DIAGNOSIS OF LUMBAR DISC HERNIA FROM IMAGES USING ARTIFICIAL NEURAL NETWORK

by

Semra KUL

A thesis submitted to

the Graduate Institute of Sciences and Engineering

of

Fatih University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Computer Engineering

October 2008
Istanbul, Turkey

APPROVAL PAGE

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Cevdet Meriç
Deputy Head of Department

This is to certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Bekir KARLIK
Supervisor

Examining Committee Members

Prof. Dr. Bekir KARLIK :

Prof. Dr. Halil Rıdvan ÖZ :

Assist. Prof. Dr. Zeynep ORHAN :

It is approved that this thesis has been written in compliance with the formatting rules laid down by the Graduate Institute of Sciences and Engineering.

Assist. Prof. Nurullah Arslan
Deputy Director

DIAGNOSIS OF LUMBAR DISC HERNIA FROM IMAGES USING ARTIFICIAL NEURAL NETWORK

Semra KUL

M. S. Thesis - Computer Engineering
October 2008

Supervisor: Prof. Dr. Bekir KARLIK

ABSTRACT

In this study, Magnetic Resonance (MR) images and Artificial Neural Network (ANN) approach are used to diagnose and classify Lumbar Disc Hernia (LDH). The classification is realized using supervised classification model. MR images are widely used in hernia classification in medicine. Sagittal (vertical) MR images are selected and used in the study. Sample MR images related with lumbar disc hernia were collected from Radiology department of Dr. Lütfi Kırdar Hospital.

In the first part of the study, preprocessing and feature extraction methods are implemented. Histogram equalization is applied to the original MR images. Previously Wavelet Transformation is used for feature extraction process. Long coefficient input vector provided by wavelet transformation directed the study to use another feature extraction technique. Average Absolute Deviation (ADD) is used in feature extraction process to reduce the number of inputs to the model. As a result, original image divided into small squared pieces (10x10 pixels) and ADD is applied to each for getting feature vector.

In the last phase of the study, Multi-Layered Perceptron Neural Network is used in the classification and minimum error rate is observed.

Keywords: Artificial Neural Network (ANN), Image Processing, Lumbar Disc Hernia, Backpropagation Algorithm.

YAPAY SİNİR AĞLARI YÖNTEMİ İLE GÖRÜNTÜ ÜZERİNDEN BEL FITIĞI TEŞHİSİ

Semra KUL

Yüksek Lisans Tezi – Bilgisayar Mühendisliği
Ekim 2008

Tez Yöneticisi: Prof. Dr. Bekir KARLIK

ÖZ

Bu tez çalışmasında, Manyetik Rezonans (MR) görüntüleri kullanılarak, Yapay Sinir Ağları yöntemi ile bel fıtığı teşhis ve sınıflandırması gerçekleştirilmiştir. Teşhiste denetimli öğrenme modeli kullanılmıştır. MR görüntüleri fıtık teşhisinde yaygın olarak kullanılan bir yöntemdir. Çalışmada sagital (dikey) MR görüntüleri kullanılmıştır. Örnek görüntüler Dr. Lütfi Kırdar Eğitim ve Araştırma Hastanesi'nden alınmıştır.

İlk aşamada, ön işlem ve öz nitelik çıkarımı gerçekleştirilmiştir. Görüntülere histogram eşitleme uygulanmıştır. Öznitelik çıkarımında öncelikle dalgacık dönüşümü kullanılmıştır. Dönüşüm sonucu elde edilen giriş vektörünün uzunluğu sebebiyle ortalama mutlak sapma yöntemi diğer bir öznitelik çıkarım yöntemi olarak uygulanmıştır. Örnek görüntüler karesel alt görüntülere (10x10) bölünmüş ve her alt görüntüye ortalama mutlak sapma uygulanarak öznitelik vektörü çıkartılmıştır.

Son aşamada, sınıflandırmada çok katmanlı yapay sinir ağı kullanılmış ve en düşük hata oranı gözlenmiştir.

Anahtar Kelimeler: Yapay Sinir Ağları, Görüntü İşleme, Bel Fıtığı, Geri Yayılım Algoritması.

To Mom

ACKNOWLEDGEMENT

I express sincere appreciation to Prof. Dr. Bekir KARLIK for his guidance and insight throughout the research.

I would like to thank Assoc. Prof. Dr. Mustafa ÖZATEŞ for his help and support in providing materials and information used in the study.

I would also thank my colleague and friend Elif ERDOĞAN for her encouragement.

I am also grateful to my husband Selami KUL for his support and endless patience during my study.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

LIST OF SYSMBOLS AND ABBREVIATIONS

SYMBOL/ABBREVIATION

ADD	Average Absolute Deviation
ADC	Apparent Diffusion Coefficient
ANN	Artificial Neural Network
CAT-CT	Computerized Tomography
CWT	Continuous Wavelet Transformation
DWT	Discrete Wavelet Transformation
ECG	Electrocardiography
EEG	Electroencephalography
EIT	Electrical Impedance Tomography
EMG	Electromyography
LBBB	Left Bundle Branch Block
LDH	Lumbar Disc Hernia
LRA	Logistic Regression Analysis
LVQ	Learning Vector Quantization
MC	Mammographic Microcalcifications
MEG	Magnetoencephalography
MLP	Multi Layer Perceptron
MNN	Modular Neural Network
MR-MRI	Magnetic Resonance Imaging
MRCP	Magnetic Resonance Cholangiopancreatography
MSE	Mean Squared Error
MS	Multiple Sclerosis
PSA	Prostate Specific Antigen
RBBB	Right Bundle Branch Block
RBF	Radial Basis Function
ROI	Regions of Interest
ROC	Receiver Operating Characteristic
SOM	Self Organization Map
SVM	Support Vector Machine

CHAPTER 1

INTRODUCTION

Artificial Neural Network (ANN) is a model in the field of Artificial Intelligent that can give various solutions to different problems. ANNs are computational paradigms based on mathematical models. Unlike traditional computing, ANNs have a structure and operation which resembles that of the mammal brain. They also called connectionist systems, parallel distributed systems or adaptive systems, because of having series of interconnected processing elements, operating in parallel. They don't have centralized control in the classical point of view, since all the interconnected processing elements change or adapt simultaneously with the flow of information and adaptive rules (Sordo, 2002).

Artificial neural network has the ability to derive meaning from complicated or imprecise data to extract patterns and detect trends that are too complicated to be noticed by humans or other computer techniques (Salim, 2004). ANN can draw interest of various fields of applications in medicine, mathematics, computer science, chemistry, economics and other fields.

Since correct and fast diagnosis takes important part in treatment process, it is the main problem in medical area. Diagnosis by human has always limitations and human expertise might be the most critical of them. One of the most important problems of medical diagnosis is the subjectivity of the specialist. This is due to the fact that, the result does not depend on a systematic and standard solution but on the interpretation of the patient's signal.

Brause (2001) emphasized some difficulties of diagnosis like insufficient number of samples, especially for the new diseases where all physicians are in the same level of experience. Artificial Intelligence techniques like ANN, can be an alternative to solve problems in diagnosis.

The development of various imaging and signaling techniques in medicine also conveys to the increasing usage of ANNs in this field. ANN can process medical images and signals which are easily measured without bothering patients. ANN can learn from the samples of any case or disease and apply it to the new cases. In that point of view, diagnosis can be named as a classification problem or pattern recognition application.

In this study, diagnosis of lumbar disc hernia from MR images is presented. Diagnosis is implemented by classification of MR images into predefined classes or categories. The ANN model can learn from the samples of MR in training phase and it is tested with different data in testing phase.

The study mainly consists of 4 steps:

- Data acquisition
- Data preprocessing
- Feature extraction
- Implementation

MR images related with lumbar disc hernia were collected from Radiology department of Dr. Lütfi Kırdar Hospital. MR images are gray level images which are generally preferred in medical diagnosis of lumbar disc hernia. Samples were collected from 20 different patients. Each patient's data consists of 10 different, randomly selected MR images. Sagittal (vertical) MR images are selected and used in the study. Initially, dimension of a sample is 291x292 pixels and it is decreased to 200x80 pixels before processing. That provides to eliminate irrelevant parts of the samples. Histogram equalization is applied to increase the contrast of the images.

For feature extraction, two different approaches are applied. First one, wavelet transformation, cannot decrease the number of input vector as desired. The second method is calculating Average Absolute Deviation (ADD). In that method, each image is divided into smaller, 10x10 pixel subdivisions and ADD of each subdivision is calculated and added to the input vector.

In the implementation phase, backpropagation, a supervised learning algorithm is used. Multi Layer Perceptron (MLP) model is used in the training of the network. In the first part of the study, preprocessing and feature extraction methods are implemented. Histogram equalization is applied to the original MR images. Previously, wavelet transformation is used for feature extraction process. Long coefficient input vector provided by wavelet transformation directed the study to use another feature extraction technique. Average Absolute Deviation (ADD) is used in feature extraction process to reduce the number of inputs to the model. As a result, original image divided into small squared pieces (10x10) and ADD is applied to each for getting feature vector.

Two types of classifications are presented in the study. In the first classification, samples are divided into 2 classes: normal and herniated samples. In the second classification, samples are categorized into 4 classes: normal, L5-S1, L5-L4 and L5-L3 level of herniated cases. After training the network with 8 samples from each patient, remaining 2 samples are used in the testing phase. Successful results are obtained and presented in the conclusion.

1.1 ORGANIZATION OF THESIS

In this thesis, medical diagnosis by using ANN classification methodology is presented. In the first part of the study, review of similar applications is given. In the subsequent parts, a new ANN classification is explained in detail.

The goal of the study is not only to discuss results but also state a new, alternative and practical solution to a real life case using real data.

Chapter 1 is an introduction to the case. Literature survey is also given in that part.

The main purpose of Chapter 2 is to introduce how ANNs perform tasks. In this chapter, ANN methodology is explained in detail related with the subject of the thesis. General structure of ANN, perceptron model and layers are explained. Learning algorithms are mentioned. Because backpropagation algorithm is used in the implementation part, it is explained in detail. ANN models are explained and MLP model is described in later part of the chapter. Finally, advantages and disadvantages of ANN models are given in the last part.

In Chapter 3, lumbar disc hernia, the subject of the thesis is explained in detail. Medical definition, symptoms and diagnosis are mentioned. MR imaging method is explained to clarify the chosen method for the implementation of the problem in diagnosis. Types of lumbar disc hernia are described in the classification part. The purpose of the chapter is to provide medical information to the subject of the thesis.

In Chapter 4, the implementation is explained. Data acquisition, preprocessing, feature extraction methods are explained. As feature extraction methods, wavelet transformation and average absolute deviation methods are explained. In the result of trials, average absolute deviation is selected as the feature extraction method. MLP model is used in modeling ANN network. Results are shown. Graphic of error rate versus iteration number is given.

Last chapter is the conclusion. The conclusion explains the results of the study.

1.2 LITERATURE SURVEY

In the study of Özkan et al. (1990), ANN is used for segmentation and classification multi-spectral MR images of normal and pathological human brain. Results show that sharp and compact segmentation of MR images can be obtained with ANN with small architecture.

Rajapakse and Acharya (1990) implemented a self organizing network multilayer adaptive resonance architecture for the segmentation of CT images of heart. Similarly, Däschlein et al. (1994) implemented a two layer ANN for segmentation of CT images of the abdomen. Application discriminates various tissues like kidney, liver, bone and pathologic tissue like renal calculus and kidney tumor.

Miller et al. (1992) trained different neural networks to recognize regions of interest (ROIs) corresponding to specific organs within electrical impedance tomography images (EIT) of the thorax. The model allows automatic selection of optimal pixels based on the number of image, over a sample period, in which each pixel is classified as belonging to a particular organ. Results using simulated EIT data indicate the possible use of neural networks in that kind of images.

Hall et al. (1992) compared ANN and fuzzy clustering methods for segmentation of MR of the brain. Both methods were applied to intelligent diagnosis. As a conclusion, it is proved to be convenient for automatic image segmentation in the context of intelligent diagnosis.

Choong (1994) applied an Entropy Maximization Network (EMN) for the prediction of metastases in breast cancer patients. The clinical and physiological features used in the analysis. The result indicates that EMN is an effective way of constructing discrete models from small data set.

Houston (1994) compared an expert system rule induction and neural network to determine the optimal diagnostic strategy for colorectal cancer using MR and tumor markers. Results showed that both methods rely on large number of samples.

Karakas (1994) used ANN for automatic screening of blood cell classification from microscope images. The network produced a binary output, indicating whether the input corresponded to a normal or a pathologic cell and network correctly classified 65 out of 82 objects.

The studies of Xing (1994) and Zheng (1994) are examples of ANN applied to pattern recognition in mammograms. Xing et al. used 14 image features extracted from mammograms by experienced radiologists. The network detects malignant tumors or clustered classifications in preprocessed mammograms. Zheng et al. used a multistage neural network for locating and classification of microcalcifications in digital mammograms. The network is trained using backpropagation algorithm and Kalman filtering.

Jervis (1994) trained a MLP network with Contingent Negative Variation to differentiate between Huntington's disease, Parkinson's disease and schizophrenia. 17 Contingent Negative Variation features were used in training. Results were promising with sensitivities greater than 0.9 being considered as clinically useful.

Hamamoto (1995) used a MLP trained with preoperative data of 54 patients with early prognosis of hepatocellular carcinoma. It turned out to be a reliable decision support tool for prognosis and assessment of hepatocellular carcinoma.

Burke et al. (1995) compared the prediction accuracy of ANN and other statistical models for breast cancer survival. The neural network model was MLP trained with backpropagation algorithm. Compared with TNM staging system (tumor size, number of nodes with metastatic disease and distant metastases), ANN was more accurate in predicting 5 year survival of 25 cases used in the study.

Sordo (2001) implemented a knowledge based neural network for classification of phosphorus, magnetic resonance spectra from normal and cancerous breast tissues. Data from 26 cases was used as input to the network. A priori knowledge of metabolic features of normal and cancerous breast tissues was incorporated into the structure of the neural network to overcome the insufficiency of available data. Classification rates of 87.36% for knowledge based neural network.

Heckerling et al. (2003) applied feed forward backpropagation network when they predicted the present or absent of pneumonia among adults with acute respiratory illness. They were able to differentiate patients with or without pneumonia with high level of accuracy of 94% to 95%.

Güler and Übeyli (2003) studied on a feature extraction method for medical diagnosis, using wavelet transformation. A MLP network was trained with internal carotid arterial Doppler signals and 96% accuracy of classification was obtained.

Michael et al. (2003) compared ANN with traditional logistical regression models for prediction of the occurrence of delayed graft function in cadaveric renal transplants. Results show that logistics regression was 36.5% sensitive and 90.7% specific compared to the ANN which was 63.5% sensitive and 64.8% specific. Matis (1995) used a MLP network on 290 liver transplantation patents to predict graft outcomes. It accurately predicted 88% of graft failures and 98% of graft survival cases.

Santos-Garcia et al. (2004) applied MLP network for the purpose of predicting cardio-respiratory morbidity after lung resection. They concluded that neural network models were able to provide more sensitive and accurate results than logistic regression.

Karlık, Şahin, Ercan, Tavli (2006) presented a paper to detect Left and Right Bundle Branch Block (LBBB and RBBB) by using ANNs and to transmit classification results of ECG signals from the patient's home to the research hospital via internet for diagnosis, because LBBB and RBBB represents an independent predictor of poor outcome in myocardial infarction.

Medical applications using ANN methodology can be classified into 3 types (Güven, 2005): Electro physical signals like EEG, EKG can be measured and after signal processing, are applied as input to ANN. In the second classification, there are tomography, ultrasonography or MR images belonging to the area of disease. After image processing, feature extraction is applied and input vector is obtained and used in ANN model. Finally, indications of disease or tests like blood pressure, blood sugar or cholesterin can be used as input.

Various diagnosis applications with ANNs using medical imaging have been developed. Wu et al. (1993) used three layered, feed forward neural networks with back propagation algorithm, trained for the interpretation of mammograms on the basis of features extracted from mammograms by experienced radiologists. With clinical cases, the performance of the neural network in merging 14 radiologist extracted features of lesions to distinguish between benign and malignant lesions was found to be higher than the average performance of attending and resident radiologists alone. The authors concluded that such networks may provide a potentially useful tool in the mammographic decision making task of distinguishing between benign and malignant lesions.

Dhawan et al. (1996) studied on analysis of mammographic microcalcifications (MC) using gray level image structure features. In digitized mammograms MCs has been widely recognized as an early sign of breast cancer in women. Jiang Y et al. (1999), Mudigonda et al. (2001), Yu et al. (2006), Pereira et al. (2007) and Karahaliou et al. (2007) studied with mammographs for computerized diagnosis of breast cancer.

Magnetic Resonance Imaging has been widely used in medical diagnosis in artificial neural network applications. Tombropoulos et al. (1993) designed a decision support tool called decision aid for diagnosing liver lesions to aid radiologists in the diagnosis of hepatic lesions seen on MR.

Somorjai et al. (1995) introduced a new classification strategy called computerized consensus diagnosis and applied to proton magnetic resonance spectra of human thyroid biopsies to classify into normal and carcinoma types. Linear discriminant analysis, a neural net based method, and genetic programming were used as the classifiers.

Kari et al. (1995) studied on evaluation of dermatomyositis, a muscle disease, using artificial neural network analysis of MR spectroscopy data.

Goldberg-Zimring et al. (1998) studied on automated detection and characterization of multiple sclerosis (MS) lesions in brain MR images. Artificial neural with back propagation was used for removal of artifacts by differentiating them from true MS lesions. The sensitivity was 0.87 and the specificity 0.96, in 34 images, 100% of the lesions was detected.

Bakken et al. (1999) applied neural network analyses to in vivo 1H magnetic resonance spectroscopy of epilepsy patients. Correct classification of spectra was obtained in 66 out of 67 cases, disregarding from which side of the brain the spectra were recorded.

Deng et al. (1999) used ANN in the measurement study of brain of Alzheimer's disease using MR and a completely new pattern discriminating method is adopted. Using ANN to the same regions and data, both the sensitivity and accuracy were found higher than using the traditional discrimination function analysis method; the indices of amygdala, hippocampus, parahippocampal gyrus, temporal lobe, and temporal horn.

Chiu et al. (2001) studied on tissue segmentation assisted analysis of Functional MRI for human motor response: an approach combining artificial neural network and fuzzy C means.

Rivière et al. (2002) described a complete system allowing automatic recognition of the main sulci of the human cortex. This system relies on a preprocessing of magnetic resonance images leading to abstract structural representations of the cortical folding patterns. Neural network in the study supplied recognition rate of 86% for the learning base and 76% for a generalization base.

Research of Degenhard et al. (2002) for screening in women at genetic risk of breast cancer aims to assist in detecting and diagnosing malignant breast lesions. Three layered, feed forward, back propagation neural network as an artificial radiological classifier using receiver operating characteristic (ROC) curve analysis shows that the neural network approach to clinical diagnosis has considerable potential and warrants further development.

Szabó et al. (2004) developed an artificial neural network based segmentation method for dynamic contrast enhanced MR imaging of the breast and compared with quantitative and empiric parameter mapping techniques.

Poulakis et al. (2004) worked to develop and test an artificial neural network for predicting biochemical recurrence based on the combined use of pelvic coil magnetic resonance imaging, prostate-specific antigen (PSA) measurement, and biopsy Gleason score, after radical prostatectomy and to investigate whether it is more accurate than logistic regression analysis (LRA) in men with clinically localized prostate cancer.

García-Gómez et al. (2004) presented a pattern-recognition approach to the soft tissue tumors benign/malignant character diagnosis using MR. ANN, support vector machine and k-nearest neighbor were used as classifier methods. 88-92% efficiency was obtained in a not-viewed set of tumors using the pattern-recognition techniques. The best results were obtained with a back propagation artificial neural network.

Twellmann et al. (2005) developed an adaptive tissue characterization network for model free visualization of dynamic contrast-enhanced magnetic resonance image data used in cancer diagnosis.

Popple et al. (2006) developed a software package called Brains2 in the University of Iowa for MR images of brain segments. It utilizes MR images, the Talairach atlas, and ANN to segment brain images into substructures in a standardized manner which is critical for Radiotherapy of brain cancer, which inevitably results in irradiation of uninvolved brain.

Logeswaran (2006) proposed a detection scheme for identifying stones in the biliary tract of the body using magnetic resonance cholangiopancreatography (MRCP), a sequence of magnetic resonance imaging targeted at the pancreatobiliary region of the abdomen.

The purpose of the study of Kato et al. (2007) was to preliminarily evaluate the usefulness of a computer algorithm analysis using the finite difference method and an artificial neural network to diagnose hepatic fibrosis with MR images. MR image texture analysis performed using the computer algorithm was found to have a potential usefulness for the diagnosis of hepatic fibrosis.

Dos Santos et al. (2007) used multilayer perceptron and Kohonen SOM classifiers as an alternative to the ADC maps for the evaluation of Alzheimer's disease with MR images. The classification results are used to improve the usual analysis of the apparent diffusion coefficient map.

Döhler et al. (2008) developed a cellular neural network based method for classification of magnetic resonance images for detection of hippocampal sclerosis in the medial temporal lobe. The study demonstrated that the network allows classifying brain tissue with respect to the presence or absence of mesial temporal sclerosis.

Powell et al. (2008) presented a comparison between templates, probability, artificial neural network and support vector machine (SVM) based automated methods for segmentation of subcortical and cerebellar brain structures of interest and showed that these methods may be as reliable as manual raters.

CHAPTER 2

ARTIFICIAL NEURAL NETWORKS

ANN is modeled as the result of inspiration from biological neural systems and it has much simpler structure comparing them. Many developed ANN systems imitate some well known characteristics of biological neural networks like learning capability. Some other features are developed using engineering approach instead of neural physiological approaches (Karlık, 2007).

ANNs are modeled using many disciplines such as neurosciences, mathematics, statistics, physics, computer science and engineering. ANN applications find places in different applications like signal processing, pattern recognition, forecasting, time series analysis and control by virtue of an important property which has the ability to learn (Haykin, 1999).

First studies related with ANN fulfilled by McCulloch and Pitts. The article they published can be considered as the first step in ANN related researches (McCulloch and Pitts, 1943). Afterwards, Hebb (1949), Rosenblatt (1958), Widrow and Hoff (1960), Hopfield (1982), Kohenen (1982), Rumelhart et al., (1986) established researches on various models and algorithms related with ANN.

2.1 GENERAL STRUCTURE OF ANN

ANN primarily consists of artificial neurons connected in a hierarchical manner and working together. ANN model can be revealed as a computerized model consists of perceptrons, the structural distribution of neural network, learning rule and strategy. The main goal of ANN is to produce outputs according to the inputs supplied to the network. To generate output, the network is trained with a training set of sample inputs and desired set of outputs related with the case or problem (learning phase), and network gains the ability of generating successful outputs for the current problem. After learning process, actual test inputs are given to the network and get the network to produce correct outputs (testing phase).

There are various algorithms and models used in ANNs. In this chapter, some algorithms and models are explained which are used in related parts of the thesis.

2.1.1 Perceptron

Artificial neuron or perceptron is the main processing unit of ANN. It is much simpler than a biological neuron. A basic perceptron model is shown in Figure 2.1.

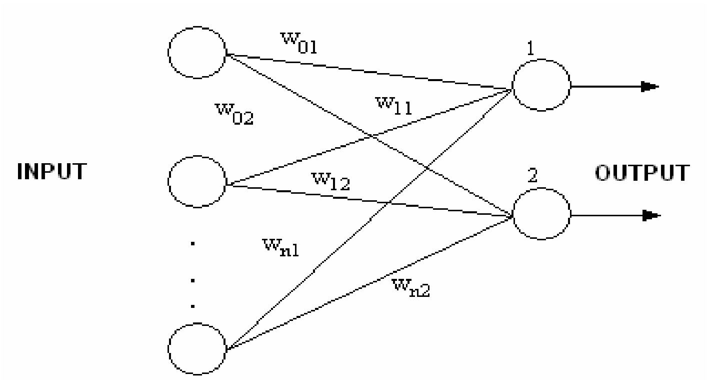


Figure 2.1 Perceptron

Basically, a perceptron consists of input signals, weights, activation function and output signals. Inputs are whether signals from external resources or outputs of other perceptrons. The weights of input signals might differ according to their strengths. Summation function is applied to inputs multiplied by weights, to find the net input. Activation function is applied to the net input to find the net output. Activation function is generally a nonlinear function.

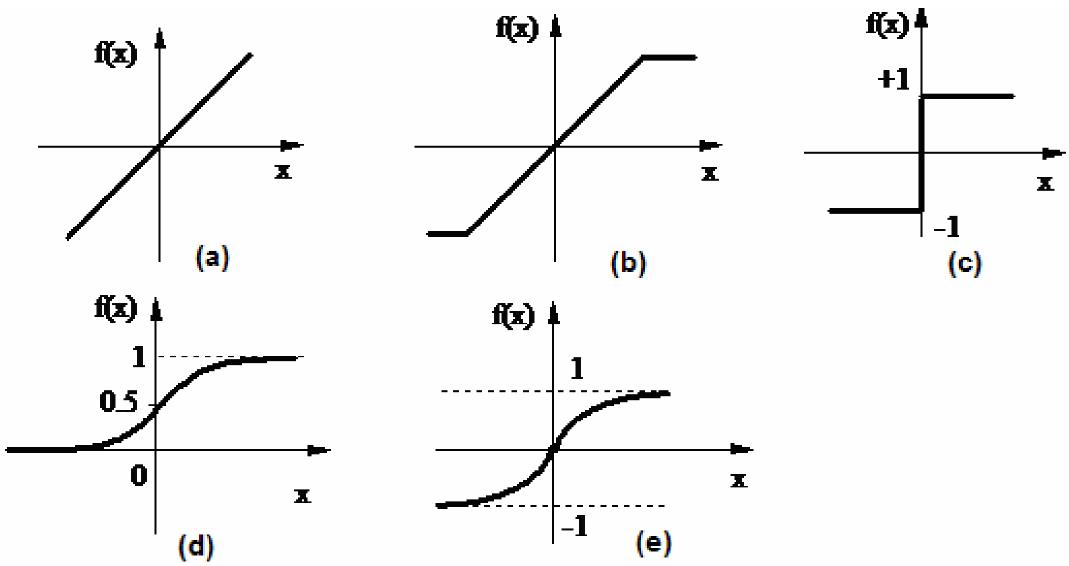


Figure 2.2 Basic activation functions (Karlık, 2007)

a-Linear, **b**-Ramp **c**-Step, **d**-Sigmoid, **e**- Tanh(x)

The choice of activation function differs according to the application (Haykin 1999, Bishop 1995). Sigmoid (2.1) and tangent hyperbolic (2.2) functions are commonly used in applications.

$$f(x)=\frac{1}{1+e^{-x}} \tag{2.1}$$

$$\tanh(x)=\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}=\frac{e^{2x}-1}{e^{2x}+1} \tag{2.2}$$

2.1.2 Layers

ANN model generally consists of 3 layers and constructs parallel network in each layer. These layers are input layer, hidden layer(s) and output layer.

Input layer is fed by signals from the external environment and transfers this information to the hidden layer. Hidden layer(s) is the main operation unit and calculates weights for each input. The final layer is the output layer. It is fed by hidden layer and it is responsible for displaying the result of neural network to the external environment.

Number of nodes in the input layer is decided by feature extraction method which will be explained in latter part of the thesis. Number of nodes in the output layer is decided by the desired number of output class. In this study, the diagnosis is considered as a classification problem and so the number of output nodes is equal to the desired number of classes.

Number of nodes in hidden layer(s) is generally decided by experience. At the beginning, it is decided randomly and gradually changed without effecting expected performance. Fewness of the number of nodes in hidden layer decreases the training time of the network. However, that might be insufficient for the network to learn the sample inputs.

Deciding number of hidden layers and number of nodes in hidden layer are important for the performance of an ANN model. In many cases, two or three layers can produce correct outcomes (Haykin, 1999). Various connections of layers result in different kind of network structures.

2.2 LEARNING ALGORITHMS

Learning in ANN can be described as calculating weight of inputs according to the desired outcomes. Learning rule is a series of mathematical operations to upgrade weights. ANN learns by sample of cases, instead of programming. According to the learning methodology, ANNs can be categorized into 2 major classes: Supervised and unsupervised learning.

In supervised learning, sample inputs and desired outputs related with the problem are supplied to the system. Main goal is to find out the weights which minimize the error, the difference between desired and realized output. Backpropagation algorithm (Widrow and Hoff 1960) is an example of supervised learning algorithm.

In unsupervised learning, desired outputs are not supplied to the system but inputs are given. ANN uses only local information, therefore it is also referred as self-organization. Unsupervised learning is mainly used for clustering problems. Kohonen (1982) is an example for unsupervised learning algorithm.

There are many learning algorithms developed. However, present algorithms are generally developed on principles of four learning algorithms: Hebb (1949), Backpropagation, Kohonen (1982) and Hopfield (1982) (Haykin 1999, Bishop 1995).

2.2.1 Hebb Learning Algorithm

Hebb learning algorithm expresses the basic idea of association learning as when two neurons are physically close and are repeatedly activated together, some chemical changes must occur in their structures, which signify the fact that two neurons fired together (Foundalis and Martinez 2007). In other words, when a node is active it tries to activate the other node connected.

2.2.2 Kohonen Learning Algorithm

That algorithm was developed by Kohonen (1982) and it is example of unsupervised learning algorithm. In that algorithm, Kohonen network is presented with data, but the correct output corresponding to the data is not specified. Each node in the network competes with the others to change weights. The node producing the higher output is the winner and it can change its weight.

2.2.3 Hopfield Learning Algorithm

Hopfield network consists of a set of interconnected neurons which update their activation values asynchronously and independently of other neurons. All neurons are both inputs and outputs in the network. In this algorithm, it is decided how to change the strength of interconnections between nodes. If both expected input and output are active/passive, then weights are increased or decreased by learning coefficient (Koçer 2007).

2.2.4 Backpropagation Learning Algorithm

Backpropagation algorithm is a common supervised learning algorithm. It is more convenient to understand compared with other algorithms. It is multi layered (MLP), feed forward, iterative algorithm which tries to minimize squared error rate between realized and target outputs (Rumelhart 1986). The error is decreased by propagating from a layer back to prior layers (from output to input).

Backpropagation algorithm can be summarized as follows:

- Present a training sample to the neural network.
- Compare the network's output with the desired output from that sample. Calculate the error in each output neuron.
- For each neuron, calculate what the output should have been, and a scaling factor, how much lower or higher the output must be adjusted to match the desired output. This is the local error.
- Adjust the weights of each neuron to lower the local error.
- Assign fault for the local error to neurons at the previous level, giving greater responsibility to neurons connected by stronger weights.
- Repeat the steps above on the neurons at the previous level, using each one's fault as its error.

In that study, backpropagation learning algorithm and MLP ANN model is implemented and backpropagation learning algorithm will be explained in detail in MLP section.

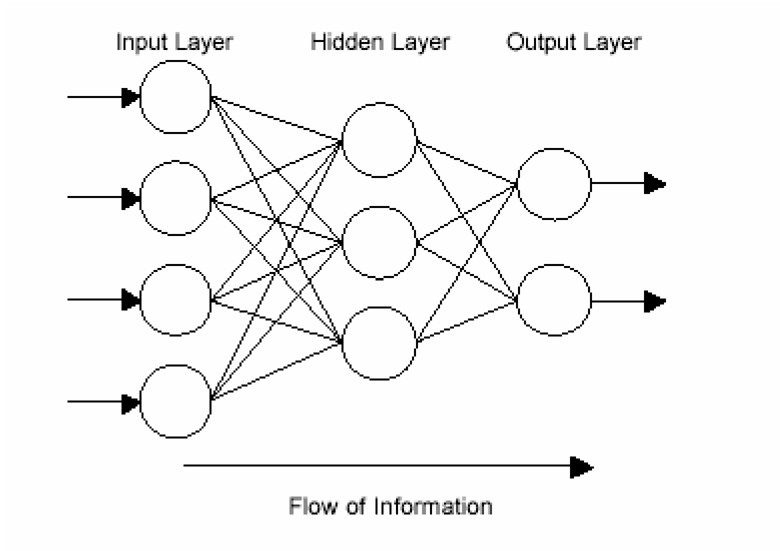


Figure 2.3 MLP Structure

2.3 ANN MODELS

ANN can be modeled as feed forward and feedback networks. In feed forward model, outputs of a layer are given as inputs to next layer (Özbay and Karlık 2001). Although, nodes are connected to other layers, they are not connected to nodes in the same layer. Input layer transmits data supplied from external environment to nodes in the hidden layer without any alteration. Input data is processed in hidden and output layers and output of ANN is generated. Multi Layer Perceptron (MLP), Modular Neural Network (MNN), Learning Vector Quantization (LVQ), Radial Basis Function (RBF) and Probabilistic Neural Network can be given as example for feed forward model (Lippmann 1987, Kohonen 1990, Hopfield 1982).

In feedback models, outputs of output or hidden layers are transmitted to input layer nodes or to prior hidden layers. That enables transmission of inputs in forward and backward direction. Feedback models have dynamic memory which make them especially convenient for prediction applications. Hopfield, Elman and Self Organization Map (SOM) networks are example of that kind of model.

2.3.1 Multi Layer Perceptron

MLP networks normally have three layers of processing elements with only one hidden layer, but there is no restriction on the number of hidden layers. The only task of the input layer is to receive the external signal and to propagate it to the next layer. The hidden layer receives the weighted sum of incoming signals sent by the input units and processes it by means of an activation function. The activation functions most commonly used are the sigmoid and hyperbolic tangent functions. The hidden units send an output signal towards the neurons in the next layer. This adjacent layer could be either another hidden layer of arranged processing elements or the output layer. The units in the output layer receive the weighted sum of incoming signals and process it using the activation function. Information is propagated forward until the network produces an output. Backpropagation learning algorithm is used extensively in MLP applications.

Algorithm starts supplying input data to first layer with small randomly selected weights. From node i to node j net input is calculated as below. *Wij* is the weight of node i to j.

$$net_j = \sum_{i=1}^{m} w_{ij}.x_i \tag{2.3}$$

Output of node j in hidden layer is calculated as,

$$y_j = f_j(net_j) \ , \quad j=1,2,...,J \tag{2.4}$$

Total input supplied to node k in output layer and nonlinear output of node k is,

$$net_k = \sum_{j=1}^{J} w_{kj}.y_j \tag{2.5}$$

$$o_k = f_k(net_k) \quad , \quad k = 1, 2, \ldots, n \tag{2.6}$$

Error rate, difference between desired and obtained output, in the output layer is calculated and mean squared error is observed.

$$e_k = (d_k - o_k) \tag{2.7}$$

$$E = \frac{1}{n}\sum_k (d_k - o_k)^2 \tag{2.8}$$

Weights are updated. **ε** is called learning coefficient and selected between 0 and 0.9.

$$\Delta w_{kj} = -\varepsilon \frac{\partial E}{\partial w_{kj}} \tag{2.9}$$

If right hand side of (2.9) is expanded this equation is obtained (Rumelhart et al., 1986)

$$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial net_k}\frac{\partial net_k}{\partial w_{kj}} \tag{2.10}$$

$$\frac{\partial net_k}{\partial w_{kj}} = \frac{\partial}{\partial w_{kj}}\sum_j w_{kj} \cdot y_j = \sum_j \frac{\partial w_{kj} \cdot y_j}{\partial w_{kj}} = y_j \tag{2.11}$$

Goal is to make $E = \sum_p E_p$ as small as possible by selecting appropriate weights. So, p is randomly selected and error at node k is calculated as,

$$\delta_o = -\frac{\partial E}{\partial net_k} \tag{2.12}$$

If (2.11) and (2.12) are replaced in 2.1

$$\frac{\partial E}{\partial w_{kj}} = -\delta_o \cdot y_j \tag{2.13}$$

$\partial E / \partial y_j$ can't be expanded directly. Effects to output layer is calculated as,

$$-\frac{\partial E}{\partial y_j} = -\sum_k \frac{\partial E}{\partial net_k}\frac{\partial net_k}{\partial y_j} = \sum_k \left(-\frac{\partial E}{\partial net_k}\right)\frac{\partial}{\partial y_j}\sum_m w_{km} y_j = \sum_k \left(-\frac{\partial E}{\partial net_k}\right) w_{kj} = \sum_k \delta_o w_{kj}$$

If (2.13) is replaced in (2.9),

$$\Delta w_{kj} = \varepsilon \delta_o y_j \tag{2.14}$$

Reducing E_p means changing weight according to $\delta_o y_j$. That is called Delta Rule. If partial derivative of (2.12) is obtained,

$$\delta_o = -\frac{\partial E}{\partial net_k} = -\frac{\partial E}{\partial o_k}\frac{\partial o_k}{\partial net_k} \tag{2.15}$$

$$\frac{\partial E}{\partial o_k} = -(d_k - o_k) \tag{2.16}$$

That gives local error of output from node k.

$$\frac{\partial o_k}{\partial net_k} = f_k^{'}(net_k) \tag{2.17}$$

If last 2 equations are replaced in (2.15) and generated equation is replaced in (2.14), (2.19) is obtained.

$$\delta_o = (d_k - o_k) f_k^{'}(net_k) \tag{2.18}$$

$$\Delta w_{kj} = \varepsilon (d_k - o_k) f_k^{'}(net_k) y_j \tag{2.19}$$

If there is a hidden layer, delta rule can also be applied to hidden layer as in (2.20) and (2.21).

$$\begin{aligned}\Delta w_{ji} &= -\varepsilon \frac{\partial E}{\partial w_{ji}} \\ &= -\varepsilon \frac{\partial E}{\partial net_j}\frac{\partial net_j}{\partial w_{ji}} = -\varepsilon \frac{\partial E}{\partial net_j} x_i \\ &= -\varepsilon \left(-\frac{\partial E}{\partial y_j}\frac{\partial y_j}{\partial net_j}\right) x_i = \varepsilon \left(\frac{\partial E}{\partial y_j}\right) f_j^{'}(net_j)\, x_i\end{aligned} \tag{2.20}$$

$$\Delta w_{ji} = \varepsilon \delta_y x_i \tag{2.21}$$

$\partial E / \partial y_j$ can't change directly. Effects on the output layer is calculated as,

$$-\frac{\partial E}{\partial y_j} = -\sum_k \frac{\partial E}{\partial net_k}\frac{\partial net_k}{\partial y_j} = \sum_k \left(-\frac{\partial E}{\partial net_k}\right)\frac{\partial}{\partial y_j}\sum_m w_{km} y_j = \sum_k \left(-\frac{\partial E}{\partial net_k}\right) w_{kj} = \sum_k \delta_o w_{kj}$$

$$\delta_y = f_j^{'}(net_j)\sum_k \delta_o w_{kj} \tag{2.22}$$

If j is in the output layer, equation (2.21) turns out to be next formula, similar to equation (2.18),

$$\delta_y = (d_y - o_y) f_j^{'}(net_j) \tag{2.23}$$

If node j belongs to the hidden layer then (2.13) is applied. Sigmoid activation function is used.

$$f(net_j) = y_j = \frac{1}{1 + e^{-net_j}} \tag{2.24}$$

If derivative of that equation is obtained (2.25) and (2.26) is generated.

$$f^{'}(net_j) = \frac{1}{1 + e^{-net_j}} \frac{1 + e^{-net_j} - 1}{1 + e^{-net_j}} \tag{2.25}$$

$$\frac{\partial y_j}{\partial net_j} = y_j (1 - y_j) \tag{2.26}$$

Similarly for layer k,

$$\frac{\partial o_k}{\partial net_k} = f_k^{'}(net_k) = o_k (1 - o_k) \tag{2.27}$$

If (2.27) is replaced in (2.18) and (2.26) is replaced in (2.22) these equations turns out to be

$$\delta_o = (d_k - o_k) o_k (1 - o_k) \tag{2.28}$$

$$\delta_y = y_j (1 - y_j) \sum_k \delta_o w_{kj} \tag{2.29}$$

Δw_{ij} must be calculated for every training set (Karlık, 2007).

2.4 ADVANTAGES OF ANN

ANNs are generally preferred in various applications working any kind of data. They can be applied to complicated problems in the field of business, finance, industry or science, which are difficult to solve by using present methods. Major features of ANNs are having parallelized distributes structure, learning ability and to apply generalization (Moody and Darken, 1989).

ANN has ability to learn. ANN can learn from similar cases and try to find a solution for a new case similar to sample ones. Because of its nonlinearity, it can handle complex problems. It develops different solution to the problem (Lippmann, 1987). In contrast with other methods, it has no database and information is stored in network connections. Information about the value of connections between nodes constitutes information about the network. ANN can work with large number of variables or parameters.

Even for dissimilar problems with sample training set, ANN can find solution by using generalization. They can handle with noisy and missing data because of having integrated memory. Noisy data in training set is distributed among weights of connections and this eliminates negative effects. They can provide general solutions with good predictive accuracy. ANN has high tolerance to errors in training and testing sets, resulting better performance.

Because of its parallel structure, ANN can process information quickly and it can be applied to real time applications like pattern recognition, signal processing, authorization and inspection.

2.5 DISADVANTAGES OF ANN

ANNs have some disadvantages make them inappropriate to apply for some problems. First of all, determining the network structure is a difficult operation and it is generally the result of many trial and errors. Network structure is critical to find acceptable solution and to get better performance from the system. Network parameters like learning coefficient, number of nodes in layers are also determined by experience.

Training is generally slow measured in thousandths of a second and result is inexact. There is no certainty about the correctness of the solution.

ANN can only work with numerical data. Therefore, if the input parameters are non numeric information, they must be converted to numeric values. Monitoring of outputs and interpreting result is also determined by the user. There is no standard for displaying outputs.

CHAPTER 3

LUMBAR DISC HERNIA

A herniated disc is an illness occurs generally in the lumbar region (low back) of the body, because the lumbar spine in the lumbar region of the body carries most of the body's weight. It is called lumbar disc hernia or disc rupture. Sometimes herniation can press on a nerve inside the spine and causes radicular pain spreading to other parts of the body (Takada et al., 2001). The amount of pain related with hernia is generally depends on the amount of material spilled into spinal canal and the amount of pressure caused by that material onto nerve. In this chapter, lumbar disc hernia is explained in detail with minimal use of medical terms as much as possible.

3.1 DEFINITION AND SYMPTOMS

Lumbar disc hernia is a common cause of low back pain and radicular leg pain. It is common in almost every society. Spine consists of 33 vertebrae (bones). In between vertebrae in spine, there are intervertebral discs. These discs help buffering body's movements and decrease the amount of pressure on the spine. The disc has a gel-like inner substance called the nucleus pulposus and a tire-like outer band called the annulus fibrosus. If the nucleus can break out the annulus, that results in a herniated disc. Herniated disc can apply pressure on nerves in spinal canal causing pains in some other parts of the body, where these nerves attain or affect. In addition to that, secretion of chemical material from the ruptured disc can also influence nerve roots and cause pain (Spineuniverse, 2008).

Symptoms of a herniated disc may include dull or sharp pain, muscle spasm or cramping, weakness, tingling, or referred pain (Spineuniverse, 2008). Referred pain means, there is pain in another part of body as a result of the disc problem. For example, if a person has herniated disc in his lower back (lumbar spine), he might have pain in his leg.

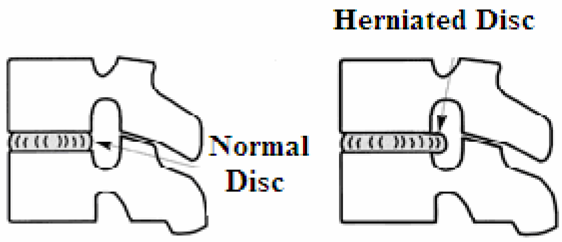


Figure 3.1 Simple vertical view of disc hernia (Yıldızhan, 2008)

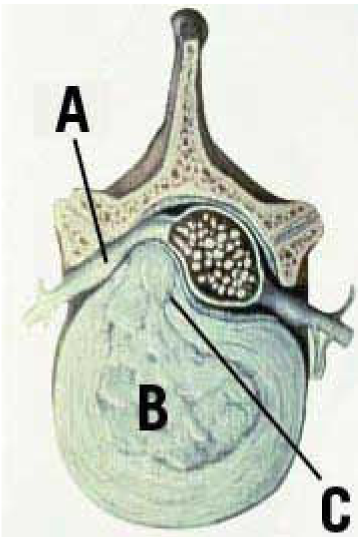


Figure 3.2 Lumbar disc hernia (Spine-health, 2008)

A: Exiting nerve root B: Disc C: Torn outer annulus

People between age of 30 and 50 are most at risk of hernia because the elasticity and water content of the nucleus pulposus decreases with age.

Lumbar disc hernia is generally called with the name of vertebrae in spine. The lower segments are called sacral segments. Above sacral segments, there are lumbar segments (bones or corpus). The naming convention is shown in Figure 3.4. If hernia occurs in intervertebral part between S1 and L5 vertebrae, it is called L5-S1 level hernia. Approximately 90% of disc hernia occurs at L4-L5 (lumbar segments 4 and 5) or L5-S1 (lumbar segment 5 and sacral segment 1), which causes pain in the L5 nerve or S1 nerve, respectively. L5 nerve pressure can cause weakness in extension of the big toe and potentially in the ankle (foot drop). S1 nerve pressure may cause loss of the ankle reflex or weakness in ankle push off, for example patients can’t do toe rises (Spine-health, 2008).

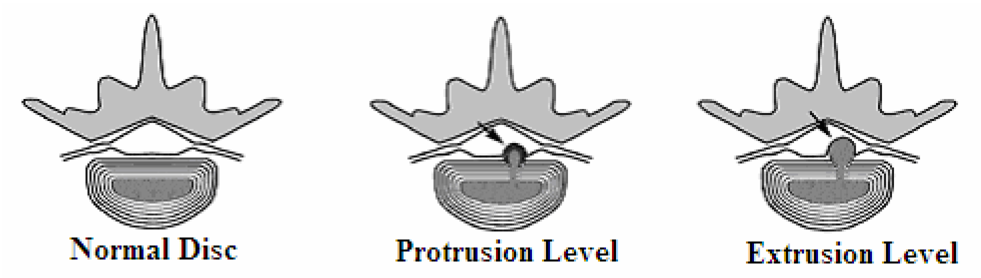


Figure 3.3 Samples of disc hernia (Yıldızhan, 2008)

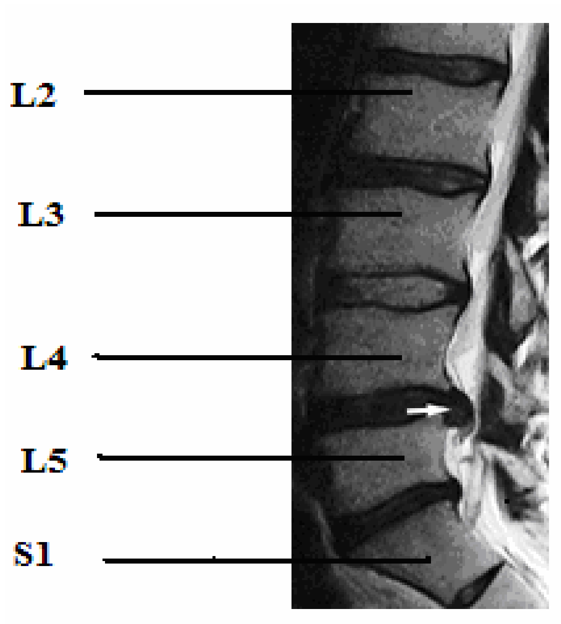


Figure 3.4 Segments of vertebrae shown on sample MR image used in the study

3.2 DIAGNOSIS

Diagnosis starts with physical examination, observing posture and range of motion. Spine specialist also conducts a neurological exam testing reflexes, muscle strength, other nerve changes, and pain spread of the patient. Radicular pain (inflammation of a spinal nerve) may increase when pressure is applied directly to the affected area. An X-ray can show a narrowed disc space, fracture or bone spur. A computerized tomography scan (a CT or CAT scan) or a magnetic resonance imaging test (MR) can show soft tissue of a disc. These tests show the stage and location of the disc hernia. If spine specialist suspects about nerve damage, an electromyography (EMG) can also be applied.

If more accurate diagnosis is needed, some extra tests can be applied. Discogram or discography is procedure in which a dye is injected into vertebral disc and observed. Bone scan technique creates computer images of bones. A very small amount of radioactive material is injected into a blood vessel then throughout the blood stream, collects in the bones and can be detected by a scanner. Lab tests can also be applied if the blood cells are normal or abnormal. Chemical changes in the blood may indicate a metabolic disorder causing back pain (Spineuniverse, 2008).

3.2.1 Magnetic Resonance Imaging

MR or MRI is a medical imaging technique primarily used in radiology to visualize the structure and function of the body (Wikipedia, 2008). It provides detailed images of the body. It is painless and safe diagnostic procedure without the use of X-rays or other radiation. Each MR image is a slice of the body area scanned and many images are created. MR scans are used for the examination of the brain, nerves, spinal cord, musculoskeletal areas (bones and joints), abdominal and pelvic organs.

In addition to its high diagnostic value, MR offers predictive information about disc hernia evolution (Splendiani et al., 2004). Assessment of the reliability of standardized MR interpretations and measurements were studied (Lurie et al., 2008) and observed that the reliability of MR relevant to patients with disc herniation is adequate. The aim of another study was assessment of the usefulness of MR, CT and Radiculography in the diagnosis of lumbar disc herniation. MR appeared to have the highest accuracy (Gościński et al., 2001). Therefore, MR imaging is selected for the study.

Two types of views are supplied by MR Imaging: Sagittal and axial. Sagittal view shows the spine in vertical orientation. In axial view, horizontal segments of disc can be observed. Sagittal and axial views of an MR imaging are shown in Figure 3.5 and Figure 3.6.

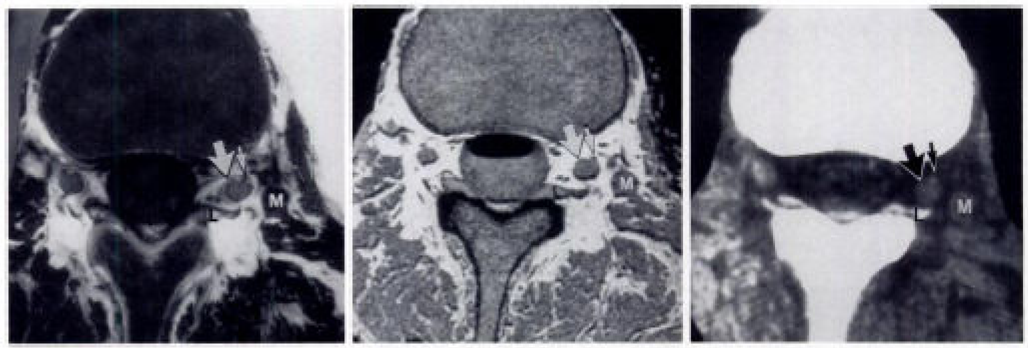


Figure 3.5 Axial view of lumbar disc MR (Kostelic et al., 1992)

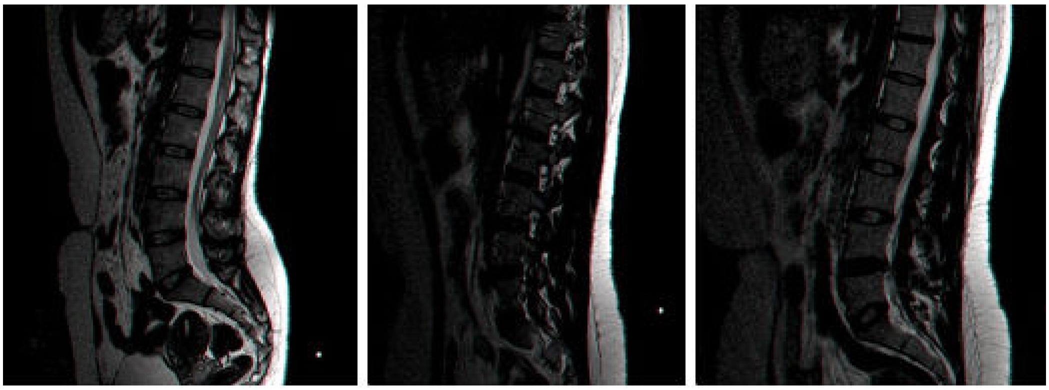


Figure 3.6 Sagittal view of lumbar disc MR (original sagittal MR sample collected from the hospital for the study)

3.2.2 Classification

Three main classifications of lumbar disc hernia are protrusion, extrusion and sequestration (Healthatoz, 2008).

Protrusion is advanced level of disc bulging, which nucleus pulposus and spinal disc remain untouched but form a structure that can press against the nerves. It is the first stage of disc hernia.

Extrusion is the second phase of disc hernia. When the outer part of the annulus fibrosus is ruptured, allowing inner gelatinous part of the disc to spill out, disc extrusion occurs.

Disc sequestration occurs when the center gelatinous part of the disc is not only spilt out, but also separated from the main part of the disc.

In Figure 4.2 normal, protrusion and extrusion type of MR are shown.

CHAPTER 4

DIAGNOSIS OF LUMBAR DISC HERNIA USING ANN

In various fields of application, Artificial Neural Network has been commonly used with rapidly increasing success as in medicine, science or commerce. Especially in medicine, where extensive amount of data and knowledge stored in clinical databases, main contributing factor influencing the development of such systems has been the additional demands for a more powerful, flexible and transparent technique (Shadabi, 2007). Neural networks have been applied to medical applications for clinical classification (Baxt 1995), image analysis and interpretation (Miller 1992, Miller 1993), signal analysis and interpretation and drug development.

Modern medicine applies variety of imaging techniques to the human body. Biomedical signal measurements comprises of electrocardiography (ECG), electromyography, EEG, magnetoencephalography (MEG), computer tomography (CT), magnetic resonance (MR), functional MRI (Acır, 2004). MR is one of them and primarily used in radiology to visualize the structure and function of the body. It provides detailed images of the body in any plane. It has much greater soft tissue contrast than CT, making it especially useful in neurological, musculoskeletal, cardiovascular and oncological imaging.

MR is widely preferred and used in diagnosis of hernia including lumbar disc hernia and it is preferred for this study. In this thesis, ANN method is used to diagnose lumbar disc hernia from MR images. The thesis proposes that ANN can provide an accurate solution for the problem. ANN model used in the implementation is a supervised learning model. First training data with desired outputs are supplied to the network. Then testing data is given and outputs are observed. Because the number of output classes is determined before the process, this application can be handled as classification problem or pattern recognition which ANN is the classifier.

Figure 4.1 General flow of process

In the first part of this chapter, data acquisition procedure is explained.

In the second part, image preprocessing is described. Description of image processing and histogram equalization is given.

Next part mentions about feature extraction procedure. 2 Dimensional discrete wavelet transformation and Average Absolute Deviation methods are explained in detail.

Structure of MLP, number of inputs and outputs, implementation process is explained in the next part of the thesis.

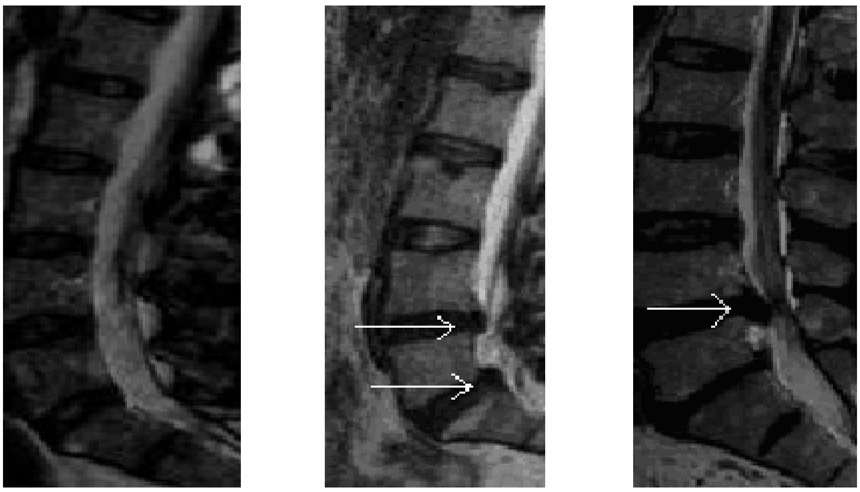
Chapter ends with the conclusion.

4.1 DATA ACQUISITION

In the diagnosis of lumbar disc hernia, two kinds of MR imaging are used: Axial (horizontal) and sagittal (vertical). Sagittal MR images are selected and used in the study for the efficiency and convenience for processing. Sample MR images related with hernia were collected from radiology department of Dr. Lütfi Kırdar Hospital. Images are gray level colored and in JPEG format.

As stated in Chapter 3, lumbar disc hernia is labeled with the segments name of backbones. The last part of backbone is called sacrum. Lumbar segments are above the sacrum and labeled starting from L5 through L1. These 5 segments (corpus) are parts of backbone which hernia is extensively observed. Figure 3.4 shows parts of backbones which MR based classification is concentrates on.

Collected MR data includes 3 types of samples: Normal cases, protrusion and extrusion type of hernia. Figure 4.2 shows samples of the cases. Figure 4.2 (a) belongs to a patient without hernia (normal case), 4.2 (b) and 4.2 (c) show protrusion and extrusion respectively. Protrusion in Figure 4.2 (b) is in the segments of L5-S1 and L4-L5. Extrusion occurs in the segment of L4-L5 in Figure 4.2 (c).



(a) (b) (c)

Figure 4.2 Sagittal MR samples collected from the hospital are marked to emphasize herniated and normal cases:
a- Normal **b**- Protrusion (L5-S1, L4-L5) **c**- Extrusion (L4-L5)

Samples are collected from 20 patients. Each MR collection from one patient consists of approximately 20 different images. 10 samples from each patient are selected randomly. 8 of them are used as input data to train ANN network, 2 of them are used for testing. Training and testing images are selected randomly out of 10 images.

In the application, 2 methods of classification are used according to distribution of samples. In the first classification, samples are divided into 2 classes: Normal and hernia. In the second classification, samples are divided into 4 categories: Normal, hernia in L5-S1, L4-L5 and L3-L4 segments.

In the first classification, extrusion and protrusion cases are classified in the herniated class. Because of the inadequate samples for extrusion, these samples are combined into one class (in herniated class).

Second classification is also decided according to the number of samples in each class. L5-S1, L4-L5, L3-L4 and normal cases are selected. Similar with the first classification, protrusion and extrusion cases are not differentiated and they are combined into the same class.

Table 4.1 ANN Classification I and II

Number of class	**Name of classes**	**Number of samples**
2	Normal, Hernia	10 patients (10 samples for each patient: 8 for training, 2 for testing)
4	Normal, L5-S1, L4-L5 and L3-L4 segments	5 patients (10 samples for each patient: 8 for training, 2 for testing)

4.2 PREPROCESSING

Images from MR collections are gray level colored and JPEG formatted. Gray level images include gray level colors between black and white. Each pixel has a numerical value from 0 to 2^8-1 = 255. In the ANN implementation, numerical values of each MR image are processed and used as input to the ANN model.

Dimensions of each MR image are initially 291x292 pixels. Each image contains unrelated parts for diagnosis. Hernia is usually located in the lower parts of the spine, through L1 to S1 segments. Therefore, top, left and right side of each image are removed from the original image. The number of pixels removed from each side is decided by observing the samples. To decrease the number of inputs given to ANN, images are cut into smaller dimension: 200x80 pixels.

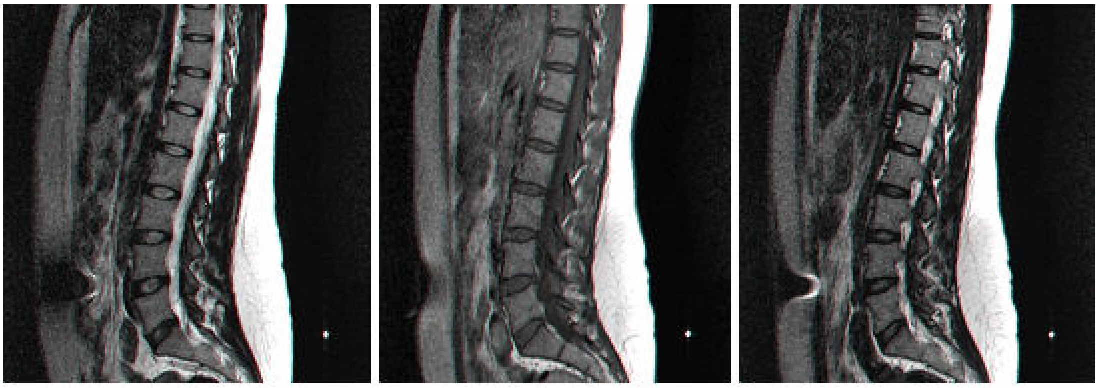


Figure 4.3 Samples of original MR images collected from the hospital

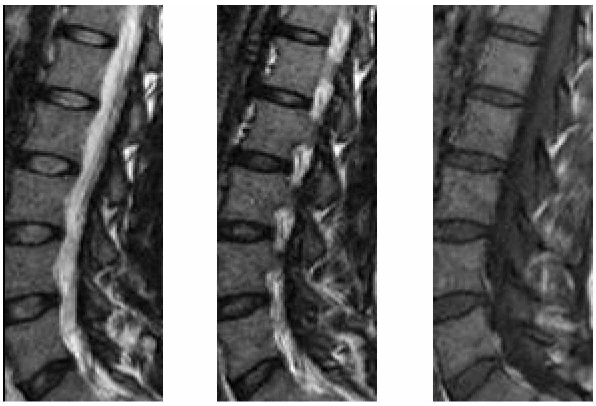


Figure 4.4 MR samples produced after resizing images in Figure 4.3

4.2.1 Histogram Equalization

Histogram equalization is an image processing method for contrast adjustment using image histogram.

Histogram is a graphic displaying the number of pixels in each gray level colors. It shows the frequency or distribution of colors in an image.

This method usually increases local contrast of the image, especially when the image contains close contrast values. Through this adjustment, the intensities of pixels are distributed in the image equally. Arrangement of local contrast doesn't affect the global contrast of image because most frequent pixel values are distributed effectively. The method is useful in images with backgrounds and foregrounds that are both bright and both dark. Histogram equalization for nxm gray level image is done with equalization in (4.1) (Umbaugh 1999). N(q) is the new pixel value after histogram equalization, k is the gray level and c(k) is the total pixel number of old gray level values until level k.

$$N(q) = \mathrm{Max}\{0, \mathrm{Round}[2L \,.\, c(k) / (m.n)] - 1\} \qquad (4.1)$$

In this study, histogram equalization is applied to images before feature extraction by wavelet transformation. That eliminates contrast variances in images and helps to find similar input vectors for the samples of same classes. MATLAB image processing tool is used for histogram equalization. In the second feature extraction method (AAD), histogram equalization is not applied because contrast variances want to be kept for calculating average absolute deviation of the original images.

(a) (b) (c)

Figure 4.5 MR samples after Histogram Equalization is applied to the resized images **a-** Normal data **b-** Protrusion (L5-S1, L4-L5) **c-** Extrusion (L4-L5)

4.3 FEATURE EXTRACTION

In ANN applications, feature extraction is an important step in the process. Feature extraction is a special form of dimensionality reduction. When the input data to the network is too large to be processed and if it is known that there is enormously redundant data, then the input data can be reduced to a small subset of features. That transformation process is called feature extraction. It is used in pattern recognition and image processing applications like image compression. Descriptive specifications of original input data are selected and transformed into numerical data. The generated input vector after feature extraction is given to the ANN.

The size of feature vector is important for the success of classification. Smaller the size of feature vector, faster the speed of execution is. On the other hand, inadequate size can be concluded in erroneous outputs.

In this study, 2 Dimensional wavelet transformation and average absolute deviation are experimented as feature extraction methods and average absolute deviation is selected for the application.

4.3.1 Wavelet Transformation

Wavelet transformation is a tool which extracts different frequency components of data. The wavelet transformation of a signal evolving in time depends on two variables, frequency change scale and time. Wavelet provides a methodology for time-frequency localization (Ruskai et al. 1992, Hees-Nielsen and Wickerhauser 1996). Wavelet transformation is the representation of a function by wavelets. Wavelet transforms have advantages over traditional Fourier transformation for representing functions that have discontinuities and sharp peaks, and for accurately deconstructing and reconstructing finite unstable signals.

Wavelet transformation is classified into Discrete Wavelet Transformation (DWT) and Continuous Wavelet Transformation (CWT). CWT operates over every possible scale and DWT uses a specific subset of scale and translation values.

Discrete wavelet transform is the inner product of a signal with the wavelet basis. That projection of discrete signals into each basic element defines a wavelet coefficient transform summation series.

Discrete wavelet transformation can be considered to pair up input values storing the difference and transmitting the sum, for an input represented by a list of $2n$ numbers. This process is repeated recursively, pairing up the sums to provide the next scale and finally results in $2n - 1$ differences and one final sum.

Images are analyzed with two dimensional discrete wavelet transformation tool in MATLAB. DWT analysis produces approximation coefficients matrix and details coefficients (horizontal, vertical, and diagonal) matrices by decomposition of the input matrix. The original matrix can be constructed from these vectors.

In this study, after histogram equalization is applied to the original images, level 1 and level 2 DWT analyses are applied and approximation coefficients are obtained. In level 1 DWT, the number of approximation coefficient is 4000. These coefficients are considered as the inputs to the ANN models and the size of the input vector should be deducted to an optimum size. Therefore, level 2 DWT is applied to the images and size is reduced to 1000.

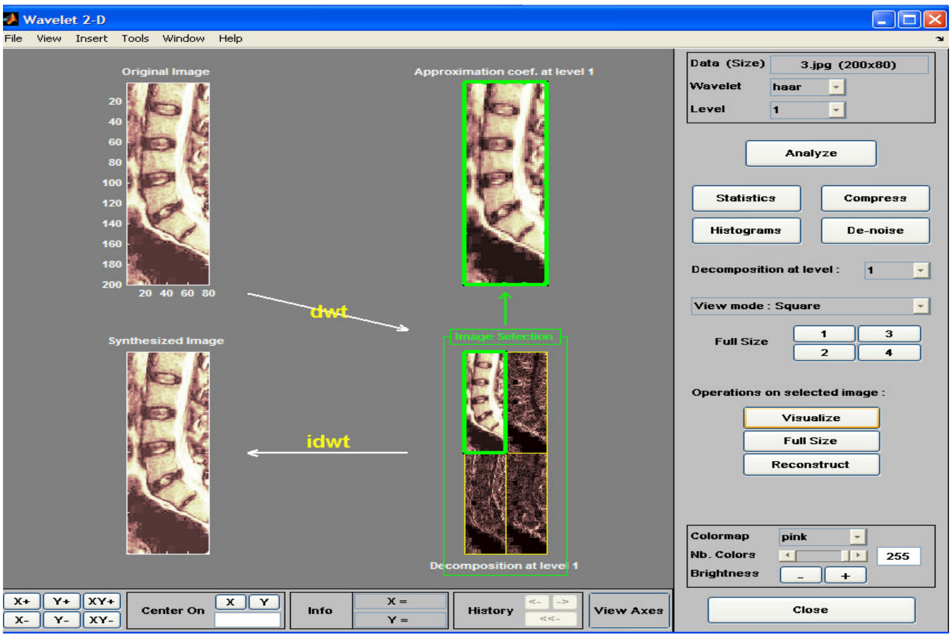


Figure 4.6 Graphical Interface for Level 1 DWT (Matlab)

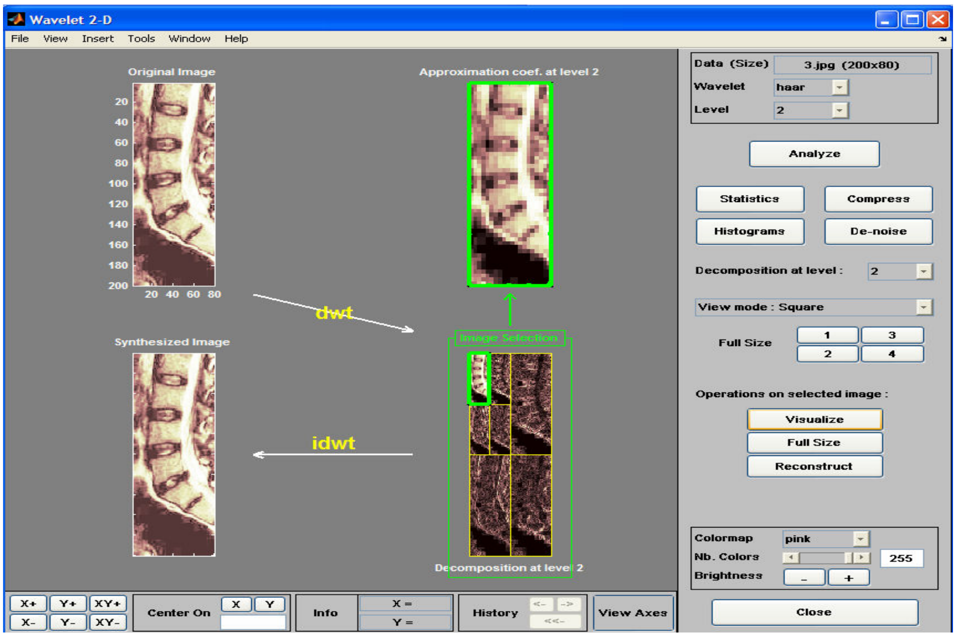


Figure 4.7 Graphical Interface for Level 2 DWT (Matlab)

Table 4.2 2D DWT Approximation Coefficients Number

Level	**Number of approximation coefficients**
1	4000
2	1000

As shown in Table 4.2, level 1 and level 2 DWT methods cannot reduce the size of feature (input) vector to appropriate size. Therefore, wavelet approach is not used in the implementation. Instead, average absolute deviation method is used to find input vector representing the original MR samples.

4.3.2 Average Absolute Deviation

In average absolute deviation method, difference of each pixel value to the mean value of overall image is calculated (MathWorld, 2008).

$$\alpha = \frac{1}{N}\sum_{i=1}^{N} | x_i - \mu | \tag{4.2}$$

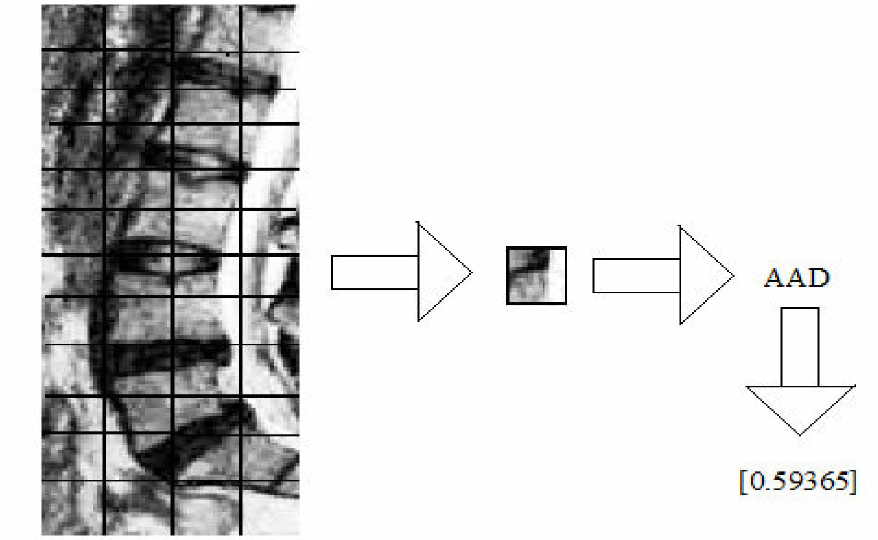
In the equation, N is the number of pixels in the image, μ is the mean of overall image, x_i is the value of pixel *i*.

In that method, original image (200x80) is divided into smaller divisions. Each division is size of 10x10 pixels. Average absolute deviation of each division is calculated. Matlab image tool is used to divide original image to subdivisions. New values are normalized between 0 and 1 by dividing by 100. The results generate the feature vector size of 160 elements.

Average absolute deviation is applied to the original images without using histogram equalization.

Table 4.3 Size of feature (input) vector

Size of original image	**Size of sub division**	**Size of feature vector**
200x80	10x10	160



Input Vector => [0.59365 0.47066 0.10019 0.0795…..]

Figure 4.8 Sub divisions to calculate Average Absolute Deviation

4.4 IMPLEMENTATION

In this study, Multi Level Perceptron (MLP) model is used in the classification of lumbar disc hernia MR images. In the real world, diagnosis is performed using supervised learning method. Therefore, the network uses a supervised learning algorithm. The ANN model is first trained with samples of the cases and after that, it is tested with some testing data. 8 samples from each patient are applied to the network. 2 samples are used in testing.

Feature vector constructed by average absolute deviation method is used in the MLP models. Because of the size of the feature vector resulted by wavelet extraction method, this method is not used in the model. AAD method conveys to a feature vector having size of 160, for each MR image. It means model has 160 nodes in the input layer.

The number of nodes in the output layer is decided by the number of classes in the model. In the first classification, we have 2 classes - 2 nodes and in the second classification, we have 4 classes – 4 nodes in the output layer.

A software program developed by Prof. Dr. Bekir KARLIK is used in the implementation. The program implements an MLP model and uses backpropagation algorithm as the learning algorithm. Sigmoid function is used as the activation function. There are 1 input layer, 1 hidden layer and 1 output layer in the model. There are 160 nodes in the hidden layer and learning rate is set to 0.95. The number of nodes in the hidden layer is generally set to the greater number of nodes whether in the output or input layer. In this case, the number of nodes in input layer is greater than in the output layer and so the number of nodes in hidden layer is set to 160.

Two classifications are modeled with ANN. Each MLP model is shown in Table 4.4.

Table 4.4 MLP models

Output classes	**Feature vectors**	**Nodes in layers**
Normal, Hernia	160	160 x 160 x 2
Normal, L5-S1, L4-L5, L3-L4	160	160 x 160 x 4

Program screenshots are given in Appendix A.

4.5 OUTPUTS OF MLP

Training set is applied to the network and iteration number is increased from 100 to 3000 iterations. Mean square error (MSE) is observed during training phase. When MSE is reached to acceptable rate, approximately 1%, training is stopped. Table 4.5, Figure 4.9 and Figure 4.10 show the relation between iteration number and error rate.

For the first classification, after 2000 iterations, acceptable error rate is achieved. Because we have just 2 output classes (normal and hernia) and more samples for training phase, successful result is reached at smaller number of iteration in the first problem.

In the second classification, we have smaller number of samples for each class to train the network and that results in higher iteration number to reach reasonable MSE.

Examples for training and testing phases are given in the appendix. In this phase, testing is performed with the data from patients whose data is also used in training step.

Table 4.5 Iteration number and error rate

Iteration	**Error rate**	
	Classification 1	***Classification 2***
100	0.109605411	0.191716668
200	0.050749756	0.095943283
300	0.045763977	0.090682968
400	0.040178059	0.081343337
500	0.036696111	0.074432670
600	0.033351036	0.067273623
700	0.031277284	0.059943008
800	0.029445962	0.053956249
900	0.027933098	0.049399514
1000	0.026885937	0.045846808
1500	0.020147041	0.035645537
2000	0.016152788	0.029648704
2500		0.024171582
3000		0.019864827

Table 4.6 Iteration number and execution time

Iteration	Execution time (min:sec)	
	Classification 1	***Classification 2***
100	00:09	00:08
200	00:18	00:16
300	00:27	00:25
400	00:36	00:34
500	00:45	00:42
600	00:54	00:51
700	01:03	00:59
800	01:11	01:08
900	01:20	01:16
1000	01:29	01:25
1500	02:13	02:06
2000	02:57	02:48
2500		03:29
3000		04:11

Table 4.7 Number of outputs (data used in both training and testing)

	Test Inputs	**Correct outputs**	**Incorrect outputs**	**Accuracy**
Classification 1	40	38	2	0.95
Classification 2	40	37	3	0.93

In the second trial, the network is tested with data from patients whose data is not used in the training phase. The correct number of outputs and accuracy for that trial is shown in Table 4.8.

Table 4.8 Number of outputs (data used only testing)

	Test Inputs	Correct outputs	Incorrect outputs	Accuracy
Classification 1	16	8	8	0.50
Classification 2	64	18	46	0.28

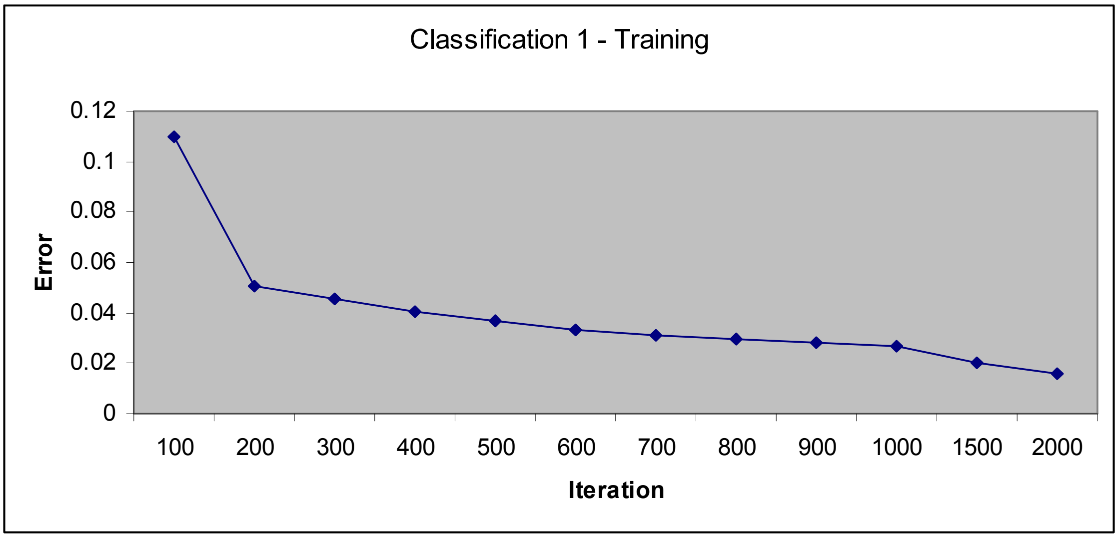


Figure 4.9 Classification I Training Graphic

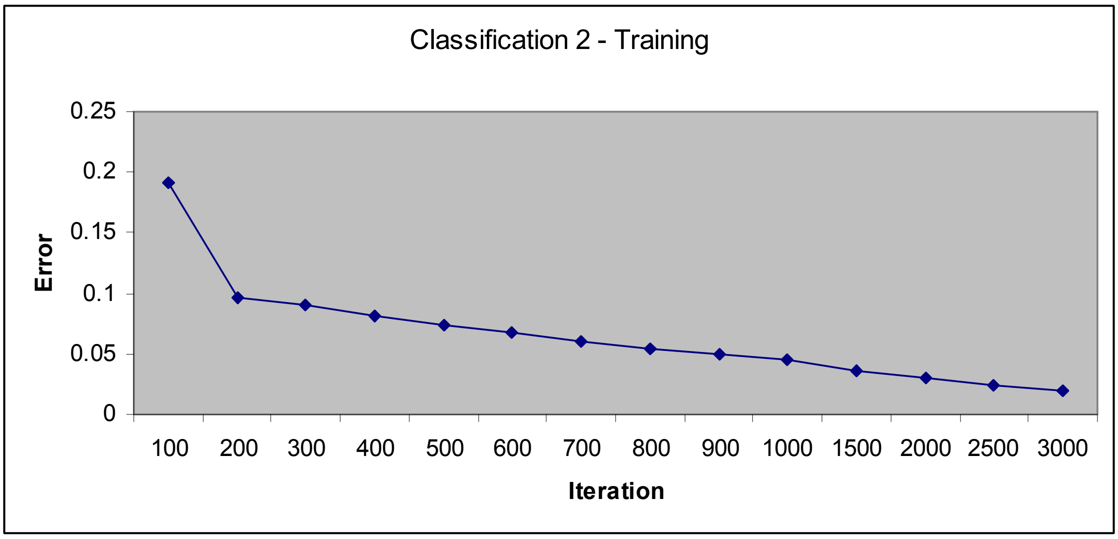


Figure 4.10 Classification II Training Graphic

CHAPTER 5

CONCLUSION

In this study, MR images are used to diagnose lumbar disc hernia with ANN methodology. The study demonstrates an application for classification of medical imaging.

MR images related with lumbar disc hernia were collected from radiology department of Dr. Lütfi Kırdar Hospital. These gray level, JPEG formatted images are 291x292 pixels and they include some unrelated parts for diagnosis. Using Matlab imaging tool, these unrelated parts are removed roughly from original images. The size of new MR images is 200x80 pixels.

Feature extraction is the next step in ANN applications. It is critical for the success of applications. Feature extraction determines the values and the size of input vectors. It also quantizes the real world problem. For example, numeric value “1” can represent for boolean value “true” and numeric value “0” can represent “false” in an ANN model. The main goal of the feature extraction is to find identifiers of the original inputs having suitable size. The input vector constructed by feature extraction method, should represent the characteristics of the original MR image and it should have smaller size (smaller than 200x80 pixels). The input vector is also 1 dimensional contrast to original 2 dimensional MR image.

In the first trial, wavelet transformation is applied as the feature extraction method. Wavelet transformation is generally used for the compression of both 1 dimensional signal and 2 dimensional images. Original data is represented by product of signal with the wavelet basis. Since the new coefficients can represent the original data, that method can also be applied in ANN applications to find input vector.

In the study, before wavelet transformation, histogram equalization is applied to the images. Histogram equalization is used for the contrast adjustment of images. After that, discrete wavelet transformation is applied. Matlab image tool is used for the calculation of approximate coefficients. The size of coefficient vector after level 1, 2 dimensional DWT is 4000 and it is 1000 after level 2, 2 dimensional DWT. Since these input sizes are unsuitable for the study, wavelet transformation method is ignored and another feature extraction method is investigated.

Deviation in images is one of the common identifiers in ANN applications and Average Absolute Deviation method is used to find input vector representing the original image in the study. It brings out the idea that, if we divide MR image into subdivisions and find average absolute deviation of each of these subdivisions, the original image can be represented with the new values. These deviations can differentiate the herniated parts and ANN can learn from these set of data. Matlab imaging tool is used to divide MR image and calculate average absolute deviation of each subdivision. MR images are divided into 10x10 pixels and average absolute deviation of each is calculated. The results are normalized between 0 and 1. These values construct input vector size of 160 input nodes.

In the study, supervised learning approach and multi layer perceptron model are applied and successful results are obtained and presented. Two classifications are tested according to the available type of data. In the first classification, there are normal and herniated classes. In the second classification, there are normal, L5-S1, L4-L5 and L3-L4 type of herniated classes.

Model is trained with 8 samples and tested with remaining 2 samples from each patient. Since there are 20 patients and 10 samples for each patient, there are 160 training and 40 testing data distributed equally for each class. The ANN is trained with various iteration numbers for both classifications and error rate is observed. Total squared error rate versus iteration number is also given for 2 types of classifications, in Figure 4.9 and Figure 4.10. In the first classification, after 2000 iterations, the error rate decreases below 1%. In the second classification after 3000 iterations, the error rate decreases below 1%. In the first classification, there are 10 patients for each class (10x8 = 80 for training, 10x2 = 20 for testing), and in the second there are 5 patients (5x8 = 40 for training, 5x2= 10 for testing) for each class. Iteration number in the second classification is higher because of less number of samples. The accuracy of first classification is 95% and accuracy of second classification is 93% (Table 4.7). The execution time for the first classification is 02:57 (min:sec) for 2000 iterations and 04:11 (min:sec) for the second classification for 3000 iterations (Table 4.6). Average absolute deviation as feature extraction method has given successful results in the study.

The model is also tested with totally different patient data, where no sample is used in training phase, the network gives inaccurate result for this trial. For the first classification, the accuracy is 50% and for the second classification the accuracy is 28%. This is the result of insufficient training samples used in implementation. If there were more than 20 patient, the network would have probably given accurate results for new patient whose data was not used in the training phase.

As a feature study, training samples for all types of lumbar disc hernia can be obtained and the model can be trained for all types.

The number of training and testing samples from each patient are selected randomly (8 for training, 2 for testing). An intelligent method can be applied to decide the optimum number of training and testing samples.

Different feature extraction methods can also be applied and results can be compared with this study. Principal component analysis (PCA) can be applied to find the input vector.

Image processing can be applied to the MR images and only intervertebral parts can be selected for the study. That would decrease the size of input vector even smaller.

ANN model presented in the study provides accurate result to the problem in a short time and proves that it can be applied diagnosis of hernia in medicine. In addition to that, with the learning ability of ANN models, adoptability to various problems, minimum data requirement and minimum processing time period helps modeling problem efficiently.

This study aims to provide alternative way for medical diagnosis and a way to decrease diagnosis time using an intelligent model. In some cases, it might help to physicians to diagnose complex cases which are difficult to perceive. Physicians can combine this opportunity and their expertise to detect early stage of diseases.

REFERENCES

Acır, N., *Automatic Pattern Recognition in EEG by using Artificial Neural Networks*, Ph.D. Thesis, Dokuz Eylül University, 2004.

Bakken, IJ., Axelson, D., Kvistad, KA., Brodtkorb, E., Müller, B., Aasly, J., Gribbestad, IS., "Applications of Neural Network Analyses to in vivo 1H Magnetic Resonance Spectroscopy of Epilepsy Patients", *Epilepsy Res.*, Vol. 35(3), pp. 245-52, 1999.

Baxt, W. G., "Application of Artificial Neural Networks to Clinical Medicine." *Lancet, 3*46. pp. 1135-7, 1995.

Bishop, C.M., *Neural Network for Pattern Recognition*, Oxford University Press, England, 1995.

Brause, R.W., Medical Analysis and Diagnosis by Neural networks. 2001. http://www.informatik.uni-frankfurt.de/asa/papers/ISMDA2001k.pdf

Chiu, MJ., Lin, CC., Chuang, KH., Chen, JH., Huang, KM., "Tissue Segmentation-Assisted Analysis of fMRI for Human Motor Response: An Approach Combining Artificial Neural Network and Fuzzy C Means", *J Digit Imaging,* Vol. 14(1), pp. 38-47, 2001.

Choong, P.L., deSilva, C.J.S., "Breast Cancer Prognosis using EMN Architecture", *Proceedings of IEEE International Conference on Neural Networks*, June 1994.

Cluett, J., *Disc Extrusion, Protrusion and Sequestration,* http://orthopedics.about.com/od/herniateddisc/g/discs.htm

Däschlein, R., Waschulzik, T., Brauer, W., et al., "Segmentation of Computertomographies with Neural Networks Based on Local Features. In Ifeachor, E., & Rosen, K. (Eds.)", *International Conference on Neural Networks and Expert Systems in Medicine and Healthcare,* pp. 240-247, 1994.

Degenhard, A., Tanner, C., Hayes, C., Hawkes, DJ., Leach, MO., "Comparison between Radiological and Artificial Neural Network Diagnosis in Clinical Screening", *Physiol Meas.,* Vol. 23(4), pp. 727-39, 2002.

Deng, X., Li, K., Liu, S., "Preliminary Study on Application of Artificial Neural Network to the Diagnosis of Alzheimer's Disease with Magnetic Resonance Imaging", *Chin Med J (Engl)*, Vol. 112(3), pp. 232-7, 1999.

Dhawan, AP., Chitre, Y., Kaiser-Bonasso, C., "Analysis of Mammographic Microcalcifications using Gray-level Image Structure Features", *IEEE Trans Med Imaging*, Vol. 15(3), pp. 246-59, 1996.

Dos Santos, WP., de Souza, RE., dos Santos Filho, PB., "Evaluation of Alzheimer's Disease by Analysis of MR Images using Multilayer Perceptrons and Kohonen SOM Classifiers as an Alternative to the ADC Maps", *Conf Proc IEEE Eng Med Biol Soc. 2007*, pp. 2118-21, 2007.

Döhler, F., Mormann, F., Weber, B., Elger, CE., Lehnertz, K., "A Cellular Neural Network based Method For Classification of Magnetic Resonance Images: Towards an Automated Detection of Hippocampal Sclerosis", *J Neurosci Methods.*, Vol. 170(2), pp. 324-31, 2008.

Foundalis, H.E. and Martinez, M., "A Generalization of Hebbian Learning in Perceptual and Conceptual Categorization", *Proceedings of the European Cognitive Science Conference*, Delphi-Greece, May 2007, pp. 312-317.

García-Gómez, JM., Vidal, C., Martí-Bonmatí, L., Galant, J., Sans, N., Robles, M., Casacuberta, F., "Benign/Malignant Classifier of Soft Tissue Tumors using MR Imaging", *MAGMA*., Vol. 16(4), pp. 194-201, 2004.

Gillard, D.M., *Disc Herniation*, 2005. http://chirogeek.com/001_Tutorial_Birth_of_HNP.htm

Goldberg-Zimring, D., Achiron, A., Miron, S., Faibel, M., Azhari, H., "Automated Detection and Characterization of Multiple Sclerosis Lesions in Brain MR Images", *Magn Reson Imaging,* Vol. 16(3), pp. 311-8, 1998.

Gościński, I., Ulatowski, S., Urbanik, A., *Comparison of the Clinical Usefulness of Magnetic Resonance (MR), Computer Tomography (CT) and Radiculography (R) in Diagnosing Lumbar Discopathy,* 2001, http://www.ncbi.nlm.nih.gov/pubmed/8130512?ordinalpos=96&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_RVDocSum

Güven, A., *Artificial Neural Network based Diagnosis of Some of the Eye Diseases by Using Ocular Electrophysiological signals*, Ph.D. Thesis, Erciyes University, 2005.

Hall, L., Bensaid, A., Clarke, L., Velthuizen, R., Silbiger, M., & Bezdek, J., "A Comparison of Neural Network and Fuzzy Clustering Techniques in Segmenting Magnetic Resonance Images of the Brain", *IEEE Transactions on Neural Networks*, Vol. 3(5), pp. 672—682, 1992

Hamamoto, I., Okada, S., Hashimoto, T., Wakabayashi, H., Maeba, T., & Maeta, H., "Prediction of the Early Prognosis of the Hepatectomized Patient with Hepatocellular Carcinoma with a Neural Network", *Comput. Biol. Med.*, Vol. 25(1), pp. 49—59, 1995.

Haykin, S., *Neural Networks A Comprehensive Foundation*, Prentice Hall, New Jersey, 1999.

Healtatoz, 2008.
http://www.healthatoz.com

Hebb, D.O., *The Organization of Behaviour: A Neurophsychological Theory*, Wiley, New York, 1949.

Heckerling, P.S., Gerber, B.S., Tape, T.G. and Wigton, R.S, "Prediction of Community-acquired Pneumonia using Artificial Neural Network", *Med. Decis. Making*, Vol. 23, pp. 112-21, 2003

Hees-Nielsen, N. and Wickerhauser, "M.V.,Wavelets and Time-Frequency Analysis". *IEEE Proc.* Vol.84, pp. 523-540, 1996.

Hopfield, J.J., "Neural Network and Physical Systems with Emergent Collective Computational Abilities", *Proceedings of the National Academy of Sciences*, Vol. 79, pp. 2254-2258, 1982.

Houston, A., Kemp, P., Macleod, M., Norswor thy, J., et al., "Analysis of a Nuclear Medicine Database for Colorectal Cancer Using Expert System Rule Induction and Neural Networks. In Ifeachor, E., & Rosen, K. (Eds.)", *International Conference on Neural Networks and Expert Systems in Medicine and Healthcare,* pp. 303-309, 1994.

Jervis, B., Yu, S., Saatchi, M., & Allen, E., "The Sensitivity of Multilayer Perceptrons for Differentiating Between CNV Waveforms of Schizophrenic, Parkinson's Disease and Huntington's Disease Subjects. In Ifeachor, E., & Rosen, K. (Eds.)", *International Conference on Neural Networks and Expert Systems in Medicine and Healthcare*, pp. 275—282, 1994.

Jiang, Y., Nishikawa, RM., Schmidt, RA., Metz, CE., Giger, ML., Doi, K., "Improving Breast Cancer Diagnosis with Computer-Aided Diagnosis", *Acad Radiol*, Vol. 6(1), pp. 22-33, 1999.

Karahaliou, A., Skiadopoulos, S., Boniatis, I., Sakellaropoulos, P., Likaki, E., Panayiotakis, G., Costaridou, L., "Texture Analysis of Tissue Surrounding Microcalcifications on Mammograms for Breast Cancer Diagnosis", *British Journal of Radiology*, Vol. 80(956), pp. 648-56, 2007.

Karakas, M., Duru, H., Tipi, V., et al., "A System for an Automated Classification of Blood Cells from Microscope Images: CELL v. 1.0. In Ifeachor, E., & Rosen, K. (Eds.)", *International Conference on Neural Networks and Expert Systems in Medicine and Healthcare*, pp. 248-256, 1994.

Kari, S., Olsen, NJ., Park, JH., "Evaluation of Muscle Diseases using Artificial Neural Network Analysis of 31P MR Spectroscopy Data", *Magn Reson Med.*, Vol. 34(5), pp. 664-72, 1995.

Karlık, B., Şahin, Y.G., Ercan, T., Tavlı, T., *Bundle Branch Blocks Diagnosis using Neural Networks for Telecardiology*, 2006, http://www.telemed.org.ua/UJTMMT/N1_06/article6.html

Karlık, B., *Artificial Neural Network Lecture Notes,* 2007

Kato, H., Kanematsu, M., Zhang, X., Saio, M., Kondo, H., Goshima, S., Fujita, H., “Computer-Aided Diagnosis of Hepatic Fibrosis: Preliminary Evaluation of MRI Texture Analysis using the Finite Difference Method and an Artificial Neural Network”, *AJR Am J Roentgenol.*, Vol. 189(1), pp. 117-22, 2007.

Koçer, H. E., *Recognition of Iris Texture by Using Artificial Intelligence Methods*, Ph.D. Thesis, Selçuk University, 2007.

Kohonen, T., “Self-organized Formation of Topologically Correct Feature Maps”, *Biological Cybernetics,* Vol. 43, pp. 59-69, 1982.

Kohonen, T., “Improved Versions of Learning Vector Quantization”, *International Joint Conference on Neural Network*, Vol. 1, pp. 545-550, San Diego, 1990.

Kostelic, J., Haughton, V. M., Sether, L., “Proximal Lumbar Spinal Nerves in Axial MR Imaging, CT, and Anatomic Sections”, *Radiology*, Vol. 183(1), pp. 239-241, 1992.

Lippmann, R.P., “An Introduction to Computing with Neural Networks”, *IEEE Acoustics,* Vol. 4, pp. 4-22, 1987.

Logeswaran, R., “Neural Networks Aided Stone Detection in Thick Slab MRCP Images”, *Med Biol Eng Comput.*, Vol. 44(8), pp. 711-9, 2006.

Lurie, JD., Tosteson, AN., Tosteson, TD., Carragee, E., Carrino, J., Kaiser, J., Sequeiros, RT., Lecomte, AR., Grove, MR., Blood, EA., Pearson, LH., Herzog, R., Weinstein, JN., “Reliability of Magnetic Resonance Imaging Readings for Lumbar Disc Herniation in the Spine Patient Outcomes Research Trial (SPORT)”, *Spine*, Vol. 33(9), pp. 991-8, 2008.

Matis, S., Doyle, H., Marino, I., Murad, R. And Uberbacher, E., “Use of Neural Networks for Prediction of Graft Failure following Liver Transplantation”, *Proceedings of the 8th Annual IEEE Symposium on Computer-Based Medical Systems: 0133*, 1995.

Mathworld, 2008. http://mathworld.wolfram.com/AverageAbsoluteDeviation.html

McCullloch, W.S. and Pitts, “A Logical calculus of the ideas immanent in nervous activity”, *Bulletin of Methematical Biophysics,* Vol. 5, pp.115-133, 1943.

Michael, E.B., Prasun, C., Klein, R.B. and Klein, J.B, “Prediction of Delayed Renal Allograft Function using an Artificial Neural Network”, *Nephrol Dial Transplant*, Vol.18, pp. 2655-2659, 2003.

Miller, A., Blott, B., and Hames, T., "Review of Neural Network Applications in Medical Imaging and Signal Processing*", Medical and Biological Engineering and Computing.*" Vol. 5, pp. 449-464, 1992.

Miller, A., *The Application of Neural Networks to Imaging and Signal Processing in Astronomy and Medicine*, Ph.D. Thesis, University of Southampton, 1993.

Moody, J. and Darken, C.J., "Fast Learning in Networks of Locally-tuned Processing Units", *Neural Computing*, Vol. 1, pp. 281-294, 1989.

Mudigonda, NR., Rangayyan, RM., Desautels, JE., "Detection of Breast Masses in Mammograms by Density Slicing and Texture Flow-Field Analysis", *IEEE Trans Med Imaging*, Vol. 20(12), pp. 1215-27, 2001.

Özbay, Y., Karlık B., "A Recognition of ECG Arrhythmia Using Artificial Neural Network", *23rd Annual International Conference IEEE Engineering in Medicine and Biology Society*, Istanbul-Turkey, 2001.

Özkan, M., Sprenkels, H., and Dawant, B., "MultiSpectral Magnetic Resonance Image Segmentation Using Neural Networks", *In International Joint Conference on Neural Networks*, Vol. 1, pp. 429-434, 1990.

Pereira, RR., Jr, Azevedo Marques, PM., Honda, MO., Kinoshita, SK., Engelmann, R., Muramatsu, C., Doi, K., "Usefulness of Texture Analysis for Computerized Classification of Breast Lesions on Mammograms", *J Digit Imaging,* Vol. 20(3), pp. 248-55, 2007.

Popple, RA., Griffith, HR., Sawrie, SM., Fiveash, JB., Brezovich, IA., "Implementation of Talairach Atlas based Automated Brain Segmentation for Radiation Therapy Dosimetry" *Technol Cancer Res Treat.*, Vol. 5(1), pp. 15-21, 2006.

Poulakis, V., Witzsch, U., de Vries, R., Emmerlich, V., Meves, M., Altmannsberger, HM., Becht, E., "Preoperative Neural Network using Combined Magnetic Resonance Imaging Variables, Prostate-Specific Antigen, and Gleason Score for Predicting Prostate Cancer Biochemical Recurrence after Radical Prostatectomy", *Urology*, Vol. 64(6), pp. 1165-70, 2004.

Powell, S., Magnotta, VA., Johnson, H., Jammalamadaka, VK., Pierson, R., Andreasen, NC., "Registration and Machine Learning-based Automated Segmentation of Subcortical and Cerebellar Brain Structures", *Neuroimage*, Vol. 39(1), pp. 238-47, 2008.

Rajapakse, J., and Acharya, R., "Medical Image Segmentation with MARA", *In International Joint Conference on Neural Networks,* Vol. 2, pp. 965-972, 1990.

Rivière, D., Mangin, JF., Papadopoulos-Orfanos, D., Martinez, JM., Frouin, V., Régis, J., "Automatic Recognition of Cortical Sulci of the Human Brain using a Congregation of Neural Networks", *Med Image Anal*., Vol. 6(2), pp. 77-92, Review, 2002.

Rosenblatt, F., "The perceptron: A probabilistic model for information storage and organization in the brain", *Psychological Review*, Vol. 65, pp. 386-408, 1958.

Rumelhart, D.E., Hinton, G.E, and Williams, R.J., “Learning Representations by Backpropagation Errors”, *Nature*, Vol. 323, pp. 533-536, 1986.

Ruskai, M.B,Beylkin, G., Coifman, R., Daubechies, I., Mallat, S.,Meyer, Y. and Raphael, L.E., Wavelets and Their Applications, Jones and Bartlett, New York, 1992.

Salim, N., *Medical Diagnosis Using Neural Networks,* 2004. http://www.generation5.org/content/2004/MedicalDiagnosis.asp?Print=1

Santos-Garcia, G., Varela, G., Novoa, N. and Jimenez, M.F. “Prediction of Postoperative Morbidity after Lung Resection using an artificial neural network ensemble”, *Artificial Intelligence in Medicine 30*, Vol. 1, pp. 61-69, 2004.

Shadabi, F., *Medical Outcome Prediction: A Hybrid Artificial Neural Networks Approach,* Ph.D. Thesis, Canberra University, 2007.

Somorjai, RL., Nikulin, AE., Pizzi, N., Jackson, D., Scarth, G., Dolenko, B., Gordon, H., Russell, P., Lean, CL., Delbridge, L., “Computerized Consensus Diagnosis: A Classification Strategy For the Robust Analysis of MR Spectra. I. Application to 1H Spectra of Thyroid Neoplasms”, *Magn Reson Med,* Vol. 33(2), pp. 257-63, 1995.

Sordo, M., Buxton, H, Watson, D., “A Hybrid Approach to Breast Cancer Diagnosis”, *In Practical Applications of Computational Intelligence Techniques. Jain, L. and DeWilde*, P. (Eds.). Kluwer Academic Publishers, 2001.

Sordo, M., *Introduction to Neural Networks in Healthcare,* 2002. http://www.openclinical.org/docs/int/neuralnetworks011.pdf

Spine-health, 2008 http://www.spine-health.com/conditions/back-pain/lumbar-herniated-disc

Spineuniverse, *Symptoms of a herniated disc*, 2008. http://www.spineuniverse.com/article/herniated-disc-symptoms-4103.html

Splendiani, A., Puglielli, E., Amicis, R.D., Barile, A., Masciocchi, C. and Gallucci, M., “Spontaneous Resolution of Lumbar Disk Herniation: Predictive Signs for Prognostic Evaluation”, *Neuroradiology*, Vol. 46(11), pp. 916-922, 2004.

Szabó, BK., Aspelin, P., Wiberg, MK., “Neural Network Approach to the Segmentation and Classification of Dynamic Magnetic Resonance Images of the Breast: Comparison with Empiric and Quantitative Kinetic Parameters”, *Acad Radiol.*, Vol. 11(12), pp. 1344-54, 2004.

Takada, E. Takahashi, M., Shimada, K., “Natural history of lumbar disc hernia with radicular leg pain: Spontaneous MRI changes of the herniated mass and correlation with clinical outcome”, *Journal of Orthopaedic Surgery*, Vol. 9(1), pp. 1-7, 2001. www.josonline.org/PDF/v9i1p1.pdf

Tombropoulos, R., Shiffman, S., Davidson, C., "A Decision Aid for Diagnosis of Liver Lesions on MRI", *Proc Annu Symp Comput Appl Med Care*, pp. 439-43, 1993.

Twellmann, T., Lichte, O., Nattkemper, TW., "An Adaptive Tissue Characterization Network for Model-Free Visualization of Dynamic Contrast-Enhanced Magnetic Resonance Image Data", *IEEE Trans Med Imaging.*, Vol. 24(10), pp. 1256-66, 2005.

Umbaugh, S.E., Computer Vision and Image Processing A Practical Approach using CVIP tools, Prentice Hall, New Jersey, 1999.

Übeyli, E.D. and Güler, I., "Neural Network Analysis of Internal Carotid Arterial Doppler Signals: predictions of stenosis and occlusion", *Expert Systems with Applications*, Vol. 25(1), pp. 1-13, 2003.

Widrow, B. and Hoff, M.E., *Adaptive Switching Circuits*, IRE WESCON Convention record, pp. 96-104, 1960.

Wikipedia, 2008.
http://wikipedia.org

Wu, Y., Giger, ML., Doi, K., Vyborny, CJ., Schmidt, RA. and Metz, CE, "Artificial Neural Networks In Mammography: Application to Decision Making In the Diagnosis of Breast Cancer", *Radiology*, Vol. 187, pp. 81-87, 1993.

Xing, G., and Feltham, R., "Pyramidal Neural Networking for Mammogram Tumour Pattern Recognition", *In IEEE International Conference on Neural Networks,* Vol. VI, pp. 3546-3551, 1994.

Yıldızhan, A., *Bel Fıtığı ve Korunma Yolları*, 2008.
http://www.belfitigi.com

Yu, S., Li, K., bHuang, Y., "Detection of Microcalcifications in Digital Mammograms using Wavelet Filter and Markov Random Field Model", *Computerized Medical Imaging and Graphics*, Vol. 30(3), pp. 163-173, 2006.

Zheng, B., Qian, W., & Clarke, L. P., "Multistage Neural Network for Pattern Recognition in Mammogram Screening", *In IEEE International Conference on Neural Networks,* Vol. VI, pp. 3537-3441, 1994.

APPENDIX A

SOFTWARE

In this study, MATLAB by Mathworks company and a program coded by Bekir KARLIK are used in the implementation.

MATLAB is a software tool, used in engineering and scientific applications requiring mass amount of numerical calculations, image processing and advanced level programming.

In the study, MATLAB is used for image processing, wavelet transformation and calculating average absolute deviation (feature extraction) of sub dimensions.

MATLAB CODES

Image resizing

```
I=imread('1.jpg')

I=im2uint8(I)
for i=66:265
for j=117:196
Ired(i-65,j-116)=I(i,j);
end
end

imwrite(Ired, '1.jpg')
clear
```

Image histogram equalization

```
I = imread('1.jpg');
J = histeq(I);
imwrite(J, '1.jpg')
```

Average absolute deviation

```
clear
totall = intmax('uint64');
total2 = intmax('uint64');
total2  = 0;
totall = 0;
p=imread('1.jpg')
I = double(p)

for i=1:20
for j=1:8
for k=(i*10)-9:(i*10)
for l=(j*10)-9:(j*10)
totall = I(k,l) + totall;
end
end

totall = totall/100;

for k=(i*10)-9:(i*10)
for l=(j*10)-9:(j*10)
total2 = total2 + abs(I(k,l) - totall);
end
end

I1(i,j) = (total2/100);
total2 = 0;
totall  = 0;
end
end;

save  I1;
```

Combinining AAD vectors into 1 matrix

```
clear
load I1
load I2
load I3
load I4
load I5
load I6
load I7
load I8
load I9
load I10

k = 0;
```

```
for i=1:20
for j=1:8
k = k+1;

Itotal(1,k)=I1(i,j)/100;
Itotal(2,k)=I2(i,j)/100;
Itotal(3,k)=I3(i,j)/100;
Itotal(4,k)=I4(i,j)/100;
Itotal(5,k)=I5(i,j)/100;
Itotal(6,k)=I6(i,j)/100;
Itotal(7,k)=I7(i,j)/100;
Itotal(8,k)=I8(i,j)/100;
Itotal(9,k)=I9(i,j)/100;
Itotal(10,k)=I10(i,j)/100;

end
end
save Itotal;
```

Level 1 - 2D wavelet transformation

```
p = imread('1.jpg');
x = double(p);
[CA,CD,CB,CH] = dwt2(x,'db1');
I1 = CA/1000
save I1;
clear;
```

Level 2 - 2D wavelet transformation

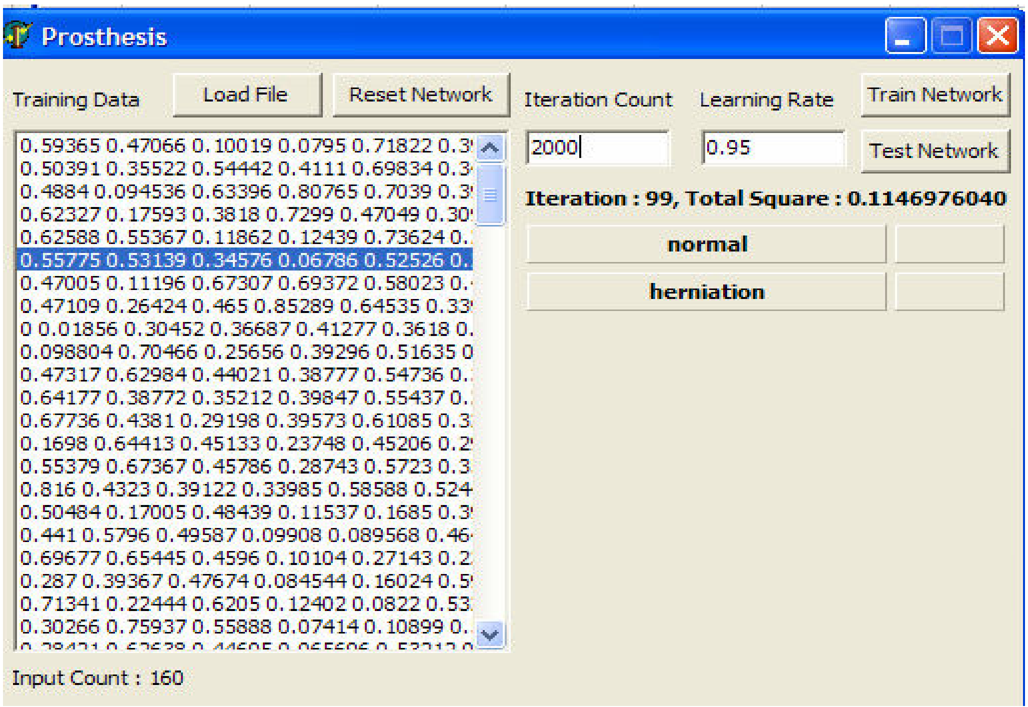
```
clear
p = imread('1.jpg');
X = double(p);
[c,s] = wavedec2(X,2,'db1');
I1 = appcoef2(c,s,'db1',2)/1000;
save I1;
```

ANN PROGRAM

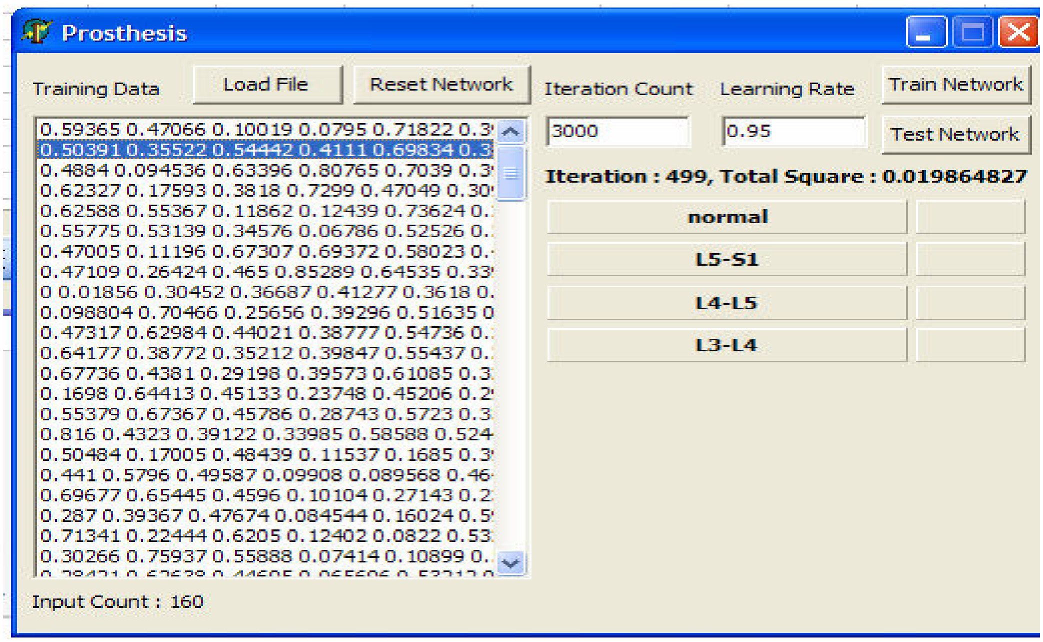
Two classifications are presented in the study. In the first classification, outputs are normal and hernia. In the second classification, outputs are normal, L5-S1, L4-L5 and L3-L4 herniated cases.

In the program, iteration number can be changed in the user interface. Training data is loaded by button "Load File" and "Train Network" button starts training. After training, again by "Load File" button testing data is loaded to the network. "Test Network" button tests the selected input data and displays the result.

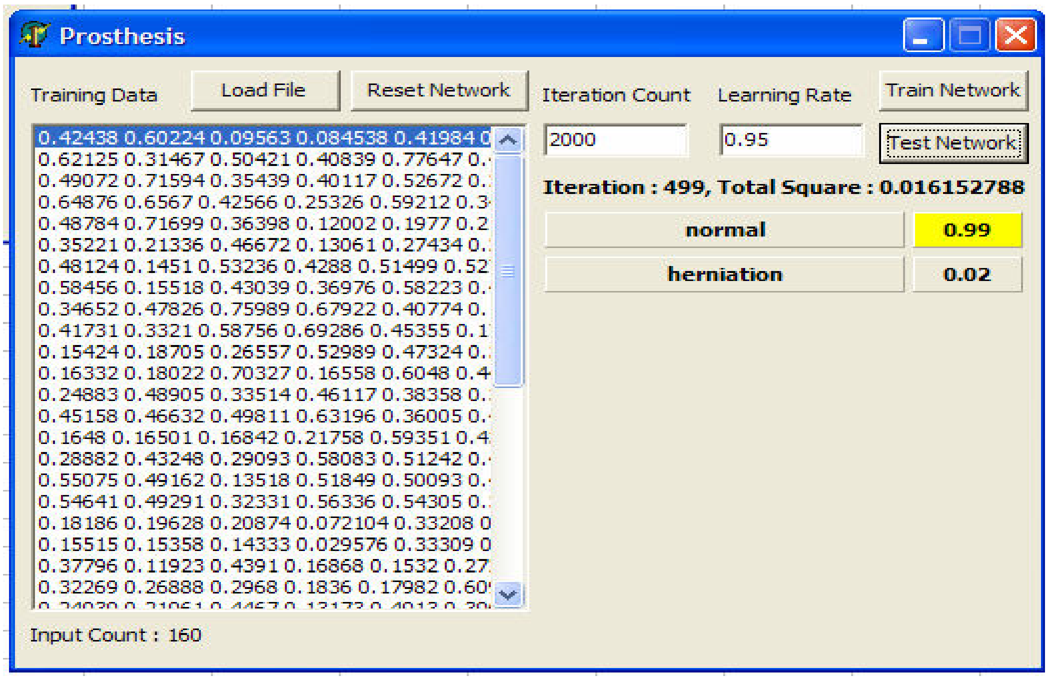
Classification I – Training



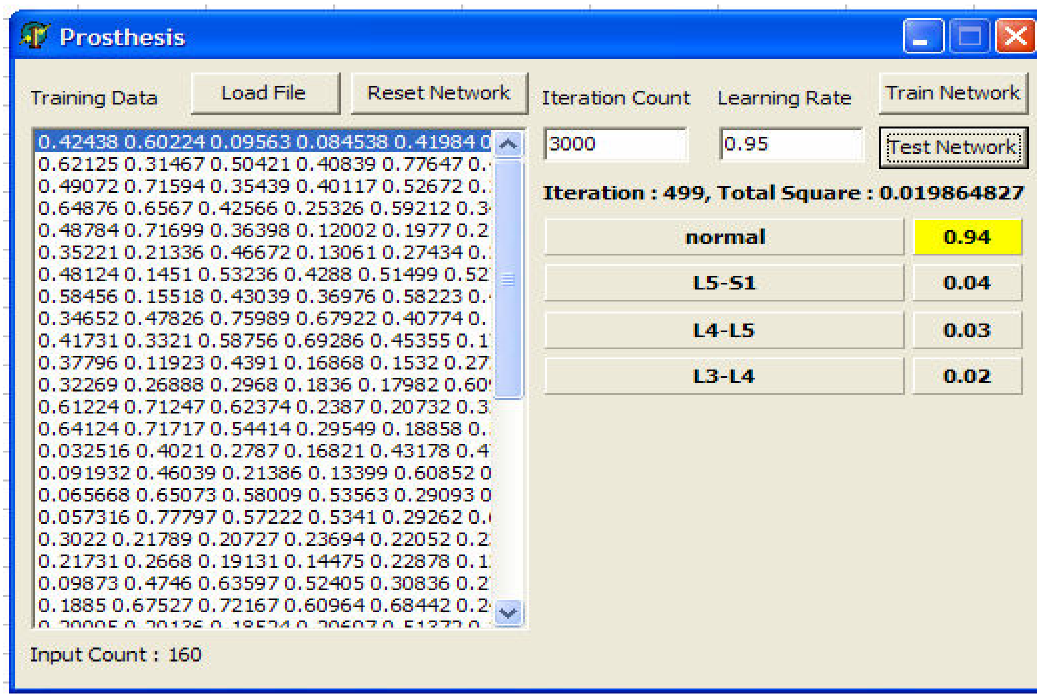
Classification II – Training



Classification I – Testing



Classification II – Testing



Sample input data for a patient

1.jpg	*0,32993*	*0,12143*	*0,349*	*0,23259*	*0,1622*	*0,30911*	*0,57733*	*0,35331*
2.jpg	*0,259*	*0,30368*	*0,49185*	*0,18398*	*0,17308*	*0,31735*	*0,46204*	*0,66914*
3.jpg	*0,37796*	*0,11923*	*0,4391*	*0,16868*	*0,1532*	*0,27297*	*0,50648*	*0,6264*
4.jpg	*0,21418*	*0,18417*	*0,4561*	*0,24706*	*0,15688*	*0,17625*	*0,57519*	*0,5956*
5.jpg	*0,12093*	*0,24172*	*0,52121*	*0,29945*	*0,15342*	*0,1398*	*0,20078*	*0,4064*
6.jpg	*0,029478*	*0,25045*	*0,29661*	*0,33773*	*0,19092*	*0,2304*	*0,23692*	*0,09264*
7.jpg	*0,40112*	*0,26661*	*0,25279*	*0,26849*	*0,18972*	*0,40319*	*0,26338*	*0,70299*
8.jpg	*0,33608*	*0,18238*	*0,31398*	*0,31161*	*0,17668*	*0,56731*	*0,17856*	*0,40659*
9.jpg	*0,32269*	*0,26888*	*0,2968*	*0,1836*	*0,17982*	*0,60902*	*0,27726*	*0,46692*
10.jpg	*0,17784*	*0,10964*	*0,41114*	*0,18808*	*0,15555*	*0,47269*	*0,21865*	*0,47136*