

DESIGN AND APPLICATION OF SIGNAL MODELING, SEGMENTATION AND CLASSIFICATION METHODS FOR HIGH-FREQUENCY ULTRASOUND BACKSCATTER SIGNALS

by
Noushin R. Farnoud

B.Sc., Shahid Beheshti University, Tehran, Iran

A thesis
presented to Ryerson University

in partial fulfillment of the
requirement for the degree of
Master of Applied Science
in the Program of
Electrical and Computer Engineering

Toronto, Ontario, Canada 2004
©Noushin R. Farnoud 2004

PROPERTY OF
RYERSON UNIVERSITY LIBRARY

UMI Number: EC53461

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.



UMI Microform EC53461
Copyright 2009 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Author's Declaration

I hereby declare that I am the sole author of this thesis.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Author's signature:

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Author's signature:

Borrower's Page

Ryerson University requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Name	Signature	Address	Date

Abstract

Design and Application of Signal Modeling, Segmentation and Classification Methods for High-frequency Ultrasound Backscatter Signals

©Noushin R.Farnoud 2004

**Master of Applied Science
Department of Electrical and Computer Engineering
Ryerson University**

In this study, we explore the possibility of monitoring program cell death (apoptosis) and classifying clusters of apoptotic cells based on the changes in high frequency ultrasound backscatter signals from these cells. One of the hallmarks of cancer is that the fail in the apoptosis mechanism in cells. Therefore this research carries the promise of designing more refined and more effective cancer therapies.

The ultrasound signals are modeled through the Autoregressive (AR) modeling technique. The proper model order is calculated by tracking the error criteria derived from statistical properties of the original and modeled signal. In the next stage, five machine learning classifiers are developed to classify backscatter signals based on their AR coefficients.

In clinical applications ultrasound backscatter signals from tissues and tumors are most likely to be non-stationary. Therefore analyzing such signals requires signal segmentation techniques. We developed recursive least square lattice filter for adaptive segmentation of ultrasound backscatter signals from multiple cell types into blocks of stationary segments and model and classify the segments individually.

In this thesis we demonstrate the accuracy of modeling, segmentation and classification techniques to detect signals from different cell pellets based on the signal processing and machine learning techniques.

Acknowledgment

I initially like to thank my supervisor Dr. Michael Kolios for his excellent guidance, helpful feed back and his support and encouragement throughout my research.

I am also indebted to a number of people who have contributed in different ways to the shaping of this thesis. In this respect I like to thank Dr. Sri Krishnan and Karthi Umapathy from Electrical Engineering department of Ryerson University for their help in signal processing, Dr. Sam Roweis from Machine Learning group at the University of Toronto; Kamyar Kazemi from G Kingston University and Shahad Ardalan from University of Waterloo for their engineering feedback . I express my sincere gratitude to the people at the Ontario Cancer Institute (OCI) at the Princess Margaret Hospital for their continuous help with this research specially Anoja Giles for help in preparing biological samples and Arthur Worthington for technical support of Ultrasound machines.

Last, but not least, I would like to appreciate my family and friends for their full-fledged support and understanding throughout my studies.

Dedicated to my father whose
sustenance and encouragement have
given me the strength to achieve my
goals ...

Contents

CHAPTER 1

INTRODUCTION	1
1.1. WHAT IS PROGRAMMED CELL DEATH?.....	1
1.2 APOPTOSIS AND CANCER	3
1.3. USING ULTRASOUND TO DETECT CELL DEATH	4
1.4. SIGNAL SEGMENTATION AND CLASSIFICATION FOR HIGH FREQUENCY ULTRASOUND BACKSCATTER	7
1.5. THESIS OBJECTIVES AND ORGANIZATION	8

CHAPTER 2

THEORY AND METHODS	10
2.1. SIGNAL MODELING FOR CLASSIFICATION AND SEGMENTATION	10
2.1.1. Autoregressive (AR) Modeling.....	11
2.1.2. AR Coefficient Determination	12
2.1.2.1. The Yule-walker Method to Estimate AR Coefficients.....	12
2.1.2.2. The Burg Method to Estimate AR Coefficients.....	12
2.1.3. AR Modeling Order Determination	17
2.1.3.1. Minimum AR Modeling Order	18
2.1.3.2. Maximum AR Modeling Order	19
2.1.4. AR Modeling Applied to Ultrasound Backscatter Signals	21
2.2. MACHINE LEARNING (ML) ALGORITHMS	22
2.2.1. Machine Learning Classification	23

2.2.2. Conditional Gaussian Classifier and its Application to Ultrasound Signal Classification	24
2.2.3. Fisher Linear Discriminant and its Application to Ultrasound Signal Classification	26
2.2.4. Naïve Bayes Classifier and its Application to Ultrasound Signal Classification	29
2.2.5. Neural Networks (NN) and Backpropagation Learning Algorithm.....	30
2.2.6. Application of Probabilistic Non-linear Neural Networks to Ultrasound Signal Classification.....	35
2.3. SIGNAL SEGMENTATION	35
2.3.1. Recursive Least Squares Lattice (RLSL) Filter	36
2.3.2. Application of RLSL algorithm on Ultrasound Backscatter Signals.....	39
2.4. SIMULATIONS OF ULTRASOUND BACKSCATTER SIGNALS	40
2.5. EXPERIMENTAL METHODS	42
2.5.1. Homogenous Cell Pellets	43
2.5.2. Non-homogenous Cell Pellets.....	45

CHAPTER 3

CLASSIFICATION & SEGMENTATION OF SIMULATED ULTRASOUND SIGNALS 47

3.1. AUTOREGRESSIVE MODELING OF SIMULATED ULTRASOUND BACKSCATTER SIGNALS.....	47
3.1.1. Choosing the Proper Model Order	50
3.1.2. Minimum Model Order.....	50
3.1.3. Maximum Model Order	52
3.2. CLASSIFICATION OF SIMULATED ULTRASOUND BACKSCATTER SIGNALS	55
3.3. SEGMENTATION OF ULTRASOUND BACKSCATTER SIGNALS.....	57

CHAPTER 4

CLASSIFICATION & SEGMENTATION OF EXPERIMENTAL

ULTRASOUND SIGNALS 61

- 4.1. AR MODELING OF ULTRASOUND BACKSCATTER SIGNALS FROM CELLS 62
- 4.2. EXPERIMENTAL ULTRASOUND BACKSCATTER SIGNAL OVER-FITTING 65
- 4.3. ULTRASOUND BACKSCATTER SIGNAL CLASSIFICATION 67
- 4.4. ADAPTIVE SIGNAL SEGMENTATION AND CLASSIFICATION OF ULTRASOUND
BACKSCATTER SIGNALS FROM CELL PELLETS USING THE RLSL FILTER 73

CHAPTER 5

CONCLUSIONS AND FUTURE RESEARCH 78

- SIGNAL MODELING 78
- ULTRASOUND BACKSCATTER CLASSIFICATION 79
- ULTRASOUND BACKSCATTER SEGMENTATION 82

FUTURE WORK..... 84

- SIGNAL CHARACTERIZATION..... 84
- CLASSIFICATION 84
- SEGMENTATION 85

APPENDIX A MAXIMUM LIKELIHOOD ESTIMATION FOR CONDITIONAL GAUSSIAN CLASSIFIER..... 84

APPENDIX B MAXIMUM LIKELIHOOD ESTIMATION FOR NAÏVE BAYES CLASSIFIER . 84

APPENDIX C ERROR MINIMIZATION AND OPTIMUM WEIGHT FOR NEURAL NETWORKS 90

APPENDIX D CLASSIFICATION RESULTS FOR EXPERIMENTAL ULTRASOUND BACKSCATTER SIGNALS..... 93

APPENDIX E LIST OF ABBREVIATIONS..... 95

REFERENCES 96

List of Figures

CHAPTER 1

FIG. 1.1: TIME-LAPSE MICROSCOPY IMAGES OF TROPHOBLAST CELL UNDERGOING APOPTOSIS.	2
FIG.1.2: HEMATOXYN AND EOSIN STAINS OF NORMAL AND APOPTOTIC AML CELLS.	5
FIG.1.3: A) IMAGE OF A MOUSE TUMOR ONE DAY BEFORE TREATED WITH CISPLATIN. B) THE IMAGE OF APPROXIMATELY THE SAME REGION ON DAY AFTER THE TREATMENT.	6

CHAPTER 2

FIG. 2.1: SCHEMATIC REPRESENTATION OF AR MODELING	11
FIG. 2.2: BLOCK DIAGRAM OF AR PROCESS.....	19
FIG.2.3: FISHER LINEAR DISCRIMINANT.....	28
FIG. 2.4: SCHEMATIC REPRESENTATION OF AN ANN TRAINING TASK.....	31
FIG. 2.5: SCHEMATIC REPRESENTATION OF A TWO-LAYER NEURAL NETWORK.....	32
FIG. 2.6: M-ORDER LATTICE PREDICTOR.....	37
FIG. 2.7: A THREE-LAYER CELL PELLET	46

CHAPTER 3

FIG. 3.1: A SIMULATED NORMAL AND APOPTOTIC BACKSCATTER SIGNAL.....	48
FIG. 3.2: COMPARISON OF THE AVERAGE PSD OF SIMULATED NORMAL ULTRASIUND BACKSCATTER SIGNALS MODELED BY DIFFERENT TECHNIQUES.	48
FIG. 3.3: COMPARISON OF THE AVERAGE PSD OF SIMULATED APOPTOTIC ULTRASIUND BACKSCATTER SIGNALS MODELED BY DIFFERENT TECHNIQUES.	49
FIG. 3.4: AVERAGE ENSEMBLE ERROR OF AR-MODELING.	51
FIG. 3.5: AVERAGE VARIANCE OF THE AR MODEL NOISE (ERROR).....	52

FIG.3.6: A) AVERAGE ENSEMBLE ERROR FOR 20 NORMAL AND 20 APOPTOTIC SIMULATED SAMPLES WITH ADDITIVE NOISE. B) OVER-FITTING OF ULTRASOUND BACKSCATTER SIGNALS	53
FIG. 3.7: A) AUTOCORRELATION OF THE ESTIMATED MODEL ERROR (NOISE) B) AUTOCORRELATION OF WHITE NOISE.....	54
FIG. 3.8: NEURAL NETWORK TRAINING ERROR WITH TANH AND SIGMOID ACTIVATION FUNCTIONS.	55
FIG.3.9: A SIMULATED ULTRASOUND BACKSCATTER SIGNAL COMPOSED OF THREE (NORMAL-APOPTOTIC-NORMAL) LAYERS.....	57
FIG.3.10: SEGMENTATION OF ULTRASOUND BACKSCATTER SIGNALS.....	58
 CHAPTER 4	
FIG.4.1: A) REPRESENTATION OF A “NORMAL” CELL PELLET. B) REPRESENTATION OF AN “APOPTOTIC” CELL PELLET.	61
FIG. 4.2: POWER SPECTRAL DENSITY OF AN ULTRASOUND BACKSCATTER SIGNAL	63
FIG. 4.3: AVERAGE ENSEMBLE ERROR	64
FIG. 4.5: AUTO-CORRELATION OF THE ESTIMATED MODEL ERROR (NOISE)	66
FIG.4.6: EFFECT OF NUMBER OF ITERATIONS IN TRAINING A NEURAL NETWORK.....	68
FIG. 4.7: NORMAL VS. APOPTOTIC CLASSIFICATION USING 3 ML CLASSIFIERS	71
FIG. 4.8: NORMAL VS. DECAY CLASSIFICATION USING 3 ML CLASSIFIERS	71
FIG. 4.9: APOPTOTIC VS. DECAY CLASSIFICATION USING 3 ML CLASSIFIERS	72
FIG. 4.10: : SEGMENTATION OF EXPERIMENTAL THREE-LAYER ULTRASOUND BACKSCATTER SIGNAL.....	75

List of Tables

TABLE 1.1: RELATION BETWEEN PCD AND DISEASE	2
TABLE 2.1: COMPARISON FOR BURG AND YULE-WALKER AUTOREGRESSIVE METHODS	16
TABLE 3.1: CLASSIFICATION ALGORITHMS ACCURACY FOR SIMULATED ULTRASOUND SIGNALS	56
TABLE 3.2: CLASSIFICATION ALGORITHMS ACCURACY FOR SIMULATED NAN SEGMENTED ULTRASOUND SIGNALS	59
TABLE 4.1: CLASSIFICATION ACCURACY FOR NORMAL VS. APOPTOTIC GROUPS	69
TABLE 4.2: AN EXAMPLE OF CLASSIFICATION RESULTS OF ONE RF LINE SEGMENTS FROM THE PELLET SHOWN IN FIG.4.8.....	76
TABLE D.1: CLASSIFICATION ACCURACY FOR NORMAL VS. APOPTOTIC GROUPS	93
TABLE D.2: CLASSIFICATION ACCURACY FOR NORMAL VS. DECAY GROUPS	94
TABLE D.3: CLASSIFICATION ACCURACY FOR APOPTOTIC VS. DECAY GROUPS	94

Chapter 1

Introduction

1.1. What is Programmed Cell Death?

Programmed cell death (PCD), is a genetically regulated type of cell death in which the cell uses specialized cellular machinery to kill itself. PCD can be generally categorized into three types: 1) apoptosis, 2) non-lysosomal; and 3) autophagy cell death [1-3]. Apoptosis is the most common type of cell death and is referred as PCD in many contexts. The name of this type of cell death is derived from the Greek word “ptosis” (meaning falling). It is a cell suicide mechanism that controls and eliminates cells that threaten the animal survival. This mechanism is essential to remove cells that are infected by viruses, cells of the immune system and cells with DNA damage such as cancerous cells. This process is also necessary for deleting structures and sculpting tissues in normal development (e.g., formation of the proper connections between neurons in the brain requires that surplus cells be eliminated by apoptosis). Perturbation of the signaling cascades regulating apoptosis, whether by internal or extra-cellular triggers, acquired germ line genetic mutations, or viral mimicry of signaling molecules, can result in a wide variety of human diseases. The main role of apoptosis and its relevant dysfunction results can be shown in the following table, adapted from [4]:

TABLE 1.1: Relation between PCD and disease

Apoptosis Function	Dysfunction
Deleting damaged cells	Cancer Autoimmune diseases
Controlling cell number	Neurodegenerative diseases
Deleting structures and Sculpting tissues	Developmental abnormalities

Apoptosis and necrosis, which is the localized death of cells in an organ or tissue, differ in their biochemical and morphological features. For example, necrosis is a mass cell death while apoptosis usually affects single cells or small groups of cells in an asynchronous fashion. Some of the changes of apoptosing cells are illustrated and explained in Fig. 1.1 (adopted from [5]).

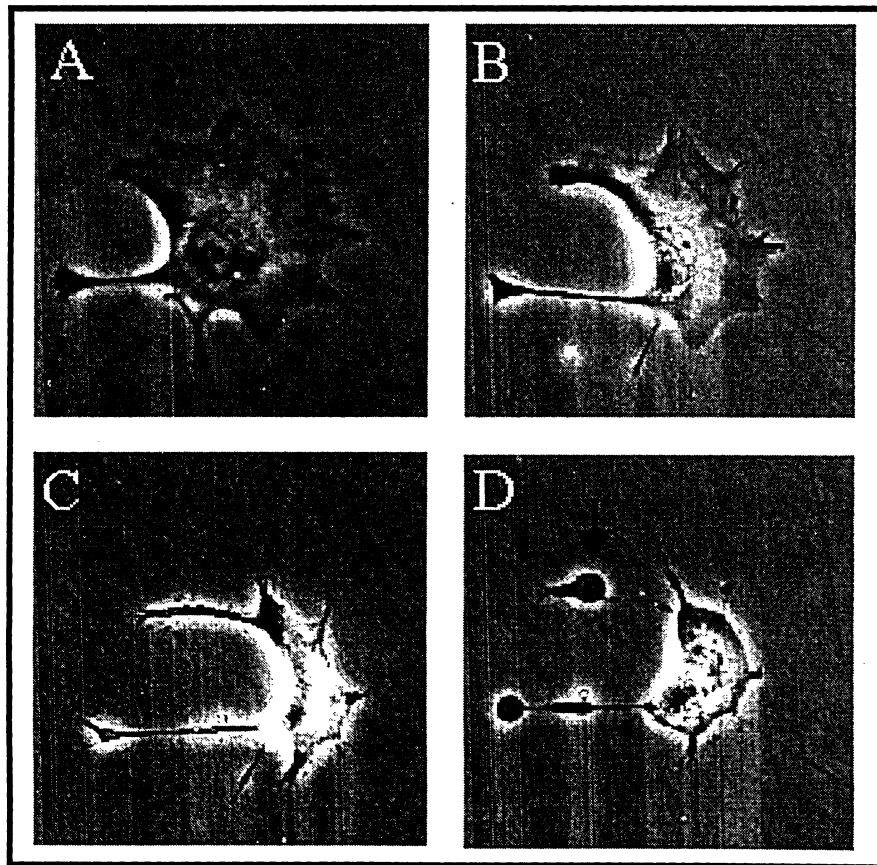


Fig. 1.1: Time-lapse microscopy images of trophoblast cell undergoing apoptosis.

Typically, the cytoplasm begins to shrink following the cleavage of lamins and actin filaments (A). Nuclear condensation can also be observed following the breakdown of chromatin and nuclear structural proteins, and in many cases the nuclei of apoptotic cells take on a "horse-shoe" like appearance (B). Cells continue to shrink (C), packaging themselves into a form that allows for easy clearance by macrophages. These phagocytic cells are responsible for removing apoptotic cells from tissues in a clean and tidy fashion that avoids many of the problems associated with necrotic cell death, such as the false up-regulation in the immunity system that can result in infection in other organs of critically ill patients. Necrotic cells swell, rupture and release their contents into the body while these changes are not seen in apoptotic cells. In order to promote phagocytosis by macrophages, apoptotic cells often undergo plasma membrane changes that trigger the macrophage response. One such change is the translocation of phosphatidylserine from the inner leaflet of the cell to the outer surface. Membrane changes can often be observed morphologically through the appearance of membrane blebs (D) or blisters which often appear towards the end of the apoptotic process. Small vesicles called apoptotic bodies are also sometimes observed (D, arrow) [5].

1.2 Apoptosis and Cancer

Resistance towards apoptosis is a hallmark of most and perhaps all types of cancer (Hanahan and Weinberg [6]). Cancer is a term of disease characterized by abnormal and uncontrolled cell division. In this disease cells exhibit genomic instability. Chromosomal rearrangements and duplications are often seen in the karyotype¹ of cancer cells. Cells normally will stop in the cell cycle if DNA is damaged through the help of several proteins have been identified that act to halt the cell cycle until DNA damage is repaired. The most known of such proteins is p53 which is a tumor suppressing gene. Cancer disrupts the apoptosis mechanism of cells by inactivating this apoptosis promoter which is normally increased in activity in response to DNA damage and leads to apoptosis (cell death) as mentioned before. Therefore the cancer cells with defect in their p53 protein pathway do not stop dividing when their DNA is damaged. This results in a pathological increase in cell number. The resulting mass, or tumor, can invade and destroy

¹ Karyotype is the chromosome profile of an individual which is useful in determining possible relationships between individuals as well as their chromosomal abnormalities and irregularities.

surrounding normal tissues by spreading of cancer cells through blood stream or lymph system and starting new cancers in other parts of the body. Based on the normal strategy of cell death (apoptosis) in response to DNA damage, radiation and chemicals are used to induce apoptosis and necrosis in many types of cancer therapies. But the main setback of such therapies is that many cancer cells develop mechanisms to inhibit the penetration of apoptosis inducing drugs into cell membrane to prevent apoptosis. For example, in the case of cancer causing viruses, one of the several human papilloma viruses (HPV) that have been implicated in causing cervical cancer produces a protein (known as “E6”) that binds and inactivates the apoptosis promoter p53; another example of such drug resistance is seen in some B-cell leukemia and lymphomas that blocks apoptotic signals they may receive by expressing high levels of Bcl-2. Therefore apoptosis-resistance mechanism is a major obstacle in many therapies that use anti-cancer agents. The other drawback of chemotherapy or other apoptosis-based cancer therapies is the suicide of normal cells which some times exceeds the number of cancerous cell death. This explains many common side effects that are seen in patients who undergo such therapies. Some of these complications are hair-loss (death of hair growing cells), pale skin color (death of melanocyte² cells).

1.3. Using Ultrasound to Detect Cell Death

Although various techniques have been developed to biochemically determine if cells undergo apoptosis, there are no techniques routinely used today that can non-invasively determine if apoptosis is occurring in cell populations [7-9]. High-frequency ultrasound (ultrasound) imaging is a powerful clinical tool that is often used in ophthalmology and dermatology. It provides high resolution images of small animal anatomy [10-12]. Ultrasound has been used for tissue classification [13], and monitoring diseases such as psoriasis³(Gupta et al.)[14], or melanoma (by Turnbull et al.) [15]. Studies have shown that in cell pellets the ultrasound backscatter signals from apoptotic acute myeloid leukemia (AML) cells differ in intensity and frequency spectrum, likely as the result of the change in size, spatial distribution and acoustic impedance of the scattering sources

² A pigment-producing cell in the skin, hair and eye that determines their color through its level of activation.

³ A chronic skin disease characterized by circumscribed red patches covered with white scales.

within the cell [16, 17]. It has been hypothesized that nucleus is the probable source of these changes in ultrasound backscatter signals. Therefore, one can assume that pulse echo data from different cell types contain distinguishable features that can be analysed using analysis techniques of the radio frequency (RF) data. These analysis techniques can be broadly classified as ultrasonic tissue characterization techniques (UTC) and have been extensively investigated for the conventional lower frequency scanners in the past. Few groups, however, have investigated the use of high-frequency ultrasound UTC, partially due to the relatively recent development of high-frequency and bandwidth ultrasound transducers and the availability of analogue-to-digital cards with sufficient speed for the very high sampling rates required for signal capture. High frequency ultrasound techniques have been successfully used in three dimensional visualization and treatment of ocular tumors (Silverman et al.[18]); analysis of different scatter morphologies in iris melanoma (Ursea et al.[19]); and characterization of myocardial edema (Dent et al. [20]).

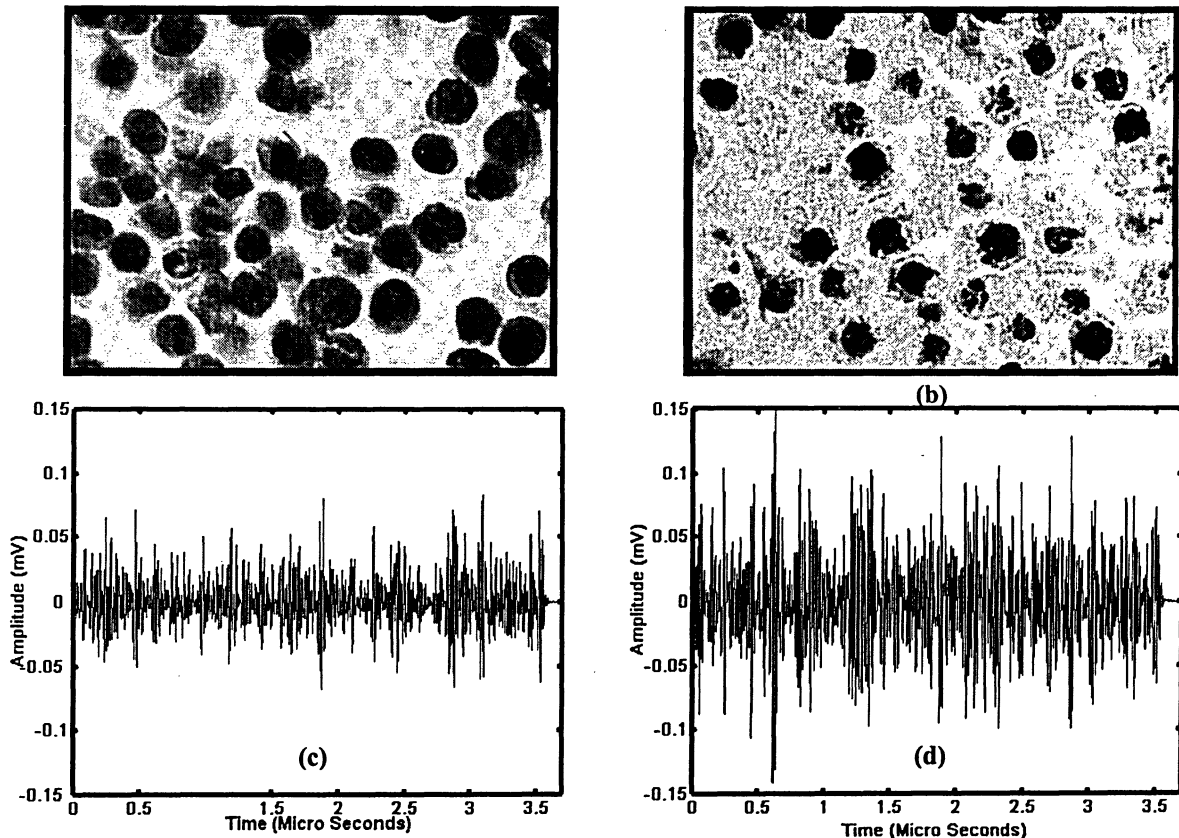
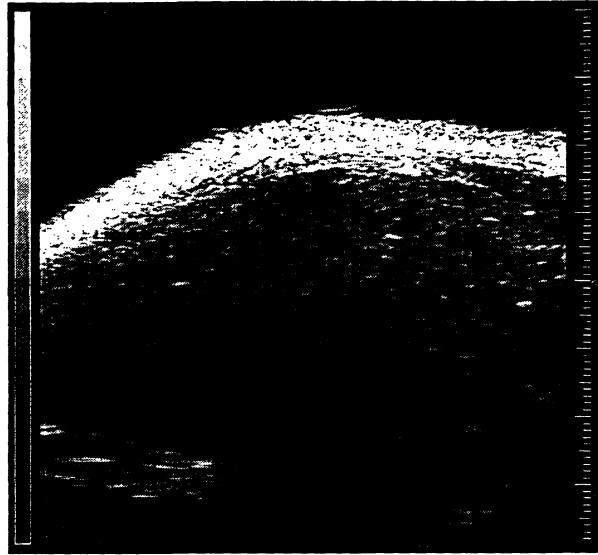
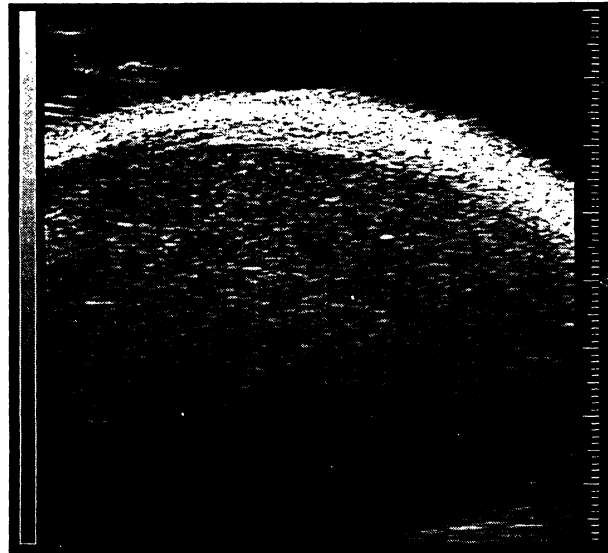


Fig. 1.2: a) Hematoxylin and Eosin (H&E) stains of normal AML cells. b) H&E stains of Apoptotic AML cells. c) Ultrasound backscatter signal from Normal cell pellet; and d) Ultrasound backscatter signal from apoptotic cell pellet.



(a)



(b)

Fig.1.3: a) Image of a mouse tumor one day before treated with cisplatin. b) The image of approximately the same region on day after the treatment.

In our experiments measured data from cell pellets the ultrasound backscatter signal from normal and apoptotic cells that are produced using the same transducer are digitized with the same 500 MHz sampling rate and used for the analysis (modeling, classification and segmentation).

1.4. Signal Segmentation and Classification for High Frequency Ultrasound Backscatter

RF backscatter analysis techniques can be used for ultrasound backscatter signal characterization and classification. Although signal processing and modeling have been widely used in biomedical applications such as EEG (Electroencephalograms) and VAG (Vibroarthrographic) signal analysis, its application for ultrasound signals was introduced later. In 1984, a research group from Philips Ultrasound Inc. and Electrical Eng. Department of the University of California published a paper which applied autoregressive (AR) modeling for characterization of ultrasound pulse echo data from two known tissue phantoms and classified the unknown pulse echo records based on their statistical model parameters [13, 21]. Similar work showed the feasibility of applying the AR techniques in modeling ultrasound pulse echo/backscatter signals. Spectral analysis techniques have also been used in image and signal processing [22, 23] that is widely used in biomedical diagnosis. An important property of biomedical signals is that many of these signals are statistically non-stationary (their statistical properties such as mean and variance change with time). This confines the application of parametric modeling such as the AR technique. Therefore, signal segmentation methods have been employed to “break” these signals into locally stationary components. To achieve this goal, adaptive signal segmentation techniques based on linear prediction were explored by Tavathia et al. [24, 25] and later by Moussavi et al.[26]. They used two dominant poles and ratio of power in the 40-120 Hz band to the total power of the segment to explore classification of knee joint vibroarthrographic (VAG) signals into four groups based on their pathological properties. In 1996 Krishnan [27, 28] developed a new adaptive segmentation method which was based on recursive least squares lattice filter. He applied this method on 35 primary VAG signals where he found an average of 8 stationary segments in each signal.

In 1960's the introduction of artificial intelligence (AI) and machine learning (ML) algorithms opened a new era in solving complex biological and medical problems due to the adaptive nature of such techniques and their ability to “learn” from a set of known samples. Based on this learning, AI and ML algorithms apply the best fit for new samples that guarantees the minimum error. For example intelligent methods have been used to construct an automated medical decision making system for identification of *tubercle*

bacilli (bacteria responsible for the tuberculosis disease) from photomicrographs of sputum smears (Veropoulos [29]). Lao et al. [30] applied ML methods to Magnetic Resonance Images (MRI) showing that brains can be correctly determined to be male or female with a successful classification rate of 97%. This proposed method also shows a high classification rate for old adults' age classification, even under difficult test scenarios. Lee et al. [31] developed "prediction rules" for predicting the degree of cell membrane disruption based on specified ultrasound parameters and measured acoustic signals. ML methods have also been used in cancer diagnosis. The results of this approach in prediction of cervical lymph node metastasis in carcinoma of the tongue have been published by Schwarzer et al. [32]. They used statistical methods (logistic regression, classification and regression tree (CART), and fuzzy inference) to predict lymph node metastasis in tongue carcinoma and compared these results that proved fuzzy inference and CART are the most accurate methods with a sensitivity of 79.2% and a specificity of 86.3%. This demonstrated that spectral analysis and machine learning techniques could be used to analyse complex biomedical problems.

1.5. Thesis Objectives and Organization

In this study high frequency ultrasound has been used to detect the structural changes cells and tissues undergo during cell death. In this thesis, I investigate whether parametric modeling techniques can characterize normal and apoptotic cancerous cells by tracking the statistics of the ultrasound backscatter signals from cell pellets. The AR method for time series modeling of ultrasound backscatter signals is used. In order to determine whether the AR technique is a suitable model to analyse ultrasound backscatter signals and choose the best algorithm to calculate the AR coefficients, the power spectral density (PSD) of the modeled signal is used as an estimator to show the similarity between the spectrum of the original and AR-modeled backscatter signals (Section 2.1.1). The proper AR model order is chosen based on statistical tests that determine the minimum and maximum range (Section 2.1.3) of the order. To determine the best model order the correlation of modeled and original signal is evaluated to show the similarity in PSDs of the two signals. In order to classify ultrasound backscatter signals, several linear and non-linear machine learning classifiers (Fisher Linear Discriminant, Conditional Gaussian

Classifier, Naive Bayes Classifier and Neural Networks with nonlinear activation functions) were implemented. The AR model coefficients, which are assumed to contain the main statistical features of the signal, are passed as the input to these classifiers and the results are shown in Chapter 3, Section 6. These classifiers were trained and tested with 15 AR coefficients from the modeled backscatter signals from each group that represent the variations of PSD of each group of signals. Section 2.2 explains the structure and mathematics behind each of these classifiers. In Chapter 2, the basic concepts of AR modeling (Sections 4 and 5) and Machine Learning algorithms that are used for classification (Sections 6-10) are explained.

The concept of signal segmentation and its application for biomedical ultrasound signals was implemented, in order to divide the non-stationary backscatter from different tissue layers into stationary segments and model them individually. To achieve this, an adaptive signal segmentation method was used (recursive least squares lattice filter), as explained in Section 2.3.1. The results of applying this method to a three layer normal-apoptotic-normal simulated and experimental system are shown in Sections 3.3 and 4.4.

Using ML algorithms with the coefficients that are derived from AR modeling of the signals as the input, I was able to show that non-linear ML classifiers such as probabilistic neural networks with sigmoid activation function provide the best accuracy in classifying normal and apoptotic cells.

Therefore, the parameters derived from modeling and classification can provide information that may lead to the diagnosis of various tissue pathologic states. The ability of ultrasonic tissue characterization techniques to extract this information about the changes in the physical characteristics of ultrasonic scatters has the potential to monitor apoptosis in population of cells with a high specificity. Therefore this study carries the promise of more refined and more effective cancer therapies.

Chapter 2

Theory and Methods

2.1. Signal Modeling for Classification and Segmentation

Biomedical signals contain large quantities of data that usually include redundancies. Therefore analyzing these data requires large scale data processing that is usually very complex. In such situations signal modeling can help to filter out the irrelevant information carried by the signal. In other words, signal modeling is the framework for signal processing. Using a modeling-based approach to perform signal processing by developing and manipulating a model of the signal source, provides a logical, coherent basis for recognizing signal types and tackles the special challenges posed by biomedical signals—including the effects of noise, changes in basic properties, or the fact that these signals contain large stochastic components.

To accomplish ultrasound backscatter signal classification, modeling is important to obtain data compression and simplify biomedical ultrasound signal analysis. By using signal modeling techniques classification can be performed with only a reduced number of model parameters instead of handling a complete series of time samples. This brings about simplicity in both computation and analysis aspects. There are different signal modeling techniques which are widely used in various applications such as moving average (MA), AR, and autoregressive moving average (ARMA) modeling techniques. AR modeling, which is a parametric spectral estimation method, is an important model that is widely used in biomedical signal processing ([33], Chapters 4 and 9), and speech signal processing [34, 35] and is explained in the following section.

2.1.1. Autoregressive (AR) Modeling

The goal of AR modeling is to describe the distribution of the power contained in a signal over a frequency range based on modeling the original signal modeled as the output of a linear system driven by white noise (Fig. 2.1).

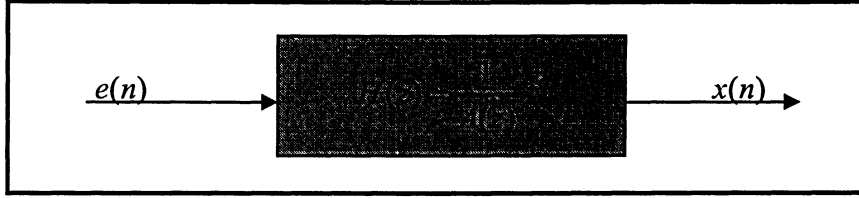


Fig. 2.1: Schematic Representation of AR modeling

In Fig. 2.1, $x(n)$ is the sequence of time sample data that are modeled as the output of a casual⁴ all-pole discrete filter excited at the input by white Gaussian noise⁵(WGN), $e(n)$, that is also referred as “excitation noise”. White noise is an idealized form of noise whose PSD is independent of the operating frequency. This noise is an inherited part of AR modeling, necessary to ensure that $x(n)$ is a wide sense stationary (WSS) random process. The effect of the filter is to “color” the white noise so as to model PSDs with several resonances [36]. The output of the AR filter and its relevant PSD are defined as following:

$$x(n) = -\sum_{k=1}^p a_k x(n-k) + e(n) \quad (2.1)$$

$$P_{xx}(f, \mathbf{a}) = \frac{\sigma^2}{|A(f)|^2} \quad (2.2)$$

$$A(f) = 1 + \sum_{m=1}^p a_m e^{-2\pi j \cdot f \cdot m}$$

Where $x(n)$ denotes AR process of order p , $e(n)$ is excitation noise that is also known as modeling error (we will refer to this error in Section 2.1.3). In Eq. (2.2), P_{xx} is the PSD estimate of AR process defined by the sampling frequency f_m , excitation white noise

⁴ A "causal" transformation uses only previous samples of the input or output signals.

⁵ White noise has equal power per hertz over the specified frequency band.

variance σ^2 and vector of AR parameters: $\mathbf{a} = [a_1, a_2, \dots, a_p]$. From Eq. (2.2) it can be concluded that the AR model is capable of producing a variety of PSDs depending on the choice of AR filter parameters. This important characteristic of AR modeling has made this technique suitable for high-resolution spectral estimation applications [37]. The AR methods tend to adequately describe spectra of data that is "peaky," that is the data whose PSD is large at certain frequencies. This makes AR modeling suitable for many applications such as speech or biomedical signal processing. Analyzing individual RF lines from ultrasound backscatter signals from cells reveals that the RF lines tend to have "*peaky spectra*" (especially backscatter from apoptotic cells). This suggests the AR modeling is an appropriate technique for characterizing these signals [13]. As explained above, AR method is another representation of the PSD by first estimating the parameters (coefficients) of the linear system that hypothetically generates the signal. The AR modeling leads to a system of linear equations which is relatively simple to solve from a computational perspective.

To apply the AR technique for modeling ultrasound backscatter signals, the all-pole filter of Fig. 2.1 models the signal from normal and apoptotic cells with the initial transmitted ultrasound pulse as the excitation noise. The order of this model is determined using the criteria explained in Section 2.1.3.

2.1.2. AR Coefficient Determination

It is evident from Eq. (2.1) that AR coefficients have a direct role in filter output prediction at each time sample (they act as weights for previous samples). These coefficients are used to estimate the PSD of the output AR process. Due to the significant role of AR coefficients in modeling accuracy, a proper algorithm must be deployed to calculate these. Once the AR coefficients of the model, (e.g., model of ultrasound backscatter signals from normal and apoptotic cells) are calculated they are fixed for the rest of analysis, new values can be compared to these parameters. For example in classification, each new signal is modeled and compared to the fixed model parameters (estimated from training data) to determine its proper class. There are various number of algorithms designed to find the best coefficients for AR modeling. Some of these

methods are the Yule-Walker [38-40], Covariance [27, 38], Burg [27, 41, 42], Modified Covariance [43, 44], Least-squares and Cholesky decomposition methods [43, 45]. In the Sections 2.1.2.1 and 2.1.2.2 two of the most common of these methods, Yule-walker and Burg, are explained in more detail.

2.1.2.1. The Yule-Walker Method to Estimate AR Coefficients

The Yule-Walker method [39, 40] for spectral estimation computes AR coefficients by calculating the autocorrelation function of the signal and solves the least-squares minimization problem to estimate the model AR coefficients [39, 40]. It uses the following formulas to calculate covariance of AR process:

$$r(k) + \sum_{i=1}^n a_i r(k-i) = 0, \quad \text{for } k > m \quad (2.3)$$

$$r(0) + \sum_{i=1}^n a_i r(-i) = \sigma^2, \quad \text{for } m = 0 \quad (2.4)$$

Where $r(k)$ is the auto-covariance sequence (ACS) at lag k and a_i is the i -th AR parameter. Combining equations (2.3) and (2.4) gives the following system of linear equations (the asterix indicates *complex conjugation*):

$$\begin{pmatrix} r(0) & r^*(1) & \dots & r^*(n) \\ r(1) & r(0) & & \vdots \\ \vdots & & \ddots & r^*(1) \\ r(n) & \dots & & r(0) \end{pmatrix} \begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (2.5)$$

Eq. (2.5) is called the Yule-Walker or Normal equation and is the basis of many other methods to find AR coefficients. In this method we first obtain the sample covariance from the given data (for $k=0$ to n) and form the autocovariance matrix and solve (2.5) for the vector of AR parameters:

$$\mathbf{a} = [a_1, a_2, \dots, a_n]^T \quad (2.6)$$

By using all but the first row of (2.5) the AR vector parameters (**a**) can be estimated using the following equation:

$$\begin{aligned} \mathbf{r}_n + \mathbf{R}_n \mathbf{a} &= 0 \\ \Rightarrow \mathbf{a} &= -\mathbf{R}_n^{-1} \cdot \mathbf{r}_n \end{aligned} \quad (2.7)$$

Where the covariance matrix in (2.7) \mathbf{R}_n is positive definite for any number of n and hence is invertible⁶. This property ensures that the solution for Eq. (2.5) exists and is unique [46]. Once the AR coefficients are calculated, the system gain σ^2 (in Eq. 2.4) can also be estimated using the first row of Eq. (2.5). The Yule-Walker method function in Matlab takes the advantage of the Hermitian⁷ Toeplitz⁸ structure of the autocorrelation matrix and uses Levinson-Durbin algorithm [47] to recursively solve (2.5) for estimated AR coefficients [48].

2.1.2.2. The Burg Method to Estimate AR Coefficients

The Burg algorithm is probably the most widely known method to estimate AR coefficients. This method was first introduced by Burg in 1967 [49, 50]. It uses a lattice filter and due to its derivation in the context of maximum entropy methods, the algorithm is sometimes designated as "MEM". The Burg method is based on minimizing the average estimates of forward and backward prediction error powers while satisfying the Levinson-Durbin equation and computes the AR coefficients directly from the data by estimating the reflection coefficients (partial autocorrelations) at successive model orders (k_m) [51, 52]. One main advantage of Burg method is producing stable AR coefficients (because $|k_m| < 1$ therefore the roots are inside or on the unit circle).

Assume we have data measurements $\{x(t)\}$ for $t = 1, 2, \dots, N$. We define the forward and backward errors as:

$$\hat{e}_{f,p}(t) = x(t) + \sum_{i=1}^p \hat{a}_{p,i} x(t-i) \quad (2.8)$$

⁶ A symmetric matrix \mathbf{R} is positive definite if it is true that for every non-zero vector \mathbf{x} , the real number which results from the multiplication $\mathbf{x}^* \mathbf{R} \mathbf{x}$ is positive.

⁷ A square complex matrix \mathbf{R} is Hermitian if it is equal to its complex conjugate transpose.

⁸ Toeplitz matrix is a matrix in which all the diagonal and sub-diagonal elements are the same.

$$\hat{e}_{b,p}(t) = x(t-p) + \sum_{i=1}^p \hat{a}_{p,i} x(t-p+i) \quad (2.9)$$

Where $e_f(t)$ and $e_b(t)$ denote forward and backward errors respectively and p is the model order. The estimated AR coefficient $\hat{a}_{p,i}$ at i -th level is defined as (we use hats to express estimated quantities):

$$\hat{a}_{p,i} = \begin{cases} \hat{a}_{p-1,i} + \hat{\Gamma}_p \hat{a}_{p-1,p-i}^* & i = 1, \dots, p-1 \\ \hat{\Gamma}_p & i = p \end{cases} \quad (2.10)$$

Where p denotes the model order and $\hat{\Gamma}_p$ is the estimated reflection coefficient at order p . This recursive algorithm calculates AR coefficients at order p based on the assumption that the AR coefficients for order $p-1$ have been previously computed. To estimate AR coefficients from Eq. (2.10), the Burg method minimizes the arithmetic mean of forward ρ_f and backward ρ_b errors known as “performance index” or ξ (Eq. 2.13). This minimization relies on finding the optimum values for reflection coefficients are as the following:

$$\xi = \min \frac{1}{2} [\hat{\rho}_f(p) + \hat{\rho}_b(p)] \quad (2.11)$$

Where

$$\hat{\rho}_f(p) = \frac{1}{N-p} \sum_{t=p+1}^N \left| \hat{e}_{f,p}(t) \right|^2 \quad (2.12)$$

$$\hat{\rho}_b(p) = \frac{1}{N-p} \sum_{t=p+1}^N \left| \hat{e}_{b,p}(t) \right|^2 \quad (2.13)$$

The prediction errors can be expressed by the following recursive equations:

$$\hat{e}_{f,p}(t) = \hat{e}_{f,p}(t) \hat{e}_{f,p-1}(t) + \hat{\Gamma}_p \hat{e}_{b,p-1}(t-1) \quad (2.14)$$

$$\hat{e}_{b,p}(t) = \hat{e}_{f,p-1}(t-1) + \hat{\Gamma}_p^* \hat{e}_{f,p-1}(t) \quad (2.15)$$

Where $\hat{e}_{f,p}(t)$ and $\hat{e}_{b,p}(t)$ denote estimated modeling forward and backward errors with modeling order p at time t respectively. Since the performance function is quadratic, we can find its minimum by differentiating with respect to reflection coefficients and

assigning this differentiation to zero ($\frac{d\xi}{d\Gamma_m} = 0$). This yields to the optimum reflection coefficients by:

$$\hat{\Gamma}_p = \frac{-2 \sum_{t=p+1}^N \hat{e}_{f,p-1}(t) \hat{e}_{b,p-1}^*(t-1)}{\sum_{t=p+1}^N \left[|\hat{e}_{f,p-1}(t)|^2 + |\hat{e}_{b,p-1}(t-1)|^2 \right]} \quad (2.16)$$

The magnitude of reflection coefficient expressed in Eq. (2.16) is less than one which confirms its stability. The table below shows the main features of different AR algorithms (adapted from [48]).

TABLE 2.1: Comparison for Burg and Yule-Walker Autoregressive Methods

	Burg	Yule-Walker
Characteristics	Does not apply window to data	Applies window to data
	Minimizes the forward and backward prediction errors in the least squares sense, with the AR coefficients constrained to satisfy the L-D recursion	Minimizes the forward prediction error in the least squares sense (also called "Autocorrelation method")
Advantages	High resolution for short data records	Performs as well as other methods for large data records
	Always produces a stable model	Always produces a stable model
Disadvantages	Peak locations highly dependent on initial phase	Performs relatively poorly for short data records
	May suffer spectral line-splitting for sinusoids in noise, or when order is very large	Frequency bias for estimates of sinusoids in noise
	Frequency bias for estimates of sinusoids in noise	
Conditions for No singularity		Because of the biased estimate, the autocorrelation matrix is guaranteed to positive-definite, hence nonsingular

The Burg method has several advantages which make it suitable for finding AR coefficients in many applications. Computational simplicity and efficiency is among these advantages. Another advantage of this technique is its capability to calculate AR coefficients with different model orders, as we can simply add or delete lattice stage(s) in recursion part of the algorithm. This “*modularity*” property of the Burg algorithm facilitates filtering the data with different model orders to find the best lattice order. This method also has a very good accuracy in estimating short data records in which the AR PSD estimates are very close to the true power values [48]. The accuracy of the Burg method is lower for high-order models, long data records, and high signal-to-noise ratios (which can cause *line splitting*, or the generation of extraneous peaks in the spectrum estimate). The spectral density estimate computed by the Burg method is also susceptible to frequency shifts (relative to the true frequency) resulting from the initial phase of noisy sinusoidal signals. This effect is magnified when analyzing short data sequences.

2.1.3. AR Modeling Order Determination

The modeling order (p) in Eq. (2.1) controls the error associated with the AR signal approximation. This parameter determines the number of previous samples that are used to predict a new sample in AR modeling technique. A small model order can not capture the main properties of the signal to be modeled, such as signal frequency or mean, and therefore it can not represent the original signal. A high model order also causes two problems: 1) it results in large-scale data processing which can be inefficient and 2) the model becomes biased towards the training signal by modeling the noise associated with the signal, and therefore it will be a weak model for testing data by having a high testing error rate (this situation is known as “over-fitting”⁹). In such a situation it is expected that testing error curve increases as we shift the model order above over-fitting boundary. The former problem is related to limitation of resources availability such as hardware facilities in storing and processing signals in digitized format that is important from practical point of view. The latter problem (over-fitting) may take place in the classification phase if the designed intelligent classifier results in poor testing accuracy because of a large number of training parameters [53-56]. Therefore a careful study on the over-fitting criterion is

⁹For more information on “over-fitting” visit: www.dmreview.com/whitepaper/WID404.pdf.

essential before proceeding with AR modeling. A common method for estimating the proper model order is the Akaike Information Criterion (AIC) [57], but applying this method would be very difficult in this work due to the peaky nature of ultrasound backscatter signals (in the AIC approach the proposed number for AR model order is twice of the number of peaks in the signal) . Instead, a novel approach was taken which uses statistical parameters to compare the correlation of original and AR-modeled signals to determine minimum model order.

2.1.3.1. Minimum AR Modeling Order

There are two methods that we use to determine the lower bound for AR modeling order: average ensemble error between the original and modeled signal spectra, and the variance of the error associated with the modeling.

The error denoted by $e(n)$ in Eq. (2.1) shows the total difference of original and AR-modeled signal. As we are interested in the frequency content of the signal rather than its value in time domain, we define the error as:

$$E = \sum_{n=1}^N |\rho_{xx}(n) - \hat{\rho}_{xx}(n)| \quad (2.17)$$

Where ρ and $\hat{\rho}$ represent the PSD of original and estimated modeled signals respectively. The modeled (estimated) signal with order p is calculated as:

$$\hat{x}(n) = -\sum_{k=1}^p a_k x(n-k) \quad (2.18)$$

The total ensemble error (Eq. 2.18) is calculated using the average Euclidean distance between the PSDs of the two signals. The PSD of the original signal is calculated using the “Welch” method (“pwelch” function in Matlab version 6.5). To calculate the PSD of the modeled signal, first the AR coefficients of the original signal are estimated and used as the poles of an AR filter (all-pole filter) excited by a WGN as its input. The output of this filter is the “*reconstructed signal*” and its PSD is estimated with the same method used for the original signals (Welch method).

The second approach estimates the variance of the AR modeling noise. As is seen in Fig. 2.1, the AR process is the output of an all-pole filter invoked by a white noise $e(n)$. This noise, which is also the difference between original and estimated signals or modeling error (Eq. 2.1 and 2.18), can be viewed as the output of the prediction error filter shown in Fig. 2.2, where $x(n)$ is the original signal and $A(z)$ is the transfer function with order p , defined by:

$$A(z) = 1 + \sum_{k=1}^p a_k z^{-k} \quad (2.19)$$

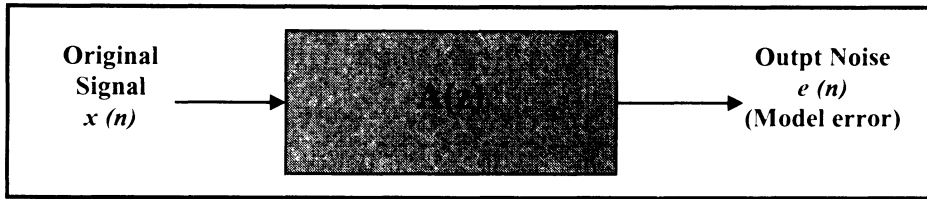


Fig. 2.2: Block diagram of AR process

Therefore it is expected that after estimating the coefficients of an AR-model, if the same filter as Fig. 2.2 is designed with a reconstructed signal estimated by Eq. (2.18) as the filter input and the filter transfer function defined by Eq. (2.19); the filter output, $e(n)$, should be approaching WGN. This can be verified by plotting the autocorrelation of the filter output. For WGN the autocorrelation is zero in all lags except a spike to 1 at zero lag. This will be used for data analysis in Chapters 3 and 4.

2.1.3.2. Maximum AR Modeling Order

Over-fitting of ultrasound backscatter signals takes place when the modeling order is high and therefore the noise associated with the signal is also modeled. Due to the randomness of this noise and despite excellent results from training data, the model usually generates high error rate on testing data when over-fitting occurs. The main obstacle in monitoring over-fitting is the unknown nature of the noise associated with the digitized ultrasound

backscatter signals. The consistency of our model to the clinical experiments can not be verified unless the nature and type of this noise is well understood.

There are several sources which affect this noise such as instrumental noise and external RF interference. Hence this noise can not be predicted and modeled. To study over-fitting in our experiments, random white noises with different variances were generated and added to the simulated ultrasound backscatter signal (additive noise). We used signal-to-noise ratio (SNR) as an estimate of the noise level associated with our signal and by using this approach we evaluated the dependency of over-fitting to the “noise level” of the signal.

The SNR which is usually taken to indicate an *average* signal-to-noise ratio is the ratio between the magnitude of a signal (meaningful information) and the magnitude of background noise. The SNR in decibels is 20 times the base-10 logarithm of the amplitude ratio, or 10 times the logarithm of the power ratio¹⁰. Because both the ultrasound backscatter and noise signals are *real* signals, we used the following property of their variance to estimate the power (Eq. 2.20). Where σ , P and μ indicate the variance, power and mean of signal x . It is evident from Eq. (2.20) that for the *white noise* where the mean is 0, the power would be equal to the noise variance.

$$\sigma_x^2 = P_x - \mu_x^2 \quad (2.20)$$

The noisy signal (with different SNRs) forms the training data of the AR method that is modeled using different orders; however the information carrying data is the ultrasound signal excluding the noise. Therefore, first the coefficients of the AR model of the noisy signal were calculated using different model orders and then the noisy signal was reconstructed using a linear prediction filter with the estimated AR coefficients as its linear prediction coefficients¹¹. The Euclidean distance of the modeled noisy and original *noise-less* signals are calculated with different model orders ranging from 1-100. This parameter forms the *testing error* which explains what model order captures the features

¹⁰ Due to the definition of decibel the SNR gives the same result independent of the type of signal which is evaluated (power, current, voltage).

¹¹ The linear prediction coefficients are equal to negative of AR coefficients and the first AR coefficient is set to zero.

parameter forms the *testing error* which explains what model order captures the features of the noise added to the original signal (in this case the difference between modeled noisy and noise-less signal is expected to increase). The results of over-fitting curves are shown in Section 3.1.3. The best modeling order for the AR method is the one that inhibits over-fitting in both of the classes that are being classified, although in some occasions this may lead to a higher training error for one class (a smaller order may be chosen to prevent one model to over-fit) and therefore the best approach is to consider a model order that achieves a relatively low error for both training and testing steps.

2.1.4. AR Modeling Applied to Ultrasound Backscatter Signals

Modeling and monitoring of ultrasound backscatter signals can help us to understand the state of cells under various conditions. Moreover it reduces the complexity of classification algorithms for diagnosis of cancerous cells (Section 2.4). Previous work has shown that AR modeling is useful for analysis of a variety of biomedical signals as well as their classification [58-66]. In modeling ultrasound backscatter signals using the AR technique, the all-pole filter (Section 2.1.1) models the signal from normal and apoptotic cells with the initial transmitted ultrasound pulse as the excitation noise. To compare different AR modeling techniques, the PSD of the modeled signals can be used as a good estimator to show the similarity between the spectrum of the original and AR-modeled ultrasound backscatter signals. The goal of PSD estimation is to describe the distribution of the power contained in a signal over its frequency band and integrating with respect to frequency will yield to average power of the signal. The results of comparing the power spectral densities of AR-modeled and original signals for simulated and experimentally measured ultrasound backscatter signals are discussed in Sections 3.1 and 4.1. The ultrasound backscatter signal is generated by a 40 MHz or 20 MHz ultrasound transducer with the signal sampled at 500MHz. The Burg method was chosen to find the AR model parameters of ultrasound backscatter signals from cells due to the overall advantages of this method over other AR estimation algorithms for ultrasound backscatter signal (Table 2.1).

2.2. Machine Learning (ML) Algorithms

Machine learning (ML) is an adaptive and robust approach in solving problems based on learning the knowledge “contained” within data examples. It often uses a probabilistic approach which is a well-known method in dealing with signals that include large amounts of data to be processed and therefore hand programming or conventional computer programming is either impossible or time consuming. Moreover ML algorithms can characterize a data set with strong statistical regularity. This is very useful in many biomedical applications; i.e. for biomedical ultrasound signals that retain some strong statistical properties but their nature is not well-understood. In the ML approach the structure of the program is written in a way that the computer can tune many internal parameters which are required to solve the problem (such in classification) in an accurate and efficient way. There are two phases in ML algorithms: “learning” or “training” the system with known data and “testing” where the system performance is tested with new (untested) data. The learning process can be categorized in to 4 general classes:

- a) **Supervised learning:** In supervised learning, the examples of the inputs and corresponding desired outputs are given to the algorithm and based on that, it is trained to predict outputs on future inputs. Examples of such learning are classification and regression [67-71].
- b) **Unsupervised learning:** In unsupervised learning, the algorithm automatically discovers representations, features, structures and many other specifics of the data. Examples of such learning are clustering, outlier detection and compression [72-79].
- c) **Reinforced learning:** In reinforced learning, given a sequence of inputs, actions from a fixed set, and scalar rewards/punishments, the system learns to select actions sequence in a way that minimizes expected cost [80, 81]. Example of such learning is its application in improving image compression techniques [82].

application is “wrapper¹² learning” (the problem of learning website wrappers from examples) where this method is used to design a system that can exploit several different representations of a document [88].

2.2.1. Machine Learning Classification

ML classification is a supervised learning where outputs (y) are the categorical parameters or class labels and inputs (x) can be the complete data or just some data features. The goal of ML classification is to select the correct class label for new input data.

There are several ML algorithm used for classification purposes such as the conditional Gaussian, neural network (NN), Fisher linear discriminant, kernel analysis and principal component analysis. These algorithms use one of the two basic approaches to solve a classification problem: the Generative or Discriminative approach.

In the **Generative** approach, the classification problem is modeled by the joint probability defined by:

$$p(\mathbf{x}, y) = p(y) \cdot p(\mathbf{x} | y) \quad (2.21)$$

Where $p(y)$ is the class prior and $p(\mathbf{x} | y)$ is the class conditional feature distribution. Maximum probability of the new data belonging to each class is calculated using the Bayes rule to expand conditional probability of Eq. (2.20).

In the **Discriminative** approach the discriminant function which directly determines the class boundaries is estimated and modeled as $f(y | \mathbf{x})$ to find the probability that the new data vector (\mathbf{x}) belongs to class (y). Examples of such algorithms are: linear discriminant analysis [89-91], logistic regression [92, 93], and support vector machines [94-96]. In the following two sections some of these algorithms and their analysis are described with the focus of their application for ultrasound backscatter signal classification from cells.

¹² Wrapper is a program that makes an existing website look like a database.

To determine the accuracy of classification algorithms, ML classifiers are trained with different sets for training and testing for both simulated (Section 2.4) and experimental ultrasound backscatter signals (Section 2.5). The simulated ultrasound samples were generated using the algorithm explained in Section 2.4 with 270 scatterers, a central frequency of 40MHz, a mean scatter spacing of $10 \mu m$; and spacing variance of $1 \mu m$ for normal and $5 \mu m$ for apoptotic cells, respectively. The increased variance for the apoptotic cell spacing is to model the observed less regular cell spacing when compared to the normal cells.

The simulated and experimental backscatter signals were first modeled using an AR technique with order 15, and these 15 coefficients were passed as the input of machine learning classifiers for training. In the testing phase each testing sample signal is also modeled with the same order and these model coefficients are sent as the testing data to the machine learning algorithms. The specification of initializing, training and testing of these algorithms are shown in the following section.

2.2.2. Conditional Gaussian Classifier and its Application to Ultrasound Signal Classification

This generative classifier fits a Gaussian function to each class and then finds its maximum likelihood fit to the data. The maximum likelihood fit of a Gaussian to some data is also a Gaussian whose mean is the data mean and its covariance is equal to the data covariance [56]. Equations (2.21)-(2.25) show the derivations for class-conditional Gaussian classifier for data vector $\mathbf{x} = \{x_1, x_2, \dots, x_D\}$ with independent features and discrete class label $y \in \{1, 2, \dots, K\}$:

$$p(y = k) = \alpha_k \quad (2.22)$$

$$p(\mathbf{x} | y = k) = (2\pi \sigma^2)^{-D/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^D (x_i - \mu_{ki})^2 \right\} \quad (2.23)$$

The maximum likelihood is calculated by differentiating the maximum likelihood equation with respect to μ_k and assigning this differentiation to zero (Appendix A). Once shared variance and shared mean are calculated the classification can be performed by evaluating $p(y|x)$ using Bayes Rule and Eq. (2.23):

$$p(y|x) = \frac{p(x|y)p(y=k)}{\sum_k p(x|y=k)p(y=k)} \quad (2.24)$$

The application of this classification technique lacks in generality as it makes strong assumptions concerning underlying distributions and is more appropriate when distributions are known and match the Gaussian assumption. Still it is successfully used in many applications such as biomedical signal classification and speech recognition [53, 97-100]. The classification results of simulated and experimental ultrasound backscatter signals using this technique are shown in Chapter 3 and 4.

To apply this algorithm for ultrasound backscatter signal classification, a Gaussian classifier was trained with a 3 dimensional matrix in which each column represents a vector of training data (15 AR coefficients) and the third index of the matrix is either 1 or 2 denoting normal and apoptotic class labels respectively ($k=2$). The training matrix had 50 samples of simulated and 100 samples of experimental ultrasound backscatter signals from two groups of normal and apoptotic signals. The second matrix with the same structure was generated for testing data with 50 samples for simulated and 100 samples for experimental ultrasound backscatter signals from each group of normal and apoptotic cells. The priors for both normal and apoptotic classes were set equal (0.5) meaning that each new data had an equal probability of belonging to either of the classes. The Gaussian classifier which tries to fit a Gaussian function to each class of normal and apoptotic signals, uses the shared variance of training data (Eq. (A.5), Appendix A) and shared mean by averaging each AR coefficients through all the samples of both classes (Eq. (A.4), Appendix A). Once these parameters are calculated the class of new testing data is determined from Eq. (2.23) (the probable class of new data is the one which

maximizes the value of Eq.(2.23)) using 50 normal and apoptotic testing samples from simulated group and 100 samples of experimentally measured signals. As for experimental case this classifier was constructed three times to implement the three classification problems: normal-apoptotic classification (already explained) and normal-decay and apoptotic-decay classification.

2.2.3. Fisher Linear Discriminant and its Application to Ultrasound Signal Classification

The linear discriminant method which is a discriminative classifier projects high-dimensional data in to a line and performs classification in a one-dimensional space to simplify classification in linear Bayes decision boundary [101, 102]. The Fisher linear discriminant performs this projection in the direction of maximum class separation by maximizing the distance between the means of the two classes while minimizing the variance within each class (Fig. 2.3). This projection defines features that are optimally discriminating. The underlying goal of Fisher's classification function is to minimize within group variance and maximize the differences between groups. Fisher linear Discriminant is based on maximizing ratio of cross-class scatter to within class scatter defined by:

$$J(w) = \frac{|\mu_1 - \mu_2|}{\sigma_1^2 + \sigma_2^2} \quad (2.25)$$

Where μ and σ are mean and variance of each class denoted by their subscripts. Equation (2.26) is also known as Fisher's Discriminant Function (FDF). In signal theory, this criterion is also known as the *signal-to-interference ratio*. The maximization of Eq. (2.26) yields a closed form solution that involves the inverse of a covariance-like matrix, and results in the in the following equations for the optimum weight and threshold of the Fisher criterion.

$$\begin{aligned} \mathbf{w}^* &= \Sigma^{-1}(\mu_0 - \mu_1) \\ w_0^* &= \frac{1}{2}(\mu_0 + \mu_1) - \mathbf{w}^T(\mu_0 - \mu_1) \left[\frac{\log p_0 - \log p_1}{(\mu_0 - \mu_1)^T \Sigma^{-1}(\mu_0 - \mu_1)} \right] \end{aligned} \quad (2.26)$$

Where \mathbf{w}^* and w_0^* are optimum weight vector and optimum threshold, μ and Σ are sample mean and covariance, p represents the class priori, and the subscripts denote the two classes. This method was not originally designed for the classification problem, but still provides an approach to decrease the dimension. The Fisher method is based on the assumption that each class has a Gaussian distribution with the same covariance. The purpose of the Fisher discriminant, like other linear classification algorithms, is to find a line or hyper-plane which best separates the two classes:

$$c(x) = \text{sign}[\mathbf{x}^T \mathbf{w} - w_0] \quad (2.27)$$

Where ‘ \mathbf{x} ’ is the test data vector to be classified, ‘ \mathbf{w} ’ is the *weight* vector perpendicular to decision boundary and w_0 is a scalar denoting the classification *threshold*. With this approach the value of $c(x)$ (either -1 or 1) is used to determine the location of new data with respect to the classifying hyper-plane. Although Fisher Discriminant is a linear classification method, it has been successfully used for problems that involve detecting and categorizing features such as face recognition systems [103] and damage detection feature for manufacturing purposes [104, 105].

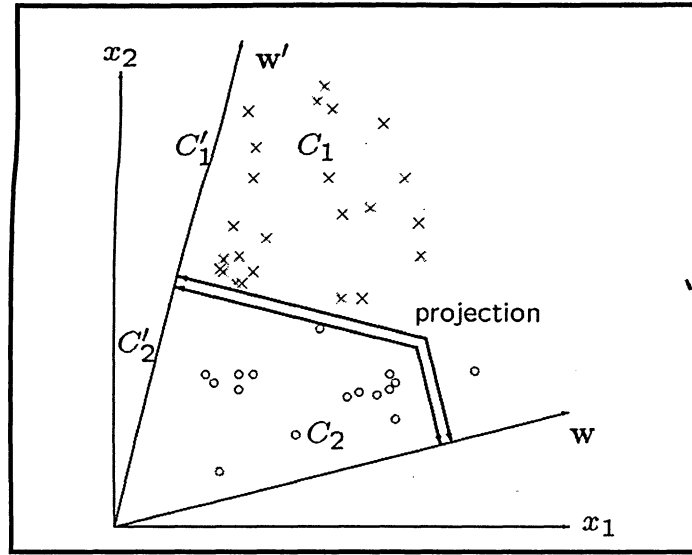


Fig.2.3: Projection on distinct vectors. The data in 2-D space is projected on two lines (discriminant functions). As seen if the data are projected on the discriminant function by weight ' w ' they two class will overlap and therefore they can not be separated. But when w' is used the data of class C_1 (denoted by ' x ') are projected on the upper part of line w' and the data of C_2 (denoted by ' o ') are projected on the bottom part of this line. Therefore the projection of C_1 and C_2 on this line gives the maximum class separation and it is chosen as discriminant function.

To classify ultrasound backscatter signals with the Fisher Linear Discriminant, the classifier was trained with the same 50 normal and 50 apoptotic training sets of simulated data and 100 samples of normal, apoptotic and decay training sets that were used in training the Gaussian classifier. The shared covariance and mean of each class (which was the mean of each AR coefficient in that class) were calculated and inserted into Eq. (2.26) to calculate optimum weight and threshold for linear Fisher's discriminant and consequently estimate the class of new data using Eq. (2.27). This defines the Fisher discriminating hyper-plane that best separates the two classes. Finally the classifier performance is tested by finding the location of the 50 normal and 50 apoptotic testing data with respect to the position of this hyper-plane in linear space. If the data falls into the right of the hyper-plane ($c(x) \geq 0$ in Eq. (2.27)), the testing data is estimated to belong to normal group and it belongs to apoptotic group otherwise ($c(x) < 0$).

2.2.4. Naïve Bayes Classifier and its Application to Ultrasound Signal Classification

The naive Bayes classifier is a generative classification algorithm which simplifies learning by assuming that features are independent given the class. It implements a simple algorithm that can achieve relatively good performance on classification tasks despite its poor assumption of data independency [106-108]. This algorithm can be used with any type of features; but it is usually used with binary features. The Naïve based approach to classify the new data is to assign the most probable target that best describes the instance $\mathbf{x} = \{x_1, x_2, \dots, x_i\}$ where $\langle x_i \rangle$ is the set of features that describe the data instance \mathbf{x} :

$$\begin{aligned} c(\mathbf{x}) &= \arg[\max_c p(c | \mathbf{x})] \\ &= \arg[\max_c p(\mathbf{x} | c) \cdot p(c)] \end{aligned} \quad (2.28)$$

In Eq. (2.28), $c(\mathbf{x})$ denotes the target value output. The two probabilities in this equation can be estimated based on the training data: $p(c)$ is easily estimated by counting the frequency that each target value “ c ” occurs in the training data. However this approach to calculate the probability is not reliable unless we have a very large training set. The Naïve Bayes classifier simplifies the estimation process by assuming that given the target value for the training data set, the probability of observing conjunction x_1, x_2, \dots, x_n is just the product of the probabilities for individual features. The details of the calculations are depicted in Appendix B.

Therefore Naïve Bayes learning involves two steps in training: calculating $p(c_i)$ and estimating $p(x_i | c_i)$ terms using their frequencies over training data set. Once learning is complete, classification of new data (testing) can be performed by using the sign of $c(\mathbf{x})$ (Eq. (B.3), Appendix B). Surprisingly, this classification algorithm despite it’s “naïve” assumption to simplify complex problems has been found effective in many practical applications, including text classification, medical diagnosis, and systems performance management [109-112]. The success of naive Bayes in the presence of

feature dependencies results from the fact that the accuracy of this method is not directly correlated with the degree of feature dependencies measured as the class conditional mutual information between the features. Instead, a better predictor of naive Bayes accuracy is the amount of information about the class that is lost because of the independence assumption and therefore an optimal classifier is obtained as long as both the actual and estimated distributions agree on the most-probable class [112].

2.2.5. Neural Networks (NN) and Backpropagation Learning Algorithm

The study of artificial neural networks (ANN) has been inspired by the nervous system of organs. The history of the ANNs stems from the 1940s, the decade of the first electronic computer. However, the first significant step took place in 1957 when Rosenblatt introduced the first concrete neural model, the “perceptron” [113]. Biological learning systems are built of very complex network of interconnected neurons and neurobiology has shown that the information-processing abilities of biological neural systems must follow from highly parallel processes operating on representations that are distributed over many neurons. Using this approach, ANNs are built out of a densely interconnected set of simple units (neurons), where each unit takes a number of real-valued inputs (possibly the outputs of other units) and produces a single real-valued output (which may become the input to many other units). Figure 2.4 shows the schematic representation of the network training task.

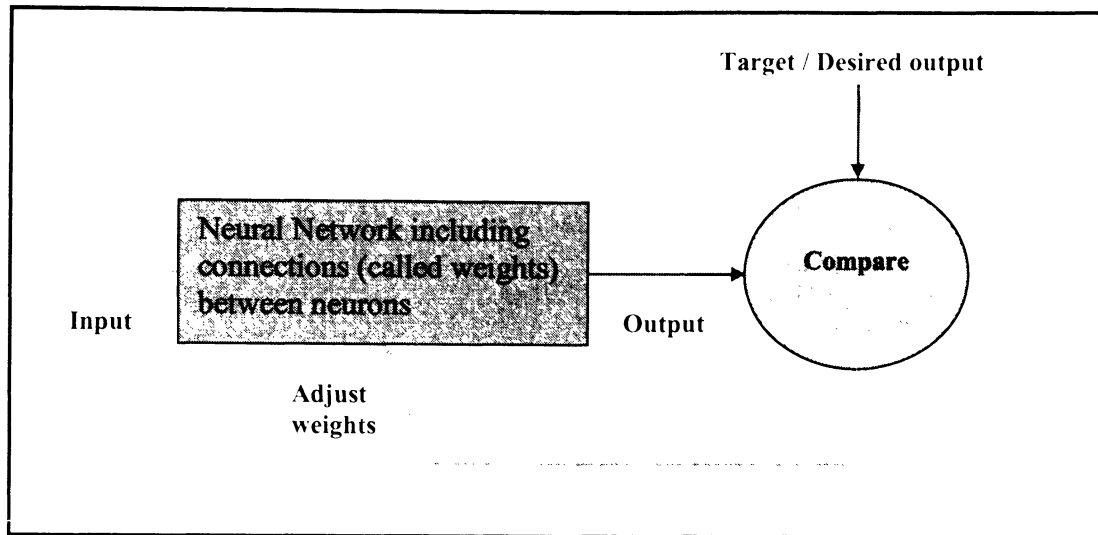


Fig. 2.4: Schematic representation of an ANN training task.

Artificial neural network algorithms became sophisticated enough for general applications in mid-1980s. Today ANNs are being applied to an increasing number of real-world problems of considerable complexity such as high performance aircraft autopilot and flight path simulation in aerospace, speech recognition, credit card activity in banking, sonar and radar image and signal processing in military, nonlinear modeling in electronics, EEG and ECG signal analysis and breast cancer cell analysis in biomedical applications[114-124]. The advantage of ANNs lies in their robustness against distortions in the input data and their capability of learning. They are often good at solving problems that are too complex for conventional technologies. They are commonly applied to biomedical problems that do not have an algorithmic solution or for which an algorithmic solution is too complex to be found. Figure 2.5 depicts a representation of a two-layer NN expressed mathematically by Eq. (2.29).

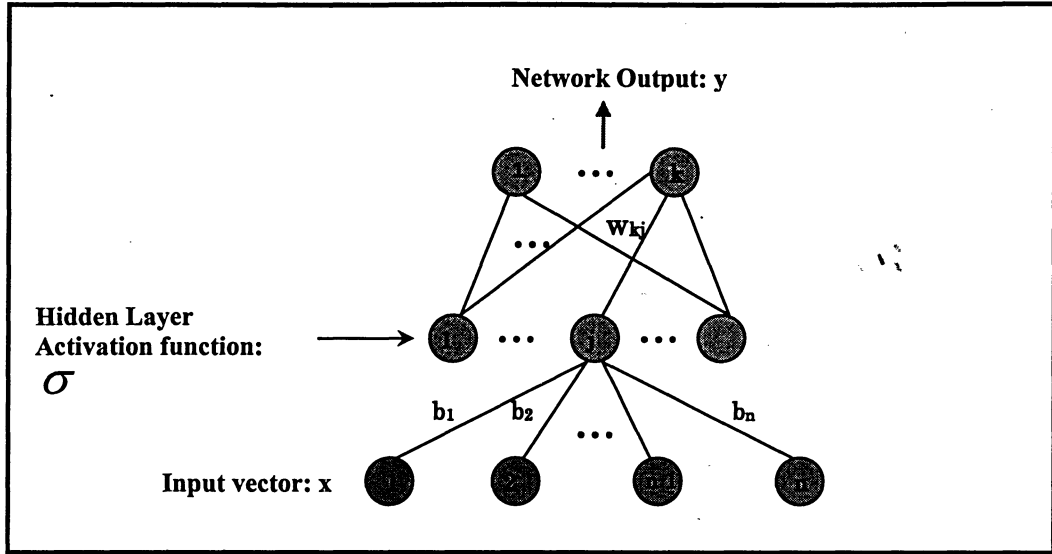


Fig. 2.5: Schematic representation of a two-layer neural network.

$$y_k = \sum_j w_{kj} \sigma(\mathbf{b}_j^T \mathbf{x}) = \mathbf{w}_k \mathbf{h} \quad (2.29)$$

where y_k is the output from unit k , \mathbf{x} is the input vector, \mathbf{b}_j^T denotes the input-to-hidden weight vector to unit j , w_{kj} is the hidden-to-output weight associated with j^{th} input to k and σ represents the activation function. Usually activation functions are invertible and differentiable. In Eq. (2.29), $h_j = \sigma(\mathbf{b}_j^T \mathbf{x})$ is known as the *hidden unit activation*. The linearity or non-linearity of a NN is directly dependant to the type of its activation function, in other words if the activation functions is linear/non-linear the ANN is referred as linear/non-linear NN. As mentioned before, the network is trained so that a particular input leads to a specific target output by adjusting the values of the connections (weights) between units. This tuning which is usually based on comparison between NN output and target (desired) output values continues until the network output matches the target within a specific error criteria defined for the model. More details on choosing this criterion is explained in the next section.

The learning rule which is also referred as *training algorithm* is a procedure for modifying the weights and biases of a NN to obtain the desired output within a preferable

range. These rules are categorized in two general groups: supervised and unsupervised learning algorithms. In *supervised learning*, the learning rule is provided with a set of examples (the *training set*) of proper network behavior and as the inputs are applied to the network, the network outputs are calculated and compared to the target outputs. The learning rule is then used to adjust the weights and biases of the network in order to move the network outputs closer to these targets. The example of such learning is the classification problem [125-129]. In *unsupervised learning*, the weights and biases are modified in response to network inputs only and no target outputs are available. Most of such algorithms perform clustering and categorizing the input patterns into a finite number of classes. Unsupervised learning is widely used in applications such as market segmentation, image classification and vector quantization [130-133]. Due to importance of supervised learning in classification problems involving ANNs, backpropagation which is the most widely used algorithm in learning neural networks.

Back propagation is the basis for training a supervised NN. Static (time independent) back propagation is used to produce an instantaneous mapping of a static input to a static output. The term “backpropagation” is used to imply a backward pass of error to each internal node within the network, which is then used to calculate weight gradients for that node. Learning progresses by alternately propagating forward the activations and propagating backward the instantaneous errors. Such networks are used to solve static classification problems such as optical character recognition (OCR) or medical diagnosis [30, 134-138]. Training continues until the network can approximate a function, associate input vectors with specific output vectors, or classify input vectors in an appropriate way as defined by the user. If properly trained, these backpropagating networks tend to give reasonable answers when presented with inputs that they have never seen (testing data). This generalization property makes it possible to train a network on a representative set of input/target pairs and get good results without training the network on all possible input/output pairs.

Gradient descent algorithm that was first introduced in 1986 (Rumelhart and McClelland) [139] is a standard backpropagation method, in which the network weights are iteratively moved along the conjugate gradient direction. This is the direction in which the

performance or cost function is decreasing most rapidly. The idea was to search the hypothesis space of possible weigh vectors to find the best weights to fit the function that result in minimum error. In order to derive a rule for these “best weights” using gradient descent approach, we should begin by specifying some kind of criteria for network training error. In the following derivations, *squared error* is used as cost (performance) function due to its simplicity:

$$E = \sum_k (y_k - o_k)^2 \quad (2.30)$$

Where y_k and o_k denotes network output and desired output from unit k respectively. In order to find the direction of the steepest descent, the derivatives of error (E) are calculated with respect to each component of weight vector. This vector is called *error gradient* and is defined by:

$$\nabla E(\vec{w}) \equiv \left[\frac{\partial E}{\partial w_0}, \frac{\partial E}{\partial w_1}, \dots, \frac{\partial E}{\partial w_n} \right] \quad (2.31)$$

After error gradient is calculated, the weight vector is updated by this gradient in each step:

$$\begin{aligned} \vec{w} &\leftarrow \vec{w} + \Delta \vec{w} \\ \Delta \vec{w} &= -\delta \nabla E(\vec{w}) \end{aligned} \quad (2.32)$$

Where $\delta > 0$ is a constant called *learning rate* that determines the step size in the gradient descent search. The negative sign in Eq. (2.32) denotes that the desired change in weight vector is in a direction that decreases the error (cost function E). In Appendix C the equations used to estimate the error gradient for a NN with sigmoid activation function are shown.

The weights in are updated in each iteration of backpropagation algorithm using Eq. (2.32) until the training error is decreased to a desired level. As seen, in conjugate gradient algorithms a search is performed along the conjugate directions of the gradient, which produces generally faster convergence than other descent directions (e.g. steepest descent) [123, 140]. Neural networks and their related learning algorithms are very extent topics and in this thesis, only the algorithms that were involved this research (to classify ultrasound backscatter signals from cells) were briefly discussed.

2.2.6. Application of Probabilistic Non-linear Neural Networks to Ultrasound Signal Classification

The probabilistic approach uses a training matrix containing 50 samples of simulated normal and apoptotic backscatter signals. For the experimental ultrasound signals 35 samples of each class are used for training the network. Three distinct NNs were trained for normal-apoptotic, normal-decay and apoptotic-decay classification respectively. Two types of probabilistic networks were designed with sigmoid and tangent hyperbolic activation functions (in both hidden and output layers) and the expected outputs of each network were set to 0.9 and 0.1 indicating the probabilities of normal and apoptotic or decay classes respectively. In total one NN was trained and tested for simulated data and 3 were trained for classification between experimental normal, apoptotic and decay ultrasound backscatter signals. The NNs were trained in 5,000 and 50,000 iterations for simulated and experimental data respectively. NNs are the only iterative classifier in this study and therefore before testing this algorithm, the average training error with different number of iterations was calculated for NNs with both tangent hyperbolic and sigmoid activation functions. This result of averaged training error is depicted in Section 3.2, Fig. 3.8. After training was completed, the networks performance were tested with 50 and 100 testing samples for simulated and experimentally measured ultrasound backscatter signals respectively.

2.3. Signal Segmentation

Non-stationary signals which their statistical characteristics (i.e. mean and variance) and frequency distribution change over time may have highly complex time-frequency features. Analyzing these signals requires developing advanced signal processing techniques such as adaptive modeling, time-frequency distribution or wavelet analysis. After segmentation, analysis of the stationary signals can be done using conventional techniques such as classical Fourier transform methods or AR modeling. Signal segmentation techniques are introduced to divide a nonstationary signal into blocks of segments that are stationary or quasi-stationary (the signal statistics change but they fluctuate around a dynamic

equilibrium). Segmentation algorithms are popular not only for simplifying signal analysis, but to determine different portions of a signal that have different statistical characteristics. This change in statistical properties may be a representation of a change in the information carried by the signal. Most biomedical signals are nonstationary [141] and this emphasizes the importance of signal segmentation techniques in studying biomedical signals. Examples of nonstationary biomedical signals are vibromyographic (muscle sounds), vibroarthrographic (VAG), vibrocardiographic (heart sounds) and ECG signals. Ultrasound backscatter signals from tissues and tumors are also likely to be non-stationary as they contain the backscatter signals from different cells, and regions which respond differently and are expected to have different statistical properties. Therefore it is important to develop signal segmentation methods in order to analyze and classify biomedical signals.

There are two main approaches in segmenting a non-stationary signal: fixed segmentation and adaptive segmentation. The *fixed segmentation* approach uses a constant length window to study the statistical changes in neighboring windows and determines the probable points of non-stationarity accordingly. In this method the choice of window length is very important as the signal portion must be short enough to be considered almost stationary, while long enough so that the desired signal modeling would converge and model parameters would be estimated. In *adaptive segmentation*, the segment length changes dynamically according to the statistical changes in the signal that enables tracking changes in signal. There are several adaptive signal segmentation methods proposed in literature such as Autocorrelation [44, 142], Spectral error measure [143], RLS-based adaptive segmentation [144, 145] and the Recursive least squares lattice based segmentation (RLSL) methods [27]. In this research we use the recursive least squares lattice filter for segmenting ultrasound backscatter signals due to its modularity, fast convergence and efficiency.

2.3.1. Recursive Least Squares Lattice (RLSL) Filter

The RLSL filter consists of two filters in each stage: a forward and backward prediction filter with different values for their reflection coefficients ($K_{f,m}(n)$ and $K_{b,m}(n)$). This structure uses a time recursion for filter forward and backward prediction errors in terms

lattice structure maintains the order update at the same time [37, 146, 147]. Therefore the filter is guaranteed for fast update as new data becomes available (Fig. 2.6).

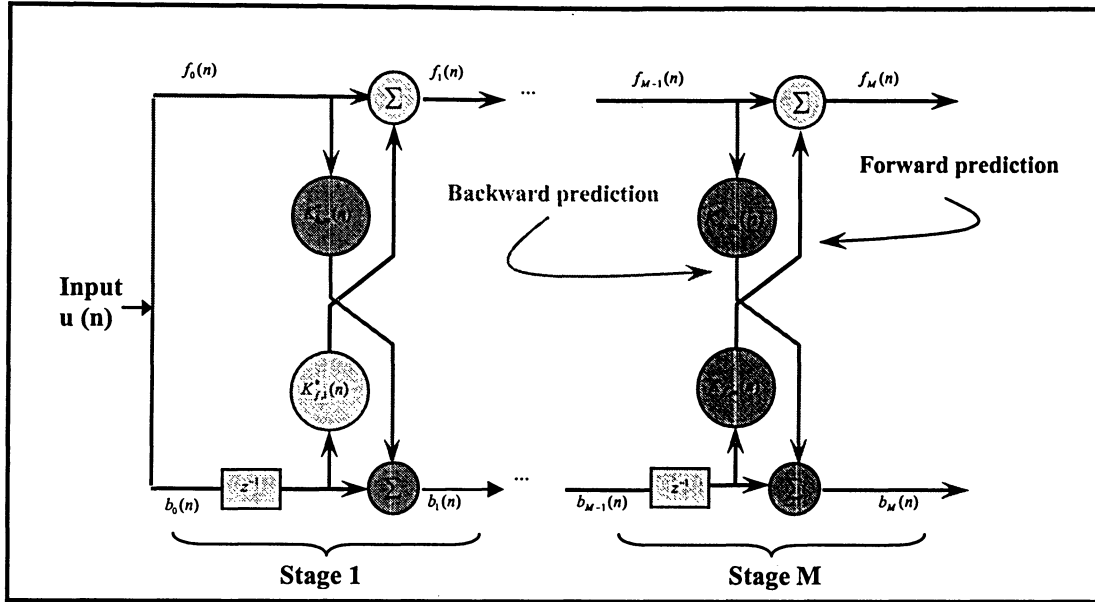


Fig. 2.6: M-order lattice predictor including of forward and backward prediction filters. This filter updates both forward and backward prediction errors $f_m(n)$ and $b_m(n)$ in each order update. The forward and backward reflection coefficients: $K_{f,m}(n)$ and $K_{b,m}(n)$ are updated with time which ensures filter adaptation.

The RLSL filter includes several properties not found in finite impulse response (FIR) filter structure such as modularity and good numerical round-off which are achieved at the expense of increased computational complexity for a given order of filter. As mentioned earlier, the main idea of using this structure is to get to a fast convergence by using both forward and backward filters. Another advantage of using lattice based adaptive algorithms is orthogonalization property of each lattice stage which leads to modularity (the backward prediction errors are determined as the result of a Gram-Schmidt orthogonalization therefore they are uncorrelated at the same time instant resulting in an *stage modularity*), good numerical conditions and ease of testing [27]. In each stage of the RLSL algorithm 8 parameters are updated recursively. The order-update recursion equations of these parameters are:

$$f_m(n) = f_{m-1}(n) + \Gamma_{b,m}(n)b_{m-1}(n-1), \quad (2.33)$$

$$b_m(n) = b_{m-1}(n-1) + \Gamma_{f,m}(n)f_{m-1}(n), \quad (2.34)$$

In Eq. (2.33) and (2.34), $f_m(n)$ is the forward prediction error in order m and time n , $b_m(n)$ is the backward prediction error; and $\Gamma_{b,m}(n)$ and $\Gamma_{f,m}(n)$ denote the forward and backward reflection coefficients respectively which are defined by:

$$\Gamma_{f,m}(n) = -\frac{\Delta_{m-1}(n)}{F_{m-1}(n)} \quad (2.35)$$

$$\Gamma_{b,m}(n) = -\frac{\Delta_{m-1}(n)}{B_{m-1}(n-1)} \quad (2.36)$$

Where $F_m(n)$ and $B_m(n)$ are forward and backward prediction error powers that are updated by Eq. (2.37) and (2.38):

$$F_m(n) = F_{m-1}(n) + \Gamma_{b,m}(n)\Delta_{m-1}(n), \quad (2.37)$$

$$B_m(n) = B_{m-1}(n-1) + \Gamma_{f,m}(n)\Delta_{m-1}(n), \quad (2.38)$$

Here $\Delta_{m-1}(n)$ is the cross correlation between the delayed backward and forward prediction errors described by:

$$\Delta_{m-1}(n) = \lambda \cdot \Delta_{m-1}(n-1) + \frac{b_{m-1}(n-1)f_{m-1}(n)}{\gamma_{m-1}(n-1)} \quad (2.39)$$

The parameter λ ($0 \leq \lambda \leq 1$) in Eq. (2.39) is called *forgetting factor*. A closer value of forgetting factor to one gives a more weight to the recent cross correlation factor. The most important result of this structure which makes it suitable for signal segmentation is the existence of parameter $\gamma_m(n)$, called *convergence factor*. The convergence factor which is also known as “*gamma factor*” provides the connecting link between different sets of a priori and posteriori estimation errors in this algorithm and is defined by the following equation:

sets of a priori and posteriori estimation errors in this algorithm and is defined by the following equation:

$$\gamma_m(n) = \gamma_{m-1}(n) - \frac{b_{m-1}^2(n)}{B_{m-1}(n)} \quad (2.40)$$

This factor may be interpreted as an approximation to a statistical likelihood variable and reflects the statistical changes that occur in the signal. As long as the signal is stationary (the input data belong to the same distribution) convergence (gamma) factor remains in the same range, but when the signal statistics (such as mean or variance) changes, abrupt changes will occur in this parameter that results in a drop below a certain threshold in the convergence factor plot. For example as long as the data belongs to a Gaussian distribution the convergence factor threshold is 1. The signal non-stationarity segments can be detected by tracking the value of convergence factor (Eq. (2.40)) with respect to a threshold and set the segment boundaries accordingly. In the following section the feasibility of RLSL segmentation method is shown by applying to ultrasound backscatter signals.

2.3.2. Application of RLSL algorithm on Ultrasound Backscatter Signals

In this work, we are interested in distinguishing between 3 groups of cells: normal, apoptotic and cells that are left to decay and therefore are dying via necrosis (Section 2.5.1.c). As mentioned in Section 2.2.1 the classification methods that are used in this research are based on ultrasound backscatter signal *model parameters*. This means that instead of using the whole signal as the input, only a selected number of coefficients are used for training and testing ML classifiers (in this research 15 coefficients are used). This is the first step in analyzing the signal by breaking it into stationary segments and using conventional spectral analysis individually on these segments. Sections 3.1 and 4.1 will discuss the details of the procedure for simulated and experimental ultrasound signals. Ultrasound signal segmentation would also be very important in clinical situations where the backscatter signal is reflected from different layers of skin and tissues and the

cancerous tumor cells may be in an unknown location from the surface of the transducer. Therefore differentiating these various layers is essential for analysis of an ultrasound signal and can be used as a method to monitor tumor progression or shrinkage during the therapy. Currently the ultrasound transducer for our experiments has a *5mm-1cm* penetration that limits its application in clinical patients but there are other investigations underway to increase this penetration depth. In order to test the feasibility of the RLSL lattice based segmentation method, an experiment was set which included a three-layer, normal-apoptotic-normal cell pellet with the apoptotic layer located between two layers of normal cells as depicted in Fig. 4.9.a. The ultrasound backscatter signal from this pellet is shown in Fig.4.8.b where the apoptotic layer is located between samples 800 to 1500 (Section 4.3).

2.4. Simulations of Ultrasound Backscatter Signals

Hunt et al. [10] showed that normal and apoptotic ultrasound backscatter signals can be modeled as resulting from regularly and randomly spaced scatterers. If scatterers are modeled as being regularly spaced, there is a large reduction in backscatter signals when compared to scatters that are randomly distributed. Predictions of this model have been supported by recent ultrasound studies on cells undergoing apoptosis, in which considerable changes in backscatter signals and frequency spectra have been observed [10, 148]. Other groups have explored the statistical properties produced from scatterers, although the exact relation between these statistics and the physical changes that occur in the cell during cell death is yet to be determined. The method which was developed by Parmar et al. [36, 148], is used as a model to generate simulated ultrasound backscatter signals for analyzing and testing the classification and segmentation techniques (Section 2.2.1 and 2.3). The computer simulations investigated the effects of spacing variations between scatterers on the ultrasound backscatter signals. The computer simulation model carried out was based on a one-dimensional discrete scattering model. Spacing between each scatterer was a random variable based on a Gaussian distribution with mean Δx and standard deviation of " σ ". The standard deviation of scatterer location determines the randomness of the scatterer spacing. For example, a variance of $\sigma = 1\%$ represented

regularly spaced scatters, while a variance of $\sigma = 50\%$ denotes a great degree of randomness in spacing. The transmitted pulse was modeled as,

$$p(t) = -t \exp(-4\beta^2 t^2) \sin(2\pi f_o t) \quad (2.41)$$

Where f_o is the center frequency and β is the bandwidth. The received backscatter signal, $s(t)$ is a superposition of echoes from all of the scatterers [149]:

$$s(t) = \sum_{k=1}^N a_k p\left(t - \frac{2x_k}{c}\right) \quad (2.42)$$

where a_k is the backscatter coefficient of the k -th scatterer (indicates how strong each scatterer scatters the sound), x_k is the location of the k -th scatterer from the origin, and c is the velocity of sound (1500m/s). Two models were used on the strength of each scatter. One assumed that the coefficient is 1 (means that all scatterers have equal strength), and the second assumes that the distribution of the scatterer strength fit to a Gamma distribution. In each simulation, a center frequency of 40 MHz with bandwidth 17 MHz was used. The simulation model assumed each cell to consist of two scatterers as the sound is scattered from two interfaces. The two reflections are 180° out of phase.

In Chapter 3 it is shown that the simulated ultrasound signals are almost stationary (the same result is evaluated for experimental ultrasound backscatter signal) as both the modeled and original signals have a very similar spectrum spread. This accuracy can not be achieved unless the ultrasound backscatter signals from cell pellets (normal, apoptotic and decay) fulfill the stationarity property of the AR modeling. Therefore we applied segmentation on a simulated non-stationary signal. This signal which is depicted in Fig. 3.9 in Section 3.3 is a three layer normal-apoptotic-normal simulated ultrasound backscatter signal (Section 2.5.2). The assumptions for generating this simulated multi-layered ultrasound signal are the focal transducer frequency of 40 MHz, high sampling frequency of 500 MHz; and backscatter variance spacing of 0.001 mm for the two normal and 0.9 mm for the single apoptotic layers. The simulation method also considers the ultrasound backscatter signal attenuation which is a characteristic of experimental ultrasound backscatter signals. The attenuation is assumed to be exponential. These values were chosen to match the frequencies of the experimental backscatter signals.

2.5. Experimental Methods

The ultrasound backscatter radiofrequency (RF) data are collected from acute myeloid leukemia (AML) cells that were grown in suspension. Cells were prepared using a cell culture system. For each ultrasound experiment, approximately 1×10^9 human acute myeloid leukemia cells (AML-5) were grown at 37°C in α -minimal-media from frozen stock samples using 200 mL of media. Cell culture growth was initiated using frozen stock cells. This cell line has a well characterized apoptotic pathway, and could produce sample cells for our experiments [150]. These cells are referred as *Normal* cells for the rest of this study. Cell pellets were made by washing cells in phosphate-buffered saline (PBS) and the preparations were subsequently pelleted in flat bottom cryotubes on a desktop swinging bucket centrifuge. Details on the biological procedure can be found elsewhere (Czarnota et al. [17]). Normal cells are then treated with chemotherapeutic agent *cis-platinum* (cisplatin) at 10 μ g/mL which induces apoptosis. Cisplatin is a DNA intercalator that causes a p53-dependent apoptosis in this cell line. Cells were treated with the drug for 0, 3, 12, 24 and 48 hours. Pellets were then immersed in PBS, which acted as the coupling medium for the ultrasound imaging and experiments were performed within a maximum of 1 hour after centrifugation. The imaging was performed using a 20MHz f2.35 or 40 MHz f2 transducer attached to a commercial ultrasound imaging unit (Visual Sonics¹³). The transducer trigger pulse (Gaussian pulse) initiated data acquisition and RF data were collected at the location of interest. Compressed RF data were transferred to a workstation for processing. Matlab version 6.5 (www.mathworks.com) was used to perform all signal analysis. The third type of cells, *decay* cells, are prepared by keeping normal AML cells in room temperature (37°C) without changing the media for 9, 12 and 24 hours. Therefore the cells in the pellet are deprived of oxygen and nutrients. The purpose of experiment on this cell type is to analyze the difference between cell death caused by apoptosis and decay (Section 4.3).

In this study we analyze the data from three groups of cells: normal, and two types of cell death: apoptosis and decaying cells (necrosis). Based on this method three sample sets were generated for each class of normal, apoptotic and decay cells. These sets

¹³ www.visualsonics.com

and their combination are used as training and testing data for ML classifiers (see Sections 2.1.4 and 2.2.6). The names and description of these data sets are presented in the following section.

2.5.1. Homogenous Cell Pellets

a) Normal Biological Samples

These three groups of normal samples are used for training and testing of normal-apoptotic and normal-decay classifiers. The results of these analyses are shown in Chapter 4, Table 4.1, and Figures 4.7 and 4.8.

a.1) Normal-ZeroHour-Pellet #1 (N-0h-P1): This is a set of 45 ultrasound backscatter signals from a pellet of normal AML cells that has been cultured and kept in normal conditions (37°C and in α -minimal-media that has been properly changed every day). This pellet is then treated with cisplatin drug which induces apoptosis and is used in groups 2 and 4-5, that are explained later in of this section. The term “ZeroHour” phrase emphasizes that this data set is not exposed to cisplatin (refers to zero hour before exposure to the drug).

a.2) Normal-ThreeHour-Pellet #1 (N-3h-P1): The 80 signals in this group are chosen from the same pellet mentioned above (pellet 1), but after it has been treated with apoptosis induced drug, cisplatin, for 3 hours. In Section 4.3 we will investigate the effect of cisplatin on these cells in this time period. It is suspected that only a small percentage of these cells become apoptotic during 3 hours exposure to cisplatin.

a.3) Normal-ZeroHour-Pellet # 2 (N-0h-P2): These 80 samples are the last set of normal testing data, which are chosen from a different cell pellet(pellet 2) that is also cultured and kept under the same normal living conditions as N-0h-P1. Similar to pellet 1 this pellet is later treated by cisplatin to analyze apoptosis in AML cells and is used in apoptotic samples group of biological samples. The purpose of choosing the second cell pellet is to verify whether the classifier results are reliable with a different

cell pellet can be extended. Later in Section 4.3 we will train the classifiers with one pellet and test it with the second one to understand if the classifier is biased towards the training set.

b) Apoptotic Biological Samples

The samples that are explained in below are used as the basis for training and testing ML classifiers that are explained in Section 2.2. The details of classification results are shown in Table 4.1, and Figures 4.7 and 4.9 in Section 4.3.

b.1) Apoptotic-48Hour-Pellet #1 (A-48h-P1): This set has 45 ultrasound backscatter samples which are selected from pellet 1 (N-0h-P1) and 48 hours after it has been treated with cisplatin. It is expected that most of the cells at this stage have become apoptotic.

b.2) Apoptotic-24Hour-Pellet #1 (A-24h-P1): These 80 ultrasound backscatter signals are prepared in the same way that previous samples (A-48h-P1) was generated, but exposed for 24 hours instead of 48 to cisplatin. These data are analyzed to give an insight into the level of apoptosis in AML cells with respect to their exposure time to cisplatin.

b.3) Apoptotic-48Hour-Pellet #2 (A-48h-P2): This set of 80 testing signals are from cell pellet 2, which was mentioned in “N-0h-P2”, after it was treated for 48 hours with cisplatin.

c) Decay Biological Samples

Necrosis which is the pathological death of a cell or group of cells happen when the ‘default’ mode of death, apoptosis cannot be initiated due to factors such as high level injury and lack of the available energy to initiate the process. Usually in such cases the tissue in an organ dies due to the lack of blood supply. To prepare such experimental samples, the cells were centrifuged and left to decay on a table. Similar to the previous

groups the three decay group sample sets (described as below) are used in training and testing the ML classifiers to test their ability in distinguishing ultrasound backscatter signals from normal-decay and apoptotic-decay groups.

c.1) Decay-12Hour-Pellet #3 (D-12h-P3): The 36 ultrasound backscatter signal samples of this group are from a third cell pellet (pellet 3) which is kept in pellet form in room temperature for 12 hours. It is expected that the majority of these cells would have gone under necrosis cell death by this time.

c.2) Decay-9Hour-Pellet #3 (D-9h-P3): Very similar to our approach in normal and apoptotic groups to choose two sample sets from the same pellet and different in their exposure time to cisplatin, the second 72 sample set of decay ultrasound backscatter signals are selected 9 hours after cell pellet 3 was left in decay conditions to monitor the resemblance of cell decay in these different time slots.

c.3) Decay-12Hour-Pellet#4(D-12h-P4): In order to study the level of classifier independency of their training data, the third 72 decay sample set is obtained from a different cell pellet (pellet 4) after it has been left 12 hours in decay process with the same condition of pellet 3 (D-12h-P3).

2.5.2. Non-homogenous Cell Pellets

To evaluate the RLSL adaptive signal segmentation technique (Section 2.3.1), we prepared a 3-layer normal, apoptotic normal cell pellet. Each homogenous normal and apoptotic layer is prepared and centrifuged separately using the procedures explained in Section 2.5.2. The three resulting layers (with known thickness) were placed into one pellet and centrifuged again to obtain the desired sample for segmentation technique. Figure 2.7 depicts the image of this three-layer pellet.

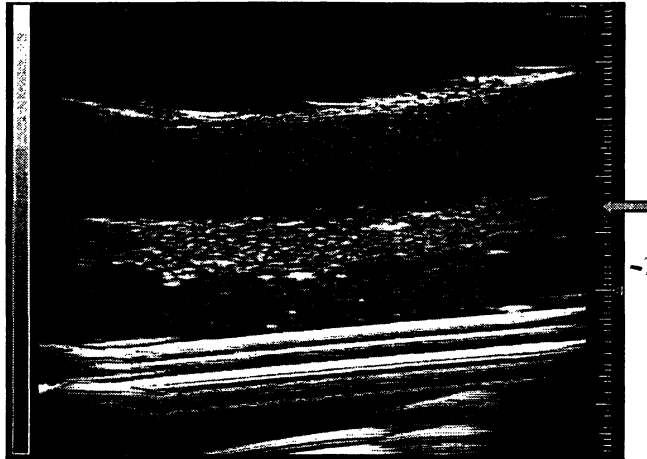


Fig. 2.7: A three-layer cell pellet. The apoptotic layer (pink arrow) is located between two normal layers.

As shown in figure 2.7 the apoptotic layer, which is brighter, is located between two layers of normal cells (which are darker). The other two bright parts of this figure are the surface and the bottom of the pellet which reflect more ultrasound backscatter signals. These data are excluded from our analysis to avoid any confusion. The ultrasound backscatter signal segmentation and classification results for this sample are shown in Section 4.4.

Chapter 3

Classification & Segmentation of Simulated Ultrasound Signals

To model, classify and segment ultrasound backscatter signals, simulated ultrasound signals were generated using the algorithm developed by Parmar et al. (Section 2.4). Although the origin of ultrasound backscatter signals is not well understood, this method has shown to be a reasonable estimate of the biomedical ultrasound backscatter signals. As explained in Chapter 2, this model convolves the incident pressure wave with scatterers that are spaced with various degrees of randomization. Figures 3.a and 3.b show resulted simulated ultrasound signals generated with a centre frequency of 40 MHz using scatterers spaced on average $10\text{ }\mu\text{m}$ apart with variances of $1\text{ }\mu\text{m}$ (for the normal) and $5\text{ }\mu\text{m}$ (for the apoptotic). It has also seen experimentally that the backscatter from the apoptotic cells is greater than that from the normal cells.

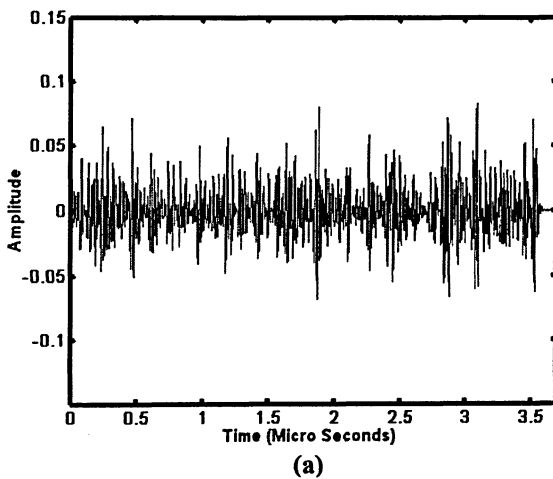
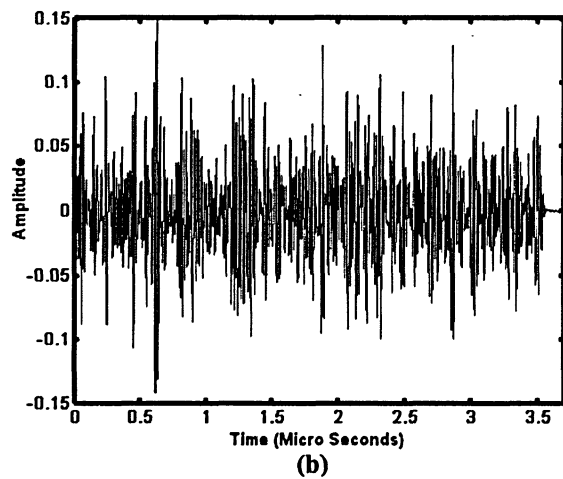


Fig. 3.1.a) A simulated normal ultrasound backscatter signal generated with scatterers that are regularly spaced $10\text{ }\mu\text{m}$ apart with a variance of $\sigma=1\%$.



b) A simulated apoptotic ultrasound backscatter signal generated with scatterers that are spaced $10\text{ }\mu\text{m}$ apart with a variance of $\sigma=50\%$.

3.1. Autoregressive Modeling of Simulated Ultrasound Backscatter Signals

The AR modeling of the ultrasound backscatter signal was applied using the Burg-lattice, Yule-walker and Covariance methods (Section 2.1.2). To compare the AR modeling techniques, a simulated normal and apoptotic signal is modeled and the PSD of the each model signal over a specific frequency band (0 to 100 MHz) is used to show the similarities between the average signal powers. Figures 3.2 and 3.3 illustrate the results from each modeling technique in PSD estimation of simulated ultrasound backscatter signals.

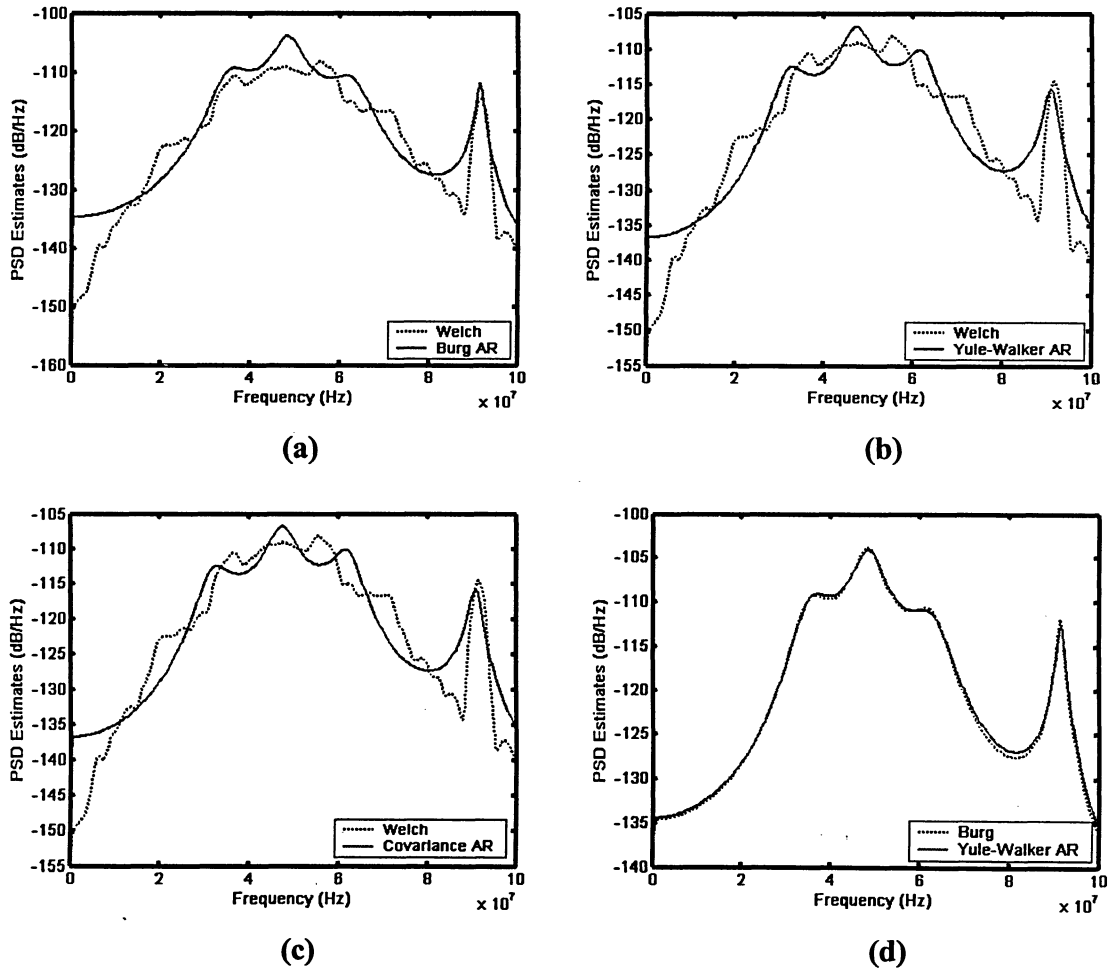


Fig. 3.2: Comparison of the average PSD of 20 simulated normal ultrasound backscatter signals calculated by: a) Burg, b) Yule-Walker; and c) Covariance techniques. The basis of this comparison is the signal PSD calculated by the Welch spectrum estimation method. d) Comparing the results of Burg and Yule-Walker PSD estimation techniques.

In these figures the Welch method, a non-parametric algorithm, is used to directly estimate the signal's PSD. This is an improved version of the periodogram and is used as the basis of our comparison. In Section 2.1.4 the accuracy of the AR modeling technique is verified by comparing the PSD of the original and modeled signals.

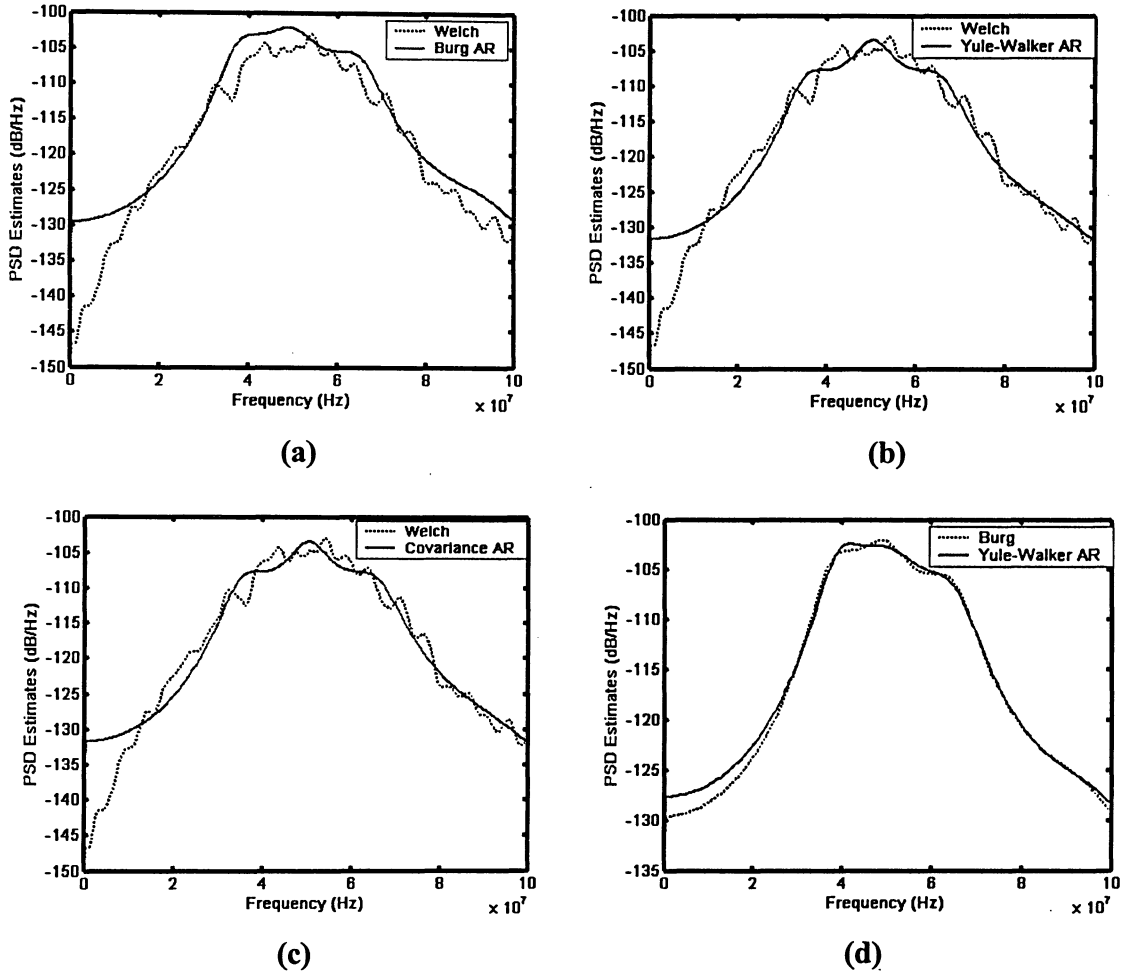


Fig. 3.3: Comparison of the average PSD of 20 simulated apoptotic ultrasound backscatter signals calculated by: a) Burg, b) Yule-Walker; and c) Covariance techniques. The basis of this comparison is the signal PSD calculated by the Welch spectrum estimation method. d) Comparing the results of Burg and Yule-Walker PSD estimation techniques.

These figures demonstrate that there is not much difference in PSD calculation of AR-modeled ultrasound backscatter signals depending on the method used. Therefore in this study the Burg-lattice method is used to find the AR-parameters due to its advantageous properties that (described in Table 2.1 on page 14).

3.1.1. Choosing the Proper Model Order

The next stage in modeling a signal after choosing an AR modeling technique is setting the model parameters. As explained in Section 2.1.3 there are two criteria for choosing the AR modeling order: the minimum (lower bound) and maximum (upper bound) permissible orders. These order boundaries must be calculated in a way that maximizes the similarity of modeled and original signals. The purpose of choosing the maximum permissible order as the upper AR order boundary is to prevent “over-fitting” in our modeling which is the result of choosing a very high model order such that the model will also “capture” the signal noise. This situation can cause problems when the AR coefficients are passed as the input for training and testing the machine learning classifiers in Section 3.2. The minimum and maximum AR modeling orders are calculated for simulated ultrasound backscatter signals and the results are shown in the following sections.

3.1.2. Minimum Model Order

Using the error criteria (Section 2.1.3), the proper range for modeling order was determined using two indicators: 1) the ensemble reconstruction error and 2) the noise (error) variance. Although examination of any one of these indicators is sufficient to reach a conclusion about modeling order, both results presented the consistency of each approach in deciding the best model order.

a) Ensemble Reconstruction Error

The errors associated with the PSD content of the reconstructed and original ultrasound signals were calculated using the Euclidean distance¹⁴ of the PSDs of the two signals (Section. 2.1.3.1). The PSD was calculated using Welch method with a window size of 1024 points and sampling frequency of 500 MHz on simulated ultrasound backscatter,

¹⁴ The Euclidean distance between points x and y in \mathbb{R}^n space is $d = |x - y| = \sqrt{\sum_{i=1}^n |x_i - y_i|^2}$.

resulting in a signal with the length of 1.5 mm. The error (difference between the PSD of original and AR-modeled signal) was averaged over 20 apoptotic and 20 normal simulated ultrasound signals (Fig.3.4).

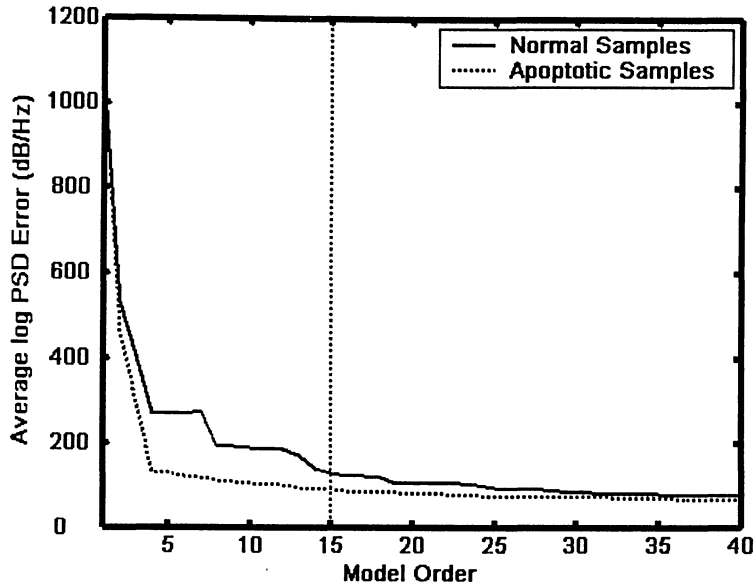


Fig. 3.4: Average Ensemble Error between the log-PSDs of AR estimated and original ultrasound signals for 20 samples of normal and 20 apoptotic signals.

Figure 3.4 shows that as predicted by theory, the similarity between the power spectral densities, which is used for signal characterization in the frequency domain, increases with modeling order. This is caused by the fact that a higher-order model uses more samples to represent the original signal and hence a better signal approximation with less modeling error is achieved. It can also be seen from Fig. 3.4 that the rate of error change does not decrease significantly after order 15 for both normal and apoptotic samples (the error drops with the rate of 88% and 91% for normal and apoptotic samples respectively).

b) Noise (error) Variance

Figure 3.5 depicts the variance of the noise estimated in the AR modeling of ultrasound backscatter signals with different model orders. It is evident from this figure that despite the abrupt drop of the noise variance between model orders 1 to 5 of approximately 4 folds, there is not a significant change in this error after order 5. But as seen in Fig. 3.4, choosing order 5 gives a relatively higher ensemble reconstruction error as compared to

order 15. Therefore we choose order 15 that satisfies the criteria for minimum error in both tests.

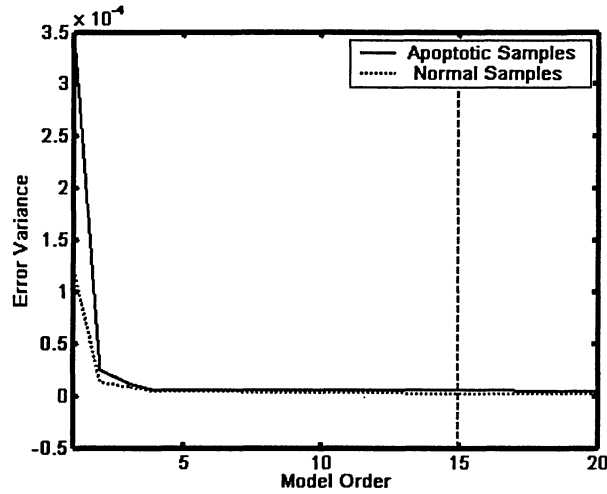


Fig. 3.5: Average variance of the AR model noise (error) for 20 normal and 20 apoptotic simulated ultrasound backscatter signals.

Figure 3.5 also confirms that a modeling order of 15 can be used as a good estimate for the AR modeling of ultrasound backscatter signals.

3.1.3. Maximum Model Order

Figure 3.6.a depicts average ensemble error calculated for 20 samples of normal and apoptotic simulated signals with SNR of 0, 3, 6 and 7 dB according to the method explained in Section 2.1.3. The effect of noise in the AR modeling error of simulated ultrasound signals is evident in this figure: a higher noise level (lower SNR) results in a higher ensemble error. Although the SNR level does not significantly affect the decrease rate of this error with modeling order (this rate is 22%, 22%, 25% and 26% for SNRs of 0, 3, 6 and 7 respectively) and therefore we see similar pattern in these errors.

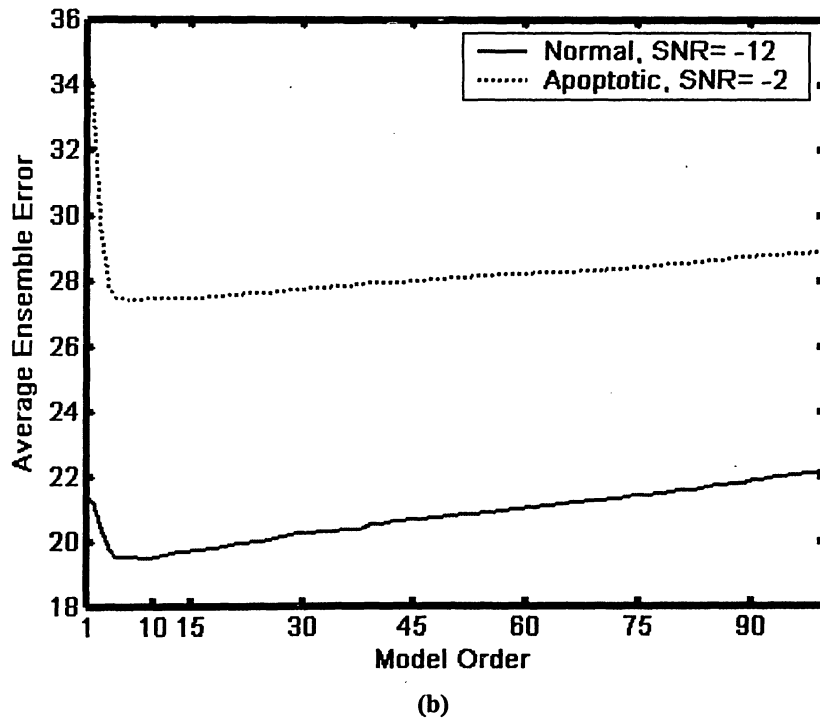
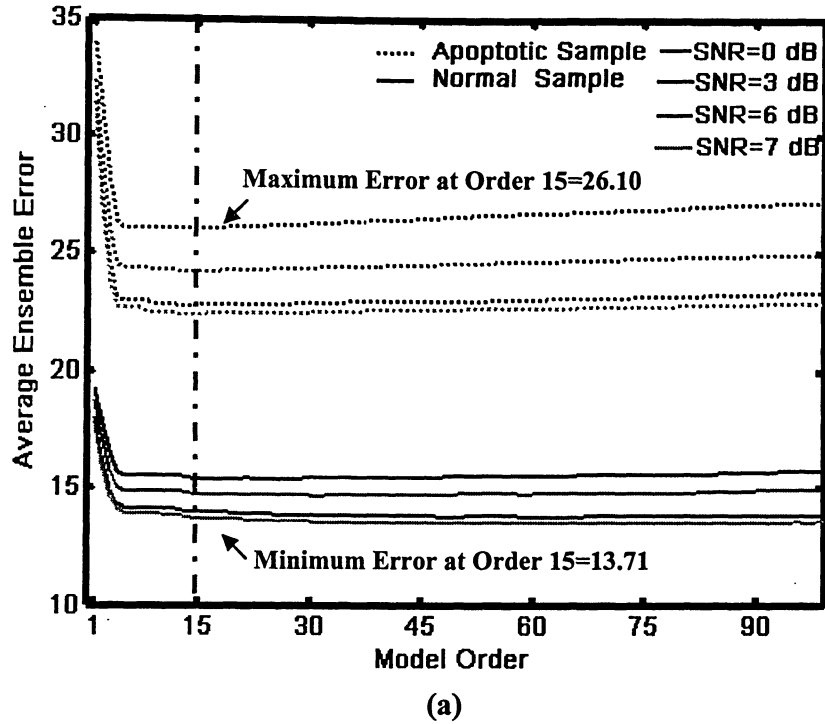


Fig.3.6: a) Average Ensemble Error for 20 normal and 20 apoptotic simulated samples with additive noise and different SNRs (SNR= 0, 3, 6 and 7 dB). Dashed lines represent Apoptotic samples.
b) Representation of over-fitting (visually detectable) for normal (SNR=-12 dB) and apoptotic samples (SNR=-2 dB).

In Fig 3.6.b two curves are shown where over-fitting can be visually detected for simulated normal and apoptotic ultrasound backscatter signals. It can be seen that over-fitting of normal and apoptotic signals starts with different model orders but the error curves increase with quite similar slopes (12% for apoptotic and 5% for normal samples). High noise variance (low SNR) will result in over-fitting even with relatively small AR modeling orders. This simulation denotes that the ensemble error does not change significantly after order 15 that is similar to the result of this analysis with “noiseless” ultrasound backscatter signals (Fig. 3.4 and 3.5).

From Figures 3.6.a-b, it can be concluded that the noise has a negligible effect and does not cause over-fitting in the AR modeling with order of 15. To verify this result, an experimental apoptotic ultrasound backscatter signal (40 MHz center frequency) was modeled with AR process of order 15, and was reconstructed using the calculated AR coefficients, the auto-correlation of the model error¹⁵ (noise) was estimated and shown in Fig. 3.7 (see Sections 2.1.3 and 2.1.4).

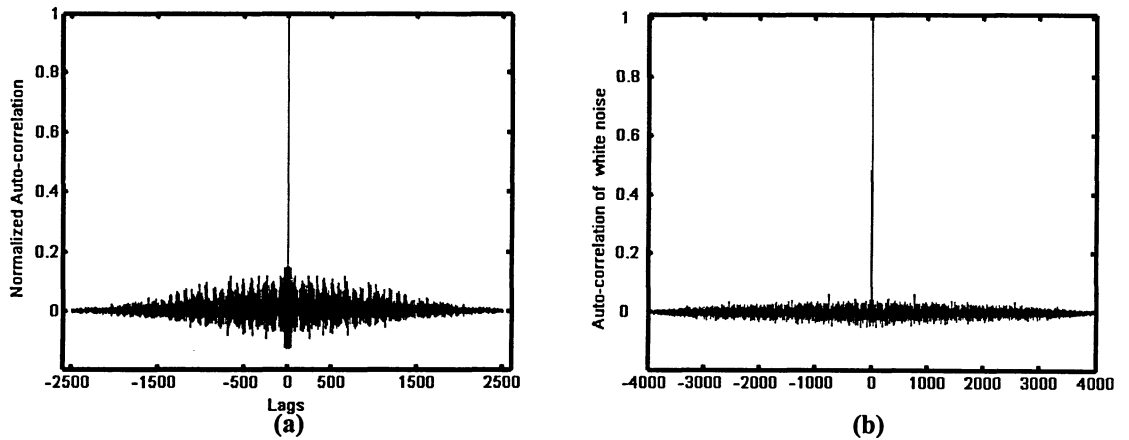


Fig. 3.7: a) Autocorrelation of the estimated model error (noise) compared to the b) autocorrelation of white noise.

The error auto-correlation in Figures 3.7.a and 3.7.b indicates the estimated error is approximately white noise. This satisfies the condition for modeling order 15 (this confirms the consistency of the results of Sections 3.1.2 and 3.1.3) for AR modeling of simulated ultrasound backscatter signals.

¹⁵ This error was assumed to be the absolute difference between original and reconstructed signals.

Most often in practice a perfect model is not necessary for performing classification of biomedical signals; a less ideal model can give almost the same classification accuracy with less computational complexity. But exploring the details of an ideal model gives an insight into the nature and statistics of ultrasound backscatter signals that opens the possibility of future developments in their clinical application.

3.2. Classification of Simulated Ultrasound Backscatter Signals

The classification algorithms are trained and tested with AR coefficients of ultrasound backscatter signals that contain the desired information about their spectral density. Therefore these classifiers will distinguish each backscatter signal based on its PSD similarity to normal or apoptotic group signals. To determine the accuracy of the classification algorithms, machine learning classifiers were trained and tested (Section 2.2) using simulated ultrasound backscatter signals (Section 2.4). For the only iterative classification method (NNs) the proper training number of iterations was estimated by evaluating the training error after each step and choosing a step in the range of the minimum error. This result is shown in Fig. 3.8 for both tangent hyperbolic and sigmoid probabilistic networks.

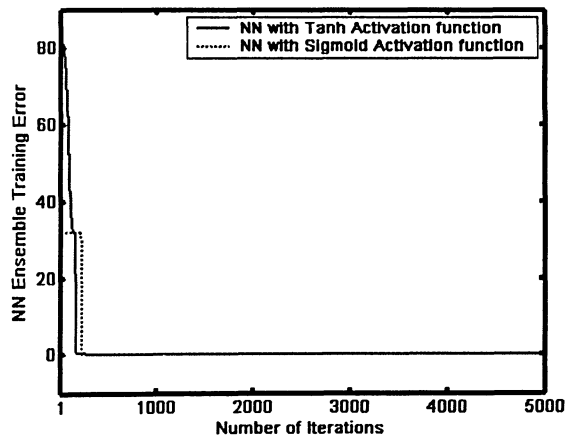


Fig. 3.8: Neural network training error with Tanh and Sigmoid activation functions.

Figure 3.8 reveals that the nature of the simulated signals is simple to classify and therefore the training, which is done in relatively small number of iterations, has resulted in almost 0% error. When the training is completed, the network weights are fixed and its

performance is tested for new unseen simulated ultrasound signals (regressed to 15 coefficients) by finding the distance between the network output to this testing data and the default normal and apoptotic class outputs (0.9 and 0.1) and assigning the testing data to the minimum distant class.

The results of classifying simulated ultrasound backscatter signals using the classifiers of Section (2.2) are shown in Table 3.1. These classifiers were trained with 100 samples of normal and 100 samples of apoptotic AR-parameters and tested with 50 samples (again in the form of a vector of 15 AR coefficients) of each group respectively. Two non-linear NNs were implemented with *sigmoid* and *tangent-hyperbolic* (Tanh) activation functions. Both networks were trained in 5,000 iterations using backpropagation algorithm. As seen in Table 3.1 all of the linear and non-linear classifiers have detected normal and apoptotic samples with 100% accuracy. This high accuracy is most likely due to the method used to simulate ultrasound backscatter signals where the normal and apoptotic backscatter signals differ only in their scatterers spacing (normal backscatter has regular and apoptotic backscatter has irregular spacing). So they may be easily classified even in a simple linear space.

TABLE 3.1
Classification Algorithms Accuracy for Simulated Ultrasound Signals

Algorithm	Normal Accuracy (%)	Apoptotic Accuracy (%)
Fisher's Linear Discriminant	100	100
Gaussian Classifier	100	100
Naïve Bayes Classifier	100	100
NN with Sigmoid activation function	100	100
NN with Tanh activation function	100	100

3.3. Segmentation of Ultrasound Backscatter Signals

The simulated 3-layer signal was generated according to the method and specifications described in Section 2.5.2 to test the segmentation algorithms. A representative three layer normal-apoptotic-normal simulated signal is illustrated in Fig. 3.9 with the known layer boundaries shown by dashed lines.

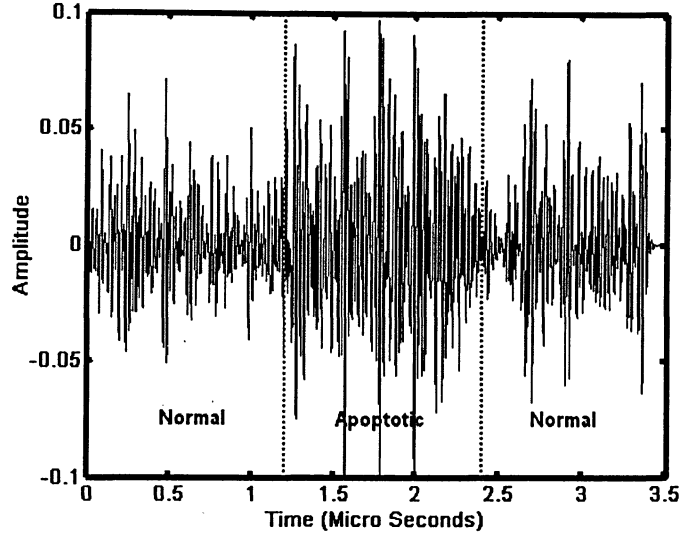
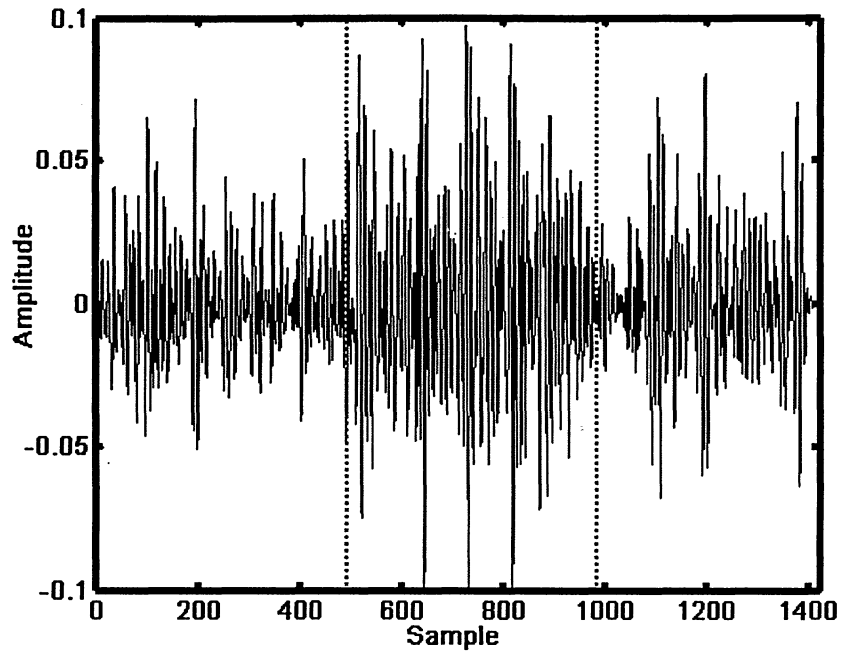


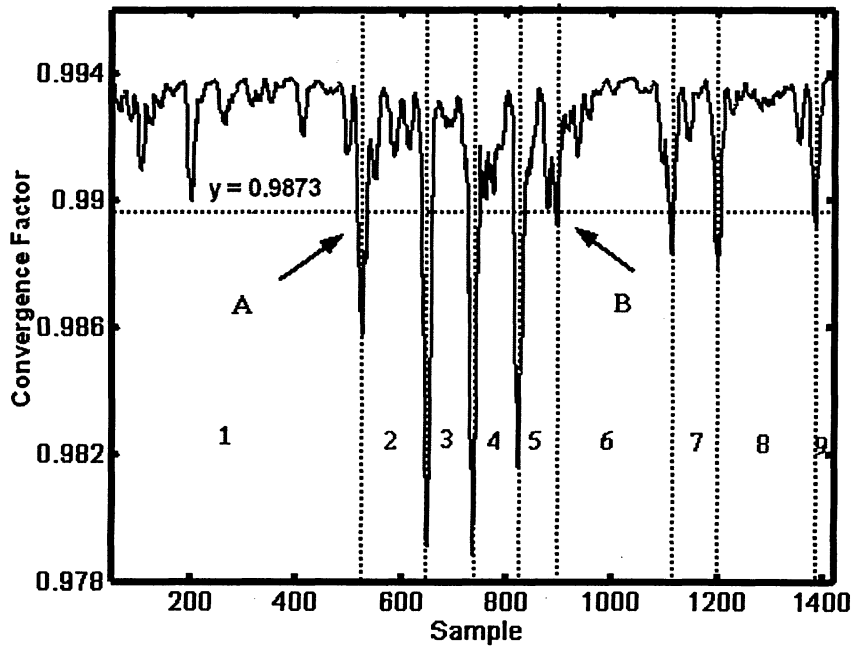
Fig.3.9: A simulated ultrasound backscatter signal composed of three (normal-apoptotic-normal) layers.

This signal is $3.5 \mu s$ in length with the two normal layers located from 0 to $1.2 \mu s$ and 2.4 - $3.5 \mu s$ (dashed lines in Fig. 3.9), resulting in the apoptotic portion slotted in between these two time points (1.2 - $2.4 \mu s$).

To segment the three-layer simulated ultrasound backscatter signal into locally stationary components, an adaptive Recursive Least Squares Lattice (RLSL) filter was implemented and applied (Section 2.3). By sending this signal to the RLSL filter, the convergence factor (a parameter in segmentation algorithm) that is used to determine the segments is tracked to inspect the changes of signal statistics and consequently determine the segment boundaries. The result of plotting the convergence (gamma) factor is tested with different orders of the adaptive RLSL filter (Section 2.3.2) to determine a proper filter order which simplifies segmentation process. Figure 3.10 shows a simulated ultrasound backscatter signal Fig. 3.10.a and its segmentation results using RLSL filter of order 14 (Fig. 3.10.b).



(a)



(b)

Fig.3.10: a) A simulated ultrasound signal from a normal-apoptotic-normal (NAN) layer system with each layer boundary denoted by a dash line. b) The plot of convergence factor of an adaptive RLSL filter of order 14 with the NAN layer simulated ultrasound signal of figure "a" used as the filter input. The arrows denote the interval that has the most abrupt variations in the convergence factor.

The stationary boundaries of RLSL signal segmentation technique are directly related to the location of the threshold, but for ultrasound simulated signals this threshold is set by visual inspection (line $y=0.9873$ in Fig. 3.10.b), because there is not a unique threshold for all different samples of ultrasound signals. This uncertainty of choosing the threshold is the main drawback of RLSL technique. In Fig. 3.10.b the noticeable difference of convergence factor perturbations between points A and B led to the choice of the location of threshold to $y= 0.9873$ which resulted in 9 signal portions denoted in this figure by vertical dashed lines and marked by labels 1 to 9. In the following section each segment is modeled individually by a 15th order AR process and the model parameters are used to classify each segment.

After the signal segmentation phase each stationary portion is sent to machine learning classifiers. These are the same classifiers in the previous section (Section 3.2) that were trained, using normal and apoptotic ultrasound simulated training data. Therefore each classifier uses its previously fixed weights and parameters to determine the possible class (normal or apoptotic) of new testing data. The classification results are shown for each individual segment in Table 3.2.

TABLE 3.2
Classification Algorithms Accuracy for Simulated NAN Segmented Ultrasound Signals

	Portion (1) Norm.	Portion (2) Apop.	Portion (3) Apop.	Portion (4) Apop.	Portion (5) Apop.	Portion (6) Norm.	Portion (7) Norm.	Portion (8) Norm.	Portion (9) Norm.	(2) to (5)
Fisher's Linear Discrimi nant	N*	A**	A	A	A	N	N	N	A	A
NN with Sigmoid function	N	A	A	A	A	N	N	N	A	A
Network Output	0.9	0.09	0.09	0.11	0.10	0.89	0.90	0.90	0.17	.09
NN with Tanh function	N	A	A	A	A	N	N	N	A	A
Network Output	0.90	0.12	0.11	0.21	0.13	0.86	0.92	0.91	0.09	.12

* Normal

** Apoptotic

These results show that portions (1) and (6-8) belong to normal class while the four middle portions (2-5) are most likely apoptotic. As seen in Table 3.2, portion 9 which is expected to be normal is classified to apoptotic group with all the algorithms. A careful insight into figure (3.10.a) reveals that the end portion of the simulated signal is almost flat (from samples 1387-1420). This simulation “artifact” is the reason for misclassification of portion (9).

The last column of Table 3.2 depicts the case where the 3 middle segments are considered as one segment. As seen in Table 3.1 all the classifiers identified this segment to be apoptotic. Comparing the boundaries of this segment to the actual apoptotic boundaries of the original signal reveals that the middle segment approximately corresponds to the apoptotic layer in the original signal (in the original signal the apoptotic layer is located between 492-983 samples; and the middle apoptotic segment in our results is located between 528-897 samples). This similarity demonstrates that the combination of multiple-layer simulated ultrasound backscatter signals segmentation and classification techniques can be used to determine the proper class of different signal layers in clinical applications. The main drawback of this technique is that in the segmentation method it is not possible to set a unique threshold for all samples of simulated ultrasound signals.

Therefore we can conclude that by applying the methods explained in Chapter 2, we achieved good results for classification and segmentation accuracy of simulated ultrasound backscatter signals from normal and apoptotic cells based on a theoretical modeling for these signals (Section 2.4) [10, 148] and a user-defined threshold for their segmentation.

In the next chapter, I will apply these methods to ultrasound backscatter signals acquired from biological experimental models to verify their sensitivity and specificity for cells.

Chapter 4

Classification & Segmentation of Experimental Ultrasound Signals

After studying the feasibility of the methods presented in Chapter 2 for modeling and classifying simulated ultrasound backscatter signals, these algorithms are tested with experimental ultrasound backscatter signals from 3 groups of cell pellets: normal, apoptotic and cells that were left to decay to verify their accuracy in clinical applications (Fig. 4.1). These experiments were conducted at the Ontario Cancer Institute (OCI) of the Princess Margaret Hospital¹⁶. The data acquisition methods are explained in detail in Section 2.5.

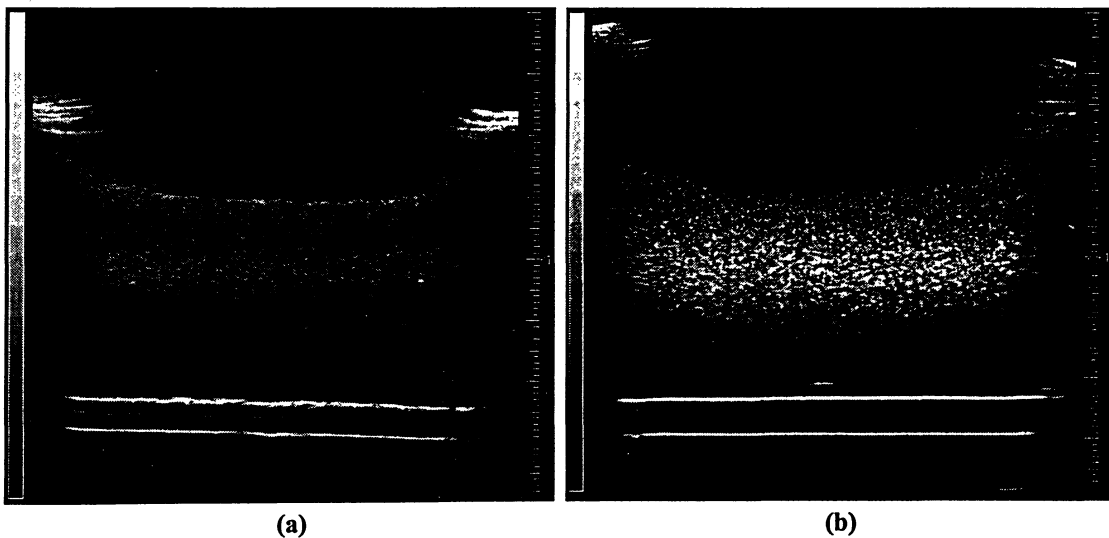


Fig.4.1. a) Representation of a “normal” cell pellet. b) Representation of an “apoptotic” cell pellet, treated with cisplatin for 24 hours.

¹⁶ <http://www.uhnres.utoronto.ca/istitutes/html/oci/mp/ocimp.html> .

This chapter focuses on characterization and classification of experimental ultrasound backscatter signals for cell pellets. The last section (Section 4.4) applies the same analysis for a multi layer cell pellet experiment which is similar to the simulations of Section 3.3.

4.1. AR Modeling of Ultrasound Backscatter Signals from Cells

In this section the stationarity of the cell backscatter signals and their proper AR modeling order are examined (Section 2.1). The experimental ultrasound backscatter signals are approximately $5\ \mu\text{s}$ in time length corresponding to a thickness of $7.5\ \text{mm}$ (since the speed of sound is $1.5\ \text{mm}/\mu\text{s}$). Since these data are relatively homogenous (Fig. 4.1) we expect that the backscatter signals from pellets are stationary or quasi-stationary (the signal statistics change but they fluctuate around equilibrium). In order to verify this, the same approach in Section 3.1 is used to confirm the stationarity of simulated ultrasound backscatter signals. The results of PSD comparison between original and AR-modeled ultrasound signals from three samples of normal, apoptotic and decay cell pellets are shown in Fig. 4.2 (using the 20 MHz transducer). The same analysis was performed on several other ultrasound backscatter signals from these three types of cells (cell pellets) and all the results confirmed that these signals meet the stationary requirement of AR modeling by showing a similar correspondence in their PSD plot to the original signals (as in Section 3.1).

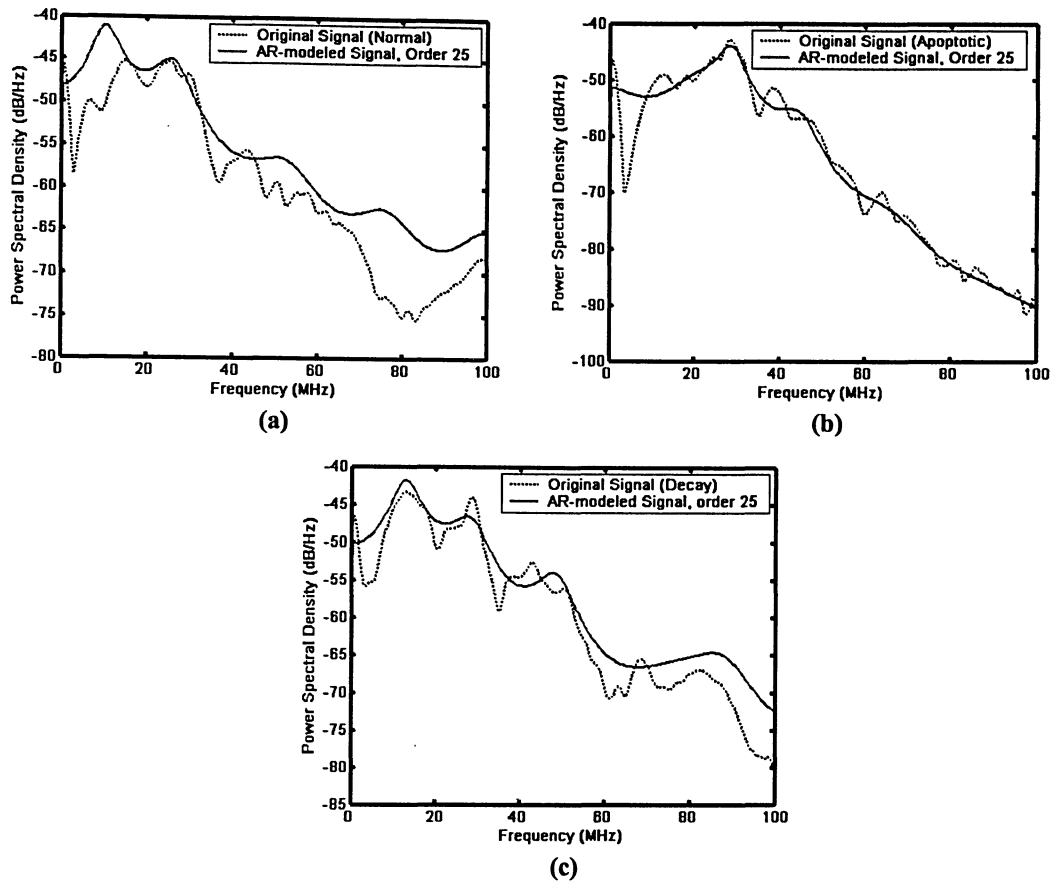


Fig. 4.2: Power Spectral Density of an ultrasound backscatter signal from a) Normal cell pellet, b) Apoptotic cell pellet; and c) Decay cell pellets. The signals are modeled using 15th order AR- Burg process compared to the PSD of the original signal calculated using the Welch method.

Using error criteria (Section 2.1.3) the errors between the PSDs of the reconstructed and original ultrasound signals were calculated and averaged over 35 normal and apoptotic sample RF lines respectively (Fig. 4.3). The variance of the estimated noise was calculated as the output of an AR filter with the calculated AR coefficients as its parameters and the original signal as its input. The result of averaging the variances of the estimated noise over 35 samples is shown in Fig. 4.4.

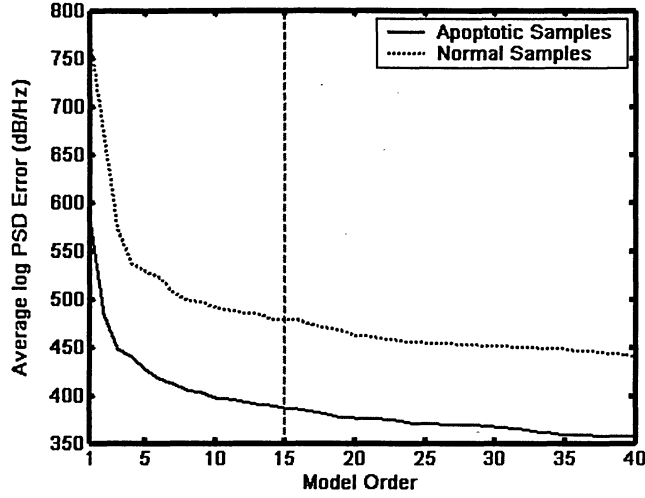


Fig. 4.3: Average Ensemble Error between the PSD of estimated and original ultrasound signal (35 samples of normal and apoptotic signals).

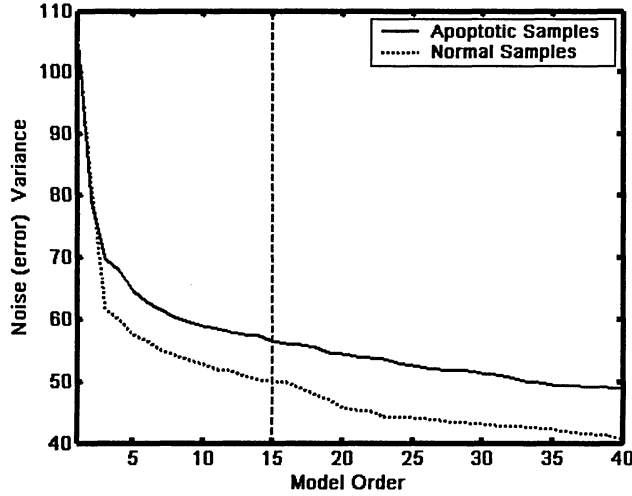


Fig. 4.4: Average variance of the estimated model noise based on the estimated AR coefficient (35 samples of normal and apoptotic signals).

In the ensemble error graph (Fig.4.3), the log-error (averaged over 35 samples from each class) has been decreased by 38% from model order 1 to 15 in comparison to only 2% decrease between model orders 16 and 40. For the normal samples, this error drops by 33% and 7% in these the two intervals mentioned above. Fig. 4.3 also demonstrates that the error curve of normal modeling falls below apoptotic modeling. This means that the AR modeling fits to normal signals better than apoptotic ones. The reason of this better fit can be the fact that the ultrasound backscatter signals from normal cells contain less perturbation in comparison to apoptotic cells. The result of comparing the modeling noise autocorrelation of normal and apoptotic signals, shown in Section 4.1, also confirms that the AR technique is a better modeling for normal backscatter signals.

Similar to its simulated counterpart, the variance of estimated model noise error does not change substantially after model order 15. There is an approximate 47% decrease in the average error variance between model order 1 to 15 for both normal and apoptotic samples but it drops with 17% and 12% rate between orders 16 to 40 (for normal and apoptotic samples respectively). Therefore, similar to the results with simulated data, the error graphs for experimental ultrasound backscatter signals modeled with different orders corroborate that order 15 ($p=15$) is a good choice for AR-modeling of high frequency ultrasound backscatter signals.

4.2. Experimental Ultrasound Backscatter Signal Over-fitting

For the normal and apoptotic simulated signals, it was found that AR modeling with order 15 does not increase the modeling error (no over-fitting) with SNR of 3, 6 and 7 dB. The over-fitting analysis in experimental ultrasound backscatter signals is based on the assumption that the noise associated with experimental signal is a random white noise with normal distribution. The same result of over-fitting analysis from simulated signals (Section 3.1.3) is used for experimental signals and modeling order 15 is chosen as a putative AR modeling order for ultrasound backscatter signals from all three cell conditions.

We first modeled 3 experimental normal, apoptotic and decay ultrasound backscatter signals with order 15 using the Burg method, and then reconstructed these signals with an AR linear filter (Section 2.1.4). The auto-correlation of modeling error (noise) is calculated (Fig. 4.5.a-c) as an estimate of the “whiteness” of the inherited noise of the AR modeling technique and compared to the autocorrelation of a generated white noise (Fig. 4.5.d). As depicted in these figures the auto-correlation of the noise that is the AR-modeling error is very similar to autocorrelation of “white noise” (with a jump at lag zero and small perturbations around zero value at other lags).

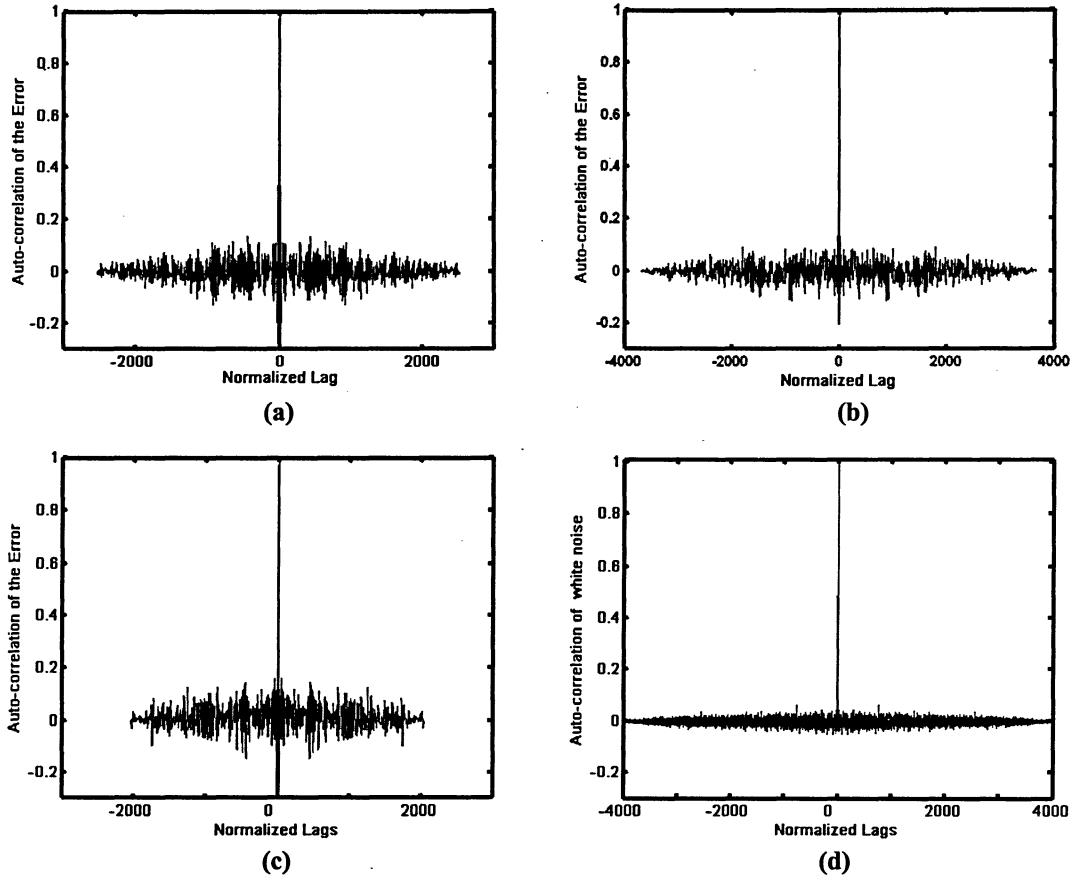


Fig. 4.5: Auto-correlation of the estimated model error (noise) for ultrasound backscatter signals from: a) Apoptotic, b) Normal, and c) Decay cell pellets compared to: d) Auto-correlation of a random white noise.

These results confirm the expectation that the AR modeling noise is a *white noise*. There is a higher correlation of the error to white noise in the case of normal backscatter signals. This can be caused by the fact that ultrasound backscatter signals from normal cells contain fewer perturbations relative to other two classes which make the statistics of these signals less variant. Therefore we expect that the backscatter signals from normal cells would better satisfy the “stationarity” assumption of the AR modeling. The other possibility is that AR-modeling fits to normal backscatters better than the other two types of signals (apoptotic and decay) for this model order. Overall these figures indicate that AR modeling technique with order 15 is a good approximation for ultrasound backscatter signals. These results confirm the conclusion of analyzing the ensemble modeling error in Section 4.1 that showed a better fit to normal signals (Fig. 4.3 and 4.4).

4.3. Ultrasound Backscatter Signal Classification

The classification methods that are used are based on ultrasound backscatter signal *model parameters*. This means that instead of using the entire signal as the input, only 15 AR coefficients that are generated using a 15th order AR-process are passed as the input of ML algorithms in the training and testing phase. The classification between normal-apoptotic, apoptotic-decay and normal-decay cell groups (Section 2.5.1) are tested with a combination of ultrasound backscatter signals from 3 different cell types from each group (Section 2.5).

The machine learning classification algorithms were implemented based on a two-class separation (a combination of normal, apoptotic and decay groups) and according to the methods explained in Section 2.2. The NN is the only classifier with an iterative algorithm to update its parameters and consequently reducing errors. Due to this unique feature of NN among all the other classification algorithms studied in this work and to estimate a proper number of iterations required for network training, the networks performance was determined by calculating the average training error with different number of iterations (Fig. 4.6.a-d). To evaluate the error, the output of NNs was set to 0.9 for normal and 0.1 for apoptotic training samples and the error was incremented by one unit when a sample from one class had an output which was closer to the default output (0.9 or 0.1) of the other class. As depicted in these figures more training iterations result in lower training error until the network reaches an equilibrium level in this error (20,000 iterations for Sigmoid NN in Fig. 4.6.a; and 10,000 iterations for Tanh NN in Fig. 4.6.b).

Another way to verify that an order of 15 is a good AR modeling order is to use testing and training error curves (Fig. 4.6.a-d). If over-fitting occurs in NNs, we would see an overall increase in “testing” error curve despite the decrease in “training” error and the two curves will eventually diverge. As seen in Fig. 4.6.c-d although there is an increase in testing error in the first thousands of iterations but this error curve is eventually decreasing with the number of iterations. This suggests that over-fitting does not happen with AR modeling order of 15.

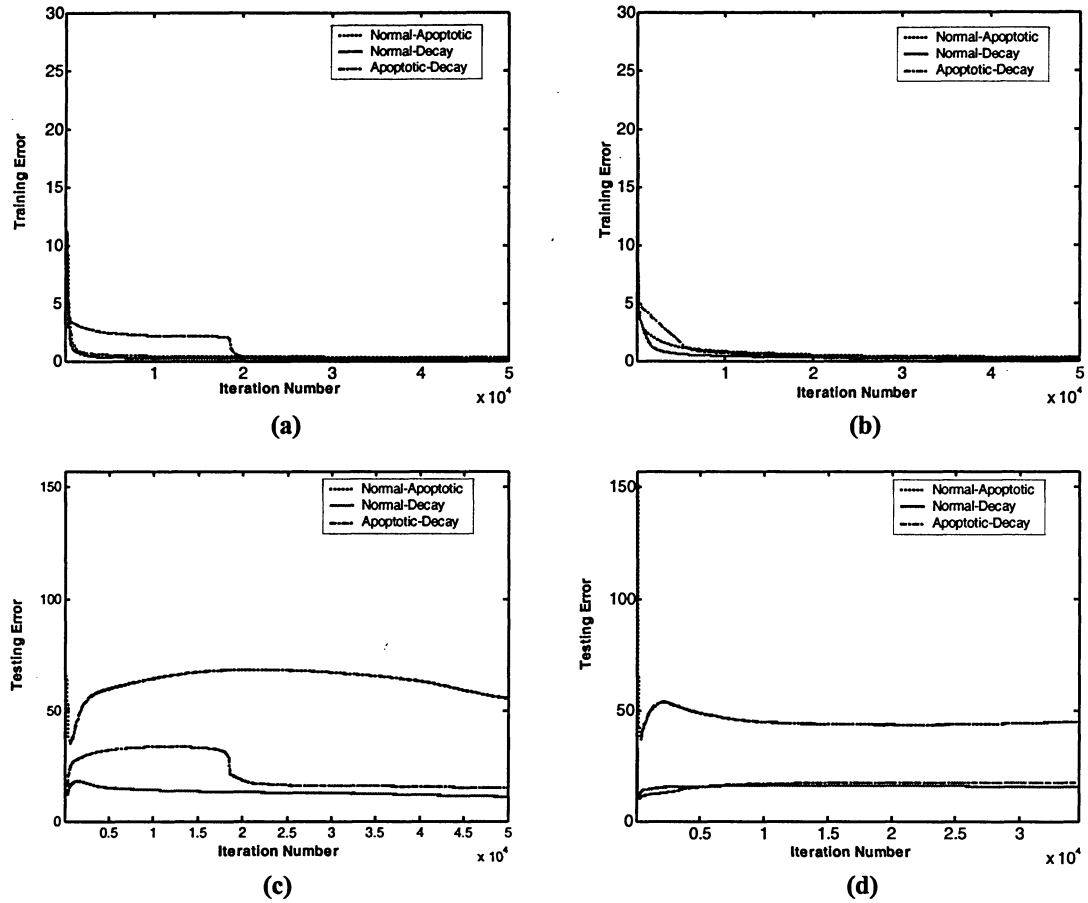


Fig.4.6: Effect of number of iterations in training a NN with different activation functions. a) Training error of NN with Sigmoid activation function. b) Training error of NN with Tanh activation function. c) Testing error of NN of part (a) with sigmoid activation function; and d) Testing error of NN of part (b) with Tanh activation function.

In the following section the classification accuracy of each machine learning classifier is determined with the same training and testing data set and the results are shown in Table 4.1 and Figures 4.7-4.9 for normal-apoptotic, normal-decay and apoptotic-decay groups. The classification experiment was performed in 2 phases.

The first 100 normal and 100 apoptotic samples with data collected using the 20 MHz transducer were used to train the 4 ML classifiers and the classification rate was evaluated using a different set of 100 normal and 180 apoptotic samples which were chosen from the same cell pellet used in the training phase (except for Gaussian classifier where 100 apoptotic samples were used for testing). The result of this analysis is shown in Table 4.1. As seen in this table, the first two linear classifiers (Gaussian and Naïve-

Bayes) show low accuracy in comparison to the NN and Fisher Discriminant. The accuracy of Fisher linear discriminant among the other linear classifiers is most likely due to its characteristic in projecting the data in a way that maximizes the two class separation.

TABLE 4.1
Classification Accuracy for Normal vs. Apoptotic groups

Algorithm	Normal Accuracy	Apoptotic Accuracy
Conditional Gaussian Classifier	40%	60%
Naive Bayes Classifier	46%	77%
Fisher's Linear Discriminant	98%	64%
NN with Sigmoid activation function	93.8%	99%
NN with tanh activation function	95.5%	99%

In the second phase and after choosing the top 3 candidate classification algorithms, the biological samples that are prepared as explained in Section 2.5.1 are used to analyze their ultrasound backscatter classification accuracy using 40 MHz ultrasonic transducer. Figures 4.7-4.9 show the chart diagram of the classification results. There were 35 normal training samples that were chosen from pellet one (P1) zero hour before it was treated with cisplatin ("0 h") (Section 2.5.1, part (a.1)); therefore we can assume that all of these samples are normal. The 35 apoptotic training samples were also chosen from pellet 1, after it was treated with cisplatin for 48 hours ("48 h") (Section 2.5.1, part (b.1)), so most of these samples are apoptotic¹⁷. Finally the 35 "decay" training samples were selected from a pellet (P3) which contained cells that were kept for 12 hours, ("12 h") (Section 2.5.1, part (c.1)).

In total, 80 samples from each of the three biological groups mentioned above (*N-0h-P1*, *A-48h-P1* and *D-12h-P3*) are generated where 35 of them are used for training the classifiers (these groups which have been used for training are labeled in "red" in the following figures). The rest of the 45 ultrasound backscatter samples are combined with

¹⁷ The state of being apoptotic is not absolute as different cells respond differently to apoptosis.

one or two sets of data from the same category (normal, apoptotic or decay) but a different sub-class (there are 3 sub-classes for each category with a different exposure time to cisplatin or decay) to evaluate the ML classifier accuracy for normal, apoptotic and decay groups with time course data (note that the testing result which is shown *excludes* the training samples). Therefore the size of the testing data depends on the number of sample groups used for analysis (35, 115 or 195 for one, two or three-group analysis).

All of the classification figures show very high accuracy between all of the normal samples. This confirms with the biology as we expect that normal cells always share the same attribute regardless of the pellet or sample used. As for the decrease in the second column of Fig. 4.9 the biological explanation is that after 3 hours treatment with cisplatin there are some cells that have been undergoing apoptosis (therefore the normal classification accuracy which reflects the population of normal cells decreases). The data corresponding to these figures are presented in Appendix D.

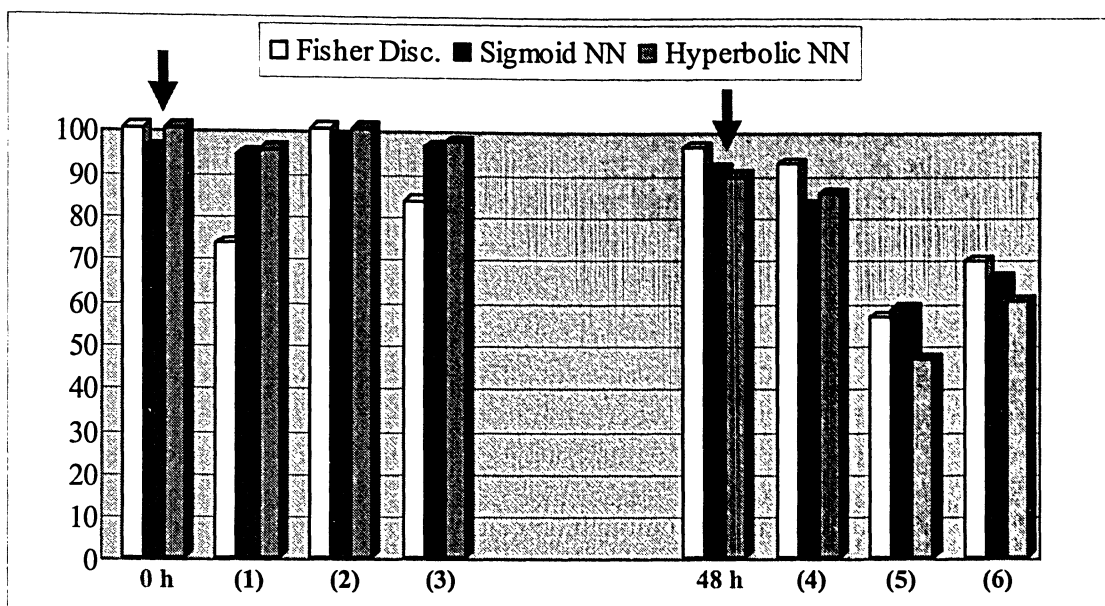


Fig. 4.7: Normal vs. apoptotic classification using 3 ML classifiers. The left and right bars represent normal and apoptotic classification accuracy respectively. The 35 training samples (different from testing) are chosen from the biological group denoted by red arrow (in Figures 4.7-4.9).

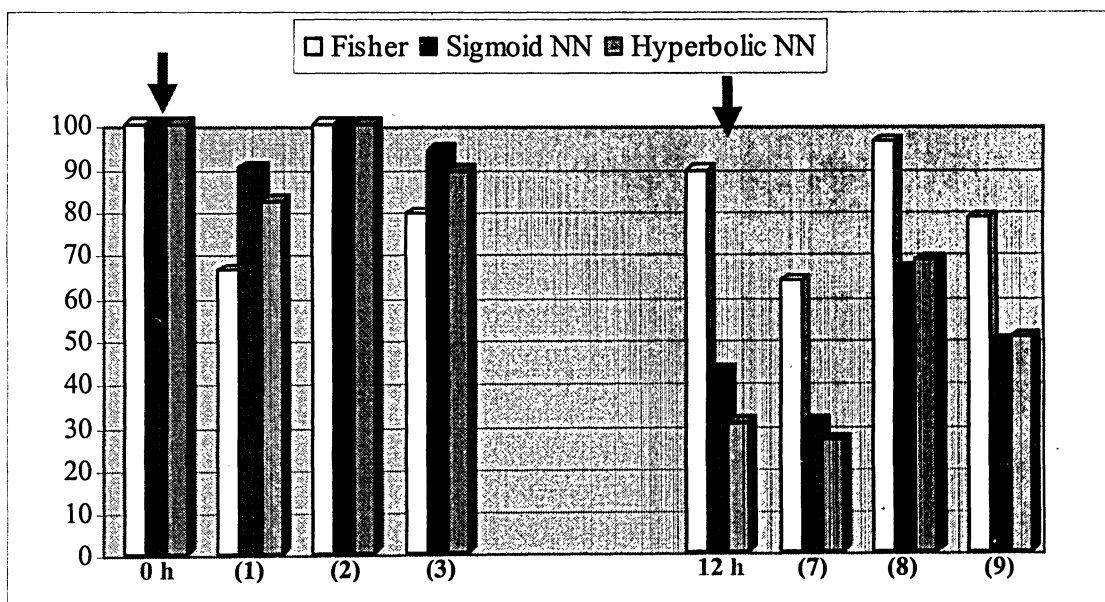


Fig. 4.8: Normal vs. decay classification using 3 ML classifiers. The left and right bars represent normal and decay classification accuracy respectively.

Normal testing sets:

0 h: Data is collected zero hours after the cells have been treated with apoptotic induced drug

(1): {0 h, 3 h} data set is the combination of 0-hour data and the data 3 hours after the treatment

(2): {0 h, 0 h (P2)} data is the combination of 0-hour data from the pellet used for training normal class and the 0-hour data from a different pellet (not used in the training)

(3): {0 h, 0 h (P2), 3h} data is the combination of groups (2) and (3)

Apoptotic testing sets:

48 h: Data is collected 48 hours after the treatment

(4): {48 h, 12 h} data is the combination of the 48-hour data (previous set) and the data 12 hours after the treatment

(5): {48 h, 48 h (P2)} data is the combination of 48-hour data from the pellet used for training and 48-hour data from a different pellet (not used for training)

(6): {48 h, 48 h (P2), 12 h} data is the combination of groups (5) and (6)

Decay testing sets:

12 h: Data is collected 12 hours after the cells have been kept without their nutrients in *decay* condition

(7): {12 h, 9 h} this data set is the combination of cells that have been undergoing *decay* for 12 (previous set) and 9 hours

(8) : {12 h, 12 h (P2)} data is the combination cells that have been undergoing decay for 12 hours from the pellet used for training decay classifier and the 12 hour decay from a different cell pellet (not used for training)

(9): {12 h, 12 h(P2), 9 h} data is the combination of groups (8) and (9)

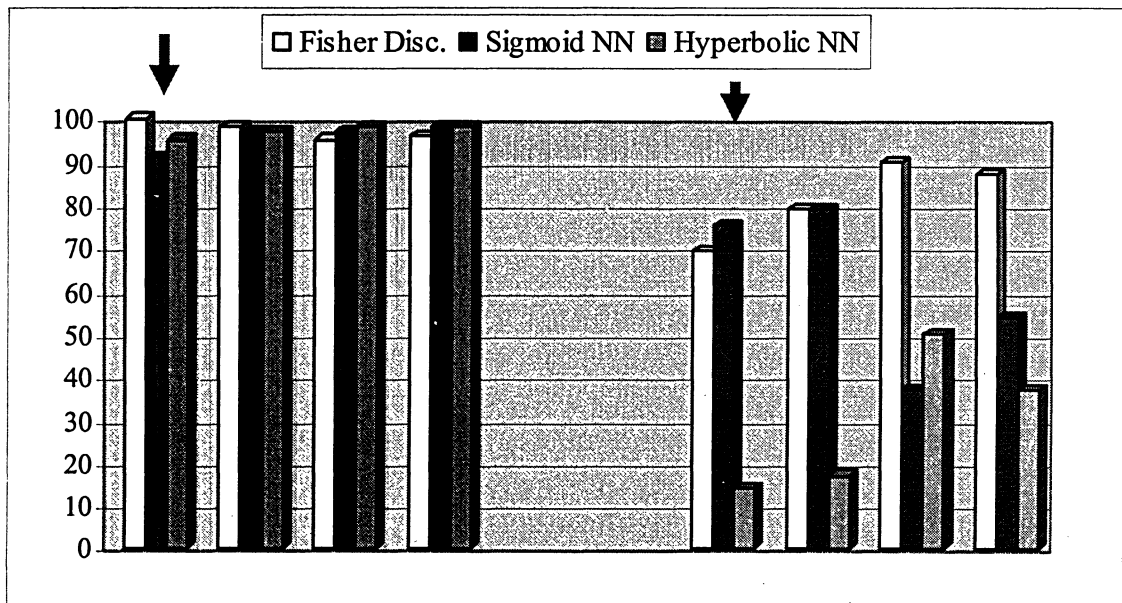


Fig. 4.9: Apoptotic vs. decay classification using 3 ML classifiers. The left and right bars represent normal and decay classification accuracy respectively.

These figures also reveal a striking result that apoptotic vs. decay classification has better accuracy compared to normal vs. decay classification. This implies that the former cells show more distinguishable features. Another result is that Fisher discriminant shows better accuracy for decay samples compared to non-linear NNs. This discriminative classifier is the only one that tries to maximize the distance between two classes instead of just learning the data pattern. So one hypothesis is that decay cells differ from normal

and apoptotic ones in such a way that Fisher discriminant is most capable their classification.

4.4. Adaptive Signal Segmentation and Classification of Ultrasound Backscatter Signals from Cell Pellets Using the RLSL Filter

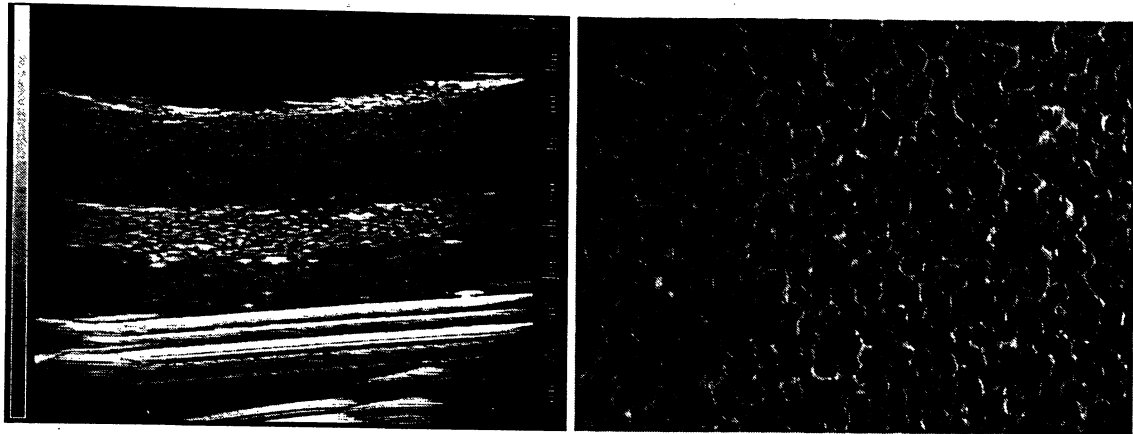
In Section 2.3.2 it was shown that signal segmentation methods can be used to separate the ultrasound backscatter signals from tissues into distinct layers and classification can be applied on each layer respectively. The purpose of such analysis is to enable processing of ultrasound backscatter signals from tissues in which there are layers of different cells with different populations of cell types. Moreover this approach can be later developed to locate tumors using ultrasound. To test algorithms developed in an experimental system, segmentation was applied in vitro and on a 3 layer normal-apoptotic-normal cell pellet (Section 2.5.2). RLSL adaptive segmentation was performed on 16 RF lines of ultrasound backscatter signals from a layered pellet with a layer of apoptotic cells in between layers of untreated cells (normal) pellet which its histology is shown in Fig.4.8.a-b.

The results of segmentation process (Fig. 4.9.c-d) show the average of 6 non-stationary segments for the ultrasound backscatter of this pellet. A striking result from segmentation process is that normal portions are not divided to any more segments which imply that these portions are stationary. These results also reveal that apoptotic backscatters unlike normal ones, show different statistical characteristics (the apoptotic layer is segmented into several portions). After the boundaries were determined each segment was classified to the normal or apoptotic group using NNs with sigmoid activation function. In this approach the model parameters which were calculated using pure normal and pure apoptotic samples in Section 4.3 were kept as default model values for the new analysis. The results also show that 100% of the first segment(s) of all these signals with the approximate depth of 900 time samples are classified into the “normal” group (only two samples are found to be normal in a shorter length with 774 and 866

time samples). The same results calculated for the last segment(s) located from 1400 time samples of the ultrasound backscatter signals also indicate that 100% of this portion is scattered from “normal” cells. It can be concluded that normal cells have similar statistical characteristics in various normal cell pellets which is consistent with the theory. The same test is performed on middle segment(s) of all the 16 signals, but this time the classification results reveal that in middle layer there is a combination of normal and apoptotic segments. The middle apoptotic layer boundaries are visually compared with the first and last boundaries of RLSL result and denoted by dashed lines in Fig. 4.9.c and d.

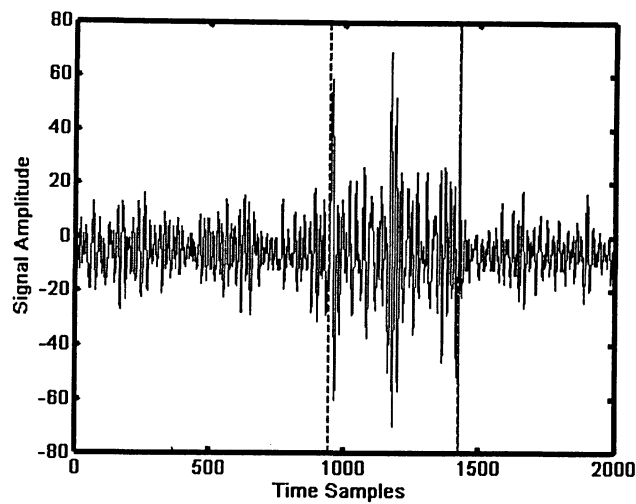
To quantify the proportion of backscatter signals from apoptotic cells in this layer, the percentage of the normal and apoptotic segments were calculated with respect to the total length of the middle layer (this layer is located approximately between 800-1400 time samples determined by visual inspection). Using this approach, it was estimated that an average of 56.5% of the middle layer consisted of “apoptotic” cell types and the other 43.5 % of this layer was formed from “normal” cells. This is consistent with the histology as not all cells exposed to the apoptotic induced drug undergo cell death. These results confirm that apoptotic layer is in fact a mixture of backscatters from normal and apoptotic cell types as theory expected. This result is very useful to monitor the effect of apoptosis induced drugs (such as cisplatin) on cancerous cells and can potentially be used as a method to quantify and compare the effect of different chemotherapeutic drugs on treating these cells.

An example of analyzing the backscatter segments for one particular RF line is shown in Table 4.2 with the original signal segmented into 6 portions and classified with ML classifiers that were trained using the normal and apoptotic samples in Section 4.3.

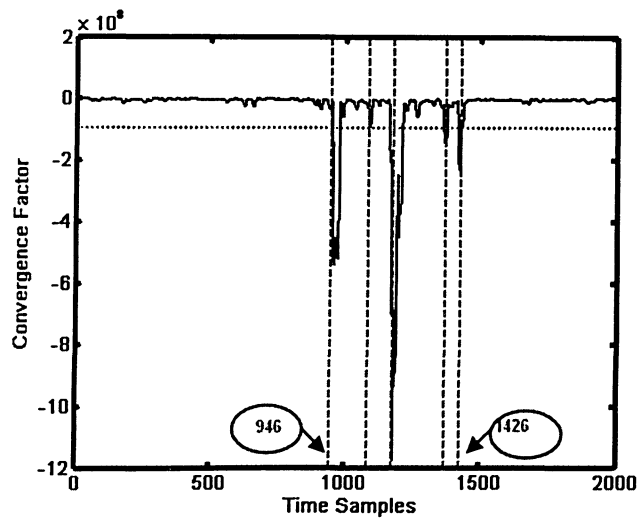


(a)

(b)









(c)



(d)

Fig. 4.10: a) A 3-layer normal-apoptotic-normal cell pellet. b) H&E stains of pellet "a": the top left part of the diagonal line illustrates untreated, the bottom right demonstrates the treated cells. c) Ultrasound backscatter signal from this cell pellet (39th RF line). d) Segmentation of the ultrasound backscatter signal based on convergence factor (RSLS algorithm). The threshold is (-1). The red circles denote the approximate segment boundary for the middle layer which corresponds to the apoptotic portion of the signal.

Table 4.2: An example of classification results of one RF line segments from the pellet shown in Fig.4.30

Seg. #	Segment shape and boundary limits(TS*)	Segment Length (mm)	Classification of segments using NN with Sigmoid activation function	
1	1:947 	2.8	Normal	Normal layer
2	948:1086 	0.42	Normal	
3	1087:1178 	0.28	Apoptotic	Apoptotic layer
4	1179:1374 	0.59	Apoptotic	
5	1375:1426 	0.16	Normal	
6	1427:2000 	1.7	Normal	Normal layer
Middle layer results	Apoptotic Density: 60 % Normal Density: 40 %			

* Time Sample

This comparison confirms the capability of the RLSL segmentation algorithm to successfully detect the non-stationarity boundaries of ultrasound backscatter signals (which separates the two normal layers from the middle one) as we expect that the backscatter signals from normal and apoptotic cells contain different signal features which result in different statistics.

Chapter 5

Conclusions and Future Research

Conclusions

Signal Modeling

In this study, we evaluated the accuracy of AR modeling for ultrasound backscatter signals using a rational approach for determining the best AR modeling order. From studying the minimum and maximum modeling order criteria we were able to demonstrate that AR order 15 is a good order preference for the backscatter signals. In earlier studies different AR modeling orders were used for ultrasound backscatter signals ranging from 3 to 150. However in this study we have shown that modeling ultrasound backscatter signals with order 15 gives an acceptable accuracy for both simulated and experimental ultrasound signals from normal and apoptotic group. We can draw the following conclusions from the analysis of AR modeling and the modeling results from experimental and simulated signals:

- 1) The common algorithms to calculate AR coefficients: Burg-lattice, Welch, Yule-walker and Covariance methods are shown to produce similar results for both experimental and simulated ultrasound backscatter signals.
- 2) The results of comparing the PSD of the original and modeled signals with increasing model orders showed that 15 AR parameters are sufficient for classification purposes.

3) The choice of modeling order also depends on the maximum permissible value, as a higher order may result in modeling the noise associated with the ultrasound backscatter signals (over-fitting). However it was shown that with current noise criteria, order 15 is a good choice to prevent over-fitting. We generated the noisy signals with different SNRs by adding a random white noise with different variances to a simulated ultrasound backscatter signal and analyzing the modeling error. This error is defined as the difference between the modeled noisy signal (reconstructed by using a linear filter with 15 AR coefficients as filter parameters) and original “noise-less” signal averaged over 20 samples of apoptotic and normal simulated ultrasound backscatter signals.

4) The autocorrelation of this error showed a high similarity to WGN that confirms the accuracy of the AR modeling for ultrasound backscatter signals (this error is the inherited noise of an AR filter which is a “white noise” by default).

5) We also showed that the AR modeling of backscatter signals from normal cells result in less error and a better fit to the data. This concludes that the ultrasound backscatter signal from normal cells is less noisy in comparison to apoptotic signals which confirms with experiments that show backscatters from normal cells have less perturbations in comparison to apoptotic cells.

Ultrasound Backscatter Classification

In this study we implemented and tested machine learning algorithms to classify ultrasound backscatter signals based on the 15 AR coefficients derived from the first part of the study (signal modeling). Each classifier was trained with the AR coefficients from normal, apoptotic or decay signals and its performance in two class separation was tested using different sets of testing data and their combination. Most work previously done to perform classification relies on using the entire signal as an input for the classifier. However the basic assumption about ultrasound backscatters signals from normal and apoptotic cell pellets is that these signals differ in their intensity and frequency spectrum. Such differences can be preserved in the AR model of the signals. Using this approach, we classified the signals by passing the AR modeling parameters of the backscatter

signals (that contain the main statistical properties of the signal). For training the classifiers with experimental data, the ultrasound backscatter signals from cell pellets that were exposed for 48 hours to apoptotic induced drug, cisplatin, were used as the apoptotic training data set. As for decay training set, the cell pellets that were kept in decay process for 12 hours were exploit.

- 1) The classification of simulated ultrasound backscatter signals resulted in 100% accuracy with all of the linear and non-linear classifiers. This is resulted from the fact that these signals are generated with the change in the spacing of scatterers (regularly for normal and randomly for apoptotic) and therefore they can be differentiated in a simple linear space.
- 2) As for experimental signals this accuracy is significantly less with linear classifiers except for fisher linear discriminant which attains a comparable accuracy with NNs.
- 3) The other striking result of our simulations on experimental data is that almost all of the classifiers showed 100 percent accuracy in detecting normal backscatter signals while their accuracy in detecting apoptotic signals ranged from 14 to 96 percent. This implies that unlike the backscatters from normal cell pellets the apoptotic backscatters consist of more random structures which can not be entirely learned by the machine learning classifier during the training phase, or that not all treated cells were apoptotic. The latter point is also confirmed by cells histology.
- 4) We also showed that the machine learning classifiers have different accuracy for normal-apoptotic, normal-decay and apoptotic-decay signal classification. Therefore one classifier does not have the best accuracy rate for all of the data sets.
- 5) Using several data sets for cells undergoing apoptosis or decay processes at different times and the combinations of these data, we studied the effect of ML classifiers on these time-course data. The results showed that after three hours exposure to cisplatin the majority of backscatters are still classified to normal class with NN classifiers (94 and 95 percent) while this accuracy is much lower for fisher linear discriminant (73 percent for

normal-apoptotic and 66 percent for apoptotic-decay classification). This implies that fisher discriminant is more sensitive towards the training data set for normal group.

6) The time course study also revealed that after 24 hours exposure to cisplatin most of the backscatters belong to apoptotic group. Comparing the proportion of the signals detected to apoptotic group between 24 and 48 hour, we conclude that there is no significant increase in the number of apoptotic cells during this period (average 7.5% increase by the three classifiers). These results also depend on the accuracy of the classification algorithm. Comparison of ML classifiers with time-course data enables us to have a better understanding about the changes that cell pellets undergo during several stages of apoptosis or decay procedures.

7) We also demonstrated the sensitivity of the ML classifiers to the training cell pellet by choosing a different pellet of the same category (normal, apoptotic or decay cells) and testing the classification accuracy. These results showed that normal classification still attains a very good accuracy which emphasizes on the fact that normal backscatter signals have less perturbation (as mentioned earlier). As for apoptotic classification there is a significant decrease in classification accuracy with the new pellet (in average 39% decrease by the three classifiers for normal-apoptotic group) which can be explained by the *noisy* structure of these signals.

8) The study of the classifier dependency to the training cell pellet showed that the ultrasound backscatter signals from decay group do not behave as expected (we expect that classifying a sample from the new pellet would have at most an equal accuracy to the training pellet while in some cases we achieved a better accuracy for decay samples from the new cell pellet or the combination of new and “old” pellets). This may be resulted from the fact that the hypothesis about the change in spectral density of the ultrasound backscatter signal from decay cells has not been yet confirmed as it has for apoptotic cell pellets and therefore backscatter signals may behave totally different in necrosis stage.

Ultrasound Backscatter Segmentation

To extend the application of ultrasound signal characterization and classification, we implemented and studied the adaptive segmentation of these signals using RLSL filters. Segmentation can be important from two perspectives: first the AR modeling and many other parametric signal modeling techniques are only applicable if the signal is stationary. We showed that the ultrasound backscatter signals of our study were almost stationary and mentioned that this may be resulted from the relative short length of the signals due to the limitations of the current technology. However there is more research in progress by Kolios et al. to increase the penetrance of ultrasound transducers. The second significance of studying signal segmentation is that in clinical applications we have backscatter signals that are reflected from different tissue layers. Analyzing such signals requires the application of signal processing tools to first divide the backscatter into stationary parts and then classifying each segment individually. We implemented such case studies for both simulated ultrasound backscatters with 3 layers and an experimental cell pellet which contained 3 layers of normal, apoptotic and normal cells. The segmentation and classification had the following results:

- 1) For both simulated and experimental ultrasound backscatter signals, we obtained more segments than three. This is mainly due to the fact that when comparing normal and apoptotic signals together the average increase in the peaks and intensity of apoptotic segment amplifies the difference between the two segments. Most of the RLSL segments were from apoptotic portion of the signal which is likely to result from the relative increase in the signal intensity and perturbation in apoptosis stage.
- 2) As for simulated signals, the middle layers which corresponded to the apoptotic boundary of the signal with $+180\ \mu m$ and $-93\ \mu m$ differences on both ends, were all classified to apoptotic portion. The other segments were also classified to normal group with 100% accuracy by all of the three classifiers.

3) Segmentation of ultrasound backscatter signals from the experimental three-layer cell pellet showed an almost perfect correspondence to the middle apoptotic portion boundaries of the original signal. The classification of normal flanking normal portions resulted in 100% accuracy. Again this confirms that there is not much change and random behavior in the statistics of the backscatter signals from normal cells. But in classification of the middle segments, we obtained a mixed result of normal and apoptotic cells. This mixed result is most likely due to the fact that clinically cells react differently to apoptosis induced drugs and therefore not all cells are in the same stage of apoptosis at the same time and also some probably do not even respond to the cisplatin treatment. This can result in the situation where the backscatters from cells that are assumed to be apoptotic contain both apoptotic and normal signals. In the example of Chapter 4.4, we found that 65% of the middle segment is apoptotic which proves that the majority of backscatters in this layer are from apoptotic cells.

Future Work

Signal characterization

In the case of signal modeling and characterization there are potential areas of development that may improve the general accuracy of classification phase or help to obtain a better insight into biomedical ultrasound signals. First we can study the noise and its role in characterization of the signals. Currently we assumed that the noise of the ultrasound backscatter signal is an additive white noise, although this assumption has not yet been proved. As the structure of this noise is not well understood we were unable to study over-fitting with experimental data. However there are adaptive signal processing methods available to remove the signal noise. These tools can be used to analyse the effect of noise in characterizing the ultrasound backscatter signals in a more precise way. Also it is worth to study the characterization of each group of normal, apoptotic and decay backscatter signals in more detail and see the variation of AR coefficient among several samples. This information can help us to better understand the ultrasound backscatter signals from these cells and the variation in their modeling.

Classification

In classification phase, it is very useful to use a combination of multiple ML classifiers. This enables us to use the best possible algorithm for each set of data being tested. Also it is worth to further develop ML algorithms and test the accuracy of classifiers which are based on Kernel analysis. Also it was shown that the length of the signal is another parameter which affects the classification result. Therefore it is worth to analyze the best data fit to each classifier. This can be used as a reference in future applications where the proper classifier to detect a signal can be chosen based on properties of the signal such as its variance, intensity or length.

Segmentation

In signal segmentation, we obtained the segment boundaries based on a user-defined threshold which was set by visual inspection. A potential development for this process is to estimate the position of this threshold automatically based on the statistics of the signal. One method to achieve this may be analyzing many individual normal and apoptotic backscatters and determining their statistics such as mean or instantaneous signal power over different time periods. This can be used as a basis for determining a threshold that differentiates possible boundaries in all backscatter signals from one class. It is also worth to analyze individual boundaries by assigning a “p-value” to each boundary. The *p-value* which is a statistical parameter is the likelihood probability that our measurement is caused by random. Based on this probability we can rank and evaluate the boundaries in a signal in a more precise way.

Finally, if proved that different multi layered signals have the same segmentation threshold, we could be able to rely entirely on signal segmentation and processing techniques to both determine the signal segments and classify them based on each groups specific statistics.

Appendix A

Maximum Likelihood Estimation for Conditional Gaussian Classifier

$$p(y = k) = \alpha_k \quad (\text{A.1})$$

$$p(\mathbf{x} | y = k) = (2\pi \sigma^2)^{-D/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^D (x_i - \mu_{ki})^2 \right\} \quad (\text{A.2})$$

α_k is the priori on class k , σ^2 is the shared variance of all features in all classes, and μ_{ik} is the mean of the feature i in class k . Assuming that all the data are independently and identically distributed (iid) the likelihood expression is:

$$\begin{aligned} l(\theta; D) &= \log p(\mathbf{x}, \mathbf{y} | \theta) = \log \prod_{m=1}^M p(x^m, y^m | \theta) \text{ since data are iid} = \\ &= \sum_{m=1}^M \log p(x^m, y^m | \theta) \quad : \text{ Since data is iid} \\ &= \sum_{m=1}^M p(x^m | y^m, \theta) p(y^m | \theta) \\ &= -\frac{MD}{2} \log(2\pi \sigma^2) - \frac{1}{2\sigma^2} \sum_{m=1}^M \sum_{i=1}^D (x_i^m - \mu_{y^m i})^2 + \sum_{m=1}^M \log \alpha_{y^m} \quad (\text{A.3}) \end{aligned}$$

The maximum likelihood is calculated by differentiating Eq. (A.3) with respect to μ_{ki} and assigning this differentiation to zero. Therefore:

$$\frac{dl}{d\mu_{ki}} = \frac{1}{\sigma^2} \sum_{m=1}^M (x_i^m - \mu_{y^m_i}) [y^m = k] \quad \text{where } [y^m = k] = 1 \text{ if } y^m = k \text{ and zero otherwise.}$$

$$\frac{dl}{d\mu_{ki}} = 0 \Rightarrow \mu_{ki} = \frac{\sum_{m=1}^M x_i^m [y^m = k]}{\sum_{m=1}^M [y^m = k]} \quad (\text{A.4})$$

Equation (A.4) describes that shared mean μ_{ki} of a specific data feature (feature i) in a specific class (k) is the mean value of the i -th feature (x_i) over all the data samples \mathbf{x} that belong to class k . The same approach is used to find the shared variance (σ^2) of the data by differentiating Eq. (A.3) with respect to σ^2 :

$$\begin{aligned} \frac{\partial l(\theta; D)}{\partial \sigma^2} = 0 &\Rightarrow \frac{1}{\sigma^2} \sum_{m=1}^M \sum_{i=1}^D (x_i^j - \mu_{y^m_i})^2 = D \times M \\ &\Rightarrow \sigma^2 = \frac{\sum_{m=1}^M \sum_{i=1}^D (x_i^j - \mu_{y^m_i})^2}{D \times M} \end{aligned} \quad (\text{A.5})$$

Equations (A.4) and (A.5) can be used to calculate $p(\mathbf{x}|\mathbf{y})$ from Eq. (A.2). This value is then used to calculate the most probable class for data set ' \mathbf{x} ' using Bayes rule:

$$p(y|\mathbf{x}) = \frac{p(\mathbf{x}|\mathbf{y})p(\mathbf{y}=\mathbf{k})}{\sum_k p(\mathbf{x}|\mathbf{y}=\mathbf{k})p(\mathbf{y}=\mathbf{k})} \quad (\text{A.6})$$

Appendix B

Maximum Likelihood Estimation for Naïve Bayes Classifier

The Naïve based approach to classify the new data is to assign the most probable target that best describes the instance $\mathbf{x} = \{x_1, x_2, \dots, x_i\}$ where $\langle x_i \rangle$ is the set of features that describe the data instance \mathbf{x} :

$$\begin{aligned} c(\mathbf{x}) &= \arg[\max_c p(c | \mathbf{x})] \\ &= \arg \left[\max_c \frac{p(\mathbf{x} | c) p(c)}{p(\mathbf{x})} \right] \\ &= \arg[\max_c p(\mathbf{x} | c) \cdot p(c)] \end{aligned} \tag{B.1}$$

In Eq. (B.1), $c(\mathbf{x})$ denotes the target value output. The two probabilities in this equation can be estimated based on the training data: $p(c)$ is easily estimated by counting the frequency that each target value “ c ” occurs in the training data. However this approach to calculate the probability is not reliable unless we have a very large training set. To calculate $p(\mathbf{x} | c)$, Naïve Bayes classifier simplifies the estimation process by assuming that given the target value for the training data set, the probability of observing conjunction x_1, x_2, \dots, x_n is just the product of the probabilities for individual features:

$$p(\mathbf{x} | c) = \prod_i p(x_i | c) \tag{B.2}$$

By substituting Eq. (B.2) in (B.1):

$$\begin{aligned}
c(\mathbf{x}) &= \arg \max_c p(c) \prod_i p(x_i | c) \\
&= \arg \max_c [\log p(c) + \sum_i \log p(x_i | c)]
\end{aligned}
\tag{B.3}$$

Using Eq. (B.3) classification is performed using the probability of a class ($p(c)$) and the probability of data given the class ($p(x|c)$). The class is determined using the sign of $c(x)$ for new testing data.

Appendix C

Error Minimization and Optimum Weight Calculation for Neural Networks

In the following derivations, *squared error* is used as cost (performance) function due to its simplicity:

$$E = \sum_k (y_k - o_k)^2 \quad (C.1)$$

Where y_k and o_k denotes network output and desired output from unit k respectively. In order to find the direction of the steepest descent, the derivatives of error (E) are calculated with respect to each component of weight vector. This vector is called *error gradient* and is defined by:

$$\nabla E(\vec{w}) \equiv \left[\frac{\partial E}{\partial w_0}, \frac{\partial E}{\partial w_1}, \dots, \frac{\partial E}{\partial w_n} \right] \quad (C.2)$$

After error gradient is calculated, the weight vector is updated by this gradient in each step:

$$\begin{aligned} \vec{w} &\leftarrow \vec{w} + \Delta \vec{w} \\ \Delta \vec{w} &= -\delta \nabla E(\vec{w}) \end{aligned} \quad (C.3)$$

Where $\delta > 0$ is a constant called *learning rate* that determines the step size in the gradient descent search. The following equations are used to estimate the error gradient for a NN

with sigmoid activation function and squared error. It is assumed that a_{ji} is the i^{th} input to unit j ; and q_j and o_j are the net input and output of unit j respectively:

$$o_j = \sigma(q_j) = \text{sigmoid}(q_j) = \frac{1}{1 + \exp(-q_j)} \quad (\text{C.4})$$

$$q_j = \sum_i w_{ji} a_{ji}$$

Taking the derivatives of equations (C.1) and (C.4) with respect to net input and output of unit j respectively results in:

$$\frac{\partial o_j}{\partial q_j} = o_j(1 - o_j) \quad (\text{C.5})$$

$$\frac{\partial E_n}{\partial o_k} = -2(y_k - o_k) \quad (\text{C.6})$$

As mentioned before the optimum weights are chosen to minimize the cost function (E). Therefore it is necessary to find the minimum of the following term resulted from equations (C.1) and (C.6):

$$\frac{\partial E_n}{\partial w_{ji}} = \frac{\partial E_n}{\partial q_j} \cdot \frac{\partial q_j}{\partial w_{ji}} = \frac{\partial E_n}{\partial q_j} \cdot a_{ji} \quad (\text{C.7})$$

In order to proceed from Eq. (C.7), the cost function rates of change with respect to the net input of all units (hidden and output) must be determined. First this rate is calculated for hidden units as:

$$\frac{\partial E_n}{\partial q_k} = \frac{\partial E_n}{\partial o_k} \cdot \frac{\partial o_k}{\partial q_k} \quad (\text{C.8})$$

Substituting Eq. (C.6) and (C.5) in (C.8) we have:

$$\frac{\partial E_n}{\partial q_k} = -2(y_k - o_k) o_k (1 - o_k) \quad (\text{C.9})$$

By substituting Eq. (C.9) in (C.7), the rate of change of error with respect to hidden unit weights is estimated as:

$$\frac{\partial E_n}{\partial w_{kj}} = -2(y_k - o_k)o_k(1 - o_k)a_{kj} \quad (\text{C.9})$$

As for output units, we assume that $\delta_j = \frac{\partial E_n}{\partial q_j}$, for every unit j . Therefore:

$$\begin{aligned} \delta_j &= \frac{\partial E_n}{\partial q_j} = \sum_i \frac{\partial E_n}{\partial q_i} \cdot \frac{\partial q_i}{\partial q_j} = \sum_i \delta_i \frac{\partial q_i}{\partial q_j} \\ &= \sum_i \delta_i \frac{\partial q_i}{\partial o_j} \frac{\partial o_j}{\partial q_j} \\ &= \frac{\partial o_j}{\partial q_j} \sum_i \delta_i w_{ij} \end{aligned}$$

Substituting Eq. (C.5) results in:

$$\delta_j = o_j(1 - o_j) \sum_i \delta_i w_{ij} \quad (\text{C.10})$$

Equations (C.9) and (C.10) are substituted in Eq. (C.3) to determine how much wiggling the weights of a network will affect the network cost function (error). The weights are updated in each iteration of backpropagation algorithm until the desired error criteria is achieved. The same approach is also used to derive the network equation when the activation function is tangent hyperbolic.

Appendix D

Classification Results for Experimental Ultrasound Backscatter Signals

The description of each data set is explained in Section 2.5. Each number corresponds to the height of the classification bar depicted in Figures 4.7-4.9.

TABLE (D.1)
Classification Accuracy for Normal vs. Apoptotic groups

Testing Sample(s) Algorithm	N. (*) N-0h-P1	A. (+) A-48h-P1	N. N-0h-P1 & N-3h-P1	A. A-48h-P1 & A-24h-P1	N. N-0h-P1 & N-0h-P2	A. A-48h-P1 & A-48h-P2	N. N-0h-P1 & N-3h-P1 & N-0h-P2	A. A-48h-P1 & A-24h-P1 & A-48h-P2
Fisher Linear Discriminant	100	96	73	92	100	56	83	69
NN with Sigmoid activation	95	91	94	82	98	58	96	65
NN with Tanh activation	100	89	95	85	100	46	97	60

TABLE (D.2)
Classification Accuracy for Normal vs. Decay groups

Testing Sample(s) Algorithm	N. N-0h-P1	D. (+) D-12h-P3	N. N-0h-P1 & N-3h-P1	D. D-12h-P1 & D-24h-P1	N. N-0h-P1 & N-0h-P2	D. D-12h-P3 & D-12h-P4	N. N-0h-P1 & N-3h-P1 & N-0h-P2	D. D-12h-P3 & D-9h-P3 & D-12h-P4
Fisher Linear Discriminant	100	89	66	63	100	96	79	78
NN with Sigmoid activation	100	42	90	30	100	66	94	49
NN with Tanh activation	100	30	82	26	100	68	89	50

TABLE (D.3)
Classification Accuracy for Apoptotic vs. Decay groups

Testing Sample(s) Algorithm	A. A-48h-P1	D. D-12h-P3	A. A-48h-P1 & A-24h-P1	D. D-12h-P3 & D-9h-P3	A. A-48h-P1 & A-48h-P2	D. D-12h-P3 & D-12h-P4	A. A-48h-P1 & A-24h-P1 & A-48h-P2	D. D-12h-P3 & D-9h-P3 & D-12h-P4
Fisher Linear Discriminant	100	69	98	79	95	90	96	87
NN with Sigmoid activation	91	75	97	79	97	37	98	54
NN with Tanh activation	95	14	97	17	98	50	98	37

* Normal Sample

+ Apoptotic Sample

* Decay Sample

Appendix E

List of Abbreviation

ACS	auto-covariance sequence
AI	artificial intelligence
AIC	akaike information criterion
AML	acute myloid leukemia
ARMA	autoregressive moving average
AR	autoregressive
DNA	deoxyribonucleic acid
FIR	finite impulse response
H & E	Hematoxylin and Eosin
IID	independantly and identically distributed
MA	moving average
ML	machine learning
NN	neural network
PCD	programmed cell death
PSD	power spectral density
RF	radio frequency
RLSL	recursive least squares lattice (filter)
Sig.	sigmoid function
SNR	signal-to-noise ratio
Tanh.	tangent-hyperbolic function
WGN	white Gaussian noise
WSS	wide sense stationary

References

1. Schweichel, J.U. and Merker, H.J., *The morphology of various types of cell death in prenatal tissues*. Teratology, 1973. 7(3): p. 253-66.
2. Marino, G. and Lopez-Otin, C., *Autophagy: molecular mechanisms, physiological functions and relevance in human pathology*. Cell Mol Life Sci, 2004. 61(12): p. 1439-54.
3. Alva, A.S., Gultekin, S.H. and Baehrecke, E.H., *Autophagy in human tumors: cell survival or death?* Cell Death Differ, 2004.
4. Gorski, S., *Apoptosis and Autophagy*. Research In Progress Seminar, Genome Sciences Center at the BC Cancer Agency, 2004.
5. Group, R.C.D.R., *Apoptosis*.
6. Hanahan, D. and Weinberg, R.A., *The hallmarks of cancer*. Cell, 2000. 100(1): p. 57-70.
7. Chinnaiyan, A.M., Prasad, U., Shankar, S., Hamstra, D.A., Shanaiah, M., Chenevert, T.L., Ross, B.D. and Rehemtulla, A., *Combined effect of tumor necrosis factor-related apoptosis-inducing ligand and ionizing radiation in breast cancer therapy*. Proc Natl Acad Sci U S A, 2000. 97(4): p. 1754-9.
8. Rao, J.Y., Apple, S.K., Jin, Y., Lin, S., Nieberg, R.K. and Hirtschowitz, S.L., *Comparative polymerase chain reaction analysis of c-myc amplification on archival breast fine-needle aspiration materials*. Cancer Epidemiol Biomarkers Prev, 2000. 9(2): p. 175-9.
9. Blankenberg, F.G., Robbins, R.C., Stoot, J.H., Vriens, P.W., Berry, G.J., Tait, J.F. and Strauss, H.W., *Radionuclide imaging of acute lung transplant rejection with annexin V*. Chest, 2000. 117(3): p. 834-40.
10. Czarnota, G.J., Kolios, M.C., Vaziri, H., Benchimol, S., Ottensmeyer, F.P., Sherar, M.D. and Hunt, J.W., *Ultrasonic imaging of viable, dead and apoptotic cells*. Ultrasound in Medicine and Biology, 1997. 23(6): p. 961-965.
11. Turnbull, D.H., Ramsay, J.A., Shivji, G.S., Bloomfield, T.S., From, L., Sauder, D.N. and Foster, F.S., *Ultrasound backscatter microscope analysis of mouse melanoma progression*. Ultrasound Med Biol, 1996. 22(7): p. 845-53.
12. Turnbull, D.H. and Foster, F.S., *In vivo ultrasound biomicroscopy in developmental biology*. Trends in Biotechnology, 2002. 20: p. 29-33.
13. Towfig, F., Barnes, C.W. and Pisa, E.J., *Tissue classification based on autoregressive models for ultrasound pulse echo data*. Acta Electronica, 1984. 26: p. 95-110.
14. Gupta, A.K., Turnbull, D.H., Harasiewicz, K.A., Shum, D.T., Watteel, G.N., Foster, F.S. and Sauder, D.N., *The use of high-frequency ultrasound as a method*

- of assessing the severity of a plaque of psoriasis. Arch Dermatol.*, 1996. **132**: p. 658-662.
15. D. H. Turnbull, J.A.R., G. S. Shivji, T. S. Bloomfield, L. From, D. N. Sauder, F. S. Foster, *Ultrasound backscatter microscope analysis of mouse melanoma progression. Ultrasound in Medicine and Biology*, 1996. **22**: p. 845-853.
 16. Kolios, M.C., Czarnota, G.J., Lee, M., Hunt, J.W. and Sherar, M.D., *Ultrasonic spectral parameter characterization of apoptosis. Ultrasound in Medicine and Biology*, 2002. **28**(5): p. 589-97.
 17. Czarnota, G.J., Kolios, M.C., Abraham, J., Portnoy, M., Ottensmeyer, F.P., Hunt, J.W. and Sherar, M.D., *Ultrasound imaging of apoptosis: high-resolution non-invasive monitoring of programmed cell death in vitro, in situ and in vivo. Br J Cancer*, 1999. **81**(3): p. 520-7.
 18. Ursea, R., Coleman, D.J., Silverman, R.H., Lizzi, F.L., Daly, S.M. and Harrison, W., *Correlation of high-frequency ultrasound backscatter with tumor microstructure in iris melanoma. Ophthalmology*, 1998. **105**: p. 906-912.
 19. Lizzi, F.L., Astor, M., Kalisz, A., Liu, T., Coleman, D.J., Silverman, R., Ursea, R. and Rondeau, M. *Ultrasonic spectrum analysis for assays of different scatterer morphologies: theory and very-high frequency clinical results. in Ultrasonics Symposium, 1996. Proceedings., 1996 IEEE.* 1996.
 20. Dent, C.I., Scott, M.J., Wickline, S.A. and Hall, C.S., *High-frequency ultrasound for quantitative characterization of myocardial edema. Ultrasound in medicine and biology*, 2000. **26**(3): p. 375-384.
 21. Nair, A., Calvetti, D. and Vince, D.G., *Regularized autoregressive analysis of intravascular ultrasound backscatter: improvement in spatial accuracy of tissue maps. IEEE Trans Ultrason Ferroelectr Freq Control*, 2004. **51**(4): p. 420-31.
 22. Y. Zimmer, R.T., S. Akselrod, *An automatic approach for morphological analysis and malignancy evaluation of ovarian masses using B-scans. Ultrasound in Medicine and Biology*, 2003. **29**(11): p. 1561-70.
 23. Zimmer, Y., Tepper, R. and Akselrod, S., *An automatic approach for morphological analysis and malignancy evaluation of ovarian masses using B-scans. Ultrasound in Medicine and Biology*, 2003. **29**(11): p. 1561-70.
 24. Tavathia, S., Rangayyan, R.M., Frank, C.B., Bell, G.D., Ladly, K.O. and Zhang, Y.T., *Analysis of knee vibration signals using linear prediction. IEEE Transactions on Biomedical Engineering*, 1992. **39**(9): p. 959-970.
 25. Tavathia, S., *Analysis of knee vibration signals using linear prediction. Master's thesis, Dept. of Electrical and Computer Eng., the University of Calgary, Calgary, AB, Canada*, 1991.
 26. Moussavi, Z.K., Rangayyan, R.M. and Frank, C., *Screening and Adaptive Segmentation of Vibroarthrographic Signals. IEEE Transaction on Biomedical Engineering*, 1996. **43**(1): p. 15-23.
 27. Krishnan, S., *M.A.Sc. thesis, Department of Electrical and Computer Engineering, University of Calgary, Alberta, Canada.* 1996.
 28. Krishnan, S., Rangayyan, R.M., Bell, G.D., Frank, C.B. and Ladly, K.O., *Adaptive filtering, modelling and classification of knee joint vibroarthrographic signals for non-invasive diagnosis of articular cartilage pathology. Med Biol Eng Comput*, 1997. **35**(6): p. 677-84.
 29. Veropoulos, K., *Machine Learning Approaches To Medical Decision Making.* 2001.

30. Lao, Z., Shen, D., Xue, Z., Karacali, B., Resnick, S.M. and Davatzikos, C., *Morphological classification of brains via high-dimensional shape transformations and machine learning methods*. Neuroimage, 2004. **21**(1): p. 46-57.
31. Lee, E.K., Gallagher, R.J., Campbell, A.M. and Prausnitz, M.R., *Prediction of ultrasound-mediated disruption of cell membranes using machine learning techniques and statistical analysis of acoustic spectra*. IEEE Transactions on Biomedical Engineering, 2004. **51**(1): p. 82- 89.
32. Schwarzer, G., Nagata, T., Mattern, D., Schmelzeisen, R. and Schumacher, M., *Comparison of fuzzy inference, logistic regression, and classification trees (CART). Prediction of cervical lymph node metastasis in carcinoma of the tongue*. Methods Inf Med, 2003. **42**(3): p. 408-423.
33. Bruce, E.N., *Biomedical Signal Processing and Signal Modeling*. 2000.
34. Gardner, W.R. and Rao, B.D., *Mixed-Phase AR Models for Voiced Speech and Perceptual Cost Functions*. Proc. of the International Conference on Acoustics, Speech and Signal Processing, Adelaide, Australia, 1994: p. 19-22.
35. Härmä, A., Juntunen, M. and Kaipio, P., *Time-varying autoregressive modeling of speech and audio signals*. In Signal Processing X: Theories and Applications, 2000: p. 2037-40.
36. Haykin, S., Sayed, A.H., Zeidler, J.R., Yee, P. and Wei, P.C., *Adaptive tracking of linear time-variant systems by extended RLS algorithms*. Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on], 1997. **45**(5): p. 1118-1128.
37. Haykin, S., *Communication systems*. 2001.
38. Makhoul, J., *Linear prediction: A tutorial review*. Proceedings of the IEEE, 1975. **63**: p. 561-580.
39. Yule, G., *On a method of investigating periodicity in disturbed series with special reference to wolfer's sunspot numbers*. Phil. Trans. Roy. Soc, London, 1927. **A**(226): p. 267-298.
40. Walker, G., *On periodicity in series of related terms*. Proc. R. Sot., London, 1931. **A**(131): p. 518-532.
41. Haykin, S., *Adaptive Filter Theory*. 4th ed. 2001: Pearson Education.
42. Orfanidis, J.S., *Optimum Signal Processing: An Introduction*. 2nd ed. 1988: New York : McGraww-Hill.
43. Marple, S.L., *Digital Spectral Analysis with Applications*. 1987: Englewood Cliffs, N.J. : Prentice-Hall.
44. Kay, S.M., *Modern Spectral Estimation: Theory and Application*. 1988: Englewood Cliffs, N.J. : Prentice Hall.
45. Lawson, C.L. and Hanson, R.J., *Solving Least Squares Problems*. 1974: p. 180-198.
46. Söderström, T. and Stoica, P., *System Identification*. 1989.
47. Levinson, N., *The wiener RMS(root-mean-square)error criterion in filter design and prediction*. Journal of Mathematics and Physic, 1947. **25**: p. 261-78.
48. *Signal processing toolbox for use with MATLAB, user's guide, version 5*. p. 3.31-3.32.
49. Burg, J.P., *Maximum Entropy Spectral Analysis*. In Proceedings of the 37 th Annual International SEG Meeting, 1967.

50. Burg, J.P., *The Relationship Between Maximum Entropy Spectra and Maximum Likelihood Spectra*, Geophysics, 1972. 37(2): p. 375-76.
51. Marple, S.L., *Digital Spectral Analysis with Applications*. 1988.
52. Proakis, J. and Monolakis, D., *Digital Signal Processing*. 1996.
53. Kohavi, R. and John, G., *Wrappers for feature subset selection*. Artificial Intelligence, 1997. 97: p. 1-2.
54. Cunningham, P., *Overfitting and Diversity in Classification Ensembles based on Feature Selection*, in *Technical Report*. 2000, Department of Computer Science, Trinity College Dublin, Padraig.
55. Domingos, P. *Bayesian Averaging of Classifiers and the Overfitting Problem*. in *Proceedings of the Seventeenth International Conference on Machine Learning*. 2000. Stanford, CA.
56. Roweis, S., *Machine Learning Lecture Notes*. Department of Computer Science, University of Toronto, Fall 2002.
57. Sakamoto, Y., Ishiguro, M. and Kitagawa, G., *Akaike Information Criterion Statistics*. 1986.
58. Anderson, C.W., Stolz, E.A. and Shamsunder, S., *Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks*. Biomedical Engineering, IEEE Transactions on, 1998. 45(3): p. 277-286.
59. Dingfei, G.E., Srinivasan, N. and Krishnan, S.M., *Cardiac arrhythmia classification using autoregressive modeling*. BioMedical Engineering OnLine, 2002: p. 1-5.
60. Guo, Z., Durand, L.G., Allard, L., Cloutier, G., Lee, H.C. and Langlois, Y.E., *Cardiac Doppler blood flow signal analysis, Part II: The time-frequency representation by using autoregressive modeling*. Medical & Biology Engineering & Computing, 1993. 31: p. 242-48.
61. Pfurtscheller, G., Neuper, C., Schlogl, A. and Lugger, K., *Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters*. Rehabilitation Engineering, IEEE Transactions on [see also IEEE Trans. on Neural Systems and Rehabilitation], 1998. 6(3): p. 316-325.
62. Hadjileontiadis, L.J. and Panas, S.M. *Autoregressive modeling of lung sounds using higher-order statistics: estimation of source and transmission*. in *Higher-Order Statistics, 1997. Proceedings of the IEEE Signal Processing Workshop on*. 1997.
63. Schlögl, A., Flotzing, D. and Pfurtscheller, G., *Adaptive Autoregressive Modeling used for Single-trial EEG Classification*. Biomedizinische Technik, 1997. 42: p. 162-67.
64. Berger, R.D., Saul, J.P. and Cohen, R.J., *Assessment of autonomic response by broad-band respiration*. Biomedical Engineering, IEEE Transactions on, 1989. 36(11): p. 1061-1065.
65. Husby, O., Lie, T., Lang, T., Hokland, J. and Rue, H., *Bayesian 2D Deconvolution: A Model for Diffuse Ultrasound Scattering*. 1999.
66. Goel, V., Brambrink, A.M., Baykal, A., Koehler, R.C., Hanley, D.F. and Thakor, N.V., *Dominant frequency analysis of EEG reveals brain's response during injury and recovery*. Biomedical Engineering, IEEE Transactions on, 1996. 43(11): p. 1083-1092.

67. Rendell, L. and Cho, H., *Empirical Learning as a function of Concept Character*. Machine Learning, 1990. 5(267-298).
68. Jordan, M.I. and Rumelhart, D.E., *Forward models: Supervised learning with a distal teacher*. Cognitive Science, 1992. 16: p. 307-354.
69. Vladutu, L., Papadimitiou, S., Mavroudi, S. and Bezerianos, A. *Ischemia detection using supervised learning for hierarchical neural networks based on kohonen-maps*. in *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*. 2001.
70. Ghahramani, Z. and Jordan, M.I., *Supervised learning from incomplete data via an EM approach*. In *Advances in Neural Information Processing Systems 6*, 1994.
71. Sykacek, P., *Bayesian inference for reliable biomedical signal processing*. 2000.
72. Makeig, S., Bell, A.J., Jung, T.P. and Sejnowski, T.J., *Independent component analysis of electroencephalographic data*. *Advances in Neural Information Processing Systems*, 1996. 8: p. 145-51.
73. Lee, T.W., Lewicki, M.S. and Sejnowski, T.J., *ICA mixture models for unsupervised classification and automatic context switching*, *International workshop on Independent component analysis and blind signal separation*. Aussois, France, 1999: p. 209-14.
74. Schizas, C.N., Pattichis, C.S., Livesay, R.R., Schofield, I.S., Lazarou, K.X. and Middleton, L.T. *Unsupervised learning in computer aided macroelectromyography*. in *Computer-Based Medical Systems, 1991. Proceedings of the Fourth Annual IEEE Symposium*. 1991.
75. Christodoulou, C.I. and Pattichis, C.S., *Unsupervised pattern recognition for the classification of EMG signals*. *Biomedical Engineering, IEEE Transactions on*, 1999. 46(2): p. 169-178.
76. Rezek, I.A. and Roberts, S.J., *Unsupervised Clustering using Metric Space Connectedness*. Research Report TR-97-4, 1997.
77. Yang, J.Y. and Ersoy, O.K., *Combined Supervised and Unsupervised Learning in Genomic Data Mining*. ECE Technical Reports 2003, 2003.
78. Akay, M., *Nonlinear Biomedical Signal Processing, Volume 1: Fuzzy Logic, Neural Networks, and New Algorithms*. 2000: p. 29-30.
79. Sinkkonen, J., *Learning Metrics and Discriminative Clustering*. 2003.
80. Yiu, K.K., Mak, M.W. and Li, C.K. *Probabilistic decision-based neural networks for speech pattern classification*. in *Signal Processing Proceedings, 1998. ICSP '98. 1998 Fourth International Conference on*. 1998.
81. Moriarty, D.E., Schultz, A. and Grefenstette, J., *Evolutionary Algorithms for Reinforcement Learning*. *Journal of Artificial Intelligence Research*, 1999: p. 241-76.
82. Xu, W., Nandi, A.K. and Zhang, J., *Novel fuzzy reinforced learning vector quantisation algorithm and its application in image compression*. *IEEE Proceedings of Vision, Image and Signal Processing*, 2003. 150(5): p. 292-98.
83. Wong, W.K., Moore, A., Cooper, G. and Wagner, M., *Rule-based Anomaly Pattern Detection for Detecting Disease Outbreaks*. *Proceedings of the 18th National Conference on Artificial Intelligence*, 2002.
84. Kwedlo, W. and Kretowski, M., *An Evolutionary Algorithm for Cost-Sensitive Decision Rule Learning*. *European Conference on Machine Learning, ECML'01*. Freiburg, Germany, 2001.
85. Zheng, Z., *Constructing New Attributes for Decision Tree Learning*. 1996.

86. Fürnkranz, J., *Separate-and-Conquer Rule Learning*. Artificial Intelligence Review, 1999. 13(1): p. 3-54.
87. Fürnkranz, J., *A pathology of bottom-up hill-climbing in inductive rule learning*. Proceedings of the 13th European Conference on Algorithmic Learning Theory (ALT-02), 2002.
88. Cohen, W., Hurst, M. and Jensen, L.S., *A Flexible Learning System for Wrapping Tables and Lists in HTML Documents (HTML)*. World Wide Web-2002, 2002.
89. Zhao, W., Chellappa, R. and Phillips, P., *Subspace linear discriminant analysis for face recognition*. Tech. rep. CAR-TR-914, 1999.
90. Balakrishnama, S., Ganapathiraju, A. and Picone, J. *Linear discriminant analysis for signal processing problems*. in *IEEE Proceedings of Southeastcon '99*. 1999.
91. Aladjem, M., *Linear discriminant analysis for two classes via removal of classification structure*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1997. 19(2): p. 187-192.
92. Zhu, J. and Hastie, T., *Classification of gene microarrays by penalized logistic regression*. Biostatistics, 2004. 5(3): p. 427-43.
93. Hosmer, D.W. and Lemeshow, S., *Applied Logistic Regression*. 2000.
94. Lee, Y.L. and Lee, C.K., *Classification of multiple cancer types by multiclass support vector machines using gene expression data*. Bioinformatics, 2003. 19(9): p. 1132-39.
95. Cristianini, N. and Shawe-Taylor, J., *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. 1st Edition ed. 2000: Cambridge University Press; 1st edition.
96. Schölkopf, S., Burges, J.C. and Smola, A.J., *Advances in Kernel Methods: Support Vector Learning*. 1999.
97. Martinez, F., Tapias, D. and Alvarez, J. *Towards speech rate independence in large vocabulary continuous speech recognition*. in *Acoustics, Speech, and Signal Processing, 1998. ICASSP '98. Proceedings of the 1998 IEEE International Conference on*. 1998.
98. Russek, E., Kronmal, R.A. and Fisher, L.D., *The effect of assuming independence in applying Bayes' theorem to risk estimation and classification in diagnosis*. Comput Biomed Res, 1983. 16(6): p. 537-52.
99. Romero, P., Obradovic, Z., Kissinger, C.R., Villafranca, J.E., Garner, E., Guillot, S., and Dunker, A.K., *Thousands of proteins likely to have long disordered regions*. Pac Symp Biocomput, 1998: p. 437-48.
100. Jones, M.J. and Rehg, J.M., *Statistical Color Models with Application to Skin Detection*. Proc. IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, 1999: p. 274-280.
101. Bishop, C.M., *Neural Networks for Pattern Recognition*. 1995, Oxford, UK: Oxford University Press.
102. Fisher, R.A., *The Use of Multiple measurements in Taxonomic Problems*. Ann. Eugenics, 1936. 7(2): p. 179-88.
103. Senior, A.W. *Face and Feature Finding for a Face Recognition System*. in *Audio- and Video-based Biometric Person authentication '99*. 1999. Washington D. C. USA.
104. Farrar, C.R., Nix, D.A., Duffey, T.A., Cornwell, R.J. and Pardo, G.C. *Damage Identification with Linear Discriminant Operators*. in *17th International Modal Analysis Conference*. 1998. Kissimmee, FL.

105. Goodlin, E.B., Boning, D.S. and Sawin, H.H. *Simultaneous Fault Detection and Classification for Semi-Conductor Manufacturing Tools*. in *201st meeting of the electrochemical society, international symposium on plasma processing xiv*. 2002. Philadelphia, PA.
106. Farrar, C.R., Nix, D.A., Duffey, T.A., Cornwell, P.J. and Pardoen, G.C. *Damage Identification with Linear Discriminant Operators*. in *17th International Modal Analysis Conference*. 1998. Kissimmee, FL.
107. Hilden, J., *Statistical diagnosis based on conditional independence does not require it*. *Comput Biol Med*, 1984. **14**(4): p. 429-35.
108. Langley, P. and Thompson, K. *An analysis of Bayesian classifiers*. in *Tenth National Conference on Artificial Intelligence*. 1992. San Jose, CA: AAAI Press.
109. Domingos, P. and Pazzani, M., *On the optimality of the simple Bayesian classifier under zero-one loss*. *Machine Learning*, 1997. **29**: p. 103-130.
110. Mitchell, T.M., *Machine Learning*. 1997: McGraw-Hill.
111. Hellerstein, J., Thathachar, J. and Rish, I. *Recognizing end-user transactions in performance management*. in *AAAI-2000*. 2000. Austin, Texas.
112. Friedman, N., Geiger, D. and Goldszmidt, M., *Bayesian network classifiers*. *Machine Learning*, 1997. **29**: p. 131-163.
113. Rosenblatt, F., *The perceptron: a probabilistic model for information storage and organization in the brain*. *Psychol Rev*, 1958. **65**(6): p. 386-408.
114. Gurney, J.W. and Swensen, S.J., *Solitary pulmonary nodules: determining the likelihood of malignancy with neural network analysis*. *Radiology*, 1995. **196**(3): p. 823-9.
115. Boone, J.M., *Sidetracked at the crossroads*. *Radiology*, 1994. **193**(1): p. 28-30.
116. Astion, M.L. and Wilding, P., *Application of neural networks to the interpretation of laboratory data in cancer diagnosis*. *Clin Chem*, 1992. **38**(1): p. 34-8.
117. Davis, G.E., Lowell, W.E. and Davis, G.L., *A neural network that predicts psychiatric length of stay*. *MD Comput*, 1993. **10**(2): p. 87-92.
118. Farmer, R.M., Medearis, A.L., Hirata, G.I. and Platt, L.D., *The use of a neural network for the ultrasonographic estimation of fetal weight in the macrosomic fetus*. *Am J Obstet Gynecol*, 1992. **166**(5): p. 1467-72.
119. Dawson, A.E., Austin, R.E. and Weinberg, D.S., *Nuclear grading of breast carcinoma by image analysis. Classification by multivariate and neural network analysis*. *Am J Clin Pathol*, 1991. **95**(4 Suppl 1): p. S29-37.
120. Lawrence, J., *Neural Nets Improve Treatment Quality*. *Medical Electronics*, 1991.
121. Bremner, F.J. and Gotts, S., *Comparing two methods of visual abstraction to human data*. *Behavior Research Methods, Instruments, & Computers*, 1993. **25**(2): p. 212-14.
122. Krusinska, E., Babic, A., Chowdhury, S., Wigertz, O., Bodemar, G. and Mathiesen, U., *Integrated approach for designing medical decision support systems with knowledge extracted from clinical databases by statistical methods*. *Proc Annu Symp Comput Appl Med Care*, 1991: p. 353-7.
123. Gurney, J.W., *Neural networks at the crossroads: caution ahead*. *Radiology*, 1994. **193**(1): p. 27-8.
124. Furlong, J.W., Dupuy, M.E. and Heinsimer, J.A., *Neural network analysis of serial cardiac enzyme data. A clinical application of artificial machine intelligence*. *Am J Clin Pathol*, 1991. **96**(1): p. 134-41.

125. Carpenter, G.A., Grossberg, S. and Reynolds, J.H., *ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network*. Neural Networks, 1991. 4: p. 565-88.
126. Hepner, G.F., Logan, T., Ritter, N. and Bryant, N., *Artificial neural network classification using a minimal training set: comparison to conventional supervised classification*. Photogrammetric Engineering and Remote Sensing, 1990. 56: p. 469-73.
127. Basham, A.M., Pinder III, J.E. and Kroh, G.C., *A comparison of Landsat Thematic Mapper and SPOT multi-spectral imagery for the classification of shrub and meadow vegetation in northern California, USA*. International Journal of Remote Sensing, 1997. 18(18): p. 3719-28.
128. Bendiktsson, J.A. and Sveinsson, J.R., *Multisource data classification and feature extraction with neural networks*. International Journal of Remote Sensing, 1997. 18(4): p. 727-40.
129. Anthony, M. and Bartlett, P.L., *Neural Network Learning: Theoretical Foundations*. 1999, Cambridge, New York: Cambridge University Press.
130. Becker, S., *Unsupervised learning procedures for neural networks*. International Journal of Neural Systems, 1991. 2: p. 17-33.
131. Becker, S. and Plumbley, M., *Unsupervised Neural Network learning procedures for feature extraction and classification*. Journal of Applied Intelligence, 1996. 6: p. 1-21.
132. Hammer, B. and Villmann, T., *Generalized Relevance Learning Vector Quantization*. Neural Networks, 2002. 15: p. 1059-68.
133. Gersho, A. and Gray, R.M., *Vector Quantization and Signal Compression*. Boston: Kluwer Academic Publishers, 1992.
134. Schwarzer, G., Vach, W. and Schumacher, M., *On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology*. Stat Med, 2000. 19(4): p. 541-61.
135. Schwarzer, G., Nagata, T., Mattern, D., Schmelzeisen, R. and Schumacher, M., *Comparison of fuzzy inference, logistic regression, and classification trees (CART). Prediction of cervical lymph node metastasis in carcinoma of the tongue*. Methods Inf Med, 2003. 42(5): p. 572-7.
136. Hertz, K., Krough, A. and Palmer, R., *Introduction to the Theory of Neural Computation*. Lecture Notes, vol.1, Santa Fe Institute Studies in The Sciences of Complexity, 1991.
137. Fateman, R., Tokuyasu, T., Berman, B. and Mitchell, N., *Optical Character Recognition and Parsing of Typeset Mathematics*. Journal of Visual Communication and Image Representation, 1996. 7(1): p. 2-15.
138. Brown, E.W., *Applying Neural Networks to Character Recognition*. 1992.
139. McClelland, J.L., Rumelhar, D.E. and PDP research group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1 : Foundation*. A Bradford Book. 1986, Cambridge, MA: MIT Press.
140. Yao, X. and Liu, Y., *A new evolutionary system for evolving artificial neural networks*. Neural Networks, IEEE Transactions on, 1997. 8(3): p. 694-713.
141. Akay, M., *Time-Frequency and Wavelet Methods with Engineering and Applications*, <http://web.umar.edu/~annie/annie96/fnl/node17.html>: Rutgers University, Piscataway, NJ.

142. Ffrench, P.A., Zeidler, Z.A. and Ku, W.H., *Enhanced Detectability of Small Objects in Correlated Clutter Using an Improved 2-D Adaptive Lattice Algorithm*. IEEE Proceedings of On Image Processing, 1996. 5.
143. Morf, M., Lee, D., Nickolls, J. and Vieira, A. *A classification of algorithms for ARMA models and ladder realizations*. in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '77*. 1977.
144. Stetson, P.F. and Jensen, J.A. *Real-time blood flow estimation using a recursive least-squares lattice filter*. in *IEEE Proceedings of Ultrasonics Symposium*. 1997.
145. Oppenheim, A.V. and Schaffer, R.W., *Digital Signal Processing*. 1975, Englewood Cliffs, New Jersey: Prentice-Hall.
146. Welch, P.D., *The Use of c for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms*. IEEE Transaction on Audio Electroacoust, 1967. 15: p. 70-3.
147. Xie, M., *Signal Decomposition for Nonstationary Processes*. 1995, Virginia Tech.: Blacksburg, VA.
148. Parmar, N. and Kolios, M.C. *Modeling Ultrasound Echo Signals from Regularly and Irregularly Spaced Scatterers*. in *International Conference for Upcoming Engineers (ICUE)*. 2004. Ryerson University, Toronto.
149. Beaufays, F., *Orthogonalizing adaptive algorithms: RLS, DFT/LMS, and DCT/LMS*. 1995, Department of Electrical Engineering, Stanford University: Prentice Hall.
150. Kolios, M.C., Czarnota, G.J., Lee, M., Hunt, J.W. and Sherar, M.D., *Ultrasonic spectral parameter characterization of apoptosis*. Ultrasound Med Biol, 2002. 28(5): p. 589-97.

6

BL-104 - 1