energy measures (TEM), and multiresolution fractal features for the detection of normal liver, HCC and cirrhosis cases [23]. Authors concluded that among the extracted features, GLCM gave the best performance for differentiating liver diseases. Kyriacou et al. used fractal dimension, GLCM, GLDM, GLRLS and FOS to differentiate among normal, fatty and cirrhosis [75]. Pavlopoulos et al. used texture features namely GLCM, GLDM, fractal dimension, GLRLS and FOS to differentiate between normal and fatty liver cases [76]. Horng et al. extracted texture feature co-occurrence matrix, GLCM, fractal and statistical feature matrix to differentiate three liver classes i.e. normal, cirrhosis and Hepatoma [77]. Lee et al. extracted m-band wavelet and fractal features from normal, cirrhosis and Hepatoma ultrasound images [78]. Vicas et al. presented a CAD system to discriminate between stages of fibrosis using FOS, GLCM, GLDM, multiresolution fractal, Gabor filter, texture edge co-occurrence matrix, Laws' energy measures and phase congruency [79]. Mitrea et al. analyzed higher order GLCM and texture edge co-occurrence matrix for classifying HCC and cirrhotic liver parenchyma on which HCC has evolved [80]. Authors further evaluated third, fifth and seventh-order GLCM for HCC evolved on normal liver and HCC evolved on cirrhosis. Singh et al. developed a CAD system using features from GLCM, Laws' texture energy, fourier power spectrum, GLDM, fractal dimension and statistical feature matrix for normal and fatty liver image classification using ultrasound images [81]. Kalyan et al. extracted features from histogram, GLCM and GLRLM to differentiate between four liver classes [82]. Santos et al. extracted first order statistics, GLCM, GLRLM, gabor filter, Law's filter, fractal, hepatorenal coefficient and attenuation [83].

Balodi et al. extracted FOS, GLCM, GLDM, SFM, Laws TEM, FPS and FDTA and noticed that among these feature GLCM gave highest accuracy in differentiating ultrasound images [84]. Wang et al. extracted GLCM, GLDM and local binary pattern to differentiate normal and cirrhosis ultrasound images [85]. Authors reported that combination of GLCM and GLDM gave the best accuracy and concluded that their combination expresses the texture characteristics very well. Additionally, it is evident from the summary of various research works provided in table 2.1 that the second-order statistics i.e. GLDM and GLCM are used frequently to classify liver tissue types. However, in these research works, the higher-order relationships between the pixels are not captured. Several studies have utilized GLDM method, which is based on local derivatives between pair of gray-levels, to extract texture features for characterization of liver diseases. These local derivatives are able to provide information about the roughness of liver surface and the inhomogeneous echotexture. Hence, if the objective is the characterization of different stages of liver disease by using GLDM method, it is necessary to compute the local derivatives as accurately approximated as possible. Mostly in literature the implementation of GLDM method is based on the absolute differences between two neighboring pixels. However, the use of pixel pair only, in calculation of local derivatives, may not be able to provide the best approximation of calculated features from GLDM method and therefore proper characterization of different stages may suffer. Involvement of more neighborhood pixels than the traditional GLDM method will increase the accuracy of calculated local derivatives; this inference is based on the fact that the increase in order of numerical differentiation reduces the error of derivative approximation. Hence, the present work is carried out with the objective to calculate derivatives with pixels greater than two and analyze their effectiveness in the characterization of different stages of liver disease.

Table 2.1 The summary of previous studies carried out on the computer aided diagnosis systems using ultrasound liver images

Authors	Liver image type (No. of patients)	Feature Extraction Methods	Feature Selection Strategy	Classifier	Classification Accuracy (%)
Wu et al. [23]	Nor (15), Cirr (15), HCC (15)	GLCM, GLDM, Laws' energy, FPS, multiresolution fractal	Hotelling trace criterion	Bayes	GLCM and GLDM: 84.4 Fractal:88
Kyriacou et al. [75]	Nor, Fatty, Cirr	GLCM, GLDM, GLRLS, FOS, fractal	-	k-NN	82.2
Pavlopoulos et al. [76]	Nor, Fatty, Cirr	FOS, GLDM, GLCM, fractal	-	Fuzzy NN	82.7
Horng et al.[86]	Nor (30), Cirr (30), Hepatoma (30)	TFCM, GLCM, fractal, statistical feature matrix		Maximum likelihood	86.7
Lee et al.[87]	Nor (50), Cirr (50), Hepatoma (50)	M-band WT, fractal		Bayes, ANN	Bayes: 90 ANN: 92
Horng et al.[77]	Nor (13), Chronic Hepatitis (20), Cirr (7)	GLCM, TFCM, CM-score	Correlation	Radial function Network	92.5
Huang et al.[88]	Nor, Fatty	FOS, GLDM, GLCM		PNN	87.2
Vicas et al.[79]	Chronic Hepatitis C	FOS, GLCM, GLDM multiresolution fractal, gabor filter, TFCM, Law's energy, EOCM, phase congruency	-	Logistic regression	69.5
Mitrea et al. [80]	Cirr with HCC, HCC	Mean, 3rd,5th , 7th order GLCM, edge-based statistics, WT, Laws' energy	Correlation	SVM, ANN, Bagging+RF, Bagging+SVM, AdaBoost+RF	Bagging + RF : 81.8
Andrade et al.[89]	Nor, Fatty	FOS, GLCM, GLRLM, fractal, Law's filter	Stepwise regression	ANN, SVM, kNN	ANN: 76.9 SVM: 79.7 k-NN: 74.0
Wu et al.[90]	Nor (90), Cirr (176), HCC (166)	GLCM, WT, Gabor WT	Genetic algorithm	k-NN, fuzzy k-NN, PNN, SVM	96.6

Lee [28]	Nor (90), Cirr (176), Hepatoma (166)	M-band Wavelet transform, Gabor wavelet	-	Minimum distance, FKNN, PNN, SVM, ensemble	Ensemble:95.69
Ribeiro et al.[91]	Nor, Chronic hepatitis (30), Compensated Cirr (34), Decompensated Cirr (36)	GLCM, WT, Autoregressive Model Laboratory and Clinical Features	-	SVM, k-NN	SVM: 76.14 k-NN: 80.68
Virmani et al.[92]	Nor (15), Cirr (16), HCC (25)	WPT	Genetic algorithm	SVM	88.8
Singh et al.[81]	Nor (80), Fatty (100)	GLCM, GLDM, FOS, FPS, statistical matrix, Law's filter, fractal dimension	Correlation	Linear classifier	95.0
Kalyan et al.[82]	Nor (30), Fatty (10), Cirr (10), Heptomegaly (10)	Intensity histogram, GLCM, GLRLM, Invariant moments	Random search, genetic search	BPNN	92.5
Santos et al.[83]	Nor (68), Fatty (52)	FOS, GLCM, GLRLM, Gabor filter, Law's filter, fractal dimension, lacunarity, hepatorenal coefficient, attenuation	Stepwise regression	ANN, SVM, kNN, Bayes, Decision Tree	79.0
Wang et al.[85]	Nor (80) Cirr (80)	GLCM, GLDM, LBP		BPANN	94
Krishnan et al.[93]	Cirr (65), Fatty (41), Hepatitis (10)	WT, GLRLM		RF	91

Nor: Normal, Cirr: Cirrhosis, HCC: Hepatocellular carcinoma, GLDM: Gray-level difference matrix, FOS: First order statistics, FPS: Fourier power spectrum, GLRLM: Gray-level run length matrix, WPT: Wavelet packet transform, WT: Wavelet transform, GLCM: Gray-level co-occurrence matrix, TFCM: Texture feature co-occurrence matrix, CM: Computer morphometry, EOCM: Edge orientation co-occurrence matrix, LBP: Local binary pattern, SFS: Sequential forward selection, k-NN: k- Nearest neighbor, ANN: Artificial Neural Network, Fuzzy NN: Fuzzy Neural Network, PNN: Probabilistic Neural Network, BPNN: Back propagation neural network, RF: Random forest, BPANN: BP_Adaboost neural network, SVM: Support vector machine

Further, the adjustable parameters of the ultrasonic device generate an incoherent texture features on original liver images and lead to variant texture analysis of chronic liver. Ranklets, the invariant texture analysis method, have been recently introduced and applied for face detection, synthetic-aperture radar image classification and mass classification in mammograms. The texture analysis using ranklet transform are less sensitive to different ultrasonic devices and fairly adopted for designing a robust CAD system [94]. Masotti et al. introduced the concept of ranklet transform for texture classification [24]. Invariance to linear/non-linear monotonic gray-scale transformations were obtained by ranklet transform (ranklets) where analysis was based on relative rank of pixels rather than on their gray values. Bianconi extended ranklets to colour images by considering inter-and intra- channel ranklet features[95]. Yang et al. extracted gray-scale invariant features by using ranklet transform to diagnose breast tumor on ultrasound images[96]. Authors compared these extracted features with wavelet transform and concluded that diagnostic performance of texture analysis via ranklets obtained best and consistent performance. Cai et al. extracted ranklets, GLCM and LBP features and observed that accuracy for differentiating malignant and benign tumors in breast ultrasound images was in order of ranklets, GLCM and LBP[97]. Thus, in the present research work, texture features are extracted from multi-resolution ranklet transformed images and by using GLDM and GLCM methods.

Selection of Handcrafted features

The combined application of all features that are obtained from GLCM, ranklets and extended GLDM may increase computation complexity and at the same time, relevancy of features is not known in advance. The feature space may also contain redundant or irrelevant features that have limited ability to discriminate [98], [99]. This emphasizes the need to adopt an important step, i.e., feature selection to reduce the computation complexity, data understanding and improve learning performance of classifiers. Table 2.2 provides the advantages and disadvanatges of feature selection methods. Mitrea et al. used correlation feature selection (CFS) to find relevant texture set for differentiating HCC evolved on normal and cirrhotic liver images [80]. CFS is simple filter method that assign ranks to features but it is not capable to identify strong interactions [100]. Chi-square is a filter method which is computationally efficient but do not show good results with non-linear database. Wu et al. used ReliefF as feature selection method and found an increase in accuracy [101]. The ReliefF is multivariate filter method that can handle multiclass problems and is capable of dealing with noisy data. The Relieff method includes interactions among features and can capture local dependencies [102]. Lu et al. used ReliefF to rank features for breast CAD system and found that this method helped in reduction of redundant and irrelevant features [103].

The stochastic methods such as PSO and GA were used by Kalyan and Virmani to select an optimal feature set [82], [92]. Although these methods efficiently capture feature redundancy and interaction; are computationally expensive. The most commonly used greedy wrapper methods such as sequential forward selection (SFS) and sequential backward selection (SBS) have less computational time. The SFS is simple method and avoid overfitting but the selected feature set will be specific to classifier under consideration. Santos et al. used SVM, ANN and k-NN for normal and fatty liver image classification [83]. The results showed that after applying SFS, the classification accuracy improved by 12.81%, 14.38% and 0.25% for ANN, SVM and k-NN respectively. Pudil et al. have suggested a floating search methods i.e. sequential forward/ backward floating search (SFBS, SFFS) that are based on greedy search with backtracking [104]. But, a recent study has demonstrated that SFFS is not superior to SFS and in case of feature set having more than hundred features then SFBS is not feasible for selecting features [105].

Some authors have proposed hybrid methods by taking the benefits of both filter and wrapper methods[106], such methods are CFS and GA, mutual information and GA, and chi-square method and optimization method [107]–[109]. Zhang et al. introduced a hybrid feature selection (HFS) method by combining Relieff and minimum redundancy maximum relevance [110]. Authors observed that HFS method improved the effectiveness of gene selection, and as a result, provided better discrimination. The rationale behind implementation of HFS method is that, first filter method is used to select a candidate feature set after that wrapper method is used to find the optimal feature subset from this selected feature set. Further, the choice of feature selection method depends on data characteristics i.e. data type, size and noise. In this work, liver ultrasound dataset is considered which do not show linear characteristics, overlapping of disease characteristic leads to noise and there are four liver classes. To handle all these characteristics and obtain an optimal feature set hybrid feature selection using ReliefF and SFS was proposed. The ReliefF filter method best handled the data characteristics, i.e. data with multiple classes, noise and non-linear and SFS provide optimal set with best classification accuracy. Here, features are ranked, by assigning them weights, with ReliefF method. Then, pre-selection is done by considering features which have weights (given by ReliefF method) above the threshold and from these pre-selected features an optimal subset is selected by SFS.

The features extracted from the same pattern by different feature extraction methods may contain complementary information. In several pattern recognition systems, multiple feature sets are used to produce new fused feature set, this can be defined as feature fusion [111]. The commonly used feature fusion schemes are serial feature combination and serial feature fusion [112]. The serial feature combination is integration of multiple feature sets. The serial feature fusion is selecting relevant features from the serial feature combination set. Evolutionary-based hierarchical feature fusion was presented by Wu et al. to classify normal, cirrhosis and HCC liver images and obtained 96.62% classification accuracy [90]. Singh et al. fused the best features from different feature domains along with their Fisher's Discrimination Ratio weights and 'weighted z-score' [81]. Chandrashekar et al. have given an overview of feature fusion schemes for pattern classification [113]. It is evident from the literature that the dimension of feature space impacts the performance of a classifier. It also increases the training time of a classifier and risk of overfitting [26]. The instance-based classifiers, such as, k-NN are very susceptible to irrelevant features. The performance of SVM classifier may also suffer with high-dimensional feature space where many features are irrelevant [26]. Hence, efficient feature selection and fusion schemes can improve the performance of classifiers and consequently the computer-aided system that have to be designed.

2.4.3 Deep Learned Features

Deep learning is the form of machine learning wherein a model directly learns from text, sound, or images to perform classification tasks. The training of deep learning models is usually done by using large sets of labelled data and features are learnt directly from the data without the requirement of ‘handcrafted’ or manual feature extraction [114]. It learns multiple levels of representation and abstraction. Currently, in computer vision, CNNs are one of the best performing deep feature learning methods for various tasks [115]–[117]. Their success in computer vision has broadly encouraged investigations in medical domain, resulting in several research papers in recent years, this proves the effectiveness of CNNs for solving variety of problems in medical imaging domain [118]–[122].

Litijens et al. summarized deep learning methods applied for solving medical imaging problems [120]. Based on the literature the transfer of knowledge from natural images to medical imaging domain can be divided into 2 major approaches. First approach comprises of researches in which a pre-trained CNN is used to generate features. Here, the outputs of certain layers, that is deep learned features, are extracted and utilized to train a classifier to solve the problem in hand [123]. The second approach comprises of researches wherein a pre-trained CNN model is fine-tuned [124]–[126]. This can be done by shallow fine-tuning or deep tuning. In shallow fine-tuning only last layer of pre-trained CNN model is trained. This type of tuning is usually used in case of limited data however the desired performance is not obtained. Whereas in deep tuning most of the convolutional layers in a pre-trained CNN model are trained. This type of tuning may well adapt the pre-trained CNN model to the required task but may need large imaging data.

AlexNet proposed by Krizhevsky et al. is a CNN architecture, trained on the massive ImageNet database to address the problem of image classification [74]. It was presented that trained AlexNet model can be utilized as feature extractor (without any amendment) to successfully perform numerous tasks, including domain adaptation and data classification. Yu et al. demonstrated that a deep learning-based algorithm using pre-trained AlexNet CNN can classify liver fibrosis stages [127]. Girshick et al. used a pre-trained CNN model for initialization and then model was fine-tuned to obtain a new model for object recognition task [128]. Authors also showed that fine-tuning of pre-trained CNN model on the new labelled target data can boost the performance. Particularly, authors fine-tuned AlexNet model for semantic segmentation and obtained results were better. Several researches in the remote sensing domain, also observed the advantages of fine-tuning pre-trained CNN model [129], [130]. Boufenar et al. used deep learned features by using both approaches: pre-trained AlexNet and fine-tuned AlexNet to recognize Arabic characters and observed that fine-tuned model gave higher accuracy compared to pre-trained model [131]. Ashnil et al. used an ensemble of CNN as feature extractors that learned image features for classifying medical images [132]. Kooi et al. discriminated cysts from solid lesion in mammogram images using deep learned features from a deep CNN trained model [133]. Zhang et al. combined deep and handcrafted features to classify medical images [134]. Antropova et al. extracted features using pretrained VGGNet to diagnose breast cancer[135]. Huynh et al. extracted features from pre-trained AlexNet model to distinguish breast lesions[123].

In this research, feature extraction is done with both approaches that is by utilizing pre-trained AlexNet and fine-tuning AlexNet to classify four liver diseases. Here, AlexNet architecture is utilized for transfer learning as this model is simple and one of the currently used CNN architecture. In comparison to other complex and deep CNN architectures i.e.

GoogLeNet, VGG-16, Alexnet is a simple CNN architecture which is easy to train and optimize.

2.5 CLASSIFICATION

Classifier plays an important role in designing computer-aided classification system for disease diagnosis. After the extraction and selection of features, these features are given as input to train the classifier for four liver classes. The primary aim of classifier is to assign one of the possible classes to the new data with minimal rate of misclassification. As the classification of disease classes becomes complex for simple linear systems, more efficient classifiers are required. The solution to such classification problem is by using classification approaches such as k-NN, SVM, ANN, decision tree (DT), linear discriminant analysis (LDA) and Bayesian networks [136], [137]. A classifier builds the separation boundaries among different classes by learning from examples. Supervised learning is that where machine is trained by using the 'labeled' data. There are several published researches on classification methods related to the classification of images. Out of wide range of classifiers, which one would perform better normally remains a question to the researchers in the process to design an efficient classification system. There are number of applications of k-NN, NN, SVM and Bayes for the evaluation of their efficiency in real diagnostic applications. Among all the classification methods, k-NN and SVM have received lots of attention due to their demonstrated performance. The k-NN classification method has been employed by several researchers to characterize soft tissues and it has reported that it outperforms Bayes and decision tree.

The k-NN has gained wide acceptance as it is simple classifier to comprehend and easy to implement as compared to neural networks. This classifier does not consider any assumptions about data before implementation. The classifier adapts easily when new training data is given. Much of the classifiers are easily implemented for binary classes task and require effort to implement them for tasks having classes more than two whereas k-NN adapt to multi-class without any extra effort.

In addition to k-NN, SVM has also emerged as a powerful tool for classification. SVM proposed by Vapnik et al. is a pattern classification method [138]. It is supervised learning model. and one of the most successful machine learning models. SVM works by constructing hyperplanes in a multidimensional space that separates instances of different class labels [139]. In SVM, initially input is mapped into a higher-dimensional feature space by using a kernel function such as radial basis function (RBF), sigmoid and polynomial function. RBF function is most widely used and it is known to perform well on a large variety of problems. This kernel effectively maps the data onto an infinite-dimensional feature space [140]. Pal et al. highlighted SVM accuracy on the dimensionality of the data set and suggested to undertake a feature selection analysis prior to classification [141].

The published studies for classification of liver ultrasound images are described next. Pavlopoulos et al. developed a technique to classify normal, fatty, and cirrhosis liver images with five-layered feed forward neural network, and 82.67% accuracy was obtained [76]. Horng et al. categorized normal, hepatitis, and cirrhotic liver and an accuracy of 86.7% was obtained by using Maximum likelihood classifier [77]. Lee et al. differentiated normal, cirrhosis and hepatoma and accuracy of statistic classifiers was 90.7% and with ANN was 92% [78]. Lee et al. proposed an ensemble-based data fusion method for characterizing normal, cirrhosis, and hepatoma [87]. The algorithm selected adequate classifier with high recognition rate and diversity. An accuracy of 95.67% was achieved. Cao et al. classified fibrosis and normal liver images by using Fisher linear discriminant and SVM, they obtained accuracies of 85.2% and 84.4%, respectively [142]. Cao et al. distinguished among normal, fibrosis, and cirrhosis of liver and obtained an average accuracy of 92.62% for normal liver, 86.6% for fibrosis, and 95% for cirrhosis [143].

Horng et al. developed ultrasonic scoring system for assessing chronic liver disease by analyzing the characteristics of liver echotexture and an accuracy of 92.5% was attained [77]. Mitrea et al. classified HCC and cirrhotic liver parenchyma on which HCC had evolved with combination of Bagging and random forest classifiers and reported a classification accuracy of 81.78% for HCC and cirrhosis on which HCC had evolved [29]. Wu et al. proposed two-stage feature fusion method to classify ultrasonic images of liver tissue in three classes: normal, cirrhosis, and hepatitis [144]. Wu et al. presented an evolution-based hierarchical feature fusion method to classify hepatoma, cirrhotic, and normal liver images [90]. Moldovanu et al. classified normal and fatty liver images with ANN classifier and classification rate of 94%, sensitivity of 96%, and specificity of 92% were obtained [145]. Singh et al. classified normal and fatty liver and an accuracy of 92% and a sensitivity of 100% were reported [146]. Minhas et al. presented a technique based on multiscale capability of WPT to detect fatty liver and heterogeneous liver by exploring features, such as echogenicity, granularity, and homogeneity of ultrasound liver images [147]. A classification accuracy of 95% was achieved by SVM. Kalyan et al. developed a technique to differentiate among fatty, cirrhotic, and hepatomegaly liver with MLP and an accuracy of 90% was obtained [82].

Virmani et al. developed a CAD system for characterizing normal, cirrhosis, and HCC evolved on cirrhotic liver, based on multiresolution texture descriptors [92]. GA-SVM was used for feature selection, and an accuracy of 88.8% was obtained. Riberio et al. built a hierarchical method, by integrating ultrasound images, laboratory tests, and clinical records, to categorize chronic liver patients [91]. Accuracies of 98.67% for normal liver, 87.45% for chronic hepatitis, and 95.71% for cirrhosis were obtained. Although, a lot of research has been done in classification of chronic liver, the developed systems may not gain from the combination of different classifiers. In medical applications the main concern is to build a model which has better performance. In recent years, to optimize the performance of CAD systems multiple-classifier system or ensemble classifier has been used [30], [148], [149]. The multiple individual classifiers whose individual predictions are fused through a combining strategy, usually a voting scheme, to classify new patterns constitute an ensemble classifier model. [150]–[153]. Wen-Li Lee applied ensemble classifier model for characterization of ultrasound liver images and observed that combining multiple classifiers with different features gave better results than the individual classifiers [28]. The early diagnosis of melanoma was done by ensemble classifier, comprised of k-NN, LDA and DT, and majority voting as a combiner [30].

The k-NN, SVM and Adaboost have been widely used for chronic liver disease classification [154]. Various factors, such as parameter settings, extracted features or feature combinations, and the quality of the experimental samples, influence classification performance. The correct classification of normal and cirrhosis liver images can be easily achieved, but difficulty lies in correctly differentiating four stages of liver i.e normal, chronic, cirrhosis and HCC evolved over cirrhosis. Generally, ensemble models can be used for improving classification performance. Earlier diagnosis of chronic liver diseases has been facilitated by ensemble model that is combining bagging and adaboost with SVM and random forest, with a voting scheme [29]. Xie et al. classified melanocytic tumours by using ensemble model composed of three classifiers having different network structures/types of ANN [155]. Catal et al. provided sentiment classification model based on vote ensemble model which utilized naïve bayes, SVM and bagging with SVM as base classifier [150]. Liu et al. used rotation forest (RF) for classification of microarray dataset and found RF outperformed Bagging and Boosting [156]. Akin et al. also found RF as effective method for Erythemato-squamous diseases diagnosis [157]. Authors used RF ensemble with different base classifiers i.e. multi layer perceptron, k-NN and decision tree. Among these base classifiers used with RF, highest classification accuracy of 97% was obtained with RF with k-NN for Parkinson disease diagnosis. It was observed from above studies and table 2.1 that SVM, k-NN and RF has been successfully used in medical applications and as ensemble classifiers. In this research work, the proposed classifier is described as a majority voting ensemble classifier that deploys different classifiers: SVM, k-NN and rotation forest. This designed ensemble model can be regarded as a combination of different strategies which enhances the generalization ability of the designed model as each component in the model learn about some part of classification problem.

2.6 CONCLUDING REMARKS

Based on the exhaustive literature survey, the following conclusions have been drawn.

- Speckle reduction methods for ultrasound image enhancement are studied by number of researchers. Mostly the researches were based on reducing the speckle noise in ultrasound images and were able to preserve edge information well. The methods fail to preserve very fine-details of texture due to the presence of fuzziness caused by echo patterns in ultrasound images. It was observed that indeterminacy information in ultrasound images is not considered in most of the traditional de-speckling methods. The NS, which are generalization of fuzzy set and can deal with indeterminant information in images. Thus, to narrow the gap image enhancement method is designed for liver ultrasound images based on NS.
- There is considerable literature on feature extraction methods both handcrafted and deep learned. It is observed that the handcrafted methods based on texture are used for the present classification study. Among these handcrafted methods, GLDM method is based on local derivatives between pair of gray-levels. These local derivatives are able to provide information about the roughness of liver surface, and the inhomogeneous echotexture. Involvement of more neighborhood pixels than the traditional GLDM method will increase the accuracy of calculated local derivatives. Hence, this gap is

taken care of and work is carried out with the objective to calculate derivatives with pixels greater than two and analyze their effectiveness in the characterization of different stages of liver disease.

- Feature selection studies based on filter, wrapper and hybrid methods are reviewed. Each approach has its advantages and limitations. In the present work, by taking the advantage of filter and wrapper method a hybrid feature selection method is designed to obtain an optimal set to reduce computation time and complexity.
- The deep learned features are also reviewed. Deep learned features can be aimed at new applications and can quickly capture new effective characteristic representation from training data. The hidden layers inside the deep neural network architecture automatically extract suitable features without user intervention to give the best texture discriminating performances in multi-class texture classification problem. Thus, in this research work, feature extraction is also done both by utilizing pre-trained AlexNet and fine-tuning AlexNet to classify four liver diseases.
- Literature is reviewed for classification of liver ultrasound images. Most of the work has not benefitted by ensemble of classifiers. The use of different types of classifier can add diversity as each classifier contains an explicit or implicit bias that leads it to favour certain generalization over others. Thus, in this research work, the designed classifier is described as a majority voting ensemble classifier that deploys different classifiers: SVM, k-NN and rotation forest.

Methodology

3.1 INTRODUCTION

Research methodology is the fundamental process that presents the methods utilized for addressing the research problem and the reasons behind their selection[158]. Generally, the methods adopted by different studies may vary with respect to the objectives of the study. The Chapter 2 highlights the significance of various steps of CAD system, such as image enhancement, feature extraction and selection, and classification, in achieving a high level of efficiency for automated diagnosis. In the wake of these inputs, this chapter provides the research methodology adopted by the present research work. It is assumed that these inputs will provide both depth and breadth for the research and shape the design of the research work. This chapter is arranged as follows: the section 3.2 provides a detailed description of the methodology adopted by the present research work, which includes a detailed schematic representation of the work and the methods used in each phase. Section 3.3 includes the description of medical ethics. The database used in the present study is detailed in Section 3.4. The narration regarding the segmented-regions-of-interest (SROIs) is given in Section 3.5 and the conclusion of this chapter is presented in Section 3.6.

3.2 FLOW OF RESEARCH

A detailed description of the flow of research is provided in this section, which includes the schematic representation, the methods adopted and the description of studies. The flow of research to achieve the objectives of each study is depicted in Figure 3.1 and the description of the study is detailed in the subsequent section.

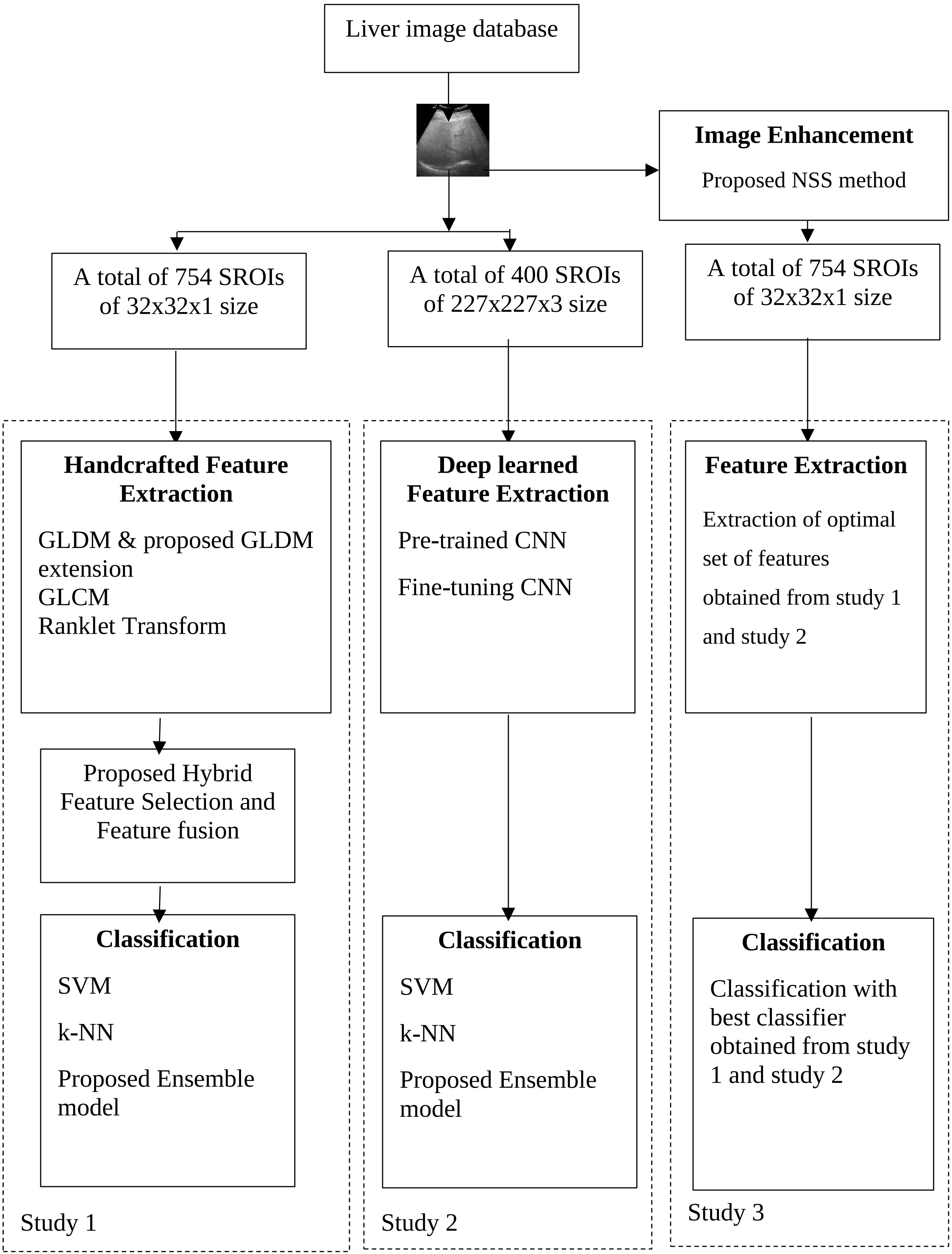


Figure 3.1 Flow diagram of research methodology

a) A clinical database of B-mode ultrasound images is acquired in a duration of one year to carry out the experiments. It is acquired from Manipal hospital, one of the best hospitals of south India and consists of one hundred eighty-nine ultrasound images of both normal and abnormal liver cases.

b) The regions-of-interest (ROIs) are normal or abnormal liver parenchyma in liver ultrasound images. The marked ROIs by radiologist on ultrasound images are segmented into small square shaped regions of fixed size. These segmented regions are defined as segmented-regions-of-interest (SROIs). These SROIs contains most diagnostic information in textural form of the specific liver tissue class. Extraction of features from these SROIs reduce the complexity and computation time as compared to extracting features from bigger region of liver image.

c) The feature extraction is key step in CAD system and characterization of liver diseases. Features are extracted from SROIs and these features are given as input to classification step. There are two ways of extracting features i.e. ‘handcrafted’ features and ‘deep learned’ features. Handcrafted feature extraction methods that is, GLDM, GLCM, and ranklet transform (*ranklets*) are used by keeping application in mind whereas, deep learned features are directly learned from images by using deep neural networks.

d) Three handcrafted feature extraction methods (i) GLDM, (ii) GLCM and, (iii) *ranklets* are used to extract features from SROIs. In this study, traditional GLDM method was extended by use of greater number of neighboring pixels in order to extract better approximations of features. In total a set of 148 features was extracted with GLDM and its extension, GLCM and *ranklets*. However, this feature space (148 feature set) may increase computation complexity as well as time. Selection of features was done to reduce computational complexity and improve performance.

In this work, HFS method was proposed by combining filter (ReliefF) and wrapper (sequential forward selection [SFS]) methods. In the proposed method two step feature reduction takes place. Firstly, ReliefF method rank features and pre-selection is done by discarding low ranked features. Secondly, from this selected features SFS method finds the optimal set of features. Advantage of proposed method is the use of (i) filter method which rapidly reduces the effective number of features under consideration, and (ii) wrapper method which avoid overfitting as they use cross-validation measures of predictive accuracy. Thereafter, to take advantage of complementary information from different feature sets, feature fusion schemes (that are: serial feature combination, serial feature fusion and, hierarchical feature fusion) are also implemented.

e) The 'deeply learned' feature extraction approaches of CNNs employed in medical image classification are described below.

- *Pre-trained CNN*: In this approach, learned features are extracted from a pretrained network, and these features are used to train a classifier, such as SVM, k-NN, and so on. In this research, a pretrained network AlexNet is used which is trained on ImageNet database. Features are extracted from different layers of AlexNet and used for classification.
- *Fine-tune a CNN*: In this approach, features learned using a large database are transferred to the new network and only the classifier part is trained with the new database. In this research a CNN architecture, AlexNet is pre-trained on the ImageNet dataset. This pre-trained CNN is then adapted to the current research problem by replacing the last three fully connected layer with a new fully connected layer for four classes.

f) The features extracted are used in classification of liver ultrasound images into four categories i.e. normal, chronic, cirrhosis and HCC evolved over cirrhosis. For rest of study

HCC evolved over cirrhosis will be written as HCC. In this research, ensemble classifier is proposed by using different base classifiers, that is k-NN, SVM and rotation forest, and combiner, that is majority voting. As the main concern of present research was to build a model which has better performance compared with the individual classifiers, ensemble model was selected. The designed ensemble model can be regarded as a combination of different strategies which enhances the generalization ability of designed model as each base classifier in the model learn about some part of classification problem.

g) A brief discussion of three studies: Study 1, Study 2, and Study 3 carried out in this research work.

- Study 1: Computer-aided diagnostic system is designed by using handcrafted features extracted from original images with proposed ensemble model. Hybrid feature selection along with feature fusion schemes are used to get the optimal feature set.
- Study 2: Computer-aided diagnostic system is designed by using deep learned features extracted from original images. The performance of three classifiers i.e k-NN, SVM and proposed ensemble model results are then compared.
- Study 3: Computer-aided diagnostic system is designed with the best architecture out of study 1 and study 2 using enhanced images. Original images were enhanced by proposed image enhancement method. This method was proposed with an intent to improve the visual details of liver ultrasound images in order to provide accurate diagnosis. The reduction of noise and enhancement of contrast were considered.

3.3 MEDICAL ETHICS

In medicine, *Medical Ethics* is the study of moral values and rulings that need to be followed in medical field. The ethical guidelines revolve around the mutual trust between radiologist and patient. Radiologists are obliged to maintain confidentiality about the patients' revelations, their own observations, the diagnosis related to patients' health and privacy, and the contents of their medical records. Moreover, medical confidentiality prevents the sharing of patient information to third parties without their prior consent. Radiologists are not supposed to collaborate on any health database if the preservation of the confidentiality of the information stored in it is not guaranteed. Radiologists must ensure that a patient's identity is not revealed in any presentation of medical documentation in any format. However, the presentation of medical data, which is gathered for teaching or research purposes, is permitted only after procuring explicit authorization or after preserving anonymity. There are numerous methods to modify the data for concealing the patients' identities. Some of these methods are coded information, linked anonymized data, and unlinked anonymized data.

This research is related to the humans and their liver ultrasound images. So, for this research authors were required to obtain ethical clearness from hospital ethics committee of the associated hospital that is Manipal Hospital, Bangalore. After going through the application of the authors, examining the research problem presentation and proposal, and enquiring about some relevant issues, the ethics committee of Manipal Hospital, Bangalore permitted the authors to pursue the research. All the images received from hospital were unlinked anonymized.

3.4 SUBJECTS AND ULTRASOUND LIVER IMAGE ACQUISITION

The research was approved by the ethical committee under the supervision of Head of Department of Radiodiagnosis, Manipal Hospital, Bangalore, Karnataka, India. Subjects are patients (men and women) who underwent a medical examination during the period of March 2013 to August 2014. In this period, database was collected from 94 patients among which 67 were male (age range: 21-70) and 27 were female (age range: 23-61). The ultrasound images were obtained with a GE LOGIQ E9 (General Electric Medical System, Milwaukee, Wisconsin). Certain etiquettes were followed during image acquisition which includes (a) fasting for four to five hours prior to examination, (b) image capturing via longitudinal scan/subcostal transverse scan/intercostal scan depend on varied patient conditions, (c) imaging with 3 to 5 MHz broadband curvilinear probe with extended width beam, (d) finer adjustments of machine settings such as focus and time-gain-compensation was done according to body habitus, and (e) image acquisition was done of right lobe through segments 7 and 8 or segments 5 and 6. A total of 189 ultrasound images of liver were collected from 94 patients. These 189 images comprise of 48 normal liver, 50 chronic liver, 50 cirrhosis, and 41 HCC. The ROIs were defined by the experienced radiologist.

3.4.1 Liver Image Collection Criteria

The collection criteria for liver images are as follows:

- The normal liver ultrasound images were collected from two types of patients (i) the patients who arrived for routine medical examination, and had normal liver in ultrasound scan and normal liver function tests; and (ii) the patients who had complaints like upper abdomen pain or uncomplicated gallbladder or kidney stones but a normal liver in ultrasound scan.
- The chronic liver ultrasound images were collected from four types of patients (i) the patients who were found positive for Hepatitis B or C infection, (ii) the patients who were known alcoholic, (iii) the patients who were taking certain drugs for a prolonged period, and (iv) in rare cases, the patients who came for check-up and were unaware of disease and were diagnosed with chronic liver.
- The cirrhosis and HCC images were collected from patients who had advanced chronic liver and cirrhosis, respectively.

3.4.2 Liver Disease Assessment Criteria

The decision regarding diagnostic quality and representativeness of each image class was made visually by a domain expert, having more than 20 years of experience in abdominal ultrasound imaging. In addition to visual interpretation of ultrasound image, clinical history of patients and affirmation of disease computed by CT/MRI/biopsy were also used as assessment criteria.

3.4.3 Database of Normal Liver and Abnormal Liver Images

The digital image database of 189 B-mode ultrasound images were taken from 94 subjects and considered for this research. The database encompasses 48 normal liver and 141 abnormal liver images. The abnormal liver images comprised of 50 chronic liver, 50 cirrhosis, and 41HCC cases. There were 67 males (12 normal, 17 chronic, 20 cirrhosis, 18 HCC) and 27 females (12 normal, 8 chronic, 5 cirrhosis, 2 HCC). Table 3.1 presents the information about patient’s age and gender for each liver class in the database.

Table 3.1 A brief information about patients

Classes	Age Group (yrs)	Male	Female	Total
Normal	21-50	12	12	24
Chronic liver	35 – 60	17	8	25
Cirrhosis	50 – 70	20	5	25
HCC	50 – 70	18	2	20

3.5 SEGMENTED REGIONS-OF-INTEREST

The region of ultrasound image that provide diagnostically relevant information is ROI. The ROI is further subjected to segmentation, where the images are demarked into fixed size squared shape regions also called SROIs to attain a comprehensive database for next step.

Researchers generally use manual, semi-automatic, fully automatic, and interactive segmentation approaches. The ultrasound images usually suffer from (i) low brightness contrast, (ii) gray level discontinuity, (iii) strong speckle noise, and (iv) attenuated artifacts, which may affect the outcomes of automatic or semi-automatic segmentation. Besides, SROIs should also be free from the portions of hepatic vessels, artifacts, and other nearby structures. This necessitates the requirement of an expert interaction and hence, the present research used the interactive segmentation where an expert radiologist who had an experience of more than 20 years in liver ultrasound imaging helped in marking the ROI seed point. The application of interactive segmentation helped the study to gather clear and specific image data.

3.5.1 Selection of the Size of Segmented Regions-of-Interest

Segmented-regions-of-interest contains the region of image with the most diagnostic information and makes it a good 'representative' of the image. Additionally, SROIs reduce the computational time for feature extraction. Thus, the selection of an appropriate size of SROIs is a crucial process and briefed in chapter 2. It was observed that the images with 75×75 pixels and 100×100 pixels were not suitable as they included blood vessels [67]. Further, SROIs of 10×10 pixels and 25×25 pixels were also not advisable, as they were too small to have sufficient information for a reliable statistical analysis, which can signpost the diffused liver disease (i.e. where whole liver gets infected). This reduced the choice to 64×64 pixels and 32×32 pixels. However, it was observed that both 32×32 pixel SROIs and 64×64 pixel SROIs contained similar information. Thus, 32×32 pixel non-overlapping SROIs were used to build a large dataset.

Additionally, 227×227 pixel SROIs were used in deep learning as the input image size was 227×227x3 pixels as per the requirement of architecture that is Alexnet. Here, all the grayscale images were converted to RGB images by replicating single channel to obtain a 3-channel RGB image.

3.5.2 Datasets

The collection of SROIs is defined as dataset. In this research work, three datasets were formed to optimize the design of CAD system. First dataset consists of 754 SROIs of size 32x32 pixels and is termed as original1 dataset. The resulting 754 SROIs comprise of 192 normal, 200 chronic, 200 cirrhosis, and 162 HCC. Second dataset consists of 400 SROIs of size 227 x 227 pixels and is termed as original2 dataset. The resulting 400 SROIs comprise of 100 SROIs of each liver class. Third dataset consists of ultrasound images on which enhancement was performed. This dataset is termed as enhanced dataset and has 754 SROIs of size 32 x 32 pixels. The resulting 754 SROIs comprise of 192 normal, 200 chronic, 200 cirrhosis, and 162 HCC. These enhanced SROIs were segmented from the same location of image as that of original1 SROIs.

The performance of classifiers was evaluated by the repeated holdout method i.e. 15 times 10- hold out method was employed. The holdout method uses one-third for testing and two-third for training. An overall error rate is calculated by averaging the error rates on different iterations. The major advantage of this methodology is that the reliability of the estimated holdout is improved, as the process is repeated with different sub-samples.

3.6 CONCLUDING REMARKS

The research methodology used in present work to achieve the objectives is presented in this chapter. The chapter begins with a brief introduction before moving to the first main topic of research methodology. The section describes the flow of study with the help of a flow diagram. Further, the methods used in each phase of the study are briefed. The next section details the steps taken to ensure that medical ethics were preserved in the research work. The details of ultrasound liver image database, which include the protocol for acquiring ultrasound images, image collection criteria, and disease assessment criteria, are documented in fourth section. Final section presents the detail about SROIs and datasets developed for this research and their distribution among liver classes.

Ultrasound Image Enhancement

4.1 INTRODUCTION

This chapter presents a method, designed in this work, for the enhancement of ultrasound B-mode images. The method is designed to improve the visualization of characteristic features of liver tissues present in an ultrasound image and to get high diagnostic accuracy during routine checkup of patient by radiologists. Further, poor image quality of clinically acquired ultrasound images, due to speckle and low contrast, often impacts feature extraction and image analysis steps thereby impacting the effective designing of CAD systems. Thus, there is a necessity of an image pre-processing task which enhances the image contrast and reduces speckle without destroying the important diagnostic features of liver ultrasound images. The literature review in chapter 2 highlights the importance of indeterminate information present in image to be considered for image enhancement. Thus, the proposed enhancement method is based on neutrosophic similarity score (NSS) scaling that can handle indeterminate information. A liver ultrasound image is represented in the neutrosophic set (NS) domain, and NSS is defined under multi-criteria. The NSS is employed to measure the belonging degree to the true texture and is used to obtain the enhanced coefficient. The resultant coefficient is applied to scale the input image to get an enhanced image. The liver ultrasound images were enhanced based on proposed methods and state-of-the-art method. Both subjective and objective criteria were used to evaluate and compare the proposed method's performance. This chapter describes the proposed method, evaluation measures, the experimental results and discussions and finally conclusions are drawn.

4.2 NEUTROSOPHIC IMAGE

Neutrosophy, a branch of philosophy, introduced by Smarandache [159], which studies the neutralities' origin, nature, and scope. It represents every entity ‹Y›, the opposite ‹Anti-Y›, and the neutralities ‹Neut-Y› that is neither ‹Y› nor ‹Anti-Y›.

An image is defined in the NS as: let U be a universe of discourse, B_p be a bright pixel set in U, and an image $f(i,j)$ defined by using NS is called neutrosophic image $f_{NS}(i,j)$. The neutrosophic image $f_{NS}(i,j)$ is characterized by three membership sets *T, I, F* i.e. true, indeterminacy and false set. A pixel *P* in the image is described as *P (T, I, F)* and belongs to Bp in the subsequent manner: it is *t%* true in the Bp set, *i%* indeterminate, and *f%* false, where *t* varies in *T*, *i* varies in *I*, and *f* varies in *F*. The membership values can be defined under different criteria. At the *intensity image criterion* c_{in} they are defined as shown in equations (4.1) – (4.3).

$$T_{Cin}(i,j) = \frac{f(i,j)-f_{min}}{f_{max}-f_{min}} \quad (4.1)$$

$$I_{Cin}(i,j) = 1-\frac{Gd_i(i,j)-Gd_{i\,min}}{Gd_{i\,max}-Gd_{i\,min}} \quad (4.2)$$

$$F_{Cin}(i,j) = 1 - T_{Cin}(i,j) \quad (4.3)$$

where $f(i,j)$ and $Gd_i(i,j)$ are the intensity value and gradient value at the position of (i,j) on the image.

4.3 NEUTROSOPHIC SIMILARITY SCORE

Neutrosophic similarity score is defined by Ye [160] to measure the similarity degree between different elements, and has been applied widely in image processing due to its capability to describe the indeterminate information such as noises and vague edges in images. A NS can be defined under different alternatives and criteria as: let $A = \{A_1; A_2; . . ; A_m\}$ be a set of alternatives in NS, and $C = \{C_1; C_2; . . .; C_n\}$ be a set of criteria. The alternative A_m at C_n criterion is denoted as $\{T_{Cn}(A_m); I_{Cn}(A_m); F_{Cn}(A_m)\}/A_m$, where $T_{Cn}(A_m)$, $I_{Cn}(A_m)$ and $F_{Cn}(A_m)$ are the membership values to the true, indeterminacy and false set at the C_n criterion.

A similarity measurement employed to evaluate the similarity degree between two elements in NS under multi-criterion is given in equation (4.4):

$$S_{Cj}(A_m, A_n) = \frac{T_{Cj}(A_m)T_{Cj}(A_n) + I_{Cj}(A_m)I_{Cj}(A_n) + F_{Cj}(A_m)F_{Cj}(A_n)}{\sqrt{T_{Cj}^2(A_m) + I_{Cj}^2(A_m) + F_{Cj}^2(A_m)}\sqrt{T_{Cj}^2(A_n) + I_{Cj}^2(A_n) + F_{Cj}^2(A_n)}} \tag{4.4}$$

In multi-criteria environment, the concept of ideal element can be used to identify the best alternative. The ideal alternative A^* is denoted as: $\{T_{Cj}^*(A_i), I_{Cj}^*(A_i), F_{Cj}^*(A_i)\}/A_i^*$. The similarity to the ideal alternative is computed as given in equation (4.5):

$$S_{Cj}(A_i, A^*) = \frac{T_{Cj}(A_i)T_{Cj}(A^*) + I_{Cj}(A_i)I_{Cj}(A^*) + F_{Cj}(A_i)F_{Cj}(A^*)}{\sqrt{T_{Cj}^2(A_i) + I_{Cj}^2(A_i) + F_{Cj}^2(A_i)}\sqrt{T_{Cj}^2(A^*) + I_{Cj}^2(A^*) + F_{Cj}^2(A^*)}} \tag{4.5}$$

The similarity score for a pixel *P(i,j)* is calculated to identify the degree to the ideal texture for a gray scale intensity image is given in equation (4.6)

$$S_{Cj}(P(i,j), A^*) = \frac{T_{Cj}(i,j)T_{Cj}(A^*) + I_{Cj}(i,j)I_{Cj}(A^*) + F_{Cj}(i,j)F_{Cj}(A^*)}{\sqrt{T_{Cj}^2(i,j) + I_{Cj}^2(i,j) + F_{Cj}^2(i,j)}\sqrt{T_{Cj}^2(A^*) + I_{Cj}^2(A^*) + F_{Cj}^2(A^*)}} \tag{4.6}$$

The similarity value is sensitive to noise on image. This can lead to the noisy regions on images getting labeled into a wrong group. In order to make the enhancing results robust to noise and enhance edges multi-criteria that are *intensity criteria* c_{in}, *local mean intensity criterion* c_{mi} and *edge detection criterion* c_e, are considered to map the image pixel to NS domain, and NSS is calculated under them.

The mean filter is the simple and intuitive method that has been applied in digital image processing for reducing noise and smoothing of images. In this filtering method intensity value of each pixel is replaced with the mean intensity value of its neighboring pixels and itself. The enhancement is affected by the size of filter; thus, experiments are conducted with varying the filter size i.e. 3×3 and 5×5 size. Thereafter, the *local mean intensity criterion* c_{mi} can be defined by using pixel values of the resulting mean filtered image and transforming them into NS domain using the definitions given in equations 4.7- 4.10.

$$T_{Cmi}(i,j) = \frac{f_m(i,j) - f_{m\,min}}{f_{m\,max} - f_{m\,min}} \tag{4.7}$$

$$f_m(i,j) = \frac{1}{w \times w} \sum_{l=i-w/2}^{i+w/2} \sum_{k=j-w/2}^{j+w/2} f(l,k) \tag{4.8}$$

$$I_{Cmi}(i,j) = 1 - \frac{Gd_m(i,j) - Gd_{m\,min}}{Gd_{m\,max} - Gd_{m\,min}} \quad (4.9)$$

$$F_{Cmi}(i,j) = 1 - T_{Cmi}(i,j) \quad (4.10)$$

where $f_m(i,j)$ and $Gd_m(i,j)$ are the intensity and gradient magnitude value at the location of (i,j) on the image after mean filter processing. $f_{m\,max}$ and $f_{m\,min}$ are maximum and minimum intensity values, respectively. $Gd_{m\,max}$ and $Gd_{m\,min}$ are maximum and minimum of the gradient value in image.

The edge detection is aimed at identifying points in an image at which the image brightness changes sharply. The edge information is extracted to evaluate the liver diseases and is utilized to define the enhancement criterion. Thus, to obtain the edge features, an edge operator applied to the image and the *edge image criterion* is defined using the edge image values and transforming into neutrosophic set domain as described in equations 4.11- 4.13.

$$T_{Ce}(i,j) = \frac{E(i,j) - E_{min}}{E_{max} - E_{min}} \quad (4.11)$$

$$I_{Ce}(i,j) = 1 - \frac{Gd_e(i,j) - Gd_{e\,min}}{Gd_{e\,max} - Gd_{e\,min}} \quad (4.12)$$

$$F_{Ce}(i,j) = 1 - T_{Ce}(i,j) \quad (4.13)$$

where $E(i,j)$ is the edge value at (i,j) computed by using Sobel operator because of its simplicity and high speed. $Gd_e(i,j)$ is the gradient value on $E(i,j)$. E_{max} and E_{min} are maximum and minimum edge values respectively. $Gd_{e\,max}$ and $Gd_{e\,min}$ are maximum and minimum of the gradient value.

4.4 PROPOSED METHOD

A proposed method based on neutrosophic to enhance ultrasound images by computing scale coefficient NSS is defined in this section. The algorithm improves contrast, remove noise and enhances edges of ultrasound liver images. The proposed method consists of three key steps: transformation to NS domain, NSS calculation and scaling.

4.4.1 Transformation to Neutrosophic Sets Domain

The NS $\{T_{Cn}(A_i); I_{Cn}(A_i); F_{Cn}(A_i)\}$ can be interpreted for specific pixel $P(i,j)$ has its own NS at a specific criterion *n*. In the proposed method, NSS calculations are performed using alternatives and multi-criteria. The multi-criteria: the intensity criterion, local mean intensity criterion and edge detection criterion are described in sections 4.2 and 4.3.

4.4.2 Neutrosophic Similarity Score Calculations

A similarity value is computed to identify the degree to the ideal object under multi-criteria:

$$S_{Cin}(P(i,j),A^*) = \frac{T_{Cin}(i,j)T_{Cin}(A^*) + I_{Cin}(i,j)I_{Cin}(A^*) + F_{Cin}(i,j)F_{Cin}(A^*)}{\sqrt{T_{Cin}^2(i,j) + I_{Cin}^2(i,j) + F_{Cin}^2(i,j)}\sqrt{T_{Cin}^2(A^*) + I_{Cin}^2(A^*) + F_{Cin}^2(A^*)}} \quad (4.14)$$

$$S_{Cmi}(P(i,j),A^*) = \frac{T_{Cmi}(i,j)T_{Cmi}(A^*) + I_{Cmi}(i,j)I_{Cmi}(A^*) + F_{Cmi}(i,j)F_{Cmi}(A^*)}{\sqrt{T_{Cmi}^2(i,j) + I_{Cmi}^2(i,j) + F_{Cmi}^2(i,j)}\sqrt{T_{Cmi}^2(A^*) + I_{Cmi}^2(A^*) + F_{Cmi}^2(A^*)}} \quad (4.15)$$

$$S_{Ce}(P(i,j),A^*) = \frac{T_{Ce}(i,j)T_{Ce}(A^*) + I_{Ce}(i,j)I_{Ce}(A^*) + F_{Ce}(i,j)F_{Ce}(A^*)}{\sqrt{T_{Ce}^2(i,j) + I_{Ce}^2(i,j) + F_{Ce}^2(i,j)}\sqrt{T_{Ce}^2(A^*) + I_{Ce}^2(A^*) + F_{Ce}^2(A^*)}} \quad (4.16)$$

The value of ideal alternative A* under specific criteria are: False set at intensity criteria, $S_{cin}^f(i,j)$ i.e. A*= {0,0,1}; True set at intensity criteria, $S_{cin}^t(i,j)$ i.e. A* = {1,0,0}; True set at local mean criteria, $S_{cmi}^t(i,j)$ i.e. A* = {1,0,0}; True set at edge detection criteria, $S_{ce}^t(i,j)$ i.e. A* = {1,0,0}. The false set overcomes the change in image brightness intensity. The true sets of all the criteria favors contrast and edges enhancement. In the proposed method, the mean value of S_{cin}, S_{cmi} and S_{ce} as the final NSS is defined as in equation 4.17:

$$NSS(i,j) = [\, S_{cin}^f(i,j) + S_{cin}^t(i,j) + S_{cmi}^t(i,j) + S_{ce}^t(i,j)\,]/4 \quad (4.17)$$

4.4.3 NSS Scaling

Finally, the resultant NSS coefficient is utilized to scale the input image pixel $f(i,j)$. The final enhanced image (f_{en}) is constructed as:

$$f_{en}(i,j) = NSS(i,j) \times f(i,j) \quad (4.18)$$

This method denoises and enhances the contrast and edges of ultrasound image. The advantage of this scaling is that it adjusts the brightness of image such that the texture features of ultrasound image are more noticeable.

4.5 EVALUATION OF THE ENHANCEMENT METHOD

In the present study, mean filter is implemented with filters of size 3x3 and 5x5. Hereafter proposed method with filter 3 x 3 and 5 x 5 are termed as template 3 and template 5 respectively.

Evaluation of proposed method was done by two ways i.e. subjective and objective. The subjective metric for image quality assessment belong to the study of relationship between physical detectable change and human perception. In the present work, senior radiologist having more than 20 years of experience rated the image based on visual image quality. The other performance assessment criteria i.e. objective metrics can be divided into two categories. The first are mathematically defined measures such as edge preservation index (EPI), universal image quality index (UIQI) and absolute mean brightness error (AMBE). The second category consider human visual system (HVS) characteristics in an attempt to include perceptual quality such as multi-scale structural similarity (MS-SSIM).

4.5.1 Subjective Evaluation

A dataset of 189 ultrasound liver images comprising of 48 normal, 50 chronic, 50 cirrhosis and 41 HCC were processed with proposed method using template 3 and 5, and state-of-the-art method [62]. One hundred and eight-nine sets of images was provided to radiologist for subjective evaluation where each set contained four images i.e. one original image and three processed images. The three processed images are (i) enhanced image with state-of-art method proposed by Shahin et al. [62], (ii) enhanced image with proposed method using template 3 and, (ii) enhanced image with proposed with using template 5. The expert radiologist was asked to score them in the range of 1 to 5. Where ‘one’ was assigned to image having worst quality and ‘five’ with best quality. The score of ‘three’ was average score. The score of ‘two’ and ‘four’ was below and above average respectively. The score marked for images was finally collected and score assigned to each image was tabulated.

4.5.2 Objective Evaluation

Numerical values were computed to quantify the quality of a processed image in comparison to the original image. The proposed method is evaluated in terms of (i) improvement in the overall quality of processed image (ii) similarity between original and processed images and, (iii) human perception. The five evaluators used here are (i) edge preservation index (EPI), its value should be near unity that depicts preservation of edges in the processed images [40], (ii) universal quality index (UQI), its value of UQI close to one depicts there is less distortion between the original and processed image [57], (iii) Absolute mean brightness error (AMBE), the value of error should be minimum for high quality images [63] and, (iv) multi-scale structural similarity (MS-SSIM), its value near to one represent high similarity in original and processed images [161].

4.6 RESULTS AND DISCUSSIONS

The proposed method with template 3 and 5 is verified on clinically acquired ultrasound image database of liver. The performance of methods is evaluated by using both subjective and objective evaluators. The scores of subjective evaluations by the expert radiologist are given in table 4.1. The total score is summation of the scores given by the radiologist on each ultrasound image. The scores of original images are lowest for all liver classes. The template 3 performs well in all liver classes except in enhancement of HCC images in which template 5 is better. The visual score for normal, chronic and cirrhosis with template 5 and Shahin et al. methods are almost same that is 3.8 and 3.9 respectively. Whereas, the overall best score of 4.3

was obtained with template 3. The expert radiologist has recommended enhancement with template 3 as the diagnostic features were more noticeable as compared to original images and other methods.

Table 4.1 Scores of subjective evaluation

Image Category (total number of images)	**The score per ultrasound image (total score)**			
	Original Image	**Proposed Template 3**	**Proposed Template 5**	**Shahin et al.**
Normal (48)	2.87(138)	4.8(216)	4.1(198)	4.1(200)
Chronic (50)	2.92(147)	4.5(225)	3.7(186)	4.0(203)
Cirrhosis (50)	3.16(156)	4.3(218)	3.8(194)	3.8(192)
HCC (41)	3.01(126)	3.9(160)	4.0(164)	3.6(148)
Total (189)	3.0(567)	4.3(819)	3.8(742)	3.9(743)

Statistically, the performance of the enhancement methods based on subjective score was also compared using analysis of variance (ANOVA). A one-way ANOVA was computed to assess differences in the mean of subjective score between four images sets i.e. original, template 3, template 5, and Shahin. Tukey post hoc test was used to evaluate differences between these sets. The level of statistical significance was set at $p < 0.05$. All statistical computations were performed using the Minitab statistics software.

The null hypothesis state that the mean subjective score values of four different image sets are equal. In the subjective score, the results indicate a significant difference in the means obtained for different enhancement methods ($F_{3,3012} = 5304.15$, $p = 0.000$). The p-value is 0.000, which is less than the significance level of 0.05, so the null hypothesis is rejected, and conclude that some of the subjective scores have different means for enhancement methods. Here, one-way ANOVA p-value is less than the significance level, we infer that some of the set's means are different, but do not which pairs of set. Thus, Tukey's grouping information table is used to determine the mean difference between which pair of set is statistically significant. In Tukey results, the statistical observations indicate that template5 and Shahin is not statistically different however, the significant differences exist with template 3.

Figure 4.1 shows original and processed images of chronic liver and HCC. The images of chronic liver and HCC are marked as 1 and 2. The ROI is marked in a rectangular region and termed as A1and A2 of original image. Similarly, B1 and B2, C1 and C2, and D1 and D2 are marking for images processed by template 3, template 5 and Shahin et.al method. It is observed that the texture information in original image are not quite visible as compared to processed images by template 3 and template 5. In case of HCC and background texture, the contrast is more noticeable in template 5 as compared to template 3. This can be the reason that HCC got more score in subjective evaluation. However, template 3 provided appropriate overall visual appearance in comparison to original images with enhanced texture contrast and no blurring effect whereas with template 5 there is more averaging of intensity which results in smoothening of features inside the HCC lesion. All these visual representations of original and processed images can be observed in the marked region of Figure 4.1. The clarity of improvements in Figure 4.1 may be less but the improvements were clearly seen when they were visualized on computer screen with black background.

The results of subjective evaluator for proposed methods are supported by objective criteria. Table 4.2 summarizes the results of four objective evaluators measured on ultrasound images. The results are shown by their mean and standard deviation calculated from 189 ultrasound images. The standard deviation values show the deviation of values around mean. The mean value of EPI is 0.9835 with template 3 which represents the texture preservation in processed image. The value of AMBE obtained by proposed methods is very small that is 0.03 which represents the preservation of brightness in the processed image. The value of UIQI is close to 1 that is 0.9777 which depicts the close representation of processed image w.r.t. original image in terms of correlation, luminance and contrast. Finally, greater the MS-SSIM

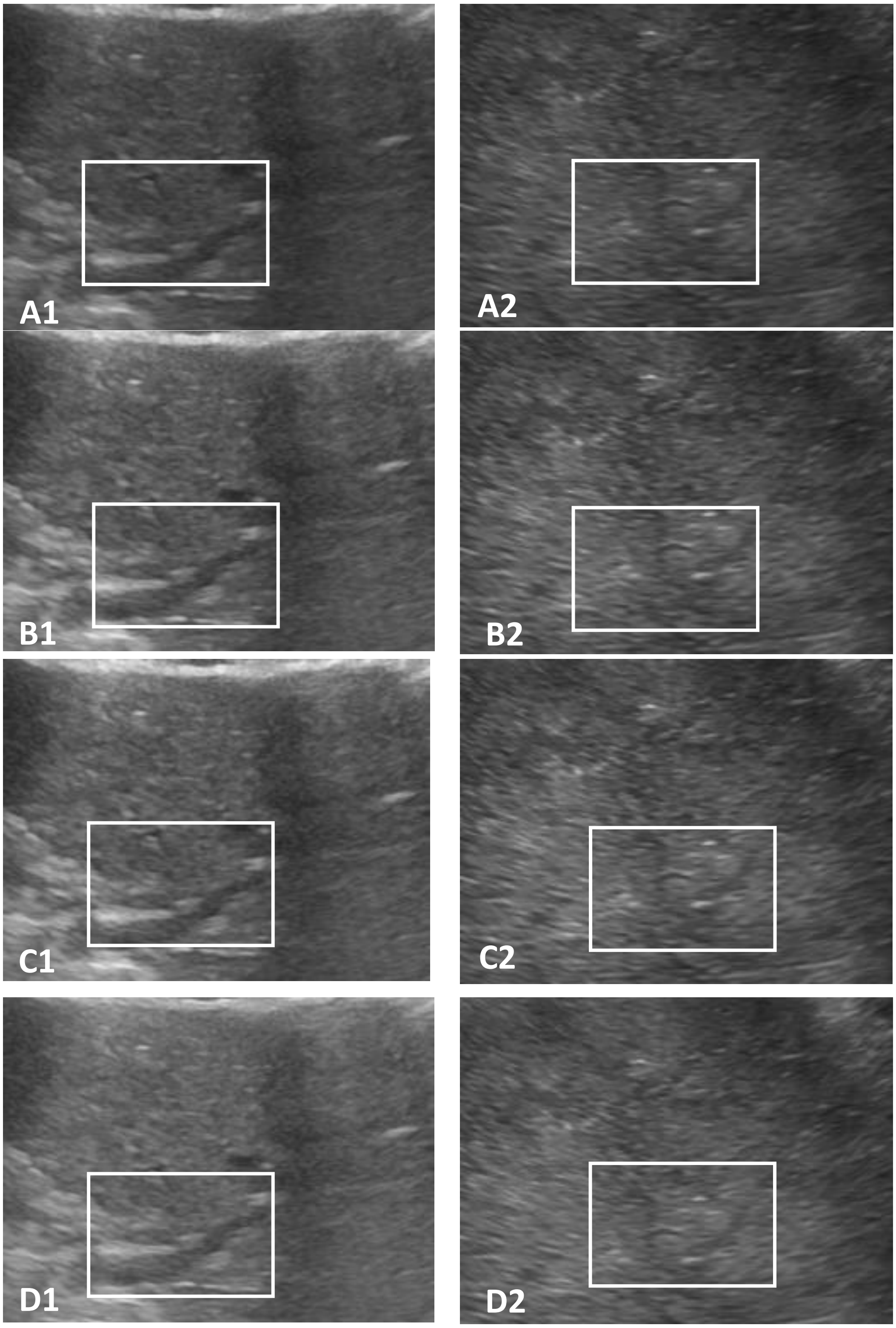


Figure 4.1 Original and processed images by proposed methods and Shahin method.

1: Chronic liver; 2: HCC; A: Original Image; B: Processed by Template 3; C: Processed by Template 5; D: Processed by Shahin et.al.

value, better the image quality is. MS-SSIM with template 3 and 5 are 0.9996 and 0.9996 which are greater than Shahin et.al. It is observed from table that the values of all objective evaluators are the highest for Template 3. The performance measures of all three methods are also tabulated separately for each liver class in tables 4.3 - 4.6. The values show the similar trend as that has been described earlier in table 4.2. This indicate that the proposed method improves the definition of point and linear features in an ultrasound image and provides better representation of diagnostic information in processed image in comparison to the original ultrasound image.

Statistically, the performances of the enhancement methods based on objective measures were also compared using ANOVA. A one-way ANOVA was computed to assess differences in the mean of EPI, AMBE, UIQI, and MS-SSIM between three methods i.e. template 3, template 5, and Shahin. Tukey post hoc test was used to evaluate differences between methods with $p < 0.05$. The null hypothesis states that the mean EPI, AMBE, UIQI, and MS_SSIM values of three different methods are equal. In AMBE, the p-value is 0.558, which is larger than the significance level of 0.05, so we accept the null hypothesis and conclude that AMBE has the same means for all methods. Whereas, in EPI, UIQI, and MS-SSIM, the p-value is 0.000, which is less than the significance level of 0.05, so we reject the null hypothesis and conclude that some of the EPI, UIQI, and MS-SSIM have different means for enhancement methods. Thus, we infer that enhanced images have significant difference in preservation of edges and image quality. In Tukey results, for EPI and UIQI, all the methods have a significant difference in their means and for MS-SSIM, Shahin and proposed methods showed significant difference. These statistical observations indicate that AMBE doesn't participate in differentiating enhancement methods. The EPI, UIQI and MSSIM provide significant difference in processed images in terms of edge preservation and quality.

Table 4.2 Results of objective performance on 189 ultrasound images by three enhancement methods

Objective Measures	**Template 3 (mean ± SD)**	**Template 5 (mean ± SD)**	**Shahin et al. (mean ± SD)**
EPI	$0.9835 \pm 4\times10^{-3}$	$0.9687 \pm 2.1\times10^{-3}$	$0.9722 \pm 1.4\times10^{-3}$
AMBE	$0.0304 \pm 8.71\times10^{-3}$	$0.0326 \pm 9.87\times10^{-3}$	$0.0401 \pm 1.53\times10^{-2}$
UIQI	$0.9779 \pm 1.08\times10^{-2}$	$0.9757 \pm 1.17\times10^{-2}$	$0.9053 \pm 3.18\times10^{-2}$
MS-SSIM	$0.9996 \pm 1.42\times10^{-4}$	$0.9996 \pm 1.63\times10^{-4}$	$0.9994 \pm 2.64\times10^{-4}$

Table 4.3 Results of objective performance by three enhancement methods in case of normal liver

Objective Measures	Template 3 (mean ± SD)	Template 5 (mean ± SD)	Shahin et al. (mean ± SD)
EPI	$0.9823 \pm 4\times10^{-3}$	$0.9650 \pm 2.1\times10^{-3}$	$0.9698 \pm 1.4\times10^{-3}$
AMBE	0.0309 ± 1.87×10-3	0.0361 ± 6.87×10-3	0.0401 ± 1.52×10-2
UIQI	0.9776 ± 1.08×10-2	0.9756 ± 1.15×10-2	0.9052 ± 3.19×10-2
MS-SSIM	0.9995 ± 1.6×10-4	0.9997 ± 1.63×10-4	0.9994 ± 2.61×10-4

Table 4.4 Results of objective performance by three enhancement methods in case of chronic liver

Objective Measures	Template 3 (mean ± SD)	Template 5 (mean ± SD)	Shahin et al. (mean ± SD)
EPI	$0.9788 \pm 4\times10^{-3}$	$0.9700 \pm 2.1\times10^{-3}$	$0.9701 \pm 1.4\times10^{-3}$
AMBE	0.0298 ± 7.83×10-3	0.0319 ± 8.87×10-3	0.0402 ± 1.50×10-2
UIQI	0.9778 ± 1.38×10-2	0.9757 ± 1.20×10-2	0.9053 ± 3.20×10-2
MS-SSIM	0.9996 ± 1.3×10-4	0.9996 ± 1.52×10-4	0.9994 ± 2.56×10-4

Table 4.5 Results of objective performance by three enhancement methods in case of cirrhosis liver

Objective Measures	Template 3 (mean ± SD)	Template 5 (mean ± SD)	Shahin et al. (mean ± SD)
EPI	$0.9815 \pm 4\times10^{-3}$	$0.9637 \pm 2.1\times10^{-3}$	$0.9742 \pm 1.4\times10^{-3}$
AMBE	0.0310 ± 1.37×10-3	0.0311 ± 2.87×10-3	0.0400 ± 1.55×10-2
UIQI	0.9787 ± 1.18×10-2	0.9757 ± 1.18×10-2	0.9054 ± 3.19×10-2
MS-SSIM	0.9997 ± 1.10×10-4	0.9996 ± 1.63×10-4	0.9994 ± 2.53×10-4

Table 4.6 Results of objective performance by three enhancement methods in case of HCC liver

Objective Measures	Template 3 (mean ± SD)	Template 5 (mean ± SD)	Shahin et al. (mean ± SD)
EPI	$0.9813 \pm 4\times10^{-3}$	$0.9665 \pm 2.1\times10^{-3}$	$0.9730 \pm 1.4\times10^{-3}$
AMBE	0.0299 ± 1.57×10-3	0.0361 ± 6.87×10-3	0.0399 ± 1.35×10-2
UIQI	0.9777 ± 1.05×10-2	0.9757 ± 1.17×10-2	0.9052 ± 3.09×10-2
MS-SSIM	0.9996 ± 1.6×10-4	0.9996 ± 1.59×10-4	0.9995 ± 2.65×10-4

Image Class **Original Image** **Enhanced Image**

NOR

CHR

CIR

HCC

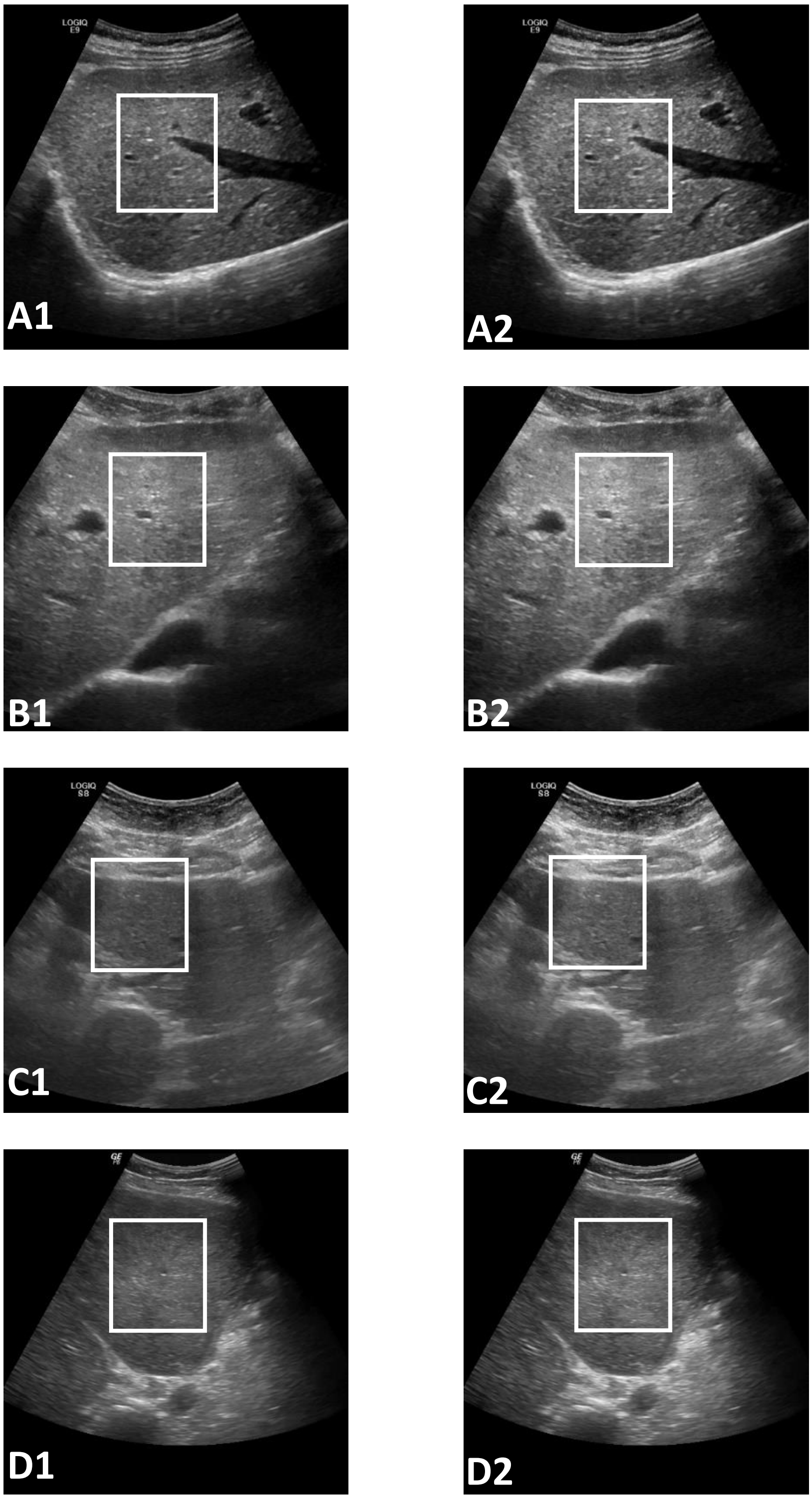


Figure 4.2 Original and Enhanced images of four liver classes

Figure 4.2 shows original and enhanced (with Template 3) images of all the four liver classes. A1, B1, C1 and D1 are original images of normal, chronic, cirrhosis and HCC. A2, B2, C2 and D2 are enhanced images by proposed Template 3 method. The diagnostically important regions where the contrast enhancement can be noted are marked in the rectangular box.

4.7 CONCLUDING REMARKS

The ultrasound images have noise and low contrast that interrupts in analyzing these images. The interpretation of ultrasound images by the experienced radiologist indicated that removal of noise and improvement in contrast should be such that diagnostically important features in the texture of ultrasound images are not affected. Thus, the foremost requirement of enhancement was texture preservation. Keeping this in mind, NSS based method was proposed with the interpretations of radiologist. The NSS was calculated over multi-criteria and final score aided to scale the image such that appropriate contrast was obtained.

In subjective evaluation, with template 3 and template 5, overall best score of 4.3 and 3.8 were obtained and that were 44% and 27% higher than the score of original images. The results of subjective evaluator were also supported by objective evaluators. The mean values of EPI were 0.9835 and 0.9687 with template 3 and template 5 which represents the texture preservation in the processed image. The values of AMBE obtained by proposed templates were very small that is 0.03 which represents the preservation of brightness in the processed image. The value of UIQIs were close to 1 that is 0.9777 and 0.9757 which depicts the close representation of processed image w.r.t original image in terms of correlation, luminance and contrast. MS-SSIM values with template 3 and 5 were 0.9996 and 0.9996 which were close to unity and represents the similarity of processed image with original image. All the results were greater than Shahin et.al. method.

It was observed that the improvement in visual quality of processed images was obtained by both the proposed templates 3 and 5. Out of these proposed templates, template 3 gave the best processed images in terms of noise removal and texture enhancement. Thus, it was concluded that the template 3 method denoises and increases contrast without blurring or smoothing texture features thereby providing more fine details of image to be visualized clearly.

Feature Extraction

5.1 INTRODUCTION

Feature extraction is the main step in the design of CAD system for the classification of liver tissues. The dataset, mentioned in chapter 3, themselves do not provide much information visually. These dataset images have to be initially processed to extract hidden information about the liver tissues. The echo patterns present in these images represent the texture of an image and can be exhibited by the texture metrics or mathematical descriptors. These mathematical descriptors are termed as features and play crucial role in differentiating the liver tissue type. The choice of features also has an important influence on accuracy of learning algorithm and the time required for execution. Various methods are available to extract features from liver images. Features can be obtained from traditional hand-designed methods termed as handcrafted features or from deep learned methods termed as deep learned features. 'Handcrafted' features are manually designed by keeping application in mind. The optimal feature set of handcrafted features is obtained via feature selection. The feature learning based on deep neural network learns from larger data and learn complex pattern. With transfer learning, learning from small number of samples is also possible. Here, features are extracted by the both approaches. In this chapter, initially the texture analysis is briefed and then feature extraction methods are described. This chapter, also describes hybrid feature selection method used to obtain optimal feature set from handcrafted features.

5.2 TEXTURE ANALYSIS

Texture of liver tissue in ultrasound images is evaluated subjectively by the radiologist. The radiologist considers various ultrasound findings of liver tissues to differentiate between normal and abnormal liver. These ultrasound findings are appearance of the texture of normal and abnormal liver tissues as discussed in chapter 1. The quantitative image analysis is required for assessment of liver disease. If the histological changes from sonograms are evaluated, then it is possible to diagnose and monitor liver diseases noninvasively. Texture analysis is an approach used to quantify patterns in images. The computed texture features are used as inputs for a classification algorithm. According to Vicas et al. texture analysis methods have better discrimination power when compared to human experts [79]. The ultrasound images show texture as one of its distinguishing attributes. The ultrasonic texture of different liver tissue type may have different appearance. The texture analysis can help in characterization of tissue on the basis of its texture content. Additionally, texture analysis techniques have been successfully applied in many medical applications, including studies on breast cancer, liver diseases classification and researchers were able to discriminate various diseases using texture analysis [20], [162]–[164]. Therefore, the present classification work is carried out with a set of texture based mathematical feature descriptors.

5.2.1 Texture of an Image

Texture is a feature in the image which can be used to describe the contents of an image. Texture may be valuable for segmenting image into ROIs and classify those regions. Texture provides the spatial arrangement of intensities in image. Russ defined texture as a descriptor of local changes in pixel brightness within regions [165] . Alternatively, texture is a measure of the variation of intensity of a surface, quantifying properties such as coarseness, smoothness and regularity. Tamura defined texture as a macroscopic region whose structure is attributed to the repetitive patterns in which elements or primitives are arranged according to a placement rule [166].

5.2.2 Texture Feature Extraction Methods

The hand-designed methods utilized to describe texture are statistical and spectral. The statistical methods characterize texture by considering the statistical properties of grey-levels of pixels and are computed from GLCM of the surface. Spectral methods are based on properties of the Fourier spectrum and describe global periodicity of the grey levels of a surface by identifying high energy peaks in the spectrum. As discussed in chapter 2, the various texture analysis methods have been proposed for efficient texture descriptors. Based on the literature, in the present work mathematical features are extracted from liver image dataset to study the possibility of numerical separation among liver classes. The “handcrafted” texture features extracted are from multiresolution ranklet transformed images, and by using GLDM and GLCM methods.

These traditional methods reported in the literature for measuring and describing the texture present in images are hand-designed. The designing of methods manually is a lengthy process, whereas the deep learning methods can learn directly from the training data and can be aimed at new applications. The possibility of learning from larger data and learning complex patterns makes deep learned features favorable for feature extraction and classification. Also, the deep learning method can overcome the problem of selecting a relevant texture features [167]. The features are automatically extracted without user interference by the hidden layers inside the deep neural network (DNN) architectures. The main characteristic of deep learning is that the features are not hand-designed, they are learned from data. In recent years, deeper and deeper networks have been proposed, for example, AlexNet, VGGNet, GoogleNet, and ResNet [168], [169]. These networks have proven their success in classifying images of both medical and nonmedical domains. Based on the literature reviewed in chapter 2, two major approaches that successfully employ CNN to medical image classification: feature extraction from pretrained CNN and fine-tuning with transfer learning, are implemented.

5.3 HANDCRAFTED FEATURE EXTRACTION METHODS AND FEATURE SELECTION

Features are called attributes or properties that are extracted from the texture of an SROI image by a feature extraction method. Feature extraction is an important step in designing a CAD system to identify a liver tissue class by a set of numerical features. The handcrafted features are described below:

5.3.1 Gray Level Difference Matrix

The GLDM characterizes the probability density of the differences between intensity levels of pixels separated by a distance $h \equiv (\Delta i, \Delta j)$ and at a rotation angle θ within the SROI. Figure 5.1 represents the image intensity function $f(i,j)$ specified by the horizontal and vertical indexes i and j respectively, where $i = 1, \dots, M,\ j = 1, \dots, N$ and $\theta = \{0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ}\}$. The difference denoted by $g_{\theta,h}(i,j)$ can be given by,

$$g_{0^{\circ},h}(i,j) = \frac{|f(i,j) - f(i,j+h)|}{h}, \tag{5.1}$$

$$g_{45^{\circ},h}(i,j) = \frac{|f(i,j) - f(i-h,j+h)|}{h}, \tag{5.2}$$

$$g_{90^{\circ},h}(i,j) = \frac{|f(i,j) - f(i-h,j)|}{h}, \tag{5.3}$$

$$g_{135^{\circ},h}(i,j) = \frac{|f(i,j) - f(i-h,j-h)|}{h}, \tag{5.4}$$

Then, $p_{\theta,h}$ is the probability density corresponding to the difference $g_{\theta,h}(i,j)$. If there are r gray levels, $p_{\theta,h}$ has the form of an r-dimensional vector i.e 1, 2... q...r whose qth component will represent the probability of occurrence with absolute difference intensity value q. In this work, value of Δi and Δj are taken as 1.

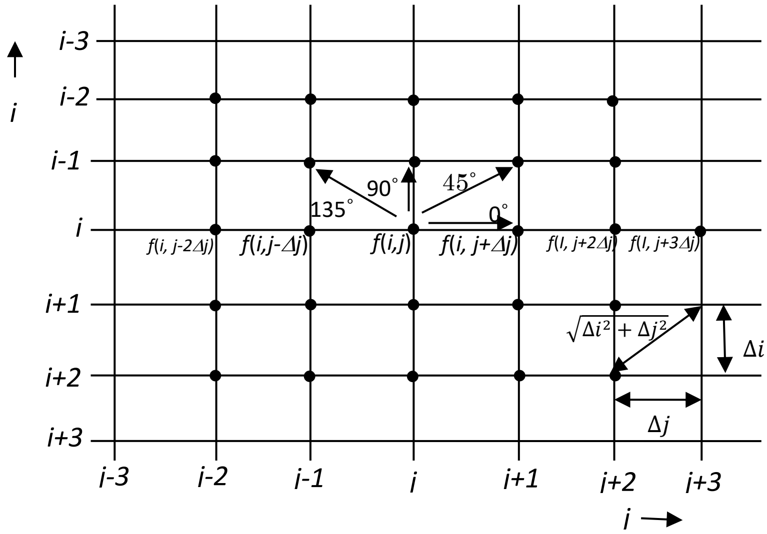


Figure 5.1 Two-dimensional structured mesh for finite difference approximations

The proposed extension of GLDM is designed by computing the absolute difference of intensities using more than two pixels i.e. three, five and seven pixels. Here, if two neighboring pixels along with the evaluation pixel are used, it is defined as 'three-pixel stencil'. If four neighboring pixels along with the evaluation pixel are being used, it is defined as 'five-pixel stencil' and so on. Each extension of GLDM can be calculated with forward, backward or central difference approximations, as defined by the symmetry of neighboring pixels. Experiment was conducted and it was observed that GLDM computed by using central difference approximation provided better results compared to forward and backward approximations [170]. These results also correlate with Khan and Ohba, where authors have also observed that central difference approximation is the best choice among the finite difference formulas [171]. The mathematical representation of extensions of GLDM using n-pixel stencils with central difference approximation is shown in Table 5.1. In this work n = 2, 3, 5 and 7 pixel stencil differences are studied.

Table 5.1 Mathematical expression of n-pixel stencil differences (where n = 2, 3, 5 and 7)

n-pixel stencil differences	Mathematical Expression	Eq. No.
Two-pixels difference	$g_{\theta,h}(i,j) = \lvert\frac{f(i,j)-f(i,j+h)}{h}\rvert$	(5.5)
Three-pixels central difference	$g_{\theta,h}(i,j) = \lvert\frac{f(i,j+h)-f(i,j-h)}{2h}\rvert$	(5.6)
Five-pixels central difference	$g_{\theta,h}(i,j) = \lvert\frac{-f(i,j+2h)+8f(i,j+h)-8f(i,j-h)+f(i,j-2h)}{12h}\rvert$	(5.7)
Seven-pixels central difference	$g_{\theta,h}(i,j) = \lvert\frac{-f(i,j-3h)+9f(i,j-2h)-45f(i,j-h)+45f(i,j+h)-9f(i,j+2h)+f(i,j+3h)}{60h}\rvert$	(5.8)

The four texture features that are contrast, entropy, angular second moment and mean are calculated from each extension of the GLDM. The features extracted from extensions of GLDM are represented as $feature\_code_{\theta}^{Method}$ for example $CON_{0°}^{2pGLDM}$. The description of the features is given below:

- Contrast (CON): Contrast is the second moment of probability density, $p_{\theta,h}(q)$. It measures the amount of local variations present in the image.

$$\text{CON} = \sum q^2 \ p_{\theta,h}(q) \tag{5.9}$$

- Angular Second Moment (ASM): Angular second moment measures the homogeneity of an image. The ASM is small when values of $p_{\theta,h}(q)$ are all same and large when some values of $p_{\theta,h}(q)$ are high and others are low.

$$\text{ASM} = \sum p_{\theta,h}(q)^2 \tag{5.10}$$

- Entropy (ENT): Entropy is the degree of disorder or randomness present in the image. The ENT is large for same values of $p_{\theta,h}(q)$ and small when $p_{\theta,h}(q)$ has unequal values.

$$\text{ENT} = -\sum p_{\theta,h}(q)\log p_{\theta,h}(q) \tag{5.11}$$

- Mean (M): Mean is small when values of $p_{\theta,h}(q)$ are concentrated near the origin and large when values of $p_{\theta,h}(q)$ are far from the origin.

$$\text{M} = \frac{1}{r}\sum q p_{\theta,h}(q) \tag{5.12}$$

Thus, a total of 64 (i.e., 4 Features × 4 Directions × 4 Neighboring Pixel Stencils) texture features are generated. Gomez et al. reported that the classification performance decreases when texture features of the same distance were averaged[172]. Thus, in this work, all the features with their specific direction are considered. Once all the features were obtained, normalization is done to balance their influence in the subsequent step.

5.3.2 Gray-Level Co-occurrence Matrix

Gray-level co-occurrence matrix illustrates how often a particular combinations of pixels values (gray levels) occur in an image [172]. The matrix, $s(i,j)$ characterizes the joint probability density of the pairs of gray levels separated by distance $h \equiv (\Delta i, \Delta j)$ and at a rotation angle θ° within the SROI. The description of the extracted features[25] is given below:

- Contrast (CON): Contrast is a measure of the degree of spread of the GLCM matrix values. The CON is high when SROI image have a large amount of intensity variations.

$$\text{CON} \equiv \sum (i-j)^2 s(i,j) \tag{5.13}$$

- Correlation (COR): Correlation measures the degree to which the rows (or columns) of the GLCM matrix are similar to each other. The COR is high when values in matrix are uniformly distributed, otherwise it is low.

$$\text{COR} \equiv -\sum \frac{[(ij)s(i,j)-\mu_x\mu_y]}{\sigma_x\sigma_y} \tag{5.14}$$

- Energy (ENE): Energy, is also known as angular second moment, is a measure of textural uniformity of an image. The ENE is high when gray level distribution is either constant or has periodic form.

$$\text{ENE} \equiv \sum[s(i,j)]^2 \tag{5.15}$$

- Homogeneity (HOM): Homogeneity, also known as inverse difference moment, is a measure of how uniformly a given region is structured with respect to its gray level variations.

$$\text{HOM} \equiv \sum_i \sum_j \frac{1}{1+(i-j)^2} s(i,j) \tag{5.16}$$

- Dissimilarity (DIS): Dissimilarity measures the local variability.

$$\text{DIS} \equiv \sum|i-j|s(i,j) \tag{5.17}$$

- Entropy (ENT): Entropy measures the degree of randomness in the SROI.

$$\text{ENT} \equiv -\sum_i \sum_j s(i,j) \log s(i,j) \tag{5.18}$$

- Cluster Prominence (CPROM): Cluster prominence is a measure of asymmetry. The value of CROM is high when intensity of SROI is less symmetric.

$$\text{CPROM} \equiv \sum\left(i-\mu_x+j-\mu_y\right)^4 s(i,j) \tag{5.19}$$

- Cluster Shade (CSHAD): Cluster shade is a measure of the skewness of GLCM matrix.

$$\text{CSHAD} \equiv \sum\left(i-\mu_x+j-\mu_y\right)^3 s(i,j) \tag{5.20}$$

- Variance (VAR): Variance is a measure of heterogeneity. The value of VAR increases when intensity values differ from their mean (μ).

$$\text{VAR} \equiv \sum_i \sum_j (i-\mu)^2 s(i,j) \tag{5.21}$$

- Max probability (MPROB): Maximum probability measures the most probable intensity pair in an ROI.

MPROB ≡ max $(s(i,j))$ (5.22)

- Autocorrelation (ACOR): Autocorrelation measures the regularity and the fineness/coarseness of texture present in image.

ACOR ≡ $\sum(i\ j)p(i,j)$ (5.23)

- Inverse difference moment normalized (IDMN)

IDMN ≡ $\sum_i \sum_j \frac{1}{1+((i-j)/N)^2} s(i,j)$ (5.24)

These are a total of 12 features representing the second order statistics of an image. Each feature can be calculated from a normalized GLCM matrix representing a specific inter-pixel spacing *d* and orientation *θ* between pixel pair. By varying *d* and *θ*, a number of normalized GLCM matrices can be generated, where each matrix will result in a different set of 12 features. In the present work, GLCM matrix characterizes the joint probability density of the pair of gray levels separated by distance *d* = 1 and at a rotation angle *θ* ° = 0°, 45°, 90°, and 135° within the SROI. Thereafter, 48 (i.e., 12 features × 4 directions × 1 distance) texture features are generated. The features extracted from GLCM are represented as $feature\_code_{\theta}^{Method}$ for example $CON_{0°}^{GLCM}$.

5.3.3 Ranklet Transform

The ranklet transform considers the relative rank of pixels instead of their gray-scale value [173]. These relative ranks, characterizes the ranklet transform as invariant to any monotonic changes of any observed pixels. The multi-resolution, nonparametric and orientation-selective properties of ranklet transform aid in defining the rank of pixel within a local region. In this work, *ranklets* are defined as a series of transformed images based on multi-

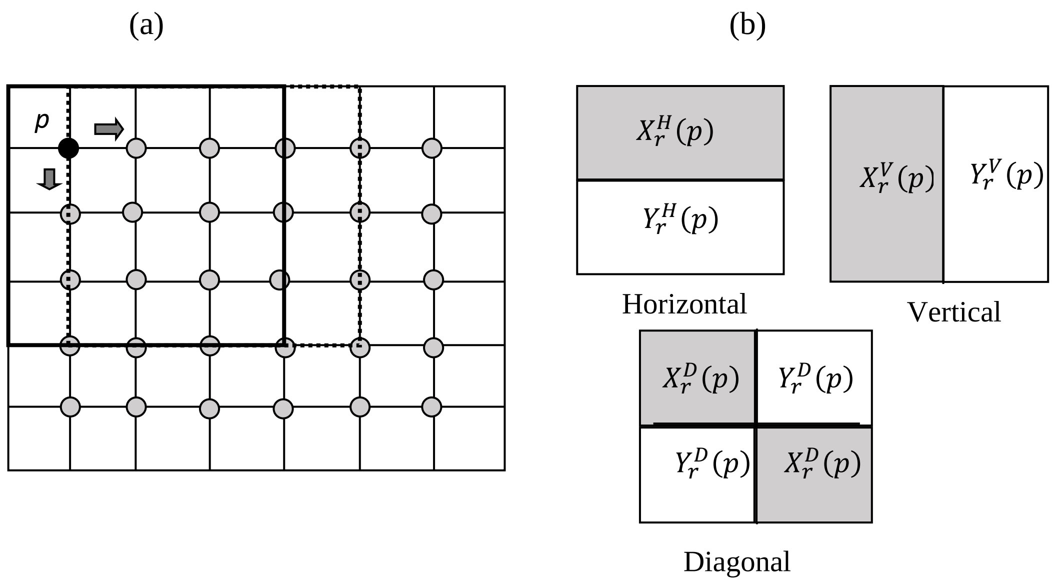


Figure 5.2 (a) The diagram depicting an arbitrary point *p* and resolution *r* in an overlapping square crop for determining ranklet transform, multi-resolution analysis. (b) The three orientations i.e. vertical, horizontal and diagonal directions of the square crop with clusters $\boldsymbol{X_r^t(p)}$ and $\boldsymbol{Y_r^t(p)}$

resolution and orientation-selective properties as shown in figure 5.2. The image is divided into overlapping square crops at different resolutions, *r*, which can be termed as multi-resolution analysis. Then, for each resolution, each crop is subdivided into two clusters, *X* and *Y*, depending on the considered orientation *t* i.e. horizontal (*H*), vertical (*V*) and diagonal (*D*). The ranklet coefficient of point *p* with specified resolution *r* and orientation *t* can be computed as follows:

- Rank the intensity values of the pixels within the observed crop (i.e., all observed gray points in $X \cup Y$).
- The rank numbers present in clusters *X* and *Y* are rearranged in ascending order to form the sorted rank descriptive vectors *RSX* and *RSY*, respectively.
- Finally, the ranklet coefficient *RC* can be defined as:

$$RC_r^t(p) \equiv \left| \frac{(\sum_{i=1}^{N/2} RSX_r^t(p,i) - RSY_r^t(p,i)}{(\frac{N^2}{4})} \right| \qquad t\ \varepsilon\ \{H, V, D\},\ r = \{4,6,8,10\} \qquad (5.25)$$

where $RSX_r^t(p,i)$ and $RSY_r^t(p,i)$ are the i-th sorted rank number in the respective subset while N (N = r^2) is the total pixels of the observed crop. The dynamic range of the ranklet coefficient is [0, +1] and it represents the contrast strength between two clusters X_r^t(p) and Y_r^t(p). Then ranklet coefficient for each target point within the SROI of liver ultrasound image is computed and ranklets $R_r^t(i,j)$ are formed. Thereafter, from these ranklets following texture features are extracted and finally, 36 (i.e., 4 Resolutions × 3 Orientations × 3 Features) texture features are computed. The features extracted from *ranklets* are represented as $feature\_code_{resolution,orientation}^{Method}$ for example $M_{4,H}^{Ranklets}$.

- *Mean (M):* Mean is defined as sum of derived rank coefficients divided by the total number of rank coefficients in the ROI of liver ultrasound image.

$$\mathrm{M} \equiv \frac{1}{N}\sum_i \sum_j R_r^t(i,j) \qquad (5.26)$$

- *Entropy (ENT):* Entropy represents the amount of randomness in rank coefficients.

$$\mathrm{ENT} \equiv -\sum_{\mathrm{i}} \sum_{\mathrm{j}} \mathrm{R}_{\mathrm{r}}^{\mathrm{t}}(\mathrm{i},\mathrm{j}) \log \mathrm{R}_{\mathrm{r}}^{\mathrm{t}}(\mathrm{i},\mathrm{j}) \qquad (5.27)$$

- *Standard deviation (STD):* Standard deviation describes the spread of rank coefficients around the mean.

$$\mathrm{STD} \equiv \sqrt{\frac{\sum_i \sum_j (R_r^t(i,j)-M)^2}{N}} \qquad (5.28)$$

5.3.4 Extracted Features

By using the above discussed hand-designed methods, a total of 148 features were extracted to distinguish the different liver tissues. The brief of extracted features is given in table 5.2.

Table 5.2 Extracted features

Feature extraction method	Number of extracted features
GLDM	64
GLCM	48
Ranklets	36

5.3.5 Feature Selection and Feature Fusion

Feature Selection

Feature selection plays significant role in training a classifier. Suppose an original feature set contains F_D number of features, the aim of feature selection is to find an optimal subset that contains only Fd ($Fd \leq FD$) features while retaining the capability of feature set to distinguish among classes. This can reduce the computational complexity and improve the classification accuracy. Here, the hybrid feature selection (HFS) method designed by combining filter method (ReliefF algorithm) and wrapper method (SFS) leverage the benefits of both. These methods have their respective weaknesses and are complementary to each other as discussed in section 2.4.1. The filter methods have low computational cost but does not consider specific classification algorithm, whereas wrapper methods tend to have superior classification accuracy but require great computational power. Figure 5.3 shows the flow diagram of proposed feature selection method.

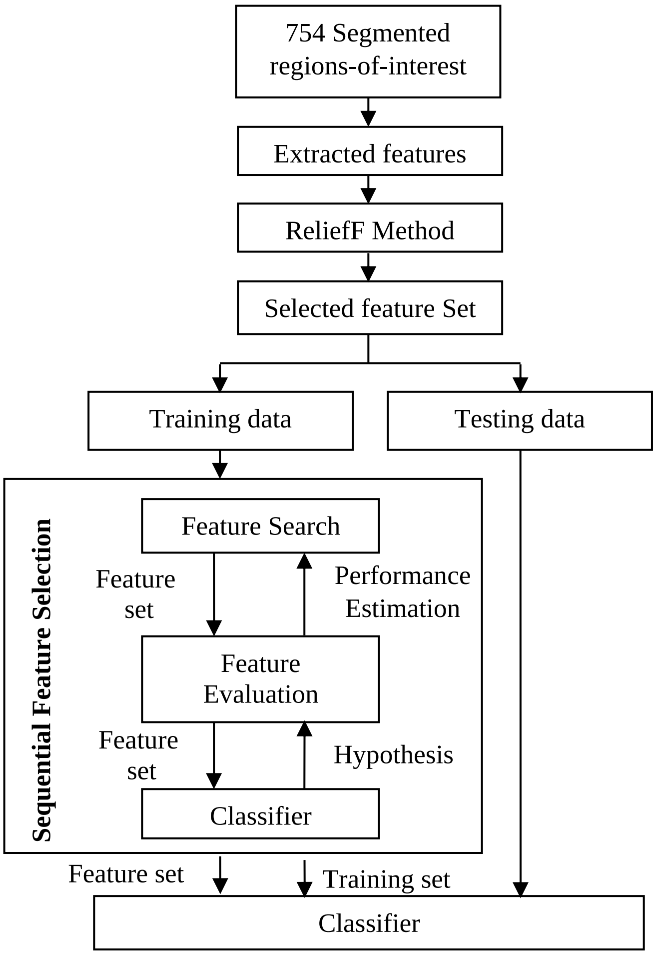


Figure 5.3 Flow diagram of hybrid feature selection method

ReliefF

The ReliefF is multi-class extension of Relief method that uses a statistical method to select the relevant features[110]. ReliefF method is multivariate, robust and noise tolerant filter method. It is a feature weight-based algorithm. Initially weights of features are assigned to zero. A data point is randomly selected from a set of training data that determines *Near Hit* and *Near Miss* data points based on a Euclidean distance. *Near Hit* is the data point having minimum Euclidean distance among all data points of the same liver class as that of the selected data point. *Near Miss* is the data point having minimum Euclidean distance among all data points of different liver class. The weights of features are updated based on an idea that a feature is more relevant if it distinguishes between a data point and its *near miss*, and less relevant if it distinguishes between a data point and its *near hit*. After selecting all data points in the training set, the final ranking is performed.

Sequential forward selection

Sequential forward selection is wrapper method [83]. It is simple and efficient greedy search method. In this method features are sequentially added to an empty candidate set. Every time one feature is added that minimize the error the most, until any further addition does not significantly minimize the error. The performance of each added feature is estimated using cross-validation measures. SFS being wrapper method interact with classifiers. In this work, SFS is combined with classifiers namely, SVM and k-NN.

Hybrid feature selection

The proposed HFS method is designed by combining ReliefF and SFS methods. In the first stage, ReliefF is applied to rank features according to the assigned weights. Thereafter, a subset of features is selected having weights greater than defined threshold. The selected high ranked features help in reducing complexity of the resulting model and computation time for the selection of optimal set. In the second stage, SFS method is applied to find the optimal set with minimum number of features and maximum accuracy. HFS makes feature selection faster since filter method rapidly reduces the effective number of features under consideration for next step and wrapper method finds the optimal set. Finally, the subset with highest accuracy, is considered the best (if two or more subsets obtain the same accuracy, the one using smallest number of features is selected).

Feature Fusion

The feature extracted from same pattern by GLDM and its extension, GLCM and *ranklets* may contain complementary information. To take the benefit of this information, feature fusion schemes (i) serial feature combination, (ii) serial feature fusion and, (iii) hierarchical feature fusion are implemented [90], [144].

- Serial feature combination: It is a process of serial combination of individual feature sets, and resulting feature set is called serial feature combination set. Suppose α and β are two feature sets where α is *k*-dimensional and β is *l*-dimensional. Then, serial combined feature set is defined by $\gamma = \binom{\alpha}{\beta}$, where γ is a (*k*+*l*)- dimensional feature set [144].
- Serial feature fusion: It is a process of selection of relevant features from serial feature combination set, and resulting feature set is called serial fused feature set [111].
- Hierarchical feature fusion: It is two stage process: (a) initially hybrid feature selection is performed on individual feature sets to obtain optimal individual feature subsets (b) and then, derived subsets are integrated to form a fused feature space, and again hybrid feature selection is applied on this fused space to obtain the optimal feature subset [90].

5.4 DEEP LEARNED FEATURES

In this section, two major approaches are described that employ CNN for feature extraction to classify liver ultrasound images: (i) feature extraction from pre-trained CNN, and (ii) feature extraction from fine-tuned CNN [174], [175]. In the first approach, the key point is to retain the learning of model in its original form and then utilize its outputs to feed the classifier [123]. This approach can be helpful if computational power is less or dataset is small [121]. In the second approach, the weights of the few layers of network are adjusted according to the new dataset. The lower layers of DNN refer to general features (problem independent), while higher layers refer to specific features (problem dependent). Usually, if dataset is small and learning parameters are large, more layers are left frozen to avoid overfitting [176]. Whereas, if the dataset is large and learning parameters are small, model can be improved by training more layers to the new task. In case there is very large dataset then model can be learnt from scratch and it will require a lot of computational power. However, in present work the dataset is limited so model is not learned from scratch.

Convolutional neural networks are so-called due to the presence of convolution layers in their architectures. These layers can detect certain local features from all locations of the input images by using filters. After every convolution layer, a non-linear layer i.e. *ReLU* layer is applied. The aim of this layer is to introduce non-linearity into the system. In CNN, to reduce the computational complexity and extract the hierarchical set of features, each convolution layer is followed by a max pooling layer. A typical architecture of CNN comprises of several pairs of convolutional and pooling layers, followed by a fully connected layer and finally, a softmax layer, to generate the desired outputs. In this research work, AlexNet model is used as it is relatively simple CNN architecture, easy to retrain and has obtained great success in classification tasks. Additionally, it has been recognized as an excellent foundational hierarchical and automatic classification model. Brief description of layers of CNN is given below:

Convolution Layer: In convolution layer the parameters consist of a set of learnable kernels or filters. These kernels help to extract features that are present throughout an image. Every kernel is small matrix which convolve all over the input image volume. For illustration, on a first layer of CNN, a kernel of size 3x3x3; represent width, height and depth. During the forward pass, each kernel is convolved across the width and height of the input image volume and dot products between the kernel and the input at any position are computed. To vary intervals of kernel the value of stride directs by how much the filter should move at each step. Suppose the stride is one then kernel move one pixel at a time. If the stride is two then the kernel jumps two pixels at a time as they convolve around. As the kernel convolve over the width and height of the input image a two-dimensional activation/feature map is generated that presents the responses of that kernel at every spatial position.

ReLU Layer: A ReLU layer execute a threshold operation to each element of the input, where any value less than zero is set to zero. ReLU function is given by $f(x) = max(0, x)$. ReLU layer is used after convolutional layer, to induce sparsity in the features and to solve the problem of vanishing gradient.

Pooling Layer: A pooling layer progressively reduce the spatial size of the feature representations, makes the features robust against noise and distortion, and also control overfitting. Pooling layer is inserted in-between successive convolution layers. This layer operates independently on every depth slice of the input and resizes it spatially, mostly max operation is used. Usually, a pooling layer with filters of size 2x2 is applied with a stride of two. Here, every max operation will choose the maximum value in the 2×2 matrix as the output.

Fully-connected layer: In a fully connected layer every neuron has full connections to all activations in the previous layer. Their activations can be computed with a matrix multiplication followed by a bias offset.

5.4.1 Feature Extraction from Pre-Trained CNN

In this approach, learned features are extracted from a pre-trained network, and these features are used to train a classifier, such as a SVM, k-NN etc. In this research, a pretrained model AlexNet is used which is trained on ImageNet database [177]. The AlexNet is trained on ImageNet database, which is a large visual database with more than 14 million images. It contains more than 20,000 categories with a typical category, such as mouse, keyboard, pencil, and animals, consisting of several hundred images.

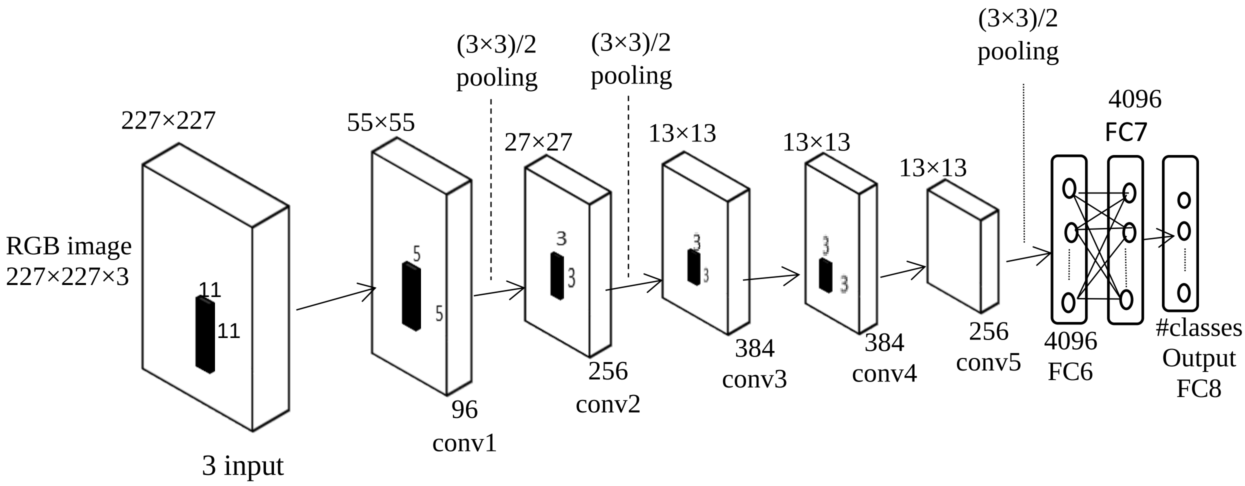


Figure 5.4 Architecture of AlexNet

AlexNet, gained popularity when it won the ImageNet Large Scale Visual Recognition contest- 2012. AlexNet received a top-5 error around 16% compared to 26.2% achieved by the second-best entry. The architecture consists of eight layers of transformations, five convolutional layers followed by two fully connected hidden layers and an output layer as shown in figure 5.4. In AlexNet architecture, the input image size is 227×227×3. The first convolutional layer, conv1, filters the input image with 96 filters. The size of each 96 filter is 11×11×3 with a stride of four pixels and zero padding. The output of conv1 layer is given as input to conv2 layer after normalizing and max pooling and conv2 layer filters the input with 256 filters. The size of each 256 filters was 5 × 5 × 48 with a stride of one pixel and with padding. The output of conv2 layer is given as input to conv3 layer after normalizing and max pooling and conv2 layer filters the input with 384 filters. The size of each 384 filters was 3 × 3 × 256 with a stride of one pixel and with padding. The conv3, conv4 and conv5 layers in the architecture are connected to one another without any intervening of normalization or pooling layers. The output of conv3 layer is given as input to conv4 layer and conv4 layer filters the input with 384 filters. The size of each 384 filters was 3 × 3 × 192 with a stride of one pixel and with padding. The conv5 layer has 256 filters of size 3 × 3 × 192. The two fully-connected layers i.e. fc6 and fc7 have 4096 neurons each. The output of the last fully-connected layer i.e. fc8 is fed to a 1000-way softmax which produces a distribution over the thousand class labels. The ReLU non-linear layer is applied to the output of every convolutional and fully-connected layer. Finally, features can be extracted from pretrained model and be used to train classifiers.

5.4.2 Feature Extraction from Fine-tuned CNN

A promising alternative to pre-trained CNN is to fine-tune a CNN that has been trained using a large labelled dataset from a different application. The key role of fine-tuning with transfer learning is to reprocess knowledge obtained in a previous training process, in order to enhance the learning process in a new task. In transfer learning, the start is from patterns that have been learned when solving a different task instead of starting the learning process from scratch. In this way, learning from previous tasks are leveraged. This is useful because learning a new task begins from scratch and requires a large amount of training data. In this approach, the features learned using a large dataset are transferred to the new network and only the classifier part is trained with the new, much smaller dataset.

In this research a CNN architecture, AlexNet is pre-trained on the ImageNet dataset. This pre-trained CNN is then adapted to the current research problem by replacing the last fully connected layer (intended for 1000 classes) with a new fully connected layer for four classes. The initial weights of CNN are derived from ImageNet dataset and then fine-tuned (optimized) through back-propagation so that they better reflect the four liver classes. The weights of the early layers of the pretrained network are transferred to the fine-tuning CNN. Initially, the learning is set high results for fast learning and then it is tapered down.

5.5 CONCLUDING REMARKS

Feature extraction is an important stage for designing a CAD system to identify a liver tissue class. The aim of this research work is to find useful features that have discriminatory information enabling one form of tissue to be differentiated from another. The liver diseases are clinically differentiated on the echo texture patterns and hence texture analysis is considered efficient for feature extraction. The three handcrafted feature extraction methods are selected on the basis of an extensive literature study for the classification of liver images. One of the methods i.e. GLDM was also extended to obtain more precise information from the liver images i.e. providing information about the roughness of liver surface and the inhomogeneous echotexture. After extraction of handcrafted features for effective application of computational intelligence in liver image classification is to select an optimal set of features. Here, feature fusion schemes are implemented to take benefit of complimentary information different feature sets have and hybrid feature selection method is implemented to select optimal feature set to classify four liver diseases. The purpose of applying hybrid feature selection method is to get a feature subset with minimum number of features and maximum accuracy.

Additionally, deep learned features via pre-trained AlexNet model and via fine-tuning of the trained model are also extracted. The key attributes of overall feature extraction process are as follows: (i) the feature extraction study is developed with the knowledge gained from the visual discrimination criteria, used by radiologist, of liver classes which helped to capture essential mathematical features; (ii) addition with visual knowledge, knowledge was gained from literature survey to distinguish ultrasound liver images; (iii) features are also learned via deep learning.

Chapter 6

Classification

6.1 INTRODUCTION

Computer based classification systems can assist radiologists in finding liver abnormalities. In the previous chapter, texture features were extracted by using hand-designed methods and deep learning methods. The success of a computer based system depends jointly on the features and classification method. An efficient set of textural features decides the correct detection of liver abnormalities and the classification method provides potential to produce accurate classification using these features. The importance of ensemble classifier in medical images classification is observed in literature studied in section 2.5. Therefore, the present chapter deals with designing of ensemble classifier to produce efficient classification using the extracted features. In beginning of this chapter, performance measures are described that are used for the evaluation of a classification system and then each component of ensemble classifier is described.

6.2 CONFUSION MATRIX AND PERFORMANCE MEASURES

The confusion matrix is also known as error matrix. It is a matrix to describe the performance of a classification model on a set of test data for which the true values are known. The number of cases correctly and incorrectly predicted by classifier are summarized in this matrix and helps to predict the various performance measure, like accuracy, sensitivity etc. of the classification system.

The classification problem in the present study is a four-class problem. The four classes are NOR, CHL, CIR and HCC. Table 6.1 shows the confusion matrix for the four-class classification system. The matrix is two-dimensional and number of rows and columns are same as number of classes. The rows represent the ground-truth or actual class and the columns represent the predictions made by classification system. Matrix contains the number of cases that have been correctly or incorrectly classified for each class. The number of cases at the diagonal indicate the number of cases that have been correctly classified for each class and off-diagonal indicate the number of cases that are incorrectly classified.

Table 6.1 Confusion matrix for four-class classification system

		Classification results (Predicted class)			
		CIR	**CHL**	**HCC**	**NOR**
Actual class (ground truth)	**CIR**	*TP_{CIR}	**$E_{CIR-CHL}$	$E_{CIR-HCC}$	$E_{CIR-NOR}$
	CHL	$E_{CHL-CIR}$	TP_{CHL}	$E_{CHL-HCC}$	$E_{CHL-NOR}$
	HCC	$E_{HCC-CIR}$	$E_{HCC-CHL}$	TP_{HCC}	$E_{HCC-NOR}$
	NOR	$E_{NOR-CIR}$	$E_{NOR-CHL}$	$E_{NOR-HCC}$	TP_{NOR}

*TP: True prediction, **E: Error made by the classification system on test dataset

The performance of a classification system is the ability of the system to correctly predict the test data to its actual class. Accuracy is the performance measure that is used for assessment of the goodness of a classification system. Accuracy for the four-class classification system is defined in terms of *class accuracy* and *overall accuracy*.

- *Class accuracy*: It is the ratio of correctly classified cases of class to the total number of cases in that class with respect to ground truth.
- *Overall accuracy*: It is overall correctness of the classification system. It is the ratio of correctly classified cases (diagonal elements) from each class to the total number of cases.

The confusion matrix for two-class classification system is shown in table 6.2. Here the entries are TP, FP, TN and FN. Where TP is the number of liver tissues correctly classified as the disease in question; FP is the number of liver tissues without disease wrongly classified as disease in question; FN is the number of liver tissues wrongly classified as without the disease; TN is the number of liver tissues correctly classified as without disease.

Table 6.2 Confusion matrix for a two-class classification system

	Predicted positive	**Predicted negative**
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Accuracy is expressed by the overall rate of correctly and wrongly classified classes and can be defined as:

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

However, for unbalanced set, accuracy may not be a good criterion for evaluating a classification system. The other measures of diagnostic test are specificity, sensitivity, positive predictive value and negative predictive value. These are defined as:

Specificity is also called true negative rate. It measures the proportion of actual negatives cases that are correctly identified.

$$Specificity = \frac{TN}{TN+FP} \tag{6.2}$$

Sensitivity is also called true positive rate or recall. It measures the proportion of actual positives cases that are correctly identified.

$$Sensitivity = \frac{TP}{TP+FN} \tag{6.3}$$

Positive predictive value (PPV) is also called precision. It measures the proportion of correctly classified positive cases among all cases which are predicted positive by the classification system in test set.

$$Positive\ predictive\ value = \frac{TP}{TP+FP} \tag{6.4}$$

Negative predictive value (NPV) is the proportion of correctly classified negative cases which are predicted negative by the classification system in the test set.

$$Negative\ predictive\ value = \frac{TN}{TN+FN} \tag{6.5}$$

6.3 ENSEMBLE CLASSIFIER

Classification is the task of learning a model that maps the feature set of input data to a pre-defined class. The systematic approach to build classification models is called classifier. Each classifier i.e. SVM, k-NN, decision tree etc. uses a learning algorithm to learn a model that best fits the feature set and labelled class of training data. The generated model should best fit both the training data as well as correctly predict the class of test data. Therefore, the main objective of the algorithm is to build models with good generalization capability.

An ensemble of classifiers is a set of classifiers whose individual predictions are combined in some way to classify data it has never seen before. In supervised learning, building a good ensemble classifier is an active area of research. Several researchers have reported that ensemble classifiers are often much more accurate than the individual classifiers which are used to make them [178], [179]. The key advantage of using ensemble classifiers is the improvement of performance; disadvantage is the amount of time it takes to finish. As the main concern of the work was to build a model which has better performance, ensemble classifier was selected. The ensemble classifier is also known as multiple classifier system wherein multiple learners are trained to solve the same classification problem. The ensemble classifiers which use a single base classifier produce homogenous ensembles. Ensembles which use multiple classifier produce heterogenous ensembles. Based on the literature discussed in section 2.5, the classifiers: k-NN, SVM and RF has been successfully used as part of ensemble classifiers in medical applications. Thus, in the proposed ensemble classifier design as shown in figure 6.1, different base classifiers i.e. k-NN and SVM and RF, and combiner i.e. majority voting are used.

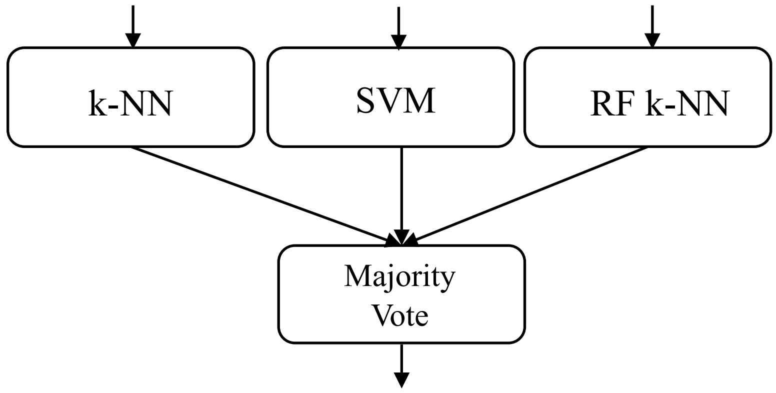


Figure 6.1 Proposed ensemble classifier

In this proposed ensemble classifier, independent training of all the base classifiers is done and finally, output of base classifiers is combined with majority voting. Here, diversity is obtained by employing different types of classifier i.e. k-NN, SVM and RF as each classifier contains an explicit or implicit bias. The trained ensemble classifier or ensemble model can be regarded as a combination of different strategies which enhances the generalization ability of designed model as each component in the model learn about some part of classification problem. The four components of proposed ensemble model, that is, k-NN, RF, SVM, and majority voting are described below:

6.3.1 k-Nearest Neighbor

k-nearest neighbor is one of the simplest classifiers in machine learning. It is supervised machine learning method that rely on labeled input data and have intense application in pattern recognition and medical image classification. It is widely used because it is non-parametric, that is, it does not make any underlying assumptions about the distribution of data. It computes the k-nearest neighboring occurrences and assigns the liver class to the test data by taking majority vote of the class of neighboring occurrences. The algorithm assumes that data points with same class exist in close proximity. The similarity in Euclidean distances between the stored training data and test data is used as the basis for majority. In this work, k-NN classifier is implemented in Matlab statistics and machine learning toolbox with k = 1 closest neighbor and its class-value is assigned to the test data[180]. The k-NN model can be implemented by following the below steps:

i. The training data is loaded
ii. The value of 'k' is initialized
iii. To predict the class of test data, iterate from 1 to total number of training data points
    a. The distance is calculated between test data and all the training data points. Here Euclidean distance is used as distance metric
    b. In an array, the calculated distance values are arranged in ascending order
    c. Get top 'k' rows from the sorted array
    d. Note the most frequent liver class of these rows
    e. Return the predicted liver class

6.3.2 Rotation Forest

Rotation forest is an ensemble classifier introduced by Rodriguez et al. in 2006 [181]. The key characteristic is transforming the subsets of features and reconstructing the whole feature space for each classifier in the ensemble. All the classifiers are trained parallelly. The following steps are considered for constructing the training set for each classifier:

i. Randomly split the feature space into subsets. The split of subsets can be disjoint or intersecting. However, to maximise the possibility for high diversity, disjoint subsets are chosen.
ii. For every subset of feature space, randomly select a non-empty subset of classes and then extract a bootstrap sample of data of size 75 % of the data count. The transformation PCA is applied on selected subset of features and selected subset of data and store the generated coefficients.

iii. Then, arrange the coefficients, obtained for every subset, in a sparse rotation matrix. Rearrange the columns of rotation matrix according to the original feature sequence. Then transformed training set for classifier is formed.

iv. Build classifier using the obtained training set.

Rotation forest encourages diversity and individual accuracy within an ensemble classifier. The whole feature set is reconstructed with all the transformed features in the ensemble. In classification phase, the predicted class of test data by all the base classifiers, in our case k-NN, is fused by majority voting and finally, the test data is assigned to the class with maximum vote. In this research work, k-NN is used as base classifier in RF ensemble, this is denoted as RF k-NN. The RF k-NN implementation in the ‘Combining pattern classifiers: Methods and algorithms’ book is used in this work [182].

6.3.3 Support Vector Machine

In early 1990s, SVM was developed as a non-linear solution for classification problems. The three key characteristics behind the success of SVM for providing reliable performance are (i) capability to learn well with a small number of features, (ii) robustness against the error of models, and (iii) efficient computation compared to other machine learning methods such as neural networks [139].

Support vector machine is supervised machine learning method that provides non-linear classification. It is based on the idea of decision planes that define decision boundaries. A decision plane separates the data points which have different class labels. In case data is not linearly separable then the idea is to transform the data to a higher dimensional space where it can be separated by a simple decision plane or hyperplane. The mapping of input data to higher-dimensional feature space is done by the using a kernel function and then a linear model is constructed in this feature space. The kernel functions often used in SVM are linear, radial basis function, polynomial, and sigmoid function. Although SVM is primarily a binary classifier, to handle multi-class cases, one may use *one*-versus-*one* (1-1) or *one*-versus-*rest* (1-r) approaches. In this work, SVM (1-1) approach was used to classify four liver stages and radial basis function kernel was used. The tradeoff parameter C was taken as 1 and the gamma of the kernel function, γ was set at 0.5. The SVM classifier is implemented using Matlab statistics and machine learning toolbox [182].

6.3.4 Majority Voting

Vote is a meta algorithm which performs the decision process by applying several classifiers. For the majority voting, the majority of liver class labels which are predicted by each classifier are used to identify the final liver class of an instance [150].

6.4 CONCLUDING REMARKS

An ensemble model for classifying liver ultrasound images into four stages, that is, normal, chronic liver, cirrhosis, and HCC has been proposed. An ensemble model is designed, by combining k-NN, SVM, and RF k-NN to increase diversity and higher consistency. The performance measures are also described that are used for the evaluation of a classification system. The proposed ensemble model is used in the designing of CAD system, and studies and results are discussed in next chapter.

Results and Discussions

7.1 INTRODUCTION

Computer-aided diagnosis systems can assist radiologist in the diagnosis of liver diseases. The present research work is carried out in three studies and, as a result, three CAD systems are designed:

(i) CAD system designed with handcrafted features extracted from original images

(ii) CAD system designed with deep learned features extracted from original images

(iii) CAD system designed with the best architecture out of study(i) and study(ii) using enhanced images

This chapter presents the three studies with their methods, and results and discussions.

7.2 STUDY 1: DESIGNING OF CAD SYSTEM USING HANDCRAFTED FEATURES WITH ORIGINAL IMAGES

Aim: Design a CAD system by using handcrafted features extracted from original1 dataset.

Method: In this study three sub-studies were conducted.

7.2.1 Sub-study 1

The primary aim of first sub-study was to examine (i) the performance of proposed GLDM extension and (ii) the effectiveness of handcrafted features for classification of four liver tissue types. In this sub-study, features were extracted from original1 dataset by using three handcrafted feature extraction methods i.e. GLDM along with its proposed extensions, GLCM and *ranklets*. GLDM extracted features, as obtained from local derivatives of gray-levels, provide information about the roughness of liver surface. GLCM features provide the measurement of statistical properties, visual characteristics, and correlation-based information of gray levels that help characterize different stages of liver. In addition, the multiresolution texture features obtained using *ranklets* have less sensitivity towards different ultrasonic devices [96].

Initially, classification accuracies with the traditional GLDM i.e. $GLDM_{2p}$ and the extensions of GLDM i.e. $GLDM_{3p}$, $GLDM_{5p}$ and $GLDM_{7p}$ were computed in order to examine the trend of accuracy with increase in number of pixels in derivative calculation. Thereafter, the effectiveness of individual features sets i.e. GLDM along with its extension, GLCM and *ranklets* was evaluated by calculating their classification accuracy separately with two classifiers i.e. SVM and k-NN.

***Results and discussion*:**

Sub-study 1(i) The prior works in literature computed local derivative by using only two-pixel stencil in GLDM, we explored computation of local derivative by using two-, three-, five- and seven- pixel stencils. As stated earlier, to obtain best results with GLDM, accurate computation of local derivative is very important. The use of higher number of neighbouring pixels improves the approximation accuracy of local derivative. The classification accuracies of feature sets from GLDM and its extensions are given in table 7.1. The accuracies obtained from $GLDM_{2p}$, $GLDM_{3p}$, $GLDM_{5p}$ and $GLDM_{7p}$ were 80%, 85%, 79% and 64% with k-NN, and 75%, 77%, 62% and 52% with SVM. Here, it was observed that, three-pixel stencil provided better results than two-pixel, five-pixel and seven-pixel stencils. This observation pointed us towards a trend that, for our dataset, increase in number of pixels for derivative calculation beyond three-pixel results in a decrease in accuracy [170].

Table 7.1 Performance in terms of accuracy of classifiers (k-NN and SVM) based on features from $GLDM_{2p}$, $GLDM_{3p}$, $GLDM_{5p}$ and $GLDM_{7p}$ in classifying four liver classes

GLDM Features	Accuracy (%)	
	k-NN	**SVM**
$GLDM_{2p}$	80	75
$GLDM_{3p}$	85	77
$GLDM_{5p}$	79	62
$GLDM_{7p}$	64	52

Sub-study 1(ii) Further, the table 7.2 shows the classification accuracy of different texture features. The classification accuracies obtained from GLDM (64 features), GLCM (48 features), and features extracted from *ranklets* (36 features) were 92.2%, 89.8%, and 91.2% with k-NN, and 91.1%, 88.7%, and 89.2% with SVM. The results show that extracted texture features are effective, as the classification accuracy of all methods (GLDM, GLCM, *ranklets*) is near 90%.

Table 7.2 The classification accuracy with texture feature methods: GLDM, GLCM and *ranklets*

Texture features	Accuracy (%)	
	k-NN	**SVM**
GLDM	92.2	91.1
GLCM	89.8	88.7
Ranklets	91.2	89.2

7.2.2 Sub-study 2

The primary aim of second sub-study was to (i) evaluate the effectiveness of proposed HFS method, (ii) obtain reduced feature sets using proposed HFS method, and (iii) evaluate the impact of feature fusion schemes.

The proposed HFS method was designed by combining filter method (ReliefF) and wrapper method (SFS), as presented in section 5.3.5. The SFS being a wrapper method depends on classifier. Here two classifiers, i.e. SVM and k-NN, were used with SFS denoted as SFS_{kNN} (SFS with kNN as classifier) and SFS_{SVM} (SFS with SVM as classifier). Consequently, in this study, after applying HFS we would obtain two sets of selected features: (i) one set would be obtained with $HFS_{k\text{-}NN}$ (using ReliefF and SFS_{kNN}) and (ii) other set would be obtained with HFS_{SVM} (using ReliefF and SFS_{SVM}).

Sub-study 2(i) The effectiveness of proposed HFS method was demonstrated by comparing the method with frequently used feature selection methods i.e. SBS and SFS. Two classifiers were used for each feature selection methods that resulted in $SBS_{k\text{-}NN}$, SBS_{SVM}, $SFS_{k\text{-}NN}$, SFS_{SVM}, $HFS_{k\text{-}NN}$ and HFS_{SVM} feature selection methods. The effectiveness of proposed method was evaluated on the combined set of all the handcrafted features i.e. 148 features. This is serial feature combination set denoted as SFCS. The comparison of the methods was conducted on the basis of number of selected features, accuracy and computation time.

Sub-study 2(ii) The reduced feature sets were obtained by applying HFS on each individual feature set, i.e. GLDM, GLCM and *ranklets*. These sets were HFS_{kNN}^{GLDM}, HFS_{SVM}^{GLDM}, HFS_{kNN}^{GLCM}, HFS_{SVM}^{GLCM}, $HFS_{kNN}^{ranklet}$and $HFS_{SVM}^{ranklet}$. The classification accuracy of these sets was computed by using k-NN and SVM classifiers separately.

Sub-study 2(iii) The impact of feature fusion schemes (discussed in section 5.3.5) on accuracy of CAD systems was observed by exploiting the complementary information extracted from different set of texture features. The flow diagram of experiment with feature fusion scheme is shown in figure 7.1. Serial feature combination scheme is implemented in order to observe the impact of all the features, taken together, on characterization of liver tissue categories. This set was denoted as SFCS. Thereafter, serial feature fusion scheme was implemented. In this fusion

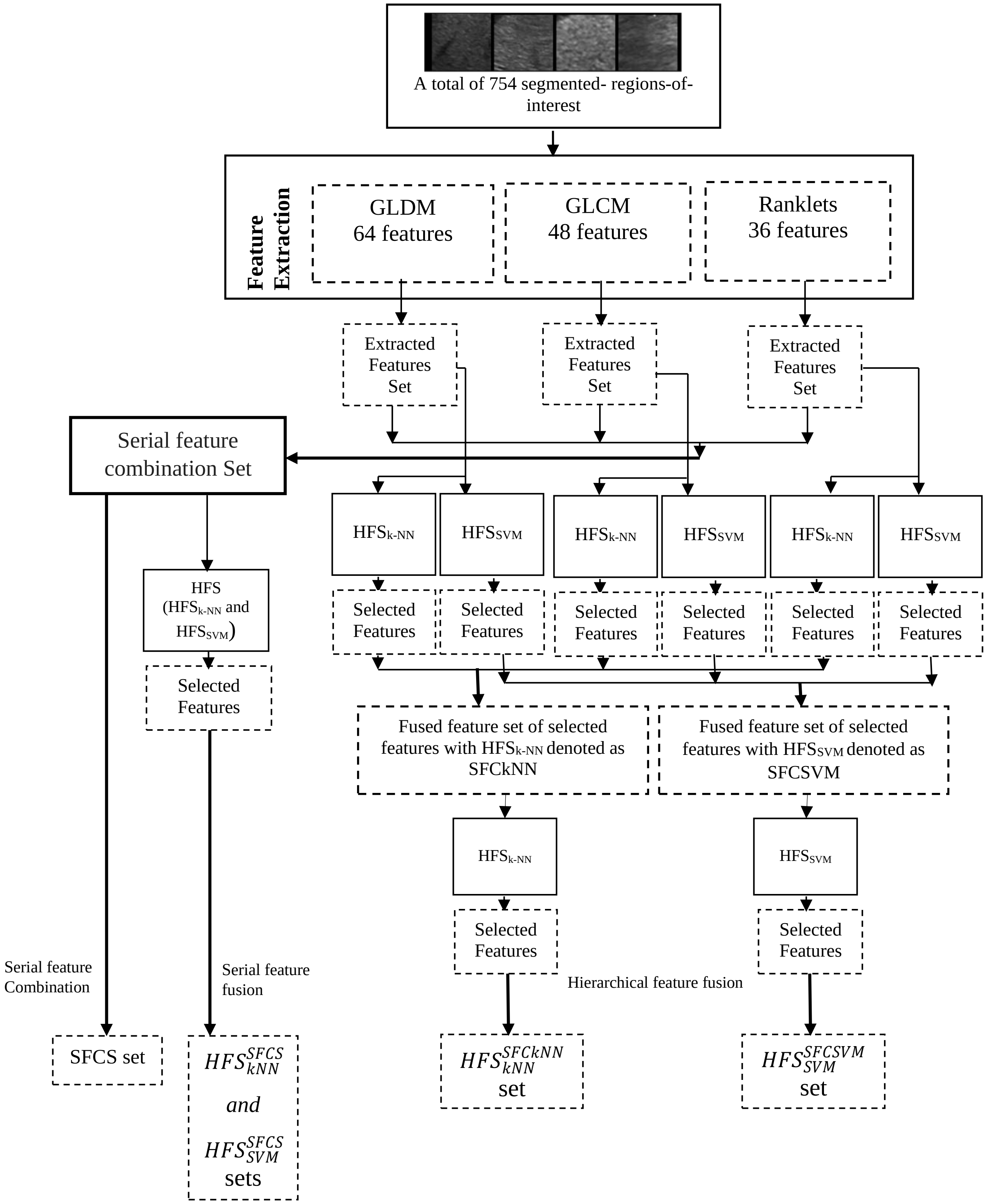


Figure 7.1 The flow diagram of experiment feature fusion schemes

scheme, HFS method was applied to SFCS set. Here, two features sets were obtained i.e. HFS_{kNN}^{SFCS} and HFS_{SVM}^{SFCS}. Finally, hierarchical feature fusion was implemented. In this fusion scheme, the selected features from individual feature sets were fused and sets were denoted as SFCkNN and SFCSVM. Then, HFS was applied on these feature sets. Final hierarchical feature fusion sets were denoted as HFS_{kNN}^{SFCkNN} and HFS_{SVM}^{SFCSVM}. The classification accuracy of selected feature sets was computed by using k-NN and SVM classifiers separately.

***Results and discussion*:**

Sub-study 2(i) The results of number of selected features, accuracy and computation time to evaluate the effectiveness of proposed HFS method are presented in figures 7.2- 7.4. The number of features selected with $SBS_{k\text{-}NN}$, SBS_{SVM}, $SFS_{k\text{-}NN}$, SFS_{SVM}, $HFS_{k\text{-}NN}$ and HFS_{SVM} were 94, 48, 15, 22, 13 and 21 respectively. The selected features with $SBS_{k\text{-}NN}$, SBS_{SVM}, $SFS_{k\text{-}NN}$, SFS_{SVM}, $HFS_{k\text{-}NN}$ and HFS_{SVM} gave classification accuracy of 93.2%, 89.5%, 93%, 90.1%, 94.5% and 91.5% respectively. The time taken to select features with $SBS_{k\text{-}NN}$, SBS_{SVM}, $SFS_{k\text{-}NN}$, SFS_{SVM}, $HFS_{k\text{-}NN}$ and HFS_{SVM} was 43, 383, 55, 105, 27 and 70 minutes. All the experiments were run on 2.67 GHz Intel i5-M480 CPU and 4.00 GB RAM. Matlab R2016a statistics and machine learning toolbox was used.

It was observed that feature space got reduced significantly with all feature selection methods. The number of features selected using $SFS_{k\text{-}NN}$ and SFS_{SVM} was 15 and 22, whereas with $HFS_{k\text{-}NN}$ and HFS_{SVM} the number was 13 and 21. The selected features were almost same but computation time for selecting features with $SFS_{k\text{-}NN}$ and SFS_{SVM} was extremely high as compared to $HFS_{k\text{-}NN}$ and HFS_{SVM}. The classification accuracy with reduced feature sets obtained by SBS_{kNN} and SBS_{SVM} was also less than $HFS_{k\text{-}NN}$ and HFS_{SVM}. Thus, it can be concluded that proposed HFS gave optimal set of features with maximum accuracy, minimum number of selected features and computational time.

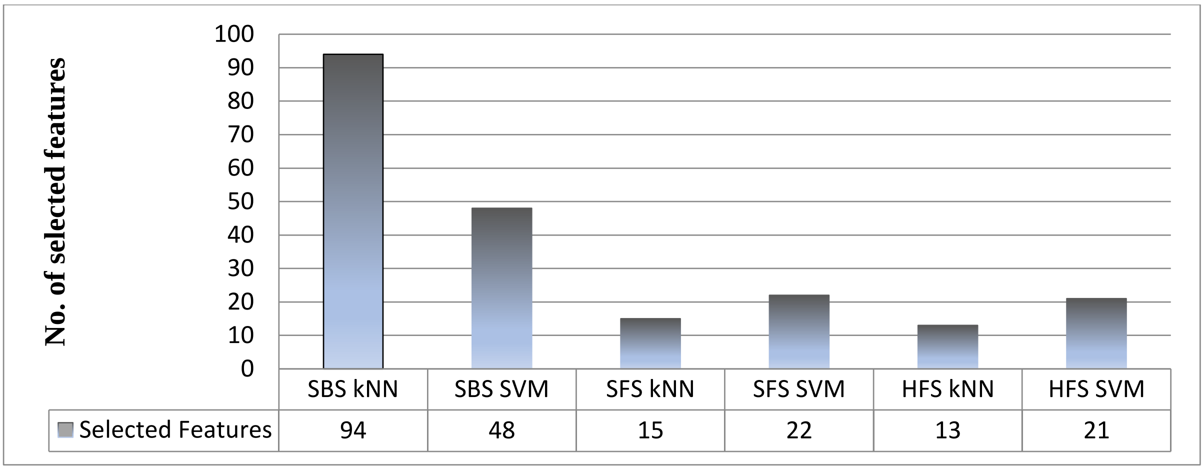


Figure 7.2 The number of selected features by SBS, SFS and proposed method (HFS) with classifiers, k-NN and SVM

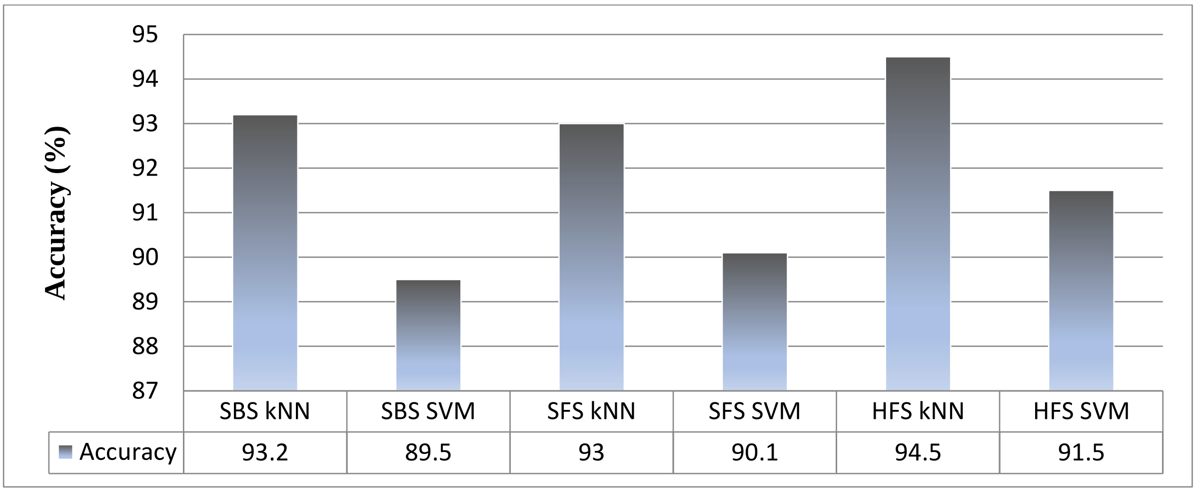


Figure 7.3 The classification accuracy obtained after applying SBS, SFS and proposed method (HFS) with classifiers, k-NN and SVM

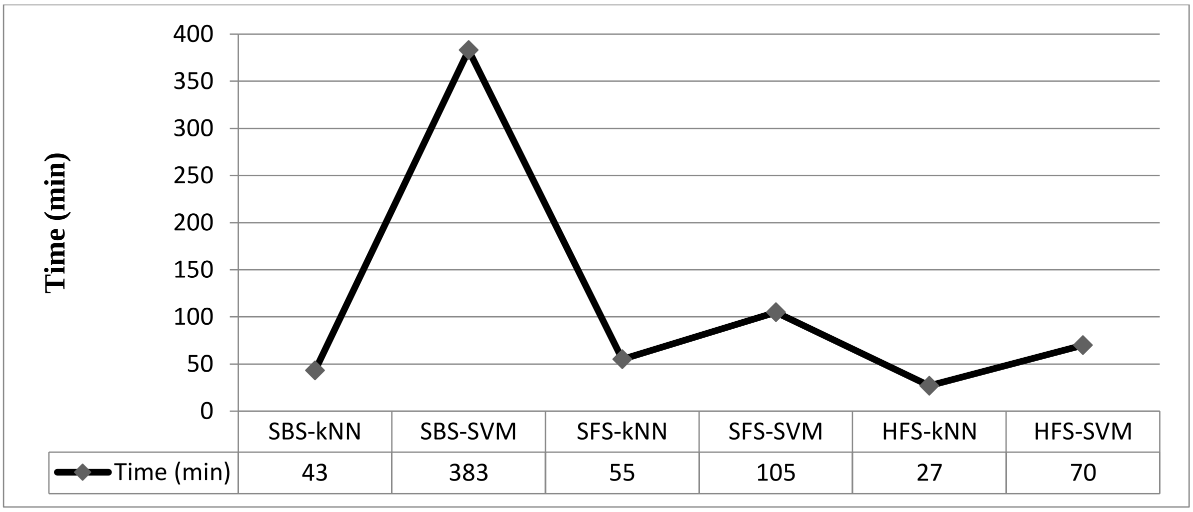


Figure 7.4 The time taken for selection by SBS, SFS and proposed method (HFS) with classifiers, k-NN and SVM

Table 7.3 The classification accuracy of the different texture features after feature selection

Texture features	Number of selected features		Accuracy	
	HFS_{kNN}	HFS_{SVM}	k-NN	SVM
GLDM	12	21	93.2%	91.2%
GLCM	09	20	90.5%	89.4%
Ranklet	13	19	92.1%	91.2%

Sub-study 2(ii) The classification accuracies results, as shown in table 7.3, obtained after applying HFS on GLDM and its extension, GLCM and *ranklets* were 93.2/90.5/ 92.1% and 91.2/89.4/91.2% with k-NN and SVM respectively. The selected number of features in sets HFS_{kNN}^{GLDM}, HFS_{SVM}^{GLDM}, HFS_{kNN}^{GLCM}, HFS_{SVM}^{GLCM}, $HFS_{kNN}^{ranklet}$ *and* $HFS_{SVM}^{ranklet}$ were 12, 21, 9, 20,13 and 19 respectively. It was noticed that feature dimensional space got reduced with the application of HFS method. For instance, the reduction in number of features for GLDM, GLCM and *ranklets* with HFS_{kNN} were recorded as 81%, 81% and 63%, and with HFS_{SVM} as 67%, 58% and 47% respectively. This reduction in dimensionality of features increases the speed and reduces the complexity for next set of experiments.

Table 7.4 The classification accuracy result after feature fusion

Feature Fusion	Feature set	No. of features	Accuracy
Serial Combination	SFCS	148	94.5 with k-NN
	SFCS	148	91.5 with SVM
Serial fusion	HFS_{kNN}^{SFCS}	13	94.9 with k-NN
	HFS_{SVM}^{SFCS}	21	91.7 with SVM
Hierarchical fusion	HFS_{kNN}^{SFCkNN}	11	95.2 with k-NN
	HFS_{SVM}^{SFCSVM}	16	92.3 with SVM

Sub-study 2(iii) The table 7.4 provides the classification accuracy results for all feature fusion schemes. The accuracy of serial combination set was recorded 94.5% and 91.5% with k-NN and SVM respectively. The accuracy of serial fusion set was 94.9% and 91.7% with k-NN and SVM respectively. The accuracy of hierarchical fusion set was 95.2% and 92.3% with k-NN and SVM respectively. It can be inferred from table 7.3 and table 7.4 that all the classification accuracy results improved after feature fusion compared to individual feature sets. The hierarchical feature fusion set is the best reduced feature set with maximum accuracy. The features of optimal sets HFS_{kNN}^{SFCkNN} *and* HFS_{SVM}^{SFCSVM} are shown in table 7.5. These sets are used in further experiments.

Table 7.5 Features selected with hierarchical feature fusion

Features of set HFS_{kNN}^{SFCkNN}
$ENT_{4,H}^{Ranklets}$, $ENT_{4,V}^{Ranklets}$, $M_{4,H}^{Ranklets}$, $ENT_{6,V}^{Ranklets}$, $VAR_{0°}^{GLCM}$, $CON_{135°}^{GLCM}$, $COR_{45°}^{GLCM}$, $CON_{3P,135°}^{GLDM}$, $ASM_{3P,135°}^{GLDM}$, $CON_{2P,0°}^{GLDM}$, $M_{5P,0°}^{GLDM}$
Features of set HFS_{SVM}^{SFCSVM}
$ENT_{4,H}^{Ranklets}$, $M_{4,V}^{Ranklets}$, $STD_{4,V}^{Ranklets}$, $M_{6,H}^{Ranklets}$, $STD_{6,D}^{Ranklets}$, $STD_{10,H}^{Ranklets}$, $COR_{45°}^{GLCM}$, $CON_{2P,0°}^{GLDM}$, $CON_{2P,45°}^{GLDM}$, $ENT_{2P,0°}^{GLDM}$, $CON_{3P,0°}^{GLDM}$, $ENT_{3P,45°}^{GLDM}$, $ASM_{5P,45°}^{GLDM}$, $ASM_{5P,90°}^{GLDM}$,$ENT_{5P,0°}^{GLDM}$, $M_{5P,135°}^{GLDM}$

7.2.3 Sub-study 3

The third sub-study was performed with aim to evaluate the performance of the proposed ensemble model.

Sub-study 3(i) The proposed ensemble model, presented in chapter 6, is evaluated by giving different inputs to each classifier of model. The different inputs are: (a) individual selected feature sets, (b) serial feature combination set of selected features, and (c) hierarchical feature fusion set.

Sub-study 3(i)(a) Individual selected feature sets are passed as input to different classifier i.e. SVM, k-NN and RF k-NN as shown in figure 7.5. The sets $HFS_{kNN}^{\mathrm{GLDM}}, HFS_{kNN}^{\mathrm{GLCM}}$ and $HFS_{kNN}^{\mathrm{Ranklet}}$ are given as input to k-NN and RF k-NN, and $HFS_{SVM}^{\mathrm{GLDM}}, HFS_{SVM}^{\mathrm{GLCM}}$ and $HFS_{SVM}^{\mathrm{Ranklet}}$ are given to SVM. Here, each classifier gets three inputs. Each classifier acts as an ensemble with different inputs and use majority voting as combiner to produce one output. Then, on the three outputs again majority voting is applied and final output is generated. Unlike traditional ensemble models that have only one level of prediction, the designed ensemble model, with three different classifier structures, has two levels of prediction.

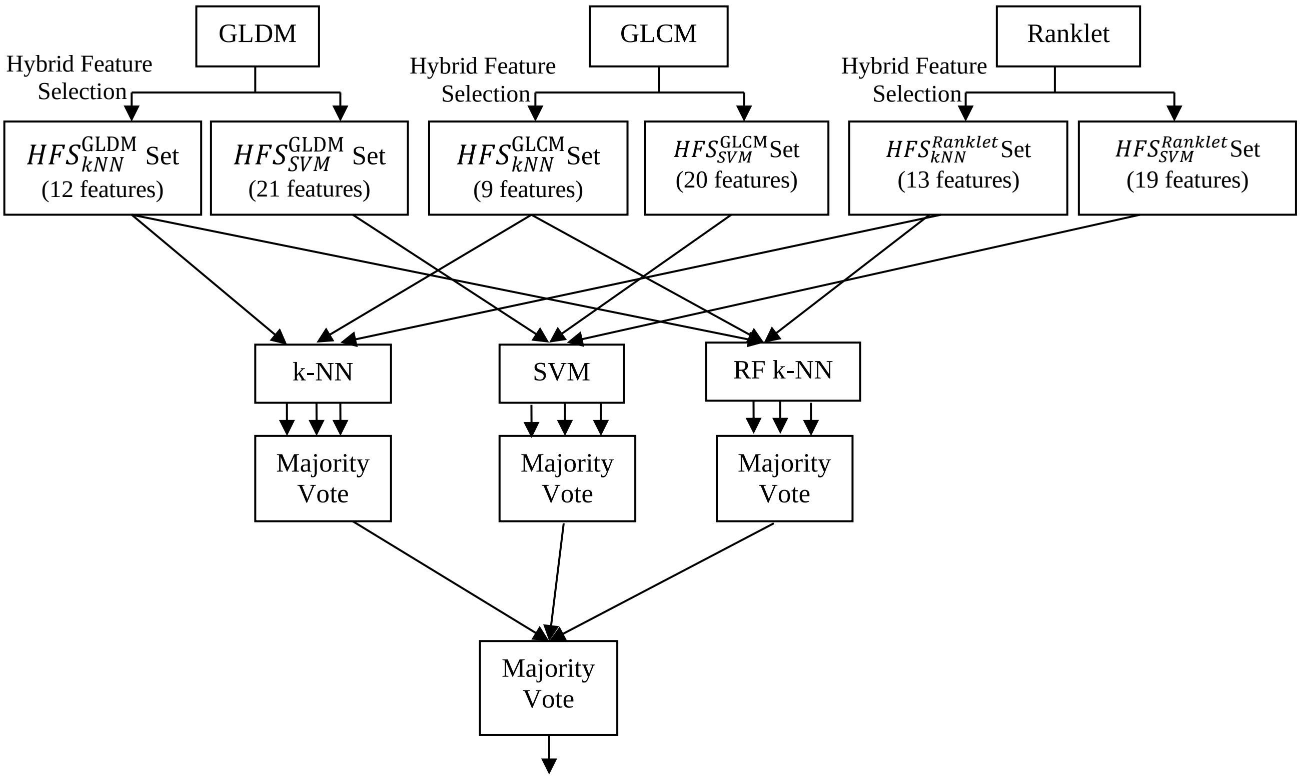


Figure 7.5 Flow diagram for sub-study3(i)(a)

Sub-study 3(i)(b) Serial feature combination set of selected features, i.e. fused feature sets of individual selected feature sets, are passed as input to different classifiers i.e. SVM, k-NN and

RF k-NN as shown in figure 7.6, on characterization of liver classes. The SFCkNN is set of 34 (12 GLDM, 9 GLCM, 13 Ranklets) features formed by fusing all the selected features with HFS_{kNN}. SFCSVM is set of 60 (21 GLDM, 20 GLCM, 19 Ranklets) features formed by fusing all the selected features with HFS_{SVM}. The SFCkNN set is given as input to k-NN and RF k-NN, and the SFCSVM set is given as input to SVM and three outputs are produced. These outputs are passed to majority voting and final output is produced.

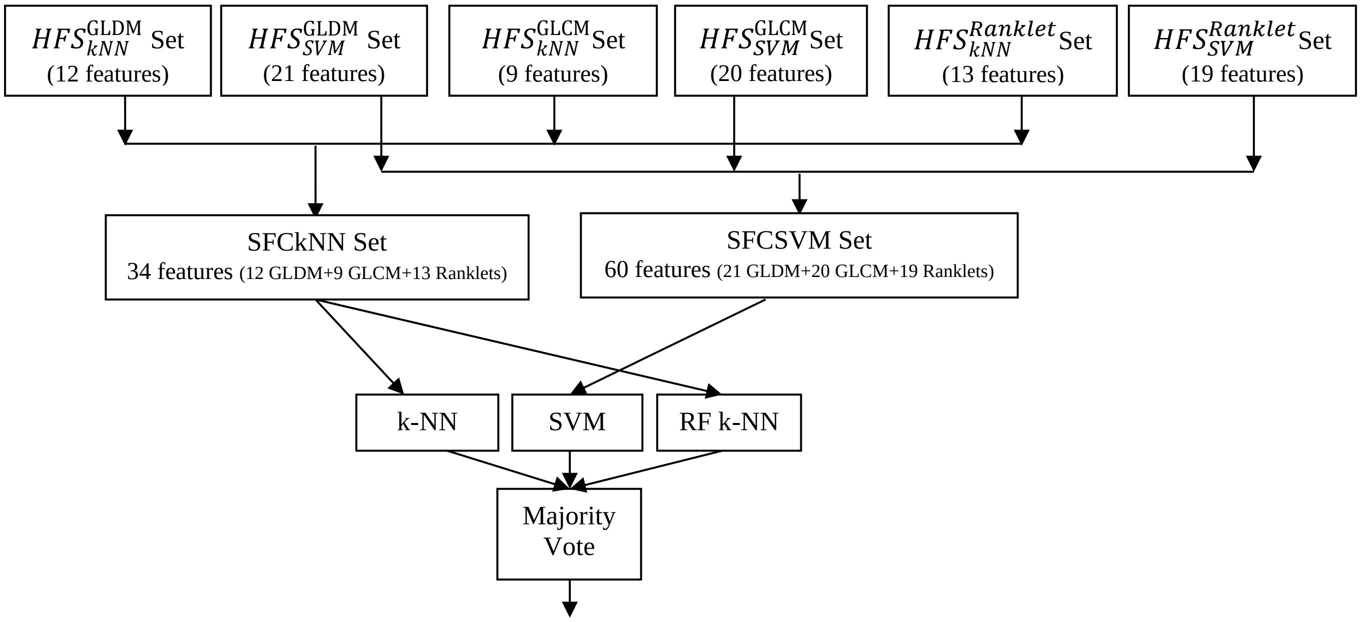


Figure 7.6 Flow diagram for sub-study3(i)(b)

Sub-study 3(i)(c) Lastly, hierarchical feature fusion sets are passed as input to different classifiers i.e. SVM, k-NN and RF k-NN, as shown in figure 7.7. The HFS_{kNN}^{SFCkNN} is set of 11 features formed after applying HFS_{kNN} on SFCkNN. The HFS_{SVM}^{SFCSVM} is set of 16 features formed after applying HFS_{SVM} on SFCSVM. HFS_{kNN}^{SFCkNN} set is given as input to k-NN and RF k-NN, and HFS_{SVM}^{SFCSVM} set is given as input to SVM. The majority voting is used to form the combined output of the proposed ensemble model.

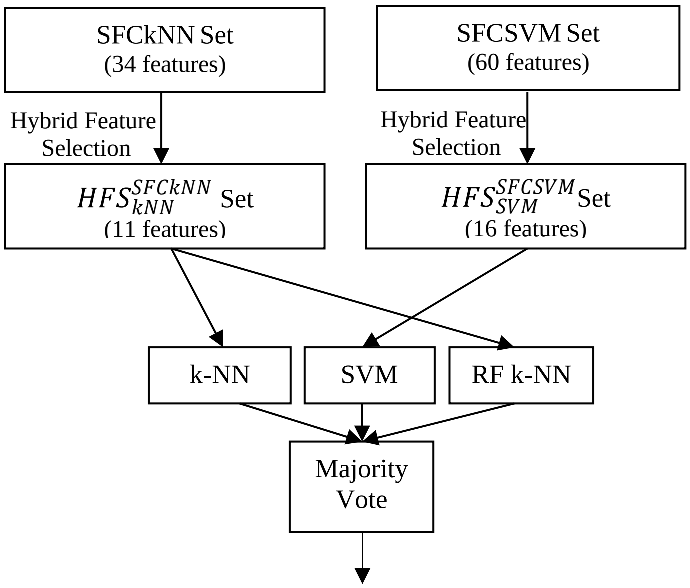


Figure 7.7 Flow diagram for sub study 3(i)(c)

Sub-study 3(ii) In this sub-study, to demonstrate the effectiveness of proposed ensemble strategy, the proposed ensemble is compared with four strategies in which one component was removed. The proposed ensemble consists of four components: hierarchal feature fusion, SVM, k-NN, RF k-NN. Table 7.6 gives a brief description of the strategies for comparison. The accuracy, sensitivity, specificity and run time of these methods were analyzed.

Table 7.6 A brief description of the strategies for comparison

Strategy	Hierarchal feature fusion	SVM	k-NN	RF k-NN
Proposed ensemble	√	√	√	√
Without fusion	-	√	√	√
Without SVM	√	-	√	√
Without k-NN	√	√	-	√
Without RF k-NN	√	√	√	-

Sub-study 3(iii) In this sub-study, proposed ensemble classifier model was compared with common classifier models mentioned in literature, including individual classifiers, i.e. k-NN, SVM, and ensemble classifiers, i.e. Adaboost with k-NN as base classifier and Adaboost with SVM as base classifier, to verify the performance of proposed ensemble model.

***Results and discussion*:**

Sub-study 3(i) The table 7.7 presents the classification accuracy results of ensemble classifier models as proposed in Figures 7.5 – 7.7 on original1 dataset. The classification accuracy obtained by ensemble models of sub-study 3(i)(a)/3(i)(b)/3(i)(c) were 96.2/96.1/96.6%. The results from table 7.3 and 7.7 indicate that classification accuracy of all the three ensembles improved relative to individual classifiers to some degree. The best classification accuracy of 96.6% was obtained with sub-study 3(i)(c). Thus, an ensemble model that is designed with voting algorithm in conjunction with three classifiers, namely k-NN, SVM and RF k-NN, and hierarchical feature fusion set as input is the best proposed ensemble model.

Table 7.7 The classification results of ensemble models on original1 dataset

	Accuracy (%)
Sub study 3 (i)(a)	96.2
Sub study 3 (i)(b)	96.1
Sub study 3 (i)(c)	96.6

Sub-study 3(ii) Table 7.8 shows the experimental results of various ensemble strategies presented in table 7.6. The classification results of the proposed ensemble strategy show a high precision in predicting the test data to its actual liver tissue class with an accuracy of 96.6%. With regards to accuracy, the performance of proposed ensemble strategy was on average 1.7% higher than that of the compared strategies. The experimental results also indicated the effectiveness of the proposed strategy with the overall average sensitivity of 96.5% revealing its ability to perform well in predicting the true positive cases in each liver tissue type correctly. The outperformance of the proposed strategy was also observed with the overall average specificity of 98.8% revealing its ability to perform well in predicting the true negative cases in each liver tissue type correctly. Here, it can also be observed that the run time without hierarchal feature fusion is highest i.e. 0.767s and with hierarchal feature fusion is 0.265s. This indicates that the inclusion of hierarchal feature fusion reduces the run time from 0.767s to 0.265s that is 34% reduction in run time. Hence, it can be emphasized that the results indicate the capability of the proposed ensemble strategy to differentiate among four liver classes and all the components i.e. hierarchal feature fusion, SVM, k-NN and RF k-NN are required.

Table 7.8 The classification results of different strategies on original1 dataset

Strategy	Accuracy (%)	Sensitivity (%)				Specificity (%)				Run Time
		Sen_{NOR}	Sen_{CHL}	Sen_{CIR}	Sen_{HCC}	Spe_{NOR}	Spe_{CHL}	Spe_{CIR}	Spe_{HCC}	(s)
Without Hierarchal fusion	95.2	94.2	98.5	94.0	96.2	99.2	97.6	98.0	99.3	0.767
Without k-NN	94.9	96.3	92.5	98.0	91.9	99.2	98.9	94.9	99.7	0.224
Without RF k-NN	94.7	95.8	92.5	98.0	91.9	99.4	98.5	95.1	99.2	0.023
Without SVM	95.9	99.4	94.0	95.0	95.1	99.2	97.2	98.1	99.6	0.212
Proposed ensemble	96.6	96.3	95.5	97.5	96.9	99.2	98.0	98.2	99.8	0.265

Sub-study 3(iii) Table 7.9 provides the classification results of the proposed ensemble classifier model and other classifier models on original1dataset. The classification results of the proposed ensemble model show a high accuracy in classifying the test data to its actual liver tissue class of 96.6%. With regards to accuracy and sensitivity, the performance of proposed ensemble model was on average 3.1% and 2.8% higher than that of the compared classifier models. Further, the model proves its capability of distinguishing four liver tissue types with the sensitivity and specificity of 96.3%/95.5%/97.5%/96.9% and 99.2%/98.0%/98.2%/99.8% on normal/chronic/cirrhosis/HCC respectively. Finally, it can be emphasized that the results of the proposed ensemble model were better as compared to other classification models.

Table 7.9 The classification results of different classifiers on original1 dataset

	Accuracy (%)	Sensitivity (%)				Specificity (%)			
		Sen_{NOR}	Sen_{CHL}	Sen_{CIR}	Sen_{HCC}	Spe_{NOR}	Spe_{CHL}	Spe_{CIR}	Spe_{HCC}
SVM	92.3	86.9	94.5	98.0	90.7	99.8	98.9	91.8	99.4
k-NN	95.2	93.7	93.5	98.5	95.6	99.4	98.3	96.7	99.1
Adaboost with k-NN	94.2	95.8	92.5	94.0	93.8	97.3	97.8	97.4	99.3
Adaboost with SVM	93.1	87.5	95.0	98.0	91.3	99.4	99.0	92.5	99.4
Proposed ensemble	96.6	96.3	95.5	97.5	96.9	99.2	98.0	98.2	99.8

7.2.4 Conclusion of Study 1

A CAD system designed by using an ensemble model and handcrafted features for classifying liver diseases, that is normal, chronic liver, cirrhosis, and HCC was proposed. The ensemble classifier that was designed with voting algorithm in conjunction with three classifiers, namely k-NN, SVM and RF k-NN, and hierarchical feature fusion set was the best proposed model having sensitivity for normal/chronic/cirrhosis/HCC was 96.3/95.5/97.5/96.9% and specificity for normal/chronic/cirrhosis/HCC was 99.2/98.0/98.2/99.8%. Finally, CAD system designed with handcrafted features and proposed ensemble model with original1 dataset gave 96.6% accuracy which show high capability of the system in predicting the test data to its actual liver tissue class.

7.3 STUDY 2: DESIGNING OF CAD SYSTEM USING DEEP LEARNED FEATURES WITH ORIGINAL IMAGES

Aim: Design a CAD system by using deep learned features extracted from original images.

Method: In this study, features were extracted from original2 dataset by using two CNN models i.e. pre-trained CNN and fine-tuned CNN. Then, deep learned features extracted from each model were passed as input to three classifiers i.e. k-NN, SVM and proposed ensemble model, and classification accuracies were calculated.

In this study, AlexNet pre-trained model was used as a feature extractor network. As the outputs of each convolutional, pooling, and fully connected layer can be extracted as features, there were 11 layers from which features could be extracted and used as input of classifiers. It was not known which layer of AlexNet in particular would be best suitable for classification of liver ultrasound images. Thus, classifiers were trained based on features of each layer to determine the optimal layer. Features from each layer were extracted and used to train SVM, k-NN, and proposed ensemble classifier. The optimal layer was then selected based on predictive performance.

AlexNet model was also fine-tuned on original2 dataset. All the layers were extracted, except the last three, from the pretrained AlexNet model. These layers were transferred to the present four-class classification problem and the remaining last three layers were replaced with a new fully connected layer, a softmax layer, and a classification output layer to fine-tune for present classification problem. The new fully connected layer has now four classes. The initial weights of the model were derived from ImageNet dataset and then fine-tuned (optimized) through back-propagation so that they better reflect the four liver classes. The network was trained for 30 epochs because the error rates for training had begun to plateau at 30 epochs with minimal improvements even when trained for more epochs. Next, the iteration number was set to 60 and the min-batch stochastic gradient descent with the batch size of 6. A learning rate was set as 0.01 and was further reduced by one-tenth every 10 epoch, this allowed the fine-tuning process to effectively learn the filter weights. Here momentum value is set to 0.9 for all epochs as pretrained CNN is used that had already been trained on large image dataset. The weight decay act as a regularization term for the gradient descent by preventing the weights from growing too large; it was also important for avoiding overfitting. Here, weight decay value of 1 was used.

***Results and discussion*:** The classification accuracy obtained with SVM, k-NN and proposed ensemble classifier by using deep learned features from all the pre-trained layers of AlexNet model from original2 dataset are presented in Figure 7.8. The 'fc7' layer was selected due to its high predictive performance in classifying four liver tissue types. The classification accuracy of SVM, k-NN, and proposed ensemble model trained and evaluated on extracted features from 'fc7' layer of pretrained CNN was 65.1%/66.3%/68.4%. The classification accuracy of SVM, k-NN, and proposed ensemble classifier model trained and evaluated on extracted features from fine-tuned CNN was 74.1%/71.3%/75.4%. The results indicate that by using the features of the fine-tuned network with proposed ensemble model provided the better results.

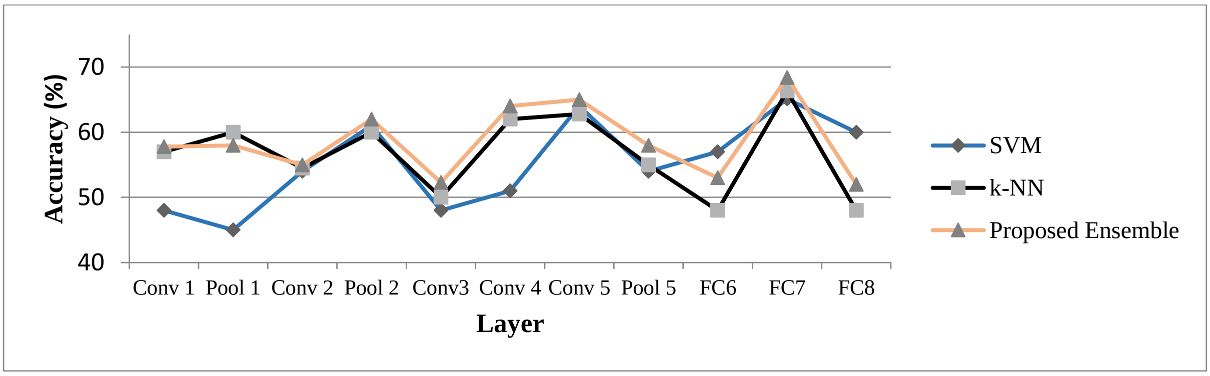


Figure 7.8 Performance in terms of accuracy of classifiers based on features from each layer of AlexNet in discriminating four liver stages

Conclusion of study 2

A CAD system was designed by using deep learned features to classify ultrasound liver images into four liver classes. Although, the classification accuracy was not high; but it was observed that the knowledge transfer from natural images to medical images is possible. Fine-tuning with transfer learning yielded better results compared to pre-trained CNN approach.

7.4 CONCUSION OF STUDY 1 AND STUDY 2

Results of study 1 outperformed results of study2. It leads to the finding that the CAD system designed with handcrafted features is more valuable than that with deep learning features in our research problem. Finally, to improve the diagnostic performance of CAD, study 3 was performed with enhanced dataset utilizing the best outcome of previous studies, i.e., handcrafted features and ensemble model.

7.5 STUDY 3: DESIGNING OF CAD SYSTEM WITH ENHANCED IMAGES

Aim: Design a CAD system with enhanced dataset by extracting the handcrafted features and using proposed ensemble model.

Method: In this study, ultrasound liver images were enhanced by using proposed image enhancement method, as presented in chapter 4. A CAD system was then designed by using handcrafted features extracted from enhanced dataset with proposed ensemble model.

Results and discussion: Figure 7.9 show two sample images with different liver diseases and their enhanced counterparts. Figure 7.9(a) and 7.9(c) are original ultrasound liver images and figure 7.9(b) and 7.9(d) are corresponding enhanced images. The marked regions of enhanced images have better appearance of texture in comparison to the original images, as stated by radiologist. The detailed analysis based on subjective and objective criteria is presented in chapter 4.

(a) Original Image (b) Enhanced Image (c) Original Image (d) Enhanced Image

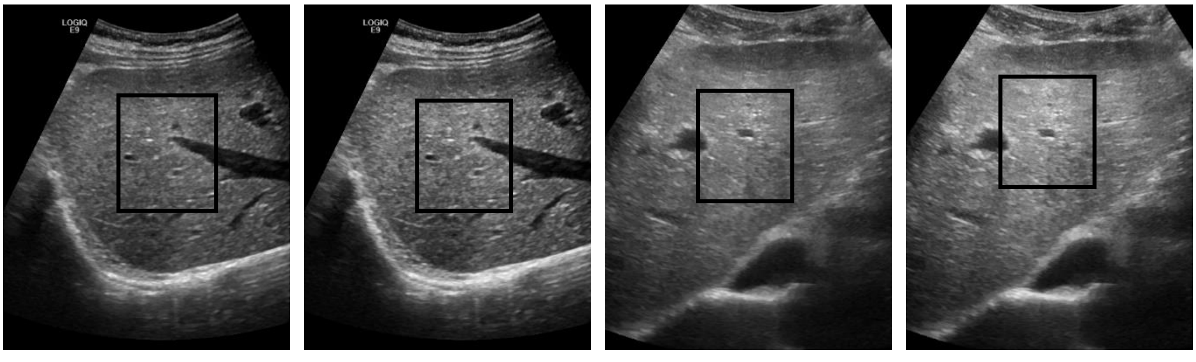


Figure 7.9 Original and enhanced ultrasound images of liver

In summary, with subjective evaluation, with proposed method, overall best score of 4.3 was obtained and that was 44% higher than the score of original images. Radiologist also mentioned that there was better visibility of diagnostic features in enhanced image. This result was also supported by objective criteria. The EPI value was 0.9835 which represents the edge preservation in the processed images. The AMBE value was 0.03, which is very close to 0.0 and reflects the preservation of brightness in the processed image. The UIQI value was 0.9777, which is very close to 1 and depicts the close representation of processed image w.r.t. original image in terms of correlation, luminance and contrast. Finally, MS-SSIM value was 0.9996 which is very close to 1 and portrays better image quality.

The results presented in table 7.10 show that the CAD designed in study 3 with enhanced dataset obtained the classification accuracy of 97.9% for classifying liver diseases. The overall average specificity of 99.1% reveals system's ability to perform well in predicting the true negative cases in each liver tissue type correctly. The overall average sensitivity of 97.4% reveals system's ability to perform well in predicting the true positive cases in each liver tissue type correctly. Further, the model proves its capability in distinguishing liver tissue classes with the sensitivity, specificity, NPV and PPV of 97.9/96.5/98/97.5%, 99.2/99.2/98.5/99.8%, 99.2/98.7/99.2/99.3 and 98.4/96.5/96.0/99.3% on normal/chronic/cirrhosis/HCC respectively.

Table 7.10 The classification results of studies with handcrafted features

	Accuracy (%)	**Sensitivity (%)**				**Specificity (%)**			
		Sen_{NOR}	Sen_{CHL}	Sen_{CIR}	Sen_{HCC}	Spe_{NOR}	Spe_{CHL}	Spe_{CIR}	Spe_{HCC}
CAD study1	96.6	96.3	95.5	97.5	96.9	99.2	98.0	98.2	99.8
CAD study3	**97.9**	**97.9**	**96.5**	**98**	**97.5**	**99.2**	**99.2**	**98.5**	**99.8**

The classification results of the proposed ensemble model using handcrafted features extracted from enhanced images, show a high precision in predicting the test data to its actual liver tissue class with an accuracy of 97.9%. It was observed that there was 1.3% increase in accuracy with respect to CAD system designed with original1 dataset. Finally, it can be emphasized that the good results indicate the effectiveness of the proposed CAD system with enhanced images.

7.6 CONCLUDING REMARKS

In this chapter, three studies were performed to find out best possible design of CAD system for our research problem. In the first study investigations are conducted to evaluate (i) performance of proposed GLDM extensions, (ii) the effectives of handcrafted features, (iii) proposed hybrid feature selection method and (iv) proposed ensemble classifier. In the second study, CAD system was designed with deep learned features to classify liver diseases. In third study, CAD system was designed with enhanced images by extracting handcrafted features. Finally, the CAD system designed with handcrafted features, proposed ensemble model and enhanced images gave promising results demonstrating the effectiveness of system in assisting radiologist in diagnosis of liver tissue type.

Conclusions

This chapter summarizes the conclusions drawn from the conducted experiments in the process of optimizing the designing of computer-aided system to diagnose liver diseases. These experiments were conducted to diagnose liver diseases, i.e. chronic, cirrhosis and HCC using liver ultrasound images. The conclusions drawn in different sections of this research work, i.e. ultrasound image enhancement, feature extraction and classification, are mentioned here. The chapter is ended with mentioning the future scopes in this research.

8.1 ULTRASOUND IMAGE ENHANCEMENT

- An ultrasound image enhancement method based on neutrosophic similarity score (NSS) was proposed. The NSS was calculated over multi-criteria and this score is used to scale the image. The NSS measures the belonging degree of pixel to the texture. The multi-criteria i.e. intensity criterion, local mean intensity criterion and edge detection criterion aided in enhancing images. The method improved contrast, removed noise and enhanced edges of the ultrasound liver images.
- The quality of enhanced images was evaluated both by subjective and objective criteria. The subjective evaluation was done by the expert radiologist. It was observed that the processed images provide better visual appearance in comparison to the original liver images in terms of enhanced contrast and reduced noise. In subjective evaluation, with template 3 and template 5, overall best score of 4.3 and 3.8 out of 5 was obtained and that were 44% and 27% higher than the score of original images. Additionally, the objective evaluators such as EPI, AMBE, UIQI and MS-SSIM were used and the results of subjective evaluator were supported by objective measures. Based on these results, it was concluded that the proposed enhancement method is suitable for extraction of diagnostic features as required in the designing of computer-aided systems for liver disease classification.

- The improvement in visual quality of processed images was obtained by both the proposed templates 3 and 5. Out of these proposed templates, template 3 gave the best processed images in terms of noise removal and texture enhancement. It was observed that improvements in visual scores and the values of objective criteria of template 5 are lesser than those of template 3. Additionally, according to the radiologist there was bit of smoothing in diagnostic features in the enhanced images, with template 5. Thus, it was concluded that the template 3 method denoises and increases contrast without blurring or smoothing texture features thereby providing more fine details of image to be visualized clearly.

8.2 FEATURE EXTRACTION

- Extension of GLDM was proposed in this work to obtain more precise feature values than the traditional GLDM method and thereby providing comparatively accurate information about the roughness of liver surface. The results with extended GLDM feature set show 12% and 16% increase in classification accuracy with k-NN and SVM, respectively. Thus, it was concluded that the features extracted with extended GLDM method were better than traditional GLDM method and can be used to build a computationally efficient CAD system for predicting liver diseases.
- A set of 148 features were obtained from GLDM (64 features), GLCM (48 features), and features extracted from *ranklets* (36 features). Results show that extracted texture features are effective, as the classification accuracy by individual handcrafted feature extraction methods (GLDM, GLCM, *ranklets*) was close to 90%. Extracted features from GLDM method represent local derivative information of gray-levels and provide information about the roughness of liver surface. GLCM measures the statistical properties, visual characteristics, and correlation-based information of gray levels, that

help characterize different stages of liver. In addition, the multiresolution texture features obtained by using *ranklets* are less sensitive towards different ultrasonic devices so these features were used in designing of CAD system.

- After the extraction of all handcrafted features, optimal feature set was obtained by proposed hybrid feature selection method. To take benefit of complimentary information from different feature sets, feature fusion schemes were also implemented. The proposed hybrid feature method outperformed in terms of time taken to select features, number of features and accuracy. It was observed that feature fusion results improved classification accuracy as compared to individual feature sets. The hierarchical feature fusion set was the optimal feature set with maximum accuracy.
- Deep learned features were used to classify ultrasound liver images into four liver classes. Although the classification accuracy was not high, it was observed that the knowledge transfer from natural images to medical images is possible. Fine-tuning with transfer learning yielded better results as compared to pre-trained CNN approach. However, the CNN specifically trained for particular tasks is expected to outperform if the given database is large.

8.3 COMPUTER-AIDED DIAGNOSIS SYSTEMS

- An ensemble classifier model for classifying liver ultrasound images into four stages, that is, normal, chronic liver, cirrhosis, and HCC was proposed. Ensemble model was designed, by combining k-NN, SVM, and RF k-NN to increase diversity and consistency. The proposed ensemble model was evaluated by giving different inputs to each classifier (i) individual selected feature sets, (ii) serial feature combination set of selected features, and (iii) hierarchical feature fusion set. The ensemble classifier that was designed with voting algorithm in conjunction with three classifiers, namely k-NN, SVM and RF k-NN, and hierarchical feature fusion set as input was the best proposed model with original1 dataset, having classification accuracy of 96.6%, sensitivity of 96.3/95.5/97.5/96.9% for normal/chronic/cirrhosis/HCC and specificity of 99.2/98.0/98.2/99.8% for normal/chronic/cirrhosis/HCC.

- A CAD system was also designed by using deep learned features. The classification accuracy was calculated with ensemble classifier from all the pre-trained layers of AlexNet model. The “Fc7” layer was selected due to its high predictive performance as observed from Figure 7.8. The classification accuracy of proposed ensemble model on extracted features from pretrained CNN and fine-tuned CNN was 68.4% and 75.4%, respectively.
- Additionally, a CAD system was designed with handcrafted features and proposed ensemble classification method with enhanced dataset. The CAD system gave 97.9% accuracy, that show the high precision capability of the system in predicting the test data to its actual liver tissue class. The experiment results also indicate the high performance of this proposed CAD system with the overall average sensitivity of 97.4%, revealing its ability to perform well in predicting the true positive cases in each liver tissue type correctly. Sensitivity, specificity, NPV and PPV of normal/chronic/cirrhosis/HCC were 97.9/96.5/98/97.5%, 99.2/99.2/98.5/99.8%, 99.2/98.7/99.2/99.3 and 98.4/96.5/96.0/99.3% respectively. It was observed that there was 1.3% increase in accuracy with the CAD system designed using enhanced dataset in comparison to the CAD system designed with original1 dataset.

Finally, it is concluded that the best proposed CAD system comprises of (i) enhanced ultrasound images, (ii) texture based on three handcrafted feature extraction methods, (iii) hybrid feature selection and hierarchal feature fusion set, and (iv) ensemble classifier. This system obtained overall classification accuracy of 97.9%. The substantial performance of experiments on clinical database supports the strong candidature of the proposed computer-aided system to assist radiologist/clinicians in the diagnosis of liver diseases.

8.4 LIMITATIONS AND SCOPE FOR FUTURE WORK

- Currently, the work focuses on categorization of normal, chronic, cirrhotic and HCC. This work can be further extended to more liver diseases and severity of stages.
- In the present work the ROIs are identified with manual interaction. In future work automatic segmentation can be implemented to identify the ROIs from ultrasound images.
- In CNN features fine-tuning with transfer learning gave the best results. It is important to note that using a pretrained CNN restricts us to use its original architecture. As, we have to fit our data to an architecture that is likely suboptimal for medical image data. Thus, more labelled medical data is required to take the advantage of CNN. Additionally, research can be carried forward with deeper neural networks.
- In this work, only liver images were used to develop CAD system. Practically, experts consider other information such as biomarkers and spleen and kidney appearance along with liver images to draw conclusions. Hence, the CAD system can be improved further by blending biomarkers and spleen and kidney information along with liver images.
- Presently, to evaluate CAD system for liver disease, there is no publicly available database. A large publicly available database would attract even more research groups into this field and help develop a CAD system that combines multiple approaches and outperform the available algorithms.

References

[1] S. S. Davidson, *Davidson ' s Medicine*. 2010.

[2] S. K. Asrani, H. Devarbhavi, J. Eaton, and P. S. Kamath, "Burden of liver diseases in the world," *J. Hepatol.*, vol. 70, no. 1, pp. 151–171, Jan. 2019.

[3] "Liver Disease in India." .

Available: https://www.worldlifeexpectancy.com/india-liver-disease.

[Accessed: 23-Jul-2019].

[4] A. Baecker, X. Liu, C. La Vecchia, and Z.-F. Zhang, "Worldwide incidence of hepatocellular carcinoma cases attributable to major risk factors.," *Eur. J. Cancer Prev.*, vol. 27, no. 3, pp. 205–212, 2018.

[5] "World Hepatitis Day 2018," 2018.

Available: https://www.who.int/campaigns/world-hepatitis-day/2018.

[Accessed: 23-Jul-2019].

[6] "Hepatitis B," *World health organization*, 2019.

Available: https://www.who.int/news-room/fact-sheets/detail/hepatitis-b.

[Accessed: 23-Jul-2019].

[7] H. A. Saleh and A. H. Abu-Rashed, "Liver biopsy remains the gold standard for evaluation of chronic hepatitis and fibrosis," *J. Gastrointest. Liver Dis.*, vol. 16, no. 4, pp. 425–426, 2007.

[8] D. V. Sahani and S. P. Kalva, "Imaging the Liver," *Oncologist*, vol. 9, no. 4, pp. 385–397, Jul. 2004.

[9] E. K. Outwater, “Imaging of the liver for hepatocellular cancer.,” *Cancer Control*, vol. 17, no. 2, pp. 72–82, 2010.

[10] V. P. B. Grover, J. M. Tognarelli, M. M. E. Crossey, I. J. Cox, S. D. Taylor-Robinson, and M. J. W. McPhail, “Magnetic Resonance Imaging: Principles and Techniques: Lessons for Clinicians.,” *J. Clin. Exp. Hepatol.*, vol. 5, no. 3, pp. 246–55, Sep. 2015.

[11] P. J. Mergo, P. R. Ros, P. C. Buetow, and J. L. Buck, “Diffuse disease of the liver: radiologic-pathologic correlation.,” *Radiographics*, vol. 14, no. 6, pp. 1291–307, Nov. 1994.

[12] J. F. Gerstenmaier and R. N. Gibson, “Ultrasound in chronic liver disease.,” *Insights Imaging*, vol. 5, no. 4, pp. 441–55, Aug. 2014.

[13] H. Tchelepi, P. W. Ralls, R. Radin, and E. Grant, “Sonography of diffuse liver disease.,” *J. Ultrasound Med.*, vol. 21, no. 9, pp. 1023–32; quiz 1033–4, Sep. 2002.

[14] “Ultrasound abdomen general.”

Available: http://www.startradiology.com/internships/general-surgery/abdomen/ultrasound-abdomen-general/.

[Accessed: 23-Jul-2019].

[15] R. G. Dantas, E. T. Costa, and S. Leeman, “Ultrasound speckle and equivalent scatterers,” *Ultrasonics*, vol. 43, no. 6, pp. 405–420, May 2005.

[16] T. E. Schouten and C. Science, “Application of neural networks for the classification of diffuse liver disease by quantitative echography,” *Utrasonic Imaging*, vol. 217, pp. 205–217, 1993.

[17] D. Fuster *et al.*, “Liver ultrasound abnormalities in alcohol use disorder,” in *Essentials*

of Abdominal Ultrasound [Working Title], IntechOpen, 2019.

[18] E. C. F. Dietrich, D. O. Mirko, V. Marie-pierre, V. Vlastimil, and P. Francesco, “EFSUMB – European Course Book,” *Ultrasound*, pp. 1–22, 2010.

[19] T. Hennedige and S. K. Venkatesh, “Imaging of hepatocellular carcinoma: diagnosis, staging and treatment monitoring.,” *Cancer Imaging*, vol. 12, no. 3, pp. 530–47, Feb. 2013.

[20] P. Bharti, D. Mittal, and R. Ananthasivan, “Computer-aided characterization and diagnosis of diffuse liver diseases based on ultrasound imaging: A Review,” *Ultrason Imaging*, vol. 39, no. 1, pp. 33–61, Jan. 2016.

[21] M. Gangwar, R. B. Mishra, and R. S. Yadav, “Intelligent computing method for the interpretation of neuropsychiatric diseases,” *Int. J. Comput. Appl.*, vol. 55, no. 17, pp. 23–31, Oct. 2012.

[22] O. Faust *et al.*, “Comparative assessment of texture features for the identification of cancer in ultrasound images: a review,” *Biocybern. Biomed. Eng.*, vol. 38, no. 2, pp. 275–296, Jan. 2018.

[23] C. M. Chung-Ming Wu, Y. C. Yung-Chang Chen, K. S. Kai-Sheng Hsieh, C. M. Wu, Y. C. Chen, and K. S. Hsieh, “Texture features for classification of ultrasonic liver images,” *IEEE Trans. Med. Imaging*, vol. 11, no. 2, pp. 141–152, Jan. 1992.

[24] M. Masotti and R. Campanini, “Texture classification using invariant ranklet features,” *Pattern Recognit. Lett.*, vol. 29, no. 14, pp. 1980–1986, Oct. 2008.

[25] R. Haralick, K. Shanmugan, and I. Dinstein, “Textural features for image classification,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3. pp. 610–621, 1973.

[26] L. Hermes and J. M. Buhmann, “Feature selection for support vector machines,” in

Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, vol. 2, pp. 712–715.

[27] S. Khan and S. Yong, “A deep learning architecture for classifying medical images of anatomy object,” no. December, 2017.

[28] W.-L. Lee, “An ensemble-based data fusion approach for characterizing ultrasonic liver tissue,” *Appl. Soft Comput.*, vol. 13, no. 8, pp. 3683–3692, 2013.

[29] D. Mitrea *et al.*, “Abdominal tumor characterization and recognition using superior-order cooccurrence matrices, based on ultrasound images,” *Comput. Math. Methods Med.*, vol. 2012, pp. 1–17, 2012.

[30] A. Sboner *et al.*, “A multiple classifier system for early melanoma diagnosis,” *Artif. Intell. Med.*, vol. 27, no. 1, pp. 29–44, Jan. 2003.

[31] J. M. Thijssen, “Ultrasonic speckle formation, analysis and processing applied to tissue characterization,” *Pattern Recognit. Lett.*, vol. 24, no. 4–5, pp. 659–675, 2003.

[32] V. Shrimali, R. Anand, and V. Kumar, “Comparing the performance of ultrasonic liver image enhancement techniques: A preference study,” *IETE J. Res.*, vol. 56, no. 1, p. 4, 2010.

[33] Yongjian Yu and S. T. Acton, “Speckle reducing anisotropic diffusion,” *IEEE Trans. Image Process.*, vol. 11, no. 11, pp. 1260–1270, Nov. 2002.

[34] S. Gupta, L. Kaur, R. C. Chauhan, and S. C. Saxena, “A versatile technique for visual enhancement of medical ultrasound images,” *Digit. Signal Process. A Rev. J.*, vol. 17, no. 3, pp. 542–560, 2007.

[35] S. Poddar, D. Sharma, A. Ghosh, S. Tewary, V. Karar, and S. K. Pal, “Non-parametric modified histogram equalisation for contrast enhancement,” *IET Image Process.*, vol. 7,

no. 7, pp. 641–652, Oct. 2013.

[36] P. Perona and J. Malik, “Scale-space and edge detection using anisotropic diffusion,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.

[37] J. Yu, J. Tan, and Y. Wang, “Ultrasound speckle reduction by a SUSAN-controlled anisotropic diffusion method,” *Pattern Recognit.*, vol. 43, no. 9, pp. 3083–3092, Sep. 2010.

[38] N. Damodaran, S. Ramamurthy, S. Velusamy, and G. Kanakaraj Manickam, “Speckle noise reduction in ultrasound biomedical B-scan images using discrete topological derivative,” *Ultrasound Med. Biol.*, vol. 38, no. 2, pp. 276–286, Feb. 2012.

[39] K. Krissian, C.-F. Westin, R. Kikinis, and K. G. Vosburgh, “Oriented speckle reducing anisotropic diffusion,” *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1412–1424, May 2007.

[40] D. Mittal, V. Kumar, S. C. Saxena, N. Khandelwal, and N. Kalra, “Enhancement of the ultrasound images by modified anisotropic diffusion method,” *Med. Biol. Eng. Comput.*, vol. 48, no. 12, pp. 1281–1291, 2010.

[41] Y. Guo, Y. Wang, and T. Hou, “Speckle filtering of ultrasonic images using a modified non local-based algorithm,” *Biomed. Signal Process. Control*, vol. 6, no. 2, pp. 129–138, Apr. 2011.

[42] P. Coupe, P. Hellier, C. Kervrann, and C. Barillot, “Nonlocal means-based speckle filtering for ultrasound images,” *IEEE Trans. Image Process.*, vol. 18, no. 10, pp. 2221–2229, Oct. 2009.

[43] C. A. Deledalle, L. Denis, and F. Tupin, “Iterative weighted maximum likelihood denoising with probabilistic patch-based weights,” *IEEE Trans. Image Process.*, vol. 18,

no. 12, pp. 2661–2672, 2009.

[44] D. Gleich, M. Kseneman, and M. Datcu, “Despeckling of TerraSAR-X data using second-generation wavelets,” *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 68–72, Jan. 2010.

[45] B. B. Saevarsson, J. R. Sveinsson, and J. A. Benediktsson, “Speckle reduction of SAR images using adaptive curvelet domain,” in *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477)*, vol. 6, pp. 4083–4085.

[46] B. Hou, X. Zhang, X. Bu, and H. Feng, “SAR image despeckling based on nonsubsampled shearlet transform,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 5, no. 3, pp. 809–823, Jun. 2012.

[47] W. Zhang, F. Liu, L. Jiao, B. Hou, S. Wang, and R. Shang, “SAR image despeckling using edge detection and feature clustering in bandelet domain,” *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 131–135, Jan. 2010.

[48] G. G. Bhutada, R. S. Anand, and S. C. Saxena, “Edge preserved image enhancement using adaptive fusion of images denoised by wavelet and curvelet transform,” *Digit. Signal Process. A Rev. J.*, vol. 21, no. 1, pp. 118–130, 2011.

[49] L. Rudin, P.-L. Lions, and S. Osher, “Multiplicative denoising and deblurring: Theory and algorithms,” in *Geometric Level Set Methods in Imaging, Vision, and Graphics*, New York: Springer-Verlag, 2003, pp. 103–119.

[50] G. Aubert and J.-F. Aujol, “A variational approach to removing multiplicative noise,” *SIAM J. Appl. Math.*, vol. 68, no. 4, pp. 925–946, Jan. 2008.

[51] J. Shi and S. Osher, “A nonlinear inverse scale space method for a convex multiplicative

noise model," *SIAM J. Imaging Sci.*, vol. 1, no. 3, pp. 294–321, Jan. 2008.

[52] D. Koundal, "Computer-aided diagnosis of thyroid nodule: A review," *Int. J. Comput. Sci. Eng. Surv.*, vol. 3, no. 4, pp. 67–83, 2012.

[53] A. Keerthivasan, J. J. Babu, and G. F. Sudha, "Speckle noise reduction in ultrasound images using fuzzy logic based on histogram and directional differences," in *2013 International Conference on Communication and Signal Processing*, 2013, pp. 499–503.

[54] A. Bayani, M. Langarizadeh, A. R. Radmard, and A. F. Nejad, "Quality improvement of liver ultrasound images using fuzzy techniques.," *Acta Inform. Med.*, vol. 24, no. 6, pp. 380–384, Dec. 2016.

[55] Y. Guo *et al.*, "Breast ultrasound image enhancement using fuzzy logic," *Ultrasound Med. Biol.*, vol. 32, no. 2, pp. 237–247, Feb. 2006.

[56] D. Koundal, S. Gupta, and S. Singh, "Neutrosophic based nakagami total variation method for speckle suppression in thyroid ultrasound images," *Irbm*, vol. 39, no. 1, pp. 43–53, Feb. 2017.

[57] S. Singh, D. Koundal, and S. Gupta, "Speckle reduction method for thyroid ultrasound images in neutrosophic domain," *IET Image Process.*, vol. 10, no. 2, pp. 167–175, Feb. 2016.

[58] Y. Guo, H. D. Cheng, and Y. Zhang, "A new neutrosophic approach to image denoising," *New Math. Nat. Comput.*, vol. 05, no. 03, pp. 653–662, Nov. 2009.

[59] M. R. Faraji and Xiaojun Qi, "An effective neutrosophic set-based preprocessing method for face recognition," in *2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, 2013, pp. 1–4.

[60] X. Qi, B. Liu, and J. Xu, “A neutrosophic filter for high-density Salt and Pepper noise based on pixel-wise adaptive smoothing parameter,” *J. Vis. Commun. Image Represent.*, vol. 36, pp. 1–10, Apr. 2016.

[61] Y. Guo and A. Şengür, “A novel image edge detection algorithm based on neutrosophic set,” *Comput. Electr. Eng.*, vol. 40, no. 8, pp. 3–25, Nov. 2014.

[62] A. I. Shahin, K. M. Amin, A. A. Sharawi, and Y. Guo, “A novel enhancement technique for pathological microscopic image using neutrosophic similarity score scaling,” *Optik (Stuttg).*, vol. 161, pp. 84–97, May 2018.

[63] Y. Guo, A. Şengür, and J. Ye, “A novel image thresholding algorithm based on neutrosophic similarity score,” *Measurement*, vol. 58, pp. 175–186, Dec. 2014.

[64] V. Chalana and Y. Kim, “A methodology for evaluation of boundary detection algorithms on medical images,” *IEEE Trans. Med. Imaging*, vol. 16, no. 5, pp. 642–652, 1997.

[65] J. A. a Noble and D. Boukerroui, “Ultrasound image segmentation: A survey,” *IEEE Trans. Med. Imaging*, vol. 25, no. 8, pp. 987–1010, Aug. 2006.

[66] H. Takizawa, T. Suzuki, H. Kudo, and T. Okada, “Interactive segmentation of pancreases in abdominal computed tomography images and its evaluation based on segmentation accuracy and interaction costs,” *Biomed Res. Int.*, vol. 2017, pp. 1–8, Aug. 2017.

[67] M. Singh, S. Singh, and S. Gupta, “Investigations on ROI selection for liver classification,” in *2014 IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2014, pp. 1–6.

[68] S. A. Azaid, M. W. Fakhr, and A. F. A. Mohamed, “Automatic diagnosis of liver

diseases from ultrasound images," *2006 Int. Conf. Comput. Eng. Syst.*, pp. 313–319, 2006.

[69] E. Kyriacou, S. Pavlopoulosl, G. Konnis, D. Koutsourisl, P. Zounipoulis, and I. Theotokai, "Computer assisted characterization of diffused liver disease using image texture analysis techniques on B-scan images," pp. 1479–1483, 1998.

[70] H. Yoshida, D. D. Casalino, B. Keserci, A. Coskun, O. Ozturk, and A. Savranlar, "Wavelet-packet-based texture analysis for differentiation between benign and malignant liver tumours in ultrasound images.," *Phys. Med. Biol.*, vol. 48, no. 22, pp. 3735–3753, Nov. 2003.

[71] N. Sriraam, J. Roopa, M. Saranya, and M. Dhanalakshmi, "Performance evaluation of computer aided diagnostic tool (CAD) for detection of ultrasonic based liver disease," *J. Med. Syst.*, vol. 33, no. 4, pp. 267–274, 2009.

[72] S. Poonguzhali, B. Deepalakshmi, and G. Ravindran, "Optimal feature selection and automatic classification of abnormal masses in ultrasound liver images," *Proc. ICSCN 2007 Int. Conf. Signal Process. Commun. Netw.*, pp. 503–506, 2007.

[73] A. Ahmadian, A. Mostafa, M. Abolhassani, and Y. Salimpour, "A texture classification method for diffused liver diseases using Gabor wavelets.," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2, no. c, pp. 1567–1570, 2005.

[74] A. Krizhevsky and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," pp. 1–9.

[75] E. Kyriacou, S. Pavlopoulos, D. Koutsouris, K. Blekas, a Stafylopatis, and P. Zoumpoulis, "Fuzzy neural network based characterization of diffused liver diseases using image texture analysis techniques on ultrasonic images."

[76] S. Pavlopoulos, E. Kyriacou, D. Koutsouris, K. Blekas, a. Stafylopatis, and P. Zoumpoulis, “Fuzzy neural network-based texture analysis of ultrasonic images,” *IEEE Eng. Med. Biol. Mag.*, vol. 19, no. 1, 2000.

[77] M. H. Horng, “An ultrasonic image evaluation system for assessing the severity of chronic liver disease,” *Comput. Med. Imaging Graph.*, vol. 31, no. 7, pp. 485–491, 2007.

[78] W.-L. Lee, K.-S. Hsieh, Y.-C. Y.-C. Chen, and Y.-C. Y.-C. Chen, “A study of ultrasonic liver images classification with artificial neural networks based on fractal geometry and multiresolution analysis,” *Biomed. Eng. Appl. Basis Commun.*, vol. 16, no. 02, pp. 59–67, 2004.

[79] C. Vicas, M. Lupsor, R. Badea, and S. Nedevschi, “Usefulness of textural analysis as a tool for noninvasive liver fibrosis staging,” *J. Med. Ultrason.*, vol. 38, no. 3, pp. 105–117, 2011.

[80] D. Mitrea, S. Nedevschi, and R. Badea, “The role of the superior order GLCM and of the generalized cooccurrence matrices in the characterization and automatic diagnosis of the hepatocellular carcinoma, based on ultrasound images,” *Proc. - 2011 IEEE 7th Int. Conf. Intell. Comput. Commun. Process. ICCP 2011*, pp. 197–204, 2011.

[81] M. Singh, S. Singh, and S. Gupta, “An information fusion based method for liver classification using texture analysis of ultrasound images,” *Inf. Fusion*, vol. 19, pp. 91–96, 2013.

[82] K. Kalyan, B. Jakhia, R. D. Lele, M. Joshi, and A. Chowdhary, “Artificial neural network application in the diagnosis of disease conditions with liver ultrasound images,” *Adv. Bioinformatics*, vol. 2014, pp. 1–14, 2014.

[83] J. Santos, J. S. Silva, A. A. Santos, and P. Belo-Soares, “Detection of pathologic liver

using ultrasound images," *Biomed. Signal Process. Control*, vol. 14, pp. 248–255, 2014.

[84] A. Balodi, M. L. Dewal, R. S. Anand, and A. Rawat, "Texture based classification of the severity of mitral regurgitation," *Comput. Biol. Med.*, vol. 73, pp. 157–164, Jun. 2016.

[85] S. Wang, Z. Pan, W. Wei, X. Zhao, and G. Wang, "The recognition of cirrhotic liver ultrasonic images of multi-feature fusion based on BP_Adaboost neural network," Springer, Singapore, 2016, pp. 106–114.

[86] M. H. Horng, Y. N. Sun, and X. Z. Lin, "Texture feature coding method for classification of liver sonography," *Comput. Med. Imaging Graph.*, vol. 26, no. 1, pp. 33–42, 2002.

[87] Wen-Li Lee, Yung-Chang Chen, and Kai-Sheng Hsieh, "Ultrasonic liver tissues classification by fractal feature vector based on M -band wavelet transform," *IEEE Trans. Med. Imaging*, vol. 22, no. 3, pp. 382–392, 2003.

[88] Y. Huang, X. Han, X. Tian, Z. Zhao, J. Zhao, and D. Hao, "Texture analysis of ultrasonic liver images based on spatial domain methods," in *2010 3rd International Congress on Image and Signal Processing*, 2010, pp. 562–565.

[89] A. Andrade, J. S. Silva, J. Santos, and P. Belo-Soares, "Classifier approaches for liver steatosis using ultrasound images," *Procedia Technol.*, vol. 5, pp. 763–770, 2012.

[90] C. Wu, W. Lee, Y. Chen, and K. Hsieh, "Evolution-based hierarchical feature fusion for ultrasonic liver tissue characterization," vol. 17, no. 5, pp. 967–976, 2013.

[91] R. T. Ribeiro, R. T. Marinho, and J. M. Sanches, "Classification and staging of chronic liver disease from multimodal data," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 5, pp. 1336–1344, May 2013.

[92] J. Virmani, V. Kumar, N. Kalra, and N. Khandelwal, "SVM-based characterization of

liver ultrasound images using wavelet packet texture descriptors," *J. Digit. Imaging*, vol. 26, no. 3, pp. 530–543, 2013.

[93] K. Raghesh Krishnan and S. Radhakrishnan, "Hybrid approach to classification of focal and diffused liver disorders using ultrasound images with wavelets and texture features," *IET Image Process.*, vol. 11, no. 7, pp. 530–538, Jul. 2017.

[94] X. Xi *et al.*, "Robust texture analysis of multi-modal images using local structure preserving ranklet and multi-task learning for breast tumor diagnosis," *Neurocomputing*, vol. 259, pp. 210–218, Oct. 2017.

[95] B. Francesco, "Robust color texture features based on ranklets and discrete Fourier transform," *J. Electron. Imaging*, vol. 18, no. 4, p. 043012, Oct. 2009.

[96] M.C. Yang *et al.*, "Robust texture analysis using multi-resolution gray-scale invariant features for breast sonographic tumor diagnosis," *IEEE Trans. Med. Imaging*, vol. 32, no. 12, pp. 2262–2273, Dec. 2013.

[97] L. Cai, X. Wang, Y. Wang, Y. Guo, J. Yu, and Y. Wang, "Robust phase-based texture descriptor for classification of breast ultrasound images," *Biomed. Eng. Online*, vol. 14, p. 26, Mar. 2015.

[98] V. Bolón-Canedo, N. Sánchez-Maroño, and A. Alonso-Betanzos, "A review of feature selection methods on synthetic data," *Knowl. Inf. Syst.*, vol. 34, no. 3, pp. 483–519, 2013.

[99] V. Sharma and K. C. Juglan, "Automated classification of fatty and normal liver ultrasound images based on mutual information feature selection," *IRBM*, vol. 39, no. 5, pp. 313–323, Nov. 2018.

[100] R. Sheikhpour, M. A. Sarram, S. Gharaghani, and M. A. Z. Chahooki, "A survey on semi-supervised feature selection methods," *Pattern Recognit.*, vol. 64, no. November

2016, pp. 141–158, 2017.

[101] W. Wu *et al.*, “Exploratory study to identify radiomics classifiers for lung cancer histology.,” *Front. Oncol.*, vol. 6, p. 71, 2016.

[102] V. Bolón-Canedo, N. Sánchez-Maroño, and A. Alonso-Betanzos, “Feature selection for high-dimensional data,” *Prog. Artif. Intell.*, vol. 5, no. 2, pp. 65–75, May 2016.

[103] W. Lu, Z. Li, and J. Chu, “A novel computer-aided diagnosis system for breast MRI based on feature selection and ensemble learning,” *Comput. Biol. Med.*, vol. 83, no. March, pp. 157–165, 2017.

[104] P. Pudil, J. Novovičová, and J. Kittler, “Floating search methods in feature selection,” *Pattern Recognit. Lett.*, vol. 15, no. 11, pp. 1119–1125, Nov. 1994.

[105] M. Bensch, M. Schröder, M. Bogdan, and W. Rosenstiel, “Feature selection for high-dimensional industrial data.,” *Esann*, pp. 375–380, 2005.

[106] S. Solorio-Fernández, J. A. Carrasco-Ochoa, and J. F. Martínez-Trinidad, “A new hybrid filter–wrapper feature selection method for clustering based on ranking,” *Neurocomputing*, vol. 214, pp. 866–880, 2016.

[107] K.-M. Osei-Bryson, K. Giles, and B. Kositanurit, “Exploration of a hybrid feature selection algorithm,” *J. Oper. Res. Soc.*, vol. 54, no. 7, pp. 790–797, Jul. 2003.

[108] Jinjie Huang, Yunze Cai, and Xiaoming Xu, “A wrapper for feature selection based on mutual information,” in *18th International Conference on Pattern Recognition (ICPR’06)*, 2006, pp. 618–621.

[109] K. M. Shazzad and Jong Sou Park, “Optimization of intrusion detection through fast hybrid feature selection,” in *Sixth International Conference on Parallel and Distributed Computing Applications and Technologies (PDCAT’05)*, 2005, pp. 264–267.

[110] Y. Zhang, C. Ding, and T. Li, "Gene selection algorithm by combining reliefF and mRMR.," *BMC Genomics*, vol. 9 Suppl 2, no. Suppl 2, p. S27, Sep. 2008.

[111] J. Yang, J. Yang, D. Zhang, and J. Lu, "Feature fusion: parallel strategy vs. serial strategy," *Pattern Recognit.*, vol. 36, no. 6, pp. 1369–1381, 2003.

[112] U. Mangai, S. Samanta, S. Das, and P. Chowdhury, "A survey of decision fusion and feature fusion strategies for pattern classification," *IETE Tech. Rev.*, vol. 27, no. 4, p. 293, 2010.

[113] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Comput. Electr. Eng.*, vol. 40, no. 1, pp. 16–28, 2014.

[114] H.-C. Shin *et al.*, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning.," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–98, 2016.

[115] A. Dubrovina, P. Kisilev, B. Ginsburg, S. Hashoul, and R. Kimmel, "Computational mammography using deep neural networks," *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.*, vol. 6, no. 3, pp. 243–247, May 2018.

[116] U. R. Acharya, H. Fujita, O. S. Lih, M. Adam, J. H. Tan, and C. K. Chua, "Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network," *Knowledge-Based Syst.*, vol. 132, pp. 62–71, Sep. 2017.

[117] 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*, pp. 1–1, 2017.

[118] K. Dmitriev *et al.*, "Classification of pancreatic cysts in computed tomography images using a random forest and convolutional neural network ensemble," Springer, Cham,

2017, pp. 150–158.

[119] N. Tajbakhsh and K. Suzuki, “A comparative study of modern machine learning approaches for focal lesion detection and classification in medical images: BoVW, CNN and MTANN,” Springer, Cham, 2018, pp. 31–58.

[120] G. Litjens *et al.*, “A survey on deep learning in medical image analysis.,” *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[121] N. Tajbakhsh *et al.*, “Convolutional neural networks for medical image analysis: full training or fine tuning?,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1299–1312, May 2016.

[122] F. Özyurt, T. Tuncer, E. Avci, M. Koç, and İ. Serhatlioğlu, “A novel liver image classification method using perceptual hash-based convolutional neural network,” *Arab. J. Sci. Eng.*, vol. 44, no. 4, pp. 3173–3182, Apr. 2019.

[123] B. Q. Huynh, H. Li, and M. L. Giger, “Digital mammographic tumor classification using transfer learning from deep convolutional neural networks,” *J. Med. Imaging*, vol. 3, no. 3, p. 034501, Aug. 2016.

[124] T. Liu, S. Xie, J. Yu, L. Niu, and W. Sun, “Classification of thyroid nodules in ultrasound images using deep model based transfer learning and hybrid features,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 919–923.

[125] E. Ribeiro *et al.*, “Exploring texture transfer learning for colonic polyp classification via convolutional neural networks,” in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017, pp. 1044–1048.

[126] P. M. Cheng and H. S. Malhi, “Transfer learning with convolutional neural networks for

classification of abdominal ultrasound images," *J. Digit. Imaging*, vol. 30, no. 2, pp. 234–243, Apr. 2017.

[127] Y. Yu *et al.*, "Deep learning enables automated scoring of liver fibrosis stages.," *Sci. Rep.*, vol. 8, no. 1, p. 16016, Oct. 2018.

[128] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," Nov. 2013.

[129] F. Zhang, B. Du, and L. Zhang, "Saliency-guided unsupervised feature learning for scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2175–2184, Apr. 2015.

[130] K. Nogueira, O. A. B. Penatti, and J. A. dos Santos, "Towards better exploiting convolutional neural networks for remote sensing scene classification," *Pattern Recognit.*, vol. 61, pp. 539–556, Jan. 2017.

[131] C. Boufenar, A. Kerboua, and M. Batouche, "Investigation on deep learning for off-line handwritten Arabic character recognition," *Cogn. Syst. Res.*, vol. 50, pp. 180–195, Aug. 2018.

[132] A. Kumar, J. Kim, D. Lyndon, M. Fulham, and D. Feng, "An ensemble of fine-tuned convolutional neural networks for medical image classification," *IEEE J. Biomed. Heal. Informatics*, vol. 21, no. 1, pp. 31–40, Jan. 2017.

[133] T. Kooi, B. van Ginneken, N. Karssemeijer, and A. den Heeten, "Discriminating solitary cysts from soft tissue lesions in mammography using a pretrained deep convolutional neural network," *Med. Phys.*, vol. 44, no. 3, pp. 1017–1027, Mar. 2017.

[134] J. Zhang, Y. Xia, Y. Xie, M. Fulham, and D. Feng, "Classification of medical images in the biomedical literature by jointly using deep and handcrafted visual features," *IEEE J.*

Biomed. Heal. Informatics, pp. 1–1, 2017.

[135] N. Antropova, H. Abe, and M. L. Giger, “Use of clinical MRI maximum intensity projections for improved breast lesion classification with deep convolutional neural networks,” *J. Med. Imaging*, vol. 5, no. 01, p. 1, Feb. 2018.

[136] Q. Huang, F. Zhang, and X. Li, “Machine learning in ultrasound computer-aided diagnostic systems: A survey,” *Biomed Res. Int.*, vol. 2018, pp. 1–10, Mar. 2018.

[137] L.-Q. Zhou *et al.*, “Artificial intelligence in medical imaging of the liver.,” *World J. Gastroenterol.*, vol. 25, no. 6, pp. 672–682, Feb. 2019.

[138] V. N. Vapnik and Vladimir, *Estimation of dependences based on empirical data*. Springer-Verlag, 1982.

[139] R. Gholami and N. Fakhari, *Support Vector Machine: Principles, Parameters, and Applications*, 1st ed. Elsevier Inc., 2017.

[140] V. Kecman, “Support Vector Machines – An Introduction,” 2005, pp. 1–47.

[141] M. Pal and G. M. Foody, “Feature selection for classification of hyperspectral data by SVM,” *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2297–2307, May 2010.

[142] G. Cao, P. Shi, and B. Hu, “Liver fibrosis identification based on ultrasound images captured under varied imaging protocols.,” *J. Zhejiang Univ. Sci. B*, vol. 6, no. 11, pp. 1107–1114, 2005.

[143] G. Cao, P. Shi, and B. Hu, “Ultrasonic liver discrimination using 2-D phase congruency,” *IEEE Trans. Biomed. Eng.*, vol. 53, no. 10, pp. 2116–2119, 2006.

[144] C.-C. Wu, W.-L. Lee, Y.-C. Chen, C.-H. Lai, and K.-S. Hsieh, “Ultrasonic liver tissue characterization by feature fusion,” *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9389–9397,

2012.

[145] S. Moldovanu, L. Moraru, and D. Bibicu, "Computerized decision support in liver steatosis investigation," vol. 6, no. 1, 2012.

[146] M. Singh, S. Singh, and S. Gupta, "A new quantitative metric for liver classification from ultrasound images," *Int. J. Comput. Electr. Eng.*, vol. 4, no. 4, pp. 605–607, 2012.

[147] F. U. A. A. Minhas, D. Sabih, and M. Hussain, "Automated classification of liver disorders using ultrasound images," *J. Med. Syst.*, vol. 36, no. 5, pp. 3163–3172, 2012.

[148] M. Woźniak and M. Graña, "A survey of multiple classifier systems as hybrid systems," *Inf. Fusion*, vol. 16, pp. 3–17, Mar. 2014.

[149] A. Takemura, A. Shimizu, and K. Hamamoto, "Discrimination of breast tumors in ultrasonic images using an ensemble classifier based on the AdaBoost algorithm with feature selection," *IEEE Trans. Med. Imaging*, vol. 29, no. 3, pp. 598–609, Mar. 2010.

[150] C. Catal and M. Nangir, "A sentiment classification model based on multiple classifiers," *Appl. Soft Comput.*, vol. 50, pp. 135–141, Jan. 2017.

[151] S. G. Mougiakakou, I. K. Valavanis, A. Nikita, and K. S. Nikita, "Differential diagnosis of CT focal liver lesions using texture features, feature selection and ensemble driven classifiers," *Artif. Intell. Med.*, vol. 41, no. 1, pp. 25–37, Sep. 2007.

[152] J. Sanz, D. Paternain, M. Galar, J. Fernandez, D. Reyero, and T. Belzunegui, "A new survival status prediction system for severe trauma patients based on a multiple classifier system," *Comput. Methods Programs Biomed.*, vol. 142, pp. 1–8, Apr. 2017.

[153] Y. Zhang, H. Zhang, J. Cai, and B. Yang, "A weighted voting classifier based on differential evolution," *Abstr. Appl. Anal.*, vol. 2014, pp. 1–6, May 2014.

[154] J. Jiang, P. Trundle, and J. Ren, “Medical image analysis with artificial neural networks,” *Comput. Med. Imaging Graph.*, vol. 34, no. 8, pp. 617–631, 2010.

[155] F. Xie, H. Fan, Y. Li, Z. Jiang, R. Meng, and A. Bovik, “Melanoma Classification on Dermoscopy Images Using a Neural Network Ensemble Model,” *IEEE Trans. Med. Imaging*, vol. 36, no. 3, pp. 849–858, Mar. 2017.

[156] K.-H. H. Liu and D.-S. S. Huang, “Cancer classification using Rotation Forest,” *Comput. Biol. Med.*, vol. 38, no. 5, pp. 601–610, May 2008.

[157] A. Ozcift and A. Gulten, “A robust multi-class feature selection strategy based on rotation forest ensemble algorithm for diagnosis of Erythemato-Squamous diseases,” *J. Med. Syst.*, vol. 36, no. 2, pp. 941–949, Apr. 2012.

[158] P. S. Daniel and A. G. Sam, *Research methodology*. Kalpaz, 2011.

[159] F. Smarandache, “Neutrosophic Logic - a Generalization of the Intuitionistic Fuzzy Logic,” *SSRN Electron. J.*, Jan. 2016.

[160] J. Ye, “Multicriteria decision-making method using the correlation coefficient under single-valued neutrosophic environment,” *Int. J. Gen. Syst.*, vol. 42, no. 4, pp. 386–394, May 2013.

[161] P. Singh and R. Shree, “A new homomorphic and method noise thresholding based despeckling of SAR image using anisotropic diffusion,” *J. King Saud Univ. - Comput. Inf. Sci.*, 2017.

[162] T. Yamaguchi and H. Hachiya, “Proposal of a parametric imaging method for quantitative diagnosis of liver fibrosis,” *J. Med. Ultrason.*, vol. 37, no. 4, pp. 155–166, 2010.

[163] F. Khursheed and A. H. Mir, “AR model based human identification using ear

biometrics," *Int. J. Signal Process. Image Process. Pattern Recognit.*, vol. 7, no. 3, pp. 347–360, Jun. 2014.

[164] L. Dandan, M. Huanhuan, L. Xiang, J. Yu, J. Jing, and S. Yi, "Classification of diffuse liver diseases based on ultrasound images with multimodal features," in *2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 2019, pp. 1–5.

[165] J. C. Russ, *The Image processing handbook*, 7th Edition. CRC Press, 2017.

[166] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Trans. Syst. Man. Cybern.*, vol. 8, no. 6, pp. 460–473, 1978.

[167] P. G. H. H. Gunasekara, J. V. Wijayakulasooriya, and H. A. C. Dharmagunawardhana, "Image texture analysis using deep neural networks," in *2017 IEEE International Conference on Industrial and Information Systems (ICIIS)*, 2017, pp. 1–5.

[168] B. Schmauch *et al.*, "Diagnosis of focal liver lesions from ultrasound using deep learning," *Diagn. Interv. Imaging*, vol. 100, no. 4, pp. 227–233, Apr. 2019.

[169] J. H. Lee *et al.*, "Deep learning with ultrasonography: automated classification of liver fibrosis using a deep convolutional neural network," *Eur. Radiol.*, pp. 1–10, Sep. 2019.

[170] P. Bharti, D. Mittal, and R. Ananthasivan, "Characterization of chronic liver disease based on ultrasound images using the variants of grey-level difference matrix," *Proc. Inst. Mech. Eng. Part H J. Eng. Med.*, vol. 232, no. 9, pp. 884–900, Sep. 2018.

[171] I. R. Khan and R. Ohba, "Taylor series based finite difference approximations of higher-degree derivatives," *J. Comput. Appl. Math.*, vol. 154, no. 1, pp. 115–124, 2003.

[172] W. Gomez, W. C. A. Pereira, and A. F. C. Infantosi, "Analysis of co-occurrence texture statistics as a function of gray-level quantization for classifying breast ultrasound," *IEEE*

Trans. Med. Imaging, vol. 31, no. 10, pp. 1889–1899, Oct. 2012.

[173] E. Angelini *et al.*, “A ranklet-based CAD for digital mammography,” pp. 340–346, 2006.

[174] R. Zhang *et al.*, “Automatic detection and classification of Colorectal Polyps by transferring low-level CNN features from nonmedical domain,” *IEEE J. Biomed. Heal. Informatics*, vol. 21, no. 1, pp. 41–47, Jan. 2017.

[175] K. R. Mendel, H. Li, D. Sheth, and M. L. Giger, “Transfer learning with convolutional neural networks for lesion classification on clinical breast tomosynthesis,” in *Medical Imaging 2018: Computer-Aided Diagnosis*, 2018, vol. 10575, p. 28.

[176] L. Saba *et al.*, “The present and future of deep learning in radiology,” *Eur. J. Radiol.*, vol. 114, pp. 14–24, May 2019.

[177] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[178] J. Stoitsis, I. Valavanis, S. G. Mougiakakou, S. Golemati, A. Nikita, and K. S. Nikita, “Computer aided diagnosis based on medical image processing and artificial intelligence methods,” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 569, no. 2 SPEC. ISS., pp. 591–595, 2006.

[179] A. Ozcift, “SVM feature selection based rotation forest ensemble classifiers to improve computer-Aided diagnosis of Parkinson disease,” *J. Med. Syst.*, vol. 36, no. 4, pp. 2141–2147, 2012.

[180] “MathWorks Announces Release 2016a of the MATLAB and Simulink Product Families - MATLAB & Simulink.”

Available: https://in.mathworks.com/company/newsroom/mathworks-announces-release-2016a-of-the-matlab-and-simulink-product-families.html. [Accessed: 12-Apr-2018].

[181] J. J. Rodriguez, L. I. Kuncheva, and C. J. Alonso, “Rotation Forest: A New Classifier Ensemble Method,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1619–1630, Oct. 2006.

[182] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms: Second Edition*, vol. 9781118315. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2014.

List of publications from present work

(i) P. Bharti, D. Mittal, R. Ananthasivan, "**Computer-aided Characterization and Diagnosis of Diffuse Liver Diseases Based on Ultrasound Imaging: A Review**," Ultrasonic Imaging, vol. 39, no. 1, pp. 33-61, 2017.

https://doi.org/10.1177/0161734616639875

Publisher: Sage

Impact factor: 2.490

(ii) P. Bharti, D. Mittal, R. Ananthasivan, "**Preliminary study of chronic liver classification on ultrasound images using an ensemble mode**," Ultrasonic Imaging, vol. 40, no. 6, pp. 357-379, 2018.

https://doi.org/10.1177/0161734618787447

Publisher: Sage

Impact factor: 2.490

(iii) P. Bharti, D. Mittal, R. Ananthasivan, "**Characterization of Chronic Liver Disease based on Ultrasound Images using the variants of Gray Level Difference Matrix**," Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, vol. 232, no. 9, pp. 884-900, 2018.
https://doi.org/10.1177/0954411918796531

Publisher: Sage

Impact factor: 1.317

(iv) P. Bharti, D. Mittal, "**Hybrid Feature Selection- based feature fusion for Liver disease Classification on Ultrasound Images**," Advances in Computational Techniques for Biomedical Image Analysis, Chapter 8, pp. 145, 2020.
Publisher: Elsevier
Book Chapter

(v) P. Bharti, D. Mittal, "**An ultrasound image enhancement method using Neutrosophic Similarity Score**," Ultrasonic Imaging. (Accepted)
Publisher: Sage
Impact factor: 1.571

Appendix A

Values of Handcrafted Features

The mean values of all the 148 handcrafted features from original images and 24 selected features from enhanced images are given in the appendix. The mean and standard deviation values of all the features for each liver class are given here.

S.No.	EXTRACTED FEATURES	ULTRASOUND LIVER TISSUE CLASS			
Handcrafted features extracted from original1 dataset					
		NOR	CHL	CIR	HCC
1	$CON_{0°}^{2pGLDM}$	0.1772±0.9995	0.0734±0.9172	0.0796±1.1390	-0.3989±0.7978
2	$CON_{45°}^{2pGLDM}$	-0.0436 ± 0.8795	0.0203±0.9164	0.2580±1.1495	-0.2919±0.9576
3	$CON_{90°}^{2pGLDM}$	-0.0134±1.0538	0.073±1.0362	0.1167±1.0615	-0.2182±0.7541
4	$CON_{135°}^{2pGLDM}$	-0.0455±1.0084	0.1655±1.0682	0.1171±1.0587	-0.2949±0.731
5	$ASM_{0°}^{2pGLDM}$	0.2027±0.9762	0.237±0.8832	0.0182±1.0175	-0.5552±0.9349
6	$ASM_{45°}^{2pGLDM}$	0.0408±1.011	0.0535±0.9634	0.1457±0.9887	-0.2942±0.9952
7	$ASM_{90°}^{2pGLDM}$	-0.119±1.0956	0.0619±1.0149	0.1674±0.9788	-0.142±0.8483
8	$ASM_{135°}^{2pGLDM}$	-0.1448±1.1072	0.2371±0.9765	0.1196±0.9142	-0.2687±0.909
9	$ENT_{0°}^{2pGLDM}$	0.1638±0.9340	0.1732±0.8852	-0.0305±1.1184	-0.3703±0.9586
10	$ENT_{45°}^{2pGLDM}$	0.0259±0.8838	0.0672±1.0248	-0.2307±0.9099	0.1713±1.1522
11	$ENT_{90°}^{2pGLDM}$	0.1683±0.9764	0.0263±1.0498	-0.2309±0.9491	0.0533±0.9835
12	$ENT_{135°}^{2pGLDM}$	0.1868±0.9712	-0.2108±1.0125	-0.1221±0.916	0.1896±1.0501
13	$M_{0°}^{2pGLDM}$	0.2812±1.0618	0.2332±0.9972	-0.0233±0.9693	-0.5925±0.6586
14	$M_{45°}^{2pGLDM}$	0.1267±1.094	0.0745±0.9741	0.0746±1.0142	-0.3342±0.8174
15	$M_{90°}^{2pGLDM}$	-0.0358±1.1079	0.0661±1.0091	0.1211±1.0243	-0.1886±0.7778
16	$M_{135°}^{2pGLDM}$	-0.0577±1.0782	0.186±1.0212	0.103±1.0006	-0.2884±0.7924
17	$CON_{0°}^{3pGLDM}$	0.1411±1.1079	0.1741±0.8090	0.0506±1.0192	-0.4446±0.9298
18	$CON_{45°}^{3pGLDM}$	0.0559±1.0579	0.1342±0.8709	0.1856±1.1888	-0.4609±0.6188
19	$CON_{90°}^{3pGLDM}$	-0.031±1.0074	0.0323±0.9964	0.214±1.1526	-0.2674±0.6877
20	$CON_{135°}^{3pGLDM}$	-0.1721±0.8962	0.0861±0.9050	0.2218±1.1341	-0.1761±0.9925
21	$ASM_{0°}^{3pGLDM}$	0.1327±1.0141	0.4575±0.7926	-0.0194±0.8005	-0.6982±1.0569
22	$ASM_{45°}^{3pGLDM}$	0.0134±0.9989	0.3473±0.9484	0.0284±0.934	-0.4796±0.9598
23	$ASM_{90°}^{3pGLDM}$	-0.0615±1.1181	-0.0244±0.9865	0.2039±0.9977	-0.1488±0.8259
24	$ASM_{135°}^{3pGLDM}$	-0.137±1.0601	0.1723±1.0166	0.1446±0.9409	-0.2288±0.9137
25	$ENT_{0°}^{3pGLDM}$	-0.0325±0.8261	-0.4314±0.8289	0.0163±0.7935	0.551±1.3047
26	$ENT_{45°}^{3pGLDM}$	0.0157±0.9134	-0.2892±1.0299	0.0817±0.8145	0.4382±1.1134
27	$ENT_{90°}^{3pGLDM}$	0.0644±1.0445	0.1184±1.0759	-0.1868±0.9434	0.0082±0.8868

Contd…

28	$ENT_{135°}^{3pGLDM}$	0.1643±0.9226	-0.1556±1.1167	-0.0763±0.9236	0.0916±0.997
29	$M_{0°}^{3pGLDM}$	0.1972±1.1063	0.403±0.9016	-0.0475±0.8614	-0.6726±0.773
30	$M_{45°}^{3pGLDM}$	0.061±1.0647	0.2287±0.8815	0.1016±1.1088	-0.4799±0.7339
31	$M_{90°}^{3pGLDM}$	-0.0363±1.0705	0.0161±0.9829	0.1965±1.0934	-0.2194±0.7402
32	$M_{135°}^{3pGLDM}$	-0.0864±1.0395	0.1344±0.9977	0.1858±1.0588	-0.2927±0.7869
33	$CON_{0°}^{5pGLDM}$	-0.4139±0.7219	0.1598±0.6102	0.1273±1.277	0.1362±1.144
34	$CON_{45°}^{5pGLDM}$	-0.4192±0.7211	0.1604±0.6109	0.1274±1.2737	0.1416±1.1456
35	$CON_{90°}^{5pGLDM}$	-0.4238±0.7234	0.1698±0.6054	0.1265±1.2748	0.1366±1.1429
36	$CON_{135°}^{5pGLDM}$	-0.4074±0.7228	0.1641±0.6169	0.1242±1.2774	0.127±1.1418
37	$ASM_{0°}^{5pGLDM}$	-0.2248±1.1232	0.2927±0.8492	0.1014±1.1034	-0.22±0.7461
38	$ASM_{45°}^{5pGLDM}$	-0.2323±1.1124	0.2741±0.8663	0.1186±1.0991	-0.2095±0.7547
39	$ASM_{90°}^{5pGLDM}$	-0.244±1.1146	0.2923±0.847	0.0926±1.074	-0.186±0.8176
40	$ASM_{135°}^{5pGLDM}$	-0.1926±1.1192	0.2649±0.8265	0.1227±1.115	-0.2502±0.773
41	$ENT_{0°}^{5pGLDM}$	-0.3573±0.9006	0.1852±0.8897	0.2202±1.0666	-0.0770±1.0340
42	$ENT_{45°}^{5pGLDM}$	0.2431±1.1757	-0.2776±0.811	-0.0812±1.0579	0.1548±0.8018
43	$ENT_{90°}^{5pGLDM}$	0.2543±1.1789	-0.293±0.7945	-0.0619±1.0369	0.1366±0.8439
44	$ENT_{135°}^{5pGLDM}$	0.2089±1.1652	-0.2769±0.7681	-0.0693±1.0842	0.1799±0.836
45	$M_{0°}^{5pGLDM}$	-0.5265±0.6142	-0.0713±0.7679	0.3727±1.1608	0.2519±1.1233
46	$M_{45°}^{5pGLDM}$	-0.4315±0.8347	0.2315±0.6109	0.0805±1.2149	0.1263±1.1193
47	$M_{90°}^{5pGLDM}$	-0.4362±0.8368	0.2413±0.603	0.0796±1.2154	0.1209±1.1179
48	$M_{135°}^{5pGLDM}$	-0.5218±0.6134	-0.0784±0.7669	0.3695±1.1632	0.2590±1.1234
49	$CON_{0°}^{7pGLDM}$	-0.4255±0.7735	0.2244±0.4933	0.0691±1.2584	0.1422±1.1853
50	$CON_{45°}^{7pGLDM}$	-0.4157±0.7744	0.2302±0.4889	0.0639±1.2633	0.1297±1.1849
51	$CON_{90°}^{7pGLDM}$	-0.4334±0.7699	0.2132±0.5111	0.082±1.2573	0.1494±1.1773
52	$CON_{135°}^{7pGLDM}$	-0.4361±0.7735	0.2272±0.5007	0.0715±1.2541	0.1482±1.1808
53	$ASM_{0°}^{7pGLDM}$	-0.1482±1.0342	-0.0134±0.9594	0.1025±1.0285	0.0656±0.959
54	$ASM_{45°}^{7pGLDM}$	-0.1382±1.0247	0.0018±0.9772	0.1048±1.0211	0.0323±0.9619
55	$ASM_{90°}^{7pGLDM}$	-0.1693±1.0391	-0.0112±0.9387	0.1074±1.0352	0.0819±0.963
56	$ASM_{135°}^{7pGLDM}$	-0.1547±1.0313	-0.0606±0.9401	0.1428±1.0174	0.0818±0.9893
57	$ENT_{0°}^{7pGLDM}$	0.1226±0.9947	0.0522±0.962	-0.0914±1.0379	-0.0968±0.9937
58	$ENT_{45°}^{7pGLDM}$	0.1211±0.9864	0.0245±0.971	-0.0921±1.034	-0.0601±1.002
59	$ENT_{90°}^{7pGLDM}$	0.1452±1.0009	0.0497±0.9462	-0.0948±1.0405	-0.1164±0.9957
60	$ENT_{135°}^{7pGLDM}$	0.1322±0.9947	0.0957±0.942	-0.128±1.0293	-0.1167±1.0154
61	$M_{0°}^{7pGLDM}$	-0.4279±0.8649	0.2901±0.4874	0.0291±1.2079	0.1132±1.1646
62	$M_{45°}^{7pGLDM}$	-0.4177±0.8661	0.2963±0.4811	0.023±1.2107	0.1009±1.167
63	$M_{90°}^{7pGLDM}$	-0.4361±0.8638	0.2769±0.5101	0.0428±1.2067	0.1224±1.154

64	$M_{135°}^{7pGLDM}$	-0.4388±0.8662	0.2932±0.4943	0.0312±1.204	0.1198±1.1582
65	$M_{4,V}^{Ranklets}$	0.2834±1.1356	0.2352±1.0482	-0.0358±0.9149	-0.582±0.4984
66	$STD_{4,V}^{Ranklets}$	0.1418±1.1261	0.0798±0.99	0.0586±1.0133	-0.3388±0.7354
67	$ENT_{4,V}^{Ranklets}$	-0.0134±1.0538	0.073±1.0362	0.1167±1.0615	-0.2182±0.7541
68	$M_{4,H}^{Ranklets}$	-0.0455±1.0084	0.1655±1.0682	0.1171±1.0587	-0.2949±0.731
69	$STD_{4,H}^{Ranklets}$	0.2027±0.9762	0.237±0.8832	0.0182±1.0175	-0.5552±0.9349
70	$ENT_{4,H}^{Ranklets}$	0.0408±1.011	0.0535±0.9634	0.1457±0.9887	-0.2942±0.9952
71	$M_{4,D}^{Ranklets}$	-0.119±1.0956	0.0619±1.0149	0.1674±0.9788	-0.142±0.8483
72	$STD_{4,D}^{Ranklets}$	-0.1448±1.1072	0.2371±0.9765	0.1196±0.9142	-0.2687±0.909
73	$ENT_{4,D}^{Ranklets}$	-0.1456±0.8395	-0.1413±0.8767	-0.0571±1.0055	0.4173±1.1884
74	$M_{6,V}^{Ranklets}$	0.0259±0.8838	0.0672±1.0248	-0.2307±0.9099	0.1713±1.1522
75	$STD_{6,V}^{Ranklets}$	0.1683±0.9764	0.0263±1.0498	-0.2309±0.9491	0.0533±0.9835
76	$ENT_{6,V}^{Ranklets}$	0.1868±0.9712	-0.2108±1.0125	-0.1221±0.916	0.1896±1.0501
77	$M_{6,H}^{Ranklets}$	0.2812±1.0618	0.2332±0.9972	-0.0233±0.9693	-0.5925±0.6586
78	$STD_{6,H}^{Ranklets}$	0.1267±1.094	0.0745±0.9741	0.0746±1.0142	-0.3342±0.8174
79	$ENT_{6,H}^{Ranklets}$	-0.0358±1.1079	0.0661±1.0091	0.1211±1.0243	-0.1886±0.7778
80	$M_{6,D}^{Ranklets}$	-0.0577±1.0782	0.186±1.0212	0.103±1.0006	-0.2884±0.7924
81	$STD_{6,D}^{Ranklets}$	0.3718±1.2633	-0.4384±0.3137	0.1178±1.105	-0.0448±0.8546
82	$ENT_{6,D}^{Ranklets}$	0.3684±1.2754	-0.4244±0.3399	0.1227±1.1168	-0.0641±0.8093
83	$M_{8,V}^{Ranklets}$	0.3763±1.2823	-0.4302±0.3265	0.1127±1.0977	-0.054±0.8296
84	$STD_{8,V}^{Ranklets}$	0.3763±1.2871	-0.4334±0.3073	0.1106±1.0911	-0.0475±0.8391
85	$ENT_{8,V}^{Ranklets}$	0.3775±1.137	-0.469±0.5035	0.1211±1.0772	-0.0178±0.9666
86	$M_{8,H}^{Ranklets}$	0.3789±1.1467	-0.4699±0.4885	0.1212±1.0765	-0.0186±0.9625
87	$STD_{8,H}^{Ranklets}$	0.3803±1.1445	-0.4736±0.4853	0.1224±1.0773	-0.0171±0.9634
88	$ENT_{8,H}^{Ranklets}$	0.3813±1.1444	-0.4816±0.4698	0.1251±1.0752	-0.0118±0.9701
89	$M_{8,D}^{Ranklets}$	-0.3651±1.1013	0.4733±0.544	-0.1053±1.0513	-0.0218±1.0272
90	$STD_{8,D}^{Ranklets}$	-0.3671±1.1046	0.4754±0.5415	-0.1065±1.0505	-0.0205±1.0235
91	$ENT_{8,D}^{Ranklets}$	-0.3679±1.1047	0.4793±0.5347	-0.1083±1.0515	-0.0222±1.0238
92	$M_{10,V}^{Ranklets}$	-0.3698±1.105	0.4878±0.5194	-0.113±1.0511	-0.0245±1.0271
93	$STD_{10,V}^{Ranklets}$	0.3778±1.2347	-0.4559±0.3438	0.1101±1.0827	-0.0208±0.9111
94	$ENT_{10,V}^{Ranklets}$	0.3807±1.2427	-0.4516±0.3557	0.1114±1.0812	-0.0312±0.8953
95	$M_{10,H}^{Ranklets}$	0.3861±1.2479	-0.4555±0.3471	0.1063±1.0713	-0.0264±0.9013
96	$STD_{10,H}^{Ranklets}$	0.3859±1.2491	-0.4602±0.3332	0.1065±1.0675	-0.0207±0.9086
97	$ENT_{10,H}^{Ranklets}$	0.2135±1.1429	0.3388±0.9964	-0.0346±0.8833	-0.6286±0.5834
98	$M_{10,D}^{Ranklets}$	0.0559±1.0579	0.1342±0.8709	0.1856±1.1888	-0.4609±0.6188
99	$STD_{10,D}^{Ranklets}$	-0.031±1.0074	0.0323±0.9964	0.214±1.1526	-0.2674±0.6877
100	$ENT_{10,D}^{Ranklets}$	-0.0626±0.98	0.1043±0.9948	0.2148±1.1392	-0.3197±0.7298
101	$CON_{0°}^{GLCM}$	0.2854±1.0581	0.3071±1.0227	-0.0295±0.8616	-0.681±0.6761
102	$ENT_{0°}^{GLCM}$	-0.0401±1.0809	0.3002±0.9345	0.0728±1.0163	-0.413±0.8022

103	$CPROM_{0°}^{GLCM}$	-0.198±0.8194	0.1286±0.8784	0.2456±1.3635	-0.2272±0.6596
104	$CSHAD_{0°}^{GLCM}$	-0.085±0.7895	-0.0042±0.9494	0.2378±1.2557	-0.1878±0.8722
105	$VAR_{0°}^{GLCM}$	-0.4112±0.7181	0.1624±0.6001	0.1211±1.2886	0.1374±1.1387
106	$IDMN_{0°}^{GLCM}$	-0.2853±1.0597	-0.3068±1.0234	0.0296±0.8609	0.6803±0.6738
107	$MPROB_{0°}^{GLCM}$	0.107±1.0294	-0.2423±0.9527	-0.0794±0.9797	0.2703±0.9718
108	$ACOR_{0°}^{GLCM}$	-0.4112±0.7181	0.1624±0.6001	0.1211±1.2886	0.1374±1.1387
109	$CON_{0°}^{GLCM}$	0.2852±1.0599	0.3067±1.0235	-0.0296±0.8608	-0.6802±0.6736
110	$COR_{0°}^{GLCM}$	-0.699±1.045	-0.021±0.8495	0.1822±0.8828	0.6295±0.7096
111	$ENE_{0°}^{GLCM}$	0.0671±1.0999	-0.2621±0.9045	-0.0757±0.9715	0.3374±0.9231
112	$HOM_{0°}^{GLCM}$	-0.2855±1.0575	-0.3072±1.0225	0.0294±0.8618	0.6813±0.6769
113	$CON_{45°}^{GLCM}$	0.143±1.0257	0.144±0.9975	0.0728±0.9821	-0.4371±0.8709
114	$ENT_{45°}^{GLCM}$	-0.1098±1.0778	0.2441±0.9112	0.1095±1.0606	-0.3063±0.8278
115	$CPROM_{45°}^{GLCM}$	-0.2117±0.7783	0.1388±0.8742	0.2526±1.3835	-0.2323±0.6617
116	$CSHAD_{45°}^{GLCM}$	-0.0965±0.7334	-0.0185±0.967	0.2416±1.2898	-0.1612±0.8469
117	$VAR_{45°}^{GLCM}$	-0.4107±0.7178	0.163±0.6004	0.1202±1.2883	0.1372±1.1394
118	$IDMN_{45°}^{GLCM}$	-0.1279±1.0278	-0.1438±1.0113	-0.0885±0.9958	0.4382±0.8285
119	$MPROB_{45°}^{GLCM}$	0.1089±1.0351	-0.2069±0.9463	-0.0874±0.9872	0.2342±0.9814
120	$ACOR_{45°}^{GLCM}$	-0.4107±0.7178	0.163±0.6004	0.1202±1.2883	0.1372±1.1394
121	$CON_{45°}^{GLCM}$	0.1265±1.028	0.1438±1.0126	0.0897±0.9971	-0.438±0.8244
122	$COR_{45°}^{GLCM}$	-0.5279±1.0104	0.2269±0.9229	0.1193±0.8886	0.1983±0.992
123	$ENE_{45°}^{GLCM}$	0.0973±1.0981	-0.2105±0.8852	-0.0912±0.9925	0.2571±0.9581
124	$HOM_{45°}^{GLCM}$	-0.1489±1.0248	-0.1437±0.9913	-0.066±0.9778	0.4353±0.8871
125	$CON_{90°}^{GLCM}$	0.0263±0.9829	0.0798±1.0242	0.1153±1.0504	-0.2719±0.8784
126	$ENT_{90°}^{GLCM}$	-0.142±1.0667	0.226±0.9144	0.1156±1.0712	-0.2533±0.841
127	$CPROM_{90°}^{GLCM}$	-0.1824±0.8159	0.1535±0.8649	0.2417±1.3768	-0.2717±0.636
128	$CSHAD_{90°}^{GLCM}$	-0.0926±0.7492	-0.0029±0.9663	0.2364±1.2926	-0.1787±0.8249
129	$VAR_{90°}^{GLCM}$	-0.4103±0.7184	0.1635±0.6006	0.1204±1.2885	0.1359±1.1387
130	$IDMN_{90°}^{GLCM}$	-0.0109±0.9717	-0.0777±1.0237	-0.134±1.065	0.2741±0.8686
131	$MPROB_{90°}^{GLCM}$	0.1453±1.0249	-0.1964±0.9472	-0.0978±1.0022	0.191±0.9795
132	$ACOR_{90°}^{GLCM}$	-0.4103±0.7184	0.1635±0.6006	0.1204±1.2885	0.1359±1.1387
133	$CON_{90°}^{GLCM}$	0.0095±0.9706	0.0775±1.0237	0.1357±1.0665	-0.2743±0.8675
134	$COR_{90°}^{GLCM}$	-0.375±1.0673	0.3287±0.8797	0.0846±0.9131	-0.0658±1.0119
135	$ENE_{90°}^{GLCM}$	0.1312±1.0965	-0.1962±0.8914	-0.0977±1.0053	0.2072±0.9468
136	$HOM_{90°}^{GLCM}$	-0.0324±0.9868	-0.0805±1.0242	-0.1078±1.0456	0.2706±0.8815
137	$CON_{135°}^{GLCM}$	0.0133±0.991	0.1287±1.047	0.1146±1.0433	-0.3161±0.8212
138	$ENT_{135°}^{GLCM}$	-0.1429±1.0593	0.2335±0.9365	0.1193±1.0688	-0.2662±0.8181
139	$CPROM_{135°}^{GLCM}$	-0.1829±0.8282	0.1401±0.8465	0.2467±1.3753	-0.2607±0.6572
140	$CSHAD_{135°}^{GLCM}$	-0.0855±0.7647	-0.0029±0.9619	0.2289±1.2801	-0.1778±0.842
141	$VAR_{135°}^{GLCM}$	-0.41±0.7186	0.1631±0.6008	0.12±1.2887	0.1366±1.1384
142	$IDMN_{135°}^{GLCM}$	0.0073±0.975	-0.1285±1.0541	-0.1364±1.0605	0.3183±0.8004
143	$MPROB_{135°}^{GLCM}$	0.1404±1.0222	-0.2018±0.9586	-0.0987±1.0039	0.2045±0.9635
144	$ACOR_{135°}^{GLCM}$	-0.41±0.7186	0.1631±0.6008	0.12±1.2887	0.1366±1.1384

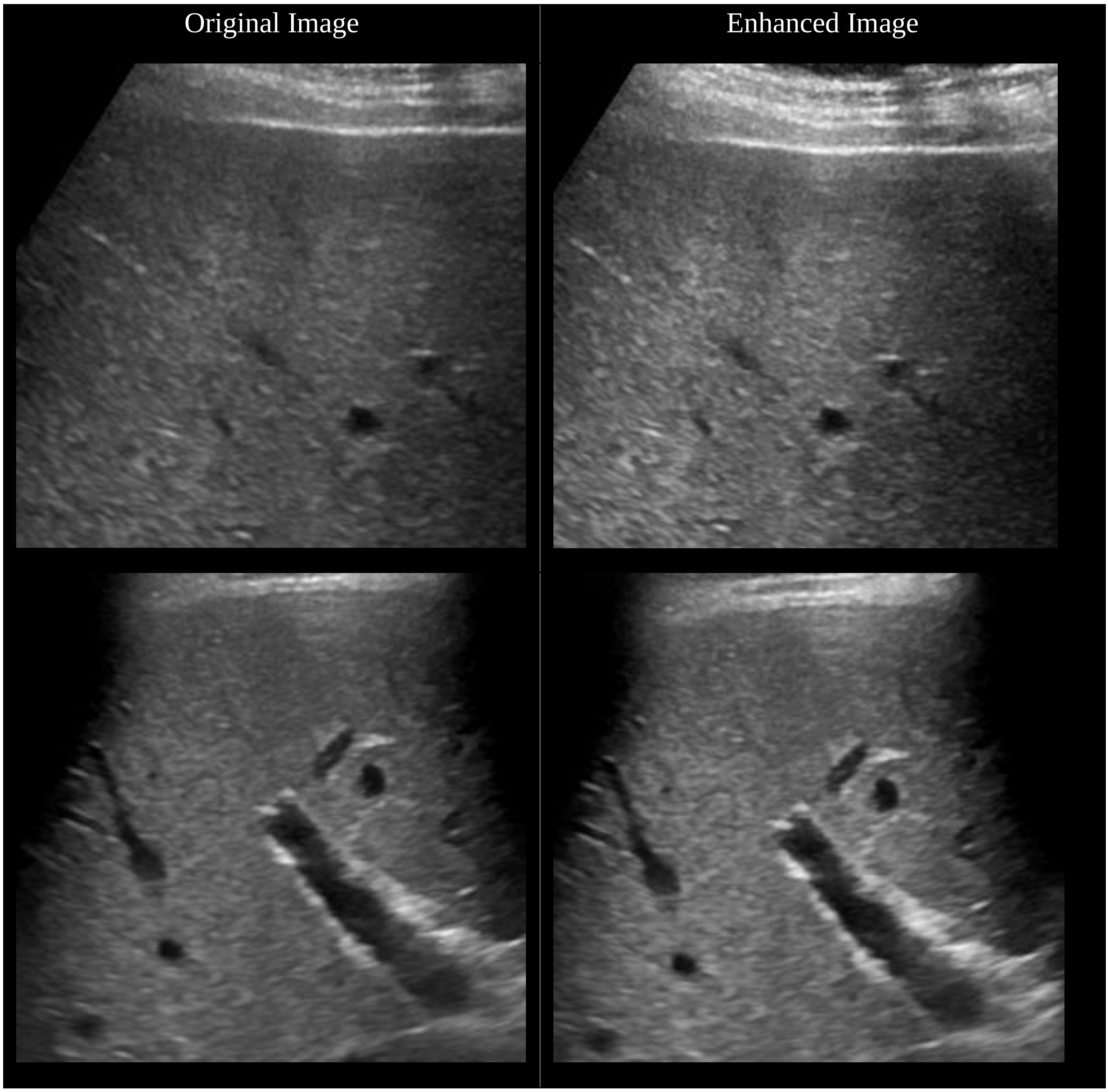
145	$CON_{135°}^{GLCM}$	-0.0092±0.9734	0.1284±1.0545	0.1384±1.0625	-0.3184±0.7982
146	$COR_{135°}^{GLCM}$	-0.36±1.0558	0.2706±0.8699	0.0663±0.9417	0.0108±1.0357
147	$ENE_{135°}^{GLCM}$	0.1279±1.0972	-0.2006±0.9073	-0.0997±1.0053	0.2192±0.9233
148	$HOM_{135°}^{GLCM}$	-0.0217±0.997	-0.1286±1.0436	-0.1056±1.0376	0.3147±0.8288
Handcrafted features extracted from enhanced dataset					
1	$CON_{135°}^{3pGLDM}$	-0.0626±0.98	0.1043±0.9948	0.2148±1.1392	-0.3197±0.7298
2	$ASM_{135°}^{3pGLDM}$	-0.137±1.0601	0.1723±1.0166	0.1446±0.9409	-0.2288±0.9137
3	$CON_{0°}^{2pGLDM}$	0.2834±1.1356	0.2352±1.0482	-0.0358±0.9149	-0.582±0.4984
4	$M_{0°}^{5pGLDM}$	-0.426±0.8343	0.2313±0.6105	0.0798±1.2173	0.1209±1.1199
5	$CON_{45°}^{2pGLDM}$	0.1418±1.1261	0.0798±0.99	0.0586±1.0133	-0.3388±0.7354
6	$ENT_{0°}^{2pGLDM}$	-0.1456±0.8395	-0.1413±0.8767	-0.0571±1.0055	0.4173±1.1884
7	$CON_{0°}^{3pGLDM}$	0.2135±1.1429	0.3388±0.9964	-0.0346±0.8833	-0.6286±0.5834
8	$ENT_{45°}^{3pGLDM}$	0.0247±0.9114	-0.3892±1.0295	0.0917±0.8165	0.3382±1.1152
9	$ASM_{45°}^{5pGLDM}$	-0.4323±1.2124	0.2991±0.8688	0.1635±1.0988	-0.3095±0.9847
10	$ASM_{90°}^{5pGLDM}$	-0.644±1.0146	0.9723±0.999	0.0899±2.074	-0.189±0.3666
11	$ENT_{0°}^{5pGLDM}$	0.2365±1.1831	-0.2992±0.7897	-0.0572±1.0698	0.1598±0.79
12	$M_{135°}^{5pGLDM}$	-0.419±0.8348	0.2362±0.6168	0.0762±1.2159	0.1111±1.1203
13	$ENT_{4,H}^{Ranklets}$	0.0418±1.021	0.0545±0.9534	0.1487±0.9897	-0.3042±0.9992
14	$ENT_{4,V}^{Ranklets}$	-0.0150±1.0528	0.080±1.0379	0.1178±1.0617	-0.2190±0.7541
15	$M_{4,H}^{Ranklets}$	-0.0465±1.0081	0.1655±1.0681	0.1161±1.0590	-0.2939±0.721
16	$ENT_{6,H}^{Ranklets}$	0.1872±0.9712	-0.2110±1.0105	-0.1220±0.916	0.1896±1.0511
17	$M_{4,V}^{Ranklets}$	0.2824±1.1346	0.2342±1.0481	-0.0362±0.9150	-0.578±0.4954
18	$STD_{4,V}^{Ranklets}$	0.1428±1.1251	0.0788±0.991	0.0536±1.0134	-0.3378±0.7353
19	$M_{6,H}^{Ranklets}$	0.2810±1.0638	0.2352±0.9992	-0.0235±0.9703	-0.5915±0.6596
20	$STD_{6,D}^{Ranklets}$	0.3716±1.2623	-0.4394±0.3127	0.1188±1.115	-0.0548±0.8446
21	$STD_{10,H}^{Ranklets}$	0.9759±2.2491	-0.3602±0.3552	0.1063±2.0675	-0.0300±0.8999
22	$COR_{45°}^{GLCM}$	-0.5289±1.0224	0.2409±1.0230	0.1213±0.8987	0.1353±0.9856
23	$CON_{135°}^{GLCM}$	-0.0192±0.9864	0.1386±1.0987	0.1356±1.0879	-0.3121±0.7851
24	$VAR_{0°}^{GLCM}$	-0.4222±0.8182	0.1734±0.6002	0.1211±2.2811	0.1874±1.1458

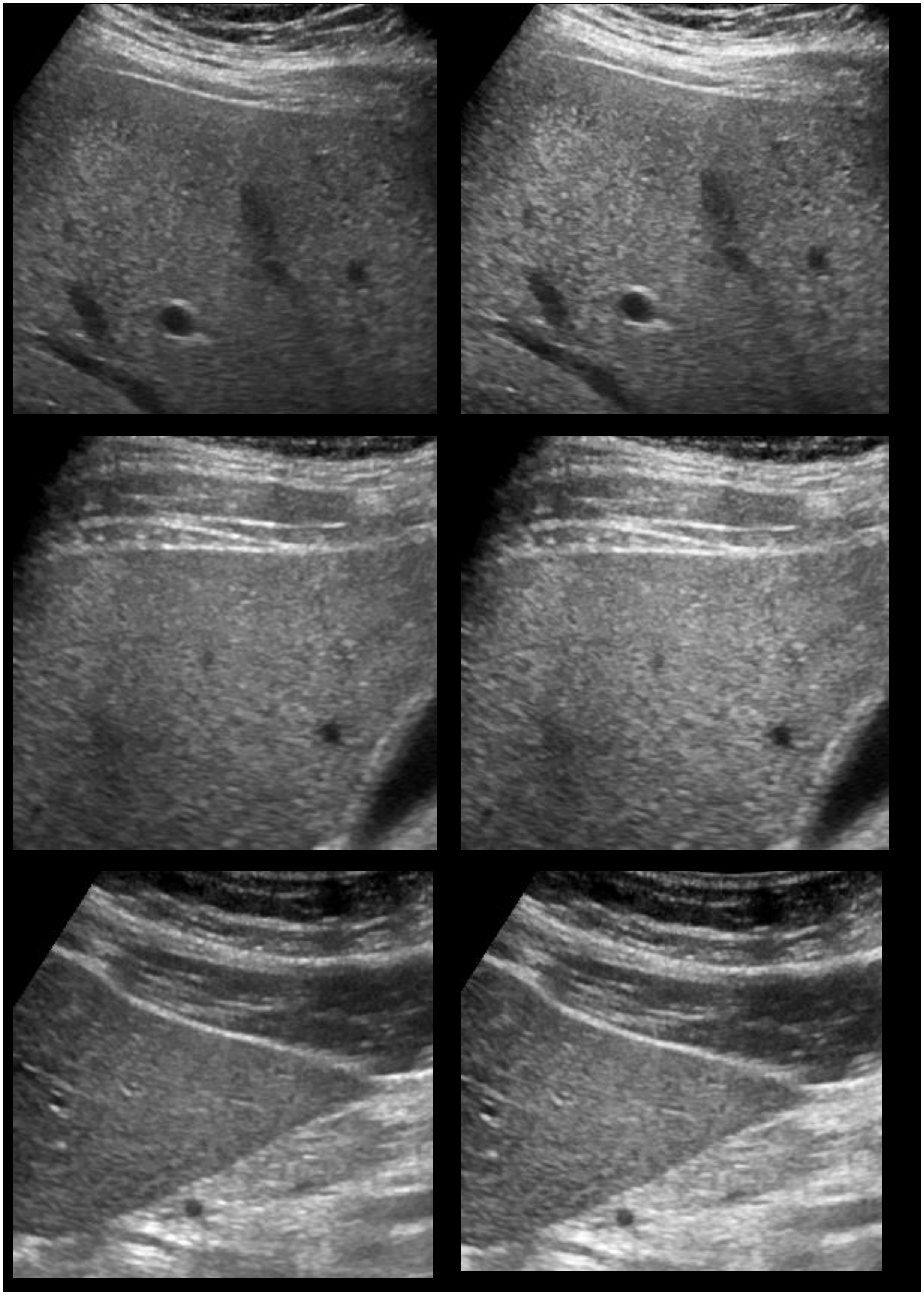
Appendix B

Sample Images

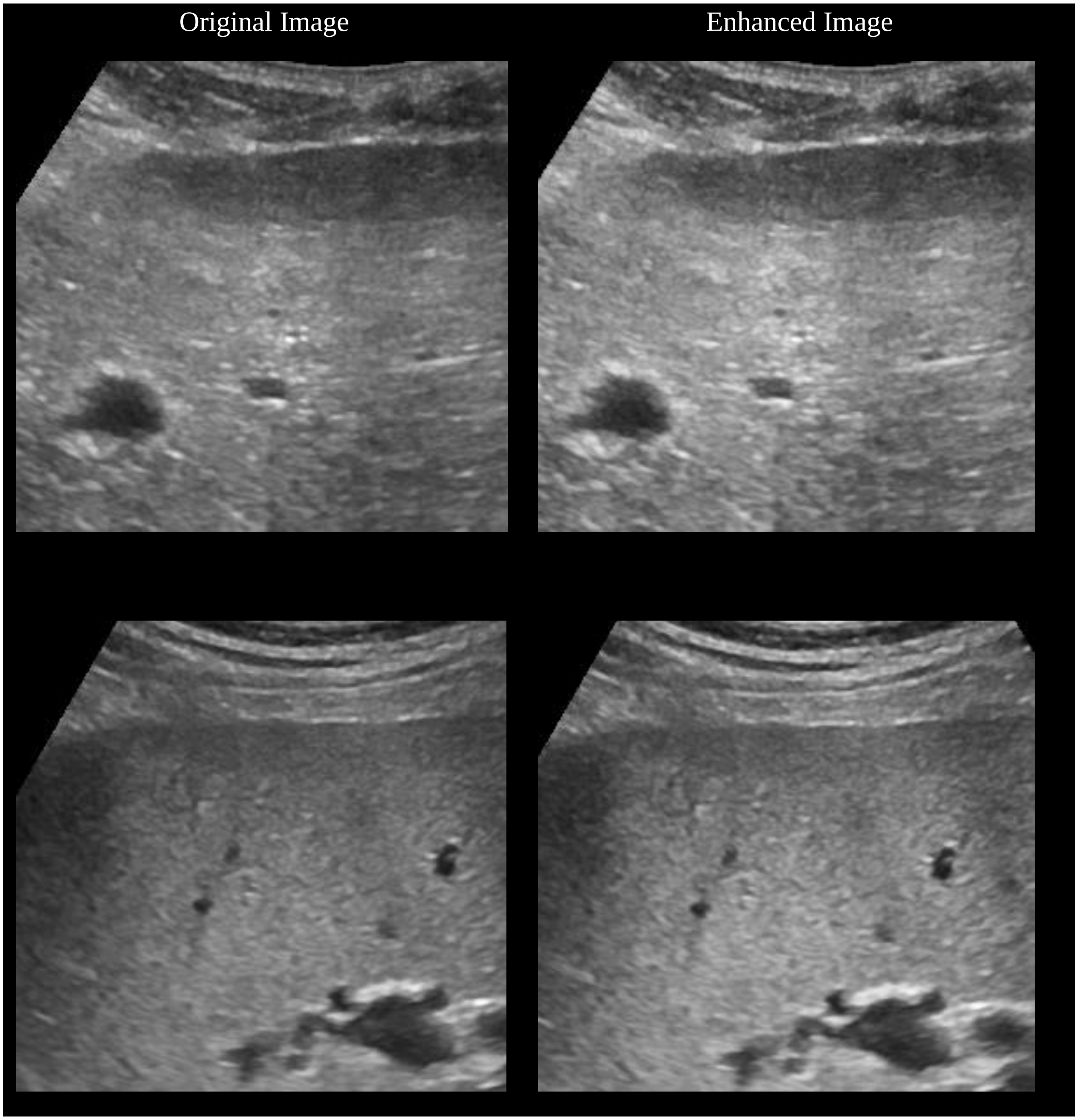
The example image from database and their corresponding results are given below:

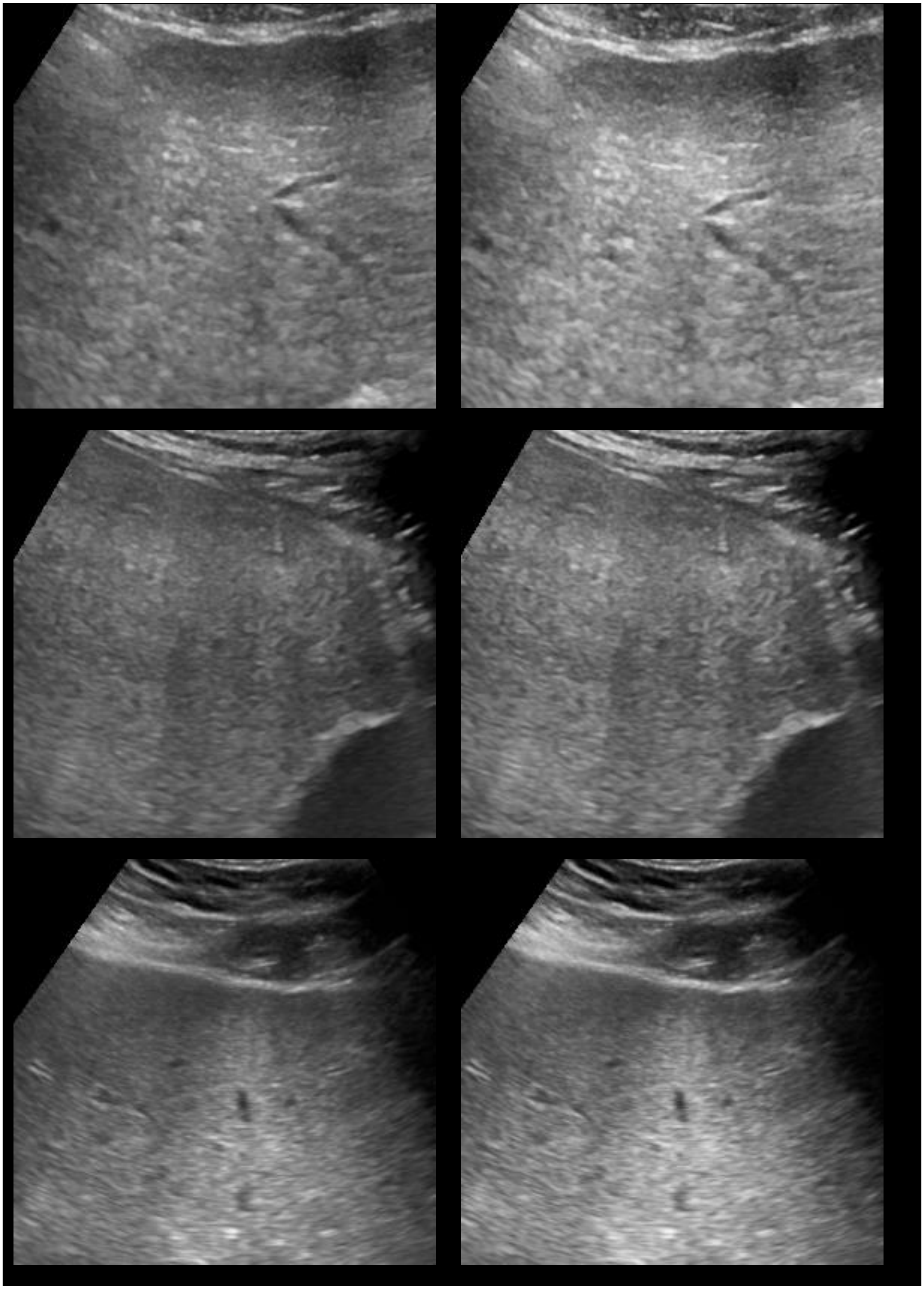
(i) Enhancement results with template 3 on sample images of normal liver are shown below:



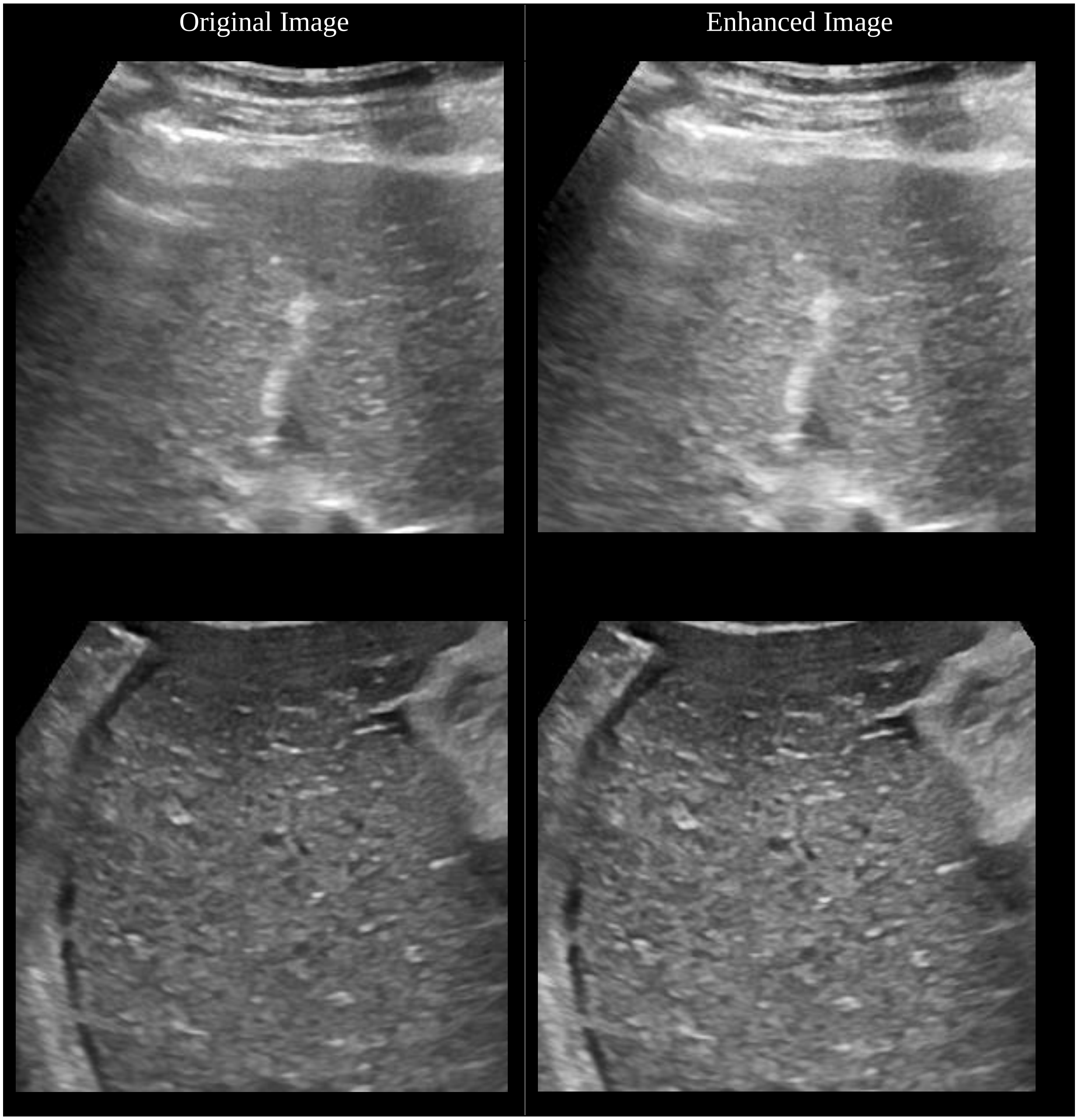


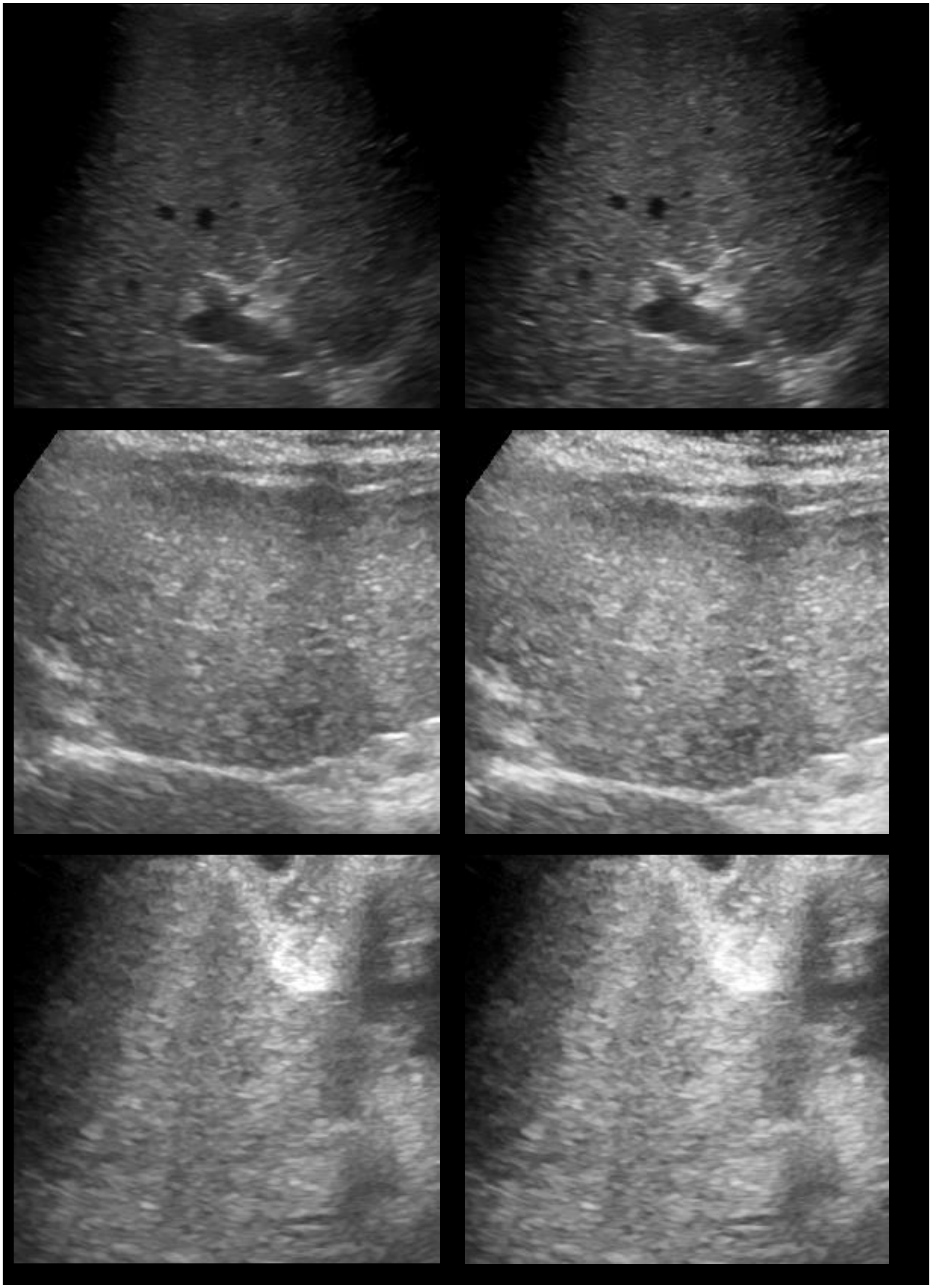
(ii) Enhancement results with template 3 on sample images of chronic liver are shown below:





(iii) Enhancement results with template 3 on sample images of cirrhosis liver are shown below:





(iv) Enhancement results with template 3 on sample images of HCC liver are shown below: