

**AN IMAGE PROCESSING DECISIONAL
SYSTEM FOR THE ACHILLES TENDON
USING ULTRASOUND IMAGES**

JAMAL SALH BEN RABHA

School of Computing, Science and Engineering
University of Salford, UK

Submitted in Partial Fulfilment of the Requirement
of the Degree of Doctor of Philosophy

2017

TABLE OF CONTENTS

TABLE OF CONTENTS	i
LIST OF TABLES	vii
LIST OF FIGURES	viii
LIST OF ABBREVIATIONS	xi
ACKNOWLEDGMENT	xiii
ABSTRACT	xv
Chapter 1 – INTRODUCTION AND MOTIVATION	1
1.1 Introduction	1
1.2 Background	2
1.2.1 Ultrasound Imaging	2
1.2.2 Ultrasound Techniques	3
1.2.3 The Achilles Tendon.....	4
1.2.4 Ultrasound Examination of the Achilles Tendon.....	6
1.2.5 Medical Image Processing Stages.....	8
1.3 Importance of the Study	12
1.4 Motivation for the Study	12
1.5 Problem of the Study.....	13
1.6 Aim and Objectives	13
1.7 Methodology	14
1.8 Contributions.....	14
1.9 Thesis Structure.....	15
1.10 Summary	15
Chapter 2 - BACKGROUND	16
2.1 Introduction	16
2.2 The Achilles Tendon	16

2.2.1 Anatomy	17
2.2.2 Achilles Tendon Abnormalities.....	19
2.2.3 Achilles Tendon Treatments.....	22
2.3 Ultrasound Imaging.....	23
2.3.1 Definition.....	23
2.3.2 Ultrasound Principles	24
2.3.3 Ultrasound Tissue Interaction.....	26
2.3.4 Ultrasound Usage and Importance in Medical Field.....	31
2.4 Ultrasound Techniques for the Achilles Tendon.....	32
2.4.1 Normal Achilles Tendon	33
2.4.2 Abnormal Achilles Tendon	35
2.5 Summary	36
Chapter 3 - LITERATURE REVIEW	37
3.1 Introduction	37
3.2 Image Pre-processing	38
3.2.1 Image Noises	39
3.2.2 Image De-noising	40
3.2.3 Comparison between different de-noising techniques.....	44
3.2.3 Image Enhancing	46
3.2.4 Summary.....	48
3.3 Image Segmentation.....	49
3.3.1 Introduction	49
3.3.2 Manual Segmentation.....	50
3.3.3 Segmentation Methods.....	50
3.3.4 Fuzzy Clustering Segmentation.....	52
3.3.5 Segmentation in Ultrasound Images.....	55

3.3.6	Determining ROI	57
3.3.7	Summary	58
3.4	Feature Extraction and Reduction	59
3.4.1	Introduction	59
3.4.2	Texture Analysis.....	59
3.4.3	Texture Analysis Approach	60
3.4.4	Texture Analysis Methods.....	62
3.4.5	Feature Reduction.....	68
3.4.6	Summary.....	72
3.5	Classification.....	73
3.5.1	Introduction	73
3.5.2	The k-Nearest Neighbour Techniques	73
3.5.3	Artificial Neural Networks Techniques	74
3.5.4	Decision Tree	75
3.5.5	Support Vector Machine	76
3.5.6	Deep Learning.....	77
3.5.7	Ensemble Classifiers	82
3.5.8	The Classification Imbalance Problem	85
3.5.9	Addressing the Imbalanced Problem	86
3.5.10	Comparison between classification techniques	88
3.5.11	Summary.....	89
3.6	Classification Evaluation	90
3.6.1	Introduction	90
3.6.2	Confusion Matrix	91
3.6.3	ROC curves.....	92
3.7	Summary	93

Chapter 4 - RESEARCH METHODS, EXPERIMENTS SETTINGS AND ANALYSIS	94
4.1 Introduction	94
4.2 Instrument and Data Acquisition.....	94
4.3 Methods Used.....	95
4.3.1 Image Pre-processing	96
4.3.2 Determining the ROI	100
4.3.3 Features Extraction	103
4.3.4 Features Reduction using KPCA.....	109
4.3.5 Classification	114
4.3.6 Cross-validation	116
4.4 Evaluation Methods.....	117
4.4.1 Image Quality Evaluation Metrics.....	117
4.4.2 ROI Detection Evaluation	119
4.4.3 Classification Evaluation.....	119
4.5 Conclusion.....	123
Chapter 5 – SYSTEM DEVELOPMENT AND IMPLEMENTATION	124
5.1 Introduction	124
5.2 The Purpose of Developing AT Decisional System.....	124
5.3 The Software Design and Functionality.....	125
5.3.1 Image Pre-processing	126
5.3.2 Features Extraction and Reduction.....	127
5.3.3 Image Classification	128
5.4 Analysis of Classification Algorithms Complexity	130
5.5 Summary	132
Chapter 6 - RESULTS AND DISCUSSION	134
6.1 Introduction	134

6.2 Image Dataset.....	135
6.3 Images Denoising.....	135
6.3.1 Approach one- Using the whole AT area.....	135
6.3.2 Approach Two- Manual Determination of the AT ROI.....	137
6.3.3 Approach Three- Dividing the AT Image into Blocks.....	139
6.3.4 Image Denoising Results Discussion.....	141
6.4 Image Enhancing.....	142
6.4.1 Approach one- the whole AT area.....	142
6.4.2 Approach Two- Manual Determination for the AT ROI.....	144
6.4.3 Approach Three- Dividing the AT Image into Blocks.....	145
6.4.4 Image Enhancing Results Discussion.....	147
6.5 Images Classification.....	148
6.5.1 Approach one- the whole AT area.....	148
6.5.1.3 Model performance measures using the first approach.....	150
6.5.2 Approach Two- Manual Determination for AT ROI.....	153
6.5.3 Approach Three- Dividing the AT Image into Blocks.....	157
6.5.4 Image Classification Analysis Results and Discussion.....	161
6.6 Summary.....	165
Chapter 7 - CONCLUSIONS AND FUTURE WORK.....	166
7.1 Introduction.....	166
7.2 Conclusions.....	168
7.2.1 Aim.....	168
7.2.2 Process of Achievement of the Research Objectives.....	169
7.2.3 Research Problem.....	170
7.3 Contribution of the Study.....	170
7.4 Research Scope and Limitations.....	170

7.4.1 Research Scope.....	170
7.4.2 Research Limitations	171
7.5 Future Work	171
7.6 Summary	173
REFERENCES.....	174
BIBLIOGRAPHY.....	190
APPNDIX.....	191
PUBLICATIONS	192

LIST OF TABLES

Table 1.1 Activities related to Achilles injuries.....	7
Table 2.1 The absorption of ultrasound by different body tissues.....	29
Table 2.2 The acoustic impedance and velocity of sound for different media.....	30
Table 3.1: The advantages and disadvantages of different de-noising techniques	44
Table 3.2: Summarize the features and limitations of different classifications techniques.....	88
Table 5.1 : Performance evaluation metrics for standard and ensemble classifiers	131
Table 6.1: Mean and standard deviation (SD) of noise metrics for various filters used with whole AT ultrasound images	136
Table 6.2: Mean and standard deviation (SD) of noise metrics for various filters used with manually determined ROI for AT ultrasound images.....	138
Table 6.3: Mean and standard deviation (SD) of noise metrics for various filters used with block divided ROI for AT ultrasound images.....	140
Table 6.4: Mean and standard deviation of noise metrics for various filters used with whole AT ultrasound images	143
Table 6.5: Mean and standard deviation (SD) of noise metrics for various filters used with the Manual determined ROI for AT ultrasound images	144
Table 6.6: Mean and standard deviation (SD) of noise metrics for various filters used with block divided ROI for AT ultrasound images.....	146
Table 6.7: A comparison between AT enhanced images for the three approaches	147
Table 6.8: Some Classifiers Evaluation (Sensitivity & Specificity) for the first approach ...	149
Table 6.9: Standard classifiers performance evaluation metrics for the first approach.....	150
Table 6.10: More classifiers performance evaluation measures using the first approach	151
Table 6.11: Some Classifiers Evaluation (Sensitivity & Specificity) for the second approach	153
Table 6.12: Standard classifiers performance evaluation metrics for the second approach ..	154
Table 6.13: More classifiers performance evaluation measures using the second approach.	155
Table 6.14: Some Classifiers Evaluation (Sensitivity & Specificity) for the third approach	157
Table 6.15: Standard classifiers performance evaluation metrics for the third approach.....	158
Table 6.16: More classifiers performance evaluation measures using the third approach ...	159

LIST OF FIGURES

Figure 1.1: The anatomy of the AT	5
Figure 1.2: Longitudinal ultrasound image of normal Achilles tendon is	6
Figure 1.3: Transverse ultrasound images of a normal Achilles tendon,	6
Figure 1.4: Ultrasound image of complete tendon tear showing the gap in the tendon	7
Figure 1.5: The general model in medical image processing	8
Figure 2.1: Right leg and ankle, posterior view with the ankle plantar flexed	17
Figure 2.2: Medial view of the foot and ankle, left	18
Figure 2.3: Tendon ultrastructure organization of collagen	19
Figure 2.4: AT Tendinosis	20
Figure 2.5: AT Partial tear	21
Figure 2.6: AT Full-thickness tear	22
Figure 2.7: Illustration of the surgical technique	23
Figure 2.8: The basic component elements in an ultrasound	25
Figure 2.9: The ultrasound wave	26
Figure 2.10: Ultrasound tissue interaction	27
Figure 2.11: Sonographic anisotropy	33
Figure 2.12: Normal AT (middle part)	34
Figure 2.13: Normal AT (lower part)	34
Figure 2.14: the Longitudinal sonographic image of a normal Achilles tendon	35
Figure 2.15: Longitudinal ultrasound image of the AT with power Doppler	36
Figure 3.1: GLCM of 4 x4 image for distance $d = 1$ and direction $\theta=0$	65
Figure 3.2: An example of GLRLM method	67
Figure 3.3 Relationships and high level schematics of different disciplines	78
Figure 3.4 Convolutional neural network	79
Figure 3.5 Recurrent neural network	81
Figure 3.6: Two ROC curves whose performance is to be compared	92
Figure 4.1: A block diagram of the proposed 6-stage AT classifier	96
Figure 4.2 Image Cropping Process	97
Figure 4.3 An original AT image	97
Figure 4.4 A cropped AT Image	98
Figure 4.5 Image De-noising Process	98

Figure 4.6 An AT image before applying any filter	99
Figure 4.7 The AT image after applying the SRAD filter	99
Figure 4.8 Image enhancing process.....	100
Figure 4.9 The AT image after contrast was enhanced using CLAHE algorithm.....	100
Figure 4.10 Ultrasound image shows the Lower AT area	101
Figure 4.11 AT ultrasound image with manual determining for ROI	102
Figure 4.12 Dividing the AT ultrasound image into equal blocks.....	103
Figure 4.13 Dividing the AT ultrasound image into six equal blocks.	103
Figure 4.14 Feature reduction KPCA reduction method	109
Figure 4.15 PCA component axes that maximize the variance	110
Figure 4.16 Linear vs nonlinear problems	111
Figure 4.17 Data scatter using PCA reduction method.....	113
Figure 4.18 Data scatter using KPCA reduction method.....	113
Figure 4.19 (a) (left table) A sample from the feature set, and (b) (right table) the reduced feature set using KPCA reduction method.....	114
Figure 4.20 <i>k</i> -fold cross-validation method	117
Figure 5.1 Research main menu.....	125
Figure 5.2 Image pre-processing selection	126
Figure 5.3 Features extraction and reduction.....	127
Figure 5.4 Features reduction menu.....	128
Figure 5.5 Images classification menu.....	129
Figure 5.6 An Example of images classification result	130
Figure 5.7 Comparison between various classifiers' results in their prediction speed	131
Figure 6.1 The result of applying filters on the whole AT ultrasound image.....	137
Figure 6.2: The result of applying filters on manually determined ROI for AT ultrasound.	138
Figure 6.3: The result of applying the filters on divided blocks represent ROI for AT	140
Figure 6.4: The result of applying CLAHE enhance algorithm filters on the whole AT	143
Figure 6.5: The result of applying CLAHE enhance algorithm on manually determined ROI	145
Figure 6.6: The result of applying the filters on divided blocks represent ROI for AT	146
Figure 6.7: Graph showing the classification evaluation metrics for different classifiers using the whole AT approach.....	152

Figure 6.8: Graph showing the classification evaluation metrics for different classifiers using the manual determining for AT ROI approach 156

Figure 6.9: Graph showing the classification evaluation metrics for different classifiers using the dividing AT into sub-images approach..... 160

LIST OF ABBREVIATIONS

AD	Average Difference
AD	Anisotropic Diffusion
ANN	Artificial Neural Network
AT	Achilles Tendon
AUC	Area Under the Curve
CAD	Computer Aided Design
CBIR	Content-Based Image Retrieval
CC	Coefficient of Correlation
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolutional Neural Network
CT	Computed Tomography
CV	Cross Validation
DT	Decision Tree
FN	False Negative
FP	False Positive
GB	Giga Byte
GLCM	Grey-Level Co-occurrence Matrix
FLICM	Fuzzy Logic Information C-Means
GLRLM	Grey Level Run-Length Matrix
G-Mean	Geometric Mean
GUI	Graphical User Interface
GWT	Gabor Wavelet Transform
HE	Histogram Equalization
HFCM	Histogram Fuzzy C-Means Clustering
IQ	Image Quality
ISFCM	Improved Spatial Fuzzy C-Means Clustering Algorithm
k-NN	k-Nearest Neighbour

KPCA	Kernel Principal Component Analysis
LDA	Linear Discriminant Analysis
LS	Least Square
LSVM	Linear Support Vector Machine
MCC	Matthews Correlation Coefficient
MSE	Mean Square Error
MRI	Magnetic Resonance Imaging
NCD	Nonlinear Coherent Diffusion
NFCM	Fuzzy C-Means Clustering Algorithm
P&M	Perona & Malik algorithm
PCA	Principal Component Analysis
PDE	Partial Differential Equation
PSNR	Peak Signal to Noise Ratio
RBF	Radius Basis Function
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristics
ROI	Region Of Interest
RUSBoost	Random Under Sampling Boosting
SD	Standard Deviation
SNR	Signal to Noise Ratio
SRAD	Speckle Reducing Anisotropic Diffusion
SSIM	Structural Similarity Index
SVM	Support Vector Machine
TP	True Positive
TN	True Negative

ACKNOWLEDGMENT

All praises to Allah, the most gracious and the most merciful, for the strengths and blessing he gave me in completing this thesis.

I would like to express my sincere gratitude to my supervisor Professor Farid Meziane for the continuous support of my PhD study, for his patience, motivation, and immense knowledge. The door to Professor Meziane office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed this thesis to be my own work but steered me in the right direction whenever he thought I needed. I could not have imagined having a better advisor and mentor for my PhD study.

Besides my supervisor, my sincere thanks also go to Dr Gillian Crofts for her help and support. Dr Crofts have provided me with all samples covered in this research dataset and enlightening me the first glance of research. Without her precious support, it would not be possible to conduct this research.

In my daily work, I have been blessed with a friendly and cheerful group of fellow students. We were not only able to support each other by deliberating over our problems and findings but also happily by talking about things other than just our researches.

I am especially thankful to my colleague Santosh More, who as a good friend was always willing to help and give his best suggestions. Many thanks to Rawan Abdeen, for helping me in collecting AT samples from the field. Special thanks go to Dr Ismaeel Agina, for his selfless support, encouragement, and also for his flawless grammatical editing of my thesis. My research would not have been possible without their help.

Last but not least; my deepest appreciation and very profound gratitude go to my parents, my wife, my daughters, my sons, my brothers and sisters for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Thank you very much, everyone!

Jamal Ben Rabha

To my Mother and Father

To my Wife

To my Daughters and Sons

ABSTRACT

The Achilles Tendon (AT) is described as the largest and strongest tendon in the human body. As for any other organs in the human body, the AT is associated with some medical problems that include Achilles rupture and Achilles tendonitis. AT rupture affects about 1 in 5,000 people worldwide. Additionally, AT is seen in about 10 percent of the patients involved in sports activities. Today, ultrasound imaging plays a crucial role in medical imaging technologies. It is portable, non-invasive, free of radiation risks, relatively inexpensive and capable of taking real-time images. There is a lack of research that looks into the early detection and diagnosis of AT abnormalities from ultrasound images. This motivated the researcher to build a complete system which enables one to crop, denoise, enhance, extract the important features and classify AT ultrasound images.

The proposed application focuses on developing an automated system platform. Generally, systems for analysing ultrasound images involve four stages, pre-processing, segmentation, feature extraction and classification. To produce the best results for classifying the AT, SRAD, CLAHE, GLCM, GLRLM, KPCA algorithms have been used. This was followed by the use of different standard and ensemble classifiers trained and tested using the dataset samples and reduced features to categorize the AT images into normal or abnormal. Various classifiers have been adopted in this research to improve the classification accuracy.

To build an image decisional system, a 57 AT ultrasound images has been collected. These images were used in three different approaches where the Region of Interest (ROI) position and size are located differently. To avoid the imbalanced misleading metrics, different evaluation metrics have been adapted to compare different classifiers and evaluate the whole classification accuracy. The classification outcomes are evaluated using different metrics in order to estimate the decisional system performance. A high accuracy of 83% was achieved during the classification process. Most of the ensemble classifiers worked better than the standard classifiers in all the three ROI approaches. The research aim was achieved and accomplished by building an image processing decisional system for the AT ultrasound images. This system can distinguish between normal and abnormal AT ultrasound images. In this decisional system, AT images were improved and enhanced to achieve a high accuracy of classification without any user intervention.

Chapter 1 – INTRODUCTION AND MOTIVATION

The introduction chapter aims to present the main themes of the study, contextualise the research and introduce the aim, objectives and the study structure. In addition, it presents a quick and concise overview of the most important aspects relevant to the subject of the study.

1.1 Introduction

Images play an important role in various aspects of our lives. Unfortunately, the human eye cannot extract some of the finest details in images which sometimes play a significant role in understanding their contents. With the current developments in computer processing capabilities, analysing and understanding these images, more importantly extracting their features, became more efficient and a routine task. Digital image processing is concerned with the automatic extraction of information from images using different computer programs. Image processing algorithms can perform different mathematical operations to extract the details needed to analyse and interpret digital images that are used in different applications and various domains. Television, photography, pattern recognition, robot vision, industrial applications and medical diagnosis are some of these important applications. Computers play an important role in all types of clinical image processing and are considerably more accurate in diagnosing diseases. For this reason, developing efficient algorithms and techniques for medical applications is the major concern for automatic analysis of medical images. There are many advantages in using computers in the medical field especially in diagnosis, surgery, and treatment planning. These include diagnosing diseases and early detections of the patients' conditions, improving the likelihood of recovery and avoiding complications. They also reduce the overall cost of treatment and in some instances, shorten the recovery time. There are different medical imaging models that include *Magnetic Resonance Imaging (MRI)*, *Computed Tomography (CT)*, *Ultrasound (US)* and *X-rays*. Each has a different purpose and uses different kinds of instruments.

1.2 Background

1.2.1 Ultrasound Imaging

Ultrasound, also called sonography, is one of the most broadly used imaging technologies in medicine. Sound waves are utilized to produce images of tissues and structures within the human body. Besides its availability and the relative ease of use, imaging is real-time and free of harmful radiation; it is portable and relatively inexpensive in comparison with the rest of imaging modalities, such as MRI and CT (Michailovich & Tannenbaum, 2006). Ultrasound offers excellent imaging of the soft tissues of the human body. The US utilizes a powerful magnetic field and radio waves to generate images of tissues, organs and additional structures within the human body that can be saved as digital images on a computer system. Pregnant women are typically instructed not to have MRI scans, with the exception of urgent situations. Furthermore, patients who have metal implants such as cochlear implants and cardiac pacemakers are instructed to avoid MRI scans because of the impact of the great magnetic field and strong radio frequency pulses (Dempsey, Condon, & Hadley, 2002). CT may be performed on any part of the human head and body. It has the ability to provide clear images of bones and tissues including muscles, organs, large blood vessels, the brain and nerves. The CT scans use X-rays, which are a form of radiation. Subjections to vast amounts of radiation are connected to the development of cancer or leukaemia usually after many years (Dempsey et al., 2002). Therefore, ultrasound imaging has been shown to be a safe procedure. It can be used to determine tissue or muscle injuries, and tendon tears due to sport related injuries. Ultrasound imaging is shown to detect important details in tissue pathology when used in clinical investigation (Minin & Minin, 2011).

Ultrasound uses a transducer to emit waves to deeper structures. A transducer is a small hand-held device used to perform the ultrasound scanning. It sends out inaudible high-frequency sound waves typically in the 2–10 MHz range (Wolbarst, Capasso, & Wyant, 2013). At the examination, the radiologist slides the transducer over the patient's skin. The probe produces sound waves at high frequencies into the respective body part and receives the echoes. The average speed of the ultrasound signal in human tissues is 1500 m/s. The depth of the sound waves penetrates the organs, and the resolution of the image is determined by the sound waves emitted by the frequency. By measuring the

returning echoes, it is possible to determine the object size, shape and nature (Chan & Perlas, 2011). The image shape displayed on the screen depends on the tissue density and the intensity of the reflected echoes. Not all the transmitted ultrasonic waves that travel through the tissues are reflected as echoes due to some of them being scattered. Actually, the tissue property determines the amount of echo reflected back to the transducer (Kaur, 2013). Due to the ultrasound waves absorbed by the tissues, the travelling waves become weaker and weaker, that is what causes noise during the acquisition stage (Ragesh, Anil, & Rajesh, 2011). Reflection, scattering and refraction form the interaction between the ultrasound device and different body organs. The reflection depends on the tissue type and structure. Each biological tissue has a different acoustic impedance which is a measure of the degree to which the medium impedes the sound waves (Minin & Minin, 2011). Higher spatial resolution in ultrasound images needs short pulses to distinguish between image objects and nearby boundaries to avoid overlap. However, the lack of clarity in ultrasound images makes them difficult to be interpreted. This problem needs to be addressed in order to refine and improve images, so it is easier to diagnose the patient's condition without having to wait a long time to get an expensive MRI, or exposure many times to X-rays.

1.2.2 Ultrasound Techniques

Ultrasonography transducers (or probes) hold numerous various piezoelectric crystals which are interconnected electronically and vibrate because of the applied electric current. The applied voltage causes an alteration in piezoelectric crystals surface and changing its thickness. These vibrations in their thickness are converted into mechanical energy inducing oscillations in piezoelectric crystals which are transmitted as ultrasound waves (Lutz & Buscarini, 2013).

There are four different ultrasound modes that are used in medical imaging. The first mode is the Amplitude Modulation (A-Mode), which is used to estimate the depth of an organ or assess the organs dimensions. The reflected sound impulses are shown on display like spikes of different heights. The distance between these heights determines the time spent by the sound wave to reflect to the transducer, and the heights determine the intensity of the returning impulse. The second mode is the Brightness Modulation (B-Mode), which is the most frequently used ultrasound mode by clinicians. The brightness is used to represent the echoes corresponding to the intensity of the echo signal. The position of the

transducer determines the position of the echo on display, where the time delay from the pulse transmission determines the bright dot and the horizontal position be determined by the location of the receiving transducer. By moving the transducer across sections of the body, it produces a sweep of the ultrasound beam and tomographic scan. The organ's tissues and boundaries are represented and constructed from the echoes which are generated by the reflection of the ultrasound waves. B-mode has a wide range of applications including imaging of the foetus, kidneys, liver, breasts, prostate and tendons (Szabo, 2004). The third mode is the Motion Modulation (M-Mode), which reflects the movement of structures over time. It is used for detecting and recording the temporal changes in echoes and provides a mono-dimensional view of the heart. The M-mode is commonly used in echocardiography for measuring the wall thickness and the heart rate (Szabo, 2004). Finally, the Doppler mode is used to measure and visualise the blood flow. The shift in frequency is calculated to trace the velocity of the blood and whether it is moving towards or away from the transducer. The Doppler shift can be either positive when the flow's direction is toward the probe, or negative when the flow moves away (Shung, 2015). Its applications are useful particularly in cardiovascular studies.

1.2.3 The Achilles Tendon

The Achilles tendon, which is also called the calcaneal tendon, is a tough band of fibrous tissue which joins the calf muscles to the heel bone calcaneus (Barfod, 2013). It originates in the middle of the lower leg at the confluence of the tendons of the gastrocnemius and soleus muscles at the gastro-soleus junction (Nunley, 2009). It is the common tendon of the gastrocnemius and soleus muscles and gives their attachment to the calcaneus, as shown in Figure 1.1. The muscle fibres are 6-8 cm long, making them capable of explosive contractions used for jumping and running (Barfod, 2013). The measurement lengthwise of the conjoined tendon is approximately 10 to 15 cm, while the gastrocnemius component is estimated to range from 11 to 26 cm, and that of the soleus element is approximated to range from 3 to 11 cm (Nunley, 2009).

The AT has been described as the largest and strongest tendon in the human body and is liable to tensile forces of around 12.5 times the body weight during sprinting and 6 to 8 times the body weight when performing athletic activities like jumping or cycling. AT plays an essential part in the foot

movement and its flexibility. As for any other organ in the human body, the AT is associated with some medical problems that include Achilles rupture and Achilles tendonitis.

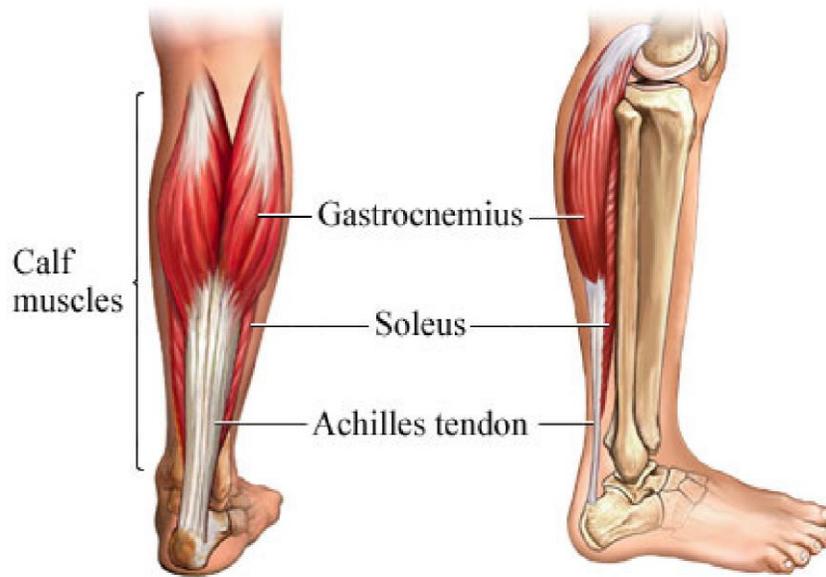


Figure 1.1: The anatomy of the AT
(Barfod, 2013)

The AT rupture symptoms occur due to the sudden increase of repetitive activity when the tissue that links the calf muscle to the heel bone tears. According to (Nunley, 2009), AT rupture affects about 1 in 5,000 people worldwide and most frequently (80% to 90% of cases) takes place 2 to 6 cm proximal to its calcaneal insertion. The incidence of proximal ruptures at the musculotendinous junction is about 10% to 15%. Despite having no AT rupture symptoms, most people experience pain and the inability to bend their foot or stand on their toes on the injured leg. Achilles tendonitis is found in about 10% of runners. However, it is also found in dancers, gymnasts, and tennis players (Barfod, 2013). According to (Mazzone & Meccue, 2002), Achilles tendonitis has symptoms such as ache on the back of the heel, difficulty walking and the pain may lead to walking being unbearable, swelling, tenderness and warmth of the AT.

Whilst reports of AT diseases are fairly extensive needs to have a better understanding of the tissue structure and develop better approaches to diagnosis, and prevention of future injuries. Doctors, after testing the lower leg, may need to perform a US scan that utilizes sound waves to create a picture

of the leg from inside. AT can be easily seen when the transducer is placed longitudinally to the tendon fibres where the normal AT is equally thick or to some extent thickened (Öhberg, 2003).

1.2.4 Ultrasound Examination of the Achilles Tendon

Surgeons use ultrasound examinations as a fast and efficient way to diagnose AT disorders. For optimal results, the AT should be scanned in longitudinal and transverse planes. Normal ATs show uniformity in their echogenicity and thickness with an organised fibrillar ultrastructure in their longitudinal position as shown in Figure 1.2. In contrast, in the transverse plane, the interior concave should be flat as shown in Figure 1.3.



Figure 1.2: Longitudinal ultrasound image of normal Achilles tendon is demonstrating parallel echogenic fibres (arrows). Anisotropy is seen in the insertional fibres, which are no longer perpendicular to the ultrasound beam (small arrows). (Courtesy of Kathryn Stevens, MD.)

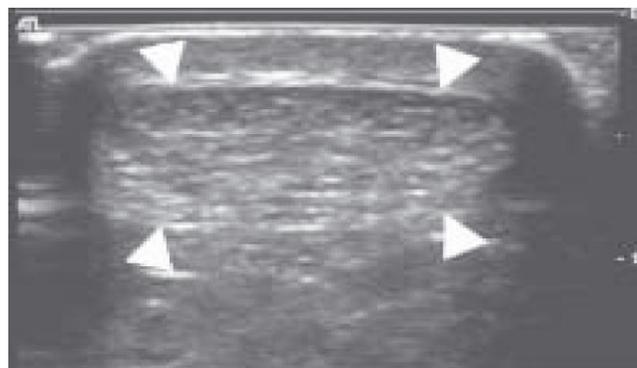


Figure 1.3: Transverse ultrasound images of a normal Achilles tendon, probe oriented at 90° to the tendon presenting a normal oval echo-bright tendon (arrowheads) (Bleakney, White, & Maffulli, 2007)

Injuries occurring on the AT are mainly due to over-exercising by young adults or overuse in elderly patients. Tendinosis appears as a swelling of the tendon with pain at the posterior heel and calf area. Tears of AT most likely happen 2 – 6 cm proximal to the calcaneal insertion (Maquirriain, 2011).

A partial and full tear could happen to the AT as shown in Figure 1.4, where in partial tears the tendon remains linked to the calf muscle, while with complete tears the tendon is entirely disjoint and detached from the calf muscle and the patient is unable to point their foot downward (Mazzone & Meccue, 2002). Table 1.1 shows most rupture sources during exhausting physical activities.

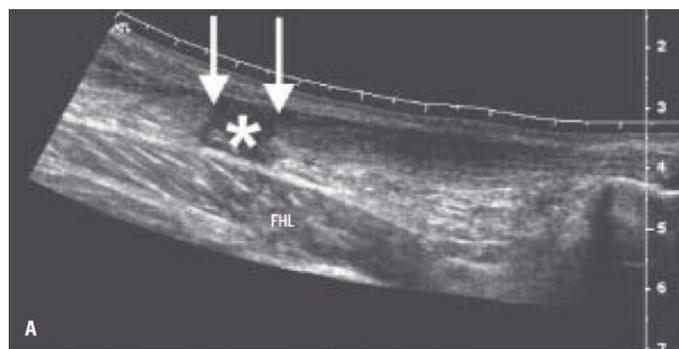


Figure 1.4: Ultrasound image of complete tendon tear showing the gap in the tendon (*) and the torn tendon ends (solid arrows) (Bleakney et al., 2007)

Table 1.1 Activities related to Achilles injuries

Tendonitis	Achilles rupture
Cross-country running	Football
Track and field	Baseball
Walking	Basketball
Cycling	Softball
Construction work	Tennis
	Racquetball

1.2.5 Medical Image Processing Stages

The general model that researchers follow in medical image processing is summarised in Figure 1.5. This model will be adapted to suit the needs of the current research. Its integration in the overall research methodology will be described in chapter 4. A summary of the different phases of this generic model will be presented in the following sub sections.

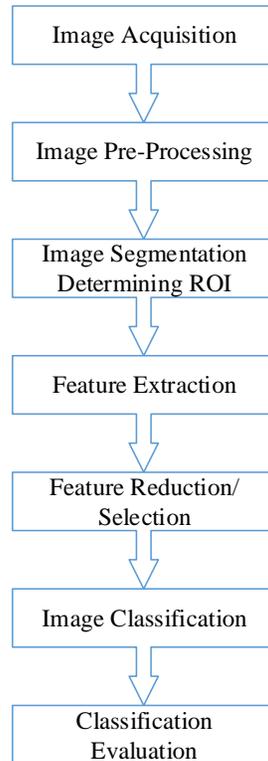


Figure 1.5: The general model in medical image processing

1.2.5.1 Pre-Processing

Various medical images suffer from the presence of noises, which hide some of their features and make the images generally difficult for the non-specialist to interpret so that often they might not be able to make decisions based on the images easily. The major challenges with ultrasound images are the presence of low contrast in images and interference with speckle noise (Cheng, Shan, Ju, Guo, & Zhang, 2010). Speckle makes the visual inspection and interpretation tricky. It results in a degradation of the image quality. Therefore, the task of the image pre-processing stage is to decrease speckle noise whilst

maintaining the vital features of AT images for diagnosis. De-speckling accomplished by various linear or nonlinear filters. Then, smoothing and sharpening details are carried out on the image to remove speckle noise and enhance edges at the same time. The goal of image enhancement techniques is to improve image contrast and brightness. This will lead to the enhancement of the overall appearance and will increase the efficiency of the image segmentation and classification processes.

1.2.5.2 Image Segmentation

Image Segmentation plays a vital role in many biomedical image applications because it represents the foundation for the remaining stages. Any shortage occurring at this stage will affect the following stages. Segmentation can be used in subdividing the objects in an image and separating different tissues from each other. Several algorithms are used for image segmentation; however, there is not a single one that can be considered to be superior due to different conditions and constraints. Successfully segmenting a medical image, allows focusing more precisely on the Region of Interest (ROI), which may improve the computational time especially on large medical datasets. Due to the complex nature of ultrasound images, segmentation is one of the most difficult tasks. However, to accomplish segmentation, there are many methods. The thresholding segmentation technique is based on distinguishing the interested object from the background by selecting an adequate threshold value. In contrast, the region-based and edge-based approaches' goal is to gather pixels with similar properties (C.-m. Chen, Lu, & Lin, 1999).

Pre-processing images using different types of filters is necessary to perform meaningful segmentation. Sometimes normalising the histogram of the image to proper distribution is followed by a filter to decrease the speckle noise and to enhance the features. For non-heavy working load, there is a simple way for image segmentation by allowing the user to identify the object boundary manually (Li, Chui, Chang, & Ong, 2011). According to (Li et al., 2011), semi-automatic computerized systems are used to allow the radiologist to adjust the segmentation parameters for optimal performance. Therefore, a robust model is needed to assist radiologists to locate the suspicious areas (Cheng et al., 2010). On the other hand, automatic segmentation in ultrasound images is a non-easy task as ultrasound images are complex in nature and deteriorated with speckle noise which makes the feature extraction phase sometimes more difficult.

1.2.5.3 Extracting Features

Extracting image features plays a crucial role in image classification in general and in this research in particular as it provides the crude data for the subsequent stages. Feature extraction is a special form of dimensionality reduction. It is performed after the pre-processing phase in an image processing system by separating the irrelevant attributes from the relevant ones. Feature reduction provides a mean for choosing the features which are best for classification, based on various criteria. It uses a classifier to recognize the selected features (Kumar & Bhatia, 2014). The decision as to whether the AT is normal or abnormal relies mainly on the features extracted during this stage. Many methodologies have been proposed to analyse and identify textures in an automated fashion (Mitra & Parekh, 2011). Most image features have been constructed on the basis of histogram and texture features (Tian, 2013). There are many techniques that have been introduced to extract texture features. They are generally classed into spatial texture feature extraction methods and spectral texture feature extraction methods based on the domain that texture is extracted from (Tian, 2013).

1.2.5.4 Features Reduction

The problem with extracting different kinds of texture features is the complexity caused by feature space dimensionality. Therefore, reducing the dimensionality of features is very important. However, dimensionality reduction should be done in such a way that only descriptive dimensions are kept. In the context of data reduction, there are several linear and nonlinear techniques available. Non-complex linear data methods, including Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), were used for years as dimensionality reduction methods. However, these techniques cannot handle successfully nonlinear representation of data (Maaten, Postma, & Herik, 2009). Recently, several nonlinear techniques for dimensionality reduction have been proposed (Maaten et al., 2009). Kernel PCA is the nonlinear reformulation technique for the traditional PCA. It embeds the data into a high dimensional space known as the feature space. In order to solve the high-dimensional mapping problem, the kernel trick was used. In medical image processing KPCA has been used to analyse the blood cells' characteristics and extract their general characteristics (Kim, Kim, Song, & Park, 2000). In (Tongsen & Ting, 2007) KPCA was used to reveal the distribution of the liver cancer cells in the B-

SCAN ultrasound image. It reaches a high level of accuracy and detects liver cancer and resists the confusion of noise.

1.2.5.5 Classification

Medical images classification is a very important stage; its objective being not only to reach a good level of accuracy but also to decide which part of the image is affected more than the others, in order to help the clinicians in their early diagnosis of the diseases. In texture classification, the objective is to assign an unknown sample image to one of a set of known texture classes. All the methods and processes that are taken in pre-processing, partitioning the image, extracting the texture features and selecting the most dominant features affect the classification process and results. The image set is divided into training and testing sets to estimate the performance of the classifiers. The n-fold approach is used to split the images into N disjoint sets, where the N-1 works as training and the Nth set are used for testing. The first phase of classification is training the classifier using a data set to determine the best set of features. The second phase is testing, where unknown data are given to the classifier for actual classification (Minavathi, S, & Dinesh, 2012). There are many techniques for the texture classification process. Non-parametric classifiers such as SVM, k-NN, and DT are used to improve classification accuracy. Choosing the good classifier affects the result, and for this reason, different classifiers are adapted, and through their results, the performance for each of them can be evaluated.

1.2.5.6 Classifiers Evaluation

The classification evaluation or post-classification stage is to measure the efficiency of the system. To estimate the efficiency and validity of the results, various techniques are used to evaluate their performance. Confusion matrix tables and ROC curves which show graphically all the approaches' results from various kinds of classifiers (Spackman, 1989). The confusion matrix is in dimension $D \times D$ where D is the number of classes. In this matrix, the number of True Positive (TP), meaning the samples that are abnormal and the test claim that they are abnormal. If some are normal because they have certain medical problems and classified by the test as abnormal, these are called False Negative (FN). There are some cases where they are abnormal and classified as normal – this is called False Positive (FP). Finally, when the images are normal and the test says that they are, this is called True Negative (TN). On the other hand, ROC curves are two dimensional graphs highlighting the relationship between

true positive rates on the Y-axis and false positive rates on the X-axis. The Area Under the ROC Curve (AUC) is often used to measure the quality of the classification model. Any realistic classifier has AUC between 0.5 and 1 (Fawcett, 2006). Random classifier has an AUC less than 0.5, while AUC for the perfect classifier is equal to 1.

1.3 Importance of the Study

This study, which is the first of its kind to deal with the AT ultrasound images, will benefit and use the information provided by health experts to distinguish between normal and abnormal AT. By building an image processing decisional system, this will help non-experts in the medical field to detect early the different diseases. This process will save time and effort on patients instead of waiting on long queues to be seen by experts. Such systems will be more useful in initial diagnosis, especially in remote villages where sometimes it is rare to find radiologist experts.

1.4 Motivation for the Study

Achilles tendon is considered as one of the most frequently injured tendons in the human body (Wren, Yerby, Beaupré, & Carter, 2001). High-resolution ultrasound is a suitable imaging tool for the initial inspection of AT. Because AT has a superficial structure, it is accessible by high-resolution ultrasound to provide more detailed images (Ying et al., 2003). According to (Fox et al., 1975) most of AT ruptures occur 2-6 cm above the heel bone and this is probably happening because the vascularity of the tendon in that area is decreasing (Lantto, 2016). This rupture affects about 1 in 5,000 people worldwide mostly during sports exercises or by stepping in a hole. Medical image processing has grown rapidly in recent years and contributed to the speed and accuracy of disease detection.

Ultrasonography is very widely used in medicine for diagnosis, as it offers real-time, portable and non-ionizing imaging capabilities. Ultrasound imaging has been shown to be a safe procedure in comparison to the other models and has considerable advantages when used in clinical procedures (Nunley, 2009). In addition, it can assist the clinician in deciding the proper pathological process, as well as the location of anomalies, reaction to treatment and assessment of the whole treatment plan. Many patients' lives are rescued due to early detection of different tumours in the brain, breasts and

other various internal organs. Discriminating between normal and abnormal cases is important to avoid unnecessary medical procedures such as surgery for the normal cases (Valentin et al., 2005). Thus, there is a critical need to develop a software model to detect the most important ROI of the AT ultrasound images and decide whether they are normal or abnormal.

1.5 Problem of the Study

AT is one of the most affected tendons in the human body (Wren et al., 2001). High-resolution ultrasound is a suitable imaging tool for initial inspection of AT. There are various medical imaging models where ultrasound is the most applied in the medical field. Unfortunately, the human eye cannot extract some of the finest details in the ultrasound images that sometimes may have an essential part in understanding the contents of the image.

As far as I know, there is a lack of studies and research that attempted to early detect and diagnosis for AT illness from US images. Moreover, the problem is exacerbated by the high cost of treatment by using different medical imaging models and length of recovery time and complications for many patients. Such problems occur in the initial diagnosis, especially in remote villages where sometimes it is rare to find radiologists as well as waiting for a long time to be seen by experts. Thus, there is an urgent need to develop a software model to detect the most important ROI for AT ultrasound images to be used to determine whether they are normal or abnormal.

1.6 Aim and Objectives

This study aims to develop a robust model that visually improves analyses and effectively classifies AT ultrasound images. This would provide medical doctors and other health professionals with discriminative visual characteristics, which will help them to diagnose better and consequently reach the right decisions as to whether the AT is normal or abnormal.

The main objectives of this study can be expressed as follows;

1. To identify and extract the most relevant features and information from AT ultrasound images.
2. To use the extracted features to classify specific characteristics of the AT image.

3. To determine the discriminatory ability of the textural features used in this research.
4. To investigate the influence of the ROI size on the classification accuracy.
5. To distinguish the ROI part from different image parts.

1.7 Methodology

In this research, machine learning techniques will be used to develop some approaches, in order to automatically distinguish between normal and abnormal tissues in AT ultrasound images. In addition, images will be pre-processed using appropriate de-noising and enhancing algorithms. Statistical-based texture features are used as they are one of the most widely used in medical image analysis (Holli et al., 2010). First and second-order description statistics are applied to extract various texture features. Grey-Level Co-occurrence Matrix (GLCM) and Grey Level Run-Length Matrix (GLRLM) are employed in this stage to extract various features. In order to overcome the dimensionality problem and reduce the features' dimensionality, the nonlinear technique Kernel Principal Component Analysis (KPCA) is used. Various classifiers will be tested in order to determine the most appropriate one for our application based on the final results. Finally, the used methods in each stage are evaluated in the evaluation stage using different evaluation metrics.

1.8 Contributions

The overall contribution of this thesis is to perform a deep research study in enhancing and classifying ultrasound images of the AT.

The following are the specific contributions of this research:

1. Developing a new system that is first for AT to determine the most appropriate ROI and classify ultrasound AT images without user intervention.
2. Identify a set of features that are specific to AT to obtain better classification accuracy.
3. Identify the best approaches to detect the ROI in ultrasound AT images.
4. Developing a new approach to identify the ROI automatically from the measures that are used by the radiologist to identify the region that is more likely to be affected by AT abnormality.
5. Evaluate the influence of imbalanced AT ultrasound dataset on the classification accuracy using standard and ensemble classifiers

1.9 Thesis Structure

The proposed study consists of six chapters; this is based on ultrasound and their contribution in the diagnosis of the AT medically. Where this chapter (**introduction**), aims to present the main themes in this study, contextualise the study and introduces the study problem, main aim, objectives and structure of the thesis.

- **Chapter 2: Background**

This chapter seeks to elaborate on the major topics that are linked to this study that includes AT anatomy and will focus on a definition of ultrasound imaging, ultrasound principles and ultrasound usage and importance in the medical field as well as ultrasound techniques for AT. In addition, this chapter will discuss how ultrasound techniques help in the field of medicine

- **Chapter 3: Literature Review**

The purpose of this chapter is to address past literature and previous researches and studies that are relevant to AT and Ultrasound and their contribution to the diagnosis of the AT medically.

- **Chapter 4: Research Methods, Experiments setting and Analysis**

This chapter presents the ways, steps and mechanisms that are related to the methods used and approaches applied and include the justification and explanation of the evaluation metrics adopted.

- **Chapter 5: Results and Discussion**

This chapter focuses on the key findings achieved in this research with a deep discussion of the results obtained.

- **Chapter 6 Conclusion** and recommendations for future work will be explored in this chapter.

1.10 Summary

This chapter focused on the main aim of the study “to develop a robust model that visually improves, analyses and effectively classifies AT ultrasound images” and the objectives of the study. In this chapter, some principal topics that are relevant to the study have been explained. In addition, this chapter focused in brief on the general models that the researcher will adopt in this study to suit the needs of the current research.

Chapter 2 - BACKGROUND

This chapter aims to elaborate on the major topics that are linked to this study and includes the AT anatomy. It will focus on the definition of ultrasound imaging, ultrasound principles, ultrasound usage and importance in the medical field and the ultrasound techniques for the AT

2.1 Introduction

The purpose of this chapter is to elaborate on the key topics that are related to this study mainly on the AT anatomy, its abnormalities and treatments. As ultrasound images are used in this research, we will proceed with the principles of performing an ultrasound examination and how it reacts with human tissues. We will then discuss how ultrasound techniques contribute to the field of medicine and elaborate on the difference between normal and abnormal AT ultrasound image from a radiologist's perspective.

2.2 The Achilles Tendon

The AT is the body's largest and strongest tendon. It connects the calf muscle of the leg to the calcaneum that makes the heel of the human foot. This helps to raise the feet up while walking and assists in running as well as tiptoeing. AT has the ability to transfer a force that is three fold greater than the body's weight when running. This provides a demonstration of the force that the body handles, which is approximately half the body's weight (Alexander, Ker, Bennet, Bibby, & Kester, 1987). This tendon is also one of the most common sites of injuries in the body (Pang & Ying, 2006). AT injuries are associated with poor ankle flexibility and strength and overusing the muscle. This muscle is also seen to be linked to medical issues like tendon rupture as well as the inflammation of the tendon called tendonitis. The tendon rupture is seen in men taking part in sports in the age groups of 30 to 40. The left side of the AT is more prone to damage than the right side (Hatstrup & Johnson, 1985). It has the tendency to occur in about 18 in 100,000 people (Gebauer et al., 2007).

The area where the AT ruptures occurs is about 2 to 6 cm from the calcaneum. This place has the smallest cross-section area as compared to the rest of the tendon (Wren, Lindsey, Beaupré, & Carter, 2003). The area where the tendon becomes inserted is capable of bearing higher strains as compared to the whole tendon. Treatment of acute AT ruptures can generally be classed as either operative or non-operative. The operative treatment is done mostly in athletes, young adults and fit patients. Percutaneous operative treatment is preferred in patients who tend to avoid the open surgery. The non-operative modality is chosen for elderly patients (Bossley, Martinelli, Maffulli, & Raisbeck, 2000).

2.2.1 Anatomy

AT is the aponeurosis of both the gastrocnemius muscle as well as the soleus muscle that is inserted to the calcaneum. The AT, as illustrated in Figure 2.1, is seen to be formed by the soleus muscle fibre as it gets inserted into the anterior portion at the end. When walking downhill, these fibres turn and twist on themselves as the medial side of the fibres gets inserted on the lateral side of the calcaneum bone (Logan & Rowe, 1995).

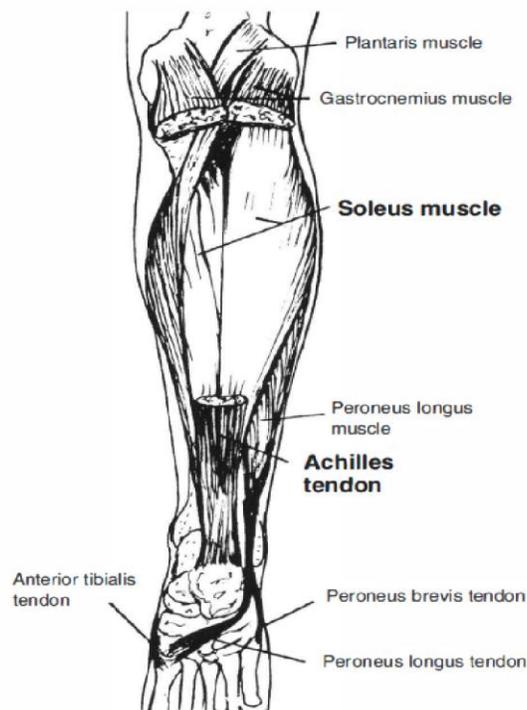


Figure 2.1: Right leg and ankle, posterior view with the ankle plantar flexed (Logan & Rowe, 1995)

The soleus muscle originates from the posterior aspect of the tibia, and the gastrocnemius originates from the distal part of the femur from the posterior aspect. The strategy behind the insertion of the gastrocnemius muscle is to assist while extending the knee and the soleus muscle to be effective while flexing the knee. The tendons from both muscles coalesce just distal to the musculotendinous junction to form the AT. This aponeurosis then has the tendency to rotate at 90 degrees shifting the gastrocnemius muscle towards the lateral side (Hamid, 2011).

The tendon gets its blood supply from the arteries and veins that are passing through the paratenon. The major bulk of the blood supply is from the posterior tibial artery (I. Ahmed, Lagopoulos, McConnell, Soames, & Sefton, 1998). These blood vessels pierce the mesotendon and supply blood to the tendon (Kader, Mosconi, Benazzo, & Maffulli, 2005). The middle part of the tendon is relatively less vascularised as compared to both of the ends of the muscle. It is one of the main reasons why the tendon ruptures about 2-6 cm above the calcaneus. The proximal part of the muscle is complimented by the arteries and veins already supplying the muscle bellies from the endotenon (Langberg, Bulow, & Kjar, 1998). The tendon and its bursa are protected from injuries and trauma by the help of Karger's fat pad that lies at the posterior part of the heel bone (Figure 2.2).

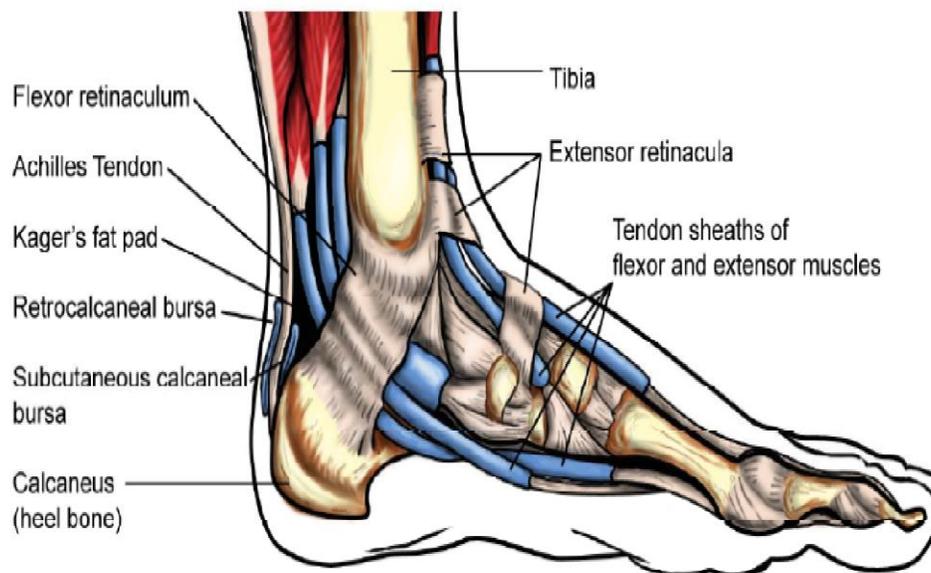


Figure 2.2: Medial view of the foot and ankle, left (Bjur, 2010)

The muscle fibres are supported and kept in place by fibrous, connective tissue covered by the help of endotenon (Figure 2.3), in which small blood vessels and lymphatic vessels and to an extent nerves are harboured, analogous to the situation in the paratenon. This gives them a free space to slide over each other (Ajis, Maffulli, Alfredson, & Almekinders, 2007). The number of nerve fibres in a tendon has yet been unknown.

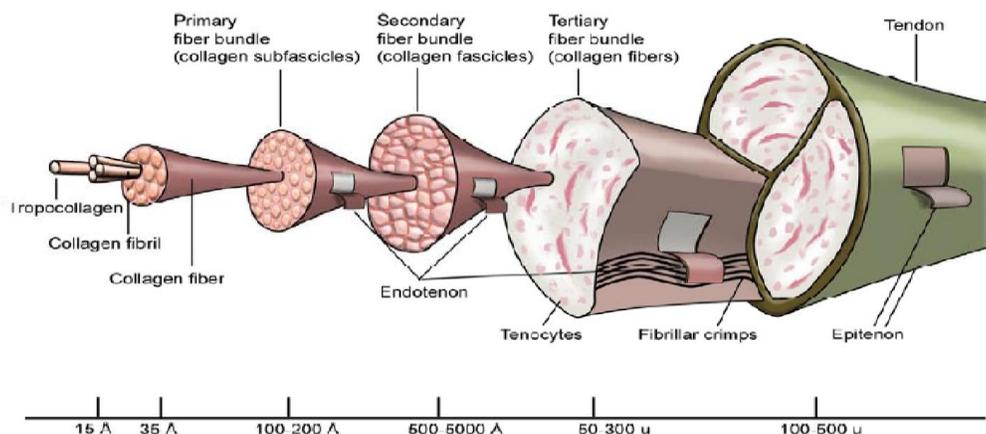


Figure 2.3: Tendon ultrastructure organization of collagen compounds from tropocollagen to the whole tendon. (Bjur, 2010)

2.2.2 Achilles Tendon Abnormalities

AT disorders are mostly occurring to the muscle as well as the degenerative changes that take place in a muscle. Tendon rupture occurs mostly 2-6 cm before its insertion in the calcaneum bone. Diseases of the tendons include tendinosis, partial and full tear of the muscle fibre. It occurs when the muscle is overused by both adolescents as well as adults (Mazzone & Meccue, 2002). In ultrasound assessment of the tendon, the tendon is mostly found to be thick and hypogenic. Earlier investigations show the thickness of tendon ranges from 7 to 16 mm as seen in patients with tendinopathy (Fornage, 1986).

2.2.2.1 Tendinosis

The inflammation of the AT is seen in about 10 percent of the patients who are athletes, dancers, tennis players, runners as well as in gymnasts ((James, Bates, & Osternig, 1978), (Brukner, 1997)). AT

tendonitis is the inflammation of the AT as illustrated in Figure 2.4. It occurs because of overuse of this muscle in the adolescents who practice sports as well as in those who are not involved in sports (R. R. Bleakney & White, 2005). The symptoms that the patient presents with are a pain at the base of calcaneum, change in gait because of tenderness at the base of the heel as well as difficulty in walking, swelling as well hyperaemia at the AT. The treatment options offered to the patients limit the activities that will cause pain and will worsen the condition, suggesting the patients to rest, prescribing Nonsteroidal anti-inflammatory drugs (NSAIDs) and offering physical therapy to these patients (McLauchlan & Handoll, 2001).



Figure 2.4: AT Tendinosis
(Body Heal, 2014)

2.2.2.2 Partial Tear

AT expansion larger than one cm in the anterior–posterior dimension or marked intrinsic tendon abnormalities, such as a hypoechoic or anechoic cleft, indicates a partial tear (Leppilahti & Orava, 1998). If the AT is torn, this is called an AT rupture. The tear may be either partial or complete. When there is a slight discontinuity in the muscle tendon fibre, it is termed as a partial tear of the tendon. This kind of tear still is attached to the calf muscle. Such tears are not categorized as complete tendon rupture as illustrated in Figure 2.5.

Partial tears typically do not expose themselves till after the body cools down from activity, where (Olsson, 2013):

- In partial tears, the tendon belly is tender to touch and appears inflamed. This, however, goes away in a couple of days.
- The tendon can appear painful and stiff if it is not used for a long period of time.

In partial tears, a person can still walk with a limp. The patient, however, complains of having severe pain when they use their toes to push off or when they touch down on their heels.



Figure 2.5: AT Partial tear
(Body Heal, 2014)

2.2.2.3 Full-thickness Tear

AT rupture and tears are identified in men among the age group of 30 to 50 years especially in those men who exercise (Mazzone & Meccue, 2002). Old people also experience tendon rupture because of degenerative changes that take place in a tendon without a history of injury to the calcaneum. The reason behind tendonitis and rupture is increasing age, poor conditioning as well as overuse while exercising. Generally, it is seen as the individual runs and performs activities like tip toeing that create a force which ruptures the tendon in a sudden break with immense pain at the heel or at the base of the foot. Intact plantar tendon is evident when there is full thickness, where AT tear and should never be misdiagnosed for intact AT fibres (Figure 2.6).



Figure 2.6: AT Full-thickness tear
(iWalk Free, 2000)

2.2.3 Achilles Tendon Treatments

Other treatment modalities offered to the patients are the conservative management of the condition the patient is suffering from. It limits the cost and risks of surgery but is essentially the cause for a person to be immobile for a longer period of time as well as immobilisation of the ankle in plantar flexion as it causes the gastrocnemius to atrophy and weakens (Nunley, 2009). If you apply cast to a patient with tendon rupture, there are chances of re-rupture as there is less pushing-off strength and less endurance when compared to tendons that have been surgically repaired. On the other hand, surgery is beneficial due to the likelihood of exerting early tension on the tendon, which will ensure a suitable orientation of the collagen fibres and enhance the power and ability of the calf muscles (Wahlby, 1978). Another added benefit of the surgical repair is that it decreases re-rupture rate, ensures a good speedy recovery of the calf muscle, restores its power and has a relatively lower chance of infection if an excellent surgical approach is chosen for the treatment. The operation is recommended for patients having no contraindications. This is offered to patients who are either athlete, adolescents as well as to those patients who have delayed ruptures, in the case of early ruptures in non-athletes, they are suggested to be treated conservatively (Leppilahti & Orava, 1998). Figure 2.7 illustrates various surgical techniques that have been carried out on patients under local anaesthesia. The tendon chosen is repaired from one side to the other by performing core sutures that are semi absorbent.

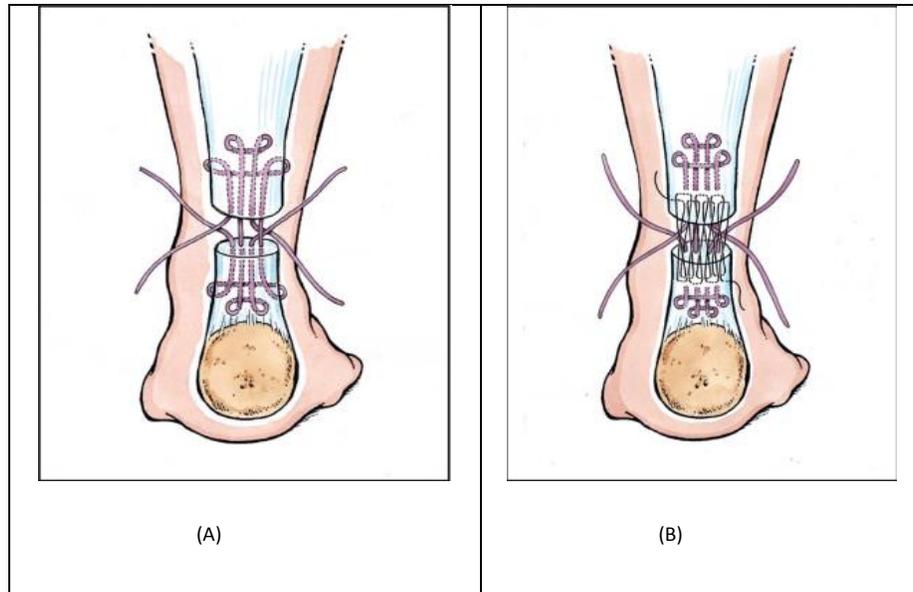


Figure 2.7: Illustration of the surgical technique.
 Image (A) show core suture and image (B) illustrate criss-cross suture. (Olsson, 2013)

2.3 Ultrasound Imaging

2.3.1 Definition

Ultrasound is the term utilized to express the sound of frequencies above 20 000 Hertz (Hz), further than the range of human hearing abilities. These frequencies are between 1-30 megahertz (MHz). The basis of ultrasound lies around SONAR (Sound Navigation and Ranging), that was created in the first half of the 20th century. It was used as a war weapon in the Second World War in the submarines. In about 1940, Commercial ultrasonic flaw detectors were created in Germany, France, Austria, UK and Japan. Upon which, the flaw detection was created for the human user in the mid twentieth century. Ultrasound occurs on the principal that emitted sound waves are used to generate images of the internal structure of the body. A better quality of the picture is obtained when shorter wavelengths or wavelengths that are inversely proportional to the frequency of the sound waves are used. On the other hand, the use of high frequencies is restricted due to their larger attenuation (loss of signal strength) in tissue and therefore shorter depth of penetration (Lutz & Buscarini, 2013).

2.3.2 Ultrasound Principles

The waves generated for the purpose of ultrasound are known as mechanical waves. When the mentioned waves pass through the tissues, they swing the muscle in a longitudinal, transmitting the mechanical wave. In response to the sound waves travelling through various tissues, the density either goes to a higher or a lower level in response to its ambient state without net motion. Sound waves are the mechanical vibrations that induce a change in refraction and cause suppression of the physical medium they pass through.

2.3.2.1 Transducers

The transducer (or probes) is the device which actually turns electrical transmission pulses into ultrasonic pulses and, contrarily, converts ultrasonic echo pulses into electrical echo signals. Ultrasound transducers have multiple piezoelectric crystals that are interconnected electronically and vibrate in response to an applied electric current. This phenomenon called the piezoelectric effect was initially reported by the Curie brothers in 1880 when they exposed a cut piece of quartz to mechanical stress generating an electric charge on the surface (Otto, 2013). In addition, they substantiated later the reverse piezoelectric effect, i.e. generation of vibration on the surface of Quartz when electricity is generated (Weyman, 1994). In response to that, they generate alternate areas of compression and rarefaction during transmitting via the human tissue. Sound waves may be defined as frequency (measured in cycles per second or hertz), wavelength (measured in millimetres) as well as amplitude (measured in decibel). Ultrasound transducers are created from thin discs made up of artificial ceramic material like lead zirconate titanate with the thickness (usually 0.1-1 mm) determining ultrasound frequency. Figure 2.8 illustrates the basic layout of the transducer. In various diagnostic applications, ultrasound is released in very short pulses in the form of a narrow beam if compared with a flashlight beam. When not generating a pulse, (mostly in 99% of the cases) the same piezoelectric crystal like a receiver used (Lutz & Buscarini, 2013).

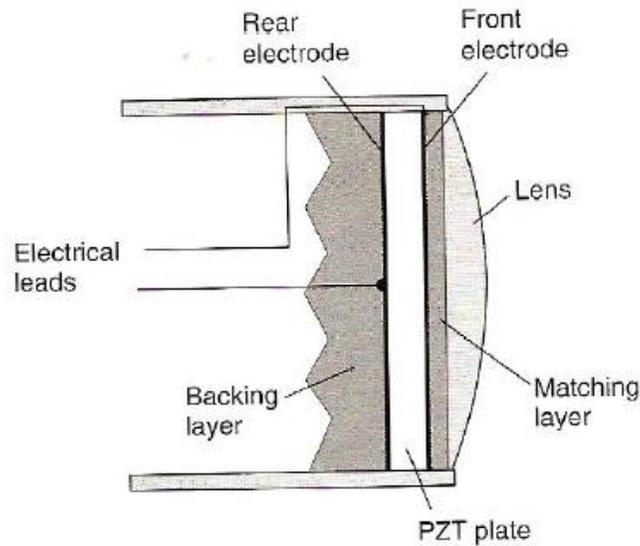


Figure 2.8: The basic component elements in an ultrasound transducer (adopted from Whittingham, 2007)

The simplest method of determining every scan lines that form a B-mode image is by physically changing the position of the transducer in order to allow the beam transverses as the pulse-echo cycle is repeated. Piezoelectric material releases waves that are detected by a transducer. The transducer then releases the ultrasonic waves and further switches to ‘receive mode’ to await the released ultrasound waves from tissue interface. This cycle is carried out a couple of times both temporally and spatially to produce images. Electronically arrayed probes are the commonest types of probes and are helpful because of their small size without moving their components. They have an array of transducer elements but are reduced to using frequencies less than 15 MHz.

2.3.2.2 Fundamentals of Acoustic Propagation

Frequency is defined as the number of ultrasound waves generated in a one-second interval. The unit for measuring it is Hertz (Hz). Humans have the tendency to hear sound waves in between 20 Hz to 20 kHz. Frequencies beyond this range are classed as ultrasound. For medical purposes, transducers are used with a frequency range between 1 to 20 MHz (Otto, 2013). Ultrasound is based on acoustic parameters like particle (or medium), pressure, velocity as particle displacement, temperature as well as density. If we apply this to another kind of medium and are displaced by external energy, it will be

called as partial or medium velocity and is the order of a few centimetres per second. Medium velocity is somewhat dissimilar from the rate at where the velocity travels via the medium. The velocity can be called phase velocity and is taken to be the velocity at which the ultrasound waves travel the medium. It is also known as sound propagation velocity and is denoted by c . The sound velocity in water is taken to be 1500 m/s. If the waves are displaced in the same direction as the medium, this kind of acoustic wave is known as a longitudinal or a compressional wave, and if the displacement is seen to be perpendicular to the direction of propagation, it is shear wave then. All these terms, when applied to the ultrasound, are shown in Figure 2.9 (Shung, 2004). Ultrasound wavelength is represented by λ , frequency by f , and velocity by c , are shown using an equation by $f\lambda=c$

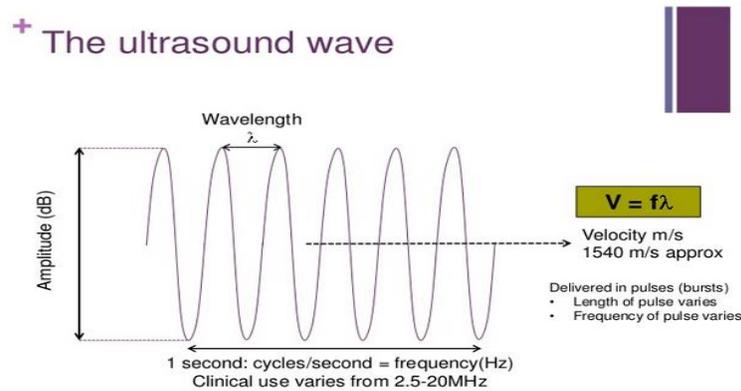


Figure 2.9: The ultrasound wave (Ferguson, 2012)

2.3.3 Ultrasound Tissue Interaction

Ultrasound is characterized by various properties as it transverse the tissues. Reflections of these waves as they pass through the tissues have various wave properties (like fat, muscle and blood). The degree of reflection is based on the magnitude of this variation. The left over waves is either reflected elsewhere or absorbed by the body tissues. As the waves interact with the body tissues, they either go deeper in the body tissues or get absorbed by the body. The waves diminish in strength as they go deeper in the tissues. This process is termed as attenuation and resembles quiet with the sound that gets fainter a person moves away from it. A few properties are shown in Figure 2.10:

+ Ultrasound/tissue interaction

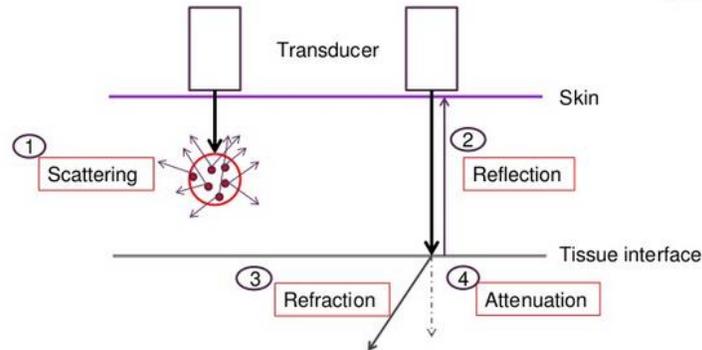


Figure 2.10: Ultrasound tissue interaction (Ferguson, 2012)

2.3.3.1 Reflection

The basis of ultrasound imaging is the reflection of the transmitted ultrasound signal from internal structures. These waves are reflected at tissue boundaries and interface and greatly are dependent on the acoustic impedance amongst two tissues and their angle of reflection. When these waves come across a large, smooth interface of two body tissues with varied acoustic impedance, the sound energy generated is reflected back to the transducer (Chan & Perlas, 2011). Boundaries of the smooth muscles act as specular reflectors. The returning echo will be missed by the transducer. It holds a relatively important role for the physician when it comes to the insertion of needles in a steep direction to reach the structures that lie deeper in the body. The waves that are deflected are constant for every interface, and the waves deflected vary with angle because the angle of incidence and reflection are equal. Thus, the optimal return of reflected ultrasound occurs at a perpendicular angle (90°). It is necessary to remember this fact for generating images when performing an ultrasound.

2.3.3.2 Scattering

When the wavelengths of the sound waves are reflected from the surfaces with irregularities, and they are similar to the irregularities in size, the energy of the wave is scattered everywhere. This is termed as scatter. For such cases, echo intensity is affected by as the echoes reflected via a large range of

angles. But, the positive result of scattering is when the return of some echo to the transducer occurs irrespective of the angle of the incidence. Various tissues appear in the images as they contain many small scattering structures. The speckle signal supplies the observable texture of organs and tissues due to the interface numerous scattered echoes generated in the volume of the incident ultrasound pulse (Hangiandreou, 2003). The echo that comes back to the transducer has a weaker intensity and can be appreciated when many scattered wavelets collectively produce an effect. Speckling is thus termed as the textured appearance of a tissue type because of the interference pattern (Palmer & Palmer, 2014).

2.3.3.3 Refraction

Refraction is termed as an alteration in the direction of the sound wave that is travelling as it strikes the tissue surface of two tissues that have varying speeds of sound transmission. It is one of the significant sources of wrong localization of a structure on an ultrasound image. The speed reduces when travelling through fatty tissues as well as soft, bulky tissues (Cox, 2013) . Refraction also occurs in various unplanned ways when it comes to forming an image thus causing “double-image” artefact (Palmer & Palmer, 2014).

2.3.3.4 Absorption

Absorption occurs as the sound waves pass through tissues. They cause the particles to transmit waves, producing vibrations as well as friction, heat generation and also a decrease in the sound energy. The rate of absorption is larger in the cases that the amount of large particles is increased in the tissue causing a tissue to have great viscosity. Absorption depends on the level of frequency of sound wave generated by much more absorption at higher frequencies. It also depends on the structure and the composition of the tissue. Taking in consideration that the tissues have increased collagen content like tendons as well as cartilages, they exhibit a higher rate of absorption as compared to other tissues having a lower rate of absorption because of increased water content. Water, urine as well as amniotic fluid and blood thus have low absorption as well as low attenuation. Table 2.1 shows various body tissues and their rate of absorption. Estimates of the contribution of absorption to attenuation are variable, but for many tissues absorption is the dominant loss mechanism (Martin, 2010).

Table 2.1 The absorption of ultrasound by different body tissues.
(Brown, Smallwood, Barber, Lawford, & Hose, 1998)

Tissue absorption (dB cm ⁻¹ MHz ⁻¹)	
Blood	0.18
Fat	0.63
Liver	1.0
Muscle	1.3-3.3
Bone	20
Lung	41

2.3.3.5 Attenuation

Attenuation is the process of sound waves passing through the tissue such that there is a loss of energy and there is a reduction in the distance travelled by the wave (Cox, 2013). Absorption is just one kind of attenuation, and it was described in detail above. Attenuation also takes place in the cases that sound waves are reflected or refracted in the means that they do not return to the receiving transducer. During the time that sound waves travel among different tissues, some of the waves are reflected. This is because of the variance in tissue density across that interface, also termed as impedance mismatch. Attenuation also is varied with various body tissues and is found to be highest in bone and smaller in muscle and solid organs, and least in blood for any provided frequency (Suri, 2008).

2.3.3.6 Acoustic Impedance

This is seen when the sound waves travel through two different kinds of tissue types having different densities. Some of the waves will be deflected some of them will continue longitudinally from the tissues. The greater variation at the tissue interface, the higher contrast of the image is formed (Hill, 1986).

Mathematically, the speed of sound is given by this equation (1):

$$c = \text{Sqrt}(k/\rho) \qquad \text{Equation (1)}$$

Where c , is the speed of sound, ρ (rho) mass density and k is the stiffness measured in units of Pascal (Pa). The acoustic impedance of a medium z is a measure of the response of the particles of the medium in terms of their velocity, to a wave of a known pressure.

Acoustic impedance is illustrated by the following equation (2):

$$z = p / v \quad \text{Equation (2)}$$

Where p , represents the local pressure and v represents the local particle velocity (Martin, 2010). It is quite similar to electrical impedance (or resistance R), and that is the ratio of the voltage (V) applied to an electrical component (the electrical driving force or pressure) to the resulting electrical current (I) which goes through it (the response), as given in Ohm's law: $R = V / I$. To determine the acoustic impedance, density (ρ) and stiffness (k) is taken into consideration.

Acoustic impedance can be described using equation (3):

$$z = \text{Sqrt}(\rho k) \quad \text{Equation (3)}$$

By combining equation (1) and equation (3), acoustic impedance (z) can be rewritten as:

$$z = \rho c \quad \text{Equation (4)}$$

Table 2.2 The acoustic impedance and velocity of sound for different media. (Shung, 2004)

	Acoustic Impedance (kg m⁻² s⁻¹)	Velocity of sound (m s⁻¹)
Air	0.0004 x 10 ⁶	330
Water at 20 °C	1.48 x 10 ⁶	1480
Soft tissue	1.63 x 10 ⁶	1540
Muscle	1.70 x 10 ⁶	1580
Bone	7.80 x 10 ⁶	4080

The sound wave while transmitting from one medium meets an interface with another medium of varying acoustic impedance, a few waves are transferred to the second medium while others are reflected backwards as well as are scattered. Ultrasound waves are usually deflected at a large interface and are scattered by a small target (Hamid, 2011).

The values of velocity of sound as they pass through various biological media, from which it could be computed that the wavelength at 1 MHz is of the order of 1.5 mm in soft tissue, are shown in Table 2.2. Acoustic impedance is the sum of the density and velocity of the sound for the medium and is therefore very low for air (Brown et al., 1998).

2.3.4 Ultrasound Usage and Importance in Medical Field

The use of ultrasound is very common nowadays in the field of medicine for diagnosing as well as for intervention. The use of ultrasound was proposed in 1940 by Gohr and Wedekind for medical diagnosis. It was later used to screen pregnant woman in the 1980s. There are various ranges of wave frequencies that are used to obtain images of the various parts of the body:

- 3–5 MHz intended for abdominal areas
- 5–10 MHz intended for small and superficial parts and
- 10–30 MHz intended for the skin or the eyes.
- Ultrasound not fits with the more traditional methods such as x-ray, but also holds special properties that are beneficial when compared with the rest of the competing modalities such as x-ray, CT and MRI.

More precisely:

- Ultrasound waves are basically nonionizing radiations that are taken to be safe for use till the latest researches.
- It is cost effective.
- It produces images in real time, unattainable at the current time through any other technique.
- It can point out resolutions in millimetres for the frequencies that are used by the clinicians these days.
- You can apply Doppler principle with it to inquire about blood flow information.
- It can be performed easily and can be performed at the bedside of the patient.

Ultrasound adds up many benefits as a diagnostic tool in various sub-specialities like gynaecology, cardiology, obstetrics, surgery, paediatrics, neurology as well as radiology (Shung, 2004). Ultrasound is very necessary for obstetrics as it is a non-invasive procedure, cheaper than other modalities and reliable for imaging capability. Ultrasound is equally important in the field of cardiology as echocardiography is learnt by every other cardiologist. Its scope in this field for the future is less as compared to obstetrics as other imaging parameters like CT and MRI are coming to improve the quality of the image.

2.4 Ultrasound Techniques for the Achilles Tendon

About 25 years ago, ultrasound imaging procedures were brought forward for diagnosing and analysing ruptures of the AT. Today that is the gold standard diagnostic tool for determining AT ruptures and tears ((Nielsen, 2004), (Gibbon, Cooper, & Radcliffe, 1999)). It has been established as an important and cost effective tool in the diagnosis of tendon problems. Also, it is a non-invasive procedure and is available at every medical setup. It is also very accurate when identifying the tears in AT and can be performed at the bedside of the patient and can be performed by the surgeon himself (Kälebo, Allenmark, Peterson, & Swärd, 1992).

Better images with spatial resolution are obtained when higher frequency musculoskeletal (MSK) ultrasound transducers are used providing a much detailed delineation of both the normal as well as abnormal tissues only to the superficial tissues as compared to the deeper tissues. Thus, MSK transducers are used in the superficial tendon assessment, being particular with the AT, its tendinopathy and rupture as well as post-treatment follow-up ((Kaplan, Matamoros Jr, & Anderson, 1990), (Lin, Fessell, Jacobson, Weadock, & Hayes, 2000)). When it comes to optimal ultrasonographic evaluation, the tendon is scanned on both their short and long axes moving the probe in a way that the waves reach towards the tendon perpendicularly (Fornage & Rifkin, 1988).

The highly ordered pattern of parallel collagen tendon fibres exhibits the highest echogenicity when tested perpendicular to ultrasound beam (Nunley, 2009). Tendons should be carefully scanned so that the beam is perpendicular to the tendon in whichever plane the imaging is done from as the sonographic beam generating a non-orthogonal sound may show that the image formed is hypoechoic

misleading the diagnosis towards tendinosis or tears because of anisotropy. At angles of insonation that is precisely perpendicular to the tendon, the reflected sound energy is at the maximum as it returns to the transducer showing tendon to be echogenic. Other than this angle of incidence, sound energy coming back to the transducer is reduced generating decreased echogenicity and artifactual appearance misleading to the diagnosis of tendinosis (Figure 2.11; (Chang & Miller, 2009), (Crass, Van de Vegte, & Harkavy, 1988))

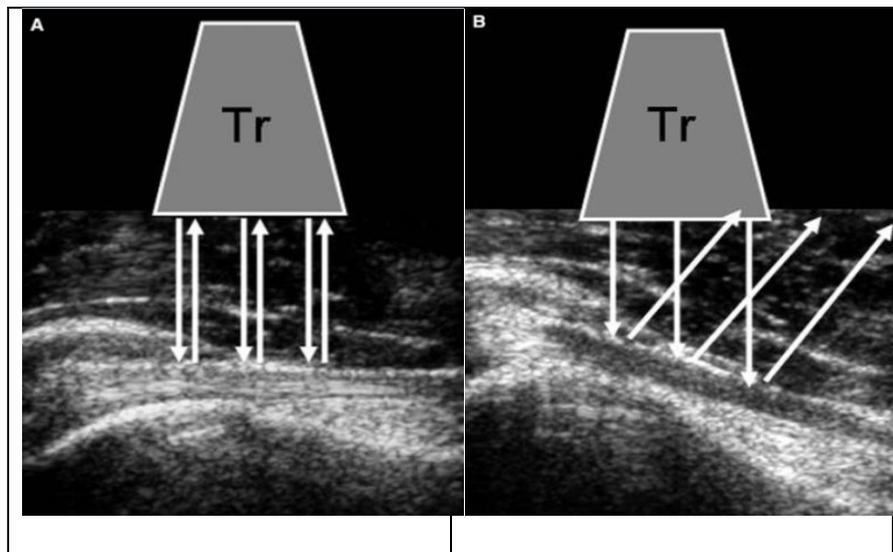


Figure 2.11: Sonographic anisotropy.

(A) when the transducer is perpendicular to the tendon, the sound waves (arrows) are mostly reflected back, and the tendon has an echogenic appearance; (B) when the transducer is oblique to the tendon, the sound waves (arrows) are not reflected back to the transducer, and the tendon has a hypoechoic appearance. (Tr) transducer. (Chang & Miller, 2009).

2.4.1 Normal Achilles Tendon

The tendon generally appears like a ribbon-like having a hypoechoic structure shown as two thin and regular echogenic bands (Figure 2.12). The thickness perceived was about 5.1 ± 0.4 mm (Maffulli, Regine, Angelillo, Capasso, & Filice, 1987).



Figure 2.12: Normal AT (middle part)

This thickness decreases as the muscle taper down before inserting in calcaneum. With both the insertion as well as thickness, two hypogenic areas were seen from which the largest one was at the front of the tendon just above the bony surface with well-marked out borders. This was seen as the deep bursa (Figure 2.13).

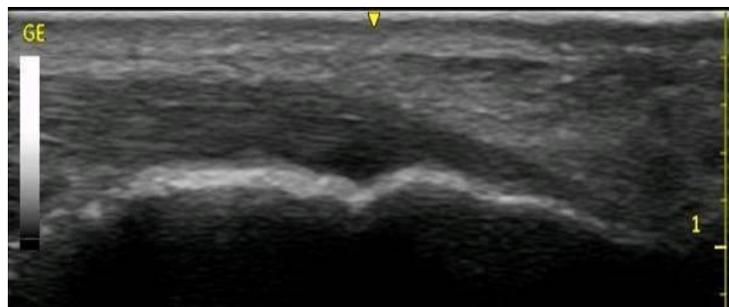


Figure 2.13: Normal AT (lower part)

The second hypoechogenic area was well marked and is the retrocalcaneal portion of the site where the bursa is inserted. AT originally has a uniform thickness as well as echogenicity in a longitudinal plane and appears flat or with concave anterior margin in the transverse margin. Normally, the tendon appears fibrillar when performing ultrasound in both short as well as a long axis as the sound bounce back towards the transducers from the collagen fibres whose intrafascicular interface acts like specular reflectors (Figure 2.14).

Normally, AT is echogenic in nature when sounds interact with parallel fibrillar lines in the longitudinal plane and appear round to avoid in contour when seen in a transverse plane. The echogenic lines are seen in a tendon when using an ultrasound probe frequency (Martinoli, Derchi, Pastorino,

Bertolotto, & Silvestri, 1993). When transverse imaging is performed, normal AT has a flat to concave anterior surface and measures about 4-6 mm in anterior to posterior (AP) diameter ((Fornage, 1986), (Khoury, Guillin, Dhanju, & Cardinal, 2007), (Kainberger et al., 1990)). 6 mm is taken as the maximum value for normal AT dimension; the values, however, are not constant and vary as per the shape of the AT. Before it gets inserted to calcaneus, AT lies above pre-Achilles fat pad which is a triangular area of fatty tissue also termed as Kager's triangle (Bleakney et al., 2007).



Figure 2.14: the Longitudinal sonographic image of a normal Achilles tendon (arrow) presenting consistent thickness and an echogenic fibrillar appearance. (Chang & Miller, 2009)

2.4.2 Abnormal Achilles Tendon

Similar to MRI, the diseased tendon may appear thicker on ultrasound (Hodgson, O'Connor, & Grainger, 2012). Calcific areas of the tendon appear as bright echogenic areas and neovascularization occurring after a tendon tear are appreciated by the help of Doppler ultrasound (Cardinal, Chhem, & Beauregard, 1998). Adding Doppler to the ultrasound has been beneficial for diagnosing blood flow and blood supply of the AT as illustrated in Figure 2.15.

By using Doppler in patellar tendinopathy, rich vascular supply was appreciated in an abnormal tendon showing the process of neovascularization (Bleakney et al., 2007). Time scale when acquiring an image is necessary as the anechoic fluid accumulates after a sudden tear in the tendon and then becomes difficult to differentiate from the associated tendon (Weinreb et al., 2014).

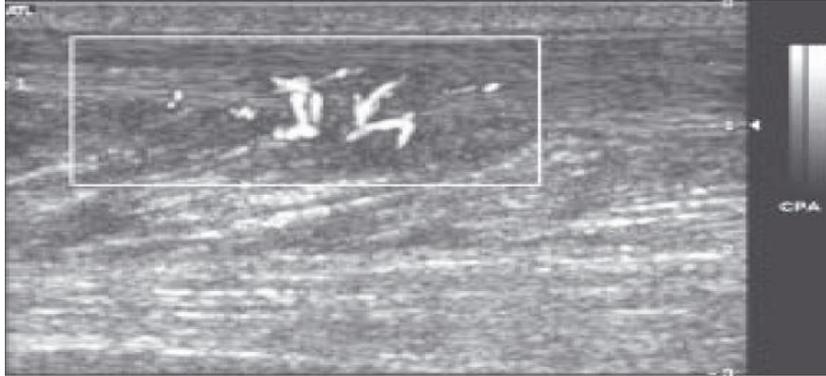


Figure 2.15: Longitudinal ultrasound image of the AT with power Doppler showing vascularity greater than before at the musculotendinous junction. (Bleakney et al., 2007)

2.5 Summary

This chapter focused on general medical background that is associated with the topic of this thesis. It gave a brief description of the anatomy of AT and the AT diseases, their ruptures as well as discussion about their treatment. The sections after that emphasized on ultrasound, and their interactions with various body tissues were discussed. At the end of the chapter, some light was thrown on the use of ultrasound in the field of medicine, and the normal as well as abnormal AT were taken into consideration and how they would appear in an ultrasound image were discussed.

Chapter 3 - LITERATURE REVIEW

This chapter will focus on previous researches that studied ultrasound and their implications in the medical field. It will also present methods and experiments that have been used by previous studies to explore and identify the most suitable for this study.

3.1 Introduction

Medical images play an important and crucial role in diagnosing diseases. They contribute to the early detections of the conditions, improving the likelihood of recovery and avoiding complications. They also contribute to reducing the overall cost of treatment and in some instances the recovery time. Digital image processing is concerned with the process of extracting information from the images using different computer programs. Computers, indeed, play an important role in all types of clinical image processing. Considerably, computers are more accurate in diagnosing the diseases. Accordingly, developing efficient algorithms and techniques in medical applications is the major concern for automatic analysis of medical images. Thus, there are many advantages of using computers in the medical field especially in diagnosis, surgery and treatment planning. There are different medical image models such as MRI, CT, ultrasound, and X-rays. Each model has a different purpose and uses different kinds of instruments. Ultrasound imaging has been shown to be a safe procedure in comparison to the other models and provides great results when used in clinical procedures, so it helps patients with different diseases (Nunley, 2009). Also, it can assist the clinician in deciding the proper pathologic process, as well as the location of symptoms, reaction to treatment and assessment of the whole treatment plan. The AT is known as the strongest tendon in the human body (Khan et al., 2005). Rupture and tendonitis are the most common medical problems affecting the AT. Doctors usually use an ultrasound scan to determine soft tissue rips, muscle injuries, or tendon tears due to injury. The present chapter provides an overview of numerous key major subjects linked to this thesis. It is categorized down into several sections. Section 3.2 gives a brief overview of the pre-processing technique in US

images. Section 3.3 gives an overview of segmentation methods used to determine ROI in US images. Section 3.4 discusses the methods available in texture feature extraction with an emphasis on texture analysis technique and feature reduction methods. Sections 3.5 and 3.6 provide a review of the classification methods implemented to images and techniques used to evaluate the classification results, respectively. Finally, Section 3.7 summarises the present chapter.

3.2 Image Pre-processing

Image pre-processing algorithms have as their main aim to convert pixel values to create an image that is more appropriate for subsequent analysis, which is needed for the precise application. (Gonzalez & Woods, 2001) Indicated that there is no standard theory for improvements of images and the observer being the judge of how good the specific techniques work. All the different pre-processing techniques involve features that will be described in the following work. These features are histogram equalization, brightness and contrast changing tool, smoothing of picture tool and filter feature. Ultrasound images have a huge role in the medical industry; the images are used in the diagnosis of many diseases, and they are often needed to decide on suitable treatment. When used for the purpose of diagnosis, an ultrasound image aims to indicate the disease source or to eliminate pathology. Through ultrasound images, the interior of the body can be seen with precision on parts like vessels, muscles and joints.

Ultrasound images need to be of very good quality to fulfil their purposes. As a result of unstable, noisy and unfinished data, pre-processing has an essential role in imaging. It is a vital step that is taken to initiate because it is needed to get high accuracy which will later be needed in other steps like classification. A common degradation is the occurrence of noise that may impact images that have different degradations. Enhancing images due to many reasons has been a focal study point in many studies over the years. During this time many algorithms for developing ultrasound images in many ways have been initiated and applied.

3.2.1 Image Noises

Various medical images suffer from the presence of noises, which hide some of their features and make ultrasound images complicated for the nonprofessional to interpret them hence not being able in some cases to decide on poor quality images. Noise often results in a degradation of images quality. Thus, image processing applications work to overcome this problem or minimize its impact.

Noise classification depends on the type of source and cause. According to (Koli & Balaji, 2013), the noise is classified into two main categories: blur noise such as Gaussian noise, and impulse noise such as salt and pepper. A blur noise, on one hand, is distributed in equal proportions in the image. It modifies all the pixels' values either upward or downward. If it is moved in a certain direction, then it is called blur motion. On the other hand, impulse noise is distributed irregularly and modifies some selected pixels values from the image and keeps the rest of them unchanged. The probability between the modified and unmodified pixels is shown in equation (5).

$$x_{ij} = \begin{cases} x_{ij} & \text{with probability } p \\ s_{ij} & \text{with probability } (1 - p) \end{cases} \quad \text{Equation (5)}$$

Where s_{ij} indicates the noiseless image pixel and x_{ij} the noise is replacing the original pixel.

There are various kinds of noise models which are classed as impulse noise. The most common ones are salt and pepper impulse noise, random valued impulse noise, Gaussian noise, additive noise and multiplicative noise (Agrawal & Verma, 2013). Speckle noise is classified as one of the multiplicative noises (Gonzalez, Woods, & Eddins, 2004) and is considered to be one of the most common noises affecting ultrasound images. Speckle is not actually noise in the usual engineering means because its texture frequently holds helpful information on the image which is viewed (Loizou & Pattichis, 2008). Therefore, the challenge is how to remove it without losing any image characteristics during the de-noising process. Speckle noise is random in an image with bright and dark spots and the interference of the returned signal with the small particles causing a granular pattern in the image. One of the main aims of speckle noise reduction (Narayanan & Wahidabanu, 2009) is to enhance the human capability to indicate normal and pathological tissue; de-speckling is the second aim; it is an elementary phase for

several image processing stages because it enhances the accuracy of subsequent tasks such as segmentation and registration.

3.2.2 Image De-noising

Given the fact that most ultrasound images deteriorate as a result of speckle noise, using de-noising algorithms is an essential step in improving these images. De-noising is the first stage in medical image processing methods. Different algorithms and filters have been introduced to enhance degraded images without losing their important features (Loizou & Pattichis, 2008). Image de-noising allows the interpreter to see details in images and perform quick and accurate visual examinations. Image de-noising algorithms are widely used with great success in the medical image field. De-speckling techniques (Loizou & Pattichis, 2008) can be categorised into linear filtering, nonlinear filtering, anisotropic diffusion filtering, and wavelet filtering. Linear filters such as Mean (Pitas & Venetsanopoulos, 1986), Wiener (Ghael, Sayeed, & Baraniuk, 1997), Lee (Lee, 1980), Frost (Frost, Stiles, Shanmugan, & Holtzman, 1982) and Kuan (Kuan, Sawchuk, Strand, & Chavel, 1985) tend to smooth the image and blur sharp edges. They remove the noise by calculating the central pixel intensity using a moving window, which calculates the intensity value from the average intensity values of the neighbouring pixels (Glasbey & Horgan, 1994).

Nonlinear filters, however, are based on nonlinear iterative algorithms for better preserving image quality (Johansen, 2011). The median filter is the most popular nonlinear filter. It overcomes the effect of impulsive noise while preserving edge information (Thivakaran & Chandrasekaran, 2010). Diffusion filtering is categorised as nonlinear filtering. It includes filters based on Anisotropic Diffusion (AD), which removes the noise from an image by solving a partial differential equation (Loizou & Pattichis, 2008). (Perona & Malik, 1990) introduced a new definition of scale-space through AD and suggested a function where the diffusion is prevented across the edge and increased parallel to it. (Yu & Acton, 2002) presented Speckle Reducing Anisotropic Diffusion (SRAD) to suppress speckle whilst maintaining edge information. SRAD's advantage in comparison to (Perona & Malik, 1990) is that it is replacing the threshold with an estimation of the standard deviation (Krissian, Westin, Kikinis, & Vosburgh, 2007). In addition, the performance of SRAD is affected by the selection of the threshold value (Lui et al., 2012). Wavelet algorithms are known as new methods for solving problems related to

image processing (Achim, Bezerianos, & Tsakalides, 2001). By transforming the image to the wavelet domain, we can apply Wavelet algorithms to remove noises. It is important in wavelet transform to select a function that has an integral is zero in time-domain. When comparing it to the traditional Fourier analysis, the wavelet transform has the ability to analyse the function in the spatial domain and time domain (Khoury et al., 2007). Wavelets have been used in a great range of applications such as signal processing, data compression, smoothing and image de-noising (Sifuzzaman, Islam, & Ali, 2009).

3.2.2.1 Linear Filtering

Linear filters too tend to smooth the image and blur sharp edges; they demolish lines and other fine image features and function weakly with the existence of signal-dependent noise (Koli & Balaji, 2013). There are many types of linear filters such as Mean, Wiener, Lee, Frost and Kuan filters as mentioned in the previous section. Mean filter or average filter is the simplest one of all filters. It replaces each pixel by the mean value of its neighbours. This leads to decreasing the amount of intensity difference between one pixel and the next. The Mean filter uses a sliding-window to replace the centre value. The window, or kernel, is usually a square that can have different dimensions (e.g. [3x3], [5x5], [7x7], etc...). Often, a 5x5 square kernel is utilised, though larger kernels (e.g. 7x7 up to 33x33 squares) may be implemented for more severe smoothing. The drawback of the Mean filter is that it sometimes blurs the edges that might be a problem if sharp edges are required in the output image (Patidar & Gupta, 2010). The Wiener filter's (Jain, 1989) objective is to decrease the amount of noise present in a signal through comparing it with an estimation of the desired noiseless signal. It takes a statistical approach to remove the noise from a signal and considers as outstanding when it comes to noise reduction or de-blurring of images. Wiener filter (Motwani, Gadiya, & Motwani, 2004) process requires the information on the spectra of noise and original signal, and its model complexity control corresponds to the window size.

3.2.2.2 Nonlinear Filtering

Noise filtering must preserve edges. Linear methods have easy mathematical calculations (Pitas & Venetsanopoulos, 1986) but distort the image edges. These drawbacks led to the use of nonlinear filtering in image processing. There are different types of nonlinear filters such as the median filter (Patidar & Gupta, 2010) and the Weight Matrix Filter (Ko & Lee, 1991). Filters using this technique

replace each pixel value with the median of the grey values in the region of the pixel. Median filters are very operative in eliminating salt and pepper and impulse noise whilst maintaining details in the image since they are not reliant on the values that are considerably different from ordinary values in the neighbourhood. However, the de-noising process using this filter may affect anything relatively small in size compared to the size of the neighbourhood because it cannot differentiate between noise and small details. The Weight Matrix is used in the weighted median filter. Their value is specified according to their position in the filter region. To compute the result, each pixel value in the original matrix times into the extended pixel vector, it is then sorted, and the median is taken from the vector centre. However, if there is intense noise or the probability of the impulse noise occurrence is high, the de-noising process comes to be similar to the standard median algorithm (Xiuqin, Yong, & Hong, 2008). Therefore, to solve this problem, many new algorithms are proposed such as “Adaptive Median Filters” (Hwang & Haddad, 1995), where a variable window size is used to reduce noise density.

3.2.2.3 Wavelet Filtering

Wavelet transforms on the basis of small waves, known as *wavelet*. The wavelet transform is one of the optimum tools in determining at which point the low-frequency area and the high-frequency area are. Wavelets have been used for de-noising in many medical imaging applications (Xu, Weaver, Healy, Jr., & Lu, 1994), (Unser & Aldroubi, 1996), (Quiroga & Garcia, 2003). According to (Pizurica, Wink, Vansteenkiste, Philips, & Roerdink, 2006) a general procedure is composed of the following steps: (i) compute the discrete wavelet transform, (ii) eliminate noise from the wavelet coefficients and (iii) reconstruct a de-noised signal or image through applying the inverse wavelet transform. The Discrete Wavelet Transform (DWT) algorithm applies a scaling function to decompose the signal into components of different resolutions (Nielsen, 2007). Since images are two dimensions DWT extends to two dimensions. By applying the forward DWT, the image is decomposed into a number of detailed coefficients, then different kinds of filters are applied between forwarding and backwards DWT (Pizurica, 2002). (Donoho, 1995) was the first to apply Wavelet Shrinkage (WS) method to denoise images, it is a signal de-noising method based on thresholding idea to the wavelet coefficient. Since most speckle noise is found in high frequencies, WS reduces the magnitude of the transform's coefficients according to certain criteria. Complex wavelet transform gives better performance than

DWT in areas of complex and modulated signals such as medical and geophysics data (Raj & Venkateswarlu, 2012). Dual-tree complex wavelet shrinks the wavelet coefficients in order to denoise the image. It is found that dual tree complex wavelet transforms results in less error than other methods and gives better results in the denoising stage by preserving the edges and textures with the minimum amount of redundancy (Selesnick, Baraniuk, & Kingsbury, 2005).

3.2.2.4 Diffusion Filtering

The isotropic nonlinear diffusion was first introduced by (Perona & Malik, 1990). It removes the noise from the image using partial differential equation (PDE) (Agrawal & Verma, 2013). Although linear smoothing filters remove noises and smooth the image they have a side effect, which is blurring or dislocating the semantically meaningful edges of the image (Tsotsios & Petrou, 2013). In order to overcome this unwanted effect, the anisotropic diffusion (AD) has been used. AD filters are implemented to successfully smooth images and give viable denoising tools with significant preservation for edges and region boundaries. There are several parameters for controlling the result of AD filtering, and it is important to estimate these crucial parameters properly. According to (Tsotsios & Petrou, 2013), these parameters are the conductance function, the gradient threshold parameter and the stopping parameter form a set of parameters that describe the behaviour and the amount of the diffusion. Determining the Coefficient of Variation is important in speckle filtering. (Yu & Acton, 2002) have proposed an edge sensitive extension of the conventional adaptive filter. They also modify Lee and Frost filters to enact the AD equation. The edge detector function in SRAD uses the instantaneous coefficient of variation to detect speckle noise in images (Li, Wang, Xiao, & Lu, 2012). At edges or in high contrast regions, this function presents a large value and a small value in the homogeneous area.

To overcome SRAD sensitivity to outliers and avoid over-smoothing (Zhi & Wang, 2008) proposed using eight-direction edges instead of four. Despite SRAD having a dynamic threshold value, it is still sensitive to the selection of that value and takes advantage of that image by having separability characteristics (Lui et al., 2012). They use the separability characteristics based on statistics to help SRAD detects edges of the ultrasound images. They divide the image into regions depending on the image intensity and modifying the diffusion coefficient of SRAD with separability coefficient to detect edges and control the amount of smooth. Unlike SRAD, Nonlinear Coherent Diffusion (NCD) (Abd-

Elmoniem, Youssef, & Kadah, 2002) combines three different models namely, isotropic linear diffusion, anisotropic diffusion and mean curvature motion. With regards to speckle extent and image anisotropy, the NCD model changes gradually from isotropic diffusion through anisotropic coherent diffusion to, finally, mean curvature motion.

3.2.3 Comparison between different de-noising techniques

A Comparison of some de-noising techniques is summarized in Table 3.1:

Table 3.1: The advantages and disadvantages of different de-noising techniques

De-noising Techniques		Advantages	Disadvantages
Spatial Domain Filtering	Mean	Simple Easy to implement Optimal for Gaussian noise	Tend to blur sharp edges Perform poor in the presence of signal dependant noise
	Wiener	Straightforward to design Optimal for Gaussian noise Optimal in that they minimize the MSE	Its model complexity control corresponds to choosing the window size
	Median	Simple More robust than mean by relies on the median of the data instead of the mean Edge preserving Less sensitive to extreme values	More expensive to compute than smoothing filters Fine details may be erased Rounding the corners
	Lee	Reduce speckle noise Based on a linear speckle noise model and minimum MSE	choosing the moved window size It is no directional It tends to ignore speckle noise in the areas closest to

			edges and lines
	Kuan	<p>Is a generalisation of the Lee filter</p> <p>Optimal for Gaussian noise</p> <p>Can effectively reduce speckle in homogeneous areas</p>	<p>Need to compute the ENL parameter</p> <p>Does not make an approximation on the noise variance within the filter window</p> <p>Do not perform well on the edges</p>
	SRAD	<p>It is nonlinear and directional</p> <p>Replaces the gradient-based edge detector with a coefficient of variation which is more suitable for speckle noise</p> <p>Preserving and enhancing feature edges</p> <p>Using variance as an edge detector</p>	Sometimes blurring the edges
Transform Domain Filtering	Discrete Wavelet Transform	<p>Removes high frequency while retaining the lower ones</p> <p>Thresholding of small coefficients without affecting significant features of the image</p> <p>Provide time and frequency localization simultaneously</p>	<p>Less efficient and natural</p> <p>Lack of directionality</p> <p>Edge information is spread across frequencies</p>

3.2.3 Image Enhancing

Enhancing ultrasound images is an essential branch in digital image processing. Through the procedure of image enhancement, an image with little awareness regarding the degradation can be improved. There are numerous methods of image enhancements that have a purpose of enhancing the quality of images, in addition to improving images visual appearance so that they are more appropriate to analyse. The standard aim of enhancing images is to alter their features to tailor it for a precise objective and throughout this procedure, there are characteristics of an image that can be enhanced. Methods of enhancing digital images can offer some options to develop the apparent quality of images. The suitable option for specific methods is vastly impacted by the modality of images and visual features.

Image enhancement knowledge used at present can be split into two sections: the enhancement technique regarding spatial domain and the enhancement technique regarding frequency domain. The former is placed in the processing of a 2D area, which is of the grey value of every pixel that is processed. Numerous algorithms for enhancing have been initiated. The enhancement technique regarding spatial domain handles the images directly in a 2D area, and this method directly deals with the image pixels. The pixel values are manipulated to reach the required enhancements. There are several enhancement methods regarding a spatial domain that works by manipulating image pixels such as the logarithmic transforms technique and the histogram equalisation technique. Enhancement methods regarding spatial domain are especially useful to modify grey levels of pixels, and thus the whole image contrast is altered. However, they tend to change the image in uniform means, and that often results in an unwanted outcome. Methods such as histogram equalisation can be very useful in many images.

3.2.3.1 Histogram Equalization

A popular alteration method used global histogram modification approach is Histogram Equalization (HE) (Nam & Choi, 1998). A key plan is to reassign the pixel intensities to redistribute intensity values so that the optimum uniformity is achieved. If $H(i)$, is considered the histogram of an image of the size $M \times N$ and a range of intestines of $[G_{\min}, G_{\max}]$ (Cheng et al., 2006). The initial image

intensity I_{org} may be planned into the consequential intensity I_{org} with the use of the HE method in the following equation:

$$I_{new} = G_{min} + (G_{max} + G_{min}) \sum_{i=0}^{I_{org}} (H(i)/(M \times N)) \quad \text{Equation (6)}$$

The HE method is easy and effectual when it comes to its function of image enhancing with low contrast if there is only one object and if there is no contrast difference between an object and its background. To improve the HE method, a technique known as the Multipeak histogram equalisation technique has been elaborated ((Wongsritong, Kittayarusiriwat, Cheevasuvit, Dejhan, & Somboonkaew, 1998), (Cheng & Shi, 2004)). The range of grey levels in this technique is $[x_0, x_L]$, at least one mid-nodes x_i ($i = 0, 1, \dots$) was determined by Mean, Median or through a number of peaks that present in the histogram. Then, the original histogram is divided into numerous parts which are equalized individually. An additional global histogram alteration is the histogram stretching ((Nam & Choi, 1998), (Wilson, Hargrave, Mitra, Shieh, & Roberson, 1998)). It makes use of a linear transfer function:

$$I_{new} = G_{min} + (I_{org} - I_{min}) \times (G_{max} - G_{min}) / (I_{max} - I_{min}), \quad \text{Equation (7)}$$

Where $[I_{min}, I_{max}]$ represents the intensity range of the original image and $[G_{min}, G_{max}]$ represents the intensity range of the resulting image. The global histogram alteration has done nothing when it comes to texture improvement, as the order of the grey levels of the original image cannot be modified.

Approaches for local processing are researched for image contrast improvement. Through the changing of pixel intensities, there are numerous methods for contrast image enhancements. One method of this is based on nonlinear mapping techniques such as the local histogram method and the bilinear method ((Nam & Choi, 1998), (Kim, Park, Song, & Park, 1997), (H. Li, Liu, Wang, & Lo, 1996), (Petrick, Chan, Sahiner, & Wei, 1996)). The application may be based on characteristics, and local features can be achieved through edge detection or through the use of local data statistics including standard deviation and mean. Moreover, the nonlinear mapping can be on the basis of “a gradient and/or statistics and the nonlinear functions” ((Nam & Choi, 1998), (Kim et al., 1997), (H. Li et al., 1996),

(Petrick et al., 1996), (Wilson et al., 1998), (Singh & Al-Mansoori, 2000)). The nonlinear mapping can also be adaptive histogram equalisations ((Wilson et al., 1998), (Pisano et al., 1998)).

3.2.3.2 Contrast Limited Histogram Equalization

The histogram in the Contrast Limited Histogram Equalization (CLHE) is cut at a certain threshold followed by applying of equalisation. CLAHE is an adaptive histogram equalization technique ((Garg, Mittal, & Garg, 2011), (Stark, 2000)). In this method, image contrast is modified and developed by application of CLHE on segments of data instead of applying it to the whole image, these segments are known as tiles. With the usage of bilinear interpolation, the surrounding tiles are then put back together (Sasi & Jayasree, 2013). In the homogeneous segments, image contrast may be limited, thus amplifying of image noise can be prevented. Effectual assessment of biomedical images has been achieved by CLAHE. CLAHE is suggested for use as a detector of abnormalities in dense mammograms (Pisano et al., 1998). This research shows the impact of CLAHE methods on images of myocardial perfusion in colour space.

3.2.4 Summary

In the literature, the low contrast and the interference with speckle noise remain the main challenge for ultrasound images. Notably, the visual observation and interpretation remain difficult because of the speckle. Reducing the speckle noise without hindering the important features of AT images that are used for the diagnosis process is the main task of the image pre-processing stage. Technically, different denoising algorithms are used for different image noises. Therefore, wavelet approaches image is transformed into wavelet domain and noise where the signal is processed at a different scale. However, this process is suffering from the increment of the time complexity.

On one hand, moving window is used to reduce the speckle noise by the linear and nonlinear filter techniques. On the other hand, the AD is performing well with the additive Gaussian noise. In this context, SRAD is introduced to handle the multiplicative noise image. Practically, SRAD is used to carry out the reduction directionally of the speckle using the edge function. The edges, in SRAD filter, are improved given that the achievement of the speckle reduction is directionally made by the function of the edge and, therefore, the function of the instantaneous coefficient of variation as an edge detector

in the speckled images. SRAD is able to eliminate the speckle without hindering the useful information of the image and the importance of the image edges. In comparison to the conventional anisotropic diffusion, SRAD provides superior performance. To improve the image contrast, CLAHE algorithm is used, which is an improved version of the AHE. Technically speaking, CLAHE partitions the images into contextual regions of pixels, which are alternatively called tiles. In CLAHE, the enhancing operation is achieved inside those tiles but not in and for the whole image. Thus, the contrast in the homogeneous regions is restricted to prevent over-amplification of the noise.

3.3 Image Segmentation

3.3.1 Introduction

Over the past years, automatic image segmentation has grown to be essential as a result of the rapidly increasing field of applications in biomedical images. The key aim of the partitioning is to divide an image area into meaningful smaller regions. However, because of complications in ultrasound imaging, segmentation is extremely not easy task. Data quality greatly impacts imaging segmentation. Image artefacts including speckle, attenuation and shadows introduce complications to segmentation. This is because of the acquisition orientation dependence which may lead to lost image objects borders. Another difficulty that adds on to the complexities is the low levels of contrast between the areas of interest. There are numerous techniques to perform image segmentation. The threshold method works by differentiating the relevant object from the background through the selection of a suitable threshold value. On the contrary, the region and edge tactics aim to gather pixels that have similar characteristics (C.-m. Chen et al., 1999). Pre-processing images with the use of various kinds of filters is an essential step that should be taken to achieve significant segmentation. On occasion, normalising an image histogram to a suitable distribution is accompanied by a filter to lessen the speckle and improve the features of the image. (Li et al., 2011) found that, for light work load, segmentation can be carried out easily by allowing the user to manually find the borders of the object.

3.3.2 Manual Segmentation

Manual segmentation is the simplest method of segmentation. It involves tracing on all sides of the interested area by hand. This algorithm of segmentation leads to the optimum results when finding structures for certain tasks. Typically, a medical practitioner specialising in anatomy uses software with a mouse to draw around or fill in the targeted region of the image. A better choice can be through the outlining of object contours, which may be achieved separately with the use of a keyboard. This leads to great accuracy; however, the speed will be low. With the mouse, the speed is higher, but the accuracy is low, comparatively. Tracing of algorithms may utilise geometrical data such as ellipses to estimate where borders of the object lie. Although the estimations are not the most accurate, this is often used for the medical purpose. Manually carried out methods are all alike in the means of the time consumed to trace objects. Also, human resources are highly costly and therefore do not sufficiently provide the required clinical use of every day. The pixels that fit in the same range of intensities can be identified manually; however, this is not ideal as images can be very large. Regardless of all the complications, manual segmentation is widely used, specifically if objects are hard to outline because of shape and low contrast. For instance, tumour segmentation is typically carried out manually. Conversely, there is no completely automated technique which can give segmentation quality as good as manual segmentation does. So a semi-automated method is simple and rapid to apply and obtain results with acceptable segmentation quality.

3.3.3 Segmentation Methods

The needed time for hand implemented delineation of organ tissue by radiologists, is a key block in the procedure of treatment scheming. However, the supervision of professionals is essential because of law and moral implications of mistakes in the procedure.

Being able to dependably and easily characterise organs and objects enables clinicians to determine and measure them. Many image analysis issues initiate with the segmentation stage and thus this stage conditions the quality of the outcomes. Whilst reliable automated segmentation is a long term aim, a semi-automatic procedure may have a direct and essential effect through enhancing efficiency and constancy in the tasks. Rapidity and simplicity of usage are vital in medical practice. However,

segmenting of images is a complicated issue. It typically demands advanced information regarding objects that are being studied. In actuality, semantically steady and high-quality segmentation is an issue that cannot be differentiated from powerful artificial intelligence and has likely no precise or even agreeable answer. This has been acknowledged for a while, and thus many segmentation techniques have been suggested by researchers (Pham, Xu, & Prince, 2000).

Threshold based segmentation is frequently utilised to divide an image into parted regions depending on their intensity value and, then, comparing each pixel with the threshold value to decide whether the pixel is considered a part of the object or from the background (Noble & Boukerroui, 2006). Multiple thresholds can be used to segment an image into more than two segments. The critical step to achieve good segmentation is the selection of an appropriate threshold value. Many methods exist to find a suitable threshold for segmentation ((Sahoo, Soltani, & Wong, 1988), (Zhang, 1996)). A histogram is often a helpful tool as a starting point in establishing a suitable threshold value (Raju & Neelima, 2012). When the peaks in the bimodal histogram are sharpened, it is easy to find an automatic threshold value. Most of the medical images suffer from different noise factors, and that makes the histogram not as sharp as we would like it to be. Therefore, it is useful to smooth the histogram in advance (Gonzalez & Woods, 2001). While in edge based segmentation an image segmented into objects depends on finding out the edges. Edge-based segmentation relies on edge detecting techniques. Edge detection is the most common approach for detecting the discontinuity in the grey level. Several edge detection approaches have been introduced (Lakshmi & Sankaranarayanan, 2010). However, Sobel, Prewitt and Laplacian (Gonzalez & Woods, 2001) operators are the most traditional algorithms used. They use a gradient operator to detect edges in the image where most of them look for maximum and minimum in the first derivative to discover edges; Laplacian method looks for zero-crossings in the second derivative. Edge linking algorithm works to remove spurious edges and bridge gaps in object boundaries where no edges are detected. If the objects in the image are simple shapes, the approach of Hough transform (Richard O. Duda & Hart, 1972) can be used for detecting straight lines in the image. However, it is often time-consuming if the shapes are complex (Olson, 1999).

In the neighbourhood search, the algorithm starts looking for possible links in the small neighbourhood and for sharp edges post-processing, which is usually needed. The highest edginess value

and backtracking are used to overcome the starting point problem in this algorithm (Maintz, 2005). Despite the relative simplicity of computational processing in edge detection algorithms, the quality of its result depends on pre and post-processing algorithms. These algorithms have an important task to smooth the images before segmentation and generate closed boundaries after it. In the region based approach, the segmentation depends on the homogeneity of the objects. Regions that have a homogeneous area are considered as one object. While merging (top-down) approach depends on the similarity criterion to merge adjacent segments, the splitting (bottom-up) depends on inhomogeneous criteria to split segments (Pratt, 2001). To overcome over-segmentation edge, strength can be evaluated to decide whether two segments should be merged. In the region growing method seeds are picked up as a set of starting pixels, merging similar neighbours to the seed before removing it from the seed set. However, formulating the stopping rule is very crucial in the region growing approach (Maintz, 2005).

3.3.4 Fuzzy Clustering Segmentation

One of the most extensively used techniques to classify correctly the pixels of an image in decision oriented applications is known as image segmentation. This method works by separating an image into equal and non-overlapping areas on the basis of a specific likeness feature (Dong & Xie, 2005). There are several applications of this method, e.g. computer vision, medical image processing, image analysis, geographical information system and remote sensing. There are two main characteristics of an image that forms the basis of image segmentation. These are 1) *intensity* values comprising of discontinuity, which signify sudden or abrupt variations in intensity as edges, and 2) *similarity*, which signifies separating a digital image into areas on the basis of a pre-established likeness standard.

The cluster analysis comprises of only the statistical classification method for identifying if the individuals of a population belong to distinct groups by performing the quantitative assessment of various characteristics. Input data is distinguished into groups or clusters based on certain similarity standard, in a way that similar data objects are part of the same group. Distance, intensity and connectivity are some of the measures that may be used in clustering. The resultant partition enhances the comprehension of human beings and facilitates in making a more knowledgeable decision.

There are two kinds of clustering methods that are usually used to examine cluster creation, i.e. Hard Clustering and Soft/Fuzzy Clustering. In Hard Clustering, data is separated into various clusters, and each data element is part of just a single cluster. On the other hand, in fuzzy clustering, data is allocated membership levels, and these levels are then used to allocate data elements to one or more clusters or groups in the image/dataset. Data elements in soft/fuzzy clustering may be involved in more than one cluster while possessing some extent of membership values. Fuzzy cluster analysis may be used in data analysis, image segmentation and pattern recognition.

Real-world images involve some degree of uncertainty; hence, their segmentation leads to fuzzy regions. Information used in clustering techniques includes brightness and spatial location of pixels. These techniques are unable to distinguish image regions that have identical pixel intensities by taking into account just their pixel intensity. There is a high association between the pixels on an image, that is, the pixels right next to each other have almost the same feature data. Hence, the spatial relationship between the pixels next to each other is a critical feature that may play an important part in image segmentation.

In the past few years, much attention has been given to image segmentation with the help of fuzzy clustering. Some examples of objective function based fuzzy clustering algorithms are: Fuzzy C-Means (FCM) algorithm, Gustafson-Kessel algorithm (GK), Gaussian Mixture Decomposition (GMD) algorithm, Fuzzy C-Varieties (FCV) algorithm, Adaptive Fuzzy-C varieties (AFC) algorithm, Fuzzy C-Shell (FCS) algorithm, Fuzzy C-Spherical Shells (FCSS) algorithm, Fuzzy C-Rings algorithm, Fuzzy C-Quadric Shells (FCQS) algorithm and Fuzzy C-Rectangular Shells (FCRS) algorithm (Wikipedia, 2018). Out of these, the most extensively used fuzzy clustering method is FCM as it has the required features to deal with uncertainty and can store a high degree of information in contrast to the clustering methods (Pham & Prince, 1999).

FCM does not include spatial information on noisy images, because of which it is sensitive to noise and other image artefacts. In addition, the sole basis of FCM cluster allocation is the distribution of pixel intensity, because of which it becomes sensitive to intensity variations in the illumination or the geometry of the object (Liew, Yan, & Law, 2005). The disadvantages of FCM can be countered by including a pre-processing image smoothing step (M. N. Ahmed, Yamany, Mohamed, Farag, &

Moriarty, 2002), (S. Chen & Zhang, 2004), (Szilagy, Benyo, Szilágyi, & Adam, 2003). In the past few studies, several image segmentation methods that employ fuzzy clustering have been put forward.

A. Robust segmentation for Noisy Medical Images Using Fuzzy Clustering with Spatial Probability

A robust segmentation method has been formulated by (Z. Beevi & Sathik, 2009) that use histogram based FCM algorithm for the segmentation of medical images. Noise is first eliminated from the images by the algorithm, after which they are segmented. De-noising is carried out using sparse 3D transform-domain collaborative filtering (Dabov, Foi, Katkovnik, & Egiazarian, 2008). The parameters of the FCM are initialized using histogram so that convergence is prevented in local minima. The objective function includes spatial probability to make the algorithm more robust against noise.

B. Fuzzy Logic Information C-Means (FLICM) Clustering Algorithm

Fuzzy Logic Information C-Means Clustering (FLICM) algorithm has been formulated by (Krinidis & Chatzis, 2010) to counter the disadvantages in algorithms that have been presented in the literature (Hathaway, Bezdek, & Hu, 2000), (Wu & Yang, 2002) and (Łęski, 2003). In FLICM, a new factor is included in the objective function of FCM, and this factor has the following features (Krinidis & Chatzis, 2010):

- Local spatial and grey level information is included in a fuzzy manner to ensure robustness and noise insensitiveness.
- It controls the impact of the neighbouring pixels based on their distance from the central pixel.
- The original image is used as input, and pre-processing steps are not carried out to preserve image details.
- It does not include any parameter selection.

C. Fuzzy C-Means Clustering Algorithm (NFCM)

A Novel Fuzzy Clustering C-Means Algorithm (NFCM) was put forward by (Kannan, Ramathilagam, Pandiyarajan, & Sathya, 2009) for intensity inhomogeneities or weighted bias estimation and segmentation of brain MRI images of the same patient. A central knowledge technique has also been put forward to decrease the operating time of the algorithm. The Centre Knowledge Algorithm reconstructs the data matrix according to its relabelling mean value and the data is divided into groups.

This is followed by the creation of the distance table to present the distance between the elements in each group. The highest distance between the groups is calculated and the mean value is obtained.

D. Improved Spatial Fuzzy C-Means Clustering Algorithm (ISFCM)

An enhanced Spatial Fuzzy C-Means algorithm has been suggested by (S. Z. Beevi, Sathik, & Sentharamaikkannan, 2010). Due to HFCM converging more rapidly because the histogram of the image rather than the entire image is clustered by it, the input parameters for ISFCM are initialized by using the histogram based FCM (HFCM). Using *a priori* probability, incorporation of the spatial neighbourhood information is done into the standard FCM. Automatically decided in the fuzzy membership algorithm (S. Z. Beevi et al., 2010), incorporation of a priori probability is done for demonstrating how the central pixel is spatially influenced by neighbouring pixels. The cluster whose members are the majority in the pixels neighbourhood are assigned with a noise pixel by a priori probability and, for getting that pixel higher membership values when their neighbouring pixels have high membership values with the corresponding cluster, there is utilization of fuzzy spatial information. The advantage of ISFCM is that it can overcome the noise sensitiveness of standard FCM. The algorithm was made robust to noise and blurred edges due to spatial information being incorporated into the clustering process.

3.3.5 Segmentation in Ultrasound Images

The major drawback of ultrasound images is the weak quality and occurrence of speckle noise. The existence of noise reduces the resolution of the ultrasound image and results in low contrast, causing problems with segmentation. Thus, image quality needs to be enhanced to attain optimum results. Techniques dependant on the categorization of pixels and tracking of boundaries are susceptible to environmental noise and image inconsistency. Semi-automated computer approaches are utilised to enable radiologists to alter the parameters in segmentation to lead to the best result (Li et al., 2011). Thus, a strong model is required to help radiologists with finding the location of suspicious regions (Cheng et al., 2010). AT ultrasound images are able to guarantee that a great amount of data regarding Achilles tendinosis status is obtained. Via ultrasound images, the location of the injured tissue can be located, in addition, measurements of the tissue size can be made, and a conclusion on whether an operation will be required or not is also made through the images.

As well as echo features in images, other features may act as a judgment detector that presents possibility rupture like the AT contour and shape. The exact boundary detection of ultrasound images will provide accuracy position for ROI area. Additionally, the features of echo, shadow and reflection of the image will degenerate the quality of the image. The degrading of quality is a result of image nature which makes it hard to identify the image edges precisely, even when it is done by a professional physician.

An essential part to segmentation is medical segmentation which includes all the prevailing methods of segmentation. Currently, the most utilized techniques of segmentation in the clinical image field are the threshold value technique and the manually functioning technique performed by medical practitioners. Despite the simplicity of the application of the threshold segmentation technique, speckle and inconsistency in images cause difficulties resulting in dissatisfying results (Tian, Bao, & Zhou, 2003). Moreover, although the manual segmentation technique is moderately simple to use with acceptable results, it is very time-consuming for patients and doctors. Application of the semi-automated segmentation technique with the use of a computer is the perfect option for segmentation in the clinical field. Due to the images having noise and the target having feeble boundaries, the technique that based on boundaries is hard to segment accurately. (Aarnink et al., 1994) applies the nonlinear Laplace filter in the automatic application of prostate image segmentation. (Fan, Santago, Riley, & Herrington, 2001) implements the nonlinear wavelet threshold technique to indicate the boundaries created via lumen-intima which is the internal wall and external membrane of implanted images. Meanwhile, (Yoshida et al., 1998) study segmentation in the use of medical images in great depths basing it on active contour models. (Lee, Yan, & Zhuang, 2001) makes use of the dynamic programming for the segmentation of varying images in the medical field and through that great results were obtained. (Zhao, Zheng, Zhang, & Tian, 2013) makes use of the active contour model that functions on gradient vector flow principle; this has attained a precise impact on soft tissue tumours when it comes to segmenting images (Yan & Zhuang, 2003). (Cvancarova, Albregtsen, Brabrand, & Samset, 2005) proposed a method of imaging segmentation known as the snake model; it functions on the basis of gradient vector flow principle (Cvancarova et al., 2005). Due to cardiac image features such as noise and inconsistency of grey scale spread, (Cheng et al., 2010) proposed polarity filter; edge sharpens feature and usage of CV Snake

model in image segmentation of extracted ventricular boundaries. (Zhao et al., 2013) proposed a technique for the image classification of fatty liver; the method works on threshold segmentation to precisely analyse features of tissue that is both lesion and non-lesion in images.

3.3.6 Determining ROI

Features extraction and classification results vary due to different factors. The main factor is determining the most suitable ROI. It is an essential factor in initiating the feature extraction phase. Following the determination of the ROI, tasks will distinguish the area and categorise into either a normal or an abnormal class. Numerous techniques in the past have talked about ROI generation ((Su, Wang, Jiao, & Guo, 2011), (Shan, Cheng, & Wang, 2012a), (Liu et al., 2009)). A computer-aided technique functioning on the basis of feature classification was suggested to indicate ROIs in the breast ultrasound images (Su et al., 2011). Initially, the technique is made up of two stages which are the feature extraction stage and the self-organized map classification phase. (Shan et al., 2012a) implemented automated seed point selection and region growth algorithms to ROI sites. (Liu et al., 2009) put forward a fully automated technique for ROI generation which functions on the basis of a controlled local texture classifier and empirical information of breast imaging. (Basset & Mestas, 1993) investigate the impact of ROI sizes on prostate imaging with the use of the mean of co-occurrence matrices; they also report the ROI size impacting the discrimination outcome. (He et al., 2004) says that the size of the ROI has an inverted link with the variance in the extracted features; it also has a positive relationship with the average of extracted features. On the other hand, noise may greatly impact the measured texture from small ROI (Osicka, Freedman, & Ahmed, 2007). When analysing texture from a small ROI, noise that may impact the precision of the measured texture parameters and lead to mistakes in diagnostic reports needs to be taken into account (Osicka et al., 2007). Conversely, less extracted features are able to reduce the vulnerability of the features to noise distortion (Al-Kadi, 2009). (Hafizah & Supriyanto, 2012) put forward an automated ROI generation method for imaging kidneys with the use of texture examination to indicate the centroid of the region prior to the cropping of the ROI. (Ulagamuthalvi & Sridharan, 2012) proposed a method that implements co-occurrence matrix features along with grey level run-length characteristics for the determination of the seed point for images of the human liver. Following the determination of the seed point, images are segmented by the application of

the area growing algorithm with the use of grey space map and Otsu algorithm. (Shan, Cheng, & Wang, 2012b) put forward a technique for segmentation which uses a novel stage to enhance the quality of images. To begin with, the ROI is made to remove the complicated background. Following the decrease process of speckle, an algorithm functioning on the phase in max-energy orientation is developed to enhance the quality of the image. The authors suggest a novel gathering method known as neutrosophic I-mean which is for the detection of the lesion boundaries. A completely automated technique for breast imaging was proposed by (Liu et al., 2009). It is made up of a generation of ROI stage which is applied for dealing with the complex structure of images and a segmentation of ROI stage. Experiments show that the technique proposed is able to do segmentation of images rapidly and precisely. An automated seed point choice algorithm was suggested by (Shan et al., 2012b). The method formulated the texture, spatial location and size of the region. The areas obtained from the iterative threshold through the empirical method formula, and middle of the winning area was chosen as the seed. Following that, a completely automated technique for segmentation was initiated on the basis of the characteristics from the spatial and frequency based domains (Liu et al., 2009).

3.3.7 Summary

Segmentation of images in medical ultrasound imaging has a vital role in computer aided diagnosis. It is a procedure in which images are partitioned into subdivisions of components. The extent of the partitioning is done depending on the problem requiring a solution, that is, the segmentation ought to end when the desired object is made isolated in an application. After the suppression of image noise and artefact, segmentation is used to locate suspicious regions (ROIs). ROI is used to establish the significant region of digital medical images. The main goal of segmentation is to attain ROIs that have abnormal characteristics and identify the locations of suspicious lesions in ROIs. AT defects like ruptures can be very complicated to find due to their radiographic and morphological features that look like features of normal AT tissue. There are numerous techniques to identify image ROI and segmentation issues. Numerous segmentation methods that are used nowadays utilize the grey scale histogram. For example, thresholding; spatial and intensity technique (e.g. Region Growing) or others using fuzzy set theoretic approaches. Active contours which are known as snakes may be utilized for automatic segmentation of objects. The segmentation of the image is carried out with the use of series of

decision. An image may be segmented with the use of other techniques of segmentation and there is no ultimate technique for the use on all types of images.

3.4 Feature Extraction and Reduction

3.4.1 Introduction

Extracting the features of an image is a significant part of medical image studies because it gives the crude data to the following phases. (Selfridge, 1955) describes the feature extraction procedure as "the extraction of critical features from a foundation of unessential detail." feature extraction is an exceptional type of dimensionality reduction. It is carried out following the pre-processing stages in image processing methodology through isolating the unneeded traits from the applicable ones. The decrease in features gives a way to pick the characteristics which are most suitable for classification, in light of different criteria. It utilises a classifier to identify the chosen characteristics (Kumar & Bhatia, 2014). Extracting the features and reduction are stages in different clinical image models abnormality detection and classification. An ideal set of features ought to have effectual and segregating characteristics, while mostly decreasing the redundant characteristic space to avoid the dimensionality problem.

The dimensionality curse proposes that the density of the samples of the training data is very little to guarantee a significant approximation of a high dimensional classification with the available finite number of training data. For several advanced classification techniques including artificial neural network and support vector machine, feature vectors dimension not just influences the performance of the classification, it additionally determine the time of training. Therefore, how to extract useful features and make a good selection of the features is a crucial task.

3.4.2 Texture Analysis

The development of the computational methods is used for extracting the visual information and automated understanding of the content of the image. This is usually achieved through understanding the properties of the textures within the images. In the analysis of the images, the texture information can be

an important factor in the applications that involve the region's segmentation, objects and the subsequent classification and recognition processed.

On one hand, the texture, according to (Kurani, Xu, Furst, & Raicu, 2004), is considered as one of the most known features that are used for analysing and interpreting the images. On the other hand, the texture, according to ((Haralick, 1979), (Bharati, Liu, & MacGregor, 2004)), has no formal or sufficient definition. Practically, the texture has been only defined through various definitions that differ from one researcher to another. For instance, the given definition of the texture by (Bharati et al., 2004) is already quoted by (Russ, Matey, Mallinckrodt, & McKay, 1994) that defines the texture as “a descriptor of local brightness variation from pixel to pixel in a small neighbourhood through an image”. However, the texture in the current thesis is defined as “the spatial distribution of pixel grey value (intensity) of B-mode images”.

3.4.3 Texture Analysis Approach

There are many classification approaches that are based on textures features have been reported in the literature. However, the utmost comprehensive classification of texture analysis method was given by ((Materka & Strzelecki, 1998), (Castellano, Bonilha, Li, & Cendes, 2004), (Materka, 2006)), and (Bharati et al., 2004). They found that, based on the method employed to assess the interrelations of the pixels, texture analysis could be summed up within these four approaches: statistical based, model based, structural based, and transform-based. Many methodologies have been proposed to analyse and recognize texture in an automated fashion (Mittra & Parekh, 2011). Most image features have been constructed based on histogram and texture features (Tian, 2013). The statistical descriptor, co-occurrence matrices for four different angles, and run length encoding were used for extracting different texture features. Also, different histogram features such as skewness and kurtosis are extracted from samples after different segmentation methods were applied. There are many techniques that have been introduced to extract texture features. They are classed into spatial texture feature extraction methods and spectral texture feature extraction methods based on the domain that the texture is extracted from (Tian, 2013).

a) Spatial-Domain Based methods

In spatial texture feature extraction methods, the texture is measured from a group of pixels by computing the pixel statistics, which is derived from local features at each pixel in a texture image (Kurita & Otsu, 1993). It includes two different types, Co-occurrence Matrices and Autocorrelation Feature. Co-occurrence Matrices such as GLCM (Grey-Level Co-occurrence Matrix or the Grey-Level Spatial Dependence Matrix) can be used to calculate numerous useful textural properties to reveal details about the image content. (Haralick, 1979) Proposed the use of co-occurrence Matrix or grey level co-occurrence matrix. Also, he defines fourteen textural features measured from the probability matrix. To reduce time (Mohanaiah, Sathyanarayana, & GuruKumar, 2013) select four significant features, Angular Second Moment, Correlation, Entropy, and the Inverse Different Moment for implementation. (Tian, 2013) Stated that spatial texture methods are easy to comprehend and extractable from any shape, but on the other hand, they are sensitive to noise and distortion.

By contrast, autocorrelation functions can evaluate linear spatial relationships between primitives. The texture primitives can be individual pixels or a region with uniform grey levels (Xie & Mirmehdi, 2008). Large primitives provide rise to coarse, and the function decreases slowly whereas it reduces rapidly if the texture composed of small primitives (Kale, Mehrotra, & Manza, 2007).

b) Model Based Methods

Model based texture analysis uses the construction of an image to describe the texture and synthesize it. It utilises fractal and stochastic models to interpret the texture of images. The fractal model can be used for texture analysis, and it has been shown to be beneficial in modelling natural textures (Materka & Strzelecki, 1998). However, it does not have orientation selectivity and is not appropriate for defining local image structure (Materka, 2002). (Zhang & Tan, 2002) Stated that the key problem for the latter model-based methods is how to estimate their coefficients which are used to characterize texture images and how to choose the suitable model for the selected texture. Examples of model-based texture analysis approach are autoregressive (AR) models, Markov random fields (MRF), and fractal models.

c) Spatial-Frequency-domain based

In spatial frequency domain method the image is transformed into the frequency domain, and then the features are computed from the transformed image. Gabor filters (Tian, 2013) are intended to sample the whole frequency domain of an image. It has been used for extracting useful features and for a variety of texture classification and segmentation tasks (Ajain & Bhattacharjee, 1992). They can be tuned to respond to particular orientation and feature scales. Content-Based Image Retrieval (CBIR) has been a very active research field in retrieving image content. It uses low-level visual features such as colour, shape, texture to index images. CBIR requires high retrieval efficiency and preserves low computational complexity (Baaziz, Abahmane, & Missaoui, 2010).

3.4.4 Texture Analysis Methods

Quantitatively, some studies in the literature have been conducted to quantitatively characterizing the echo-texture of the B-Mode images. Regarding the principality, as reported by (Morris, 1988), when such diseases processes impact the structure of the tissue, a reflection of the altered ultrasound signal from the tissue should be obtained that, in turn, reflected in the texture feature values that are different to the normal tissue. For instance, the torn tissue transformation results in the changing the tissue properties such as density, elasticity and echogenicity. According to this rule, the textural features are expected to be derived from the tissue of tore where the normal tissue will be different.

As demonstrated in (Morris, 1988), the significant difference between the placenta of smoker and non-smokers people is highly affected by the grey level co-occurrence matrix (GLCM) features that already derived from the ultrasound images. Identically, according to (Tsai & Kojima, 2005), the use of the GLCM for extracting the texture features from the ultrasound images can be used in the classification of the heart disease. Based on this finding by (Morris, 1988), those writers came to the conclusion that the texture analysis of the ultrasound images has a great effect to clinically become helpful for the computer-aided diagnosis of the cardiomyopathy.

The image texture analysis of the endoscopic ultrasound (EUS) also received a massive body of research in the literature. The recent research by (Das, Nguyen, Li, & Li, 2008), pointed out that the

initial try to characterize the cancer of the pancreatic from the EUS images was stimulated. Those researchers explored out that the texture analysis was capable of distinguishing the cancer of the pancreatic from the chronic inflammation and normal tissue. Therefore, they indicated that the area under the ROC curve of 0.93. The differentiation between the malignancy of the pancreatic and pancreatitis was an application of the texture analysis of EUS images, which is already reported by ((Norton et al., 2001), (Loren, Seghal, Ginsberg, & Kochman, 2002)). Practically, they used it to analyse the lymph node metastasis in the patients who hold the oesophageal carcinoma. Both types of research came up with encouraging and challenging outcomes. (Xian, 2010) Reported that texture analysis can be applied widely to liver disease. Information extracted from the ultrasound liver images used in distinguishing between healthy and unhealthy livers tissue.

(Fellingham & Sommer, 1984), for instance, but not limited to, used the texture analysis for characterizing the structure of the in vivo tissue of the human liver and spleen. Many other promising results to quantitatively characterizing the tissue of diffuse liver disease were reported (Kadah, Farag, Zurada, Badawi, & Youssef, 1996). Based on the experimental results (Mojsilovic, Popovic, Markovic, & Krstic, 1998), the authors proposed that the features of the wavelet-transform might be a method with reliability for the texture characterization of B-mode liver images. The researchers in their research work (Yeh, Huang, & Li, 2003), used the GLCM and the features of the wavelet for classifying six grades of liver fibrosis and they came up with the result that the accurate optimum classification of two, three, four and six classes were, respectively, 91%, 85%, 81% and 72%. Other researchers ((Layer et al., 1990), (Wang et al., 2002), (Lee, Chen, & Hsieh, 2003), (Ribeiro & Sanches, 2009)) have also used the texture analysis for implementing the diseases that have a relation with the liver.

The application of the texture analysis also covered another medical subject such as the thyroid gland. The researcher (Smutek, Šára, Sucharda, Tjahjadi, & Švec, 2003), came up with the result that the inflamed thyroid tissue could be distinguished from the healthy thyroid gland tissue and, therefore, they found a success rate of the classification process of 100%.

3.4.4.1 Intensity Histogram Features

The characteristics of an image can be expressed using histogram features. By dividing the histogram values $h(i)$ on the whole total number of the image pixels M according to the Equation (8), the approximate possibility density of appearance of the intensity levels p obtained as

$$p(i) = \frac{h(i)}{M} \quad \text{Equation (8)}$$

Where, $i=0,1,\dots,L-1$, is a grey level in the image, L is the total number of intensity levels in the image and M the total number of pixels.

Based on the most important statistical tests such as histogram as the descriptive test as well as various features such as mean, standard deviation, skewness, kurtosis, energy and entropy are computed using equations 9, 10, 11, 12, 13 and 14 respectively.

Mean μ : $\mu = \sum_{i=0}^{L-1} ip(i)$ Equation (9)

Variance σ : $\sigma^2 = \sum_{i=0}^{L-1} (i - \mu)^2 p(i)$ Equation (10)

Skewness S : $S = \sigma^{-3} \sum_{i=0}^{L-1} (i - \mu)^3 p(i)$ Equation (11)

Kurtosis K : $K = \sigma^{-4} \sum_{i=0}^{L-1} (i - \mu)^4 p(i) - 3$ Equation (12)

Energy E : $E = \sum_{i=0}^{L-1} [p(i)]^2$ Equation (13)

Entropy T : $T = -\sum_{i=0}^{L-1} p(i) \log_2 [p(i)]$ Equation (14)

Where μ is the mean average level of the image matrix, and σ shows the variation of the intensity around the mean average. The kurtosis factor is the measure of the symmetry of the histogram while entropy is used to measure the histogram uniformity.

3.4.4.2 Grey Level Co-occurrence Matrix (GLCM)

GLCM is a two-dimensional function determining the relationship between a pixel and its neighbours over an entire image in various directions. Where the co-occurrence matrix is known by a

second-order histogram statistics. In a second-order, the relationship between the groups of two pixels in the competent image is applied. A GLCM is a matrix where the number of rows and columns are the same as the number of grey levels. Different texture characteristics are removed with accordance to the number of intensity points in every group.

According to (Kurani et al., 2004), fourteen texture features are determined from the probability matrix to extract the texture features. The fourteen texture features are calculated with $d=1$

Where, d , represents the distance in pixels with four directions for θ of $0^\circ, 45^\circ, 90^\circ$ and 135° .

The GLCM matrix with distance $d=1$ and the direction $\theta=0$ is shown in Figure 3.1

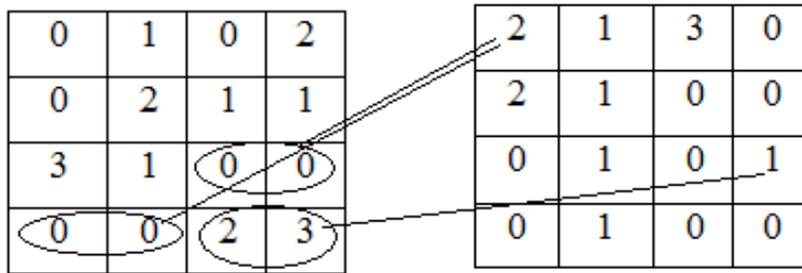


Figure 3.1: GLCM of 4 x4 image for distance $d = 1$ and direction $\theta=0$

The features that are extractable from the co-occurrence matrix are (Haralick, Shanmugam, & Dinstein, 1973):

- 1- Angular Second Moment (ASM)

$$ASM = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (p(i, j))^2 \quad \text{Equation (15)}$$

Where N_g represents the number of grey levels on the image, and $p(i, j)$ is the element at coordinates i and j in the co-occurrence matrix.

- 2- Contrast (Con)

$$CON = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (p(i, j))^2 \right\}, |i - j| = n \quad \text{Equation (16)}$$

- 3- Inverse Different Moment (IDM)

$$IDM = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{p(i, j)}{1+(i-j)^2} \quad \text{Equation (17)}$$

- 4- Entropy (ENT)

$$ENT = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i,j) \log(p(i,j)) \quad \text{Equation (18)}$$

5- Correlation (COR)

$$COR = \frac{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad \text{Equation (19)}$$

Where μ_x, μ_y, σ_x and σ_y are the mean values and standard deviations of p_x and p_y respectively.

6- Sum of Squares

$$SSQ = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (1 - \mu)^2 p(i,j) \quad \text{Equation (20)}$$

7- Sum Average (SAVE)

$$SAVE = \sum_{i=2}^{2N_g} i p_{x+y}(i) \quad \text{Equation (21)}$$

Where p_{x+y} is

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j), i + j = k, k = 2, 3, \dots, 2N_g \quad \text{Equation (22)}$$

8- Sum Entropy (SENT)

$$SENT = \sum_{i=2}^{2N_g} p_{x+y}(i) \log(p_{x+y}(i)) \quad \text{Equation (23)}$$

9- Sum Variance (SVAR)

$$SVAR = - \sum_{i=2}^{2N_g} (i - SENT)^2 p_{x+y}(i) \quad \text{Equation (24)}$$

10- Difference Variance (DVAR)

$$DVAR = \sum_{i=2}^{2N_g} (i - SAVE)^2 p_{x-y}(i) \quad \text{Equation (25)}$$

11- Difference Entropy (DENT)

$$DENT = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log(p_{x-y}(i)) \quad \text{Equation (26)}$$

$$12- p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j), |i - j| = k, k = 2, 3, \dots, N_g - 1 \quad \text{Equation (27)}$$

3.4.4.3 Grey Level Run-Length Matrix (GLRLM)

GLRLM is a statistical approach used to describe the texture information of a grey level image region. In addition, this approach is usually used for texture feature extraction. Also, it is explained and represented by run length matrix $P(i, j/\theta)$,

Where (i) is the number of the grey level that runs in length (j) in different directions, often in $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$. Moreover, the example of GLRLM method is shown in Figure 3.2. Different

texture characteristics can be obtained from the run-length matrix. It presents the total number of sequential runs of length (j) at grey level (i) in the direction θ .

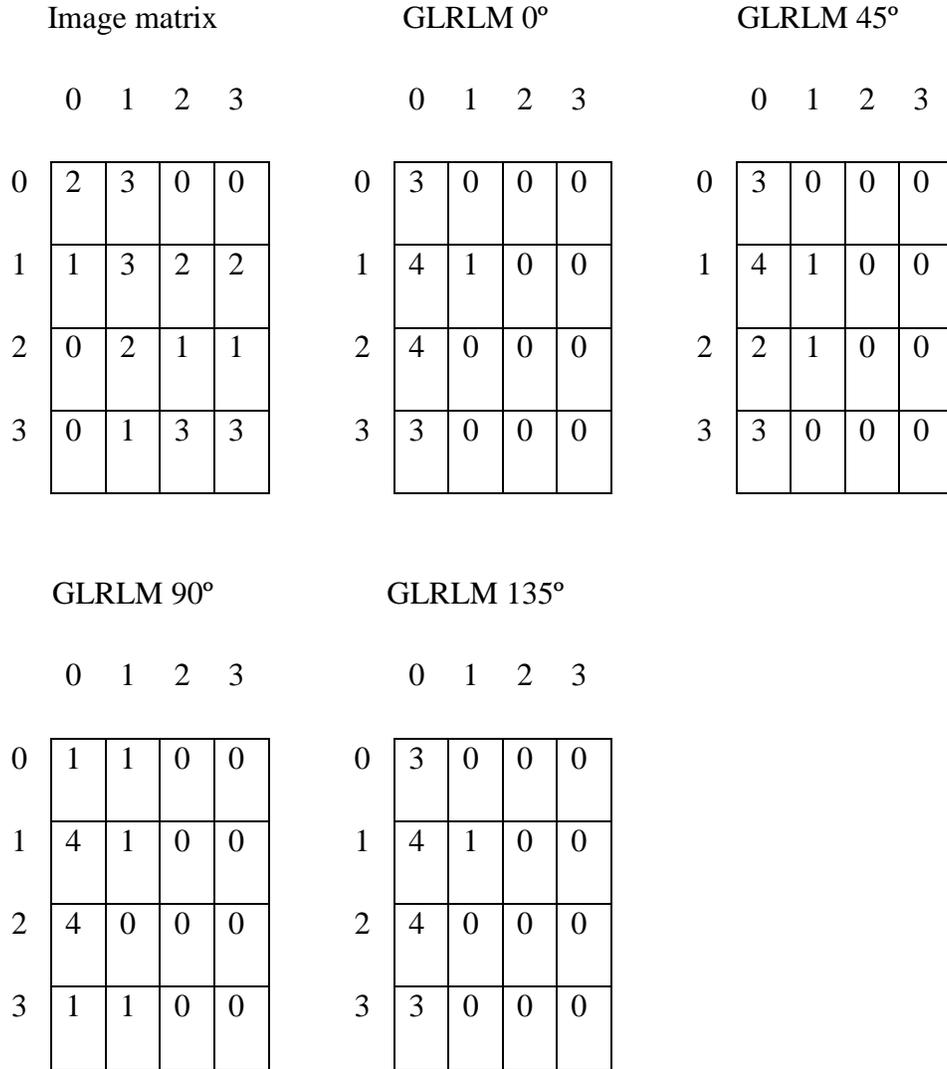


Figure 3.2: An example of GLRLM method

From the run-length matrix $p(i, j)$ a number of statistical texture features can be derived as follows (Dasarathy & Holder, 1991):

1) Short Run Emphasis (SRE):

$$SRE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{p(i, j)}{j^2} = \frac{1}{n_r} \sum_{j=1}^N \frac{p_r(j)}{j^2} \quad \text{Equation (28)}$$

2) Long Run Emphasis (LRE):

$$LRE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N p(i, j) \cdot j^2 = \frac{1}{n_r} \sum_{j=1}^N p_r(j) \cdot j^2 \quad \text{Equation (29)}$$

3) Grey-Level Nonuniformity (GLN):

$$GLN = \frac{1}{n_r} \sum_{i=1}^M \left(\sum_{j=1}^N p(i, j) \right)^2 = \frac{1}{n_r} \sum_{i=1}^M p_g(i)^2 \quad \text{Equation (30)}$$

4) Run Length Nonuniformity (RLN):

$$RLN = \frac{1}{n_r} \sum_{j=1}^N \left(\sum_{i=1}^M p(i, j) \right)^2 = \frac{1}{n_r} \sum_{j=1}^N p_r(j)^2 \quad \text{Equation (31)}$$

5) Run Percentage (RP):

$$RP = \frac{n_r}{n_p} \quad \text{Equation (32)}$$

6) Low Grey-Level Run Emphasis (LGRE):

$$LGRE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{p(i, j)}{i^2} = \frac{1}{n_r} \sum_{i=1}^M \frac{p_g(i)}{i^2} \quad \text{Equation (33)}$$

7) High Grey-Level Run Emphasis (HGRE):

$$HGRE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N p(i, j) \cdot i^2 = \frac{1}{n_r} \sum_{i=1}^M p_g(i) \cdot i^2 \quad \text{Equation (34)}$$

For large coarse and structural textures, the run lengths are predictable to be large, but to a specific range, they are small for fine textures. Whether there is a high-intensity clustering in the texture, then the long run length is also large, while, the non-uniformity features are small if the run lengths or the grey levels are similar all over the matrix.

3.4.5 Feature Reduction

The reduction of the dimensionality could be described as the transformation of the high-dimensional data into a significant representation of reduced dimensionality. Typically, the representation, which already has been reduced, should hold the dimensionality that corresponds to the intrinsic dimensionality of the data. According to (Fukunaga, 1993), the minimum number of the required parameters that needed for an accounting of the observed properties of data is the intrinsic data dimensionality. In many physical-sense applications, the dimension of the feature may easily be high as

tens of thousands. However, such an extreme dimensionality may be so hindering and detrimental to the analysis of data and the processing as well. Therefore, in the feature spaces, the high dimensionality usually means high computational complexity and power consumption both in the (offline) learning and in the (online) prediction phases. Therefore, the dimension of the extreme feature may cause the numerical process to converge prematurely to a sub-optimal solution.

According to (Kung, 2014), when the dimension of the vector, in the supervised learning, far exceeds the number of training samples, this means that the data overfitting becomes highly likely that, in turn, jeopard the generalisation capability. Therefore, dimensionality reduction can solve this problem in an effective way. It reduces the features in a way that only the most dominant features are kept. In terms of the principality, the more affordable information about every single pattern, the better learning algorithm is expected to be performed. This leads to suggest that the features should be used as many as possible for the task at hand. Practically, this is not the case that should be taken into consideration given that many algorithms have poor performance in the high dimensional space because of the small amount of the learning samples as reported by (Law, 2006).

Some of the features in the set of data are only 'noise' and, thus, do not contribute, sometimes they are even degrading, the process of learning. This complexity, in turn during the analysing of data sets, with other many features and the small number of the samples is causing the curse of dimensionality as mentioned by (Raudys & Jain, 1991). Dimensionality reduction can remedy this problem by decreasing the feature dimension before starting the training phase. This will lead to decrease the process time and save more space to store the information in the classification process. However, it is easier for domain experts to interpret models that have less number of variables.

As a visualisation tool, the reduction of the dimensionality is, therefore, invaluable where data of the high-dimensional level is transformed into two or three dimensions for the purpose of the display. Remarkably, the main disadvantage of the reduction of the dimensionality is the information loss possibility. The reduction of the dimensionality can discard useful instead of irrelevant information when it is done poorly. According to (Law, 2006), there is no technical way to recover the lost information for whatever the next process will be.

In the context of the reduction of data, many linear and nonlinear techniques are currently available. For many years, the non-complex linear data techniques such as the Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) were used as the methods of the reduction of the dimensionality. The linear techniques such as PCA are usually executed and performed in the original space as an attempt to linearly discover new axes with high variance. According to (Maaten et al., 2009), however, the linear techniques cannot be used to handle successful nonlinear representation.

Remarkably, the linear techniques, due to the availability of various data patterns and representation, may not be able to accurately separate the complex nonlinear data. During the last decade, however, many nonlinear techniques for the reduction of the dimensionality were proposed as pointed out by ((Burges, 2005), (Schölkopf, Smola, & Müller, 1998), (Hinton & Salakhutdinov, 2006)) and (Maaten et al., 2009). Technically speaking, the nonlinear techniques are actual extensions of their corresponding linear versions as they act linearly in the nonlinear spaces.

The nonlinear techniques, which are paradoxical to the traditional linear techniques, can act with the complex nonlinear data. Specifically, for data in the real-world, the nonlinear dimensionality reduction techniques can provide benefit given that the real-world data is likely to be high-nonlinear. The previous work (Maaten et al., 2009) showed that the nonlinear techniques are outperforming their counterparts on the complex artificial tasks.

Kernel Principal Component Analysis (KPCA):

KPCA is the nonlinear reformulation technique for the traditional PCA. It embeds the data into a high dimensional space called the feature space. To solve the high-dimensional mapping problem, the kernel trick was used. This trick allows us to apply the PCA in the feature space in terms of dot product and the result is stored in the kernel matrix. Considering all the features in the high-dimensional space are zero-mean, then, the principal eigenvectors of the centred kernel matrix are calculated. The data is projected onto the eigenvectors of the covariance matrix. Choosing the kernel function affects the performance of the KPCA. KPCA has been used in various successful applications such as face recognition, speech recognition and novelty detection (Maaten et al., 2009). In medical image processing, KPCA has been used to analyse the blood cells' characteristics and extract their general characteristics (Kim et al.,

2000). In (Tongsen & Ting, 2007) KPCA was used to reveal the distribution of the liver cancer cells in the B-SCAN ultrasound image. It reaches a high level of accuracy and detects liver cancer and resists the confusion of noise. By applying KPCA, the extracted features are mapped in an implicit high-dimensional (possibly infinite) feature space using Gaussian Kernel and then dimensionality reduction is done using linear PCA (Kung, 2014). It is an eigenvector-based technique and has proven to be a powerful pre-processing step before the classification stage as it preserves features with a high variance that best describe data and eliminates the others. KPCA first generates a kernel matrix out from data in its original space with the help of some predefined nonlinear function called Kernel Function. This matrix represents the inner product of data in its high-dimensional feature space without even generating the latter. This is called Kernel Trick. Second, PCA is performed in the kernel space based on the kernel matrix. Third, dominant eigenvectors corresponding to the largest eigenvalues are chosen and extracted from the output matrix of KPCA to form the transformation matrix. Finally, original extracted features are projected in a new lower-dimensional space using the transformation matrix, which will be our new features.

The kernel matrix is defined as:

$$k_{ij} = k(x_i, x_j) \quad \text{Equation (35)}$$

Where k is the kernel function, and x_i is the data points. In order to normalize the data points and ensure that the data in the feature space are zero-mean the kernel matrix is centred by:

$$k_{ij} = k_{ij} - \frac{1}{n} \sum_{k=1}^n j_{ik} - \frac{1}{n} \sum_{k=1}^n k_{jk} + \frac{1}{n^2} \sum_{1,k=1}^n k_{1k} \quad \text{Equation (36)}$$

Then the principal eigenvectors are computed and ranked to know how much the variation of the captured data is. To compute the covariance matrix C of the features with dimensions $M \times M$

$$C = \frac{1}{n} \sum_{i=1}^n k_i (k_i)^T \quad \text{Equation (37)}$$

Eigen values and eigenvectors are calculated by

$$C v_k = \lambda_k v_k \quad \text{Equation (38)}$$

Where $k = 1, 2, \dots, M$. In order to perform the mapping operation, Gaussian kernel function was chosen as a Kernel PCA mapping function.

3.4.6 Summary

Textural features are defined as the mathematical parameters calculated from the pixels distribution, that describes the texture type, and therefore, the basic structure of the objects presented in images. In addition, textural analysis describes a defined area in an ultrasound image through the elements of the texture. Generally, the textural analysis method may be simplified into three main stages which are the acquisition of the ultrasound image, the analysis of it and its classification. Moreover, texture analysis is categorised mainly based on the methods that are applied to assess the interrelationships of the pixels in the image. The greatest inclusive classification of textural analysis simplifies it into four principal sets: statistical based group, structural based group, model based group, and transform based group. Where these methods can be furthermore subdivided into several methods.

It is very important to establish the reliability of the features that are extracted with the purpose of guaranteeing the correctness of the textural analysis methods when applied in image classification. The extracted features dependability can be impacted from one of the following; the feature extraction processes, the ultrasound system and the operator. It is thusly vital to building the reliability of B-mode images.

The reduction in dimensionality handles the change of a high dimensional dataset into a low dimensional place while retaining the majority of the beneficial structure from the initial data. Dimensionality reduction began to be of high significance as result of the appearance of a large number of datasets with many features. The basic hypothesis for the decrease in dimensionality is that the information points are not randomly placed in the high-dimensional area; there is, in fact, a particular structure in the places of the data points that may be used. In addition, the useful information in the high dimensional information may be summarised through a few properties.

The basic concept of the KPCA is to change the inserted patterns to a further higher nonlinear dimensional space and thereafter, carry the process out in the new KPCA space. The KPCA is the

addition of the typical linear PCA. PCA performs a linear segmentation of information in the initial space; meanwhile, KPCA implants the information on to a high dimensional space which is known as the feature space. The mapping linking the two spaces is usually implemented via the kernel function. In the Gaussian function, the information is mapped to unlimited dimensions, rather than computing the dot outcome. Through calculating the standard eigenvectors of the kernel matrix, the feature space information is projected via the eigenvectors that demonstrate the majority of the differences.

3.5 Classification

3.5.1 Introduction

Medical images classification is an important stage. Its objective being not only to reach a good level of accuracy but also to decide which part of the image is affected more than the others to aid the medical practitioners in their early diagnosis of the diseases. In texture classification, the aim is to assign an unknown sample image to one of a set of known texture classes. All the methods and processes that are taken in pre-processing, partitioning the image, extracting the texture features and selecting the most dominant features affect the classification process and results. The classification module is the final stage in automated diagnosis. The task of classification is defined as the process of selecting a subset of image features from a very large set to distinguish between different object classes, and this is done with minimum effort with few errors (Jähne, 2002). The first phase of classification is training the classifier using a data set to determine the best set of features. The second phase is testing where the unknown data are given to the classifier for actual classification (Minavathi et al., 2012). There are many possible techniques for the texture classification process. However, there are several classification methods to classify and detect abnormalities in medical images. We only report the most common methods namely the k-nearest neighbour (k-NN), Artificial Neural Network techniques (ANN), Decision Tree (DT), Linear Support Vector Machine (LSVM), nonlinear SVM and Ensemble Classifiers.

3.5.2 The k-Nearest Neighbour Techniques

The k-Nearest Neighbour (k-NN) is a supervised non-parametric algorithm. In k-NN, classifying objects is based on closest k training examples, which are the nearest neighbours to the test example. We

used a common distance function for k-NN, which is Euclidean distance: neighbours (Dougherty, 2009). (Suguna & Thanushkodi, 2010) Introduced a novel method to improve the classification performance of k-NN and reduced the rate calculation complexity by using Genetic Algorithm. (Lashari & Ibrahim, 2013) introduce a new framework for medical image classification that based on soft set to obtain greater performance, improve accuracy, precision and computational speed.

(Baily & Jain, 1978) Used weights with classical k-NN and gives an algorithm named weighted k-NN. The Condensed Nearest Neighbour algorithm (CNN) stores the patterns one by one and removes duplicate ones (Bhatia & Vandana, 2010). CNN was improved by (Gates, 1972) through eliminating the pattern that does not affect the training data set result. (Yong, 2009) Proposed a technique that uses clustering to calculate the nearest neighbour. It does not use all training samples as in a traditional k-NN algorithm. It clusters each training set by k values and assigns weight to each cluster.

In k-NN classifier, the new feature is assigned a label depending on the most frequently occurring label in the k nearest. If we choose $k=1$, we will classify objects in a way that they will be extremely sensitive to the outlier samples. However, if we select a large value of k, then the average will be calculated including the noisy data points. In this research, choosing k is a result of testing different k values in the classification process and selecting optimal k is based on the values that show less misclassification error rate.

3.5.3 Artificial Neural Networks Techniques

Artificial Neural Networks (ANN) can provide a valuable tool that could minimise the inconsistencies in medical image interpretation. ANNs have various applications such as pattern recognition, forecasting, clinical diagnosis and more (Deepa & Aruna Devi, 2011). Artificial classifiers have very high classification accuracy with a comparison to conventional classifiers. The advantages of ANN include self-learning, associative memory, parallelism strength, and error tolerance (Gharehchopogh, Molany, & Mokri, 2013).

The back-propagation network is the most prevalent and generalised neural network presently in usage (Zhai, Wangi, & Zhang, 2007). It learns by example, gives the algorithm examples of what we want the network to do and adjusts the weights to produce the desired output. The combination of

weights that reduce the error function is thought to be an answer to the learning problem. Once the network has been trained, it should be able to recognise the noiseless and noisy patterns in the training set. The network's performance can be improved by applying the patterns in random order to the network.

ANN has an essential role in the medical imaging field. It can monitor many health indices or predict the patient responses to therapy (Khan, Zope, & Suralkar, 2013). (Y. Jiang et al., 1999) showed that ANN is most widely used in breast cancer diagnosis with some improvement corresponding to increased sensitivity. (Murata, Yoshizawa, & Amari, 1994) Investigate the problem of identifying the best number of parameters in neural networks from a statistical point of view. However, for achieving a great and reliable performance, a large number of training cases are commonly required (Sahiner, Chan, & Hadjiiski, 2008).

3.5.4 Decision Tree

A decision tree is one of the most widely known classification algorithms used in data classification. It is divided into one root node, internal branches and leaves, which are the decision nodes (Quinlan, 1993). Decision nodes are labelled with certain classes, which represent the classification decision. A decision tree is made from a training set that composed of samples and every individual one of them is described by a set of characteristics and class labels. The classification process starts at the root of the tree and the test at the node is identified corresponding to the outcome. The training set is used to induct the DT while the testing set is used to measure their accuracy (Podgorelec, Kokol, Stiglic, & Rozman, 2002). To construct DT, there are various numbers of algorithms required as well as different parameters such as splitting and stopping rules. Splitting the node is based on the splitting rule, and the stopping rule controls whether the training set is able to divide into more layers or not. In this research, the splitting criteria are *gini*, which is defined as:

$$gini(t) = \sum p_i(1 - p_i) \quad (3.29)$$

Where class *i* at node *t* is the relative frequency of p_i , any node (parent or child) represented by node *t* where separated data is provided is performed. A *gini* method attempts to discover the huge pure node (homogeneous category) from the available classes and to segregate it from the rest of the data.

Decision trees have been used in medical applications for many years. (Cremilleux & Robert, 1997) Proposed a theoretical framework for using decision trees in the medical field. Therefore, (Tsien, Fraser, Long, & Kennedy, 1998) developed a model using classification tree techniques and Logistic Regression (LR) methods for early diagnosis of Myocardial Infarction (MI). Information, such as electrocardiographic about the patient's condition while they are in ER is used. A training data set was gathered in Edinburgh, Scotland.

The model tested on separate Edinburgh data sets, in addition to data sets collected from various hospitals in the UK. This study finds that classification trees have particular benefits over LR models in the diagnosis of patients with MI. (Bonner, 2001) Explains how the decision tree approach collaborates with clinical decision-making in mental health care in the UK. The approach offers a method to support multi-professional decision-making to inform the final decision. (Letourneau & Jensen, 1998) used decision trees to find accuracy and confident decisions about wound care treatment. Data were collected from two groups: one group was measured after the first contact with a decision tree, and the other group was measured two years following the implementation. They summarised that a decision tree is able to help with decision making through guiding the nurse through the assessment and treatment choices.

3.5.5 Support Vector Machine

A Support Vector Machine (SVM) is a statistical supervised learning method. SVM uses an optimal linear separation hyper plane to divide two sets of data in their input space. It achieves this goal by maximizing minimum margins between the two sets. It involves the input as a set of training data. If in the case the classes cannot be linearly segregated in the input space, Kernel SVM will transform them into the higher dimensional space known as the feature space, where linear separation becomes easier than the input space. In feature space, the closest points, which determine the margin of the classifier, are called support vectors. In this study, comparative analysis classifiers with linear and nonlinear types of kernels are implemented using the same data set.

For many years, SVM has shown outperformance in many applications including digital handwriting recognition (Schoelkopf et al., 1996), object recognition (Massimiliano & Alessandro, 1998), face detection (Osuna, Freund, & Girosi, 1997) and text categorization (Thorsten, 1999).

However, in medical image processing applications SVM have various applications. (Magnin et al., 2009) Brought a technique on the base of the SVM that enabled automatic classification of patients with early Alzheimer's disease. It is based on structural MRI and works automatically using defined ROI. (Mesrob et al., 2012) Suggested a new multimodal MRI is combining brain grey matter concentration and mean diffusivity at the voxel level. They prove to utilise nonlinear SVM with radial basis function kernel that the discrimination between Alzheimer's disease patients and elderly control reached the high precision level when relevant ROI is chosen. (Wei, Tiebin, Rodolfo, Marta, & Muin, 2010) Used the SVM model to select sets of variables that will be used to classify between diagnosed diabetes or pre-diabetes against no diabetes. They conclude from their study that the discriminative performance of SVM models was equal to the epidemiological method. (Suykens, Gestel, Brabanter, Moor, & Vandewalle, 2002) Brought the concept of modifying the Vapnik's SVM formulation by the addition of a Least Square (LS) term to the cost function to solve the quadratic programming problem. The LS-SVM formulation transformed the problem into a set of linear equations. (Selvaraj, Selvi, Selvathi, & Gewali, 2007) Used the LS-SVM approach for automated classification of MRI slices as normal and abnormal. SVM classifier was trained in (El-Naqa, Yang, Wernick, Galatsanos, & Nishikawa, 2002) to ensure whether a Micro Calcification (MC) is found or not at every location in a mammogram. By iteratively selecting the 'most representative' MC-absent examples, they tried to improve the performance of the SVM classifier.

3.5.6 Deep Learning

Due to the fact that valuable information can be harvested from complex systems through deep learning, it plays a vital part in big data solutions (X.-W. Chen & Lin, 2014). Actually, deep learning dates back to the 1940s. However, traditional training strategies for multi-layer neural networks always result in a locally optimal solution or cannot guarantee the convergence. Although it was realized that better performance for feature and representation learning could be achieved by the multilayer neural networks, wide applications have not been given to the multi-layer neural networks. To train deep learning efficaciously, result in the first back-through of deep learning, a two stage strategy, pre-training and fine-tuning, was suggested by (Hinton & Salakhutdinov, 2006). Moreover, the prevalence of deep learning is also elevated by the increase of computing power and data size. Collection of numerous

samples can be done so that the parameters of deep learning models can be trained upon the arrival of the era of big data. Simultaneously, high-performance computing systems are needed for training a large-scale deep learning model.

The creation of AI systems is being attempted for a very long time. The relationships and high-level schematics of various fields can be seen in Figure 3.3. Efforts for explicitly programming the needed knowledge for given tasks have been made by early approaches; nevertheless, as there are many difficulties in designing all the detail needed for an AI system to achieve satisfactory results by hand, these experienced problems in dealing with complex real-work issues (Goodfellow, Bengio, Courville, & Bengio, 2016).

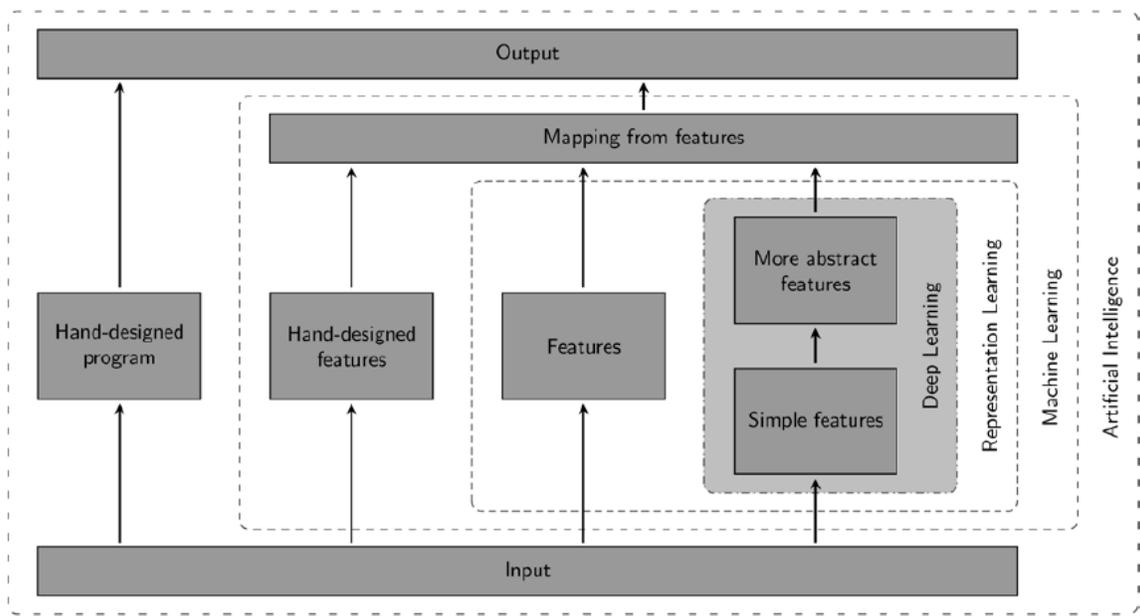


Figure 3.3 Relationships and high level schematics of different disciplines

Nonetheless, a specific type of deep, feedforward network, namely the convolutional neural network (ConvNet) (LeCun et al., 1990; LeCun, Bottou, Bengio, & Haffner, 1998) existed which was quite easy to train and generalized in a better way in comparison to networks having full connectivity between adjacent layers. During the time when neural networks were unpopular, a number of practical successes were achieved by it and the computer-vision community has widely adopted it as of late.

3.5.6.1 Convolutional Neural Network (CNN)

For large-scale image classification and recognition, the most prevalent deep learning model in feature learning is the convolutional neural network (Han et al., 2017; Karpathy et al., 2014; Krizhevsky, Sutskever, & Hinton, 2012; Maggiori, Tarabalka, Charpiat, & Alliez, 2017; Simonyan & Zisserman, 2014). As Figure 3.4 illustrates, convolutional layer, subsampling layer (pooling layer) and fully-connected layer are the three layers present in a convolutional neural network [44]. For achieving the weight sharing, the convolution operation is used by the convolutional layer and, for reducing the dimension, it uses the subsampling. A 2-dimensional image x is to be taken as an example. A sequential input $x = \{x_1, x_2, \dots, x_N\}$ is what the image is initially decomposed into. Following is the definition of the convolutional layer for sharing the weight:

$$y_i = f(\sum_i k_{ij} \otimes x_i + b_j),$$

Where j th output for the convolutional layer is represented by y_i and the convolutional kernel with the i th input map x_i is represented by k_{ij} . The bias is represented by b_j and the discrete convolution operator is represented by \otimes . Furthermore, the non-linear activation, usually a scaled hyperbolic tangent function, is represented by f .

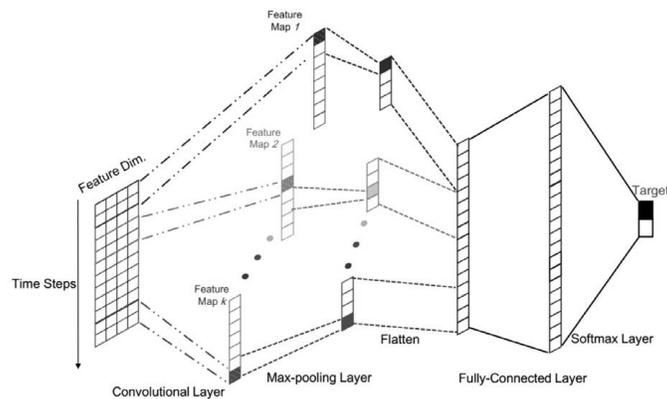


Figure 3.4 Convolutional neural network

Reducing the dimension of the feature map is the objective of the subsampling layer. An average pooling operation or a max pooling operation can usually be used for its implementation. Then, for

classification and recognition, different fully-connected layers and a softmax layer are normally put on the top layer.

For feature learning on large-scale images, a number of convolutional layers and subsampling layers are usually present in the deep convolutional neural network. Convolutional neural networks have contributed a lot in language processing and speech recognition etc, recently (Abdel-Hamid, Mohamed, Jiang, & Penn, 2012; Qian, Bi, Tan, & Yu, 2016; Swietojanski, Ghoshal, & Renals, 2014).

3.5.6.2 Recurrent Neural Network (RNN)

The conventional deep learning models like stacked auto-encoders, deep belief networks and convolutional neural networks are not appropriate for learning features for the time series data as they do not consider the time series. A natural language sentence similar to a typical time series data is to be taken as an example. There should be utilization of the previous one or more words as inputs during the use of the current word for predicting the next word as there is a close connection between each word and other words in a sentence. Clearly, due to the information of previous inputs not being stored by the feed-forward deep learning models, they are not that appropriate for this task. Features for the series data by a memory of previous inputs, which are stored in the internal state of the neural network, are learned by the recurrent neural network, which is a conventional sequential learning model. As it can be seen in Figure 3.5, there is the introduction of a directed cycle for constructing the connections between neurons.

Input units $\{x_0, x_1, \dots, x_t, x_{t+1}, \dots\}$, output units $\{y_0, y_1, \dots, y_t, y_{t+1}, \dots\}$ and hidden units $\{s_0, s_1, \dots, s_t, s_{t+1}, \dots\}$ are present in a recurrent neural network. The current sample x_t and the previous hidden representation s_{t-1} are taken as input by the recurrent neural network as it can be seen in Fig. 8, at the time step t for acquiring the present hidden representation s_t :

$$s_t = f(x_t, s_{t-1}),$$

Where the encoder function is represented by f . The utilization of vanilla recurrent neural network is prevalent which at the time step t is defined as the following forward pass:

$$s_t = f(W_{sx}x_t + W_{ss}s_{t-1} + b_s),$$

$$y_t = g(W_{ys}s_t + b_y)$$

Where the encoder and decoder are represented by f and g respectively, and the parameter set is represented by:

$$\theta = \{W_{sx}, W_{ss}, W_s; W_{ys}, b_y\}.$$

Accordingly, through the integration of the previously hidden representation s_{t-1} into the forward pass, the dependency between the current samples x_t with the previous one x_{t-1} is acquired by the neural network. Arbitrary-length dependencies can be acquired by the recurrent neural network according to a theoretical viewpoint.

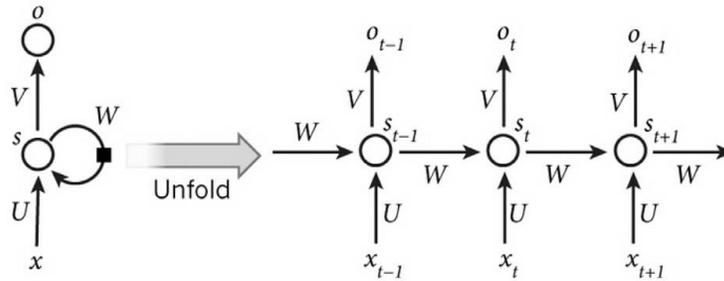


Figure 3.5 Recurrent neural network

Nevertheless, due to the gradient vanishing with the back-propagation strategy for training the parameters, there are difficulties for the recurrent neural network in capturing a long-term dependency. Through the prevention of the gradient vanishing or gradient exploding, some models, like long short-term memory, have been presented for overcoming this issue (Chung, Gulcehre, Cho, & Bengio, 2014; Hochreiter & Schmidhuber, 1997).

Stacking of many recurrent neural networks into a deep learning model is possible. In a number of applications like natural language processing, speech recognition and machine translation, super performance has been achieved by the recurrent neural network and its variants (X. Chen, Liu, Qian, Gales, & Woodland, 2016; Chien & Ku, 2016).

However, because data set we have in this research is small, and the experiments are conducted on a standard PC. Deep learning was not appropriate and was not used in this research. Deep learning will not be further referred to in this thesis.

3.5.7 Ensemble Classifiers

The strategy in ensemble systems is to create some classifiers and combine their outputs such that, the combination improves the performance of an individual classifier. This demand, therefore, that single classifiers make errors on a number of different examples. The intuition is that, whether every classifier causes different mistakes, thereafter a strategic group of these classifiers has the ability to minimise the total mistakes. Supervised classification issues fall under the category of learning from instances where each instance/pattern/example is connected to the label/class. An individual classifier like Neural Network, Decision Tree, or a Support Vector Machine is trained on a classified data set. Depending on the distribution of the patterns, it is possible that some of the patterns will be at the high level of learning by an individual classifier. Classifier implements weekly on the test set under such situations. To solve such problem a group of classifiers trained on the same problem, where the individual classifiers are known by base/weak classifiers. At the period of learning, the foundation classifiers are trained separately on the data set. A fusion method then collects the decisions produced by the principle classifiers. (Polikar, 2006) Indicated that there occur perfect numbers of fusion methods in the literature containing the majority of voting, Borda count, and algebraic combiners. The philosophy of the group classifier is that another base classifier compensates the mistakes that have been made by one base classifier.

However, training the base classifier in a straightforward way will not solve this problem. (Polikar, 2006) indicated that an ensemble classifier implements much better than its base counterpart if the base classifiers are exact and base classifier errors are unlinked. ((Tang, Suganthan, & Yao, 2006), (Brown, Wyatt, Harris, & Yao, 2005), (Kuncheva & Whitaker, 2003)) Showed that there are a large number of methods to calculate diversity containing pairwise diversity measures and non-pairwise diversity measure. A variety of ensemble classifier generation methods purposes to obtain diversity between the principal classifiers. Some of the ensemble classifiers are also improved aiming the specific problems/applications. Different base and ensemble classifiers and their applications will be explained in the following section. There is a need to answer two interrelated and basic questions.

Where is the first question: how will individual classifiers (base classifiers) be generated? And the second is: how will they differ from each other? Therefore, the strategy for creating the ensemble

members has to develop the ensemble's diversity. Moreover, according to the (Brown, 2004), indicated that ensemble algorithms do not attempt to maximise a specific diversity measure. Rather, increased diversity is usually sought through different resampling processes or chosen of the different training parameters.

3.5.7.1 Bagging

Bagging shorts for bootstrap aggregating. It is one of the most original ensembles algorithms. Furthermore, it is one of the simplest for application and comes with a good performance (Breiman, 1996). Bagging diversity is achieved through the usage of bootstrapped replicas of data for training: various data subsets are picked from the whole data, randomly while replacing. All individual subsets are utilised to train a specific classifier from the same kind. Each classifier is then mixed through proceeding with a majority voting of decisions made. The class picked by the majority of the classifiers for whichever instance, is the final ensemble decision. A variation of bagging is known as Random forests; this is due to it being structured from decision trees (Breiman, 2001). It can be produced from single decision trees that have random parameters for training. These parameters may be bootstrap replicas of the data, or they may be different characteristic subsets. Pasting small votes is different to bagging; it is intended for use with big data sets (Breiman, 1996). Big data sets are segmented into subsets known as *bites*, and each one of these bites is put to use in training classifiers. There are two variations for pasting small votes, one of these makes random data subsets, and it is known as Rvotes, and the other one makes successive datasets on the basis of the instance significance, and it is known as Ivotes.

3.5.7.2 Boosting

A weak learner is known as an algorithm which produces classifiers that are simply better functioning than random guesses. A strong learner produces classifiers that can accurately classify everything with the exception of a random small fraction of instances (Schapire, 1990). Schapire proved that a weak learner might be turned into a strong learner. Furthermore, Schapire also presents a simple algorithm to improve the performance of a weak learner to match that of a strong learner; this is known as boosting. Today, the algorithm is thought to be one of the greatest enhancements in the late history of machine learning. Like bagging, boosting produces a group of classifiers through resampling of the data that then

gets combined with majority votes. Nevertheless, the similarities end there. Resampling in boosting is purposely geared to give informing training data to all successive individual classifiers. Fundamentally, boosting makes three weak classifiers: C1, C2 and C3. C1 is put through training with an arbitrary subset of data available. C2 training data is selected as the greatest informative, given the first classifier. That is, the second classifier is trained on data that is partially misclassified and partially classified by the first classifier. C3 is trained by instances that disagree with the first and second classifiers. C1, C2 and C3 are merged with a three-way majority vote.

3.5.7.3 AdaBoost

AdaBoost was presented in 1997 by Freund and Schapire, and it attained great interest (Freund & Schapire, 1995). It is said to be a general form of boosting. In comparison with the numerous variants of AdaBoost, AdaBoost.M1 and AdaBoost.R are often used because they can deal with issues such as multiclass and regression. AdaBoost produces a group of hypotheses which are put together by major votes predicted through hypotheses. The hypotheses are produced via training a weak classifier with the usage of selected instances from distributed training data which is updated. Data is kept updated to guarantee misclassification of instances from past classifiers is likely to be part of the training data of the coming classifier. Thus, data of the successive classifier is equipped in favour of instances that are difficult to classify.

3.5.7.4 RUSBoost

RUSBoost is short for Random Under-Sampling Boosting (Seiffert, Khoshgoftaar, Van Hulse, & Napolitano, 2008). It is an algorithm for both sampling and boosting, and it is intended to enhance the performance of learners trained on imbalanced data. It makes use of an arbitrary under-sampling method that takes out data on a random basis from the majority class. To begin with, every training data chosen has the weight of $1/m$, in this m represents the number of training instances. Following that, every learner goes through the following training iteratively. Initially, arbitrary under-sampling is implemented to eliminate T% from the major class, till it turns to a minority in the new short-term training set and a new weight distribution is arranged. The new short-term training set and weight distribution are moved to the weak learner, and then the calculation of pseudo-loss is made. Following that, the weight

distribution is brought up to date through the pseudo-loss and made normal, and the weights that were updated are put to the application in the following iteration. Following N iterations, the outcome is brought from the weighted major vote.

3.5.7.5 Random Subspaces

According to (Ho, 1998), Random Subspaces constructs individual classifiers from randomly selected feature subspaces. This process will solve the problem of execrating of dimensionality and therefore, is appropriate for the high dimensional dataset. Moreover, (Ho, 1998) indicated that this process compensates for the potential shortage of accuracies and therefore, features to high ensemble diversity. In this method, feature subspaces are chosen randomly from the original feature space, and individual classifiers are produced mainly on the basis of the quality of the feature subspaces. The outcomes from every individual classifier are combined through uniform plurality voting to yield the final prediction.

3.5.8 The Classification Imbalance Problem

Within classification, a dataset is thought to be imbalanced if the number of instances that represent a class is less than instances from different classes. In addition, the class that has the least quantity of instances is typically the point of interest from the standpoint of the learning task (Chawla, Japkowicz, & Kotcz, 2004). This issue is a huge focal point due to it showing in a number of classification issues in the real world, particularly in the diagnostic clinical field ((Mazurowski et al., 2008), (Freitas, Costa-Pereira, & Brazdil, 2007)). In these instances, principle classifying algorithms have a prejudice regarding any class that has a higher quantity of instances. This is since rules that accurately forecast the instances are biased in the benefits of the accuracy metric, while specific rules that forecast examples within the minority class are typically ignored for and treated as noise since more general principles are favoured. In these means, instances of minority classes are more frequently classified incorrectly than ones from different classes. Nevertheless, the spread of skewed information is not an obstacle to the learning task when it is by itself ((Sun, Wong, & Kamel, 2009), (Benoît Magnin et al., 2009)); the problem is that typically a set of obstacles linked to this issue show up.

Small samples size: Normally, an imbalanced dataset means that there is a lack of minority class examples. Researchers have found that the rate of errors due to the imbalanced class's distribution

reduced if the quantity of minority class examples is representative, thus fixing the imbalance ratio. Therefore, patterns that are given by positive cases could be learned more efficiently although there is an imbalance in the class distribution.

Overlapping or class separability: In the case that this occurs, distinguishing rules are difficult to bring about. As a result, more general rules are induced that cause misclassification of a small number of instances (García, Mollineda, & Sánchez, 2008). When the classes do not overlap each other, every straightforward classifier can learn a suitable classifier despite the class distribution.

Small disjuncts: (Weiss & Provost, 2003) stated that small disjuncts exist in a dataset in the case that the minority class represented concept is created from sub concepts. Moreover, small disjuncts are implied in the majority of problems. The presence of sub concepts grows the complication of the issue due to the number of instances often being imbalanced among them.

3.5.9 Addressing the Imbalanced Problem

3.5.9.1 Pre-processing

As stated by ((Galar, Fernandez, Barrenechea, Bustince, & Herrera, 2012), (Khoshgoftaar, Van Hulse, & Napolitano, 2011)), a vast amount of methods have been introduced for the handling of the two class imbalanced problem, for the standard learning algorithms and ensemble methods.

These methods can be classified into the following groups:

1. Data level solutions: One of the aims of this solution is the adjustment of the class distribution balance through the data space sampling in order to reduce the impact of the imbalance, acting as an exterior method ((Batista, Prati, & Monard, 2004), (Y. Tang, Zhang, Chawla, & Krasser, 2009)).
2. Algorithmic level solutions: The objective is to adapt and strengthen particular algorithms to classifying in the direction of the positive class. Thus, they can be described as interior means of producing new algorithms or altering present ones in order to change and reflect on imbalance issues ((Barandela, Sánchez, Garcia, & Rangel, 2003), (García-Pedrajas, Pérez-Rodríguez, García-Pedrajas, Ortiz-Boyer, & Fyfe, 2012)).

3. Cost-sensitive solutions: These solutions include methods aimed at data level or/and algorithm level, taking into account great misclassifying prices for positive class examples regarding the negative class, thus, aiming to reduce pricey errors ((Domingos, 1999), (Ting, 2002)).

The benefit of solutions of data level is that they are versatile because they rely on the chosen classifier to be used in particular tasks. In addition, any data set can be pre-processed prior to usage to enable them to train for varying classifiers. In these means, data only needs to be prepared once. There are various techniques for rebalancing with which to pre-process the training data, and these can be categorized into three sets known as the under-sampling techniques, over-sampling techniques and hybrid techniques.

3.5.9.2 Using Ensemble Classifiers

In the medical field, prior to making decisions on any crucial cases, many doctors have their say and share their views. As a result, more reliable results are obtained because numerous professionals come together to make the decisions. This approach of merging the expert opinions to come to decisions may be implemented in classifying data, this is called the Classifier Ensemble (CE), and it is also known as Multiple Classifier System (MCS). CE is a mixture of various classifiers, combined to implement classification jointly. When each classifier is diverse, and thus they are not compatible with each other, their random errors cancel each other, therefore resulting in accurate decisions. Data that requires classification is taken by the classifiers as an input in order to obtain a number of predictions. The predictions which are the output are then mixed through the use of various techniques including voting and weighted voting. CE may be produced through the use of various methods that may be classified into the following classes (Kuncheva, 2001):

1. Usage of various training sets: Introduce diversity through segmenting training data sets into N subsets and training each classifier with varying subsets.
2. Usage of various characteristic subsets: Introduce diversity through the training of each classifier using various subsets of characteristics.
3. Usage of various classifier models: Introduce diversity through the mixture of various classifiers.
4. Usage of various combination schemes: Introduce diversity through the usage of various combination schemes.

Regular ensemble techniques function powerfully in the cases that the input datasets have balance in the distribution. At present, many applications have an imbalance in the distribution of data. This may lead to the classifier being biased to the majority class that has a larger quantity of instances than the other class. Due to this, a requirement of changes in the current ensemble is raised for the imbalanced data to be dealt effectively.

3.5.10 Comparison between classification techniques

A comparison between the applied classification techniques is summarized in Table 3.2

Table 3.2: Summarize the features and limitations of different classifications techniques

Algorithm	Features	Limitations
k-Nearest neighbour	<ul style="list-style-type: none"> Zero cost of the learning process Well suited for multimodal classes Easy to implement More efficient if training data is large Highly adapted behaviour to new unlabelled data New examples require only updating the distance database 	<ul style="list-style-type: none"> Time to find the nearest neighbours in a large training data set can be expensive It is sensitive to noisy or irrelative attributes The precision of algorithm decreases if the amount of data less. Calculating measures where variables have different measurement scales To classify new instances, the entire training data needs to be scanned
Support Vector Machine	<ul style="list-style-type: none"> It predicts accurate results for most of the classification and prediction problems High accuracy Contains a non-linear transformation Provides a good generalization capability Reduction in computational complexity The problem of overfitting is eliminated. 	<ul style="list-style-type: none"> For obtaining good results, it requires a very large number of records Parameter tuning Performance and accuracy depends upon the hyperplane selection and kernel parameter Training is time-consuming Determination of optimal parameters is not easy when there is non-linearly separable training data.
Artificial Neural	<ul style="list-style-type: none"> Work well even if the data is not linearly separable in the base feature space 	<ul style="list-style-type: none"> Speed and size requirement both in training and testing more

Network	<p>It is easy to use with few parameters to adjust</p> <p>Simple to implement</p> <p>Applicable to wide range of problems</p> <p>It is a non-parametric classifier</p> <p>It is data driven self-adaptive technique.</p>	<p>Difficult to know how many neurons and layers are necessary</p> <p>Learning can be slow</p> <p>High complexity and extensive memory requirements for classification in many cases</p> <p>Difficult in choosing the type network architecture</p> <p>Problem of overfitting</p>
Decision Tree	<p>Easy to generate and understand</p> <p>Reduce problem capacity</p> <p>Based on the hierarchical rule and can handle non-parametric training data</p> <p>Does not require an extensive design</p> <p>The computing time is less and statistical errors are avoided</p>	<p>Training time is so expensive with large data set</p> <p>Suffer from overfitting</p> <p>Become complex calculation when various values are undecided and/or when various outcomes are correlated</p> <p>The accuracy depends fully on the design of the decision tree and the selected feature</p>
Ensemble Classifiers	<p>Lower error</p> <p>Less sensitive to overfitting</p> <p>Reduce the bias</p> <p>Reduce the variance</p> <p>Improve the predictive performance</p> <p>More stable models</p> <p>More accurate with imbalanced data set</p> <p>Able to identify noise and outliers</p> <p>Using random sample more suitable for data streams</p>	<p>Difficult to measure correlation between classifiers from different types of learners</p> <p>Learning time and memory constraints</p> <p>Needs to be adapted to handle data stream</p> <p>Has memory limitations</p>

3.5.11 Summary

There are several numbers of classification methods applied to distinguish and indicate irregularities in the images in the medical images field. Non-parametric classifiers including SVM, k-NN, and DT are classification techniques applied to enhance correctness and precision in classification.

Recently, the field of ensemble classifiers has been developing in which the features and benefits of all individual classifiers have been mixed to output a better prediction.

Ensemble systems have increasingly caught attention due to their adaptable properties. Initially, with ensemble itself, numerous classifiers may have a preferable solution to that of an individual classifier. Several studies have researched on ensemble models and demonstrated that it could attain a mean average of prediction errors and minimise bias and change of errors. Furthermore, the majority of new ensemble models have somewhat the same learning methods, base learning algorithm and voting, but have varying approaches in every phase. In addition, every phase supplies an opportunity to enhance the model to classify minority classes.

In the imbalanced dataset, the class which has a higher instance number is known as the major class, and the class which has a comparatively lower instance number is known as the minor class. Medical diagnosis prediction applications of uncommon but serious disease are very substantial than ordinary treatment. In these cases, the majority of the classifiers are biased to the major classes and consequently display deficient rates of classification on minor classes. Also, the classifiers may predict all as a major class and disregard the minor class. Different methods are suggested to solve the issues that are linked to an imbalance in classes.

3.6 Classification Evaluation

3.6.1 Introduction

Assessment of performance is essential in numerous phases of developing classifiers. The procedure of designing a classification algorithm to extract a particular model within the data is usually repetitive. Every repetition will make a major modification on the classifier, thus it has to be re-assessed to identify how the performance was influenced. When the development procedure comes to an end, it is essential to demonstrate that the last classifier obtains a satisfactory performance level and that shows a major enhancement over other classifiers. To estimate the efficiency and validity of the results, various techniques have been used. There are two common evaluation techniques that are used, confusion matrix and ROC curves. The confusion matrix is of dimension $D \times D$ where D is the number of classes. In this

matrix, the number of True Positives (TP), represent the proportion that is abnormal and the classifier claims that they are abnormal. If some are normal because they do not have certain medical problems and classified by the test as abnormal, these are called False Negative (FN). There are some cases where they are abnormal and classified as normal – this is called False Positive (FP). Finally, when the images are normal, and the test says that they are, this is called True Negative (TN). However, ROC curves are two-dimensional graphs highlighting the relationship between true positive rates on the Y-axis and false positive rates on the X-axis. The Area Under the ROC Curve (AUC) is often utilised to measure the quality of the classification model. Any realistic classifier has AUC between 0.5 and 1 (Fawcett, 2006). A random classifier has an AUC less than 0.5 while AUC for the perfect classifier is equal to 1.

3.6.2 Confusion Matrix

The confusion matrix is utilised to describe the classification model performance. Through this matrix, we can determine the number of correct and incorrect classified data. The basis for whichever evaluation or visualisation of a 2-class classifier’s performance is the numbers in the confusion matrix. The bold inner box is a two by two confusion matrix at which the rows represent the actual class of an instance, and the columns represent the predicted class. Sensitivity and specificity are the most important characteristics for medical tests. According to (Deepa & Aruna Devi, 2011), there are two measures to evaluate the results: sensitivity and specificity. Sensitivity measures the percentage of images that are correctly identified as having the condition. Specificity, conversely, measures the percentage of images that are correctly identified as not having the condition. The higher value of both sensitivity and specificity means better performance for each classifier. Also, the accuracy metric is used to measure the performance of the classifier.

		predicted	
		positive	negative
truth	positive	<i>tp</i>	<i>fn</i>
	negative	<i>fp</i>	<i>tn</i>

$$\text{Sensitivity} = \text{TP}/(\text{TP}+\text{FN})$$

$$\text{Specificity} = \text{TN}/(\text{TN}+\text{FP})$$

$$\text{Accuracy} = (\text{TP}+\text{TN})/(\text{TP}+\text{FP}+\text{TN}+\text{FN}),$$

Where TP is the number of true positive, TN shows the number of true negatives, FP is the number of false positive and FN is the number of false negatives.

3.6.3 ROC curves

Receiver Operating Characteristics (ROC) provides the nuanced details about the behaviour and sensitivity of the classifier. ROC plots the rate of true positive to the rate of false positive where space measures the performance of the classifier under the curve (Figure 3.3). A good classifier has a larger space under the curve. However, for a bad classifier, its ROC curve covers a very small area.

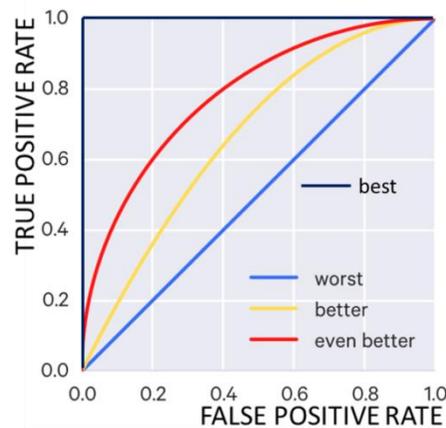


Figure 3.6: Two ROC curves whose performance is to be compared (Analytic, 2017)

(Altman & Bland, 1994) Indicated that ROC is especially beneficial in cases of comparing two or more measures. Moreover, (Morrison, Coughlin, Shine, Coull, & Rex, 2003) indicated that ROC curves had been widely utilised in the field of biomedical sciences to evaluate the ability and effectiveness of diagnostic examinations in distinguished healthy and unhealthy cases. This is support by Morrison in his study on 2005, Morrison identified that ROC curves evaluate the value of diagnostic

examinations by supplying a standard measure of the capability of an examination to accurately implement classification of subjects. In addition, (Morrison, 2005), found that an area under the curve (AUC) near to 1 demonstrates a great discriminatory power/capability of the detector variable. Meanwhile, AUC which is near to 0.5 demonstrates that the variable has weak discriminatory power.

Moreover, the ROC curve is frequently plotted through the usage of the true positive rate (TPR) against the False Positive Rate (FPR) for various cut-points of a diagnostic examination, beginning with the coordinates (0, 0) and ending with the coordinates (1, 1). FPR (1 – specificity) is represented via the *x*-axis, and TPR (sensitivity) is represented via the *y*-axis. Therefore, the ROC curve is the plotting of the test sensitivity against 1-specificity. The explanation of ROC curve is from an individual point in the ROC area, where, the nearer a point on the curve is to the desired coordinates, the higher the correctness of the test is. The test is less accurate if the points on the ROC curve was nearer the diagonal. Also, the test results are more beneficial when the curve is the faster approach the optimum point. The area under ROC curve (AUC) gives a method of measuring the correctness and precision of a diagnostic examination for example in the medical field. The greater region, the more accurate the diagnostic examination is. AUC of ROC curve can be computed by this equation;

$$t = (1 - \text{specificity}) \text{ and ROC } (t) \text{ is sensitivity} \quad \text{Equation (39)}$$

3.7 Summary

There are no studies so far that are related to the AT and the automatic identification of the medical condition and classification using US images. Hence, this research will use the findings from the current and reviewed literature to identify the best algorithms and processes that will produce the best classification system of the AT. We can summarize the findings from the literature as follows:

1. De-noising is very important, and SRAD is shown to be the best filter for US image pre-process. Hence, in this research, we adapted SRAD as the algorithm to be used in the denoising phase.
2. The region of interest improves the classification process and hence, because no previous research performed on the AT, this study will explore the different manner to define the best ROI to the AT US images.

Chapter 4 - RESEARCH METHODS, EXPERIMENTS SETTINGS AND ANALYSIS

This chapter will focus on the methods and data processing techniques that have been applied in this study including sample collection and preparation. This chapter mainly focuses on data acquisition and the image pre-processing as well as on determining the ROI and features extraction and finally will discuss the features reduction using KPCA, present the classification and evaluation methods.

4.1 Introduction

The aim of the present research is to develop an automated system that is able to differentiate between normal and abnormal human AT. Its purpose is to assist medical doctors and radiologists in their diagnosis tasks for the AT when using ultrasound images. AT ultrasound images are first acquired from different normal and abnormal patients using specified ultrasound instruments. The task of the pre-processing stage is to reduce the amount of speckle noise without affecting the image features. However, to focus on the most vulnerable AT area, various segmentation methods were proposed to determine the ROI. Subsequently, different texture features were extracted on the ROI using GLCM and GLRM algorithms. These sets of features along with histogram features and region properties undergo a process of dimensionality reduction using KPCA algorithm that selects the most dominant features. In the classification stage, various classifiers have been tested to determine the most appropriate one for our application based on the final results. For this reason, we are going to apply SVM, k-NN, DT and Ensemble classifiers, and through their results, we will be able to evaluate the performance of each one of them. Finally, the methods used in each stage will be evaluated using different evaluation metrics.

4.2 Instrument and Data Acquisition

A portable Venue 40 musculoskeletal ultrasound system (GE Healthcare, UK) with a 5–13 MHz wideband linear array probe with 12.7 mm _ 47.1 mm footprint area was used for scanning. The scans were performed independently by different operators according to a reliable scan protocol and within the

same session (Crofts, Angin, Mickle, Hill, & Nester, 2014). Operator order was dependent upon logistics. Good contact was maintained between probe and skin without applying excessive pressure. Three assessments were taken at each site of the Achilles with the probe removed between each recording. Each subject lay in the prone position for scanning.

The AT images used for this research were collected from various clinics by an expert in radiology from the School of Health at the University of Salford. Due to the lack of research on ultrasound AT images, it was not possible to find medical AT image datasets in the commonly used machine learning repositories or on the Internet. For this reason, a total of 57 AT ultrasound images were taken for the normal and abnormal cases in JPEG format. Each of these images is already diagnosed by medical experts and labelled as normal or abnormal for training and testing the classifiers in later stages. From the whole set of 57 AT ultrasound images, 43 of them are labelled as normal and the other 13 as abnormal. The AT US images with a size of 157 x 600 pixels were gathered. Each pixel in the grayscale AT image represents the pixel intensity in the range of [0, 255] (8-bit). The following sections in this chapter will discuss the modelling of the proposed system. All the images figures used in this chapter are taken from the images sample collected during this study.

4.3 Methods Used

The proposed technique is for improving AT ultrasound images diagnosis. It involves six subsequent stages, namely, Image Acquisition, Pre-processing, Determining the ROI, Feature Extraction & Reduction, Classification, and finally, Evaluation, as shown in Figure 4.1. The first stage, image pre-processing techniques and algorithms are applied for the purpose of image de-noising (using the SRAD filter), and the image is enhancing (using CLAHE algorithm). The second stage, determining the ROI (using manual and automatic algorithms). The third stage is feature extraction (using GLCM and GLRM algorithms), followed by feature reduction (using KPCA algorithm). The last stage is the classification process using various classifiers. Later, all these algorithms used in this research are evaluated using different evaluation metrics. In this research, the image processing functions, texture analysis techniques and used classifiers are applied using the MATLAB Image Processing and Statistics Toolboxes.

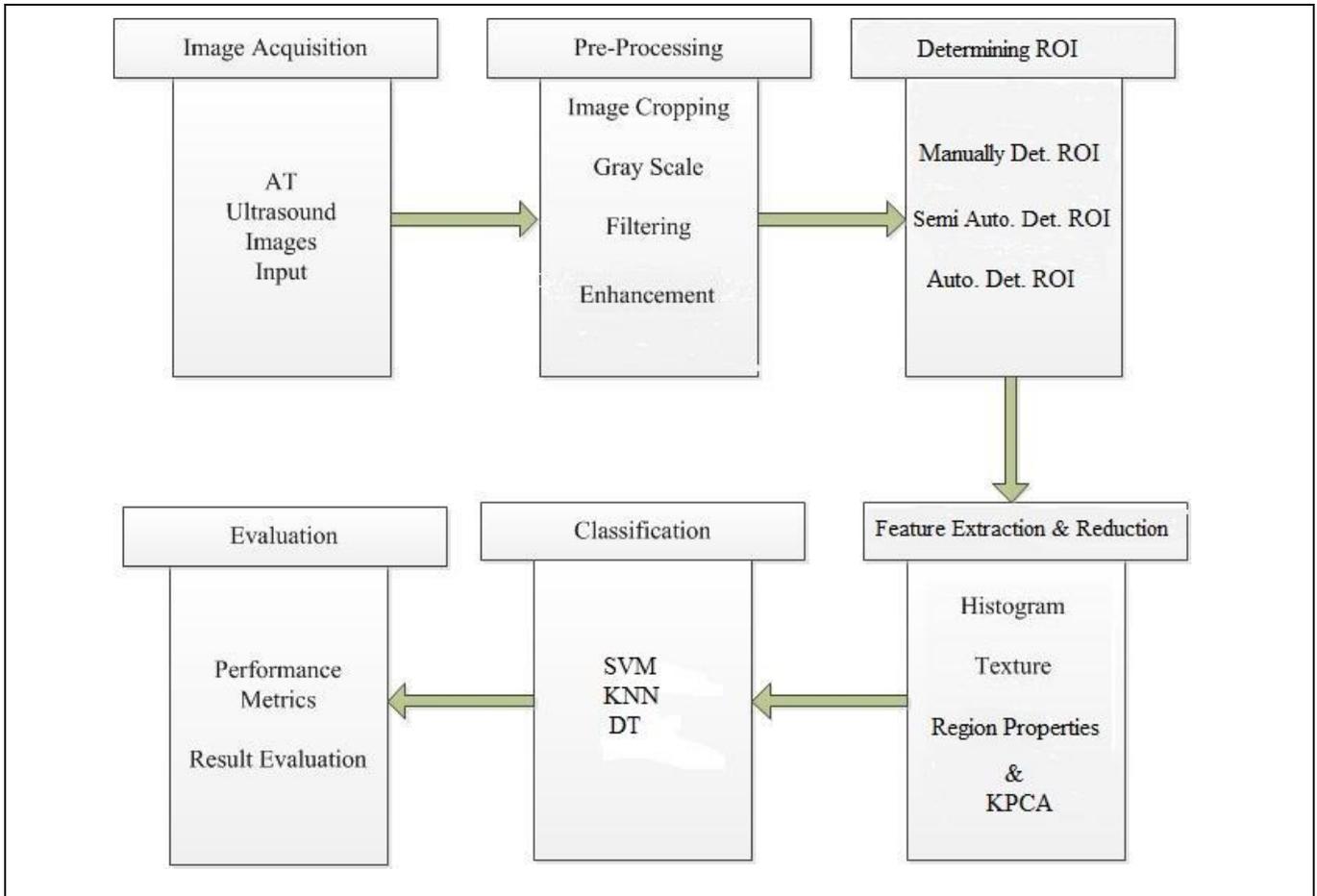


Figure 4.1: A block diagram of the proposed 6-stage AT classifier

4.3.1 Image Pre-processing

US AT Images typically contains artefacts in the form of random bright and dark spots and interference of the returned signal. These artefacts are usually causing a granular pattern in the image. The major problem with the precise segmentation of the ROI is due to the existence of such artefacts, which may cause trivial segmentation algorithms to fail. The AT pre-processing stage indicated in Figure 4.1 involves noise removal and image enhancement in order to suppress the speckle noise and preserve the edges of the AT images.

Another purpose of AT pre-processing is to improve the reliability and robustness of the AT segmentation, as discussed in the following sections.

4.3.1.1 Image Cropping

Cropping is the process of taking a full image and removing the parts that are not of interest or do not contain interesting information. This process is summarized in Figure 4.2.

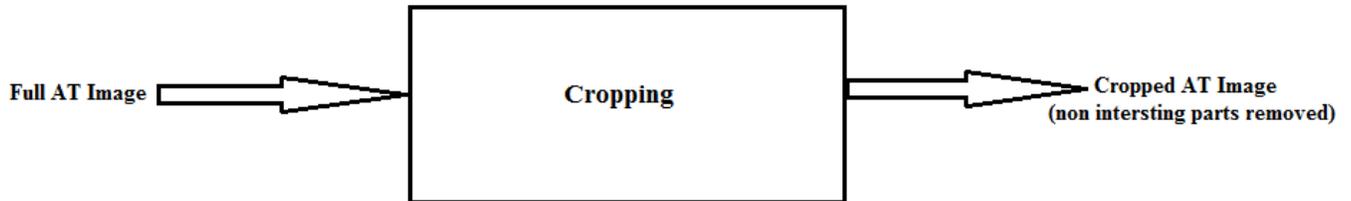


Figure 4.2 Image Cropping Process

To illustrate this example, the image given in Figure 4.4, is the result of cropping the full image given in Figure 4.3. The whole informative area is located in the top portion of the AT image where its height, according to image scale, is about 1cm according to the image scale. The remaining area below the top 1cm scale is out of AT tissues and is considered insignificant in the process of diagnosing AT diseases according to the experts.



Figure 4.3 An original AT image

Therefore, to speed up the diagnosis process and give more concentration on the ROI, all dataset images need to be cropped under the expert supervision to eliminate the area that is unrelated to the AT. A sample of a cropped AT image is shown in Figure 4.4.

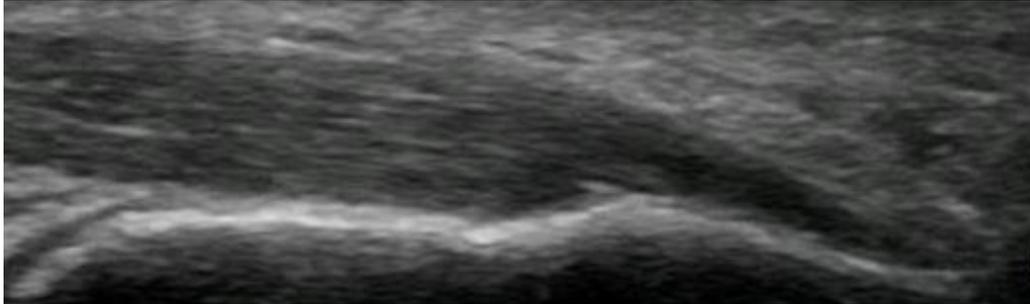


Figure 4.4 A cropped AT Image

4.3.1.2 Image Resizing

The size of the cropped image was 157 x 529 pixels and needs to be normalised. In order to preserve image information and to avoid losing details in the resizing process, we used the bilinear interpolation algorithm (Rowland, 1979). Which consider the closest four cell centres from the input and then it takes the average of theses 4 pixels to calculate the output value. The new image size is 157 x 600 pixels. The image will be divided into equal blocks at later stages, as mentioned in Section 4.2.2.3.

4.3.1.3 Image De-noising

As discussed in Section 3.2.2 SRAD can effectively solve the speckle noise without distorting useful image information and without destroying the important image edges. Therefore, to overcome the problem of having speckle noise in the AT image, the SRAD de-speckle method was used. It is considered better than the classical anisotropic reduction method when the image contains speckle noise. This process of using a de-noising algorithm (SRAD) is summarized in Figure 4.5.

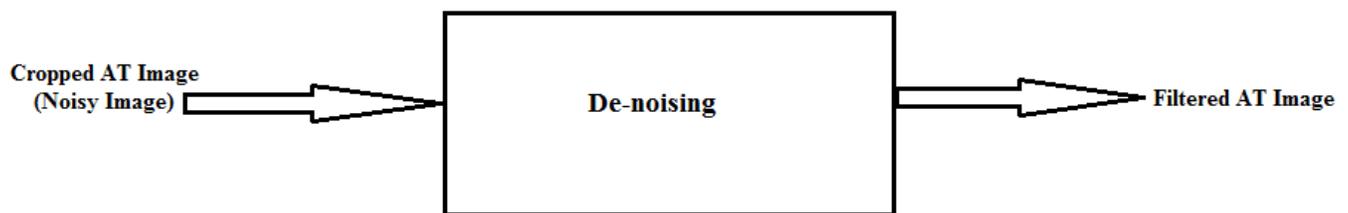


Figure 4.5 Image De-noising Process

SRAD filter can preserve the edges also, smoothen the rest of the image while reducing the noise (Yu & Acton, 2002). A sliding window with a number of iterations is defined as a parameter for the SRAD method in order to obtain good results. From the experiments, it is shown that a large number of iterations produce more smoothing, while a small number is not recommended. Figures 4.6 and 4.7 show an AT image before and after applying the SRAD image respectively.

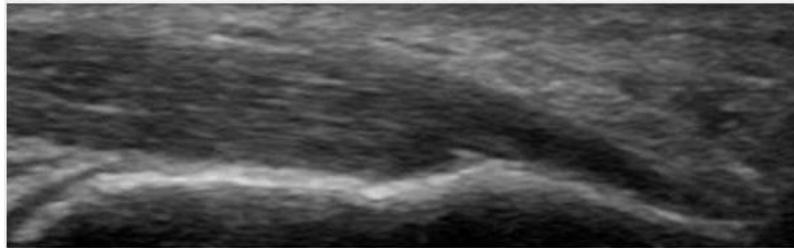


Figure 4.6 An AT image before applying any filter

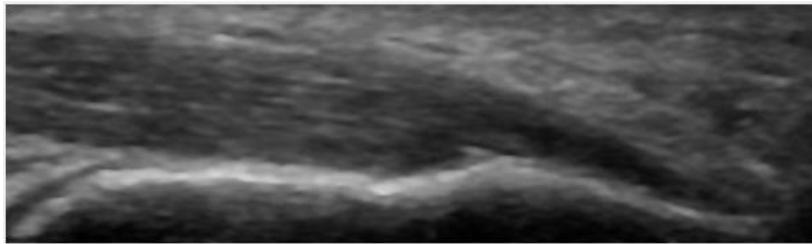


Figure 4.7 The AT image after applying the SRAD filter

4.3.1.4 Image Enhancement

To improve the image contrast without over-enhancing noise content or introducing artefacts to the processed image we will use Contrast-Limited Adaptive Histogram Equalization (CLAHE). It works on small regions of the image called tiles rather than the entire image. By this way, the contrast in the homogeneous region can be limited to its minimum, and it also limits the appearance of artefacts and noises. Figure 4.8 shows the process of enhancing the filtered AT image, and a sample of enhanced filtered AT image after applying the CLAHE histogram enhancement shown in Figure 4.9. It can be seen from the figure that the image details and edges are preserved.

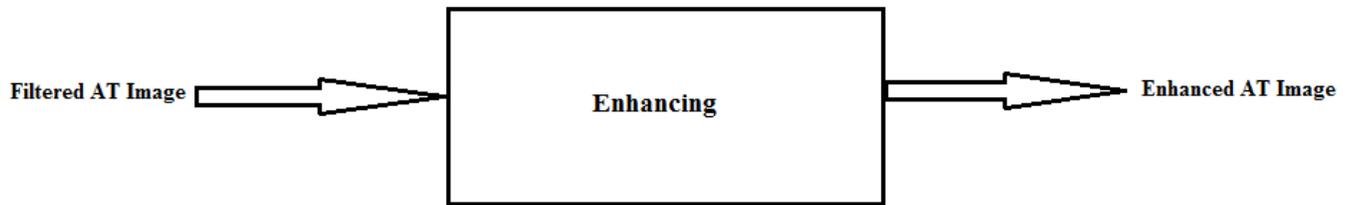


Figure 4.8 Image enhancing process

Histogram equalisation enhances image contrast through attaining a uniform histogram. It assigns pixels values intensities in the image input so that an outcome will have a uniform intensity distribution. This method may be utilised on an image as a whole or a part of an image. With Adaptive Histogram Equalization, ultrasound image contrast is improved through the change of intensity values. The method of histogram equalisation is also known to improve an image by enhancing the histogram equalisation of an image as a whole (Yu & Bajaj, 2004). Moreover, Adaptive Histogram Equalization tries to overcome the restrictions of global linear minimum-maximum windowing through offering the data which is considered the most desired in an image that can be made with no use of manual intervention (Polesel, Ramponi, & Mathews, 2000). In contrast Histogram Equalization functions by working on parts of an image separately. This makes the technique effectual, and therefore it is widely used in enhancing the contrast of grey scale and colour images.

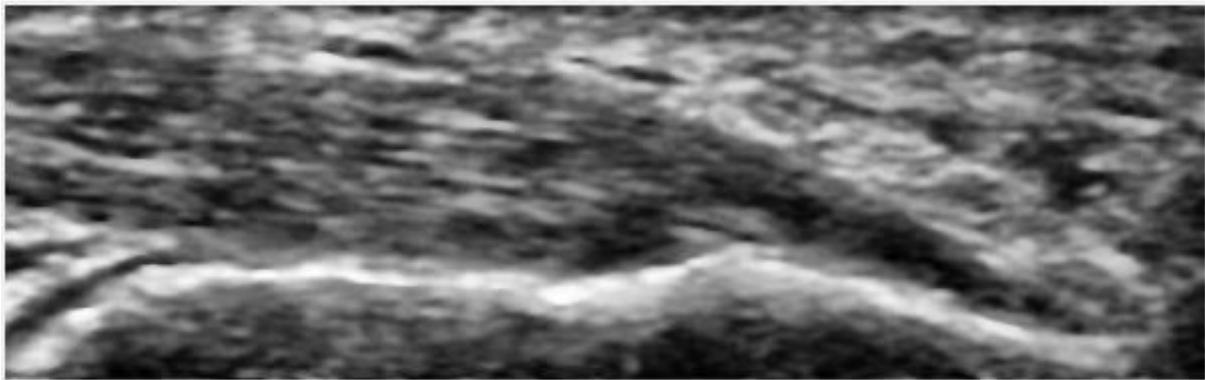


Figure 4.9 The AT image after contrast was enhanced using CLAHE algorithm

4.3.2 Determining the ROI

Image Segmentation plays a vital role in many biomedical image applications because it represents the foundation for the following stages. If any shortage occurs at this stage, it will affect the

following stages. We intended to locate the suspicious area in AT ultrasound image by subdividing the image into component regions. There are many algorithms used for image segmentation whereas no single one can be considered as superior due to the different conditions and constraints. There are various kinds of segmentation methods and approaches available such as manual, semi-automatic, or fully automatic segmentation. In this work, different algorithms will be applied on AT images, and the results will be reported to understand which one is more suitable for segmenting the ROI effectively. In the present section, we present the different approaches adopted in this research in order to achieve the research goals with minimum intervention from the radiologists. With each approach, the way and size of determining the ROI are different. The results of each approach will be evaluated and discussed through the classifiers results (more detailed results are reported in Chapter 5).

4.3.2.1 Approach one- The whole AT area

Initially, instead of going through different methods to determine the ROI, the raw data is used. After the image cropping is performed, as mentioned in Section 4.3.1.1, the entire AT area and the surrounding tissues are captured where the ROI area boundaries are attached upward to the calf muscles from the insertion in heel bone as shown in Figure 4.10. This delineation process of ROI is done manually under the supervision of the clinical expert in the cropping image process.

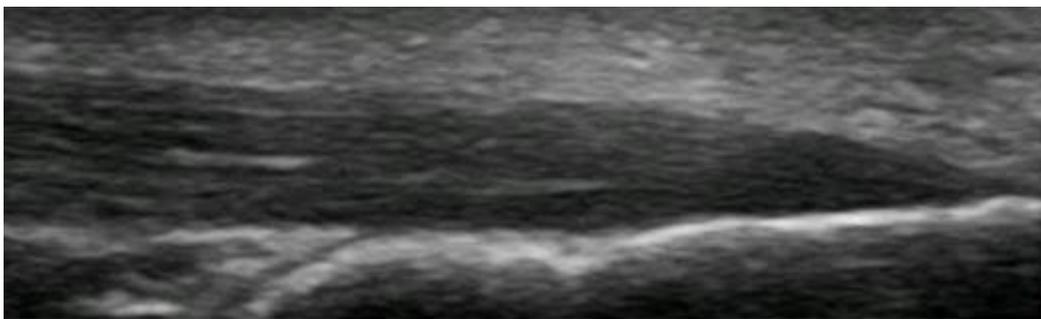


Figure 4.10 Ultrasound image shows the Lower AT area

To consider the whole AT image as the ROI can reduce the clinical expert interventions and can save their time in determining the ROI more precisely. After de-noising and enhancing the selected ROI, different texture features will be extracted as mentioned in Section 4.3.3. This feature matrix is reduced and fed to the classifiers to discriminate between normal and abnormal images. Finally, the whole approach was estimated using various evaluation techniques (more details in Section 4.4).

4.3.2.2 Approach two- Manual determination for AT ROI

According to the radiologists and the literature (Nunley, 2009), 80% to 90% of the AT ruptures occurs 2 to 6 cm proximal to its calcaneal insertion. Owing to this, this region will be the focus of the second approach as it is the most vulnerable area in AT. A window with dimensions of 4 cm width and 1 cm height has been selected from this area as shown in Figure 4.11. This window is determined as follows: at the beginning, the calcaneal insertion point is assigned and a line is drawn from this point along the AT width. Next, we draw another perpendicular line on the first line up to approximately 2 cm length; at that point, we start drawing the mentioned window to cover our ROI. This window with fixed coordinates is applied to all dataset samples. Tissue texture features and other features are extracted from the area in the specified window and explained in Section 4.3.3.

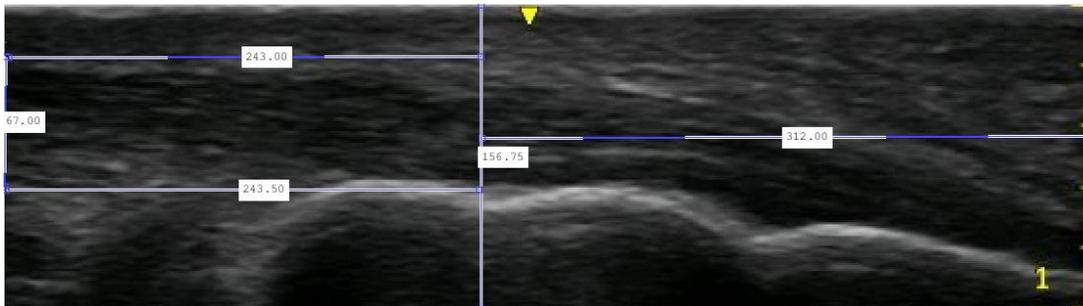


Figure 4.11 AT ultrasound image with manual determining for ROI

4.3.2.3 Approach three- Dividing the image into blocks

In this approach, we are going to achieve one of the research goals which is to automate the identification of the ROI area fully. In the first instance, the system divides the AT image into six equal non-overlapping blocks as shown in Figure 4.12. These blocks are equal in size and are in adjacent positions. Dividing the AT images is done automatically, and each block is saved as a separate image. Hence, our dataset is now composed of 215 blocks obtained from the 57 AT images after excluding 114 untargeted blocks and 13 distorted blocks. In Figure 4.13 the first block (a) starts from the Achilles insertion in the calcaneal followed by the second block (b) and together will occupy approximately 2.0 cm, starting from the calcaneal insertion point. These blocks are untargeted blocks in all AT images. The last four blocks starting from (c) to (f) in Figure 4.13 contain the most vulnerable AT area according to

(Yu & Bajaj, 2004). More focus is placed on these blocks in extracting various features and obtaining the classification results.



Figure 4.12 Dividing the AT ultrasound image into equal blocks

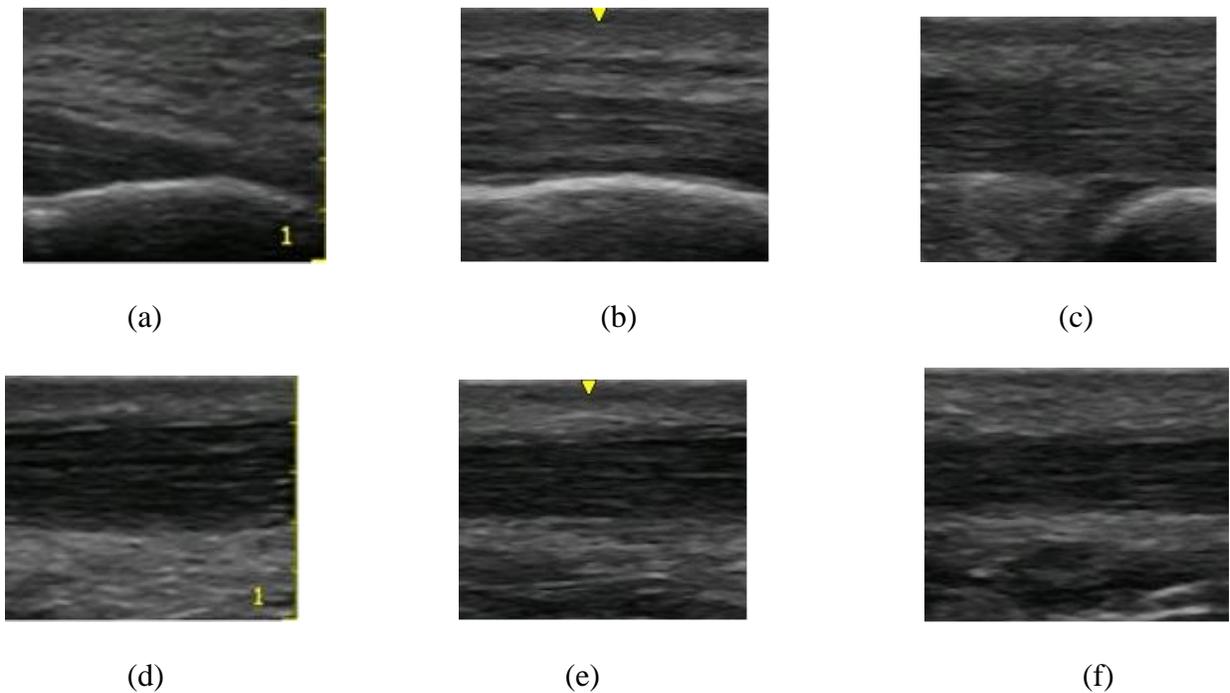


Figure 4.13 Dividing the AT ultrasound image into six equal blocks.

(a), (b), and (c) are AT ultrasound image lower part. (d), (e) and (f) AT ultrasound image middle part

4.3.3 Features Extraction

Extracting image features plays a crucial role in this research as it provides the crude data for the subsequent stages. The decision of whether the AT is normal or abnormal mainly relies on the features

extracted during this stage. In this study, various features are extracted from the AT ultrasound image using different feature extraction methods. Many methodologies have been proposed to analyse and recognise textures in an automated fashion (Mitra & Parekh, 2011). Most image features have been constructed based on histogram and texture features (Tian, 2013). The statistical descriptor, co-occurrence matrices for four different angles, and run length encoding were used for extracting different texture features. Besides, different histogram features such as skewness and kurtosis are extracted from samples after different segmentation methods were applied. The area, major axis, minor axis and other region properties are extracted. However, all these features are subject to filtering using reduction methods at later stages.

Feature extraction is an essential element of the classification process. Methods of extraction are majorly utilised for obtaining descriptions of specific image characters in a specific region. The main objective is to gather the characters in a number of regions of an image and utilise them for image recognition and classification. This work used texture features to classify images by the most prominent feature in the ultrasound images. In general, image texture can be described with the usage of both the structured approach and the statistical approach (Tuceryan & Jain, 1993).

A structured approach is not generally suitable for ultrasound image processing because ultrasound images are rich in texture and textural processing works well with disease classification (Tian, 2013). Also, the structured method is only utilised in the cases that objects that need to be recognised in the image. Therefore, the statistical method is implemented in this work since it is easier and more specific to use for textural classification. Currently, there are numerous methods in computer vision that handle the extraction of features from ultrasound images. Textural characteristics along with statistical characteristics are vastly important heuristic methods for the identification of patterns in imaging. Commonly used methods in textural assessments mostly function on the basis of statistical characteristics of the intensity histogram. An example of a frequently implemented method of characteristic extraction which is based on texture is autocorrelation function of textures GLCMs. Numerous techniques of extracting statistical features are present in the literature. Some of the current techniques are wavelets, edge frequency and run length. Another form of techniques is the geometrical

techniques; they analyse textural structure through the detection of the basis elements (Gonzalez & Woods, 2001).

In the last few years, GLCMs have been a successful approach in identifying cancer in various implementations in the medical field. As there is a great rise in the amount of GLCMs publicised implementations in digital mammography, it is simple to consider the GLCMs significance when comparing it to the rest of the methods. Therefore, the study by (Haralick et al., 1973) implements GLCMs in extracting texture feature from AT ultrasound images with the use of texture descriptors.

The extracted features in the current research will be split into three groups that are: histogram, GLCM and GLRLM.

4.3.3.1 Intensity histogram features

Histogram features is a visual feature which is greatly utilised in the classification and retrieval of images. An image histogram f_n is a N –dimensional vector $H(f_n, i) = 0, 1, 2, \dots, N - 1$ in which N is the quantity of grey levels and $H(f_n, i)$ is the number of pixels with grey level value i . The histogram property is somewhat robust to difficulties in backgrounds. Moreover, it is not sensitive to change in the rotation or size that has a small or no influence on grey level distribution in ultrasound images. Thus, image features may be expressed using the following measurements: Mean, Standard Deviation (SD), Skewness and kurtosis (Chitrakala, Shamini, & Manjula, 2009).

Mean shows the general image brightness. An image which is bright ought to have a great mean and an image which is dark ought to have a small mean. The mean is useful as it also distinguishes calcifications. The variance and the SD can demonstrate contrast in images; an image which has fine contrast will have high variance. Also, SD can distinguish the cluster. Skew is a value of imbalance in the grey level distribution. An ultrasound image that has bimodal histogram spread needs to have a big value of variance and low skew spread. Thus it should have a peak on both sides of the mean. Energy is similarly linked to skew. When there is a high skew distribution, there is typically a high-energy value. Entropy gives a mean of bits to code each grey level. Entropy has a reverse relation with skew and energy. When there is a high skew distribution, low entropy is typically yielded. In the ROI, an image distribution of histogram is calculated, after, there are six features which are computed for classification.

4.3.3.2 Grey Level Co-occurrence Matrix (GLCM)

GLCM is a popular textural descriptor that is commonly utilised; it has demonstrated that the attained outcomes from the co-occurrence matrices exceed that of different textural discriminations techniques ((Davis, 1975), (Huang, Wunsch, Levine, & Jo, 2008)). It calculates statistic characteristics on the basis of the ultrasound image grey level intensity. The features of GLCM like these are effective when it comes to the recognition of texture in images (Walker, Jackway, & Longstaff, 1997), segmentation of ultrasound images (Sahoo, 2011), analysis of colour, retrieval of images, classification (de Almeida, de Souza, & Candeias, 2010), identification of objects and textual analysis techniques (Lo & Don, 1989). Numerous approaches of statistics are implemented to indicate varying characteristics including first order parameters like variance and skewness and second order parameters such as correlation and contrast.

First statistical parameters such as mean, variance, skewness, and kurtosis are computed through the values of intensities with no thought to the neighbouring pixel spatial connections. Second order statistical parameters do however put into consideration the relationships among the neighbour intensities. Thus, to attain the spatial dependence correlation, the second order parameter is utilised. In a statistical analysis of textures, the features were calculated based on the statistical pixel intensity distributed at a specific point respective to the rest in a matrix of a pixel representing the image. The first order statistical parameters, the second order statistical parameters or higher order statistical parameters all work on the basis of the number of pixels in individual combinations. GLCM extraction of features is the second statistics which may be utilised to analysing image as a texture. GLCM is done by tabulating frequencies; a count of the occurrence of a combination of pixel brightness numbers in an image (Hall-Beyer, 2000).

4.3.3.3 Grey Level Run-Length Matrix (GLRLM)

The texture is described as a sequence of grey pixel intensities in a specific direction from the reference pixels. The GLRLM technique is commonly utilised for retrieval and classification of images, where it is used to extract the texture descriptor within the grey ultrasound images (Galloway, 1975). GLRLM is known as a matrix that is used in the extraction of textural features used for texture analysis (Thangavel, Karnan, Sivakumar, & Mohideen, 2005). It is an approach to analysing the image, across

known direction, for runs of pixels which are of an equal grey level value. The run length is a value of adjacent pixels which have an equal grey intensity in a specific route. GLRLM is a two-dimensional matrix in which individual elements are the number of elements j with intensity i in the θ direction. Therefore, with a known direction, the run length matrix can measure for every grey level value, the quantity of the runs. For instance, two successive pixels with an equal value. Following that, the same is done for three then four, five successive pixels, etc. Numerous matrices of run length can be calculated for an individual image, one for every picked direction (More explanations in Section 3.4.4.3). GLRLM works on the basis of calculating the number of grey level runs of different measures (Kim et al., 1997). A grey level run is a group of successive pixel points that lie on the same line and are of an equal grey level value. The run length is the number of pixel spots on the run. There are seven textural characteristics that are extractable from GLRLM. The features utilise the pixel grey level in succession and are aimed to differentiate texture which has an equal value of SRE and LRE but has variance in the grey level distribution.

4.3.3.4 Region Properties through Region Growing Algorithm

(Oghli, Fallahi, & Pooyan, 2010) Stated that region growing is a method of segmenting images; it initiates with a single pixel of a specific area and expands it through the addition of adjacent similar pixels until pixels that are used for comparison are unlike. With the majority of the methods for region growing, the seed point is picked with manual means (Kanwal, Girdhar, & Gupta, 2011). Within this stage, numerous little regions are combined in accordance with the similarity restrictions. In order to begin the procedure, the selection of a seed pixel is chosen and it is compared to the nearby pixels. At the start of the procedure, the area begins with the seed pixel, and from that point, the nearby pixels are observed. The variance between the intensity pixel value and the mean of the region is needed for use as a comparison value. The procedure comes to an end when the variance between the new pixel value and the mean of the region grows to be bigger than a specific threshold. The pixel incorporated is considered the regional border pixel and the analysis to incorporate into the region is implemented once more. The procedure is constantly repeated until there is no pixel included into the region.

With the purpose of extracting numerous region characteristics from the ROI, these stages are performed as follows:

- 1- In the middle of the image, pick a seed point manually; it should be within the area of interest and be a part of the ROI.
- 2- The growth of the region is implemented repetitively through the comparison of the region to the individual unassigned surrounding pixels.
- 3- The variance between the mean of the region and the intensity pixel value is utilised to determine the presence of similarities. When the pixel with the least variance is found, it is assigned to a particular region.
- 4- The process ends when the variance between the new pixel value and the mean of the region grows to be bigger than a specific threshold.
- 5- Implement morphological processes through a hole filling procedure, after that a morphological opening which is one step takes place (erosion followed by dilation).
- 6- The next step consists of constructing a matrix that allocates a particular number to individual neighbouring objects and computing the features of all the linked regions.
- 7- Then, region features are extracted, including the area, solidity and perimeter.

From the segmented areas of the image, there are three varying characteristics to be computed:

- Area:

The area is required to allow comparison of areas of regular and irregular ROI in AT images. This takes place when AT irregular areas transform with regards to abnormality reason. The area is computed as a scalar which identifies the number of pixels found in the region and expanded as a column in the matrix property.

- Perimeter:

Return the scalar which determines the length around the border of the area. Calculate the perimeter through the measure of the distance between all adjacent couples of pixels surrounding the boundary of the area.

- Solidity:

Shape values may be utilised to differentiate regular and irregular AT images. Regular images tend to have a smooth shape, and thus they form a normal shape in elastograms and ultrasound images.

Meanwhile, irregular images have abnormal shapes in elastographic imaging. The variance may be attained by means of a property known as solidity.

4.3.4 Features Reduction using KPCA

At this stage, the extracted features from the previous stage will be reduced into meaningful representations through a dimensionality reduction method. By this reduction, the curse of dimensionality is mitigated, and we might eliminate some features as just noise. Figure 4.11 Summarize the feature reduction process where the full original features are used as input to the feature reduction method, and the output are the transformed features.

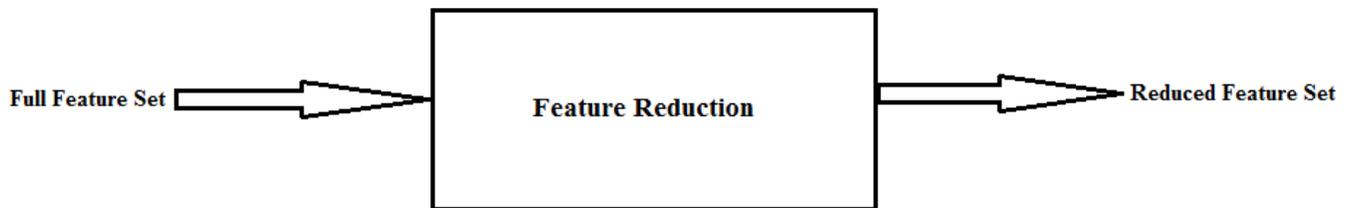


Figure 4.14 Feature reduction KPCA reduction method

In the case of complex nonlinear data, nonlinear techniques have shown superiority over their linear counterpart. The SVMs and the PCA are commonly known examples of linear classifiers. The majority of real life information, however, needs nonlinear techniques to perform tasks that consist of analysing and detecting patterns effectively. Many detection systems make use of the method based on eigenvector in dimensionality reduction. Numerous current recognition systems based on eigenvector make use of PCA (Ogata, Tan, & Ishikawa, 2006). In this study, we use KPCA as kernel based version from PCA. This algorithm is thought to be an addition to the linear edition to nonlinear distribution. KPCA has an aim of decreasing the data dimensionality, and at the same time, it maintains the most it can of the variation found in the initial set of data. The optimum low dimension place may be established through the optimum (largest) eigenvectors of the covariance matrix. The focal point of this part is to outline the Kernel based PCA technique with a Gaussian radius basis function (RBF) kernel that is used to perform nonlinear dimensionality reduction.

4.3.4.1 Principal Component Analysis (PCA)

The most important aim of PCA is to analyse data in order to determine patterns that analysis the data (Figure 4.15). The principal components may be comprehended like new data set axes which increase the variance over those axes. Therefore, PCA has a purpose of determining the axes that have the most variance along where the data is maximally distributed.

PCA:
component axes that
maximize the variance

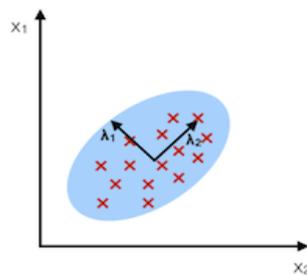


Figure 4.15 PCA component axes that maximize the variance
(Raschka, Linear, Gaussian, & LLE, 2014)

4.3.4.2 PCA and linear dimensionality reduction

A popular implementation of PCA is to decrease dimensionality of data sets while preventing information from going missing. The whole dataset (d dimensions) is presented in a subspace (k dimensions where $k < d$). This method of projection is effective in reducing costs of computing and the error of approximation of parameter.

The principle PCA method may be summed in the following steps:

- i) First, calculate the covariance matrix of the initial d -dimensional dataset x
- ii) Then, calculate the dataset eigenvectors and eigenvalues
- iii) Organize the eigenvalues so that they are arranged in a descending order
- iv) Select the k eigenvectors which match up to the k biggest eigenvalues where k is the number of dimensions for the new feature subspace

v) Build the projection matrix W of the k chosen eigenvectors

vi) Change the initial dataset X to attain the k -dimensional feature space Y , $Y=Wt \cdot X$

4.3.4.3 Nonlinear dimensionality reduction

The standard PCA method outlined above is a method of linear projection which functions greatly when the data is divisible linearly. Nevertheless, when data cannot be linearly separated, a nonlinear method needs to be implemented when dataset dimensionality needs to be reduced (Figure 4.16).

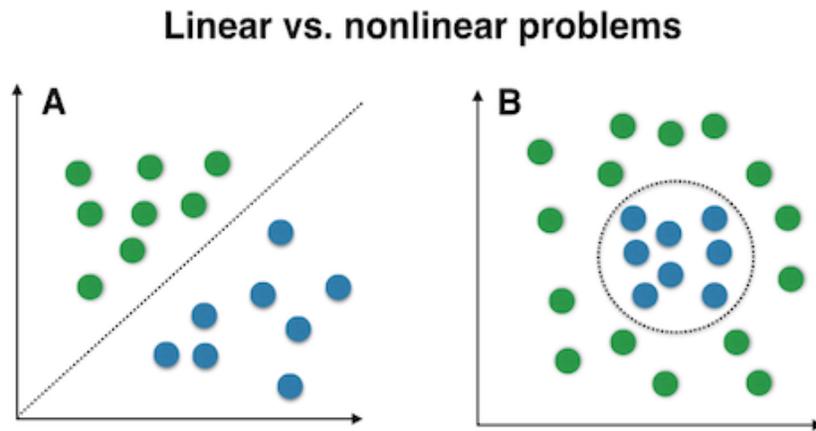


Figure 4.16 Linear vs nonlinear problems
(Raschka et al., 2014)

4.3.4.4 Kernel Functions and Kernel Tricks

The basic concept of handling linearly data that cannot be separated is to project it on a plane of a higher dimension at which it can be separated. The nonlinear mapping function is given by ϕ ; sample x will thus have a mapping given by $x \rightarrow \phi(x)$, this is known as the kernel function. Therefore, “kernel” represents the function which computes the dot product of the sample x image under ϕ .

$$k(x_i, x_j) = \phi(x_i)\phi(x_j)^T \quad \text{Equation (40)}$$

Thus, ϕ maps the initial d -dimensional characteristics into a bigger, k -dimensional characteristic place through producing nonlinear mixtures of the initial characteristics. For instance, when x has two features:

$$x = [x_1 \quad x_2]^T \quad x \in \mathbb{R}^d \quad \text{Equation (41)}$$

↓ ϕ

$$x' = [x_1 \quad x_2 \quad x_1x_2 \quad x_1^2 \quad x_1x_2^3 \quad \dots]^T \quad x \in IR^k \quad (k \gg d) \quad \text{Equation (42)}$$

Typically, the RBF kernel is mathematically defined and given by:

$$k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|_2^2) \quad \text{Equation (43)}$$

Where $\gamma = \frac{1}{2\sigma^2}$ is defined as a free parameter that will be optimized.

4.3.4.5 Gaussian radial basis function (RBF) kernel PCA

In the linear-based PCA method, the standard elements which increase the difference in the dataset are a point of interest. It is implemented through the extraction of eigenvectors which are the standard elements that fit with the biggest eigenvalues based on the covariance matrix:

$$Cov = \frac{1}{N} \sum_{i=1}^N x_i x_i^T \quad \text{Equation (44)}$$

This method was made general by Bernhard Scholkopf (1998) for information which is mapped to a bigger space of dimension through the following kernel function:

$$Cov = \frac{1}{N} \sum_{i=1}^N \phi(x_i) \phi(x_i)^T \quad \text{Equation (45)}$$

When implemented, however, the covariance matrix in the higher dimensional space is not computed explicitly, and this is known as the kernel trick. Thus, with the usage of RBF kernel PCA will not yield the principal component axes (unlike the classic PCA); however, the attained eigenvectors may be comprehended to be projections of information on the principal components.

Figures 4.17 and 4.18 highlight the difference between data scatter using linear PCA and nonlinear PCA, or kernel PCA, respectively. It is obvious that data in the linear PCA are overlapped making the classification process more difficult. In this study, we adopt Kernel PCA as it is considered to be a nonlinear technique in dimensionality reduction.

In this study, a matrix with 57 different texture features was extracted from the whole samples using the feature extraction methods mentioned previously. These features that were extracted by different extracting methods are gathered in one matrix where the row represents the samples dataset, and the columns represent the extracted features for each sample.

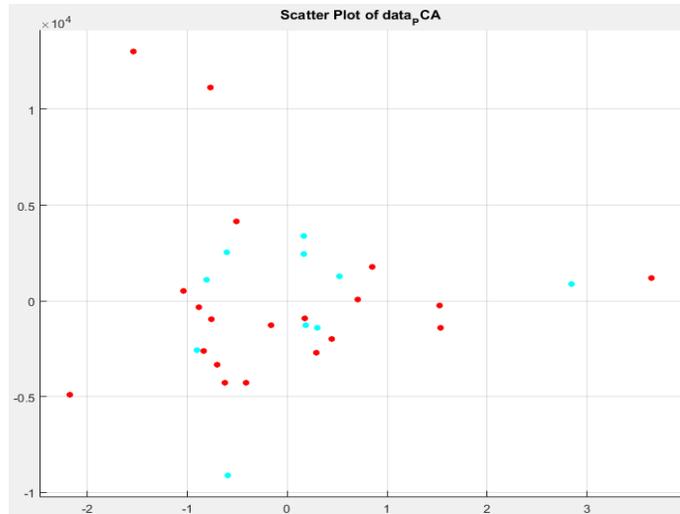


Figure 4.17 Data scatter using PCA reduction method

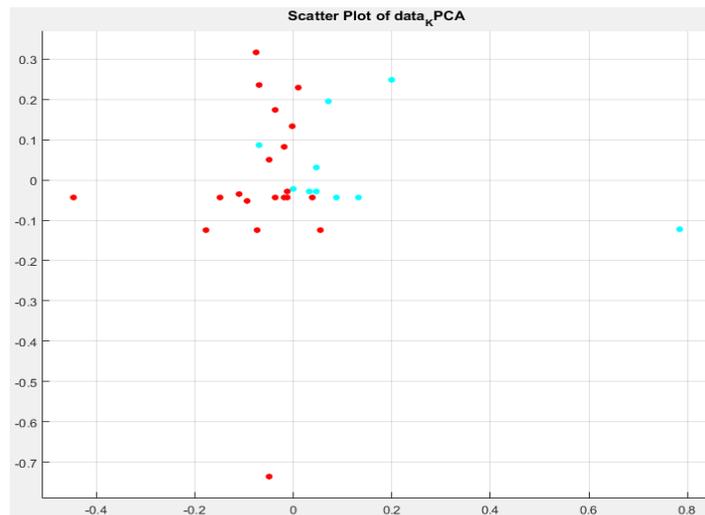


Figure 4.18 Data scatter using KPCA reduction method

By using the KPCA as a reduction method with Gaussian kernel function these features are reduced to a minimum number that corresponds to the largest eigenvalues in order to be an input for various classifiers as shown in Figure 4.19. For the supervised classification methods, we need to add a column to the reduced matrix as a label to distinguish between normal and abnormal samples.

	CJ	CK	CL	CM	CN	CO	CP	CQ	CR	CS	CT	CU	CV	CW	CX	CY	CZ	DA	DB	DC	DD	DE	DF	DG	DH	DI	DJ	A	B		
1	63.73974	63.45719	62.16652	63.12991	61.66274	63.85141	62.5428	63.50209	64.65116	63.01742	64.02898	62.23454	64.09453	63.03855	62.22125	63.5779	63.67869	62.85952	61.84746	63.38404	63.5824	63.37426	63.12737	63.10372	63.98549	63.42944	63.53338	1	-0.0075	0.053219	
2	62.51094	61.51173	57.99609	60.09845	57.90935	61.1477	59.86005	60.89566	63.81236	60.32267	62.18048	57.35199	63.0148	60.47586	58.96641	60.86377	62.01574	59.6605	58.47852	59.27268	61.86077	59.68585	60.88739	60.2367	63.47566	59.48935	61.86045	2	-6.4E-11	3.83E-10	
3	0.165788	0.17127	0.198707	0.177846	0.210573	0.16367	0.190283	0.170387	0.149273	0.180165	0.160358	0.197156	0.159152	0.179226	0.197459	0.168906	0.167054	0.187749	0.206142	0.172718	0.168819	0.172913	0.177898	0.178383	0.161163	0.171818	0.169774	3	-6.4E-11	3.83E-10	
4	0.027486	0.029333	0.039485	0.031629	0.044341	0.026788	0.036208	0.029032	0.022282	0.032459	0.025715	0.038871	0.025329	0.032302	0.03899	0.028529	0.027907	0.035247	0.042495	0.029832	0.0285	0.029899	0.031648	0.031821	0.025974	0.029521	0.028823	4	-0.00808	0.0486	
5	-0.08844	-0.10414	-0.14384	-0.12873	-0.15293	-0.11503	-0.12094	-0.11656	-0.06915	-0.12624	-0.09653	-0.14744	-0.07376	-0.12463	-0.13991	-0.11554	-0.09271	-0.13076	-0.14438	-0.12822	-0.09365	-0.12762	-0.1099	-0.12116	-0.06579	-0.12886	-0.09649	5	-0.00808	0.0486	
6	0.659356	0.62956	0.569344	0.616346	0.595042	0.625323	0.614063	0.63291	0.661983	0.634467	0.66591	0.509793	0.670638	0.633213	0.614417	0.62634	0.648884	0.617427	0.6104	0.594181	0.654212	0.571031	0.622922	0.5754	0.656907	0.565129	0.655552	6	-0.00808	0.0486	
7	0.86876	0.825425	0.858691	0.847112	0.8095216	0.859531	0.762276	0.851383	0.847323	0.830944	0.84945	0.842194	0.836441	0.806472	0.80611	0.823895	0.880545	0.804249	0.801306	0.855753	0.815825	0.851592	0.794603	0.867849	0.851144	0.8652	0.80849	7	0.004613	0.06595	
8	0.666434	0.651758	0.599391	0.649978	0.614403	0.658857	0.632319	0.656114	0.675083	0.657693	0.678129	0.556534	0.672794	0.660297	0.636905	0.664819	0.663627	0.647725	0.633328	0.630893	0.669695	0.604945	0.649903	0.614771	0.666318	0.60208	0.668449	8	0.004613	0.06595	
9	64.05224	63.15273	62.28082	63.22074	61.74381	64.0345	62.61968	63.65489	64.80736	62.84038	64.23446	62.16677	64.11979	62.85205	62.07979	63.81284	63.88426	62.40042	61.78002	63.41705	63.39407	63.36978	63.27317	63.23806	64.12632	63.51977	63.42879	9	0.004613	0.06595	
10	62.82344	61.20727	58.11039	60.18928	57.99043	61.33079	59.93692	61.04846	63.96856	60.14563	62.38597	57.28422	63.04007	60.28936	58.82495	61.10771	62.22631	59.4014	58.37308	59.30569	61.67245	59.68137	61.03319	60.37104	63.6165	59.57389	61.75385	10	0.004613	0.06595	
11	0.159929	0.17738	0.19611	0.175996	0.208617	0.160256	0.188606	0.167415	0.146613	0.183874	0.156609	0.1987	0.15869	0.183627	0.200701	0.164228	0.163052	0.193428	0.209478	0.172063	0.172519	0.173002	0.174937	0.175646	0.158571	0.170154	0.17183	11	0.004613	0.06595	
12	0.025577	0.031464	0.038459	0.030975	0.043521	0.025682	0.035572	0.028028	0.021495	0.03381	0.024526	0.039482	0.025183	0.033719	0.040281	0.026971	0.026586	0.037414	0.043881	0.029606	0.029763	0.02993	0.030603	0.030851	0.025145	0.028952	0.029526	12	0.004613	0.06595	
13	-0.08388	-0.10354	-0.14406	-0.12893	-0.15301	-0.11398	-0.12061	-0.11515	-0.06871	-0.12109	-0.0856	-0.13855	-0.07366	-0.12042	-0.14146	-0.11755	-0.09057	-0.13115	-0.14475	-0.09287	-0.12824	-0.11145	-0.09287	-0.12824	-0.11145	-0.07046	-0.12783	-0.09653	13	-0.00595	0.056714
14	0.652833	0.614961	0.561535	0.605439	0.583862	0.614534	0.603731	0.622454	0.649697	0.617036	0.658933	0.500774	0.63917	0.614213	0.59824	0.609532	0.638659	0.59723	0.590043	0.579015	0.637359	0.555676	0.615633	0.564074	0.649617	0.553276	0.64228	14	0.005487	0.044528	
15	0.890923	0.79514	0.871976	0.856734	0.816527	0.86958	0.770808	0.863339	0.862435	0.796861	0.864678	0.799628	0.842321	0.767956	0.783518	0.862968	0.895693	0.772228	0.79581	0.86612	0.785439	0.855018	0.821088	0.872023	0.881196	0.817359	0.795833	15	0.005487	0.044528	
16	0.68129	0.644078	0.608785	0.653976	0.620781	0.665468	0.640117	0.663254	0.685095	0.647884	0.694276	0.546883	0.678633	0.648645	0.627489	0.67119	0.675654	0.634204	0.622189	0.631061	0.659521	0.604096	0.662699	0.608763	0.681471	0.60682	0.667983	16	4.19E-10	-5.1E-11	
17	63.63755	63.20768	62.16788	62.95073	61.74525	63.67754	62.53621	63.55538	64.46006	62.86685	63.84327	62.2277	63.48521	62.9206	61.99067	63.41453	63.52347	62.51529	61.73401	63.35241	63.10527	63.21489	62.8297	63.18304	63.85639	63.3703	63.03961	17	4.19E-10	-5.1E-11	
18	62.40876	61.26222	57.99745	59.91927	57.99186	60.97383	59.85345	60.94894	63.62125	60.1721	61.99478	57.34515	62.40549	60.35791	58.75383	60.7004	61.86553	59.51627	58.36507	59.24105	61.38365	59.52648	60.58972	60.31602	63.34657	59.43022	61.36667	18	0.00538	0.044841	
19	0.16775	0.176261	0.198676	0.181553	0.208582	0.166979	0.190427	0.169345	0.152594	0.183315	0.163823	0.197312	0.170718	0.182184	0.202777	0.172113	0.169968	0.190887	0.208852	0.173488	0.178351	0.176115	0.1841	0.176762	0.163576	0.172992	0.179705	19	0.00538	0.044841	
20	0.02814	0.031068	0.039472	0.032961	0.043507	0.027887	0.036283	0.028678	0.023285	0.033804	0.026838	0.038932	0.029145	0.033191	0.041116	0.029623	0.028889	0.036438	0.043619	0.03005	0.031809	0.031016	0.033893	0.031245	0.026756	0.032294	0.032294	20	4.19E-10	-5.1E-11	
21	-0.07992	-0.10453	-0.14219	-0.12437	-0.14957	-0.11121	-0.11729	-0.11496	-0.0615	-0.12519	-0.09021	-0.14343	-0.07237	-0.12254	-0.13723	-0.11401	-0.0894	-0.12932	-0.1424	-0.12754	-0.08991	-0.12508	-0.10599	-0.11665	-0.06281	-0.1271	-0.09177	21	-0.01047	0.01166	
22	0.66113	0.626696	0.565741	0.615714	0.596668	0.626414	0.616789	0.634315	0.664039	0.633299	0.667721	0.510111	0.66883	0.632247	0.611539	0.625057	0.64879	0.615496	0.607536	0.594281	0.652156	0.568924	0.62283	0.575111	0.656404	0.564827	0.654544	22	-0.01612	-0.00884	
23	0.834059	0.796858	0.849963	0.810861	0.797763	0.830827	0.745005	0.843409	0.800548	0.806487	0.812939	0.822522	0.767221	0.77594	0.748197	0.796355	0.846162	0.769851	0.763935	0.848376	0.739991	0.823645	0.748236	0.855471	0.817286	0.848729	0.734208	23	-0.01612	-0.00884	
24	0.668637	0.644675	0.603553	0.644594	0.612798	0.654945	0.633738	0.654633	0.67681	0.649469	0.680318	0.555487	0.65996	0.653743	0.621571	0.660105	0.630676	0.623333	0.627799	0.655013	0.600177	0.646898	0.610656	0.605002	0.599171	0.657089	24	-0.01612	-0.00884		
25	64.03136	63.66835	62.20648	63.20838	61.65558	63.95608	62.52168	63.5997	64.8503	63.21728	64.22967	62.52543	64.39718	63.21484	62.24854	63.65223	63.87771	62.798	61.88934	63.40519	63.74654	63.35553	63.17659	63.16006	64.13284	63.47021	63.83375	25	0.003676	0.044194	
26	62.80256	61.72289	58.03604	60.17692	57.90219	61.25238	59.83892	60.99327	64.01149	60.52254	62.38118	57.64288	63.31745	60.65215	58.9937	60.93811	62.21976	59.79898	58.5127	59.29383	62.02492	59.66712	60.93661	60.29304	63.62301	59.53013	62.16082	26	0.003676	0.044194	
27	0.160314	0.167156	0.197795	0.178247	0.210747	0.16171	0.190746	0.169483	0.14589	0.176066	0.156695	0.190663	0.153702	0.176116	0.19684	0.167467	0.163175	0.184774	0.205285	0.172298	0.165658	0.173286	0.178693	0.17723	0.158452	0.171013	0.164003	27	0.003676	0.044194	
28	0.025701	0.027941	0.039123	0.031063	0.044414	0.02615	0.036384	0.028386	0.021284	0.030999	0.024553	0.038652	0.023624	0.031017	0.038746	0.028045	0.026626	0.034141	0.042142	0.029687	0.027443	0.030028	0.031291	0.03141	0.025107	0.029545	0.026897	28	0.003676	0.044194	
29	-0.08449	-0.10712	-0.14511	-0.12915	-0.15505	-0.11555	-0.12138	-0.11642	-0.06727	-0.12951	-0.0952	-0.15066	-0.07886	-0.12683	-0.14558	-0.1171	-0.0911	-0.1351	-0.14872	-0.13027	-0.09576	-0.12926	-0.10982	-0.12118	-0.06688						

The aim of a classification system is to assign a pattern to a class via the feature vector. There are numerous classifiers that are implemented and used in classifying patterns. The complications in classifying issues rely on the changes in feature values for patterns in the same class relative to the variance between feature values for patterns that are not in the same class, so, the dataset determines the attainment of the best performance of a classifier. Thus, obtaining the best performance for a pattern recognition system is not inevitably dependable with achieving the optimal performance of an individual classifier. In practice, a researcher may encounter a case in which, no individual classifier is able to create a classification with an appropriate level of accuracy. In these situations, it can be more beneficial to compile the results from various classifiers to make the best decision. Each classifier is able to function correctly on many varying means of the input feature vector. Under suitable hypothesis, gathering numerous classifiers might drive to enhanced generalization performance during the comparison with the individual component classifiers.

Following the extraction and the reduction of the features, these are input to a classifier to classify the images into either, normal or abnormal classes. In pattern classification, a supervised learning algorithm analysis the training data and creates a deduced function called a classifier. Throughout the stage of testing, the deduced function is applied to categorise the correct class membership for all valid pairs of input features. Consequently, the deduced function predicts the class membership of the testing samples. (Cortes & Vapnik, 1995) Confirmed that this needs the learning machine to generalise with the use of the training data to class unobserved samples in a suitable pattern. To be able to carry out the supervised learning project for a binary classification task, a labelled dataset of training patterns has to be supplied. To achieve this object, the training patterns are required together with the associated labels pointing out to their class membership. There is a large number of classifying techniques. In this study, the researcher uses popular and widespread classifying procedures to classify and detect abnormalities in medical images. SVM, k-NN, DT and Ensemble classifying procedures are known as nonparametric classifiers, and they are used to enhance accuracy in classification.

To estimate the performance of the classifiers, the medical image class is separated into two sets which are the training set and the testing set. In addition, the N-fold cross-validation method is utilised

to divide the medical images into N disjoint sets, in which $N-1$ functions are for training, and the N th set is utilised for testing.

4.3.6 Cross-validation

The k -fold Cross-Validation (CV) is a method where the dataset D is arbitrarily divided in k mutually exclusive folds F_1, F_2, \dots, F_k all have the same size. In a particular CV case known as *stratified cross-validation*, the folds are arranged in a way that allows them to include a proportion of samples which is equal to that of the original dataset. The learning algorithm is trained and tested k times, namely for each time $t \in ((1, 2, \dots, k))$ it is trained on $D \setminus D_t$ and tested on D_t as shown in Figure 4.15 (Kohavi, 1995).

The main benefit of the CV method with regard to the holdout procedure is that every pattern within the dataset is applied for training and testing. Meanwhile, the true error is predicted as the mean average error on the test patterns, consequently solving the problem initiating from unequal splits. In addition, (Kohavi, 1995) indicated that the rate of mean average error for the k -fold CV could be computed with the usage of the following expression:

$$e_{avg} = \frac{1}{k} \sum_{i=1}^k e_i \quad \text{Equation (46)}$$

Where k is the number of folds, and i is the true error rate for all the individual k -folds.

(Kohavi, 1995) Has drawn numerous interesting considerations in connection with the option of the right number of folds k . Firstly; if the number of folds k is big, the bias of the true error approximation is small, so, the estimator might be thought of as perfect. Unfortunately, because of the great amount of repetition, the difference between the true rate estimator and the computing time is predicted to be big. Secondly; if the number of folds k is minimised, the bias of the true error approximate is in general big, therefore the estimator might be thought of as conservative, or as bigger than the true error rate.

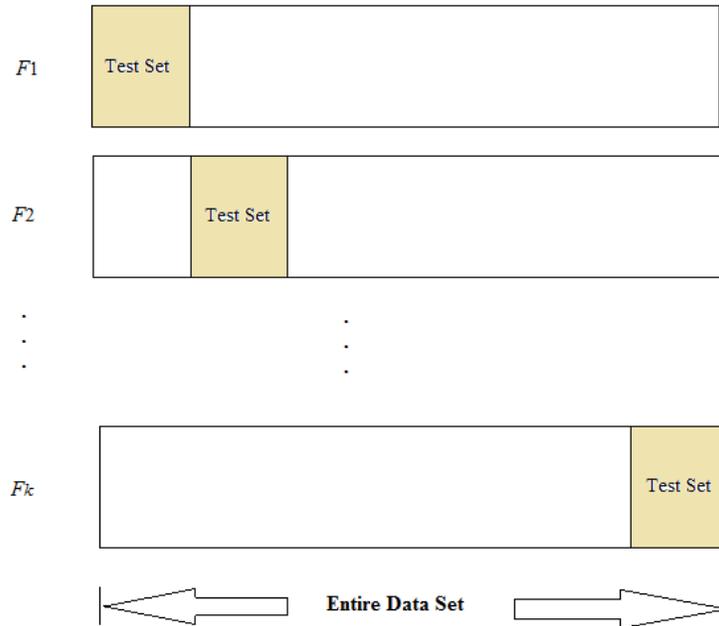


Figure 4.20 k -fold cross-validation method

In these cases, as a result of the decreased number of repetitions, the difference of the true error rate estimator and the computing time is usually small. In practice, the selection of the CV folds' number highly relies on the dataset size. With big datasets, a 3-fold CV can even be quite accurate. With thinly distributed datasets, it can be important to split the dataset into a great number of folds to train on a possible number of patterns. A popular option for k -fold cross validation is usually, $k = 10$

4.4 Evaluation Methods

4.4.1 Image Quality Evaluation Metrics

Many speckle reduction filters have been suggested in the ultrasound image processing field. Evaluating the performance of these filters has been acquired by a different number of procedures and methods. In addition, common techniques for evaluating relative performance are quantitative image quality metrics, and qualitative inspection and the choice is left to the researchers. During recent years, a big deal of effort has been made to improve image quality metrics that linked with realizing quality measurement. Test data and information for evaluation encompass clinical and phantom images also simulated the US that allows evaluation of relatively filtering to an ideal speckle free reference. Therefore, (Jeong, Kim,

Kim, & Kim, 2011) indicated that the measurement of US image enhancement is hard and there is no individual algorithm available to measure the enhancement of US image. Resulting images of this study will be evaluated using different quality evaluation metrics such as *Average Difference* (AD), the *Coefficient of Correlation* (CC), *Root Mean Square Error* (RMSE), *Signal to Noise Ratio* (SNR) and *peak signal to noise ratio* (PSNR). Moreover, in this study, the performance of each filter will be estimated quantitatively for AT ultrasound image with speckle noise utilizing the quality metrics like *Mean Square Error* (MSE), *Signal-to-Noise Ratio* (SNR), and *Image Quality Index* (IQ).

Let x and y denote the original and de-noised image.

SNR: Signal to Noise Ratio (SNR) (Sakrison, 1977) compares the level of the desired signal to the level of background noise. The higher the ratio, the less obtrusive the background noise is.

$$SNR = 10 \log_{10} \frac{\sum_{i=1}^M \sum_{j=1}^N (x_{i,j}^2 + y_{i,j}^2)}{\sum_{i=1}^M \sum_{j=1}^N (x_{i,j} - y_{i,j})^2} \quad \text{Equation (47)}$$

Where M and N are the width and height of the image. The larger SNR values correspond to the good quality image.

RMSE: The Root Mean square error (RMSE), is given by (Gonzalez & Woods, 2001):

$$RMSE = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (x_{i,j} - y_{i,j})^2} \quad \text{Equation (48)}$$

PSNR: Peak Signal to Noise Ratio (PSNR) is computed using (Sakrison, 1977):

$$PSNR = 10 \log_{10} \left(\frac{g^2 \max}{MSE} \right) \quad \text{Equation (49)}$$

Where $g^2 \max$ is the maximum intensity in the unfiltered images. The PSNR is higher for a better transformed image.

IQ: The universal Quality Index (Z. Wang, Bovik, Sheikh, & Simoncelli, 2004):

$$IQ = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \frac{2\bar{y}\bar{x}}{(\bar{y})^2 + (\bar{x})^2} \frac{2\sigma_y \sigma_x}{\sigma_x^2 + \sigma_y^2} \quad \text{Equation (50)}$$

Where \bar{x} and \bar{y} represent the mean and σ_x, σ_y the standard deviation of original and de-speckled images. σ_{xy} represents the covariance between the original and de-speckled images.

4.4.2 ROI Detection Evaluation

(Jiang, Marti, Irniger, & Bunke, 2006) proposed two approaches for the experimental evaluation for image segmentation, feature based and task-based. The feature based approach measures the performance of the algorithm based on the quality of the extracted features under consideration such as edges and regions. The task-based approach uses a much different philosophy where the image segmentation represents only one step in achieving the high-level goal of a vision system such as object recognition, and the most important and final interest is the overall performance of the system. In conclusion, they stated that “instead of abstractly comparing the performance of segmentation algorithms, it may be thus more meaningful to conduct an indirect comparison based on their influences on the final performance of the entire system”.

4.4.3 Classification Evaluation

There are many ways of evaluating and measuring a classifier’s performance. For supervised learning with two possible classes, all measures of performance are mainly depending on four ratios that are acquired from using the classifier on the test dataset. These ratios are known as true positives tp , false positives fp , true negatives tn and false negatives fn as explained in more details in Section 3.6.2.

The researcher trains the classifier on the training set and uses the test set for its evaluation and then measure the performance through comparing the predicted labels with the correct labels. It is important to split the dataset into two parts, the training and the testing subsets. As a rule of thumb, the dataset is usually divided into training set and test set using a 70%/30% ratio respectively. Moreover, this process (split the dataset into training and testing subsets) is usually randomly accomplished to guarantee that both subsets should be random samples from the same distribution. It is reasonable to perform stratified sampling, to ensure that each class is present in the same proportions in both the training and test subsets.

To determine the discriminatory ability of textural features, ROC curve analysis has been applied and it is found to be satisfactory. In addition, (Fawcett & Provost, 1997) introduced the ROC analysis to the data process and machine learning communities. Furthermore, (Swets, 1988) confirmed that the analysis of the ROC had been expanded to be applied in visualising and analysing the patterns of the diagnostic

systems. Furthermore, Altman and Bland, (1994) indicated that the ROC is particularly useful when comparing two or more measures and this view is supported by (Morrison et al., 2003) that showed that ROC had been applied widely in the biomedical field to assess the effectiveness of diagnostic tests in distinguishing between healthy and diseased individuals. (Spackman, 1989) Was one of the earliest adopters of ROC curves in machine learning and explained the value of ROC graphs in estimating and comparing algorithms. ROC graphs are particularly useful for domains with skewed class distribution (Fawcett, 2004). They estimate the value of a diagnostic test by presenting a level measure of the ability of a test to correctly classify subjects. (Morrison, 2005) Showed that an AUC adjacent to 1 points out to a strong discriminatory power/ability of the index variable while an AUC adjacent to 0.5 points out to the variable having a small discriminatory power. In this study, ROC graph analysis was applied to determine the discriminatory ability of the tested textural features in discriminating both normal and abnormal groups.

The most commonly reported measure of a classifier is accuracy. This measure evaluates the overall efficiency of an algorithm. However, as illustrated earlier, predictive accuracy can be a misleading evaluation measure when the data is imbalanced. This is because in such cases, more weights are placed on the majority class than on the minority class making it more difficult for a classifier to perform well in the minority class. Sensitivity, another performance measure, measuring the accuracy of positive cases whereas specificity measures the accuracy of negative cases. Sensitivity assesses the effectiveness of the classifier on the positive/minority class while specificity assesses the classifier's effectiveness on the negative/majority class. For any given analysis, there is usually a trade-off between sensitivity and specificity. Precision, on the other hand, is a measure of a model's exactness. A higher precision value for a classifier is an indication of a good classifier.

Regardless of the improvements of the previous techniques on the model performance through the parameter tuning, it is clear to see that, the best performing model usually is chosen based on the testing subset but based on the performance measure of the training subset. The distribution of the testing data might be varying from that of the training data, and the true misclassification cost may be unknown during the learning time. Moreover, the testing data needs to be consistent and consider the state of nature of the real data in order to produce perfect estimates for future events.

Therefore, sampling techniques cannot be applied to test data to fully balance the class distribution. In such situations, it is the duty of the researcher or practitioner to determine an appropriate performance measure to use when choosing between different classifiers. (Akosa, 2017) Found that a number of assessment metrics are adopted in the machine learning community for learning problems with imbalanced data sets. In this study, seven measures will be used as the performance metrics.

- **Precision**

Precision is a measure of a model's exactness. A higher precision value for a classifier is considered as an indication of a perfect classifier

$$Precision = \frac{TP}{TP+FP} \quad \text{Equation (51)}$$

- **Geometric Mean**

It is possible to define the Geometric Mean (G-Mean) as a metric that measures the equilibrium between minority classes and classification performances on both the majority. A low G-Mean is an indication of a bad performance in the classification of the positive situations even if the negative situations are accurately classified as such. This measure is substantial in the avoidance of overfitting the both of under appropriate the positive class and negative class.

$$G - Mean = \sqrt{\text{Sensitivity} \times \text{Specificity}} \quad \text{Equation (52)}$$

- **F-Measure**

The F-Measure conveys the balance between the precision and sensitivity. The measure is zero when either the precision or the sensitivity is zero. The formula for this measure is given by equation

$$F - Measure = \frac{2 \times \text{sensitivity} \times \text{precision}}{\text{sensitivity} + \text{precision}} \quad \text{Equation (53)}$$

When the data is balanced this measure, however, performs well.

- **Balanced Accuracy**

The balanced accuracy is defined as the mean average between the sensitivity and the specificity, which can measure the average accuracy acquired from both the minority and majority classes. This quantity decreases to the traditional accuracy if a classifier performs equally well in either class. Conversely, if the high value of the traditional accuracy is consequent to the classifier taking

advantage of the classification of the majority class, then the balanced accuracy will be minimized compared to the accuracy.

$$\text{Balanced Accuracy} = 1/2(\text{sensitivity} + \text{specificity}) \quad \text{Equation (54)}$$

- **Matthew's Correlation Coefficient**

The Matthews correlation coefficient (MCC) is less influenced by imbalanced data. It is a correlation coefficient between the observed and predicted classifications. The value of correlation (r) ranges from -1.0 to +1.0 with a value of +1 representing a perfect prediction, zero as no correlation (no better than random prediction) and -1 the worst possible prediction.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad \text{Equation (55)}$$

- **Cohen's Kappa (or Kappa)**

Kappa takes into account the accuracy that would be generated purely by chance. The form of the measure is:

$$kappa = \frac{\text{total accuracy} - \text{random accuracy}}{1 - \text{random accuracy}} \quad \text{Equation (56)}$$

Where

$$\text{total accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad \text{Equation (57)}$$

And

$$\text{random accuracy} = \frac{(TN+FP)(TN+FN) + (FN+TP)(FP+TP)}{(TP+TN+FP+FN)^2} \quad \text{Equation (58)}$$

In a similar method to the MCC, Kappa takes on values from -1.0 to +1.0, with a value of zero means there is no agreement between the actual and classified classes. A value of 1.0 reference to a perfect concordance of the model prediction and the actual classes and a value of -1.0 indicates disagreement between prediction and the actual.

- **Youden's Index**

Youden's index evaluates the ability of a classifier to avert misclassifications. In addition, it sets equivalent weights on a classifier's performance on both cases (positive and negative).

Thus:

$$\text{Youden's index } (\gamma) = \text{sensitivity} - (1 - \text{specificity}) \quad \text{Equation (59)}$$

A higher value of γ is an indication of a good performing classifier.

4.5 Conclusion

The main objective of estimating an algorithm is to realize its behaviour in dealing with different categories of images, and/or help in evaluating the best factors implementation for various applications. Eventually, this may involve some comparison with similar algorithms, in order to rank their performance and presents guidelines for selecting algorithms on the foundations of the application domain. Estimating the performance of any algorithm in image processing is difficult in order that, the performance is mainly depending on many implementation factors.

Metrics could be used to monitor or to make the most effective use image quality and estimate algorithms of the image classification. In addition, there are more than one single metric that correlates quite with image quality as understood by the human visual system. Chosen of a convenient evaluation methodology is mainly depending on the aim of the task. Given some acceptable performance measures, the researchers are faced with the problem of selecting a specific one in an estimation task. However, it is significant and important to recognize that, the performance measures might be themselves biased in specific and particular situations. Instead of using a single measure, a collection of measures an overall implementation measure must be considered. In general, such a combination approach has received relatively less attention in the literature so far. In addition, it is believed that it will get hold of better behaviour by avoiding the bias of the individual measures. The performance measures presented in this chapter of the study are applied in the next chapter in order to estimate the study approaches.

Chapter 5 – SYSTEM DEVELOPMENT AND IMPLEMENTATION

This chapter deals with the design and implementation of the image processing decisional system, including various processing algorithms and classification techniques. We also present the implementation model and its related system components with explanation and description of the user interface and operation.

5.1 Introduction

The main objective of this chapter is to offer some insight into the implementation phase of the research, based on the methods and algorithms discussed in the previous chapter. Most of these image processing methods have been provided to meet the requirements of the research needs and all techniques and algorithms presented in the previous chapter are implemented using MATLAB. The system could be implemented using any powerful programming language, but at this stage, MATLAB can achieve this purpose, in an easier and faster way.

This chapter is organised as follows: Section 5.2 presents the purpose of the development of the AT decisional system while the software design and functionality are explained in Section 5.3. A review and analysis of the classification algorithms complexity are given in Section 5.4, and Section 5.5 presents a summary.

5.2 The Purpose of Developing AT Decisional System

Developing the AT decisional system part fulfils the research objectives and addresses the image classification challenges. The proposed system gives the research ability to test and evaluate the proposed algorithms. Also, it can give the researcher measurable results by implementing different algorithms on the collected AT ultrasound images. Such a system should meet and achieve the research requirements.

The wider aim of this work is to address the classification problem of AT ultrasound images. Different classification techniques have been implemented and tested, and the results from various standard and ensemble classifiers have been registered. By implementing such a system, a comparison between these classifiers is made easier with a suitable user interface chosen for the present system. The user interface has been implemented to be a user-friendly interface and to provide all image processing methods that have been used in this work.

5.3 The Software Design and Functionality

The Graphical User Interface (GUI), chosen for the present system, is a widely used method to accommodate interaction between machines and human users. User interfaces can take many forms, but always accomplish two fundamental tasks: communicating information from the user to the machine (System Input Interface) and communicating information from the machine to the user (System output Interface). The main screen of the system is shown in Figure 5.1.

The system user interface has been implemented to be a user-friendly interface and to provide all image processing methods that have been used in this work and explained in Chapter 4. Most of these image processing methods have been provided to meet the requirements of the research needs.

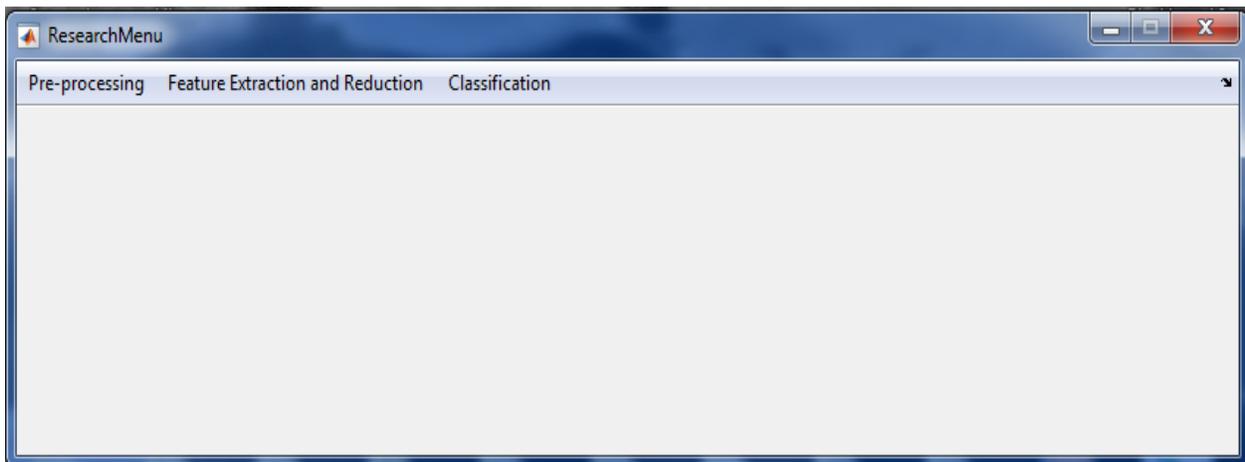


Figure 5.1 Research main menu

5.3.1 Image Pre-processing

Choosing a “Pre-processing” option from the main menu as shown in Figure 5.1 provides access to several options of pre-processing methods. Figure 5.2 shows the GUI of pre-processing selection. The pulldown menu shows all pre-processing methods where a researcher can apply them to all dataset images. The researcher can choose one of the pre-processing methods to apply to the AT ultrasound images, and image cropping, de-noising and enhancing are the available options. All dataset images are in JPG format and located in the “Research Dataset” folder. As mentioned in section 4.3.1.1, the cropping process removes the unrelated area to speed up the diagnosis process and give more focus on the ROI. Cropping coordinates are taken under the expert's supervision to avoid cropping the ROI, and all the cropped images are saved in the “Cropped Images” folder.

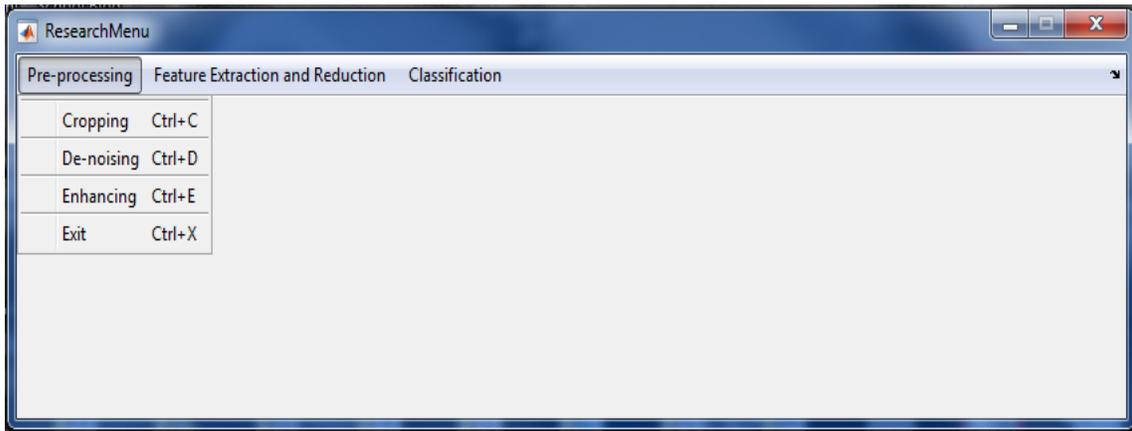


Figure 5.2 Image pre-processing selection

De-noising option calls SRAD de-noising function, which starts working on AT ultrasound images by removing the speckle noise and preserving the ROI edges as described in section 4.3.1.3. SRAD combined both the ordinary nonlinear anisotropic diffusion proposed by P&M as well as Lee and Frost Filters. Also, SRAD algorithm replaces the gradient-based edge detector in an original anisotropic diffusion PDE with the instantaneous coefficient of variation that is suitable for speckle filtering. All the de-noised AT images are stored in “De-noised Images” folder.

To emphasise some image features such as edges, contours and contrast image enhancing algorithm was applied. By choosing ‘Enhancing,’ option CLAHE function is called to enhance the de-

noised AT ultrasound images. Histogram equalization algorithm is very popular for image enhancement because of its simplicity and better performance. CLAHE operates on small regions in the image, called *tiles*, rather than the entire image. Each tile's contrast is enhanced so that the histogram of the output region approximately matches the histogram specified by the distribution parameter, which is 'uniform' flat histogram. The neighbouring tiles are then combined using bilinear interpolation to eliminate artificially induced boundaries. The contrast, especially in homogeneous areas, can be limited to avoid amplifying any noise that might be present in the image.

5.3.2 Features Extraction and Reduction

As Figure.5.3 shows, the “Features Extraction and Reduction” selection is divided into two main sub-choices: “Extract All Features” and “Feature Reduction KPCA”. The first sub-choice gives the researcher the ability to extract textual features to form data, and the second sub-choice is provided to give users the ability to reduce the extracted textual features using kernel PCA method.

The features extraction choice extracting the four different statistical approaches to texture analysis are as follows: First order intensity histogram features, second-order histogram using GLCM and GLRLM methods to determine the relationship between a pixel and its neighbours over an entire image in various directions. All these approaches were explained in detail in the previous chapter in Section 4.3.3. The extracted features are saved in a separate Excel sheet in the “Extracted Features” folder.

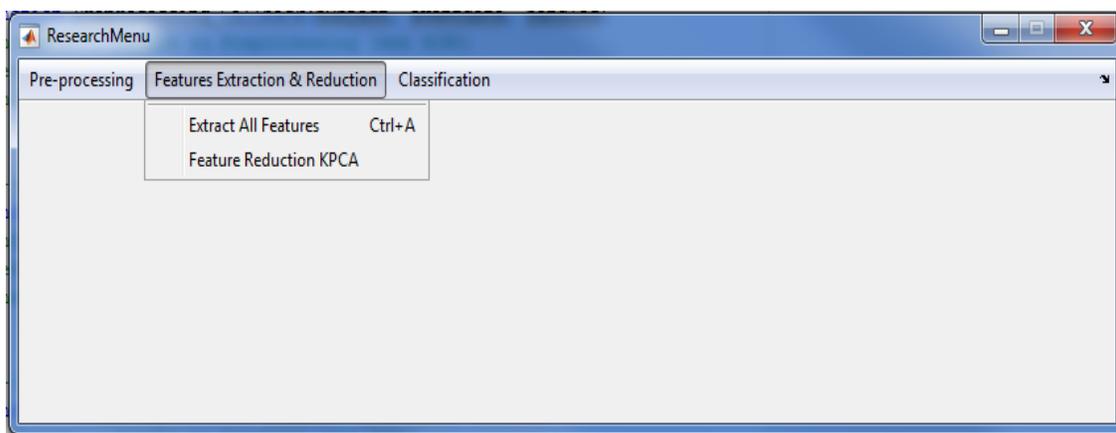


Figure 5.3 Features extraction and reduction

Dataset feature reduction can be used to increase the performance of machine learning algorithms when applied to different datasets of an image. In choosing the “Feature Reduction” sub-choice, the extracted features are projected on a plane of a higher dimension at which it can be separated. Kernel PCA algorithm is used and the excel sheet contains the extracted features uploaded through a standard dialogue box. The new reduced and transformed features will be stored in a new Excel sheet through a new dialogue box. The user specifies the place and a name to save the reduced features. Figure 5.4 shows the GUI of the feature reduction selection (Van Der Maaten, Postma, & Van den Herik, 2009).

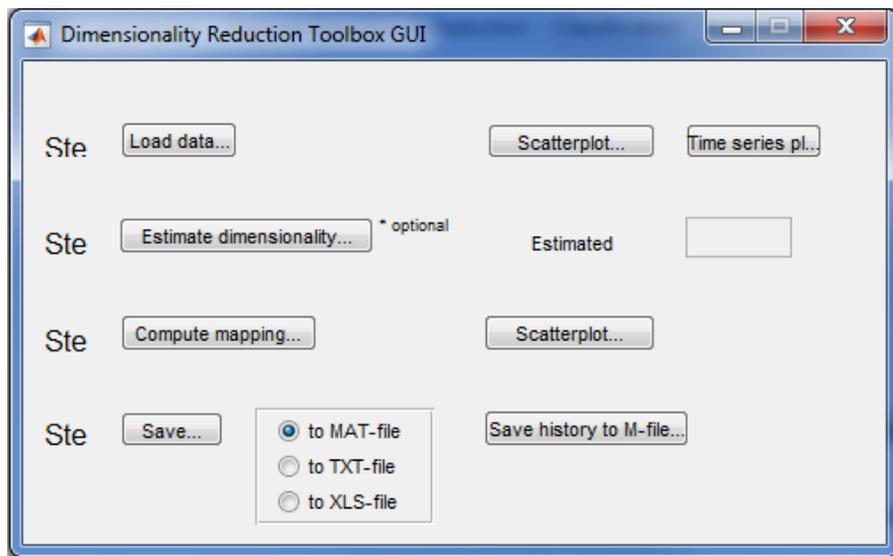


Figure 5.4 Features reduction menu

5.3.3 Image Classification

In this part of the system, the researcher employs many classifiers in order to compare their accuracy and efficiency using certain metrics mentioned in Section 4.4.3. Each classifier should go through the training stage first to build a classifier model that can be used with any new feature data. All supervised learning methods start with an input data matrix, where each row represents one observation, and each column represents one variable or predictor. Eight different classifiers have been adapted in this research, four are standard, and the rest are ensemble classifiers. Figure 5.5 shows the ‘Classification Menu’ with the loading and classifier buttons.

In this work, the training data was divided into two sets with 70% of the data used to train the chosen classifier, and the remaining 30% of the data used for testing 10 times and calculating some measurements through the confusion metrics: True positives rate (TPR), False positives rate (FPR), Accuracy (ACC), True negatives rate (TNR), Positive predictive value (PPV), and Negative predictive value (NPV).

Before applying any classifier, the research data should be loaded by ‘Load Data’ option. The samples’ features, which is reduced through the feature reduction process along with sample labels, will feed to the classifiers. The standard dialogue box appears to give the user the ability to choose the text file containing training data which has already been generated using the previous selection “Feature Reduction”, which was explained previously. After the user loads the research data and chooses a classifier among the available standard and ensemble classifiers (e.g. Decision Tree), the trained classifier starts work on the uploaded data as shown in Figure 5.6.

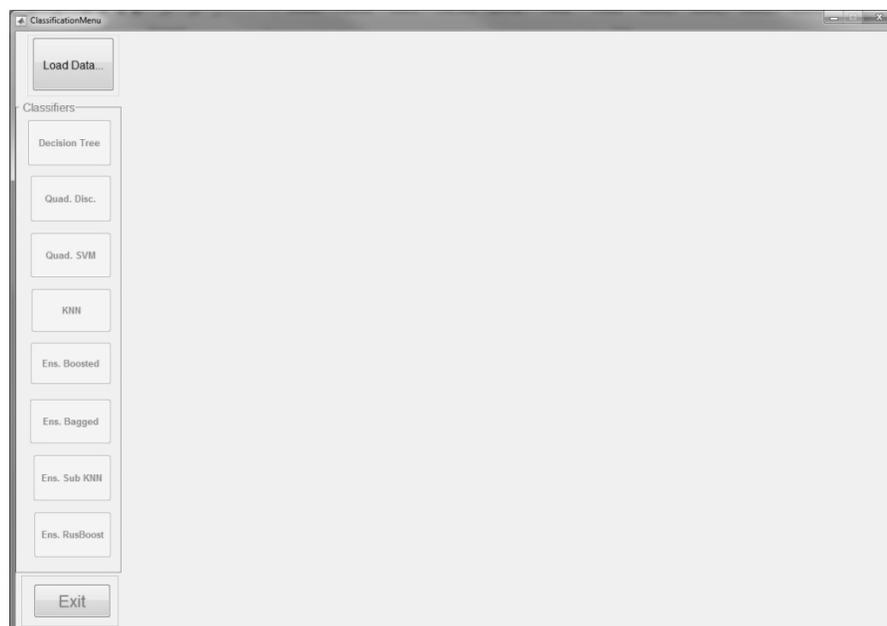


Figure 5.5 Images classification menu

The result of the “Classifier” choice is dependent on the outcome of all the previous operations: Pre-processing, Extract Image Features, Feature Reduction and Training the classifier. The researcher can

easily choose any classifier from the list of classifier models shown in the system and record the classifier results in order to determine the most proper classifier for the new data.

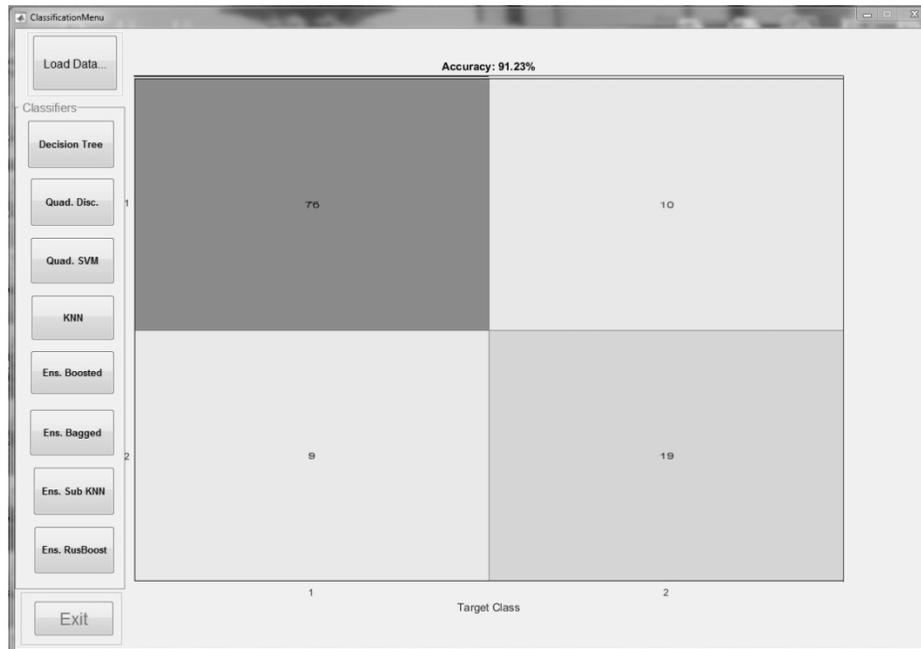


Figure 5.6 An Example of images classification result

5.4 Analysis of Classification Algorithms Complexity

The aim of supervised machine learning is to build a model that makes predictions based on evidence in the presence of uncertainty. As adaptive algorithms identify patterns in data, a computer learns from the observations. When exposed to more observations, the computer improves its predictive performance.

There are several characteristics of algorithms that can be used to estimate the classification algorithm complexity, such as: classification accuracy, training time and prediction speed.

The experiments are conducted on a Personal Computer with the following specifications:

Processor and Speed: Intel Core (i7), 3.60 GHz

RAM: 16.0 GB

And MATLAB version 2017a is used.

Table 5.1 : Performance evaluation metrics for standard and ensemble classifiers

Classifiers	Classification Accuracy (%)	Prediction Speed (Obs/Sec)	Training time (Secs)	Learners Numbers
Decision Tree	82.46	11000	0.253	1
Quadratic Discriminant	74.38	5900	0.245	1
Quadratic SVM	62.80	8900	4.558	1
KNN	81.42	8300	0.192	1
Ensemble (Boosted Trees)	82.46	1200	2.418	30
Ensemble (Bagged tree)	83.34	1200	1.731	30
Ensemble (Subspace KNN)	83.88	590	1.956	30
Ensemble (RUSBoost Trees)	78.76	1400	1.801	30

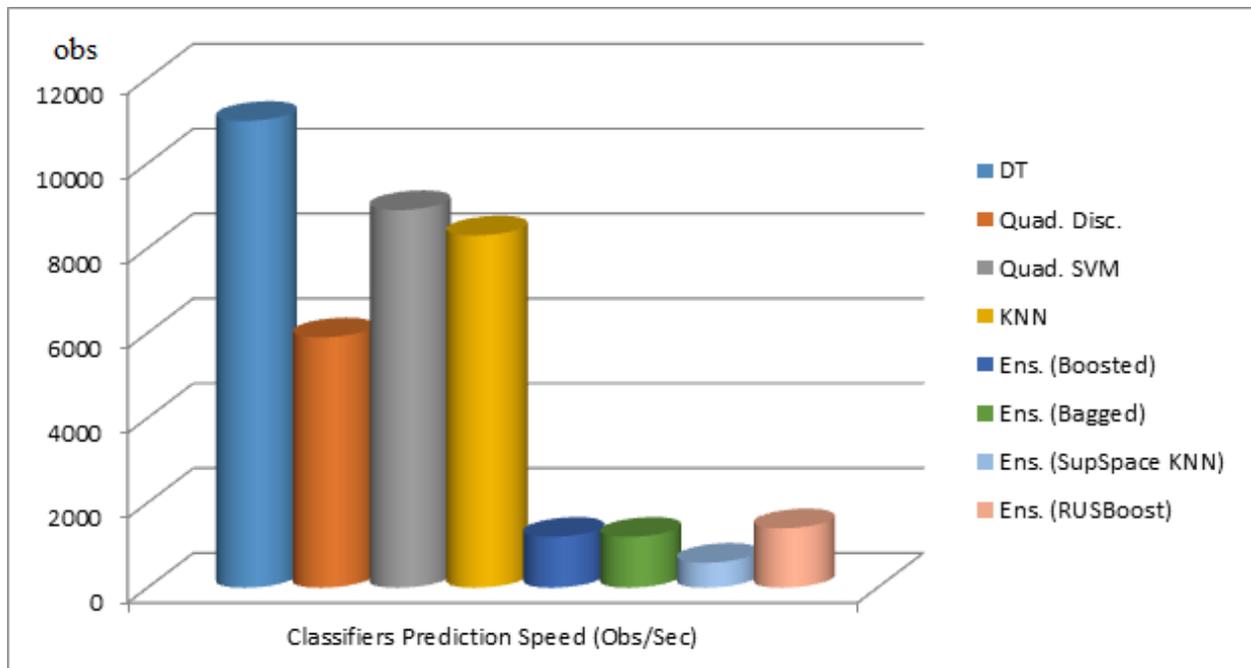


Figure 5.7 Comparison between various classifiers' results in their prediction speed

Table 5.1 shows the comparison of various metrics for estimating the classifier complexity. Figure 5.7 shows the graph for the prediction speed of the applied classifiers, where the y-axis represents the number of observations per second. It is clear that the standard classifiers have a large number of obs/sec in comparison to the ensemble classifiers. The reason behind such a result is due to the number of classifiers that participate in the classification process in the ensemble classifiers. Among all standard classifiers the ‘Decision Tree’ produced the best prediction speed of 1100 obs/sec and ‘Quadratic Discriminant’ records the lowest prediction speed with 5900 obs/sec. The ensemble classifiers Ensemble (RUSBoost) has slightly better results than Ensemble (Bagged) and Ensemble (Boosted) with values of 1400 obs/sec, 1200 obs/sec and 1200 obs/sec respectively, while ‘SubspaceKNN’ has the weakest prediction speed with the value of 590 obs/sec.

Regarding classification accuracy, the four ensemble classifiers generally had good values of accuracy and small error rate with values of 82.46, 83.34 and 83.88 for Ensemble (Boosted Trees), Ensemble (Bagged tree) and Ensemble (Subspace KNN) correspondingly. Decision Tree and KNN had good values of accuracy, but they did not exceed those of the ensemble classifiers. The Quadratic SVM classifier had the lowest values of classification accuracy with 62.80 percentages.

Also, the training time showed the best values for Decision Tree and Quadratic Discriminant with 0.253 and 0.245 respectively. Most of the ensemble classifiers had good values in the training time with an average value of 1.823 sec recorded for Ensemble (Bagged tree), Ensemble (Subspace KNN) and Ensemble (RUSBoost). The highest training time had been registered for Quadratic SVM and Ensemble (Bagged Tree) with values 4.558 and 2.418 respectively.

5.5 Summary

To conclude this chapter, the researcher offers a collection of the proposed approaches and algorithms for each stage in one platform package. This package has been designed and implemented to make it user-friendly; also, it gives the researcher an opportunity to evaluate and test various choices. In addition, it gives the researcher the ability to compare different classification techniques. One of the

main purposes of implementing this package is to find a base system in order to help the researcher to determine the system requirements and to achieve a complete system that meets all needs.

Chapter 6 - RESULTS AND DISCUSSION

This chapter deals with the comparison of de-noising techniques namely, linear and nonlinear filtering. We also evaluate the classification performance of the proposed approaches on AT ultrasound images. Experiments' results are provided in this chapter followed by a thorough discussion of these results.

6.1 Introduction

The results reported in this chapter are subdivided into image de-noising/enhancing results and image classification results. Each set of results is presented in different tables using various metrics. A discussion will be included at the end of the chapter about the research findings, quality of the results and the extent to which the goals of the research have been achieved. The selection of the de-noising technique is application dependent; in this study, we have compared a set of de-noising techniques to select the most suitable technique for the current application. Different image quality metrics were used in this research as mention in Section 4.4.1 to compare between various de-noising algorithms. The results obtained when applying these metrics will be used as the quantitative standard for comparison. In addition, a set of experiments were carried out to evaluate the influence of the ROI size on the classification accuracy. In this research, for all the proposed classifiers we use the same features data matrix, which is extracted from AT ultrasound images. For each ROI determined, as mentioned in Section 4.3.2, we apply the whole procedures starting from pre-processing up to the classification stage. The classification process is carried out in two steps: learning, and testing. During the learning stage, the classifier is trained on the extracted features to create a model, which is then used in the testing stage. In the training stage, the input vectors are named a 'training set'. After the learning stage is completed, the classifier is ready to be tested on new data that have not been used during the training phase. The efficiency of the classification stage results depends on two important matters. First, the effectiveness of the algorithms that were chosen in the previous stages before classification. All these stages have significant impacts on the classification algorithm results. Second, choosing the correct

classifier is a critical part of the pattern classification system, because different classification results may be obtained depending on the choice of certain classifier(s). A comparison between different classifiers' performance in various approaches regarding accuracy and AUC will be explored in this chapter.

6.2 Image Dataset

The research dataset is composed of 57 ultrasound AT images, where 14 are determined as abnormal according to the radiologist experts and 43 as normal. Images were taken by a portable Venue 40 musculoskeletal ultrasound system, as mentioned in Section 4.2. Images are taken on three AT positions separately. The first image is taken from the lower AT part starting from the calcaneal insertion point; the second image focuses on the middle part of the AT, while the third image covers the higher part of the AT upward to the calf muscles. In this research, we are focusing on the lower and middle parts of the AT.

6.3 Images Denoising

In this stage, the AT ultrasound image will be denoised using different image processing filters, and these filters will be applied on various AT image ROI sizes. The results from three different approaches will be explored. The effect of these filters will be evaluated using eight metrics as listed in Table 6.1.

6.3.1 Approach one- Using the whole AT area

In this approach, the whole AT area will be considered as the ROI, and the filters will be applied to the whole AT image. Table 6.1 shows the comparison of various average noise metrics for filtering the whole AT ultrasound images dataset. Most of the evaluation metrics indicate that SRAD is more efficient in the denoising stage than the other filters. The data indicate that, in general, SRAD has the best values against five metrics namely SNR, RMSE, MSE and CC of the whole eight evaluation metrics which are investigated in this study. In addition, SRAD had the best second values in both PSNR (82.31), where the first was the Lee filter (88.62) and SSIM (0.923), where the first was the median filter (0.935; Table 6.1).

The mean and median had good values against the majority of the metrics, particularly against PSNR, RMSE, AD, IQ and SSIM. The data shows that the Lee filter had the highest value; with slightly greater than 88 against the PSNR, as well as it was the weakest filter compared to the others with the lowest values for the seven-remaining metrics.

In addition, both Kuan and Wavelet filters were the weakest with low values against all filters, particularly against both PSNR and SNR. As well as some differences, which have been recorded at the remaining metrics. Furthermore, data shows a relatively normal distribution with values of standard deviation which is fairly close.

The original and filtered images are shown in Figure 6.1 where it is found that SRAD is outperforming in de-noising procedures without losing or changing the useful information such as AT edges. Higher SNR and lower MSE indicate that SRAD is more suitable for de-speckling AT images. Despite that, Mean, Median and Lee filters had a better overall IQ indicator than that of SRAD, it is clear that the CC and SSIM for SRAD are higher than the others.

Table 6.1: Mean and standard deviation (SD) of noise metrics for various filters used with whole AT ultrasound images

Metrics & Filters		PSNR	SNR	RMSE	MSE	AD	IQ	CC	SSIM
Mean	Mean	74.1935	70.23439	0.05206	0.002918	0.00340	0.905401	0.921407	0.90266
	SD	±2.74	±3.91	±0.014	±0.00	±0.00	±0.02	±0.04	±0.02
Median	Mean	74.8260	70.86694	0.05359	0.003475	0.00479	0.938414	0.903048	0.93485
	SD	±5.2	±6.25	±0.03	±0.00	±0.00	±0.02	±0.08	±0.02
Lee	Mean	88.6188	69.92624	0.06105	0.003947	0.00439	0.913066	0.887393	0.88875
	SD	±12.99	±3.09	±0.02	±0.00	±0.00	±0.02	±0.05	±0.02
Kuan	Mean	59.3874	55.42827	0.27982	0.081666	0.19161	0.041009	0.497293	0.04296
	SD	±1.87	±0.31	±0.06	±0.03	±0.03	±0.03	±0.15	±0.03
Wavelet	Mean	57.9853	54.02621	0.05206	0.002918	0.00340	0.905401	0.921407	0.52034
	SD	±1.94	±1.12	±0.07	±0.04	±0.07	±0.11	±0.09	±0.10
P & M	Mean	69.0702	65.1111	0.09095	0.008475	0.01111	0.514054	0.746222	0.65199
	SD	±1.45	±2.55	±0.015	±0.00	±0.00	±0.05	±0.10	±0.03
SRAD	Mean	82.3098	78.35076	0.02129	0.000593	0.00503	0.878737	0.992169	0.92266
	SD	±3.19	±2.76	±0.01	±0.00	±0.02	±0.03	±0.00	±0.01

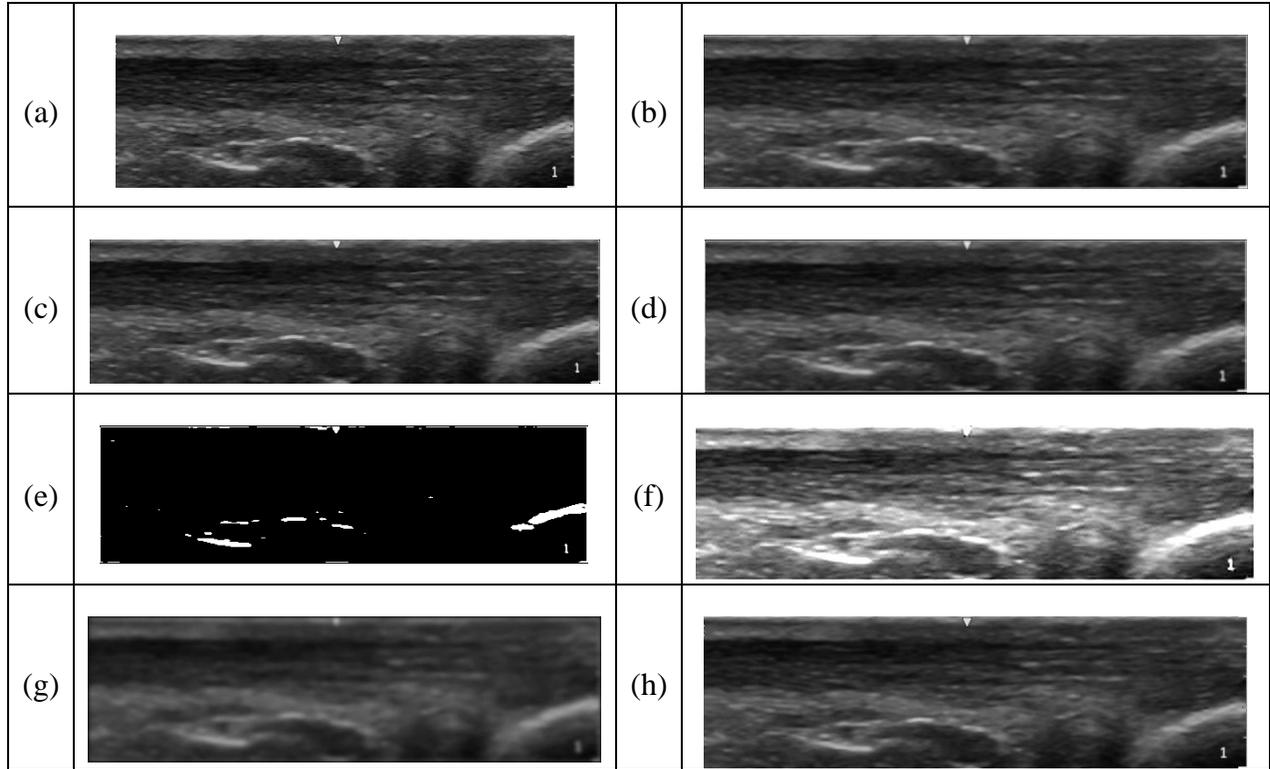


Figure 6.1 The result of applying filters on the whole AT ultrasound image where (a) is the original image and (b), (c), (d), (e), (f), (g) and (h) are the AT image after applying Mean, Median, Lee, Kuan, Wavelet, P & M and SRAD filters respectively.

6.3.2 Approach Two- Manual Determination of the AT ROI

In this approach, the ROI is determined manually by radiologist experts. The entire suspicious area for both lower and middle AT parts are cropped and saved as separate image files. Then, various de-noising filters are applied on all dataset images to reduce the impact of speckle noise on ultrasound images without deteriorating their valuable information. Table 6.2 presents the result of applying these filters and the evaluation metrics used to evaluate their effectiveness in the de-noising process. In this approach, Mean and Median filters have slightly better results in most evaluation metrics than SRAD and the other de-noising filters. Figure 6.2 shows the visual output after applying these filters.

Table 6.2: Mean and standard deviation (SD) of noise metrics for various filters used with manually determined ROI for AT ultrasound images

Metrics & Filters		PSNR	SNR	RMSE	MSE	AD	IQ	CC	SSIM
Mean	Mean	80.5972	70.40165	0.024255	0.000609	0.003043	0.901465	0.968658	0.895907
	SD	±1.73	±5.44	±0.00	±0.00	±0.00	±0.02	±0.01	±0.02
Median	Mean	85.49176	75.29621	0.013802	0.000199	0.000851	0.940883	0.988968	0.953228
	SD	±1.56	±6.14	±0.00	±0.00	±0.00	±0.02	±0.01	±0.01
Lee	Mean	80.89076	66.80266	0.036954	0.001443	0.004289	0.90017	0.926161	0.87781
	SD	±8.62	±5.55	±0.01	±0.00	±0.00	±0.02	±0.03	±0.02
Kuan	Mean	60.27218	50.07663	0.259943	0.073486	0.202871	0.027593	0.448528	0.026652
	SD	±2.89	±5.29	±0.08	±0.04	±0.05	±0.03	±0.23	±0.03
Wavelet	Mean	59.79756	49.60201	0.02455	0.000609	0.003043	0.901465	0.968658	0.558596
	SD	±2.98	±5.13	±0.09	±0.05	±0.09	±0.13	±0.07	±0.11
P & M	Mean	72.73636	62.54081	0.06027	0.00379	0.013358	0.475294	0.799089	0.631115
	SD	±1.97	±5.27	±0.01	±0.00	±0.00	±0.06	±0.08	±0.04
SRAD	Mean	70.63706	60.44151	0.087683	0.009891	-0.06233	0.843961	0.985481	0.873687
	SD	±5.17	±9.90	±0.05	±0.01	±0.06	±0.09	±0.01	±0.08

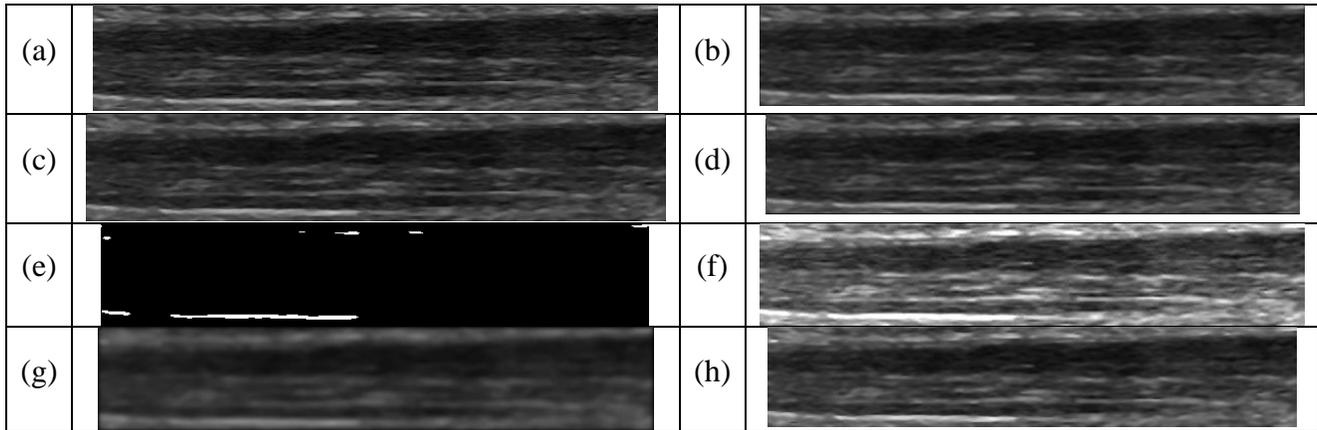


Figure 6.2: The result of applying filters on manually determined ROI for AT ultrasound image, where (a) is the original image and (b), (c), (d), (e), (f), (g) and (h) are the AT image after applying Mean, Median, Lee, Kuan, Wavelet, P & M and SRAD filters respectively.

The data indicates that there are some clear differences in the results obtained. Mean, Medium and Lee have the best values against PSNR, SNR, MSE and IQ, with similar results on the remaining evaluation metrics. Some differences have been recorded for PSNR, SNR and SSIM where, the Median has the best values at PSNR (85.49), SNR (75.30), IQ (0.94), CC (0.989) and SSIM (0.95).

Both SRAD and P&M de-noising filters have good and slightly similar values against most of the metrics, with some advantages for SRAD compared to P&M at IQ (0.84: 0.47), CC (0.99:0.80) and SSIM (0.87:0.63; Table 6.2). Furthermore, the de-noising Kuan filter has the lowest result against most metrics compared to the other filters, particularly PSNR, SNR and SSIM where, the values did not exceed 60.2, 50.1, 0.02, respectively (Table 6.2). This is followed by the Wavelet de-noising filter which had some low values at PSNR (59.8), SNR (49.6) and SSIM (0.55; Table 6.2). Overall in this approach, Mean and Median filters have slightly better results in most evaluation metrics than other filters, followed by SRAD which exceeded the results of many filters as shown in Figure 6.2.

6.3.3 Approach Three- Dividing the AT Image into Blocks

In this approach, the AT ultrasound is divided into six equal adjacent blocks starting from the calcaneal insertion up to the end of the middle part of the AT, as explained in Section 4.3.2.3. These blocks are considered as separate images and are saved in individual files. As shown in Table 6.3 from all the evaluation metrics used, SRAD gave better results than all the other different filters at many metrics. SRAD had the highest (65.67; Table 6.3) value at the SNR metric in comparison with all the other filters. SRAD also had the best value at RMSE (0.033), MSE (0.0015) and CC (0.99; Table 6.3).

In addition, against the PSNR metric, SRAD is the second-best filter (79.08), where the highest value was obtained by Lee filter by just 0.45. SRAD is also the second-best filter against the SSIM metric (0.92), where the highest value was the Median filter (0.93). Using the IQ indicator, the Mean, Median and Lee filters had the highest values in comparison with other filters, with values of 0.905817, 0.940198 and 0.908153, respectively. Similarly, these three filters along with the Wavelet filter had good values using the AD metric (0.0040, 0.0045, 0.0042, and -0.279, respectively). Figure 6.3 shows a sample image from the dataset before and after applying the various de-noising filters.

Table 6.3: Mean and standard deviation (SD) of noise metrics for various filters used with block divided ROI for AT ultrasound images

Metrics & Filters		PSNR	SNR	RMSE	MSE	AD	IQ	CC	SSIM
Mean	Mean	74.47221	61.05866	0.050552	0.002766	0.004024	0.905817	0.916514	0.89317
	SD	±0.80	±3.96	±0.01	±0.00	±0.00	±0.02	±0.05	±0.01
Median	Mean	75.74957	62.33603	0.049136	0.002995	0.004545	0.940198	0.904608	0.933946
	SD	±5.43	±6.56	±0.02	±0.00	±0.00	±0.02	±0.90	±0.02
Lee	Mean	79.536	59.8986	0.061347	0.003998	0.004299	0.908153	0.875366	0.882746
	SD	±11.76	±3.36	±0.02	±0.00	±0.00	±0.02	±0.06	±0.02
Kuan	Mean	59.22872	45.81517	0.286362	0.086136	0.19847	0.041937	0.469757	0.034616
	SD	±2.08	±0.34	±0.06	±0.04	±0.03	±0.04	±0.19	±0.02
Wavelet	Mean	57.91876	44.50522	0.333247	0.116553	-0.27998	0.440144	0.847681	0.494896
	SD	±2.14	±1.17	±0.07	±0.05	±0.08	±0.12	±0.10	0.1
P & M	Mean	68.97258	55.55903	0.092228	0.008766	0.014167	0.505921	0.705058	0.625481
	SD	±1.58	±2.56	±0.02	±0.00	±0.00	±0.05	±0.11	±0.03
SRAD	Mean	79.08382	65.67028	0.033133	0.001541	0.010306	0.88155	0.990442	0.917612
	SD	±4.63	±3.70	±0.02	±0.00	±0.03	±0.03	±0.00	±0.01

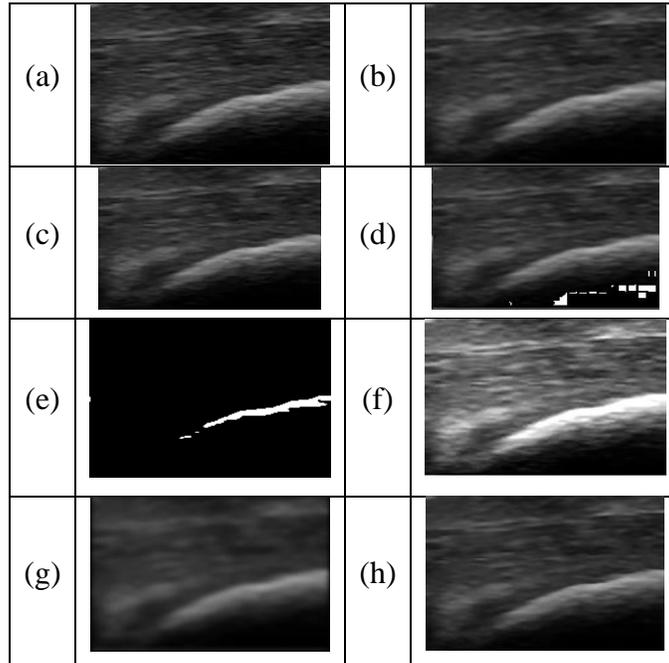


Figure 6.3: The result of applying the filters on divided blocks represent ROI for AT ultrasound image, where (a) is the original image and (b), (c), (d), (e), (f), (g) and (h) are the AT image after applying Mean, Median, Lee, Kuan, Wavelet, P & M and SRAD filters respectively.

6.3.4 Image Denoising Results Discussion

Denoising ultrasound images is a major step in this study in order to mitigate the effect of speckle noise. In the process of choosing the most convenient filter in this study, various denoising filters that are used in medical image processing have been tested. A comparison between these filters is performed to decide which filter is more appropriate than the others for the three approaches that have been used and experimented with in the study (Dangeti, 2003). From the results obtained and shown in Tables 6.1, 6.2 and 6.3, this study has concluded that some filters such as (SRAD, Mean and Median) are working better than the others according to the measuring metrics that have been applied.

Median and SRAD as nonlinear filters are performing well in all approaches; this is supported by (Thivakaran & Chandrasekaran, 2010) where they identified that Median could overcome the effect of impulsive noise while preserving edge information. This is also supported by (Yu & Acton, 2002) who indicated that the SRAD filter is based on anisotropic diffusion to suppress speckle while maintaining edge information. These findings are highly compatible with the main results of this study and indicated that the effect of SRAD filtering was more obvious on full AT images and divided blocks than manually determining ROI.

Moreover, it was evident that all quality metrics presented here were equally important for image quality evaluation. Specifically, five out of the eight quality metrics (RMSE, MSE, AD, IQ and SSIM) had better measures obtained with the median filter with manually determined ROI, where only one of the eight quality metrics (SSIM) had better measures obtained for a median filter with the whole AT determination for ROI.

On the other hand, three of the eight quality metrics (PSNR, SNR and CC) have the best results for SRAD, with the full determination of the ROI. Furthermore, it is important to note that a higher PSNR (or equivalently, a lower RMSE) does not necessarily imply a higher subjective image quality, although they do provide some measure of relative quality.

Figure 6.3 shows that the divided images were visually better, although the quality metrics in Table 6.2 were rated better for the manually determined ROI of the images. Noise measures may be therefore

misused when evaluating image quality and when the image is corrupted by degradation other than the additive noise.

6.4 Image Enhancing

In the conducted three experiments, the CLAHE enhancing algorithm was used after the image denoising process. CLAHE was originally developed for medical imaging and has proven to be successful for the enhancement of low contrast images such as portal films (Jeong et al., 2011). The CLAHE algorithm partitions the images into contextual regions and applies the histogram equalization to each one. It works on small regions in the image called tiles rather than the entire image. This way, the contrast in the homogeneous region can be limited to its minimum, and it also limits the appearance of artefacts and noises.

6.4.1 Approach one- the whole AT area

As shown in Table 6.4 SRAD had the highest (63.49) value at the PSNR metric in comparison with all the other filters. SRAD also had the best value at the SNR metric with a value of 59.53. In addition, the evaluation metrics RMSE and MSE showed the best values with the SRAD filter, with 0.171234 and 0.029569, respectively (Table 6.4). At the AD evaluation metric, the best value (-0.114) was obtained by the P&M filter, followed by the SRAD filter which had the second-best value (-0.122). The Kuan and Wavelet filters had overall weak values with the majority of the evaluation metrics, where the Kuan filter had the weakest value of 0.12 at the SSIM evaluation metric and the weakest value of 0.49 at the CC metric (Table 6.4). The Mean, Median and Lee filters had overall good values, but generally, they did not exceed the SRAD filter values. Figure 6.4 shows the original and filtered images using CLAHE enhancing algorithm using various filters.

Table 6.4: Mean and standard deviation of noise metrics for various filters used with whole AT ultrasound images

Metrics & Filters		PSNR	SNR	RMSE	MSE	AD	IQ	CC	SSIM
Mean	Mean	62.9319	58.9727	0.18270	0.03365	-0.1278	0.584715	0.825099	0.618105
	SD	±0.80	2.79	0.02	±0.01	±0.03	±0.06	±0.03	±0.05
Median	Mean	62.8356	58.8765	0.18488	0.03451	-0.1268	0.56820	0.80282	0.60916
	SD	±0.87	±2.87	±0.02	±0.01	±0.03	±0.07	0.06	0.06
Lee	Mean	62.7846	58.8255	0.18571	0.03473	-0.1253	0.58345	0.78896	0.60924
	SD	±0.74	±2.72	±0.02	±0.01	±0.03	±0.06	±0.04	±0.05
Kuan	Mean	59.7559	55.7968	0.26871	0.07557	0.1783	0.04176	0.49729	0.12426
	SD	±1.95	±0.25	±0.06	±0.03	±0.03	±0.03	±0.15	±0.04
Wavelet	Mean	57.7435	53.7843	0.33113	0.10997	-0.2669	0.27616	0.73051	0.30261
	SD	±0.47	±2.46	±0.02	±0.01	±0.02	±0.08	±0.06	±0.07
P & M	Mean	62.8528	58.8937	0.18401	0.03401	-0.1143	0.514568	0.671535	0.55649
	SD	±0.58	±2.47	±0.01	±0.00	±0.02	±0.06	±0.08	±0.04
SRAD	Mean	63.4956	59.5364	0.17123	0.02956	-0.1225	0.613118	0.876039	0.679943
	SD	±0.80	±2.73	±0.02	±0.01	±0.03	±0.06	±0.02	±0.05

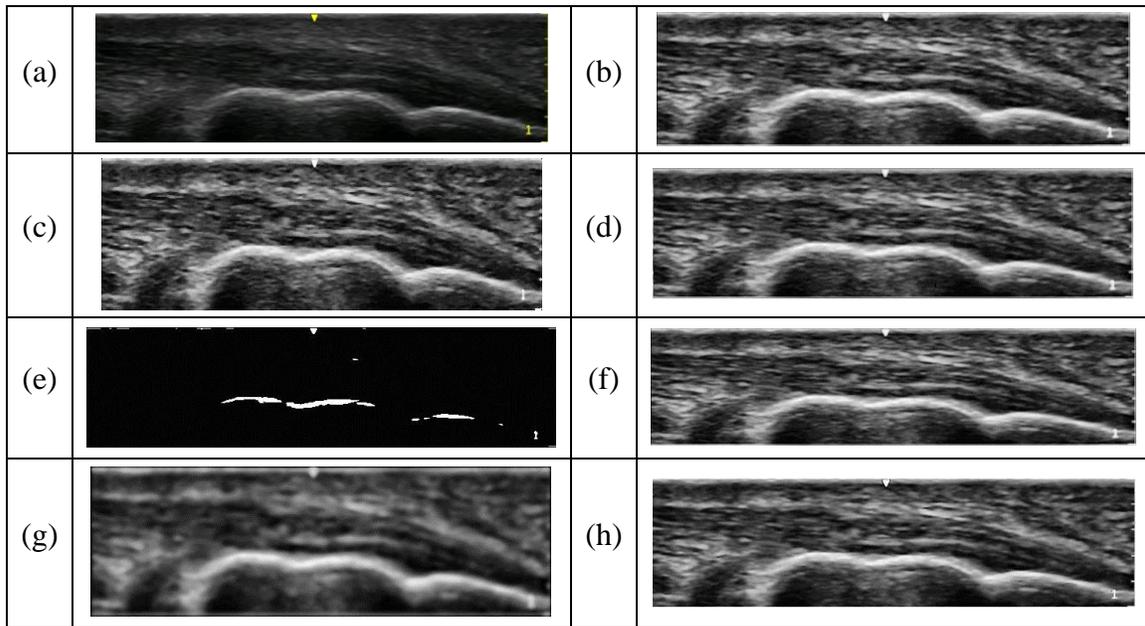


Figure 6.4: The result of applying CLAHE enhance algorithm filters on the whole AT ultrasound image, where (a) is the original image and (b), (c), (d), (e), (f), (g) and (h) are the AT image after applying Mean, Median, Lee, Kuan, Wavelet, P & M and SRAD filters respectively

6.4.2 Approach Two- Manual Determination for the AT ROI

The P&M filter overall had the best values at many of the evaluation metrics as shown in Table 6.5. P & M had the highest (63.95) value at both, the PSNR metric and the SNR metric (53.76). Furthermore, the evaluation metrics RMSE and MSE showed the best values with the P&M filter, with 0.162 and 0.026, respectively. Also, the AD evaluation metric had a good value (-0.115) at the P&M filter (Table 6.5). The Mean, Median and Lee had good values at PSNR 63.4, 63.77 and 63.71335, respectively and SNR 53.27, 53.57 and 53.51, respectively. These three filters also had good values at the IQ, CC and SSIM metrics.

Table 6.5: Mean and standard deviation (SD) of noise metrics for various filters used with the Manual determined ROI for AT ultrasound images

Metrics & Filters		PSNR	SNR	RMSE	MSE	AD	IQ	CC	SSIM
Mean	Mean	63.4690	53.2735	0.1722	0.0300	-0.1319	0.5585	0.8609	0.5781
	SD	±1.10	±6.54	±0.02	±0.01	±0.03	±0.09	±0.04	±0.07
Median	Mean	63.7709	53.5754	0.1665	0.02816	-0.1293	0.5481	0.8835	0.5933
	SD	±1.12	±6.53	±0.02	±0.01	±0.03	±0.08	±0.03	±0.08
Lee	Mean	63.7133	53.5178	0.1673	0.0283	-0.1213	0.5622	0.8254	0.5820
	SD	±0.97	±6.47	±0.02	±0.01	±0.03	±0.08	±0.05	±0.07
Kuan	Mean	60.769	50.5734	0.24696	0.066	0.1883	0.02872	0.3955	0.1143
	SD	±3.10	±5.24	±0.08	±0.04	±0.05	±0.03	±0.26	±0.06
Wavelet	Mean	58.065	47.8694	0.3191	0.10212	-0.2650	0.2658	0.7500	0.2947
	SD	±0.47	±6.10	±0.02	±0.01	±0.02	±0.10	±0.10	±0.10
P & M	Mean	63.959	53.7640	0.1624	0.026	-0.1151	0.5016	0.7425	0.54305
	SD	±0.90	±6.43	±0.02	±0.01	±0.02	±0.07	±0.08	±0.05
SRAD	Mean	61.358	51.1621	0.2252	0.0540	-0.1769	0.4618	0.85643	0.51525
	SD	±2.22	±7.67	±0.06	±0.03	±0.07	±0.13	±0.04	±0.13

Following the results obtained for the Mean, Median and Lee filters, the SRAD filter scored at the IQ (0.46), CC (0.85) and SSIM (0.51; Table 6.5) metrics. The P&M filter also had good values similar to SRAD against many of the metrics. The Kuan and Wavelet filters had overall weak values at the majority of the evaluation metrics as shown in Figure 6.5.

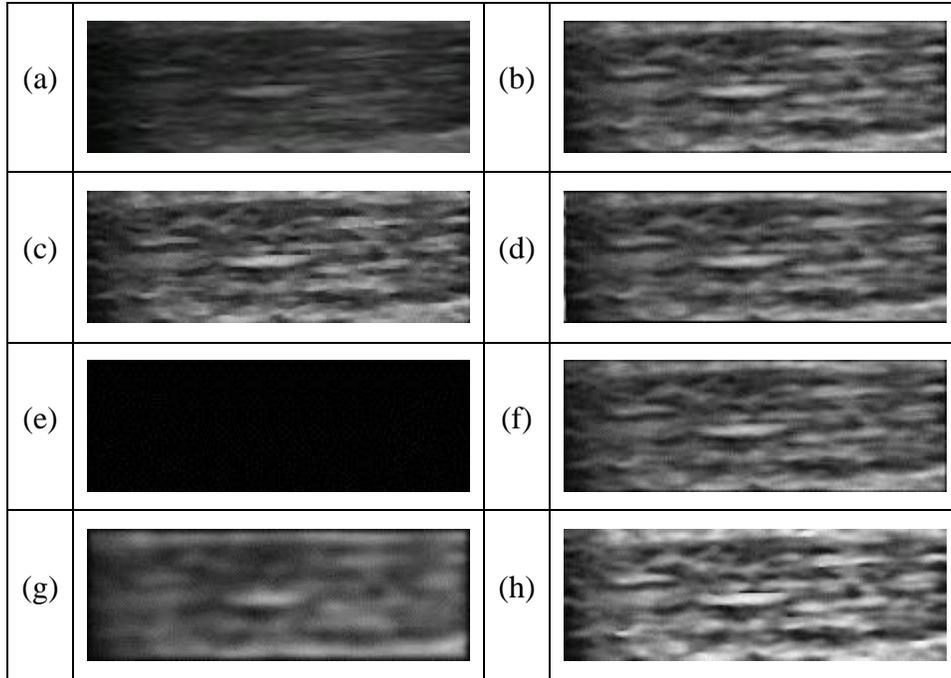


Figure 6.5: The result of applying CLAHE enhance algorithm on manually determined ROI for AT ultrasound image, where (a) is the original image and (b), (c), (d), (e), (f), (g) and (h) are the AT image after applying Mean, Median, Lee, Kuan, Wavelet, P & M and SRAD filters respectively.

6.4.3 Approach Three- Dividing the AT Image into Blocks

The filter with the best values of the majority of the evaluation metrics is the SRAD filter; at five of the eight evaluation metrics, SRAD had the best values. SRAD had the highest (67.65) value at the SNR evaluation metric in comparison with all the other filters. SRAD also had the best value at both of the RMSE and the MSE metrics with values of 0.030 and 0.00096, respectively.

In addition, the evaluation metrics CC and SSIM showed the best values for the SRAD filter, with 0.99 and 0.67 respectively (Table 6.6). The values of the median filter exceeded all the other filters against the IQ evaluation metric with 0.98. The median filter had good values at many other filters such as AD with 0.0035. However, its values are not better than the values of the SRAD filter against the majority of the evaluation metrics (Table 6.6). The Kuan and Wavelet filters had overall weak values with the majority of the evaluation metrics. Figure 6.6 shows a sample image from the dataset before and after applying the CLAHE enhancing algorithm with various denoising filters.

Table 6.6: Mean and standard deviation (SD) of noise metrics for various filters used with block divided ROI for AT ultrasound images

Metrics & Filters		PSNR	SNR	RMSE	MSE	AD	IQ	CC	SSIM
Mean	Mean	73.9051	64.0086	0.0524	0.0028	0.0048	0.9691	0.9678	0.6104
	SD	±1.67	±2.36	±0.01	±0.00	±0.00	±0.01	±0.02	±0.05
Median	Mean	76.8412	66.9447	0.0410	0.0020	0.0036	0.9852	0.9754	0.6154
	SD	±4.27	±4.91	±0.02	±0.00	±0.00	±0.00	±0.00	±0.06
Lee	Mean	78.6214	62.0631	0.0677	0.0047	0.0050	0.9677	0.9470	0.6054
	SD	±12.03	±1.83	±0.01	±0.00	±0.00	±0.01	±0.02	±0.06
Kuan	Mean	57.8281	47.9316	0.3279	0.1077	0.1135	0.2071	0.7561	0.1129
	SD	±0.44	±0.90	±0.02	±0.01	±0.04	±0.07	±0.06	±0.04
Wavelet	Mean	54.7422	44.8457	0.4708	0.2250	-0.3960	0.4952	0.8804	0.2864
	SD	±1.11	±0.51	±0.06	±0.05	±0.05	±0.09	±0.07	±0.08
P & M	Mean	66.4753	64.2733	0.1216	0.0149	0.0184	0.6571	0.8217	0.5348
	SD	±0.85	±0.73	±0.01	±0.00	±0.00	±0.04	±0.05	±0.04
SRAD	Mean	78.4512	67.6532	0.0308	0.0010	0.0058	0.9514	0.9908	0.6712
	SD	±1.22	±1.10	±0.00	±0.00	±0.00	±0.01	±0.00	±0.05

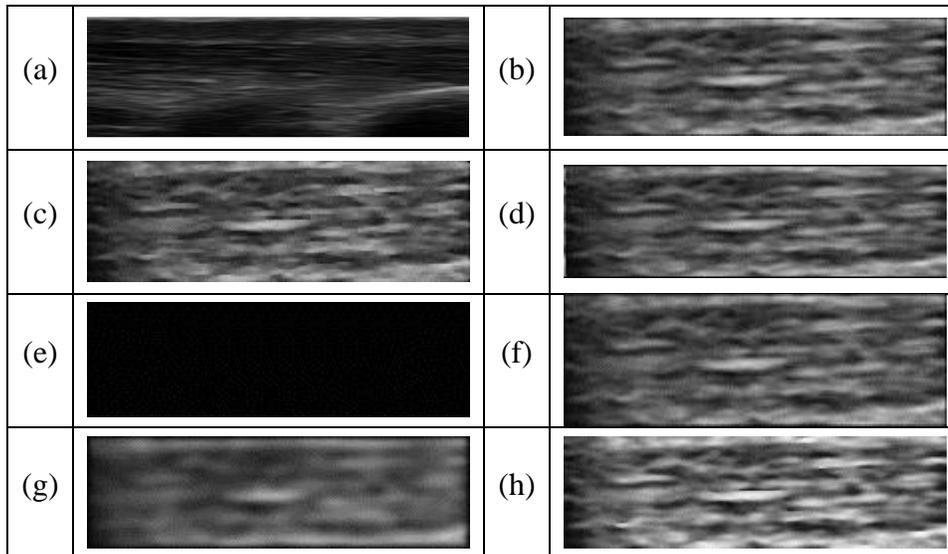


Figure 6.6: The result of applying the filters on divided blocks represent ROI for AT ultrasound image, where (a) is the original image and (b), (c), (d), (e), (f), (g) and (h) are the AT image after applying Mean, Median, Lee, Kuan, Wavelet, P & M and SRAD filters respectively.

6.4.4 Image Enhancing Results Discussion

In ultrasound, medical image classification process, image enhancing is an essential part. Using the appropriate enhancing method affects the accuracy of the classification results. In addition, the precision of the measured texture parameters may be influenced by noises, and it leads to mistakes in diagnostic reports. It was observed that changing the ROI size affects the image enhancing results. (Basset & Mestas, 1993), in their study about the influence of the ROI size on ultrasonic images of the prostate using the mean/average of the co-occurrence matrices, have indicated that the size of ROI affects the discrimination results. Moreover, performance can be improved by creating a new sub-image which consists entirely of the ROI at the region of interest. A comparison between the three approaches shows that auto dividing images into blocks achieve better performance. Such a method would reduce computational overhead of the ROI. Not only that, but if the sub-image is small, then the algorithm does not waste time looking for ROI and extracts all the interesting features (Won, Park, & Park, 2002).

Figures 6.4, 6.5 and 6.6 show that the boundary of the AT edges after performance enhancing through the CLAHE algorithm is very smooth and the details are well preserved. Most of the metrics shown in Tables 6.4, 6.5 and 6.6 indicate an improvement in the targeted images, especially when using the SRAD filter. In order to evaluate the performance of the enhanced images, a comparison between the three approaches is presented in Table 6.7.

Table 6.7: A comparison between AT enhanced images for the three approaches

App.	PSNR	SNR	RMSE	MSE	AD	IQ	CC	SSIM
Full	SRAD	SRAD	SRAD	SRAD	P & M	SRAD	SRAD	SRAD
	63.50	59.54	0.171	0.030	-0.114	0.613	0.876	0.680
Man.	Lee	P & M	P & M	P & M	SRAD	Lee	Median	Median
	63.71	53.76	0.162	0.026	-0.177	0.562	0.884	0.593
Auto	Lee	SRAD	SRAD	SRAD	Median	Median	SRAD	SRAD
	78.62	67.65	0.031	0.001	0.004	0.985	0.991	0.671

Ultrasound image enhancing proves that dividing the whole image into sub-images could improve the pre-processing stage in both denoising and enhancing steps (Gottumukkal & Asari, 2004). Small ROI affects the enhancing results by avoiding the unrelated area in the images which is an easier and more

accurate way (Won et al., 2002). These enhanced images can be processed further to detect the abnormality with high accuracy. Such enhancement is beneficial for diagnosis.

6.5 Images Classification

In classification analysis, this study evaluates the classifiers using the confusion matrix. Tables 6.8 - 6.16 show the results of eight classifiers that have been applied in this study; the columns represent the classifier's predictions, and the rows are the actual classifiers metrics values.

- TP (True Positive) is the number of positive cases correctly classified as such.
- FN (False Negative) is the number of positive cases incorrectly classified as negatives.
- FP (False Positive) is the number of negative cases that are incorrectly identified as positive cases
- TN (True Negative) is the number of negative cases correctly classified as such
- Sensitivity is found by $TP/(TP+FN)$
- Specificity is found by $TP/(TN+FP)$

6.5.1 Approach one- the whole AT area

In this approach, we consider the whole AT area as our ROI, and all features that are mentioned in Section 4.3.3 were extracted. After gathering all these features in one matrix, these features are reduced using KPCA method in order to project the samples on an axis of two dimensions. Various classifiers are tested in order to estimate the whole process model.

6.5.1.1 Model performance metrics for the first approach

Based on the most important classifiers that are listed in Table 6.8 and applied in this study to classify AT ultrasound images through the extracted and reduced features. For the TP, it is clear that, the quadratic discriminant had the highest value (82 out of 86), with the lowest value of FP and it had the lowest percentage of sensitivity (slightly less than 77%) in comparison with the sensitivities of the other classifiers, which ranged from about 80% (quadratic SVM) to slightly less than 89% (ensemble

RUSboost). Both classifiers of the ensemble (bagged tree) and ensemble (subspace KNN) had the same relatively high value (80) of TP and the same value of FP, but slightly different values of sensitivity.

Table 6.8: Some Classifiers Evaluation (Sensitivity & Specificity) for the first approach

Classifiers	TP (86)	TN (28)	FP	FN	Sensitivity	Specificity
DT	78	16	8	12	86.7%	66.7%
Quadratic Discriminant	82	3	4	25	76.6%	42.9%
Quadratic SVM	60	12	26	16	78.9%	31.6%
KNN	78	15	8	13	85.7%	65.2%
Ensemble (Boosted Trees)	78	16	8	12	86.7%	66.7%
Ensemble (Bagged tree)	80	15	6	13	86.0%	71.4%
Ensemble (Subspace KNN)	80	16	6	12	87.0%	72.7%
Ensemble (RUSBoost Trees)	71	19	15	9	88.8%	55.9%

On the other hand, for the true negative (TN) it can be concluded that the data is relatively close, except the second classifier (quadratic discriminant), where only 3 out of 28 were recorded as TN, and the remaining 25 were recorded as FN (Table 6.8).

In general, the ensemble (RUSBoost) classifier had the best values for sensitivity and the negatives (true and false). Meanwhile, the quadratic discriminant classifier had the worst recorded values in the same model performance metrics (sensitivity and the negatives (true and false), this is followed by the quadratic SVM classifier, which had low values at specificity and the positives (true and false).

6.5.1.2 Accuracy, precision, ROC/AUS and the percentages of error rate for the first approach

The results from the first approach for the model performance metrics are presented in Table 6.9, where the accuracy, precision, error rates and ROC/AUC are shown for eight different classifiers. Four of the classifiers are ensemble (Ensemble (Boosted Trees), Ensemble (Bagged tree), Ensemble (Subspace KNN) and Ensemble (RUSBoost Trees)); the four generally had good values for accuracy (82.46, 83.34, 83.88 and 78.76, correspondingly). They all so had general good values of precision, with 0.91, 0.93, 0.93 and 0.83, respectively. Ensemble (Subspace KNN) and Ensemble (Bagged tree) had the smallest error rates, with values of 16.12% and 16.66%, respectively. With ROC/AUC, the highest four

values came from the four ensemble classifiers with values of 0.82, 0.83, 0.78 and 0.85 (Table 6.9). The DT and KNN classifiers had good values for accuracy, precision, error rates and ROC/AUC but they did not exceed those of the ensemble classifiers in the majority of the results. The highest error rate and the lowest percentage of accuracy were recorded at the Quadratic SVM classifier.

Table 6.9: Standard classifiers performance evaluation metrics for the first approach

Classifiers	Accuracy %	Precision	ROC/AUC	Error Rate %
DT	82.46	0.91	0.77	17.54
Quadratic Discriminant	74.38	0.95	0.55	25.62
Quadratic SVM	62.80	0.70	0.71	37.20
KNN	81.42	0.91	0.73	18.58
Ensemble (Boosted Trees)	82.46	0.91	0.82	17.54
Ensemble (Bagged tree)	83.34	0.93	0.83	16.66
Ensemble (Subspace KNN)	83.88	0.93	0.78	16.12
Ensemble (RUSBoost Trees)	78.76	0.83	0.85	21.24

6.5.1.3 Model performance measures using the first approach

The results from the first approach for the model performance measures are presented in Table 6.10. Where, Kappa, Youden’s index, G-Mean, MCC, Balanced Accuracy and F-Measure values are recorded for the eight different classifiers. In general, the F-measure had the highest values against all the classifiers, where, the values ranged between 0.74 (Quadratic SVM) and 0.90 (Ensemble subspace KNN). All the values within this range were close and were between 0.85 (for the Quadratic Discriminant classifier) and 0.89 (for three different classifiers) (Table 6.10).

The values of F-measure are followed by the values of the G-Mean and Balanced accuracy, which were relatively high as they ranged from 0.50 (Quadratic SVM) to 0.80 (Ensemble suspense KNN). The lowest value (0.50) was recorded at the Quadratic SVM classifier, which was also the classifier with the lowest value against the F-measure (0.74). In addition, the ensemble classifiers

(Boosted Trees, Bagged tree, Subspace KNN and RUSBoost Trees) showed generally good values, in comparison with the other classifiers. Comparisons between different classifiers' performance in this approach in terms of accuracy, AUC and various classification evaluation metrics are shown graphically in Figure 6.7.

Table 6.10: More classifiers performance evaluation measures using the first approach

Classifiers	Kappa	Youden's index	G- Mean	MCC	Balanced Accuracy	F- Measure
DT	0.49	0.54	0.76	0.51	0.77	0.89
Quadratic Discriminant	0.26	0.34	0.58	0.11	0.60	0.85
Quadratic SVM	0.12	0.11	0.50	0.12	0.56	0.74
KNN	0.46	0.51	0.75	0.47	0.76	0.88
Ensemble (Boosted Trees)	0.49	0.54	0.76	0.47	0.77	0.89
Ensemble (Bagged tree)	0.50	0.57	0.78	0.52	0.79	0.89
Ensemble (Subspace KNN)	0.53	0.60	0.80	0.55	0.80	0.90
Ensemble (RUSBoost Trees)	0.48	0.45	0.71	0.47	0.73	0.86

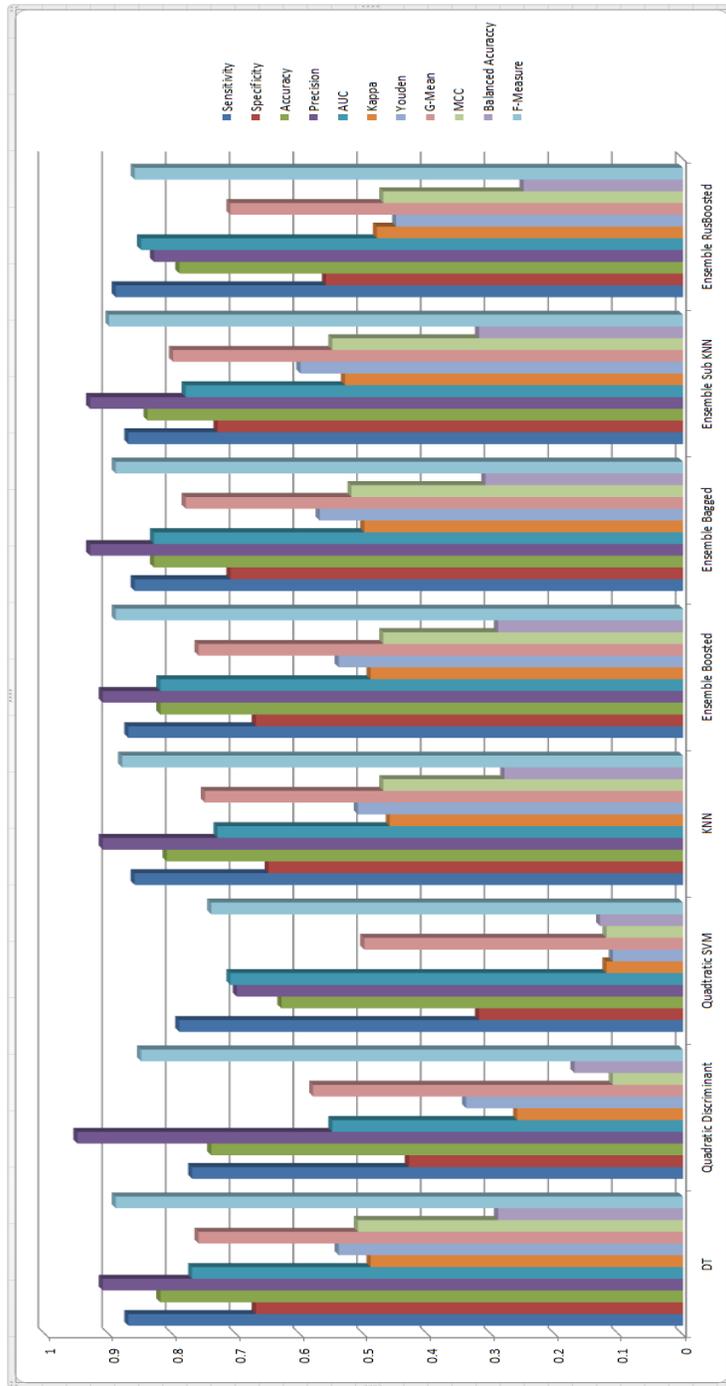


Figure 6.7: Graph showing the classification evaluation metrics for different classifiers using the whole AT approach

6.5.2 Approach Two- Manual Determination for AT ROI

In order to enhance the different classifiers’ results and to concentrate more on the ROI, in this approach we are determining the ROI manually by an expert from the most vulnerable area in the AT image. According to medical reports, this area is responsible for 80% to 90% of the AT abnormality. Therefore, by focusing more on this area a lot of speckle noise deteriorating the other image regions can be avoided. The window is located approximately 2 cm away from the calcaneal insertion point, with 1 cm width and 0.5 high. Features are extracted and reduced as mentioned in Sections 4.3.2.2.

6.5.2.1 Some model performance metrics for the second approach

The results from the model performance metrics for the second approach is summarised in Table 6.11. In general, the data (Table 6.11) shows some different and ab-normal results in comparison with the results of the first approach. For the TP, data shows that the ensemble (subspace KNN) had the highest value (80 of 86), with the lowest value of FP (6 of 86).

Table 6.11: Some Classifiers Evaluation (Sensitivity & Specificity) for the second approach

Classifiers	TP (86)	TN (28)	FP	FN	Sensitivity	Specificity
DT	75	14	11	14	84.3%	56.0%
Quadratic Discriminant	76	5	10	23	76.8%	33.3%
Quadratic SVM	50	12	36	16	75.8%	25.0%
KNN	74	9	12	19	79.6%	42.9%
Ensemble (Boosted Trees)	76	15	10	13	85.4%	60.0%
Ensemble (Bagged tree)	76	13	10	15	83.5%	56.5%
Ensemble (Subspace KNN)	80	10	6	18	81.6%	62.5%
Ensemble (RUSBoost Trees)	72	17	14	11	86.7%	54.8%

It also had a relatively high percentage of sensitivity (slightly less than 82%) in comparison with the sensitivities of the other classifiers which ranged from about 76% (quadratic SVM) to slightly less than 87% (ensemble RUSBoost), which are similar results to the first approach. Both ensemble bagged

tree and ensemble boosted trees classifiers, as well as a quadratic discriminant, had relatively high values (76) for TP, with relatively different values for sensitivity.

On the other hand, for the true negative (TN) it can be concluded that the data is relatively close for six of the eight classifiers, except for the quadratic discriminant and the KNN classifiers, where only 5 and 9 of 28 have been recorded, respectively (Table 6.11). For (FN), the data shows that the highest value (23) is found at the classifier of the quadratic discriminant, it was followed by 19 and 18 which were recorded for KNN and ensemble (subspace KNN) respectively.

In general, the ensemble (RUSBoost) classifier had the best values at three of the six model performances; sensitivity and the false negative as well as the true negative. It was followed by the ensemble (subspace KNN) which had good values at TP, FP and specificity. Meanwhile, the quadratic SVM classifier had the worst recorded values in the model performance metrics of TP, FP and sensitivity as well as in specificity (Table 6.11).

6.5.2.2 Accuracy, precision, ROC/AUS and the percentages of error rate for the second approach

The results from the second approach for the model performance metrics are presented in Table 6.12, where the accuracy, precision, error rates and ROC/AUC for eight different classifiers are recorded. The four ensemble classifiers (Ensemble (Boosted Trees), Ensemble (Bagged tree), Ensemble (Subspace KNN) and Ensemble (RUSBoost Trees)); had the best values of accuracy (80.18, 78.26, 79.10 and 78.08, respectively). In addition, they all had the best error rates comparatively, with 19.82, 21.74, 20.90 and 21.92, respectively.

Table 6.12: Standard classifiers performance evaluation metrics for the second approach

Classifiers	Accuracy %	Precision	ROC/AUC	Error Rate %
DT	77.54	0.87	0.75	22.46
Quadratic Discriminant	71.76	0.88	0.56	28.24
Quadratic SVM	53.86	0.58	0.50	46.14
KNN	70.54	0.86	0.55	29.46
Ensemble (Boosted Trees)	80.18	0.88	0.71	19.82
Ensemble (Bagged tree)	78.26	0.88	0.77	21.74
Ensemble (Subspace KNN)	79.10	0.93	0.64	20.90
Ensemble (RUSBoost Trees)	78.08	0.84	0.77	21.92

Furthermore, they generally had good values of precision and ROC/AUC. The classifiers DT and Quadratic Discriminant had relatively good values of accuracy, precision, error rates and ROC/AUC but they did not exceed those of the ensemble classifiers in most of the model performance metrics. The Quadratic SVM classifier had the highest (46.14%) percentage of error rate and had the lowest values for accuracy, precision and ROC/AUC. The KNN classifier generally had good values for precision and accuracy. However, it had one of the highest error rates (29.46%).

6.5.2.3 Model performance measures using the second approach

The results from the second approach for the model performance measures are presented in Table 6.13. Where, Kappa, Youden’s index, G-Mean, MCC, Balanced Accuracy and F-Measure values are recorded for the eight different classifiers which are listed in Table 6.13. Similarly, to the first approach, the F-measure generally had the highest values against all the classifiers, where, these values were relatively close and ranged between 0.82 (Quadratic Discriminant) and 0.88 (Ensemble (Boosted Trees)), except the quadratic SVM which had the lowest F-measure value of 0.66.

Table 6.13: More classifiers performance evaluation measures using the second approach

Classifiers	Kappa	Youden’s index	G-Mean	MCC	Balanced Accuracy	F-Measure
DT	0.39	0.40	0.69	0.39	0.70	0.83
Quadratic Discriminant	0.10	0.10	0.50	0.08	0.55	0.82
Quadratic SVM	0.0	0.01	0.44	0.01	0.51	0.66
KNN	0.15	0.23	0.59	0.26	0.62	0.83
Ensemble (Boosted Trees)	0.47	0.45	0.71	0.44	0.72	0.88
Ensemble (Bagged Tree)	0.37	0.41	0.69	0.37	0.71	0.86
Ensemble (Subspace KNN)	0.34	0.45	0.72	0.36	0.73	0.87
Ensemble (RUSBoost Trees)	0.42	0.42	0.69	0.43	0.71	0.86

This was followed by the values of the Balanced accuracy and G-Mean, where, they were relatively high in comparison to other performance measures as they ranged from 0.44 (quadratic SVM) to 0.73 (ensemble (subspace KNN)). In addition, the ensemble classifiers (Boosted Trees, and Subspace KNN and RUSBoost Trees) showed good values in general, in comparison with the other classifiers. Comparisons between different classifiers’ performance in this approach in terms of accuracy, AUC and various classification evaluation metrics are shown graphically in Figure 6.8.

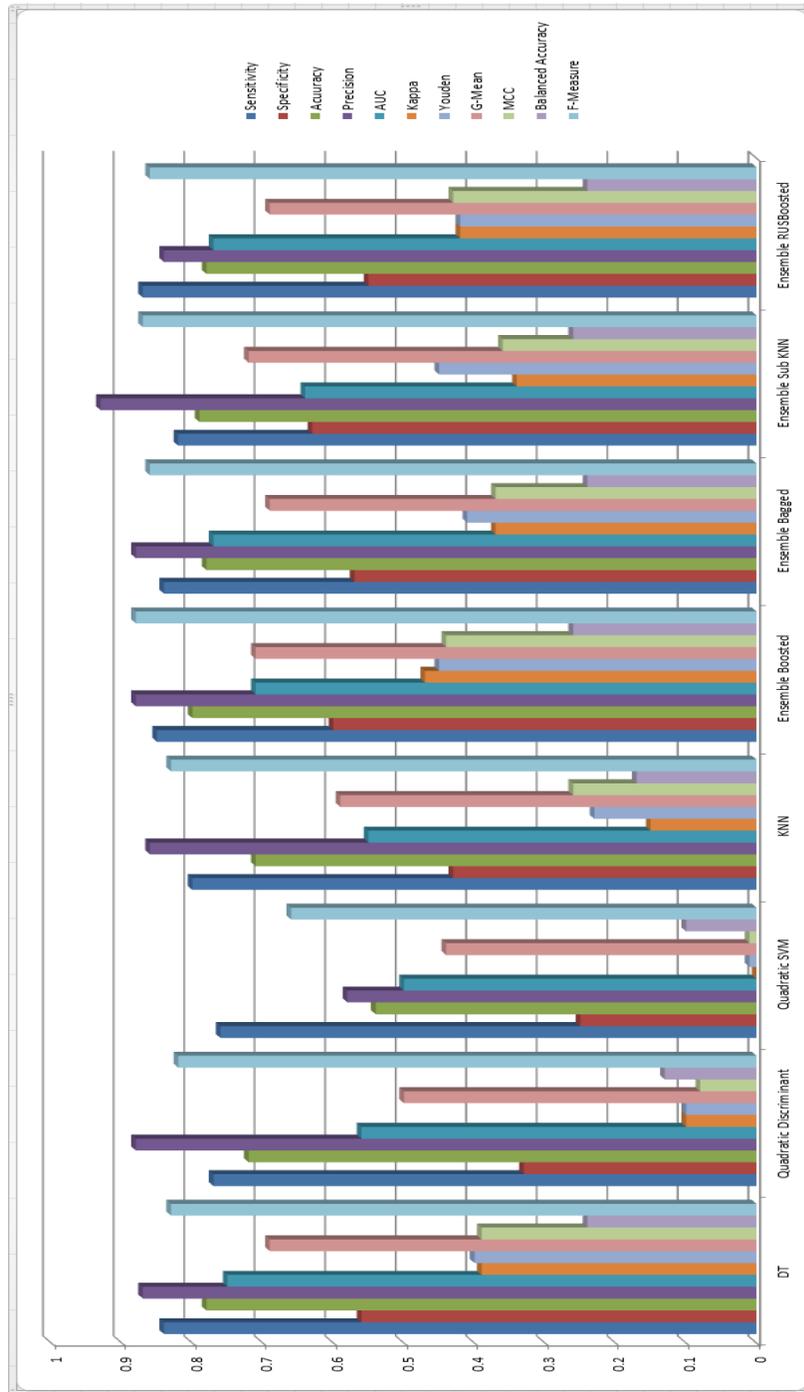


Figure 6.8: Graph showing the classification evaluation metrics for different classifiers using the manual determining for AT ROI approach

6.5.3 Approach Three- Dividing the AT Image into Blocks

In this approach, our goal is to achieve one of the research objectives, by preventing the user intervention in determining the ROI area, automating the model by dividing the AT image into six equal non-overlapped blocks. These blocks are equal in their size and in adjacent positions, where the first block starts from the calcaneal insertion point with approximately 2 cm width and 1cm height. Then, the next block started at the point where the first block ends with the same width and height, while the remaining four blocks started after the second block. This means that the third block started at 4 cm away from the calcaneal insertion point as explained in Section 4.3.2.2. To estimate the efficiency of this approach, features are extracted from the third block, which represents the ROI. These features are reduced using nonlinear reduction method KPCA, and then the returned predictors are considered as input for the classifiers.

6.5.3.1 Model performance metrics for the third approach

For the TP, it is clear that data is relatively similar; it ranges from 126 to 147 out of a total of 159. Both quadratic classifiers (SVM and discriminated) had the highest values of TP (146 and 147 out of the total of 159). Meanwhile, these classifiers had the lowest values of FP (13 and 12, respectively), in addition, they had the lowest percentages of sensitivity (73% and 75, respectively) in comparison with the sensitivities of the other classifiers which ranged from 73% (Quadratic Discriminant) to slightly less than 89% (ensemble RUSBoost).

Table 6.14: Some Classifiers Evaluation (Sensitivity & Specificity) for the third approach

Classifiers	TP (159)	TN (56)	FP	FN	Sensitivity	Specificity
DT	140	29	19	27	83.8%	60.4%
Quadratic Discriminant	146	2	13	54	73.0%	13.3%
Quadratic SVM	147	7	12	49	75.0%	36.8%
KNN	138	26	21	30	82.1%	55.3%
Ensemble (Boosted Trees)	138	33	21	23	85.7%	61.1%
Ensemble (Bagged Tree)	145	33	14	23	86.3%	70.2%
Ensemble (Subspace	144	15	15	41	77.8%	50.0%
KNN)						
Ensemble (RUSBoost	126	40	33	16	88.7%	54.8%
Trees)						

On the other hand, for the true negative (TN) it can be concluded that the data is relatively mixed and divergent, where it ranged from 2 at the quadratic discriminant classifier to 33 at two of the ensemble classifiers (boosted trees and bagged trees, Table 6.14). In general, the ensemble (bagged tree) classifier had the greatest values of sensitivity, specificity and TN. Meanwhile, the ensemble (RUSBoost trees) classifier generally had lower values, in comparison with the other classifiers.

6.5.3.2 Accuracy, precision, ROC/AUS and the percentages of error rate for the third approach

The results from the third approach for the model performance metrics are presented in Table 6.15, where the accuracy, precision, error rates and ROC/AUC are shown for eight different classifiers. The data shows that the ensemble (bagged tree) classifier had the highest value of accuracy of slightly less than 83% and the lowest error rate of 17.2%. It was followed by the ensemble (boosted trees) classifier with slightly less than 80% and the DT classifier with about 79%, where they both also had relatively low rates of error (20.3% and 21.2 %, respectively). Meanwhile, the quadratic discriminant and the quadratic SVM classifiers had the same highest (0.92) value of precision, followed by both classifiers of the ensemble (subspace KNN) and ensemble (bagged tree) which also had relatively high (0.91) and equal values of precision. With ROC/AUC, the highest values came from two classifiers of the ensemble; bagged tree and RUSBoost Trees with values of 0.82, and 0.81, respectively (Table 6.15). In general, the ensemble (bagged tree) classifier had the best value of accuracy, ROC/AUC and error rate.

Table 6.15: Standard classifiers performance evaluation metrics for the third approach

Classifiers	Accuracy	Precision	ROC/AUC	Error Rate
	%			%
DT	78.80	0.88	0.75	21.20
Quadratic Discriminant	69.02	0.92	0.56	30.98
Quadratic SVM	71.90	0.92	0.59	28.10
KNN	76.46	0.87	0.68	23.54
Ensemble (Boosted Trees)	79.72	0.87	0.77	20.28
Ensemble (Bagged tree)	82.80	0.91	0.82	17.20
Ensemble (Subspace KNN)	73.58	0.91	0.71	26.42
Ensemble (RUSBoost Trees)	77.00	0.79	0.81	23.00

6.5.3.3 Model performance measures using the third approach

The results from the third approach for the model performance measures are presented in Table 6.16. Where, Kappa, Youden’s index, G-Mean, MCC, Balanced Accuracy and F-Measure values are recorded for the eight different classifiers. In general, the F-measure had the highest values against all the classifiers, where, the values ranged between 0.81 (Quadratic Discriminant) and 0.88 (Ensemble (Bagged tree)). It was followed by the values of the G-Mean and Balanced accuracy, where, they were relatively high as they ranged from 0.43 for Balanced accuracy (Quadratic Discriminant) to 0.78 (Ensemble (Bagged tree)), with the exception of the lowest value (0.31) which was recorded to G-Mean at the Quadratic Discriminant classifier; this was also the classifier with the lowest value against the F-measure. In addition, the ensemble classifiers (Boosted Trees, Bagged tree, Subspace KNN and RUSBoost Trees) generally showed good values, when compared to the other classifiers. Negative low values are recorded at the quadratic discriminant against Kappa, Youden’s index and MCC. Comparisons between different classifiers’ performance in this approach in terms of accuracy, AUC and various classification evaluation metrics are shown graphically in Figure 6.9.

Table 6.16: More classifiers performance evaluation measures using the third approach

Classifiers	Kappa	Youden’s index	G-Mean	MCC	Balanced Accuracy	F-Measure
DT	0.43	0.44	0.71	0.42	0.72	0.86
Quadratic Discriminant	-0.07	-0.14	0.31	-0.08	0.43	0.81
Quadratic SVM	0.07	0.08	0.50	0.07	0.56	0.84
KNN	0.36	0.37	0.67	0.35	0.69	0.84
Ensemble (Boosted Trees)	0.47	0.47	0.72	0.46	0.74	0.87
Ensemble (Bagged tree)	0.53	0.56	0.78	0.53	0.78	0.88
Ensemble (Subspace KNN)	0.21	0.28	0.62	0.22	0.64	0.84
Ensemble (RUSBoost Trees)	0.45	0.44	0.70	0.47	0.72	0.84

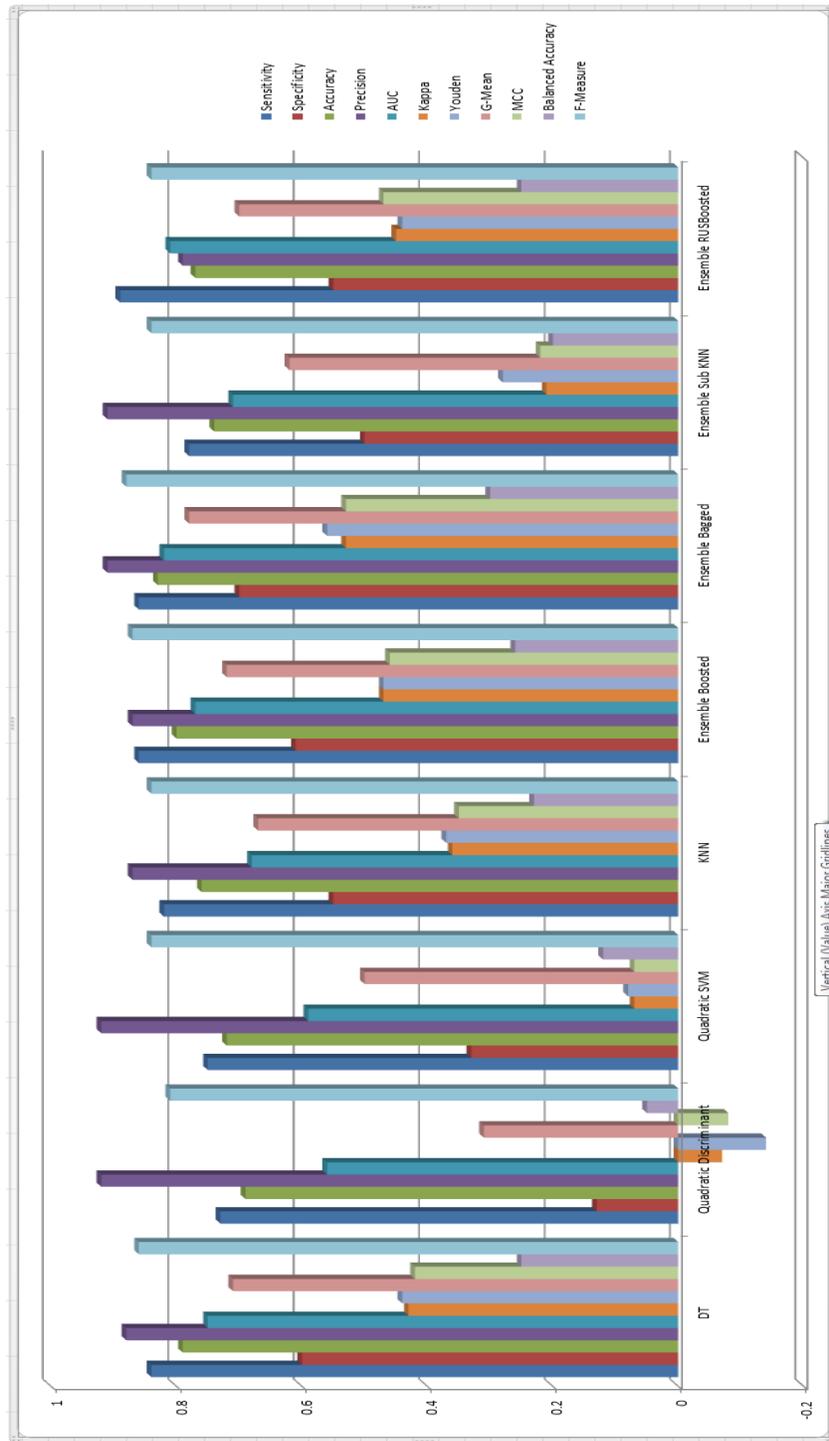


Figure 6.9: Graph showing the classification evaluation metrics for different classifiers using the dividing AT into sub-images approach

6.5.4 Image Classification Analysis Results and Discussion

Ensemble methods ((Polikar, 2006), (Rokach, 2010)) are designed to increase the accuracy of a single classifier by training several different classifiers and combining their decisions to output a single class label. Recent experimental studies show that classifier ensemble may improve the classification performance in the case of the imbalanced dataset. The idea behind ensemble methods is to gather several classification algorithms into one meta-classifier, which should provide better generalization. However, because of the imbalanced dataset in our research, ensemble classifier results are outstanding in comparison to the standard classifiers. These classifiers are performing well with a balanced dataset where there is no bias toward the majority class. Moreover, in general, machine learning applications, and classification performance are often measured using accuracy. Unfortunately, accuracy is not suitable for medical decision support systems if the dataset is imbalanced. Hence, we consider a combination of various performance measures instead of relying only on one. These metrics are shown in Tables (6.10, 6.13 and 6.16) for the approaches one, two and three respectively.

6.5.4.1 Approach one- the whole AT area

The main discussion of approach one results shown in Tables (6.8 – 6.10) indicates that the ensemble subspace KNN algorithm has produced good performance values, which is better than the other selected ensemble classifiers. On the contrary, quadratic SVM has recorded the lowest performance values, due to the combined, interrelated and imbalanced dataset. Additionally, with nonlinear quadratic function SVM algorithm was unable to reach an accuracy of more than 63%. DT has produced the highest accuracy among all other standard classifiers. However, DT accuracy almost reached 83%, and F-Measure metrics stand with the value of 0.89, which shows that DT is prime in comparison to other standard classifiers.

In ensemble classifiers, the AUC recorded all the value ranges between 0.78 (Subspace KNN) and 0.85 (RUSBoost Trees), as these values are far away from randomization. To contrast, the standard classifiers range is between 0.55 to 0.77, which suggests the ensemble classifiers are more efficient and reliable than standard classifiers.

In general, precision values seem nearer to most of the classifiers with an average of 0.92, though there are a couple of exceptions of RUSBoost Trees and Quadratic SVM whose values resulted in 0.83 and 0.70 respectively.

To conclude, in the first approach we considered the whole AT area as our ROI. Texture features are extracted from the whole AT regions instead of targeting a specific region in the AT. Balanced and imbalanced metrics have been used to evaluate the performance of various classifiers. Most of the ensemble classifiers recorded accuracy higher than the standard classifiers in most used metrics. However, in this approach with a minimum user intervention, the model achieved an accuracy of 83% and 0.90 for the F-measure metric.

6.5.4.2 Approach Two- Manual Determination for AT ROI

In this Section, the produced results from approach two have been discussed and concluded by the researcher. The performance metrics in Tables (6.11 to 6.13) illustrates that the ensembled classifiers are more efficient than standard classifiers. The Subspace KNN and Boosted Trees are almost equivalent to each other, and these two ensemble classifiers results are similar in most of the metrics cases. However, the results between Bagged Tree and RUSBoost Trees are recorded as being different to some extent. The Kappa metrics for Boosted Trees are more efficient than any other classifiers with 0.47 result value and recorded 0.0, which is the lowest for Quadratic SVM. Thus, the researcher concludes that ensemble classifiers are more efficient and Kappa metric supports focus on providing better results in this research approach.

After selecting a particular area as the ROI, the best accuracy recorded is for ensemble Boosted Trees with 80.18% while considering standard classifiers Quadratic SVM are recorded as their low accuracy with a value of 53.86%. AUC produced better results with ensemble classifiers such as Bagged Tree, and RUSBoost Tree results are a similar value of 0.77, which is the maximum as compared to other all classifiers. In contrast, standard classifier DT produced the result value of 0.75, which is very high in standard classifiers and similar to the high value of ensembled classifiers.

The precision works better with Subspace KNN with the high value of 0.93 in comparison to the lowest precision value of 0.58 obtained by the Quadratic SVM in ensemble classifiers. However, in the standard classifier, Quadratic SVM produced a very low percentage of accuracy (53.86) in this study.

The sensitivity and specificity obtained using ensemble classifiers are 0.87 of sensitivity for RUSBoost Trees, and a 0.63 maximum value was recorded for Subspace KNN.

The researcher concluded that the specificity ability of standard classifiers to detect the normality is very poor as compared to the ensemble classifiers. In general, detecting normal samples is more efficient than recognizing the abnormal samples. However, because determining the ROI has been done manually in this approach, not all AT areas are targeted for extracting texture features and classifications. Thus, the ability of most classifiers is affected by the manual determined for ROI, especially in classifying the abnormal samples.

6.5.4.3 Approach Three- Dividing the AT Image into Blocks

In this approach, AT ultrasound images have been divided into sub-images in order to focus more on ROI and to avoid the influence of differences between experts in determining ROI manually. In terms of classifiers, accuracy and precision both standard and ensemble classifiers have similar values ranging between 69% (Quadratic Discriminant) and 83% (Ensemble Bagged Tree) for accuracy, and between 0.79 (Ensemble RUSBoost) and 0.92 (Quadratic Discriminant) for precision. Within the ensemble classifiers, the bagged tree algorithm shows good results in all the evaluation metrics used in this research, while Quadratic SVM records the lowest results in general (Table 6.16). Most of the classifiers did not return high accuracy with regards to an abnormal case, as shown in the specificity results except for the bagged three, which recorded 0.70 compared with the lowest result recorded as 0.13 (Table 6.14) for quadratic discriminant. F-Measure shows the average value of all classifiers is approximately 0.84, which is affected by the similarity of sensitivity for all classifiers.

To conclude, the results of this approach, which divided the AT ultrasound images into sub-images, which have subsequently increased the number of samples of the research dataset, but in actual fact keep the imbalance ratio the same, explain the effectiveness of ensemble classifiers compared with the standard classifiers. Most of the imbalanced performance measures shown in Table 6.16 agreed with the performance metrics shown in Table 6.15 because the research dataset is not highly imbalanced.

6.5.4.4 Comparison between the Three Approaches

In this study, a set of experiments was carried out to evaluate the influence of the ROI size on the classification accuracy. From the results shown previously in the three approaches, shown in Tables (6.8 - 6.16), it can be concluded that the performance of most ensemble classifiers remains the same regardless of the ROI size and shape, while other classifiers produced different results with respect to the different approaches considered for this unique research study.

In approaches two and three, parts from the AT have been excluded. These parts may have some percentage of abnormality reaching 20% according to the medical literature. This probability might explain the best results obtained by approach one.

In approach two, determining ROI is subjective, and it depends on the expert experience and understanding of the AT anatomy. In addition, this is not aligned with the research aim, which was to automate the process and the researcher does not want to waste the doctor's time by performing this extra tasks.

In approach three, determining the ROI is achieved by dividing the AT images into various sub-images and excluding unrelated sub-images from the targeted images. Although this process achieved an accuracy of 83%, it takes more time by dividing the image into blocks compared to approach one where we used the whole AT image and record a slightly better accuracy of 84%. Additionally, increasing the number of sub-images will lead to unnecessarily processing more images, thus increasing the computational time.

In summary, the results reported for all three approaches indicate that in the AT ultrasound images ruptures could happen in the most vulnerable area according to the medical reports or anywhere outside that area. To avoid misleading performance measures, we did not rely on only one performance measure to choose between candidate approaches. A combination of different evaluation metrics was used to avoid or minimize imbalance-biased performance estimates.

6.6 Summary

To conclude this chapter, the researcher conducted three experiments in order to build a model that will be able to distinguish between normal and abnormal AT ultrasound images without any user intervention. The produced results from these experiments are shown, described and discussed in this chapter. In the beginning, the denoising stage was performed using various filters, and their results are compared and discussed to select the most suitable candidate filter for our research model. Subsequently, to improve the denoising results, AT images were enhanced using CLAHE algorithm and results from this process for the three approaches have been recorded and discussed in detail. In addition, different texture features have been extracted from the AT ultrasound images and reduced using KPCA algorithm in order to reduce computational complexity. Finally, various standard and ensemble classifiers have been used to determine the most accurate classifier for the given dataset. To avoid the imbalanced misleading metrics, different evaluation metrics have been adapted to compare different classifiers and evaluate the whole classification accuracy.

Chapter 7 - CONCLUSIONS AND FUTURE WORK

The last chapter of this thesis summarises the outcomes and findings of the research study based on the data analysis performed on the available dataset. The produced results are related to the aim and objective presented in the introductory chapter. The chapter ends with enumerating the limitations of the current study together with the directions of future research.

7.1 Introduction

Ultrasound images play an important role in medical diagnosis processes today. There is widespread use of ultrasound images because of their unique features compared to the other types of image model techniques. Features such as non-invasive, portability nonionizing and low cost made ultrasound techniques on the focus in many research studies, in order to build decisional systems to help in early disease detection.

AT, as the strongest tendon in the human body is exposed to different kinds of ruptures during daily activity. AT ruptures affect about 1 in 5,000 people worldwide and most frequently (80% to 90% of cases) take place 2 to 6 cm proximal to its calcaneal insertion. Ultrasound examination is used for detecting AT ruptures and tendinitis. To detect early stages of partial ruptures is important to avoid unnecessary medical procedures such as surgery.

To the best of the researcher's knowledge, there is a lack of research that looks into attempts for early detection and diagnosis for AT abnormalities from ultrasound images. We built a complete model which enables one to crop, denoise, enhance, extract the important features and classify the AT ultrasound images. This could mitigate the treatment cost by early detection and reduce the recovery time for the potentially affected patients. Moreover, with this model, we can save the expert's time as well as saving patients from bearing the pain while waiting for a long time to be seen by the experts.

Many Computer Aided Diagnosis (CAD) systems have been built by various researchers to explore the computer technology in the medical field. These CAD systems are customised to solve special problems using certain computer algorithms. However, in the literature, most of these CADs focus on breast, brain tumour and other parts of the disease because of their risk to the lives of the human being as well as their high rates worldwide. AT tissue and shape are different from most of the organs targeted by previous research. Hence, it is difficult to evaluate or compare our model with the other CAD models reported in the literature.

Speckle noise deteriorates ultrasound images and hinders the important features in the image. Reducing the speckle noise without losing any useful information is not an easy task in many research fields. In this study, different denoising algorithms have been tested and evaluated in order to mitigate the influence of the speckle noise on the AT ultrasound images. Additionally, different evaluation metrics are used to evaluate the efficiency of these filters. SRAD has been adopted in this research as the denoising filter for its capability in handling the multiplicative noise image. Also, it can suppress the speckle noise while maintaining the edge information.

Enhancing ultrasound images is an important step in the medical image processing. There are various enhancing algorithms already proposed to improve the image quality and enhance the visual appearance of the images. Enhancing algorithm can improve the image contrast and histogram which enhance the accuracy of the subsequent stages. To improve the image contrast in our research, CLAHE algorithm is used. It works on small areas called tiles rather than the entire image, so, the image contrast in the homogenous area can be limited to a minimum.

Most image features have been constructed based on histogram and texture features. The texture is considered as one of the most important features when analysing the medical images. There are many techniques that have been proposed to extract texture features. In this study, different histogram and texture features have been extracted. Features such as skewness and kurtosis are extracted from samples after different segmentation methods were applied. Also, second-order description statistics are applied to extract various texture features. GLCM and GLRLM Algorithms are employed to extract different texture features. A total of 57 texture features were extracted from the AT ultrasound images. These sets

of features along with five histogram features and three region properties undergo a process of dimensionality reduction using KPCA algorithm in order to select the most dominant features.

Several linear and nonlinear feature reduction techniques are available. These techniques provide a means to reduce the dimensionality of the extracted features to solve the high dimension problems. Kernel PCA has been used in this research to reduce the number of the extracted features. The reduction can discard the irrelevant and noisy information when it is performed successfully. We applied the KPCA with Gaussian kernel function on the extracted features. Features have higher variances that have been preserved, which is best in describing the data. Then, the original extracted features are projected in a new lower-dimensional space using the transformation matrix, which provided the new features for this research.

Medical image classification is a very important step in medical decisional systems. Classification accuracy depends on the accuracy in the previous stages and the selection of a good classifier. Various classifiers have been adapted in this research to improve the classification accuracy. Also, because the research dataset was imbalanced, different ensemble classifiers are explored to achieve higher classification accuracy. Most of the ensemble classifiers works better than the standard classifiers in all the ROI determination approaches.

7.2 Conclusions

7.2.1 Aim

This study aimed to develop a robust model that visually improves analyses and effectively classifies AT ultrasound images. This aim was achieved and accomplished by building an image processing decisional system for the AT. This system can distinguish between normal and abnormal AT ultrasound images. In this decisional system, AT image was improved and enhanced to achieve high accuracy of classification without any user interventions.

7.2.2 Process of Achievement of the Research Objectives

- **To identify and extract the most relevant features and information from AT ultrasound images**

The researcher determined and identified the best AT ultrasound texture features that can be used effectively in the classification process. First statistical parameters such as mean, variance, skewness, and kurtosis are computed through the values of intensities. The second order statistical parameters do, however, take into consideration the relationships among the neighbouring intensities. Thus, to attain the spatial dependence correlation, the second order parameter is utilised.

- **To use the extracted features to classify specific characteristics of the AT image**

The extracted features have been reduced in order to avoid the problem of dimensionality. Various classifiers are trained and tested using the new reduced features to estimate the accuracy of the classifier. High classification accuracy has been recorded especially with the ensemble classifiers.

- **To determine the discriminatory ability of the textural features used in this research**

Texture features are extracted in this research using different extraction algorithms. These features are chosen and extracted among different features because they can reflect the statistical characteristics of the ultrasound image grey level intensity. It is proven that features of GLCM and GLRLM are effective when it comes to the recognition of texture in images. All these features have been tested in the classification stage by the task-based evaluation method.

- **To investigate the influence of the ROI size on the classification accuracy**

The researcher conducted three different experiments to investigate the influence of ROI on the classification accuracy. A whole AT area, manually determining the ROI and the AT image being divided into sub-images. All these experiments in determining ROI have been evaluated and tested with various classifiers. The results discussed in detail in chapter 5 show that using the whole AT area as the ROI leads to higher accuracy and can save time as well as an effort for the experts.

- **To distinguish the ROI part from different image parts**

Determining the AT ultrasound image ROI is a major part of this research. We conduct experiments to investigate the most vulnerable ROI in the AT according to the medical doctor's reports. Different

classification accuracies have been achieved with different ROIs. A comparison and justification of all these results have been explored and discussed in more detail in chapter 6.

7.2.3 Research Problem

AT is considered as one of the most affected tendons in the human body. Moreover, ultrasound images are the most applied image model in the medical field. There is a critical need to develop a software model to detect the most important ROI of the AT ultrasound images and decide whether they are normal or abnormal. The image processing decisional system has been designed for the AT ultrasound images to solve this problem with high accuracy and less user intervention. With early detection of AT ruptures and tendonitis, the recovery cost and time will be reduced for most of the patients. Using this model may help non-experts to detect the AT abnormality, which will decrease the effort and minimise the expert interventions.

7.3 Contribution of the Study

This study contributes to the development of a new decisional system, which is the first for AT ultrasound image classification. This research study contributes to deciding better results of AT. This result is rectified based on different ROI regions that have been tested and evaluated using various classifiers. The influence of ROI size and position has been explored. Also, the most important features affecting the classification accuracy have been identified. For this system, standard and ensemble classifiers are evaluated using different evaluation metrics.

7.4 Research Scope and Limitations

7.4.1 Research Scope

This research targets ultrasound images for the AT, from the medical stream and systematic image processing in the computer science field. It does not deal or work with different kinds of medical image techniques such as MRI, CT scan or X rays. The research dataset contains only two-dimensional grey AT ultrasound images. The researcher focussed only on the lower and middle parts of the AT, which started from the calcaneal insertion point upward the two heads of the gastrocnemius. The goal of this

research is to distinguish between normal and abnormal ultrasound AT images without determining precisely the type of abnormality.

7.4.2 Research Limitations

The first limitation of the study is the research dataset composed of 57 patients' images, and this number of images is acceptable for the research because it produced stable results. However, the total number of images is small to allow more accurate and reliable outcomes. The second limitation is that the dataset of images is imbalanced in their distribution between normal and abnormal categories. This imbalanced distribution sometimes affects the system classification accuracy and performance. The third limitation is that some of the dataset images were noisier than others. This noisiness occurred because it is influenced by the speckle noise and the limited capability of the used portable ultrasound scan device. Portable ultrasound devices have limited resolution, and the quality of the scanned images depends on the radiologist experiences. Finally, the lack of CAD and image processing decisional systems targeting the ultrasound AT in the literature limit the evaluation and comparison process in our system.

7.5 Future Work

- **Measure AT Thickness**

AT thickness is different for normal and abnormal. This difference in the thickness measures could be utilised for the sake of classification processes. Using Hough algorithm and different thickness measure algorithms may help in this field. Applications have been used in measuring the lumen diameter, distension and intima-media thickness in the study of arterial characterisation that may be applied to AT.

- **Applying Segmentation Methods**

To determine more precisely the ROI in the AT ultrasound images certain segmentation methods may be used. Such methods may be utilised in discarding unrelated areas, which can reduce the noise and alleviate the classification accuracy.

- **Using Feature Selection Ranking**

Different features extracted from the AT ultrasound images can be reduced using different reduction algorithms. Instead of using feature reduction algorithms feature selection ranking may be used and their effect estimated on the classification accuracy. Various feature selections are found in the literature, and different applications have been developed in the medical image processing field. The feature selection method can be performed in three steps: screening, ranking and selecting. In the end, the most important subset features are identified. Such algorithms will contribute towards widening the theory of the subject area.

- **Using Unsupervised Classifiers**

Classification procedures that use labelled samples to train the classifier are called supervised, while classification procedures which use only unlabelled samples are named unsupervised. In this research, unsupervised algorithms are not applied in the present research. Algorithms such as k-mean for clustering or a priori algorithm for association rule learning problems can be used.

- **Real-time Three-dimensional Ultrasound**

Real-time three-dimensional ultrasound images have attracted much more attention amongst medical researchers because it helps clinicians acquire high-quality images. These images are sometimes necessary for investigating the automatic analysis or quantitative measurements. Such techniques could be applied on AT to perform more successful segmentation and classification process.

- **Non-Ultrasound**

Studies and applications of different medical image techniques such as MRI and CT scan might be conducted. Many ultrasound limitations can be overcome with these image techniques such as speckle noise. This variety of medical image models could be explored for future studies to build an image processing decisional system for the AT.

7.6 Summary

This chapter provides the conclusions, contributions, limitations and openings to the further research studies. It also shows how the researcher achieved his aim by describing the process of objectives. The researcher indicated the future perspectives for this research.

REFERENCES

- Aarnink, R., Giesen, R., Huynen, A., De La Rosette, J., Debruyne, F., & Wijkstra, H. (1994). A practical clinical method for contour determination in ultrasonographic prostate images. *Ultrasound in Medicine & Biology*, 20(8), 705-717.
- Abd-Elmoniem, K. Z., Yousef, A.-B. M., & Kadah, Y. M. (2002). Real-Time Speckle Reduction and Coherence Enhancement in Ultrasound Imaging via Nonlinear Anisotropic Diffusion. *IEEE Trans on Biomedical Engineering*, 49(9), 997-1014.
- Abdel-Hamid, O., Mohamed, A.-r., Jiang, H., & Penn, G. (2012). *Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition*. Paper presented at the Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on.
- Achim, A., Bezerianos, A., & Tsakalides, P. (2001). Novel Bayesian Multiscale Method for Speckle Removal in Medical Ultrasound Images. *IEEE Transactions on Medical Imaging*, 20(8), 772-783.
- Agrawal, P., & Verma, J. S. (2013). A Survey of Linear and Non-Linear Filters for Noise Reduction. *International Journal of Advance Research in Computer Science and Management Studies*, 1(3), 18-25.
- Ahmed, I., Lagopoulos, M., McConnell, P., Soames, R., & Sefton, G. (1998). Blood supply of the Achilles tendon. *Journal of orthopaedic research*, 16(5), 591-596.
- Ahmed, M. N., Yamany, S. M., Mohamed, N., Farag, A. A., & Moriarty, T. (2002). A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging*, 21(3), 193-199.
- Ajain, A. K., & Bhattacharjee, S. K. (1992). Address Block Location on Envelopes Using Gabor Filters. *Pattern Recognition*, 25(12), 1459-1477.
- Ajis, A., Maffulli, N., Alfredson, H., & Almekinders, L. C. (2007). Tendinopathy of the main body of the Achilles tendon. *The Achilles tendon*. London: Springer-Verlag, 59-69.
- Akosa, J. S. (2017). Predictive Accuracy: A Misleading Performance Measure for Highly Imbalanced Data. *Схема документа: <http://support.sas.com/resources/papers/proceedings17/0942-2017.pdf>*.
- Al-Kadi, O. S. (2009). Assessment of texture measures susceptibility to noise in conventional and contrast enhanced computed tomography lung tumour images. *Computerized Medical Imaging and Graphics*, 34(6), 494-503. doi:10.1016/j.compmedimag.2009.12.011
- Alexander, R., Ker, R., Bennet, M., Bibby, S., & Kester, R. (1987). The spring in the arch of the human foot. *Nature*, 325(6100), 147-149.
- Altman, D. G., & Bland, J. M. (1994). Statistics Notes: Diagnostic tests 2: predictive values. *Bmj*, 309(6947), 102.
- Analytic, B. (2017). Assessing Expected Goals Models. Retrieved from <http://business-analytic.co.uk/blog/evaluating-expected-goals-models/>
- Baaziz, N., Abahmane, O., & Missaoui, R. (2010). Texture Feature Extraction in The Spatial-Frequency Domain for Content-Based Image Retrieval. *Computer Vision and Pattern Recognition*.
- Baily, T., & Jain, A. K. (1978). A Note on Distance-Weighted k-Nearest Neighbor Rules. *IEEE Transaction on Syaytems, Man, and Cybernetics*, SMC-8(4).
- Barandela, R., Sánchez, J. S., Garcia, V., & Rangel, E. (2003). Strategies for learning in class imbalance problems. *Pattern Recognition*, 36(3), 849-851.
- Barfod, K. W. (2013). Achilles Tendon Rupture; Assessment of Non-operative Treatment. *Danish Med Journal*.
- Basset, O., & Mestas, J.-L. (1993). Texture Analysis of Ultrasonic Images of the Prostate by Means of Co-Occurrence Matrices. *ULTRASONIC IMAGING*. doi:10.1006/uimg.1993.1014

- Batista, G. E., Prati, R. C., & Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. *ACM Sigkdd Explorations Newsletter*, 6(1), 20-29.
- Beevi, S. Z., Sathik, M. M., & Senthamaraiannan, K. (2010). A robust fuzzy clustering technique with spatial neighborhood information for effective medical image segmentation. *arXiv preprint arXiv:1004.1679*.
- Beevi, Z., & Sathik, M. (2009). A robust segmentation approach for noisy medical images using fuzzy clustering with spatial probability. *methods*, 29(37), 38.
- Benoît Magnin, Lilia Mesrob, Serge Kinkingnéhun, Mélanie Pélégrini-Issac, Olivier Colliot, Marie Sarazin, . . . Benali, H. (2009). Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. *Diagnostic Neuroradiology*, 51(2), 73-83.
- Bharati, M. H., Liu, J. J., & MacGregor, J. F. (2004). Image texture analysis: methods and comparisons. *Chemometrics and intelligent laboratory systems*, 72(1), 57-71.
- Bhatia, N., & Vandana. (2010). Survey of Nearest Neighbor Techniques. *International Journal of Computer Science and Information Security*, 8(2), 302-305.
- Bjur, D. (2010). *The human Achilles tendon: innervation and intratendinous production of nerve signal substances-of importance in understanding the processes of Achilles tendinosis*. Umeå university.
- Bleakney, White, L. M., & Maffulli, N. (2007). Imaging of the Achilles Tendon. 284.
- Bleakney, R. R., & White, L. M. (2005). Imaging of the Achilles tendon. *Foot and ankle clinics*, 10(2), 239-254.
- Body Heal. (2014). Achilles Tendon Injuries: Tendonitis, Tendinosis, Tear & Rupture Discussed. Retrieved from <http://www.bodyheal.com.au/blog/achilles-tendon-injuries-tendonitis-tendinosis-tear-rupture>
- Bonner, G. (2001). Decision making for health care professionals: use of decision trees within the community mental health setting. *Journal of Advanced Nursing*, 35, 349-356.
- Bossley, C. J., Martinelli, B., Maffulli, N., & Raisbeck, C. C. (2000). Rupture of the Achilles tendon. *JBJS*, 82(12), 1804.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123-140.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Brown. (2004). *Diversity in neural network ensembles*. University of Birmingham.
- Brown, Smallwood, R., Barber, D., Lawford, P., & Hose, D. (1998). *Medical physics and biomedical engineering*: CRC Press.
- Brown, Wyatt, J., Harris, R., & Yao, X. (2005). Diversity creation methods: a survey and categorisation. *Information Fusion*, 6(1), 5-20.
- Brukner, P. (1997). Sports medicine. Pain in the Achilles region. *Australian family physician*, 26(4), 463-465.
- Burges, C. (2005). Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers, chapter Geometric Methods for Feature Selection and Dimensional Reduction: A Guided Tour. *Kluwer Academic*, 1, 5.
- Cardinal, É., Chhem, R. K., & Beauregard, C. G. (1998). Ultrasound-guided interventional procedures in the musculoskeletal system. *Radiologic clinics of North America*, 36(3), 597-604.
- Castellano, G., Bonilha, L., Li, L., & Cendes, F. (2004). Texture analysis of medical images. *Clinical radiology*, 59(12), 1061-1069.
- Chan, V., & Perlas, A. (2011). Basics of Ultrasound Imaging. In S. N. Narouze (Ed.), *Atlas of Ultrasound-Guided Procedures in Interventional Pain Management* (pp. 13-19).
- Chang, A., & Miller, T. T. (2009). Imaging of tendons. *Sports health*, 1(4), 293-300.
- Chawla, N., Japkowicz, N., & Kotcz, A. (2004). Editorial: special issue on learning from imbalanced data sets, SIGKDD Explor. Newsl. 6 (1)(2004) 1-6.

- Chen, C.-m., Lu, H. H.-S., & Lin, Y.-c. (1999). An Early Vision-Based Snake Model for Ultrasound Image Segmentation. *Ultrasound in Medicine & Biology*, 26(2), 273-285.
- Chen, S., & Zhang, D. (2004). Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 34(4), 1907-1916.
- Chen, X.-W., & Lin, X. (2014). Big data deep learning: challenges and perspectives. *IEEE access*, 2, 514-525.
- Chen, X., Liu, X., Qian, Y., Gales, M., & Woodland, P. C. (2016). *CUED-RNNLM—An open-source toolkit for efficient training and evaluation of recurrent neural network language models*. Paper presented at the Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on.
- Cheng, Shan, J., Ju, W., Guo, Y., & Zhang, L. (2010). Automated Breast Cancer Detection and Classification Using Ultrasound Images: A survey. *Pattern Recognition*, 43(1), 299-317. doi:10.1016/j.patcog.2009.05.012
- Cheng, & Shi, X. (2004). A simple and effective histogram equalization approach to image enhancement. *Digital signal processing*, 14(2), 158-170.
- Cheng, Shi, X. J., Min, R., Hu, L. M., Cai, X. R., & Du, H. N. (2006). Approaches for automated detection and classification of masses in mammograms. *Pattern Recognition*, 39(4), 646-668. doi:10.1016/j.patcog.2005.07.006
- Chien, J.-T., & Ku, Y.-C. (2016). Bayesian recurrent neural network for language modeling. *IEEE transactions on neural networks and learning systems*, 27(2), 361-374.
- Chitrakala, S., Shamini, P., & Manjula, D. (2009). *Multi-class enhanced image mining of heterogeneous textual images using multiple image features*. Paper presented at the Advance Computing Conference, 2009. IACC 2009. IEEE International.
- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297.
- Cox, B. (2013). Acoustics for ultrasound imaging. *Lecture Notes, University College London*.
- Crass, J. R., Van de Vegte, G., & Harkavy, L. (1988). Tendon echogenicity: ex vivo study. *Radiology*, 167(2), 499-501.
- Cremilleux, B., & Robert, C. (1997). A theoretical framework for decision trees in uncertain domains: Application to medical data sets. *Lecture Notes In Artificial Intelligence*, 1211, 145-156.
- Crofts, G., Angin, S., Mickle, K. J., Hill, S., & Nester, C. (2014). Reliability of ultrasound for measurement of selected foot structures. *Gait & posture*, 39(1), 35-39.
- Cvancarova, M., Albrechtsen, F., Brabrand, K., & Samset, E. (2005). *Segmentation of ultrasound images of liver tumors applying snake algorithms and GVF*. Paper presented at the International Congress Series.
- Dabov, K., Foi, A., Katkovnik, V., & Egiazarian, K. (2008). *Image restoration by sparse 3D transform-domain collaborative filtering*. Paper presented at the Image Processing: Algorithms and Systems VI.
- Das, A., Nguyen, C. C., Li, F., & Li, B. (2008). Digital image analysis of EUS images accurately differentiates pancreatic cancer from chronic pancreatitis and normal tissue. *Gastrointestinal endoscopy*, 67(6), 861-867.
- Dasarathy, B. V., & Holder, E. B. (1991). Image characterizations based on joint gray level-run length distributions. *Pattern Recognition Letters*, 12, 497-502.
- Davis, L. S. (1975). A survey of edge detection techniques. *Computer graphics and image processing*, 4(3), 248-270.

- de Almeida, C. W., de Souza, R. M., & Candeias, A. L. B. (2010). *Texture classification based on co-occurrence matrix and self-organizing map*. Paper presented at the Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on.
- Deepa, S. N., & Aruna Devi, B. (2011). A Survey on Artificial Intelligence Approaches for Medical Image Classification. *Indian Journal of Science and Technology*, 4(11).
- Dempsey, M. F., Condon, B., & Hadley, D. M. (2002). MRI Safety Review. *Seminars in Ultrasound, CT, and MRI*, 23(5), 392-401.
- Domingos, P. (1999). *Metacost: A general method for making classifiers cost-sensitive*. Paper presented at the Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining.
- Dong, G., & Xie, M. (2005). Color clustering and learning for image segmentation based on neural networks. *IEEE transactions on neural networks*, 16(4), 925-936.
- Donoho, D. L. (1995). De-noising by Soft-Thresholding. *IEEE Trans on Information Theory*, 41(3), 613-627.
- Dougherty, G. (2009). *Digital Image Processing for Medical Applications*. Cambridge: Cambridge University Press.
- Duda, R. O., & Hart, P. E. (1972). Use of the Hough Transformation To Detect Lines and Curves in Pictures. *Graphics and Image Processing*. doi:10.1145/361237.361242
- Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification*. 2nd. Edition. New York, 55.
- El-Naqa, I., Yang, Y., Wernick, M. N., Galatsanos, N. P., & Nishikawa, R. M. (2002). A Support Vector Machine Approach for Detection of Microcalcifications. *IEEE Transactions on Medical Imaging*, 21(12).
- Fan, L., Santago, P., Riley, W., & Herrington, D. M. (2001). An adaptive template-matching method and its application to the boundary detection of brachial artery ultrasound scans. *Ultrasound in Medicine & Biology*, 27(3), 399-408.
- Fawcett. (2004). ROC graphs: Notes and practical considerations for researchers. *Machine Learning*, 31(1), 1-38.
- Fawcett. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874. doi:10.1016/j.patrec.2005.10.010
- Fawcett, & Provost, F. (1997). Adaptive fraud detection. *Data mining and knowledge discovery*, 1(3), 291-316.
- Fellingham, L. L., & Sommer, F. G. (1984). Ultrasonic characterization of tissue structure in the in vivo human liver and spleen. *IEEE transactions on sonics and ultrasonics*, 31(4), 418-428.
- Ferguson, A. (2012). Basic Ultrasound Physics. Retrieved from <https://www.slideshare.net/fergua/ultrasound-11592611>
- Fornage, B. (1986). Achilles tendon: US examination. *Radiology*, 159(3), 759-764.
- Fornage, B., & Rifkin, M. (1988). Ultrasound examination of tendons. *Radiologic clinics of North America*, 26(1), 87-107.
- Fox, J. M., Blazina, M. E., Jobe, F. W., Kerlan, R. K., Carter, V. S., Shields Jr, C. L., & Carlson, G. J. (1975). Degeneration and rupture of the Achilles tendon. *Clinical orthopaedics and related research*, 107, 221-224.
- Freitas, A., Costa-Pereira, A., & Brazdil, P. (2007). Cost-sensitive decision trees applied to medical data. *Data Warehousing and Knowledge Discovery*, 303-312.
- Freund, Y., & Schapire, R. E. (1995). *A decision-theoretic generalization of on-line learning and an application to boosting*. Paper presented at the European conference on computational learning theory.
- Frost, V. S., Stiles, J. A., Shanmugan, K. S., & Holtzman, J. C. (1982). A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, PAMI-4(2).
- Fukunaga, K. (1993). *Statistical pattern recognition*. Paper presented at the Handbook of pattern recognition & computer vision.

- Galar, M., Fernandez, A., Barrenechea, E., Bustince, H., & Herrera, F. (2012). A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(4), 463-484.
- Galloway, M. M. (1975). Texture analysis using gray level run lengths. *Computer graphics and image processing*, 4(2), 172-179.
- García-Pedrajas, N., Pérez-Rodríguez, J., García-Pedrajas, M., Ortiz-Boyer, D., & Fyfe, C. (2012). Class imbalance methods for translation initiation site recognition in DNA sequences. *Knowledge-Based Systems*, 25(1), 22-34.
- García, V., Mollineda, R. A., & Sánchez, J. S. (2008). On the k-NN performance in a challenging scenario of imbalance and overlapping. *Pattern Analysis and Applications*, 11(3-4), 269-280.
- Garg, R., Mittal, B., & Garg, S. (2011). Histogram equalization techniques for image enhancement. *Int. J. Electron. Commun. Technol*, 2(1), 107-111.
- Gates, G. W. (1972). The Reduced Nearest Neighbor Rule. *IEEE Trans on Information Theory*, 431-433.
- Gebauer, M., Beil, F. T., Beckmann, J., Sárváry, A. M., Ueblacker, P., Ruecker, A. H., . . . Meenen, N. M. (2007). Mechanical evaluation of different techniques for Achilles tendon repair. *Archives of orthopaedic and trauma surgery*, 127(9), 795-799.
- Ghael, S., Sayeed, A. M., & Baraniuk, R. G. (1997). Improved Wavelet Denoising Via Empirical Wiener Filter *Proc. SPIE, Wavelet Applications in Signal and Image Processing* (Vol. 3169, pp. 389-399).
- Gharehchopogh, F. S., Molany, M., & Mokri, F. D. (2013). Using Artificial Neural Network in Diagnosis of Thyroid Disease: a Case Study. *International Journal on Computational Sciences & Applications*, 3(4), 49-61. doi:10.5121/ijcsa.2013.3405
- Gibbon, W., Cooper, J., & Radcliffe, G. (1999). Sonographic incidence of tendon microtears in athletes with chronic Achilles tendinosis. *British journal of sports medicine*, 33(2), 129-130.
- Glasbey, C. A., & Horgan, G. W. (1994). *Filters Image Analysis for The Biological Sciences*.
- Gonzalez, R. C., & Woods, R. E. (2001). *Digital Image Processing*.
- Gonzalez, R. C., Woods, R. E., & Eddins, S. L. (2004). Intensity Transformations and Spatial Filtering *Digital Image Processing Using MATLAB* (pp. 65-107): Pearson Prentice-Hall.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1): MIT press Cambridge.
- Gottumukkal, R., & Asari, V. K. (2004). An improved face recognition technique based on modular PCA approach. *Pattern Recognition Letters*, 25(4), 429-436.
- Hafizah, W. M., & Supriyanto, E. (2012). Automatic generation of region of interest for kidney ultrasound images using texture analysis. *International Journal of Biology and Biomedical Engineering*, 6(1), 26-34.
- Hall-Beyer, M. (2000). GLCM texture: a tutorial. *National Council on Geographic Information and Analysis Remote Sensing Core Curriculum*.
- Hamid, B. A. (2011). *Image texture analysis of transvaginal ultrasound in monitoring ovarian cancer*. Cardiff University.
- Han, Y., Kim, J., Lee, K., Han, Y., Kim, J., & Lee, K. (2017). Deep convolutional neural networks for predominant instrument recognition in polyphonic music. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, 25(1), 208-221.
- Hangiandreou, N. J. (2003). AAPM/RSNA physics tutorial for residents: topics in US: B-mode US: basic concepts and new technology. *Radiographics*, 23(4), 1019-1033.
- Haralick. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5), 786-804.
- Haralick, Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Trans Syst Man Cybern*.

- Hathaway, R. J., Bezdek, J. C., & Hu, Y. (2000). Generalized fuzzy c-means clustering strategies using L/sub p/norm distances. *IEEE Transactions on Fuzzy Systems*, 8(5), 576-582.
- Hattrup, S. J., & Johnson, K. A. (1985). A review of ruptures of the Achilles tendon. *Foot & ankle*, 6(1), 34-38.
- He, Z., Huston, D. R., Grimm, S., Jannicky, E., Garra, B. S., Wagner, R. F., & K.A., W. (2004). Dependence of tissue characterization features on region of interest (ROI) size: studies on phantoms and simulations. *IEEE Ultrasonics Symposium, 2083*, 2082-2085.
- Hill, C. R. (1986). Physical principles of medical ultrasonics. *Chichester*.
- Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786), 504-507.
- Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on pattern analysis and machine intelligence*, 20(8), 832-844.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- Hodgson, R., O'Connor, P., & Grainger, A. (2012). Tendon and ligament imaging. *The British journal of radiology*, 85(1016), 1157-1172.
- Holli, K. K., Harrison, L., Dastidar, P., Wäljas, M., Liimatainen, S., Luukkaala, T., . . . Eskola, H. (2010). RTeesexartchu arreteicl eanalysis of MR images of patients with Mild Traumatic Brain Injury. *BMC Medical Imaging*, 10(8).
- Huang, D.-S., Wunsch, D. C., Levine, D. S., & Jo, K.-H. (2008). *Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence: Fourth International Conference on Intelligent Computing, ICIC 2008 Shanghai, China, September 15-18, 2008, Proceedings* (Vol. 5227): Springer.
- Hwang, H., & Haddad, R. A. (1995). Adaptive Median Filters: New Algorithms and Results. *IEEE Trans on Image Processing*, 4(4), 499-502.
- iWalk Free. (2000). Achilles Tendon Rupture – Symptoms, Causes, Treatments, Surgery and Recovery
- Retrieved from <http://iwalk-free.com/injury-resource-center/achilles-tendon-rupture/>
- Jähne, B. (2002). Digital image processing: IOP Publishing.
- Jain, A. K. (1989). *Fundamentals of Digital Image Processing*: Prentice-Hall.
- James, S. L., Bates, B. T., & Osternig, L. R. (1978). Injuries to runners. *The American journal of sports medicine*, 6(2), 40-50.
- Jeong, C. B., Kim, K. G., Kim, T. S., & Kim, S. K. (2011). Comparison of image enhancement methods for the effective diagnosis in successive whole-body bone scans. *Journal of Digital imaging*, 24(3), 424-436.
- Jiang, Marti, C., Irniger, C., & Bunke, H. (2006). Distance measures for image segmentation evaluation. *EURASIP Journal on Applied Signal Processing*, 2006, 209-209.
- Jiang, Y., Nishikawa, R., Schmidt RA, Metz CE, Giger ML, & K, D. (1999). Improving Breast Cancer Diagnosis With Computer-Aided Diagnosis. *Academic Radiology*, 6(1), 22-33.
- Johansen, M. (2011). *Nonlinear Filtering in Digital Image Processing*. (Master of Science in Electronics), Norwegian University of Science and Technology.
- Kadah, Y. M., Farag, A. A., Zurada, J. M., Badawi, A. M., & Youssef, A.-B. (1996). Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images. *IEEE Transactions on Medical Imaging*, 15(4), 466-478.
- Kader, D., Mosconi, M., Benazzo, F., & Maffulli, N. (2005). Achilles tendon rupture *Tendon Injuries* (pp. 187-200): Springer.
- Kainberger, F. M., Engel, A., Barton, P., Huebsch, P., Neuhold, A., & Salomonowitz, E. (1990). Injury of the Achilles tendon: diagnosis with sonography. *AJR. American journal of roentgenology*, 155(5), 1031-1036.

- Kale, K. V., Mehrotra, S. C., & Manza, R. R. (2007). *Advances in Computer Vision and Information Technology*. India: I. K. International Publishing House Pvt. Ltd.
- Kälebo, P., Allenmark, C., Peterson, L., & Swärd, L. (1992). Diagnostic value of ultrasonography in partial ruptures of the Achilles tendon. *The American journal of sports medicine*, 20(4), 378-381.
- Kannan, S., Ramathilagam, S., Pandiyarajan, R., & Sathya, A. (2009). Fuzzy clustering Approach in segmentation of T1-T2 brain MRI. *Int. Journal of Recent Trends in Engineering*, Vol2(1), 157-160.
- Kanwal, N., Girdhar, A., & Gupta, S. (2011). *Region based adaptive contrast enhancement of medical x-ray images*. Paper presented at the Bioinformatics and Biomedical Engineering,(iCBBE) 2011 5th International Conference on.
- Kaplan, P. A., Matamoros Jr, A., & Anderson, J. C. (1990). Sonography of the musculoskeletal system. *AJR. American journal of roentgenology*, 155(2), 237-245.
- Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. (2014). *Large-scale video classification with convolutional neural networks*. Paper presented at the Proceedings of the IEEE conference on Computer Vision and Pattern Recognition.
- Kaur, K. (2013). Digital Image Processing in Ultrasound Images. *International Journal on Recent and Innovation Trends in Computing and Communication*, 1(4), 388-393.
- Khan, Fick, D., Keogh, A., Crawford, J., Brammar, T., & Parker, M. (2005). Treatment of Acute Achilles Tendon Ruptures. A Meta-Analysis of Randomized, Controlled Trials. *The Journal of Bone Joint Surgery*, 87(10), 2202-2210. doi:10.2106/JBJS.D.03049
- Khan, Zope, P. H., & Suralkar, S. R. (2013). Importance of Artificial Neural Network in Medical Diagnosis Disease Like Acute Nephritis Disease and heart disease. *International Journal of Engineering Science and Innovative Technology*, 2(2), 210-217.
- Khoshgoftaar, T. M., Van Hulse, J., & Napolitano, A. (2011). Comparing boosting and bagging techniques with noisy and imbalanced data. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 41(3), 552-568.
- Khoury, V., Guillin, R., Dhanju, J., & Cardinal, É. (2007). *Ultrasound of ankle and foot: overuse and sports injuries*. Paper presented at the Seminars in musculoskeletal radiology.
- Kim, Kim, P. K., Song, J. J., & Park, Y. C. (2000). *Analyzing blood cell image to distinguish its abnormalities*. Paper presented at the Proceedings of the eighth ACM international conference on Multimedia, New York, NY, USA.
- Kim, Park, J. M., Song, K. S., & Park, H. W. (1997). Adaptive mammographic image enhancement using first derivative and local statistics. *IEEE Transactions on Medical Imaging*, 16(5), 495-502.
- Ko, S.-J., & Lee, Y. H. (1991). Center Weighted Median Filters and Their Applications to Image Enhancement. *IEEE Trans on Circuits and Systems*, 38(9), 984-993.
- Kohavi, R. (1995). *A study of cross-validation and bootstrap for accuracy estimation and model selection*. Paper presented at the Ijcai.
- Koli, M., & Balaji, S. (2013). Literature Survey on Impulse Noise Reduction. *Signal & Image Processing*, 4(5), 75-95. doi:10.5121/sipij.2013.4506
- Krinidis, S., & Chatzis, V. (2010). A robust fuzzy local information C-means clustering algorithm. *IEEE Transactions on image processing*, 19(5), 1328-1337.
- Krissian, K., Westin, C.-f., Kikinis, R., & Vosburgh, K. G. (2007). Oriented Speckle Reducing Anisotropic Diffusion. *IEEE Trans on Image Processing*, 16(5).
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *Imagenet classification with deep convolutional neural networks*. Paper presented at the Advances in neural information processing systems.

- Kuan, D. T., Sawchuk, A. A., Strand, T. C., & Chavel, P. (1985). Adaptive Noise Smoothing Filter for Images with Signal-Dependent Noise. *IEEE Transaction on Pattern Analysis and Machine Intelligence, PAMI-7*(2).
- Kumar, G., & Bhatia, P. K. (2014). *A Detailed Review of Feature Extraction in Image Processing Systems*. Paper presented at the Fourth Intranational Conference on Advanced Computing & Communication Technologies.
- Kuncheva, L. I. (2001). Combining classifiers: Soft computing solutions. *Pattern recognition: From classical to modern approaches*, 427-451.
- Kuncheva, L. I., & Whitaker, C. J. (2003). Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. *Machine Learning, 51*(2), 181-207.
- Kung, S. Y. (2014). *Kernel Methods and Machine Learning*. United Kingdom: Cambridge University.
- Kurani, A. S., Xu, D.-H., Furst, J., & Raicu, D. S. (2004). *Co-occurrence matrices for volumetric data*. Paper presented at the 7th IASTED International Conference on Computer Graphics and Imaging, Kauai, USA.
- Kurita, T., & Otsu, N. (1993, 92). *Texture Classification by Higher Order Local Autocorrelation Features*. Paper presented at the Proc. Asian Conference on Computer Vision.
- Lakshmi, S., & Sankaranarayanan, V. (2010). A Study of Edge Detection Techniques for Segmentation Computing Approaches. *Computer Aided Soft Computing Techniques for Imaging and Biomedical Applications*.
- Langberg, H., Bulow, J., & Kjar, M. (1998). Blood flow in the peritendinous space of the human Achilles tendon during exercise. *Acta physiologica scandinavica, 163*(2), 149-154.
- Lantto, I. (2016). ACUTE ACHILLES TENDON RUPTURE.
- Lashari, S. A., & Ibrahim, R. (2013). A Framework for Medical Images Classification Using Soft Set. *Procedia Technology, 11*, 548-556. doi:10.1016/j.protcy.2013.12.227
- Law, H. C. (2006). *Clustering, dimensionality reduction, and side information*. Retrieved from
- Layer, G., Zuna, I., Lorenz, A., Zerban, H., Haberkorn, U., Bannasch, P., . . . R ath, U. (1990). Computerized ultrasound B-scan texture analysis of experimental fatty liver disease: influence of total lipid content and fat deposit distribution. *ULTRASONIC IMAGING, 12*(3), 171-188.
- LeCun, Y., Boser, B. E., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. E., & Jackel, L. D. (1990). *Handwritten digit recognition with a back-propagation network*. Paper presented at the Advances in neural information processing systems.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278-2324.
- Lee. (1980). Digital Image Enhancement and Noise Filtering by Use of Local Statistics. *IEEE Transaction on Pattern Analysis and Machine Intelligence, PAMI-2*(2).
- Lee, Chen, Y.-C., & Hsieh, K.-S. (2003). Ultrasonic liver tissues classification by fractal feature vector based on M-band wavelet transform. *IEEE Transactions on Medical Imaging, 22*(3), 382-392.
- Lee, Yan, J.-y., & Zhuang, T.-g. (2001). *A dynamic programming based algorithm for optimal edge detection in medical images*. Paper presented at the Medical Imaging and Augmented Reality, 2001. Proceedings. International Workshop on.
- Leppil hti, J., & Orava, S. (1998). Total Achilles tendon rupture. *Sports Medicine, 25*(2), 79-100.
-  ęski, J. (2003). Towards a robust fuzzy clustering. *Fuzzy Sets and Systems, 137*(2), 215-233.
- Letourneau, S., & Jensen, L. (1998). Impact of a decision tree on chronic wound care. *Wound Ostomy Continence Nurs, 25*, 240-247.

- Li, Chui, C. K., Chang, S., & Ong, S. H. (2011). Integrating Spatial Fuzzy Clustering with Level Set Methods for Automated Medical Image Segmentation. *Comput Biol Med*, 41(1), 1-10. doi:10.1016/j.compbimed.2010.10.007
- Li, Wang, Y.-N., Xiao, C.-Y., & Lu, X. (2012). A New Speckle Reducing Anisotropic Diffusion for Ultrasonic Speckle. *Acta Automatica Sinica*, 38(3), 412-418. doi:10.1016/s1874-1029(11)60301-7
- Li, H., Liu, K. R., Wang, Y., & Lo, S.-C. B. (1996). *Nonlinear filtering enhancement and histogram modeling segmentation of masses for digital mammograms*. Paper presented at the Engineering in Medicine and Biology Society, 1996. Bridging Disciplines for Biomedicine. Proceedings of the 18th Annual International Conference of the IEEE.
- Liew, A.-C., Yan, H., & Law, N.-F. (2005). Image segmentation based on adaptive cluster prototype estimation. *IEEE Transactions on Fuzzy Systems*, 13(4), 444-453.
- Lin, Fessell, D. P., Jacobson, J. A., Weadock, W. J., & Hayes, C. W. (2000). An illustrated tutorial of musculoskeletal sonography: part I, introduction and general principles. *American Journal of Roentgenology*, 175(3), 637-645.
- Liu, B., Cheng, H., Huang, J., Tian, J., Liu, J., & Tang, X. (2009). Automated segmentation of ultrasonic breast lesions using statistical texture classification and active contour based on probability distance. *Ultrasound in Medicine & Biology*, 35(8), 1309-1324.
- Lo, C.-H., & Don, H.-S. (1989). 3-D moment forms: their construction and application to object identification and positioning. *IEEE Transactions on pattern analysis and machine intelligence*, 11(10), 1053-1064.
- Logan, A. L., & Rowe, L. J. (1995). *The foot and ankle: clinical applications*: Jones & Bartlett Learning.
- Loizou, C. P., & Pattichis, C. S. (2008). Despeckle Filtering Algorithms and Software for Ultrasound Imaging. *Synthesis Lectures on Algorithms and Software in Engineering*, 1(1), 1-166. doi:10.2200/s00116ed1v01y200805ase001
- Loren, D. E., Seghal, C. M., Ginsberg, G. G., & Kochman, M. L. (2002). Computer-assisted analysis of lymph nodes detected by EUS in patients with esophageal carcinoma. *Gastrointestinal endoscopy*, 56(5), 742-746.
- Lui, S., Wei, J., Feng, B., Lu, W., Denby, B., Fang, Q., & Dang, J. (2012). An Anisotropic Diffusion Filter for Reducing Speckle Noise of Ultrasound Images Based on Separability. *Signal & Information Processing Association Annual Summit and Conference*.
- Lutz, H., & Buscarini, E. (2013). *Manual of diagnostic ultrasound* (Vol. 2): World Health Organization.
- Maaten, L. V. d., Postma, E., & Herik, H. V. d. (2009). *Dimensionality reduction: A comparative review*. Retrieved from
- Maffulli, N., Regine, R., Angelillo, M., Capasso, G., & Filice, S. (1987). Ultrasound diagnosis of Achilles tendon pathology in runners. *British journal of sports medicine*, 21(4), 158-162.
- Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017). Convolutional neural networks for large-scale remote-sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2), 645-657.
- Magnin, B., Mesrob, L., Kinkingnehun, S., Pelegrini-Issac, M., Colliot, O., Sarazin, M., . . . Benali, H. (2009). Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. *Neuroradiology*, 51(2), 73-83. doi:10.1007/s00234-008-0463-x
- Maintz, T. (2005). Segmentation *Digital and Medical Image Processing*.
- Maquirriain, J. (2011). Achilles tendon rupture: avoiding tendon lengthening during surgical repair and rehabilitation. *The Yale journal of biology and medicine*, 84(3), 289.
- Martin, K. (2010). Introduction to B-mode imaging. *Diagnostic ultrasound: physics and equipment*, 1-22.

- Martinoli, C., Derchi, L., Pastorino, C., Bertolotto, M., & Silvestri, E. (1993). Analysis of echotexture of tendons with US. *Radiology*, 186(3), 839-843.
- Massimiliano, P., & Alessandro, V. (1998). Support Vector Machines for 3D Object Recognition. *IEEE Transactions on pattern analysis and machine intelligence*, 20(6), 637-646. doi:10.1109/34.683777
- Materka, A. (2002). *Texture Analysis Methodologies for Magnetic Resonance Imaging*. Paper presented at the Biological Interface Conference, France.
- Materka, A. (2006). What is the texture. *Texture analysis for magnetic resonance imaging, Med4Publishing, Prague*, 11-44.
- Materka, A., & Strzelecki, M. (1998). Texture analysis methods—a review. *Technical university of lodz, institute of electronics, COST B11 report, Brussels*, 9-11.
- Mazurowski, M. A., Habas, P. A., Zurada, J. M., Lo, J. Y., Baker, J. A., & Tourassi, G. D. (2008). Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. *Neural Networks*, 21(2), 427-436.
- Mazzone, M. F., & Meccue, T. (2002). Common Conditions of the Achilles Tendon. *Medical College of Wisconsin*.
- McLauchlan, G., & Handoll, H. (2001). Interventions for treating acute and chronic Achilles tendinitis. *Cochrane Database Syst Rev*, 2(2).
- Mesrob, L., Sarazin, M., Hahn-Barma, V., Souza, L. C. d., Dubois, B., Gallinari, P., & Kinkingnéhun, S. (2012). DTI and Structural MRI Classification in Alzheimer's Disease. *Advances in Molecular Imaging*, 2, 12-20.
- Michailovich, O. V., & Tannenbaum, A. (2006). Despeckling of Medical Ultrasound Images. *IEEE Trans on Ultrasonics, Ferroelectrics, and Frequency Control*, 53(1).
- Minavathi, S, M., & Dinesh, M. S. (2012). Classification of Mass in Breast Ultrasound Images using Image Processing Techniques. *International Journal of Computer Applications*, 42(10), 29-36.
- Minin, I., & Minin, O. (2011). *ULTRASOUND IMAGING – MEDICAL APPLICATIONS* (First ed.). Janeza Trdine 9, 51000 Rijeka, Croatia: InTech.
- Mitra, A. K., & Parekh, R. (2011). Automated Detection of Skin Diseases Using Texture Features. *International Journal of Engineering Science and Technology*, 3(6), 4801-4808.
- Mohanaiah, P., Sathyanarayana, P., & GuruKumar, L. (2013). Image Texture Feature Extraction Using GLCM Approach. *International Journal of Scientific and Research Publications*, 3(5), 1-5.
- Mojsilovic, A., Popovic, M., Markovic, S., & Krstic, M. (1998). Characterization of visually similar diffuse diseases from B-scan liver images using nonseparable wavelet transform. *IEEE Transactions on Medical Imaging*, 17(4), 541-549.
- Morris, D. (1988). An evaluation of the use of texture measurements for the tissue characterisation of ultrasonic images of in vivo human placentae. *Ultrasound in Medicine & Biology*, 14(5), 387-395.
- Morrison, A. M. (2005). *Receiver Operating Characteristic (ROC) Curve Analysis of Antecedent Rainfall and the Alewife/Mystic River Receiving Waters*: Massachusetts Water Resources Authority, Environmental Quality Department.
- Morrison, A. M., Coughlin, K., Shine, J. P., Coull, B. A., & Rex, A. C. (2003). Receiver operating characteristic curve analysis of beach water quality indicator variables. *Applied and environmental microbiology*, 69(11), 6405-6411.
- Motwani, M. C., Gadiya, M. C., & Motwani, R. C. (2004, 109). *Survey of Image Denoising Techniques*. Paper presented at the Proc. of GSPx 2004, Santa Clara Convention Center, Santa Clara, CA.
- Murata, N., Yoshizawa, S., & Amari, S.-i. (1994). Network Information Criterion-Determining the Number of Hidden Units for an Artificial Neural Network Model. *IEEE Trans on Neural Networks*, 5(6), 865-872.

- Nam, S. H., & Choi, J. Y. (1998). *A method of image enhancement and fractal dimension for detection of microcalcifications in mammogram*. Paper presented at the Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE.
- Narayanan, S. K., & Wahidabanu, R. S. D. (2009). A View on Despeckling in Ultrasound Imaging. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2(3), 85-98.
- Nielsen. (2004). Musculoskeletal ultrasound in an European Journal. *Ultraschall in der Medizin*, 1(6), 533-534.
- Nielsen. (2007). *Real-Time Wavelet Filtering on the GPU*. (Master of Science in Computer Science), Norwegian University of Science and Technology.
- Noble, A., & Boukerroui, D. (2006). Ultrasound Image Segmentation: A Survey. *IEEE Transactions on Medical Imaging*, 25(8), 987-1010.
- Norton, I. D., Zheng, Y., Wiersema, M. S., Greenleaf, J., Clain, J. E., & DiMagno, E. P. (2001). Neural network analysis of EUS images to differentiate between pancreatic malignancy and pancreatitis. *Gastrointestinal endoscopy*, 54(5), 625-629.
- Nunley, J. A. (2009). *The Achilles Tendon*. USA: Duke University Medical Center.
- Ogata, T., Tan, J. K., & Ishikawa, S. (2006). High-speed human motion recognition based on a motion history image and an eigenspace. *IEICE TRANSACTIONS on Information and Systems*, 89(1), 281-289.
- Oghli, M. G., Fallahi, A., & Pooyan, M. (2010). *Automatic region growing method using GSmap and spatial information on ultrasound images*. Paper presented at the Electrical Engineering (ICEE), 2010 18th Iranian Conference on.
- Öhberg, L. (2003). *The Chronic Painful Achilles Tendon Sonographic Findings and New Methods for Treatment*. Umea University Med Dissertations.
- Olson, C. F. (1999). Constrained Hough Transforms for Curve Detection. *Computer Vision and Image Understanding*, 73(3), 329-345.
- Olsson, N. (2013). *Acute Achilles Tendon Rupture. Outcome, Prediction and Optimized treatment*.
- Osicka, T., Freedman, M. T., & Ahmed, F. (2007). Characterization of pulmonary nodules on computer tomography (CT) scans: The effect of additive white noise on features selection and classification performance. *Medical Imaging, Spie-Int Society Optical Engineering*, 6512, 51245-51245.
- Osuna, E., Freund, R., & Girosi, F. (1997). *Training support vector machines: Application to face detection*. Paper presented at the Computer Vision and Pattern Recognition, Puerto Rico.
- Otto, C. M. (2013). *Textbook of Clinical Echocardiography E-Book*: Elsevier Health Sciences.
- Palmer, L. S., & Palmer, J. S. (2014). *Pediatric and adolescent urologic imaging*: Springer Science & Business Media.
- Pang, B. S., & Ying, M. (2006). Sonographic measurement of Achilles tendons in asymptomatic subjects. *Journal of ultrasound in medicine*, 25(10), 1291-1296.
- Patidar, P., & Gupta, M. (2010). Image De-noising by Various Filters for Different Noise. *International Journal of Computer Applications*, 9(4), 45-50.
- Perona, P., & Malik, J. (1990). Scale-Space and Edge Detection Using Anisotropic Diffusion. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 12(7).
- Petrick, N., Chan, H.-P., Sahiner, B., & Wei, D. (1996). An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection. *IEEE Transactions on Medical Imaging*, 15(1), 59-67.
- Pham, D. L., & Prince, J. L. (1999). An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities. *Pattern Recognition Letters*, 20(1), 57-68.
- Pham, D. L., Xu, C., & Prince, J. L. (2000). Current methods in medical image segmentation. *Annual review of biomedical engineering*, 2(1), 315-337.

- Pisano, E. D., Zong, S., Hemminger, B. M., DeLuca, M., Johnston, R. E., Muller, K., . . . Pizer, S. M. (1998). Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. *Journal of Digital imaging*, *11*(4), 193-200.
- Pitas, I., & Venetsanopoulos, A. N. (1986). Nonlinear Mean Filters in Image Processing. *IEEE Trans on Acoustics, Speech, and Signal Processing*, *ASSP-34*(3), 573-584.
- Pizurica, A. (2002). *Image Denoising Using Wavelets and Spatial Context Modeling*. (PhD), University in Ghent, Belgium.
- Pizurica, A., Wink, A. M., Vansteenkiste, E., Philips, W., & Roerdink, J. B. T. M. (2006). A Review of Wavelet Denoising in MRI and Ultrasound Brain imaging. *Current Medical Imaging Reviews*, *2*(2), 247-260.
- Podgorelec, V., Kokol, P., Stiglic, B., & Rozman, I. (2002). Decision trees: an overview and their use in medicine. *Journal of Medical Systems*, *26*(5), 445-463.
- Polesel, A., Ramponi, G., & Mathews, V. J. (2000). Image enhancement via adaptive unsharp masking. *IEEE Transactions on image processing*, *9*(3), 505-510.
- Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and systems magazine*, *6*(3), 21-45.
- Pratt, W. K. (2001). *Digital Image Processing*: John Wiley & Sons, Inc.
- Qian, Y., Bi, M., Tan, T., & Yu, K. (2016). Very deep convolutional neural networks for noise robust speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *24*(12), 2263-2276.
- Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning* (Vol. 16): Morgan Kaufmann
- Quiroga, R. Q., & Garcia, H. (2003). Single-Trial Event-Related Potentials with Wavelet Denoising. *Clinical Neurophysiology*, *114*(2), 376-390. doi:10.1016/s1388-2457(02)00365-6
- Ragesh, N. K., Anil, A. R., & Rajesh, R. (2011). *Digital Image Denoising in Medical Ultrasound Images: A Survey*. Paper presented at the ICGST AIML-11 Conference.
- Raj, V. N. P., & Venkateswarlu, T. (2012). Denoising of Medical Images Using Dual Tree Complex Wavelet Transform. *Procedia Technology*, *4*, 238-244. doi:10.1016/j.protcy.2012.05.036
- Raju, P. D. R., & Neelima, G. (2012). Image Segmentation by Using Histogram Thresholding. *IJCSET*, *2*(1), 776-779.
- Raschka, S., Linear, P., Gaussian, R., & LLE, L.-L. E. (2014). Kernel tricks and nonlinear dimensionality reduction via RBF kernel PCA. *Blog*, September.
- Raudys, S. J., & Jain, A. K. (1991). Small sample size effects in statistical pattern recognition: Recommendations for practitioners. *IEEE Transactions on pattern analysis and machine intelligence*, *13*(3), 252-264.
- Ribeiro, R., & Sanches, J. (2009). *Fatty Liver Characterization and Classification by Ultrasound*. Paper presented at the IbPRIA.
- Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review*, *33*(1), 1-39.
- Rowland, S. (1979). Computer implementation of image reconstruction formulas. *Image reconstruction from projections*, 9-79.
- Russ, J. C., Matey, J. R., Mallinckrodt, A. J., & McKay, S. (1994). The image processing handbook. *Computers in Physics*, *8*(2), 177-178.
- Sahiner, B., Chan, H. P., & Hadjiiski, L. (2008). Classifier Performance Estimation Under The Constraint of A Finite Sample Size: Resampling Schemes Applied to Neural Network Classifiers. *Neural Networks*, *21*(2-3), 476-483. doi:10.1016/j.neunet.2007.12.012
- Sahoo. (2011). *Biomedical image fusion & segmentation using GLCM*. Paper presented at the International Journal of Computer Application Special Issue on "2nd National Conference—Computing, Communication and Sensor Network" CCSN.
- Sahoo, Soltani, S., & Wong, A. K. C. (1988). A Survey of Thresholding Techniques. *Computer Vision, Graphics and Image Processing*, *41*, 233-260.

- Sakrison, D. (1977). On the role of the observer and a distortion measure in image transmission. *IEEE Transactions on Communications*, 25(11), 1251-1267.
- Sasi, N. M., & Jayasree, V. (2013). Contrast limited adaptive histogram equalization for qualitative enhancement of myocardial perfusion images. *Engineering*, 5(10), 326.
- Schapire, R. E. (1990). The strength of weak learnability. *Machine Learning*, 5(2), 197-227.
- Schoelkopf, B., Sung, K., Burges, C., Girosi, F., Niyogi, P., Poggio, T., & Vapnik, V. (1996). Comparing Support Vector Machines with Gaussian Kernels to Radial Basis Function Classifiers. *IEEE Transactions on Signal Processing*. doi:10.1109/78.650102
- Schölkopf, B., Smola, A., & Müller, K.-R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*, 10(5), 1299-1319.
- Seiffert, C., Khoshgoftaar, T. M., Van Hulse, J., & Napolitano, A. (2008). *RUSBoost: Improving classification performance when training data is skewed*. Paper presented at the Pattern Recognition, 2008. ICPR 2008. 19th International Conference on.
- Selesnick, I. W., Baraniuk, R. G., & Kingsbury, N. G. (2005). The Dual-Tree Complex Wavelet Transform. *IEEE Signal Processing Mag*, 123-151.
- Selfridge, G. (1955). *Pattern Recognition and Modern Computers*. Paper presented at the Western Joint Computer Conference, Massachusetts Institute of Technology, Lexington, Mass.
- Selvaraj, H., Selvi, S. T., Selvathi, D., & Gewali, L. (2007). Brain MRI Slices Classification Using Least Squares Support Vector Machine. *International Journal of Intelligent Computing in Medical Sciences & Image Processing*, 1(1), 21-33. doi:10.1080/1931308X.2007.10644134
- Shan, J., Cheng, H., & Wang, Y. (2012a). Completely automated segmentation approach for breast ultrasound images using multiple-domain features. *Ultrasound in Medicine & Biology*, 38(2), 262-275.
- Shan, J., Cheng, H., & Wang, Y. (2012b). A novel segmentation method for breast ultrasound images based on neutrosophic I-means clustering. *Medical physics*, 39(9), 5669-5682.
- Shung, K. K. (2004). Ultrasound and tissue interaction *Encyclopedia of Biomaterials and Biomedical Engineering* (pp. 1-9): Taylor & Francis.
- Shung, K. K. (2015). *Diagnostic ultrasound: Imaging and blood flow measurements*: CRC press.
- Sifuzzaman, M., Islam, M. R., & Ali, M. Z. (2009). Application of Wavelet Transform and its Advantages Compared to Fourier Transform. *Journal of Physical Sciences*, 13.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Singh, S., & Al-Mansoori, R. (2000). Identification of regions of interest in digital mammograms. *Journal of Intelligent Systems*, 10(2), 183-217.
- Smutek, D., Šára, R., Sucharda, P., Tjahjadi, T., & Švec, M. (2003). Image texture analysis of sonograms in chronic inflammations of thyroid gland. *Ultrasound in Medicine & Biology*, 29(11), 1531-1543.
- Spackman, K. A. (1989). *Signal detection theory: Valuable tools for evaluating inductive learning*. Paper presented at the Proceedings of the sixth international workshop on Machine learning.
- Stark, J. A. (2000). Adaptive image contrast enhancement using generalizations of histogram equalization. *IEEE Transactions on image processing*, 9(5), 889-896.
- Su, Y., Wang, Y., Jiao, J., & Guo, Y. (2011). Suppl 1: Automatic Detection and Classification of Breast Tumors in Ultrasonic Images Using Texture and Morphological Features. *The open medical informatics journal*, 5, 26.
- Suguna, N., & Thanushkodi, K. (2010). An Improved k-Nearest Neighbor Classification Using Genetic Algorithm. *International Journal of Computer Science and Information Security*, 7(4), 18-21.

- Sun, Y., Wong, A. K., & Kamel, M. S. (2009). Classification of imbalanced data: A review. *International Journal of Pattern Recognition and Artificial Intelligence*, 23(04), 687-719.
- Suri, J. S. (2008). *Advances in diagnostic and therapeutic ultrasound imaging*: Artech House.
- Suykens, J. A., Gestel, T. V., Brabanter, J. D., Moor, B. D., & Vandewalle, J. (2002). Least squares support vector machines. *World Scientific*.
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *science*, 240(4857), 1285-1293.
- Swietojanski, P., Ghoshal, A., & Renals, S. (2014). Convolutional neural networks for distant speech recognition. *IEEE Signal Processing Letters*, 21(9), 1120-1124.
- Szabo, T. L. (2004). *Diagnostic ultrasound imaging: inside out*: Academic Press.
- Szilágyi, L., Benyo, Z., Szilágyi, S. M., & Adam, H. (2003). *MR brain image segmentation using an enhanced fuzzy c-means algorithm*. Paper presented at the Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE.
- Tang, Suganthan, P. N., & Yao, X. (2006). An analysis of diversity measures. *Machine Learning*, 65(1), 247-271.
- Tang, Y., Zhang, Y.-Q., Chawla, N. V., & Krasser, S. (2009). SVMs modeling for highly imbalanced classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(1), 281-288.
- Thangavel, K., Karnan, M., Sivakumar, R., & Mohideen, A. K. (2005). Ant colony system for segmentation and classification of microcalcification in mammograms. *International Journal on Artificial Intelligence and Machine Learning*, 5(3), 29-40.
- Thivakaran, T. K., & Chandrasekaran, R. (2010). Nonlinear Filter Based Image Denoising Using AMF Approach. *International Journal of Computer Science and Information Security*, 7(2), 224-227.
- Thorsten, J. (1999). *Transductive Inference for Text Classification using Support Vector Machines*. Paper presented at the roceedings of the Sixteenth International Conference on Machine Learning, San Francisco, CA, USA.
- Tian. (2013). A Review on Image Feature Extraction and Representation Techniques. *International Journal of Multimedia and Ubiquitous Engineering*, 8(4), 385-395.
- Tian, Bao, S. L., & Zhou, Q. M. (2003). Medical image processing and analysis. *Electronic Industry Press*, 96-114.
- Ting, K. M. (2002). An instance-weighting method to induce cost-sensitive trees. *IEEE Transactions on knowledge and data engineering*, 14(3), 659-665.
- Tongsen, H., & Ting, G. (2007). *Characteristics Preserving of Ultrasound Medical Images Based on Kernel Principal Component Analysis*. Paper presented at the Medical Imaging and Informatics.
- Tsai, D.-Y., & Kojima, K. (2005). Measurements of texture features of medical images and its application to computer-aided diagnosis in cardiomyopathy. *Measurement*, 37(3), 284-292.
- Tsien, C. L., Fraser, H. S. F., Long, W. J., & Kennedy, R. L. (1998). Using Classification Tree and Logistic Regression Methods to Diagnose Myocardial Infarction *Studies in Health Technology and Informatics* (Vol. 52, pp. 493 - 497).
- Tsiotsios, C., & Petrou, M. (2013). On The Choice of The Parameters for Anisotropic Diffusion in Image Processing. *Pattern Recognition*, 46(5), 1369-1381. doi:10.1016/j.patcog.2012.11.012
- Tuceryan, M., & Jain, A. K. (1993). Texture analysis. *Handbook of pattern recognition and computer vision*, 2, 235-276.
- Ulagamuthalvi, V., & Sridharan, D. (2012). *Automatic identification of ultrasound liver cancer tumor using support vector machine*. Paper presented at the International Conference on Emerging Trends in Computer and Electronics Engineering.
- Unser, M., & Aldroubi, A. (1996). A Review of Wavelets in Biomedical Applications. *Proc of The IEEE*, 84(4), 626-638.

- Valentin, L., Ameye, L., Testa, A., Lécuru, F., Bernard, J.-P., Paladini, D., . . . Timmerman, D. (2005). Ultrasound characteristics of different types of adnexal malignancies. *Gynecologic Oncology*. doi: 10.1016/j.ygyno.2005.11.015
- Van Der Maaten, L., Postma, E., & Van den Herik, J. (2009). Dimensionality reduction: a comparative. *J Mach Learn Res*, 10, 66-71.
- Wählby, L. (1978). *Achilles tendon injury: aspects on muscle structure and strength*. Umeå universitet.
- Walker, R., Jackway, P., & Longstaff, I. (1997). *Recent developments in the use of the co-occurrence matrix for texture recognition*. Paper presented at the Digital Signal Processing Proceedings, 1997. DSP 97., 1997 13th International Conference on.
- Wang, Itoh, K., Taniguchi, N., Toei, H., Kawai, F., Nakamura, M., . . . Ono, T. (2002). Studies on tissue characterization by texture analysis with co-occurrence matrix method using ultrasonography and CT imaging. *Journal of Medical Ultrasonics*, 29(4), 211-223.
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on image processing*, 13(4), 600-612.
- Wei, Y., Tiebin, L., Rodolfo, V., Marta, G., & Muin, J. K. (2010). Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. *Medical Informatics and Decision Making*. doi:10.1186/1472-6947-10-16
- Weinreb, J. H., Sheth, C., Apostolakis, J., McCarthy, M.-B., Barden, B., Cote, M. P., & Mazzocca, A. D. (2014). Tendon structure, disease, and imaging. *Muscles, ligaments and tendons journal*, 4(1), 66.
- Weiss, G. M., & Provost, F. (2003). Learning when training data are costly: The effect of class distribution on tree induction. *Journal of Artificial Intelligence Research*, 19, 315-354.
- Weyman, A. (1994). Physical principles of ultrasound. *Principles and practice of echocardiography*, 3-25.
- Wikipedia. (2018). Fuzzy Clustering. Retrieved from www.wikipedia.com
- Wilson, M., Hargrave, R., Mitra, S., Shieh, Y.-Y., & Roberson, G. H. (1998). *Automated detection of microcalcifications in mammograms through application of image pixel remapping and statistical filter*. Paper presented at the Computer-Based Medical Systems, 1998. Proceedings. 11th IEEE Symposium on.
- Wolbarst, A. B., Capasso, P., & Wyant, A. R. (2013). Diagnostic Ultrasound Contrast from Differences in Tissue Elasticity or Density Across Boundaries *Medical Imaging: Essentials for Physicians*: Wiley-Blackwell.
- Won, C. S., Park, D. K., & Park, S.-J. (2002). Efficient use of MPEG-7 edge histogram descriptor. *ETRI journal*, 24(1), 23-30.
- Wongsritong, K., Kittayaruasiriwat, K., Cheevasuvit, F., Dejhan, K., & Somboonkaew, A. (1998). *Contrast enhancement using multipeak histogram equalization with brightness preserving*. Paper presented at the Circuits and Systems, 1998. IEEE APCCAS 1998. The 1998 IEEE Asia-Pacific Conference on.
- Wren, T. A., Lindsey, D. P., Beaupré, G. S., & Carter, D. R. (2003). Effects of creep and cyclic loading on the mechanical properties and failure of human Achilles tendons. *Annals of biomedical engineering*, 31(6), 710-717.
- Wren, T. A., Yerby, S. A., Beaupré, G. S., & Carter, D. R. (2001). Mechanical properties of the human achilles tendon. *Clinical Biomechanics*, 16(3), 245-251.
- Wu, K.-L., & Yang, M.-S. (2002). Alternative c-means clustering algorithms. *Pattern Recognition*, 35(10), 2267-2278.
- Xian, G.-m. (2010). An identification method of malignant and benign liver tumors from ultrasonography based on GLCM texture features and fuzzy SVM. *Expert Systems with Applications*, 37(10), 6737-6741.
- Xie, X., & Mirmehdi, M. (2008). A Galaxy of Texture Features. In M. Mirmehdi, X. Xie, & J. Suri (Eds.), *Handbook of Texture Analysis* (pp. 375-407).

- Xiuqin, D., Yong, X., & Hong, P. (2008). A New Kind of Weighted Median Filtering Algorithm Used for Image Processing. *International Symposium on Information Science and Engineering*, 738-743. doi:10.1109/isise.2008.288
- Xu, Y., Weaver, J. B., Healy, D. M., Jr., & Lu, J. (1994). Wavelet Transform Domain Filters: A Spatially Selective Noise Filtration Technique. *IEEE Trans on Image Processing*, 3(6), 747-758.
- Yan, J., & Zhuang, T. (2003). Applying improved fast marching method to endocardial boundary detection in echocardiographic images. *Pattern Recognition Letters*, 24(15), 2777-2784.
- Yeh, W.-C., Huang, S.-W., & Li, P.-C. (2003). Liver fibrosis grade classification with B-mode ultrasound. *Ultrasound in Medicine & Biology*, 29(9), 1229-1235.
- Ying, M., Yeung, E., Li, B., Li, W., Lui, M., & Tsoi, C.-W. (2003). Sonographic evaluation of the size of Achilles tendon: the effect of exercise and dominance of the ankle. *Ultrasound in Medicine & Biology*, 29(5), 637-642.
- Yong, Z. (2009). An Improved KNN Text Classification Algorithm Based on Clustering. *Journal of Computer*, 4(3), 320-337.
- Yoshida, H., Keserci, B., Casalino, D. D., Coskun, A., Ozturk, O., & Savranlar, A. (1998). *Segmentation of liver tumors in ultrasound images based on scale-space analysis of the continuous wavelet transform*. Paper presented at the Ultrasonics Symposium, 1998. Proceedings., 1998 IEEE.
- Yu, & Acton, S. T. (2002). Speckle reducing anisotropic diffusion. *IEEE Trans Image Process*, 11(11), 1260-1270. doi:10.1109/TIP.2002.804276
- Yu, & Bajaj, C. (2004). *A fast and adaptive method for image contrast enhancement*. Paper presented at the Image Processing, 2004. ICIP'04. 2004 International Conference on.
- Zhai, J.-h., Wang, X.-z., & Zhang, S.-f. (2007, 95-0). *Rough-Neural Image Classification Using Wavelet Transform*. Paper presented at the the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong.
- Zhang. (1996). A Survey on Evaluation Methods for Image Segmentation. *Pattern Recognition*, 29(8), 1335-1346.
- Zhang, & Tan, T. (2002). Brief Review of Invariant Texture Analysis Methods. *Pattern Recognition*, 35, 735-747.
- Zhao, J., Zheng, W., Zhang, L., & Tian, H. (2013). Segmentation of ultrasound images of thyroid nodule for assisting fine needle aspiration cytology. *Health information science and systems*, 1(1), 5.
- Zhi, X., & Wang, T. (2008). An Anisotropic Diffusion Filter for Ultrasonic Speckle Reduction. *The Institution of Engineering and Technology*, 327-330.

BIBLIOGRAPHY

Loizou, Christos (2015). Image Despeckle Filtering Toolbox (<https://uk.mathworks.com/matlabcentral/fileexchange/54044-image-despeckle-filtering-toolbox?focused=5710051&tab=function>). MATLAB Central File Exchange. Retrieved May 20, 2016.

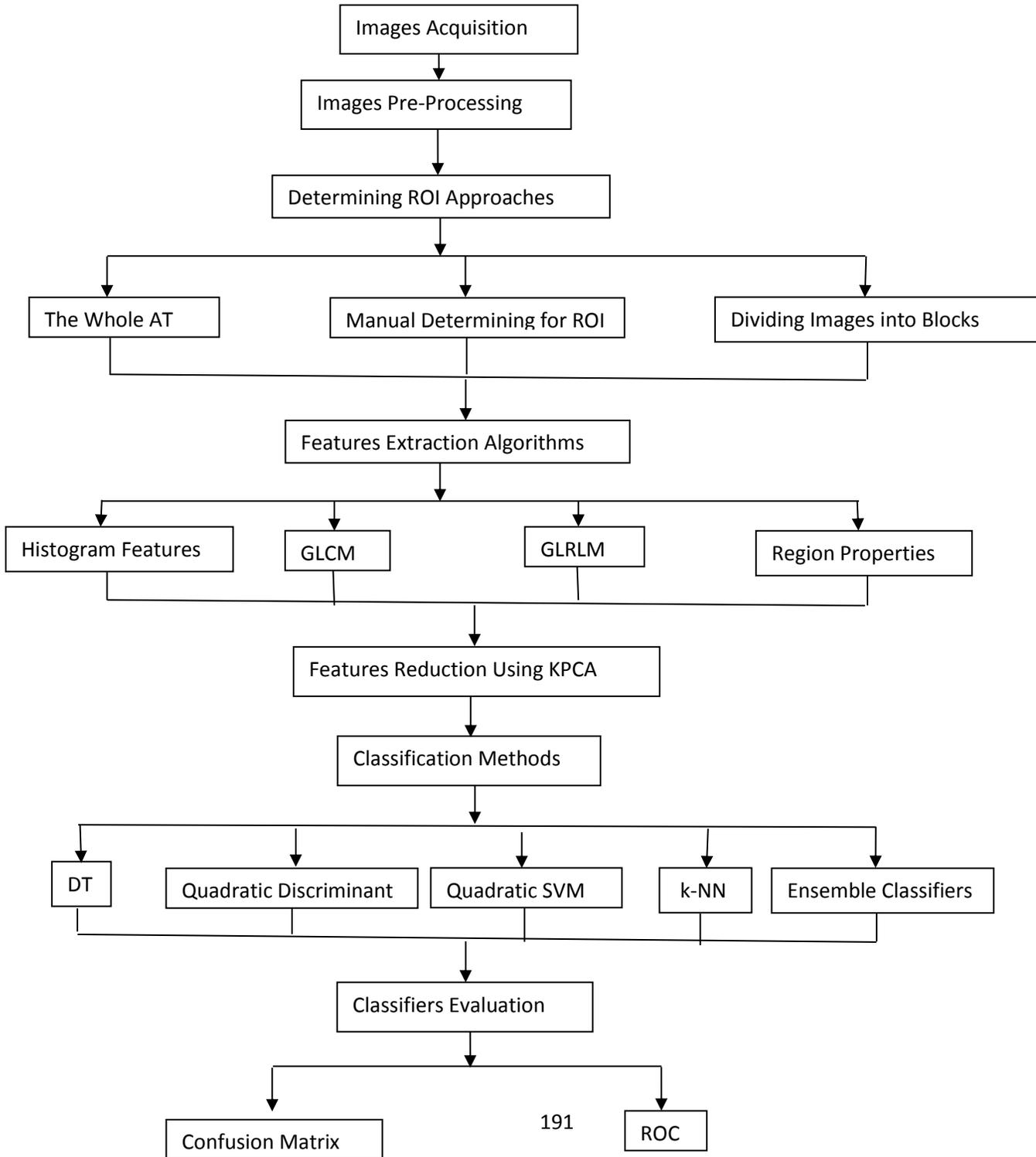
Maaten, E.O. Postma, and H.J. van den Herik. Dimensionality Reduction: A Comparative Review. Tilburg University Technical Report, TiCC-TR 2009-005, 2009.

MATLAB and Image Processing Toolbox Release 2016a, The MathWorks, Inc., Natick, Massachusetts, United States.

Uppuluri, Avinash (2008). GLCM Texture Features (<https://uk.mathworks.com/matlabcentral/fileexchange/22187-glcm-texture-features>), MATLAB Central File Exchange. Retrieved September 14, 2016.

APPENDIX

Research Steps Diagram:



PUBLICATIONS

Benrabha, J., & Meziane, F. (2017). *Automatic ROI detection and classification of the Achilles tendon ultrasound images*. Paper presented at the ACM–International Conference Proceedings Series (ICPS).