

# Artificial Neural Networks for Microwave Detection

by

Ahmed Ashoor

A thesis  
presented to the University of Waterloo  
in fulfillment of the  
thesis requirement for the degree of  
Master of Applied Science  
in  
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2012

© Ahmed Ashoor 2012

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Abstract

Microwave detection techniques based on the theory of perturbation of cavity resonators are commonly used to measure the permittivity and permeability of objects of dielectric and ferrite materials at microwave frequencies. When a small object is introduced into a microwave cavity resonator, the resonant frequency is perturbed. Since it is possible to measure the change in frequency with high accuracy, this provides a valuable method for measuring the electric and magnetic properties of the object. Likewise, these microwave resonators can be used as sensors for sorting dielectric objects.

Techniques based upon this principle are in common use for measuring the dielectric and magnetic properties of materials at microwave frequencies for variety of applications. This thesis presents an approach of using Artificial Neural Networks to detect material change in a rectangular cavity. The method is based on the theory of the perturbation of cavity resonators where a change in the resonant frequencies of the cavity is directly proportional to the dielectric constant of the inserted objects. A rectangular cavity test fixture was built and excited with a monopole antenna. The cavity was filled with different materials, and the reflection coefficient of each material was measured over a wide range of frequencies.

An intelligent systems approach using an artificial neural network (ANN) methodology was implemented for the automatic material change detection. To develop an automatic detection model, a multi-layer perceptron (MLP) was designed with one hidden layer and gradient descent back-propagation (BP) learning algorithm was used for the ANN training. The network training process was performed in an off-line mode, and after the training process was accomplished, the model was able to learn the rules without knowing any algorithm for automatic detection.

## Acknowledgements

First and foremost I would like to thank my academic advisor, Professor Omar Ramahi for his tremendous help and guidance. Professor Ramahi gave me the academic flexibility and freedom that is necessary for a healthy academic research. His invaluable support and advice provided a guiding light at critical junctures in my academic and personal life. Thank you so much Professor Ramahi. Many thanks also to my committee members, Professor Mansour and Professor Canizares, for their suggestions and critical review of this thesis.

Additionally, I would like to offer my sincere gratitude and thanks to my family. My wife Anssaf, and my children, Zainab and Batool have been my biggest cheerleaders. Anssaf has been extremely patient with me during the long hours and lost weekends that this research has required; without her this would not have been possible.

Finally, I would like to greatly thank my parents for the emotional support that they have given me. My parents always insisted on excellence and in doing well my educational pursuits. I could not have achieved this without their amazing support. I am forever grateful to my parents for their love, blessings and support. Their confidence on my abilities helped me to remain focused and motivated me to give my best.

I would also gratefully acknowledge the financial support from the Ministry of Higher Education in Saudi Arabia.

## **Dedication**

This thesis is dedicated to my mother, father, and wife.

# Table of Contents

List of Tables	viii
List of Figures	ix
<b>1 Motivation and Overview</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Intelligent Systems . . . . .	1
1.3 Research Objective and Overview . . . . .	3
1.4 Thesis Organization . . . . .	4
<b>2 Background</b>	<b>6</b>
2.1 Microwave Principles . . . . .	6
2.2 Cavity Perturbation Theory . . . . .	6
<b>3 Artificial Neural Networks</b>	<b>9</b>
3.1 Introduction . . . . .	9
3.1.1 How the Human Brain Learns . . . . .	10
3.2 Structure of Artificial Neural Networks . . . . .	11
3.3 Type of Activation Functions . . . . .	12
3.3.1 Linear Activation Function . . . . .	13
3.3.2 Non-Linear Activation Function . . . . .	13

3.4	ANNs Topologies . . . . .	14
3.5	ANN Learning Algorithms . . . . .	15
3.6	Multi-Layer Perceptrons . . . . .	16
3.7	BackPropagation Learning Algorithm . . . . .	17
3.8	Determination of the Network Layers and Elements . . . . .	18
	3.8.1 Number of Hidden Layers . . . . .	19
	3.8.2 Number of Hidden-Layer Elements . . . . .	19
3.9	Feature selection . . . . .	19
3.10	Types of Cross-Validation . . . . .	21
	3.10.1 Holdout Cross-Validation . . . . .	21
	3.10.2 k-Fold Cross-Validation . . . . .	21
	3.10.3 The Overall Holdout . . . . .	21
3.11	Over-fitting . . . . .	22
<b>4</b>	<b>Experimental Setup and Results</b>	<b>23</b>
4.1	Experimental Setup . . . . .	23
4.2	Pre-Processing Algorithm and Feature Selections . . . . .	27
4.3	Neural Networks Analysis . . . . .	32
<b>5</b>	<b>Conclusion and Future Work Directions</b>	<b>46</b>
	<b>References</b>	<b>47</b>

# List of Tables

3.1	Terminology and Analogies of Artificial Neural Networks . . . . .	10
4.1	First Stage: Neural Networks Specification . . . . .	35
4.2	First Stage: Neural Networks Results . . . . .	35
4.3	Second Stage: Neural Networks Specification . . . . .	38
4.4	Second Stage: Neural Networks Results . . . . .	38
4.5	Third Stage: Neural Networks Specification . . . . .	43
4.6	Third Stage: Neural Networks Results . . . . .	43



# List of Figures

3.1	Structure of human neuron [1]. . . . .	11
3.2	A typical scalar product of an artificial neuron . . . . .	12
3.3	Linear activation function . . . . .	13
3.4	Log sigmoid activation function . . . . .	14
3.5	Tan-sigmoid activation function . . . . .	15
3.6	General multilayers feedforward network. . . . .	17
3.7	Backpropagation of errors for a single neuron . . . . .	18
4.1	Experimental Setup . . . . .	23
4.2	Experimental Setup . . . . .	24
4.3	$S_{11}$ of Cavity filled with Corn Boxes . . . . .	25
4.4	$S_{11}$ of Cavity filled with Corn Boxes . . . . .	25
4.5	$S_{11}$ of Cavity filled with Corn Boxes . . . . .	26
4.6	$S_{11}$ of Cavity filled with Metal . . . . .	26
4.7	$S_{11}$ of Cavity filled with Water Boxes . . . . .	27
4.8	$S_{11}$ of Cavity filled with Water Boxes . . . . .	27
4.9	$S_{11}$ of Cavity filled with Wood Sheets . . . . .	28
4.10	$S_{11}$ of Cavity filled with Wood Sheets . . . . .	28
4.11	$S_{11}$ of Cavity filled with Metal Sheets . . . . .	29
4.12	$S_{11}$ of Cavity filled with Metal Sheets . . . . .	29
4.13	$S_{11}$ of Cavity filled with Corn Sheets . . . . .	30

4.14	$S_{11}$ of Cavity filled with Corn Sheets . . . . .	30
4.15	$S_{11}$ of Cavity filled with Water Sheets . . . . .	31
4.16	$S_{11}$ of Cavity filled with Water Sheets . . . . .	31
4.17	$S_{11}$ of Cavity filled with 3/4 Corn Boxes and 1/4 empty . . . . .	32
4.18	$S_{11}$ of Cavity filled with 3/4 Corn Boxes and 1/4 empty . . . . .	32
4.19	$S_{11}$ of Cavity filled with 3/4 Corn Boxes and 1/4 Metal . . . . .	33
4.20	$S_{11}$ of Cavity filled with 3/4 Corn Boxes and 1/4 Metal . . . . .	33
4.21	$S_{11}$ of Cavity filled with 3/4 Corn Boxes and 1/4 Water . . . . .	34
4.22	$S_{11}$ of Cavity filled with 3/4 Corn Boxes and 1/4 Water . . . . .	34
4.23	Peaks Selected of $S_{11}$ of Cavity filled with Water Sheets . . . . .	36
4.24	Peaks Selected of $S_{11}$ of Cavity filled with 3/4 Corn Boxes and 1/4 Metal .	36
4.25	Peaks Selected of $S_{11}$ of Cavity filled with 3/4 Corn Boxes and 1/4 Water .	37
4.26	Peaks Selected of $S_{11}$ of Cavity filled with Metal Sheets . . . . .	37
4.27	First Stage: Confusion Matrix of Training Dataset . . . . .	39
4.28	First Stage: Confusion Matrix of Testing Dataset . . . . .	39
4.29	First Stage: ROC of Training Dataset . . . . .	40
4.30	First Stage: ROC of Testing Dataset . . . . .	40
4.31	Second Stage: Confusion Matrix of Training Dataset . . . . .	41
4.32	Second Stage: Confusion Matrix of Testing Dataset . . . . .	41
4.33	Second Stage: ROC Matrix of Training Dataset . . . . .	42
4.34	Second Stage: ROC Matrix of Testing Dataset . . . . .	42
4.35	Third Stage: Confusion Matrix of Training Dataset . . . . .	44
4.36	Third Stage: Confusion Matrix of Testing Dataset . . . . .	44
4.37	Third Stage: ROC of Training Dataset . . . . .	45
4.38	Third Stage: ROC of Testing Dataset . . . . .	45

# Chapter 1

## Motivation and Overview

### 1.1 Introduction

In a wide variety of applications from security to material characterization, the main task is to identify a small perturbation to the material inside a given region. Fields of physics, engineering, material science, and even biology and medicine have greatly benefited from microwave measurement techniques based on cavity perturbation theory [2, 3, 4, 5, 6, 7].

In security applications, for instance, the perturbation can be a hidden gun, concealed narcotics, rotten food or even impurity in the chemical composition of certain substance. Thus, to detect such a change, the detection can be performed using near-field or far-field sensors in an open region. In such methods, the scattered field is analyzed, thus limiting the application since if a material to be detected is shielded by a different object, detection becomes difficult.

The method that is proposed here works based on change in the resonant frequencies of a cavity that contains the material under test. Unlike the detection modalities based on the scattered field, here the whole region of the cavity is covered by the resonator and microwave field penetrates inside the whole material.

### 1.2 Intelligent Systems

An intelligent system is composed of a knowledge base, an inference engine, a working memory, a user interface and an explanation based subsystems [8]. There are several intelligent systems approaches that are widely used. Such approaches include but are not

limited to rule based reasoning, fuzzy logics, artificial neural networks, genetic algorithms, case based reasoning. Intelligent systems approaches have been practised in various applications for many decades but recently have received a considerable amount of attention in microwave applications [9, 10, 11, 12, 13, 14, 15, 16, 17, 18].

It was demonstrated in [19, 20] that an intelligent system is able to learn the rules without the need to know any algorithms. However, it is an absolute necessity to improve the networks by improving its structure in order to build an accurate intelligent system. Therefore, applying better selections of normalization procedure, signal/data pre-processing, and training the network with an accurate data can lead to having a more accurate system in terms of learning to obtain desired results.

In recent years, Artificial Neural Network (ANN) has been recognized as one of the most powerful techniques for modelling highly complex systems. ANN is an information processing system, which can learn from observation and then generalize an arbitrary multidimensional non-linear input-output relationship [21]. The ability of the ANN models to detect non-linear relationships between parameters is due to their unique training method and structure. Thus, ANN models have been utilized in many area of engineering and scientific applications such as remote sensing, communication, power systems, energy systems, biometric, controls, robotics, pattern recognition, weather forecasting, medicine, manufacturing, optimization and social/psychological sciences, etc.

On the other hand, there are some potential limitations associate with ANN. ANNs are in a sense the ultimate 'black boxes', and the final output of it is a trained network that provides no equations or coefficients defining a relationship beyond it's own internal mathematics. Thus the ability of systems to behave intelligently as well as to learn does not increase by simply using many interacting computing units. This is because intelligent behaviour depends heavily on the algorithmic of the particular architecture. Thus with any approach, intelligent neural network architectures require much engineering work. This engineering work cannot be essentially reduced by using only connectionist approaches.

Despite the fact that Artificial Neural Network has been studied for many years and applied to other fields, the application of ANN is relatively new to electromagnetics [22]. The work in [22] also demonstrates that errors reported when using these Neural Network models are well within acceptable limits, which clearly suggests that artificial Neural Networks can be used for modeling in other fields, such as electromagnetics.

Some related work in the field of electromagnetics includes the work that was done by Stankovic et al. in [23]. In this work, Neural Network techniques were used to model microwave properties of a raised dielectric slab in a cylindrical cavity. Existing partial knowledge about the resonant frequency behaviour of the loaded cavity is incorporated

in the model using a knowledge-based network. Also, ANN with the use of microwave means in the medical field was done by M. El-Shenawee [24]. Interpreting Artificial Neural Networks for Microwave Detection of Breast Cancer, El-Shenawee, M. In this work, a simplified model of the breast containing a tumor was used to determine the scattering of electromagnetic waves in the microwave band. Then the obtained data was used to train and test a ANN in order to predict whether a tumor is present or not.

In addition, Penaranda Foix F. L. in [25] presented a practical expression for frequency deviation due to the insertion of a dielectric material in a cylindrical cavity. This technique is often used to measure the permittivity of the inserted material, which is similar to the material change classification problem. Similarly, Farahani A. V. in [26] presented a finite-difference time-domain calculation of cavity resonant frequencies in the case of non-uniform internal magnetic field distributions. In [27], Hill presented a comprehensive theoretical analysis of cavity resonance in his book *Electromagnetic Fields in Cavities: Deterministic and Statistical Theories*. Raveendranath U. in [28] also presented a perturbation technique for measuring complex permeability of ferrite materials. A rectangular waveguide cavity was used in [28] to measure the magnetic properties of ferrite materials.

One more application of Neural Networks in the field of electromagnetics was done by Ma J. F. in [29]. He addressed the problem of detecting low-dielectric constant cavities buried deep in lossy ground by using finite-difference time-domain methods in conjunction with Neural Networks for extrapolation and object identification. According to [29], it was found that the Neural Network approach was good at approximating the relationship between the field scattered by the object and its depth.

### 1.3 Research Objective and Overview

The main purpose of this research work was to develop an automatic method that could assist developing a detection system by electromagnetic means.

In this thesis, a change in resonant frequencies technique has been employed to measure the small change in resonant frequencies for different materials using a rectangular cavity and monopole antenna. The effect of any change in the material inside a cavity will be analyzed. Then presenting a solution to what I call "the cavity material change classification technique", i.e. the problem of classifying material changes in a cavity. The broader aims of this work focus on the analysis of rectangular cavities excited by only one type of antenna. A Neural Network classification method is applied to a set of experimental data where the goal of the classifier is to distinguish between different material perturbations

using the changes in the resonant frequencies of the cavity. The Neural Network classifier is required because small changes in the location of the resonant frequencies may not be apparent to the human eye but could be detected using a classifier like the Neural Network.

The Artificial Neural Network is not the only possible choice of classification method, and the proper selection of classification method is very important. The non-linear nature of the proposed problem and because of the fact that the distributions of classes are unknown, linear classifiers or Bayesian classifiers are hard to use. Therefore Neural Network classifiers, Fuzzy classifiers and Nearest Neighbour classifiers are excellent candidates. In this thesis, a Neural Network classifier is used because of its flexibility and generality in situations where less information is available about the nature of the problem.

In this thesis, a Neural Network classifier is applied to the selected resonant frequencies of the experimental data to classify small changes in the resonant frequencies. These changes in resonant frequencies may be noticeable for identical objects and may not be noticeable when testing different objects. Thus the purpose of building a neural network model is to distinguish between these changes and then assign them the right class. In the used Neural Networks, multi-layer perceptron neural networks (MLP) feedforward was considered. The number of the hidden layer was determined during the ANN training process. The back-propagation (BP) learning algorithm was used for training process in an off-line mode. The performance of the designed networks was evaluated by the test database in terms of Recognition Rate (RR), i.e. ratio of the number of input data correctly classified to the total number of inputs and a specific error rate. Several advantages that multilayer feedforward with backpropagation learning algorithm has above other NN types. One advantage is that it can be used as a classifier as well as a pattern recognition which will serve the propose on this thesis. Also, its incremental learning algorithm where at any stage during the training process, training can be stopped and the network would still serve as a model of function being learned. Not only that but also it can be used in developing empirical models based on experimental knowledge and gives reasonable accuracy even if the given data is noisy or incomplete. Finally, its hidden layers perform feature extraction on the presented patterns at the input layer. All these advantages serve the classification purpose of this thesis.

## 1.4 Thesis Organization

Chapter 2 presents brief theoretical background on cavity perturbation theory in electromagnetics. A general discussion of cavity perturbation is also presented in this chapter. In Chapter 3, Artificial Neural Networks is discussed in general. Various neural network

structures and learning algorithms are presented. The focus of this Chapter 3 is mostly on previously developed methods and algorithms. Chapter 4 presents the experimental setup for the proposed detection problem. Development and analysis of the ANN model is presented. Finally, in Chapter 5, a summary of the thesis is presented, highlighting the contributions of the proposed detection technique using ANN model. Further research and some suggestions for future research are presented in this chapter as well.

# Chapter 2

## Background

### 2.1 Microwave Principles

The microwave spectrum ranges from 300 MHz to 300 GHz with wavelengths ranges from 1 meter to 1 mm [30]. Microwaves can deeply penetrate into dielectric and composite materials. This interaction called material energy interaction. The loss factor associated with dielectric materials and the frequency at which the experiment is conducted are the two main factors affecting the material energy interaction. The key advantages given to the microwave techniques are that they are considered as a non contact technology where no need for a couplant and high power [30].

Measurement of complex permittivity and permeability of dielectric and magnetic material plays an important role in microwave technology. Several techniques such as cavity perturbation [31], open-ended coaxial probes [32], free space [33], waveguide transmission line [34] and dispersive fourier transform spectrometer [35] have all been developed for material permittivity and permeability measurement. Nevertheless, the cavity perturbation method has a reputation as one of most simple and accurate techniques for measuring dielectric and magnetic properties of material.

### 2.2 Cavity Perturbation Theory

The earliest treatment of cavity perturbation theory was introduced by Bethe and Schwinger [31]. In many applications, cavity resonators are modified by making small changes in their



shape, or by introducing small pieces of dielectric material. By measuring the shift in the resonant frequency, the material's dielectric constant can be determined.

In some cases, one can exactly calculate the change of the cavity's response, but more often approximations must be used. The cavity perturbation method provides a formalism for this by first assuming that the eigen-modes of the fields in the cavity are unchanged from those of the unperturbed cavity which is similar to the first order perturbation theory in quantum mechanics.

Furthermore, since the majority of the microwave energy is concentrated in the cavity, the fields distribution inside the cavity will not be disturbed significantly by the tip-sample interaction. Thus the cavity perturbation theory is a good approach for calculating the frequency shift caused by changes in the dielectric environment. With this theory, one can relate the resonant frequency shift of the cavity to a change in the permittivity  $\Delta\epsilon$ , or permeability  $\Delta\mu$  in the volume occupied by sample. This suggests that change in the resonant frequencies of a cavity could be used to identify material change.

To demonstrate this change in the resonant frequencies, an air filled rectangular cavity with a resonant mode at frequency  $\omega_0$  is considered. If the dielectric constant inside a region  $\Delta\tau$  is changed to  $\epsilon_r$  then change in the resonant frequency is given by the following [36]:

$$\frac{\Delta\omega}{\omega_0} = -\frac{\int \int \int (\Delta\epsilon \vec{E} \cdot \vec{E}_0^* + \Delta\mu \vec{H} \cdot \vec{H}_0^*) d\tau}{\int \int \int (\epsilon \vec{E} \cdot \vec{E}_0^* + \mu \vec{H} \cdot \vec{H}_0^*) d\tau} \quad (2.1)$$

where  $\vec{E}_0$  and  $\vec{H}_0$  are the unperturbed electric and magnetic fields, respectively,  $\vec{E}$  and  $\vec{H}$  are the perturbed electric and magnetic fields inside  $\Delta\tau$  in the presence of the dielectric material, respectively, and  $\Delta\epsilon$  and  $\Delta\mu$  are the changes in the electric permittivity and magnetic permeability in the material due to the perturbation, respectively. Taking  $\Delta\epsilon$  and  $\Delta\mu$  to be small and approximating  $\vec{E} = (1/\epsilon_r)\vec{E}_0$ , we have the following [36]:

$$\frac{\Delta\omega}{\omega_0} \approx -\frac{\int \int \int_{\Delta\tau} (\epsilon_r - 1) \vec{E} \cdot \vec{E}_0^* d\tau}{2 \int \int \int \epsilon |\vec{E}_0|^2 d\tau} \quad (2.2)$$

A rectangular cavity is only considered with a thin dielectric slab positioned along one of the sides of the cavity. This consideration gives the following [36]:

$$\frac{\Delta\omega}{\omega_0} \approx -\frac{(\epsilon_r - 1) \bar{\omega}_e^\tau}{2\epsilon_r \bar{\omega}_e} \quad (2.3)$$

where  $\bar{\omega}_e^\tau$  and  $\bar{\omega}_e$  are the stored electric energies in  $\Delta\tau$  and the entire cavity, respectively (both are calculated in the absence of the dielectric media). Therefore, the percentage of change in the resonant frequency is proportional to the dielectric constant of the inserted material. The factor  $\bar{\omega}_e^\tau/\bar{\omega}_e$  however, depends on the size of the dielectric slab. If the dielectric slab is small compared to the wavelength, a change in its size will change all the frequency shifts of all resonant modes by the same factor. Nevertheless, since the dielectric constant of most materials changes with frequency, a change in material of the slab will have different effects on different resonant frequencies. This allows the possibility of defining the material of a small slab using changes in the resonant frequencies.

The ratio  $\bar{\omega}_e^\tau/\bar{\omega}_e$  also depends on the position of the dielectric slab relative to the electric field pattern inside the cavity. This ratio is maximized when the object is placed in the location of maximum electric field strength and is zero when the object is placed at an electric field null. Since the electric field pattern changes for different modes, it may also be possible to define the position of the object using change in resonant frequencies.

The above discussion explains the overall picture of the conducting experiment. This can be used in many applications from security to material characterization where the main task is to identify a small perturbation to the material inside a given region. The perturbation can be a hidden gun, concealed narcotics, rotten food or even impurity in the chemical composition of certain substance. Whereas in the medical field, for instance, this method can be used to detect tumor in breast cancer.

In this thesis, this method is used to detect material change in a rectangular cavity with the use of Neural Networks. It is based on the theoretical knowledge that a change in the resonant frequencies of the cavity is directly proportional to the dielectric constant of the inserted material. A rectangular cavity test fixture was constructed and excited with a monopole antenna and the reflection coefficient measured over a range of frequencies. Different materials were inserted into the cavity including metal, water, wood and air. The experimental setup will be discussed in great details in chapter 4.

# Chapter 3

## Artificial Neural Networks

### 3.1 Introduction

The nervous system [37] is composed of millions of interconnected nerve cells where each one performs a specific task. Since the interconnections between those cells are complex, the outcome of these multi simple task's accomplishments is the execution of a bigger more complicated task. In a similar manner, Artificial Neural Networks (ANNs) are composed of a large number of interconnecting neurons. The essence of ANNs are in the mimicking of how the nervous system in the human processes information.

In 1940, the first neural networks computing model was introduced by McCulloch and Pitts, and since then the evolution of this field has been began. Connectionist systems, parallel distributed systems, or adaptive systems are other names in which the neural networks have been called. Neural networks can be simply stated as a processing technique, mainly based on mimicking the human brain in which massive interconnected layers process information simultaneously and adapt over the course of multiple runs. Thus a neural network attempts to generalize from a known data to new unknown data. Table 3.1 shows the common terms used in reference to neural networks and their analogous on brain expressions.

Interest in the artificial neural networks subsided between years 1950 to 1980 due to lack of hardware capability which was needed to endorse a learning machine. In 1982, a paper was presented by John Hopfield to the National Academy of Sciences which motivated researchers to revisit their interests in the field. His paper presented the benefits of using bidirectional interconnects instead of the traditional unidirectional networks. At that

<b>ANN Terminology</b>	<b>Brin Equivalent</b>
Node	Neuron
Interconnect	Synapse
Weight	Simulation

Table 3.1: Terminology and Analogies of Artificial Neural Networks

time, the hardware had developed to the point where neural networks could be utilized in practical and more efficient ways.

Artificial Neural Networks (ANNs) have been used in a wide variety of application. ANNs are excellent tools for approximating any given function of single or multi-variable inputs and outputs. Due to the flexibility inherent in ANNs' configurability, they can be used in many decision making processes. Pattern recognition and numerical classifications are ones of those typical applications that use ANNs [38]-[39].

### 3.1.1 How the Human Brain Learns

Artificial Neural Networks success at system modelling for highly complex physical processes can be attributed to the original architecture on which they are based, the human brain. At present, brain function is not fully understood. A brain neuron collects signals from other neurons of the Central Nervous System (CNS), through structures called dendrites as shown in Figure 3.1. The neuron sends out spikes of electrical activity through a long thin strand called an axon. This axon splits into thousands of branches. At the end of a branch, a structure called a synapse converts the activity from the axon into electrical effects that may excite or inhibit activity in the connected neurons.

When a neuron receives an excitatory input that is sufficiently large compared with its inhibitory input, it sends a spike of electrical activity down its axon. Learning occurs by changing the effectiveness of the synapses so that the influence of one neuron on another changes [1].

The structure of the human brain neuron is the template for artificial learning. However, lack of knowledge leads to approximations and assumptions of the general architecture of an artificial neural network. The knowledge of neurons is incomplete and computing power is limited so models are often idealisations of real networks of neurons. Figure 3.2) is a typical representation of a scalar product neuron.

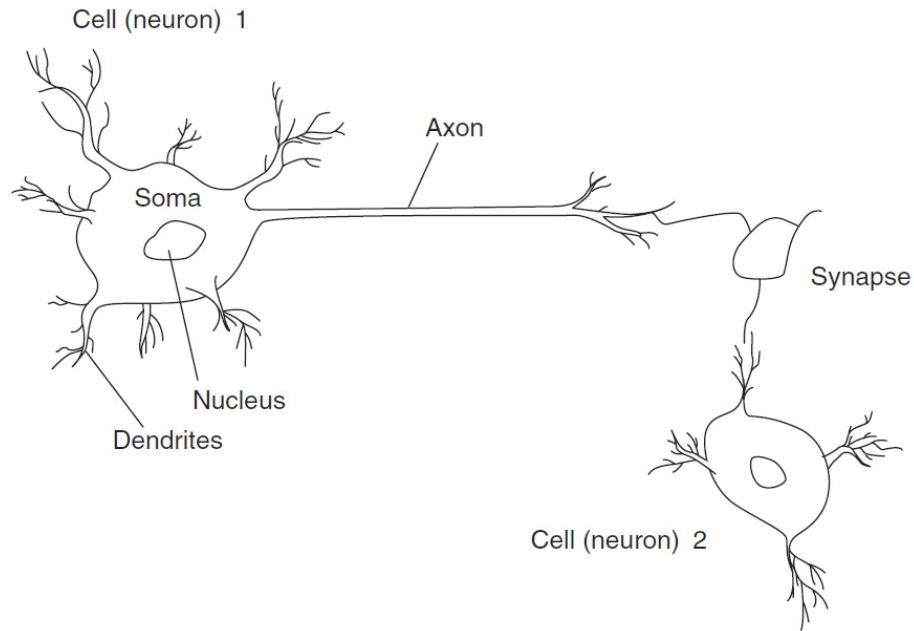


Figure 3.1: Structure of human neuron [1].

## 3.2 Structure of Artificial Neural Networks

An artificial neural network is comprised of an input layer, number of hidden layers and output layer. The main responsibility of the input layer is to receive the external measurements 'data' and propagate it forwards to the hidden layer or layers. In each layer there are number of nodes called neurons, and each neuron is capable of communicating with its neighbours through weights and is capable of changing its present state depending on the signal it receives. Figure 3.2 shows a scalar product of one artificial neuron. There are more than one neuron in the network, and neurons can simultaneously change state while maintaining the global state or purpose of the network. The neurons are connected to each other in layers with the first layer being the input layer and the final layer being the output. Whatever is in between those two layers are called the hidden layers. The objective of a neural network is, for a given set of inputs, to give a set of desired or expected outputs. The weights or what it is called synaptic weights are updated using a certain learning algorithm until the error between the actual output of the neural network and the desired output is reduced to an acceptable level which will be depending on the function being approximated.

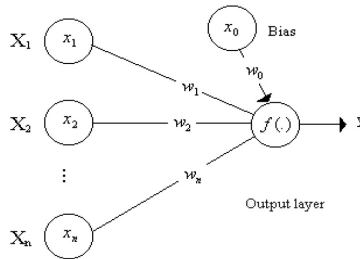


Figure 3.2: A typical scalar product of an artificial neuron

The number of hidden layers is not restricted in to a certain number. However, Kolmogorov’s Mapping Neural Network has proven that one hidden layer capability performance with proper configuration is identical to multiple hidden layers [40]. An activation function is presented at each node in the hidden layer in which manipulating the output that is passed to the next layer. An activation function is also presented at the output layer to limit the output under certain utilization. One more important parametric structure of the network is the number of nodes being included in the hidden layer or layers. The number of these nodes affects the accuracy of the results to some degree. It is usually beneficial to have a large number of hidden nodes while serving extremely complex networks. Determining the suitable number of nodes is not an easy task, but typically it is determined by repeat testing the network until getting the desired one. However, a trade-off is involving, the more complex the network is, the longer training time is required. Training effort exponentially increases as the number of layer increases which will increase the computation required for training. Not only that but also it may negatively impact in the results. Thus the use of more than three hidden layers is rare in most systems.

### 3.3 Type of Activation Functions

The artificial neuron like the biological neuron described in Figure 3.1 is a processing element. An output for this artificial neuron is calculated by multiplying its inputs by a weight vector. The results are then added together and an activation function is applied to the sum. The activation function is a function used to transform the activation level of a unit or rather a neuron into an output signal. Typically, activation functions have a squashing effect; they contain the output within a range.

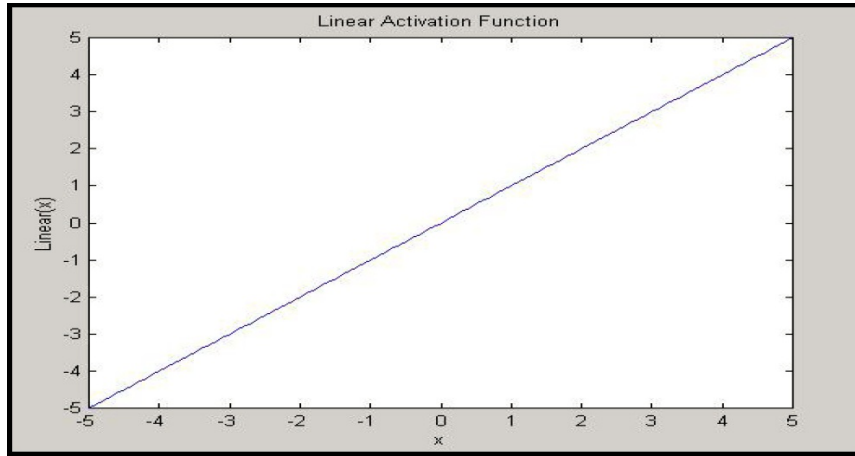


Figure 3.3: Linear activation function

### 3.3.1 Linear Activation Function

There are many activation functions that can be applied to neural networks; three main activation functions are dealt with in this project [41]. The first is the linear transform function which is shown in Figure 3.3, or pure-line function. It is defined as follows

$$f(x) = x \quad (3.1)$$

Neurons of this type are used as linear approximators on the function being approximated.

### 3.3.2 Non-Linear Activation Function

There are several types of non-linear activation functions; the two most common are the log-sigmoid transfer function and the tan-sigmoid transfer function. Plots of these differentiable, non-linear activation functions are illustrated in Figures 3.4 and 3.5. They are commonly used in networks that are trained with back-propagation.

The first sigmoid activation function which is called logistic activation function is shown in Figure 3.4 and defined as follows

$$\text{logsig}(x) = \frac{1}{1 - \exp(\beta(x))} \quad (3.2)$$

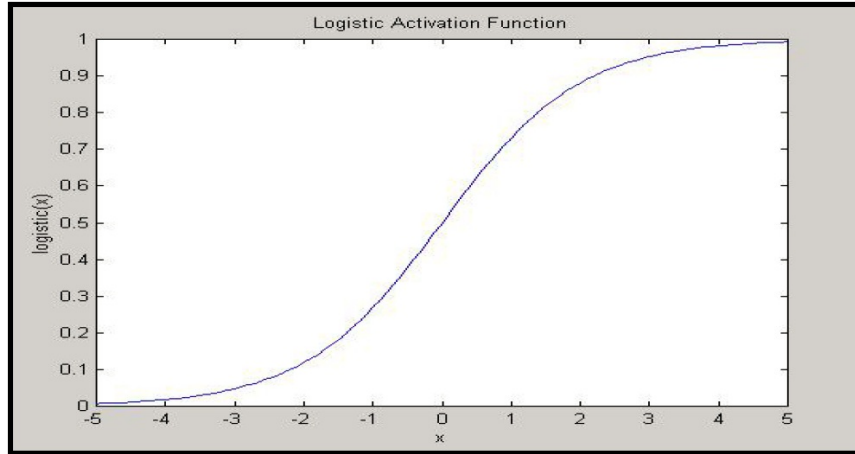


Figure 3.4: Log sigmoid activation function

The value of  $\beta$  can be changed in which it changes the shape of the sigmoid. As  $\beta$  tends toward infinity it behaves more and more like a hard-limiter where the slope of the sigmoid is zero. In the case when the slope is not zero, the output range is squeezed between 0 and 1.

The second mentioned activation function is called tan-sigmoid activation function. Unlike log-sigmoid activation function, this function squeezes the output between -1 and 1. Tan-sigmoid activation function is shown in Figure 3.5 and defined as follows

$$tansig(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \quad (3.3)$$

During the initial testing, different activation functions were tested. The activation function which is used then in this thesis is the Tan-sigmoid function. The Tan-sigmoid function is used in both the hidden and output layers of the built network. It shows a reasonable accuracy with the non-linear proposed problem comparing to other functions.

### 3.4 ANNs Topologies

ANN can be typically classified into two group of topologies based on connection patterns between their units and data propagation between them. These include feedback neural networks and feed-forward neural networks.



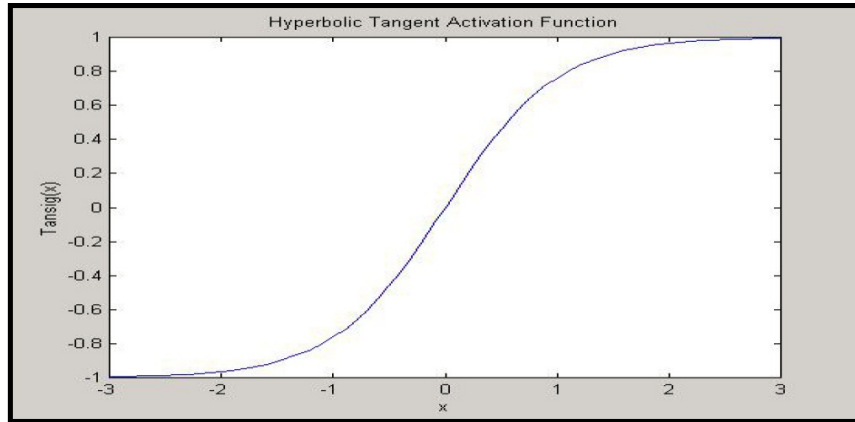


Figure 3.5: Tan-sigmoid activation function

**Feedback neural networks:** This type of network has feedback connections i.e. cycles or loops are present within this type of network. In such networks, data flows from input to output units or vice-versa. Feedback occurs in a dynamical system when the output of an element in the system influences its particular input, which give rises to one or more closed paths for signal transmission around the system [42]. They are also referred to as the recurrent networks. Kohonen networks, and Hopfield networks are examples of feedback neural networks.

**Feed-forward neural networks:** In this type of network, the data usually flows from input to output units in feedforward fashion. This is one of the simplest ANN where the information moves in only one direction from input to hidden (if present) to output nodes.

There is not any kind of cycles or loops within this type of network. Perceptron and Adaline are the classical examples of this kind of network. This research work uses the feed-forward networks.

### 3.5 ANN Learning Algorithms

After defining the structure of the ANNs and its topology, a training algorithm must be chosen to train the network. ANN can be trained in various ways depending upon the network structure and learning process. The basic learning rules include: error-correction learning, memory-based learning, Hebbian learning, competitive learning, and Boltzmann learning. However, all these learning methods applied for training ANN can be classified into two main classes:

**Supervised learning:** This type of learning process is same as learning with a teacher. In this, each output pattern is provided with its desired response for its corresponding input pattern. During the learning process, the system gradually adjusts with its environment by adjusting its network parameters. Error-correction learning, reinforcement learning, and stochastic learning are some of the types of supervised learning. For this type of learning process error convergence is very important. Then the error has to be minimized between the desired response and the actual response of the network. Usually, supervised learning process is carried out off-line. ANN training applied for this research work uses supervised learning process.

**Unsupervised learning:** This type of learning is similar to learning without a teacher. It is also called a self-organisation map, i.e. it self-organises data presented to it and detects their emergent collective properties. Hebbian learning, and competitive learning are examples of unsupervised learning. This type of learning is usually carried out in online mode.

## 3.6 Multi-Layer Perceptrons

There is a wide variety of ANNs, but the most commonly used NN type is the Multi-Layer Perceptron (MLP), alternatively called feedforward neural network. The most commonly learning algorithm that is used within the feedforward MLP is called error back-propagation.

The MLP with back-propagation algorithm is generally characterized by the presence of an input layer, one or more hidden layers, and an output layer as shown in Figure 3.6. The input signal propagates through the network in a forward direction, on a layer-by-layer basis [42]. MLP have been successfully applied to solve various diverse and complex problems [42] by using a supervised learning process.

Finally, a set of outputs are produced as the actual response of the network. This generates a difference (error) between the output of the network and the desired output. The synaptic weights of all networks are all fixed during the forward pass. Similarly, during the backward pass, error signal is computed and it propagates backward through the network. Finally, the synaptic weights are all adjusted in accordance with an error-correction rule. This is how this particular algorithm received its name from 'back-propagation'.

Training time can range from seconds to days depending on the problem domain, the amount and representation of data, and the implementation of the network. Once trained,

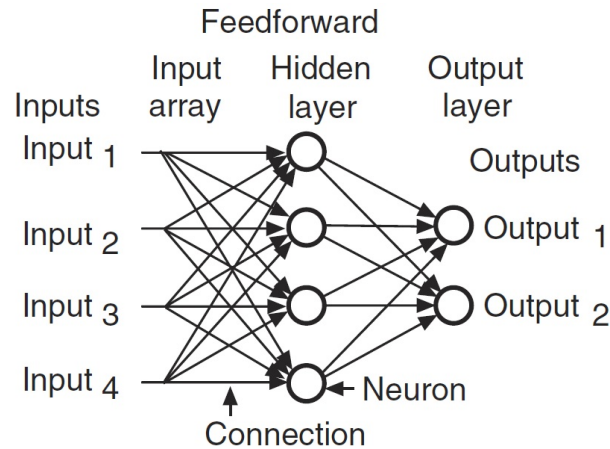


Figure 3.6: General multilayers feedforward network.

however, neural networks are orders of magnitude faster when it comes to evaluating new examples.

### 3.7 BackPropagation Learning Algorithm

Backpropagation which, more descriptively, can be called back-error propagation, is the most widely used supervised learning algorithm in neural computing [43]. A backpropagation network is very easy to implement, and it includes one or more hidden layers with its input and output layers. This type of network is considered feedforward because there are no interconnections between the output of a processing element and the input of a node in the same layer or in a preceding layer. Externally provided correct patterns are compared with the neural networks output during (supervised) training, and feedback or precisely a back propagate is used to adjust the neurons connecting weights until the network has categorized all the training patterns as correctly as desired. The error tolerance must be set in advance .

Starting with the output layer, errors between the actual and desired outputs are used to correct the weights for the connections to the previous layer (see Figure 3.7).

The learning algorithm includes the following procedures:

1. Initialize weights with random values and set other parameters.
2. Read in the input vector and the desired output.

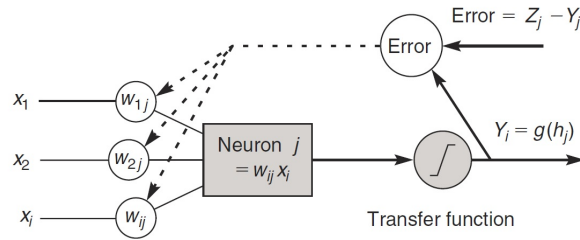


Figure 3.7: Backpropagation of errors for a single neuron

3. Compute the actual output via the calculations, working forward through the layers.
4. Compute the error.
5. Change the weights by working backward from the output layer through the hidden layers.

This procedure is repeated for the entire set of input vectors until the desired output and the actual output agree within some predetermined tolerance. Given the calculation requirements for one iteration, a large network can take a very long time to train; therefore, in one variation, a set of cases are run forward and an aggregated error is fed backward to speed up learning. Sometimes, depending on the initial random weights and network parameters, the network does not converge to a satisfactory performance level. When this is the case, new random weights must be generated, and the network parameters, or even its structure, may have to be modified before another attempt is made.

### 3.8 Determination of the Network Layers and Elements

Although ANN can be designed and trained in various ways, there are several issues that must be resolved first during the training process. These issues include the following [44]:

1. Determining the correct number of hidden layers.
2. Determining the correct number of neurons in each hidden layer.
3. Finding a globally optimal solution that avoids local minima.

4. Converging to an optimal solution in a reasonable period of time.
5. Validating the neural network to test for over-fitting.

### **3.8.1 Number of Hidden Layers**

Once the desired output is achieved, the weights between the nodes are locked, and the network is considered trained at this stage. Then the network is capable to receive a new unseen data that will pass through the trained network and be classified based on it. This step is called the testing stage where the network is able to generalize the training to the new unseen data and correctly making accurate predictions or classifications. The number of hidden layers is selected such that the comparison with prestructured networks would be of apples to apples. The prestructured networks have a single hidden layer, so the fully connected networks to compete with them must have the same type of resource allocation in the form of hidden layers. (More than one hidden layer in multilayer perceptrons is believed to help discriminate among multiple classification criteria to be learned.)

### **3.8.2 Number of Hidden-Layer Elements**

The number of elements per hidden layer (sometimes called network size) is determined through an experimental procedure called k-fold cross-validation [29]. In this method, the generalization performance of any particular instantiation (a network with a certain number of hidden-layer elements) is compared with the performances of various other instantiations. (The only difference distinguishing all the network instantiations is the number of elements in one hidden layer; all other parameters are kept constant for this experiment.) The question arises as to which two network sizes ought to be compared first, and how much should network size be altered for subsequent comparisons.

## **3.9 Feature selection**

Feature selection (also known as subset selection) is a process commonly used in machine learning, wherein a subset of the features available from the data is selected for application of a learning algorithm. The best subset contains the least number of dimensions that most contribute to accuracy; the remaining and unimportant dimensions were disregarded. This is an important stage of pre-processing and is one of the two ways of avoiding the curse of dimensionality (the other is feature extraction). In general, there are two main approaches:

1. Forward selection: Starting with no variables and including them one by one, at each step encompassing the one that decreases the error the most, until any further insertion does not significantly decrease the error.
2. Backward selection: Starting with all the variables and removing them one by one, at each step removing the one that decreases the error the most (or increases it only slightly), until further removal increases the error significantly

Feature selection is the problem of choosing a small subset of features ideally necessary and sufficient to perform the classification task, from a larger set of candidate features. Feature selection has long been one of the most important topics in pattern recognition and it is also an important issue in supervised learning. If one could select a subset of variables, one could reduce the size of the data used to training a supervised method, the amount of data to process, the training time, and possibly increase the generalization performance.

In the case of NNs, direct estimation methods are preferred because of the computational complexity of training an NN. Inside this category we can perform another classification: methods based on the analysis of the training set [45].

There are many different methods to define feature selection on a trained neural network. Some of them are specifically based on the analysis of a trained feedforward network and the others are general methods which have the ability of doing feature selection in all kinds of trained networks.

For example, according to the definition of relevance of an input unit  $S_i$ , one input  $I_i$  is considered more important if its relevance  $S_i$  is larger. Also relevance  $s_{ij}$  of a weight  $w_{ij}$  connected between the input unit and the hidden unit is defined. The relation between  $s_i$  and  $s_{ij}$  is in equation 3.4, where  $N_h$  is the number of hidden units.

$$s_i = \sum_{i=1}^{N_h} s_{ij} \quad (3.4)$$

The criteria for defining weight relevance are varied. Some of them are based on direct weight magnitude. As an example, the criterion proposed by Belue [46] is in equation 3.5.

$$s_{ij} = (w_{ij})^2 \quad (3.5)$$

Other criteria of weight relevance, introduced by Tekto [47], is based on an estimation of the change in the MSE (Mean Square Error), when setting the weight to 0.

Another proposed method by Utans [48] focuses on the MSE, and calculates its increment when substituting an input for its mean value. One input is considered more relevant if the increment of MSE is higher.

## **3.10 Types of Cross-Validation**

There are two variants of cross-validation, known as Holdout Cross-Validation and Multi-fold (k-Fold) Cross-Validation.

The experimental design for the present research uses k-fold cross-validation for parameter estimation (designing the network, also called model selection) and a single holdout for performance estimation (evaluating the designed network, or model assessment).

### **3.10.1 Holdout Cross-Validation**

In Holdout Cross-Validation, the available data are divided into three subsets: the estimation subset, the validation subset, and the test set. The first two are collectively known as the development data set (also called design data in this document); the estimation subset is commonly called the training set; and the test set is alternately known as the holdout set or the out-of-sample test set.

### **3.10.2 k-Fold Cross-Validation**

In k-Fold Cross-Validation, there is still a design set and a test set, but the design set is divided into k equal subsets. Each of the k subsets (folds) is set aside as a temporary validation set while training is conducted using the other k-1 subsets. Thus, k separate trainings take place, with significant overlap in training data but no overlap in validation data. The performance estimate is given from the average of the k validation results.

### **3.10.3 The Overall Holdout**

For the validity of the comparison of generalization estimates for the prestructured and fully connected networks (the comparison that is central to this dissertation), an additional holdout set is needed. These data are set aside prior to any training of neural network, and are in addition to any holdout set discussed in the standard techniques of Holdout Cross-Validation and k-Fold Cross-Validation.

### 3.11 Over-fitting

Learning is not that easy, one has to experiment with choosing the model, architecture and layers. There are many parameters that may cause oscillation in addition to the many training sessions that might be needed which may lead to something called over-fitting. Over-fitting occurs when the error on the training set is driven to a very small value but when new data is presented to the network the error is large. In that case, the network has memorised the training examples but has learned not to generalise to new situations. If over-fitting were to occur early stopping could be implemented. In this thesis, the over-fitting has been considered, and the final built model is tested and no over-fitting has occurred .



# Chapter 4

## Experimental Setup and Results

### 4.1 Experimental Setup

A rectangular cavity resonator was built with dimensions of  $19 \times 28 \times 60$  cm to conduct the experiment and to test whether or not the obtained resonant frequencies of the rectangular cavity could serve as a detection technique to identify and detect any material changes within the cavity.

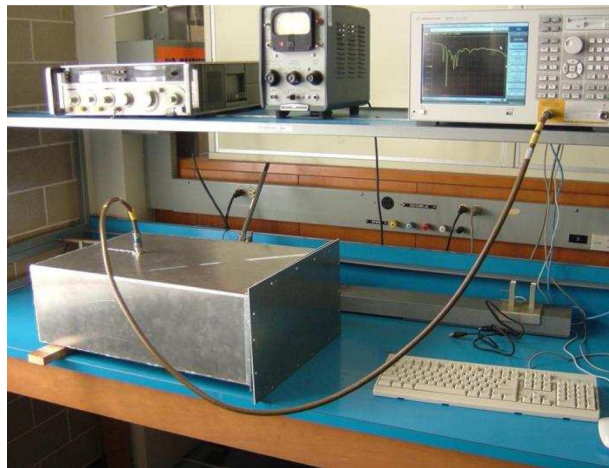


Figure 4.1: Experimental Setup

Also, a monopole antenna was placed near the center of the top face of the rectangular cavity as shown in Fig. 4.1.



Figure 4.2: Experimental Setup

The source of microwave signal and the scattering parameters measurement equipment is an Agilent Vector Network Analyzer (VNA).

In order to find the resonant frequencies, the cavity was excited by the monopole antenna. Then the reflection coefficient  $S_{11}$  was measured over a wide range of frequencies 3 MHz to 2 GHz. The reflection coefficients had been obtained for the different objects inserted inside the cavity. These objects including metallic boxes (each of size  $2 \times 3 \times 5$  cm), wooden boxes (each of size  $2 \times 5 \times 38$  cm), water containers (each of size  $4 \times 7 \times 10$  cm), and boxes of dry corn (each of size  $4 \times 7 \times 10$  cm) are used in the experiment. Figures 4.2 and 4.2 illustrate the conducted experiment Three experimental stages has been conducted. These stages are as follows:

### 1. First Stage:

The main purpose of this stage is to study the effect of changing the position of different objects as a whole on the resonant frequency. In this stage, the four objects (Corn, metal, water, and wood) are inserted into the empty container (one at a time) and the reflection coefficients  $S_{11}$  for each object are recorded. This procedure is repeated for each object in different random places inside the cavity. Then the resonant frequencies are used for the classification problem. Figures 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9 and 4.10 are samples of the obtained resonant frequencies of each object.

### 2. Second Stage:

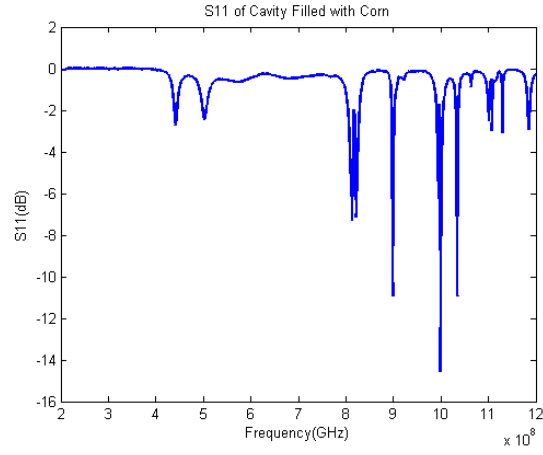


Figure 4.3:  $S_{11}$  of Cavity filled with Corn Boxes

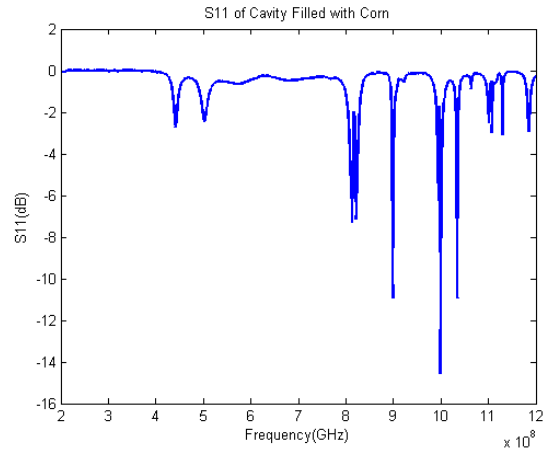


Figure 4.4:  $S_{11}$  of Cavity filled with Corn Boxes

Unlike the first stage which presents the effect of changing the position of the whole object inserted in different positions for different materials, the main purpose in this stage is to distinguish between changes in resonant frequencies due to the change in the position of the boxes and the materials inserted in the cavity. In this stage, a metal sheet of size  $9 \times 5 \times 3$  cm is inserted in many different places in the cavity and the resonant frequencies are measured. Then the same procedure is repeated for the corn, water, empty boxes (one at a time). Figures 4.11, 4.12, 4.13, 4.14, 4.15, and 4.16 are samples of the obtained resonant frequencies of each object

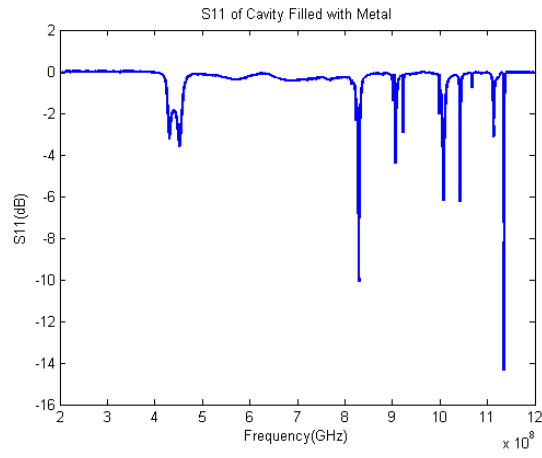


Figure 4.5:  $S_{11}$  of Cavity filled with Corn Boxes

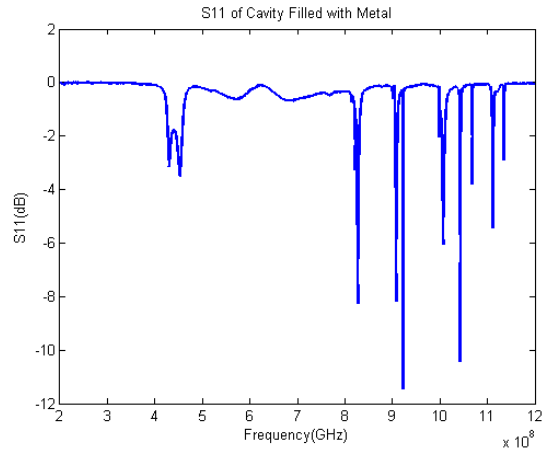


Figure 4.6:  $S_{11}$  of Cavity filled with Metal

### 3. Third Stage:

In this stage a comparison between four different cases was conducted. These cases are: the cavity full of boxes of corn, the cavity with 1/4 of the corn boxes taken out, 3/4 corn filled cavity with a metal sheet in a random place, and a the 3/4 corn filled cavity with 1/4 boxes of water at a random place. Sample of the obtained resonant frequencies are shown in figures 4.17, 4.18, 4.19, 4.20, and 4.21

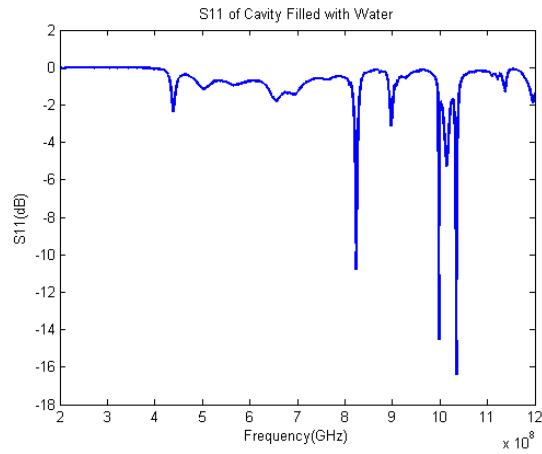


Figure 4.7:  $S_{11}$  of Cavity filled with Water Boxes

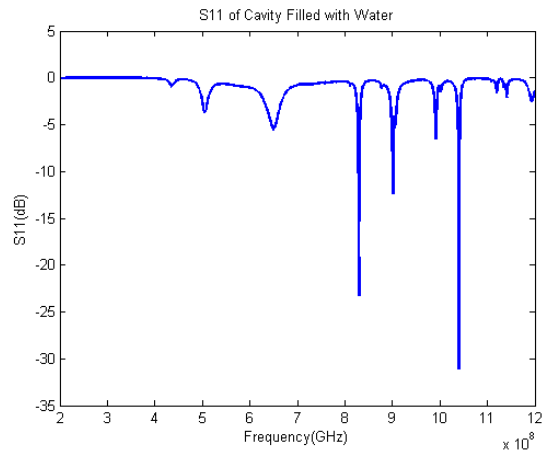


Figure 4.8:  $S_{11}$  of Cavity filled with Water Boxes

## 4.2 Pre-Processing Algorithm and Feature Selections

A pre-processing algorithm was applied to the obtained experimental data before using them as input to the Neural Network. This pre-process is required in order to autonomously detect the resonant frequencies of the reflection and eliminate the noisy data. The pre-processing algorithms are important because the effectiveness of the Neural Network is contingent on the proper selection of a characteristic set of data of small dimensionality [49].

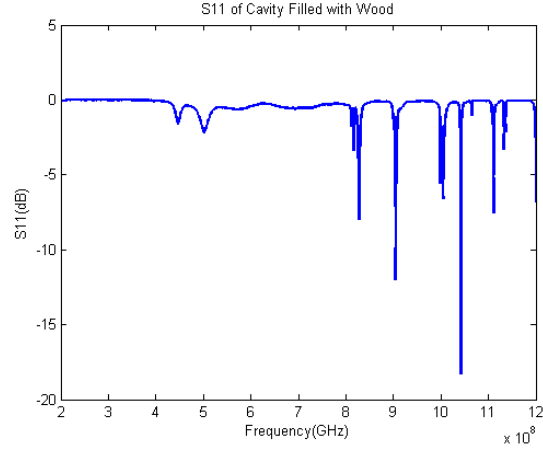


Figure 4.9:  $S_{11}$  of Cavity filled with Wood Sheets

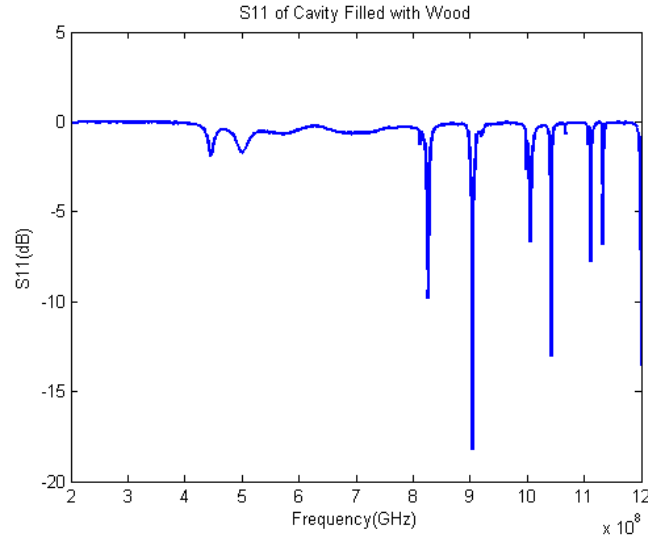


Figure 4.10:  $S_{11}$  of Cavity filled with Wood Sheets

In this work, the Peak Interpolation with Sigma Filter algorithm is used. This algorithm prepares the Neural Network input by performing the following steps:

1. In the step the data is half-sampled by selecting the maximum value of each data pair. Mathematically,  $S_{11,1}(f_j) = \max(S_{11,0}(f_{2j}), S_{11,0}(f_{2j+1}))$  where  $j = 0, 1, 2, \dots, (J - 1)/2$ .

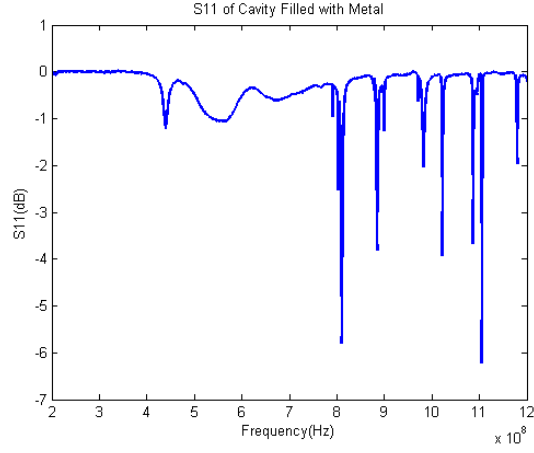


Figure 4.11:  $S_{11}$  of Cavity filled with Metal Sheets

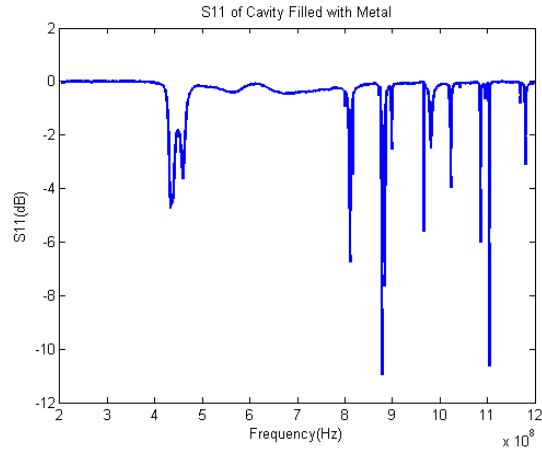


Figure 4.12:  $S_{11}$  of Cavity filled with Metal Sheets

2. Zero out values lower than the absolute average of the data set. Mathematically,  $a = 2/(J - 1) \times \sum_{j=0}^{(J-1)/2} |S_{11,1}(f_j)|$ . Then  $S_{11,2}(f_j) = S_{11,1}(f_j) < a ? 0.0 : S_{11,1}(f_j)$  where  $j = 0, 1, 2, \dots, (J - 1)/2$ .
3. Select the peak values of each section in the data set, where a section is defined as the domain enclosed by two zero values with no zero within the domain.
4. At the location of each  $M$  peaks  $p_m, n = 0, 1, 2, \dots, M - 1$  on the frequency axis, place an exponential function of the form  $g_n(f_j) = S_{11,3}(f_j) \times \exp[-k_1(|m - j| + k_2)]$  for

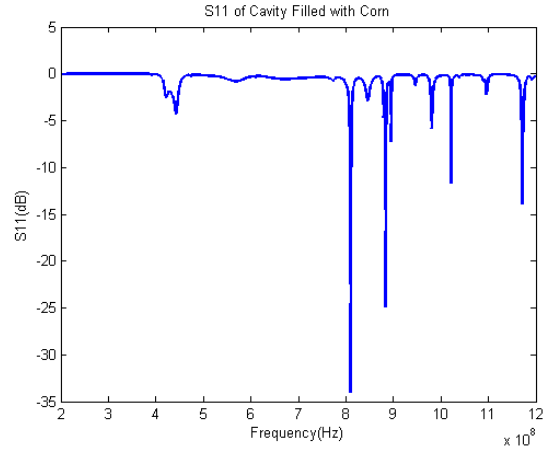


Figure 4.13:  $S_{11}$  of Cavity filled with Corn Sheets

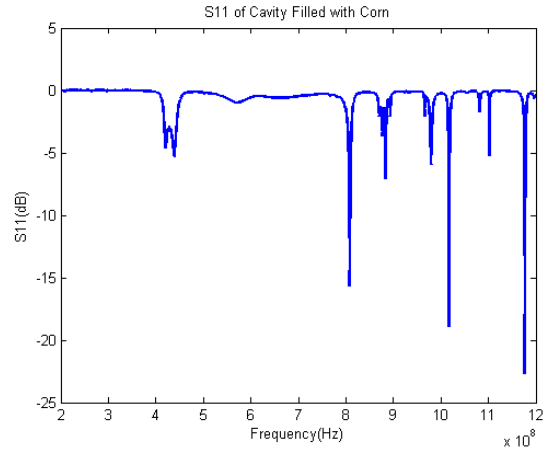


Figure 4.14:  $S_{11}$  of Cavity filled with Corn Sheets

some constants  $k_i$  and find the summation of all  $g_n$  across each  $j = 0, 1, 2, \dots, (J-1)/2$ .

5. Apply a sigma filter sufficiently until data jitter has dissipated.
6. Select the resonant frequencies by locating the local maximum by comparing subsequent data samples.

In terms of notation, the reflection parameter data are  $S_{11,i}(f_j)$  where  $j = 0, 1, 2, \dots, n-1$ , where  $f_j$  is the frequency at sample  $j$  and  $S_{11,i}$  is the reflection parameter after step  $i$ .



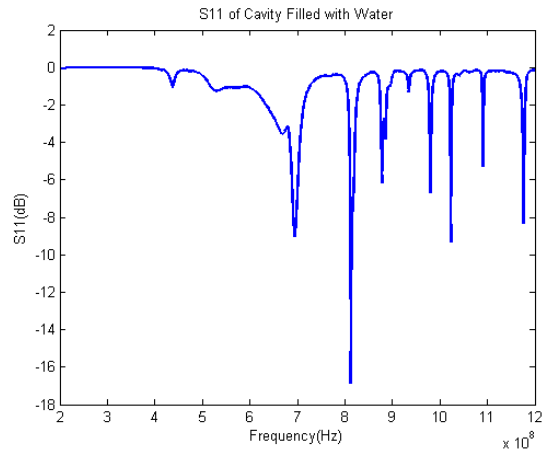


Figure 4.15:  $S_{11}$  of Cavity filled with Water Sheets

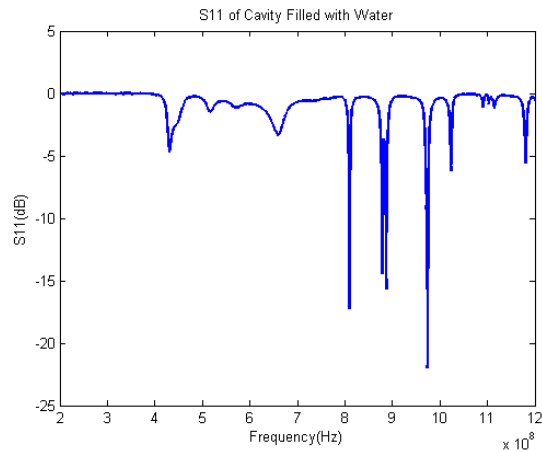


Figure 4.16:  $S_{11}$  of Cavity filled with Water Sheets

The following figures (Figures 4.26, 4.24, 4.25, and 4.23) are shown the selected frequency points for some of samples after applying the filter algorithm.

After applying the above filter algorithm, a sequential forward and backward selection were then applied to the data obtained from the filter. This algorithm beside the filter one eliminate a lot of the noise in the experimental data.

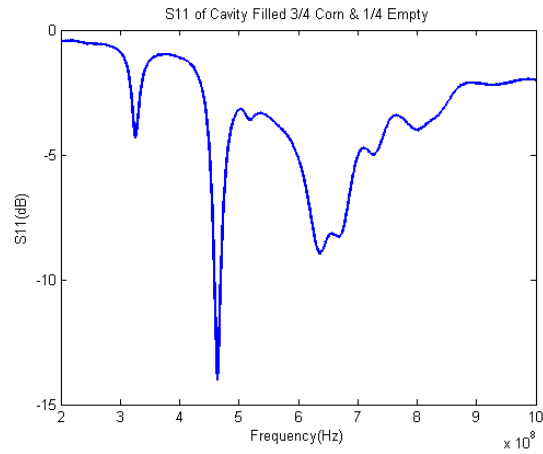


Figure 4.17:  $S_{11}$  of Cavity filled with 3/4 Corn Boxes and 1/4 empty

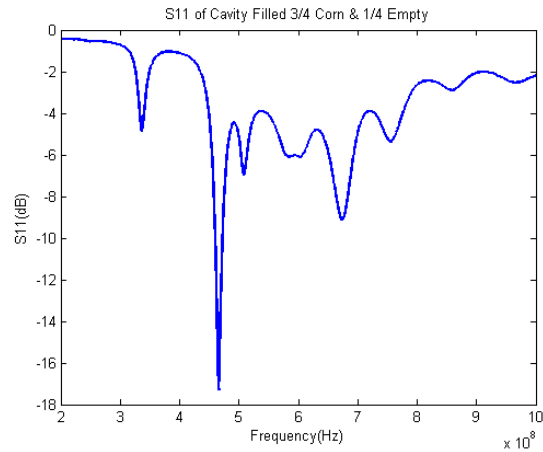


Figure 4.18:  $S_{11}$  of Cavity filled with 3/4 Corn Boxes and 1/4 empty

### 4.3 Neural Networks Analysis

After selecting the most impacting features, a Neural Networks classifier was then applied to the selected resonant frequencies of the experimental data from the previous section in order to detect small changes in the resonant frequencies. The feed-forward neural network was used and trained through the back-propagation with the gradient descent learning algorithm. In order to select the number of internal nodes and the right activation function, many tests were conducted. The convergence and classification rate of the neural

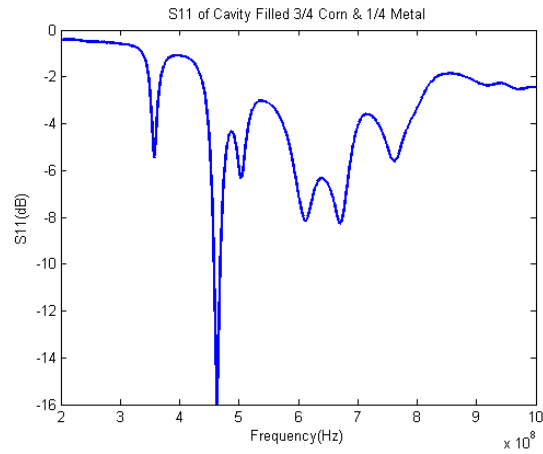


Figure 4.19:  $S_{11}$  of Cavity filled with 3/4 Corn Boxes and 1/4 Metal

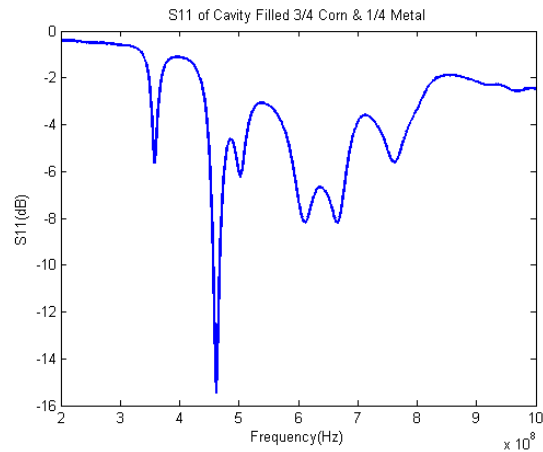


Figure 4.20:  $S_{11}$  of Cavity filled with 3/4 Corn Boxes and 1/4 Metal

network while varying these parameters are obtained thus an optimal choice was made.

The overall accuracy of the built neural network models is not as sufficient as needed. This is because of the limited number of samples for each object which will also have a great impact on validating the models. Also, another preprocessing and more advanced feature selections techniques may consider in the future work in order to get more accurate models.

The following are the neural networks specifications and results for the three proposed

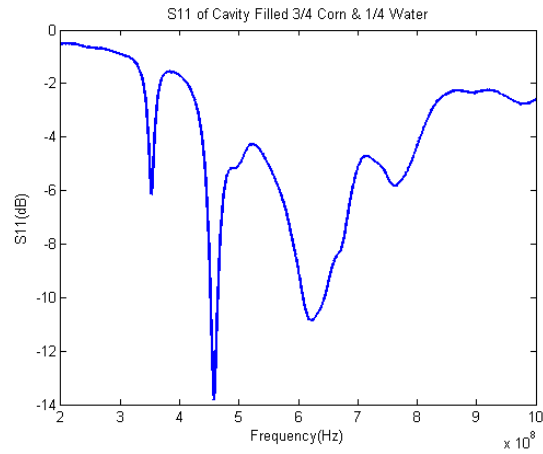


Figure 4.21:  $S_{11}$  of Cavity filled with 3/4 Corn Boxes and 1/4 Water

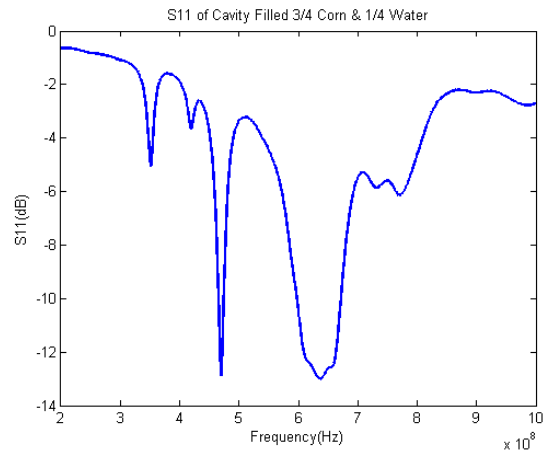


Figure 4.22:  $S_{11}$  of Cavity filled with 3/4 Corn Boxes and 1/4 Water

stages:

1. **First Stage:** Table 1 summarizes the Neural Network specifications for the this stage while table 1 shows the obtained results.

Figure shows the confusion matrix which shows the percentages of correct and incorrect classifications. Correct classifications are the green squares on the matrices diagonal. Incorrect classifications form the red squares. If the network has learned to classify properly, the percentages in the red squares should be very small, indicating

Number of hidden layers	1
Number of neurons in the input layer	6
Number of neurons in the hidden layer	10
Number of neurons in the output layer	2
Hidden Layer Activation Functions	Tan-Sigmoid
Output Layer Activation Functions	Tan-Sigmoid
Learning Rate	0.06
Momentum Rate	0.8
Maximum Number of Epochs	10000
Mean Squared Error goal	0.04

Table 4.1: First Stage: Neural Networks Specification

Percentage Correct Classification on training dataset	73.75%
Percentage Incorrect Classification on training dataset	26.25%
Percentage Correct Classification on testing dataset	75%
Percentage Incorrect Classification on testing dataset	25%
The Mean Recognized of Training Data	80
Mean Train Accuracy	100%
The Mean Recognized of Testing Data	20
Mean Test Accuracy	100%

Table 4.2: First Stage: Neural Networks Results

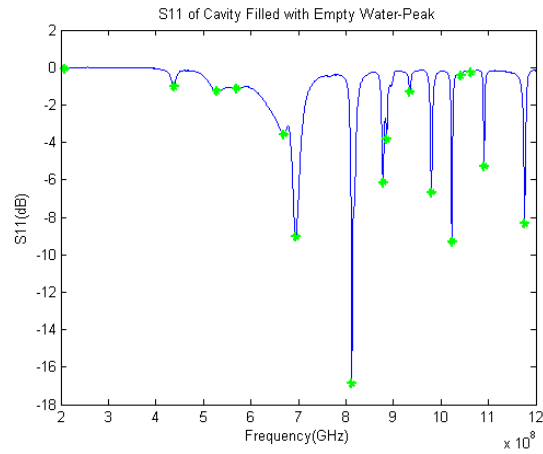


Figure 4.23: Peaks Selected of  $S_{11}$  of Cavity filled with Water Sheets

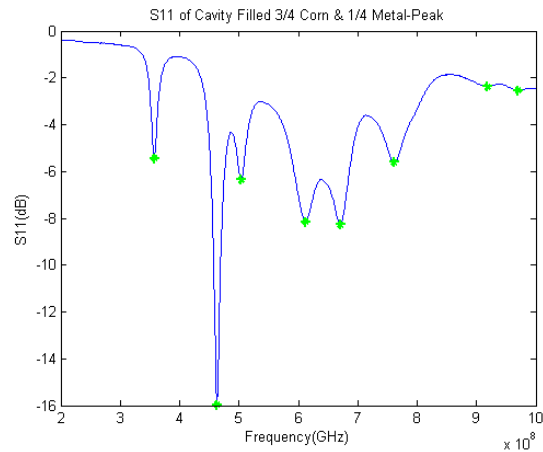


Figure 4.24: Peaks Selected of  $S_{11}$  of Cavity filled with 3/4 Corn Boxes and 1/4 Metal

few miss-classifications. Figures 4.27 and 4.36 show the confusion matrices of the training and testing dataset obtained in this stage.

Another measure of how well the neural network has fit data is the receiver operating characteristic (ROC) plot. This shows how the false positive and true positive rates relate as the thresholding of outputs is varied from 0 to 1. The farther left and up the line is, the fewer false positives need to be accepted in order to get a high true positive rate. The best classifiers will have a line going from the bottom left corner,

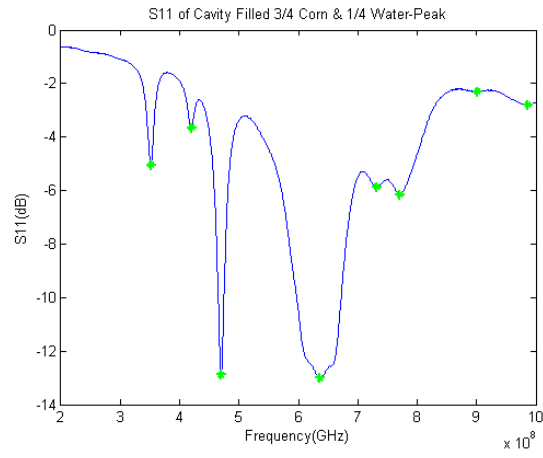


Figure 4.25: Peaks Selected of  $S_{11}$  of Cavity filled with 3/4 Corn Boxes and 1/4 Water

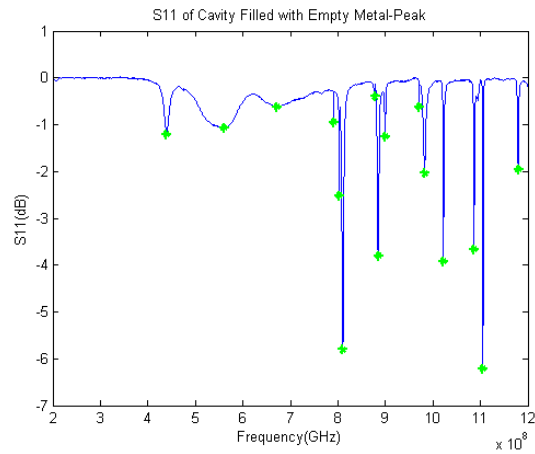


Figure 4.26: Peaks Selected of  $S_{11}$  of Cavity filled with Metal Sheets

to the top left corner, to the top right corner, or close to that.

Figures 4.27 and 4.30 show the ROC plots of the training and testing dataset obtained in this stage

2. **Second Stage** Table 2 summarizes the Neural Network specifications for the this stage while table 2 shows the obtained results.

Figures 4.31 and 4.32 show the confusion matrices of the training and testing dataset obtained in this stage

Number of hidden layers	1
Number of neurons in the input layer	6
Number of neurons in the hidden layer	20
Number of neurons in the output layer	2
Hidden Layer Activation Functions	Tan-Sigmoid
Output Layer Activation Functions	Tan-Sigmoid
Learning Rate	0.06
Momentum Rate	0.6
Maximum Number of Epochs	10000
Mean Squared Error goal	0.04

Table 4.3: Second Stage: Neural Networks Specification

Percentage Correct Classification-training dataset	81.067961%
Percentage Incorrect Classification-training dataset	18.932039%
Percentage Correct Classification-testing dataset	84.067961%
Percentage Incorrect Classification-testing dataset	15.932039%
The Mean Recognized of Training Data	408
Mean Train Accuracy	99.0291%
The Mean Recognized of Testing Data	102
Mean Test Accuracy	99.0291%

Table 4.4: Second Stage: Neural Networks Results



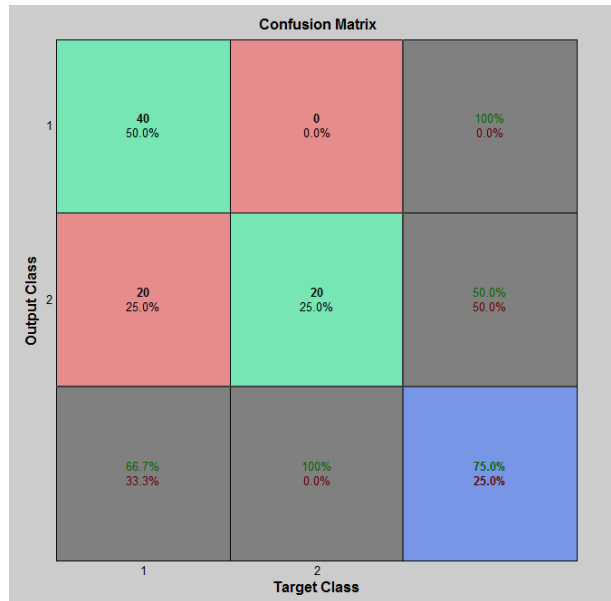


Figure 4.27: First Stage: Confusion Matrix of Training Dataset

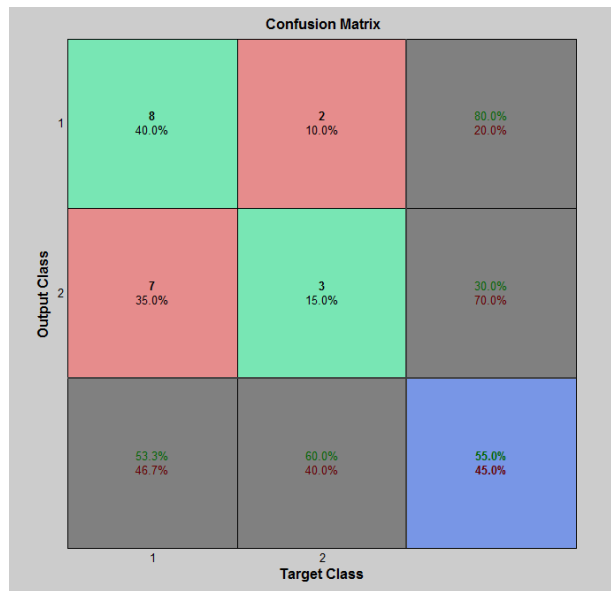


Figure 4.28: First Stage: Confusion Matrix of Testing Dataset

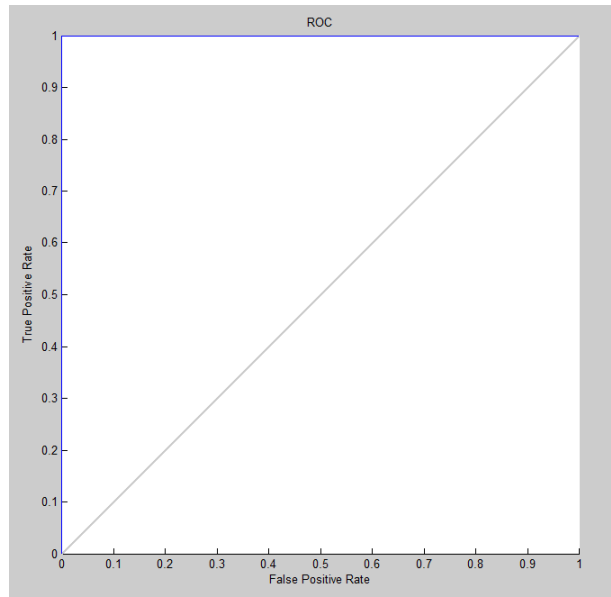


Figure 4.29: First Stage: ROC of Training Dataset

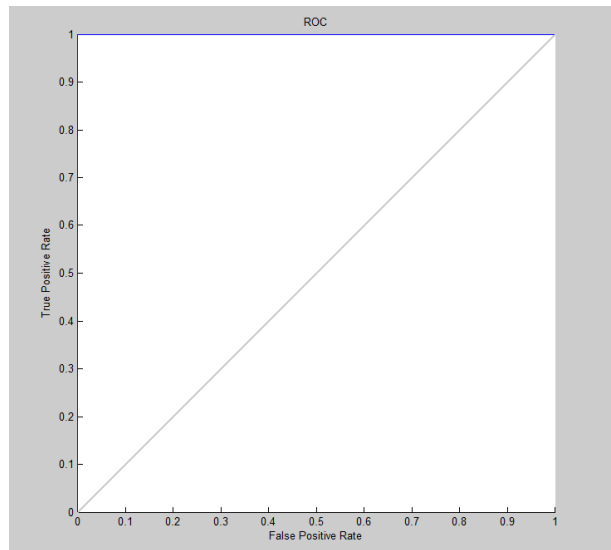


Figure 4.30: First Stage: ROC of Testing Dataset

Figures 4.33 and 4.34 show the ROC plots of the training and testing dataset obtained in this stage.

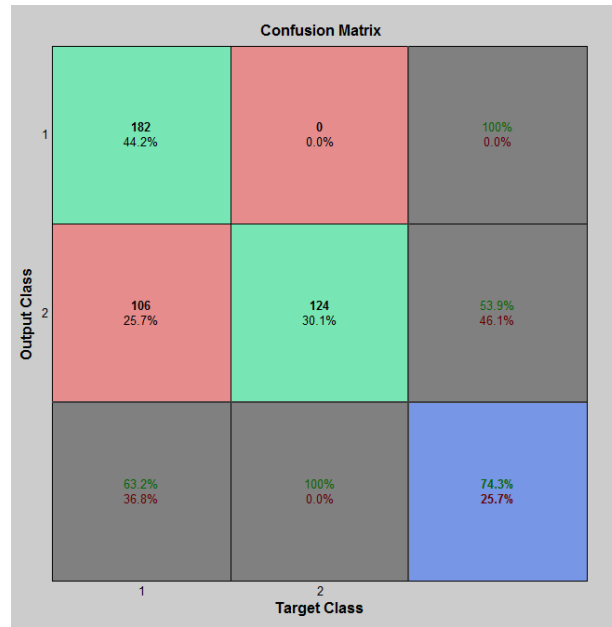


Figure 4.31: Second Stage: Confusion Matrix of Training Dataset

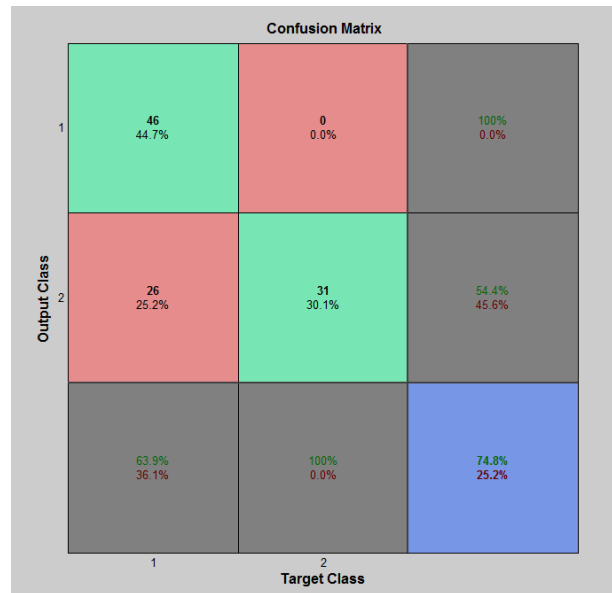


Figure 4.32: Second Stage: Confusion Matrix of Testing Dataset

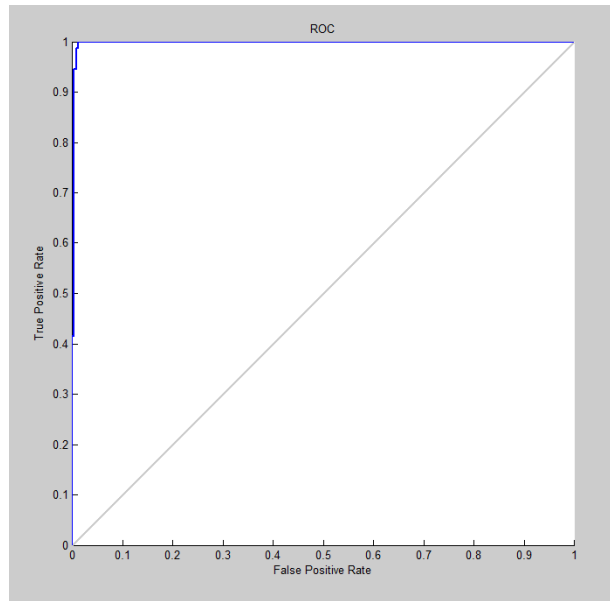


Figure 4.33: Second Stage: ROC Matrix of Training Dataset

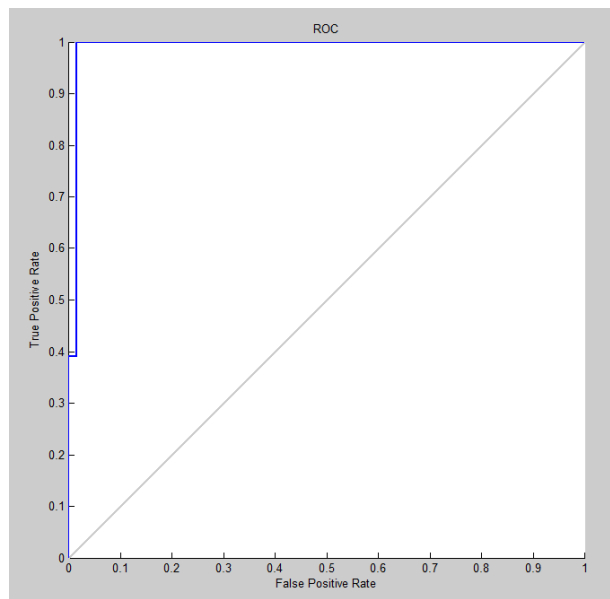


Figure 4.34: Second Stage: ROC Matrix of Testing Dataset

Number of hidden layers	1
Number of neurons in the input layer	6
Number of neurons in the hidden layer	10
Number of neurons in the output layer	2
Hidden Layer Activation Functions	Tan-Sigmoid
Output Layer Activation Functions	Tan-Sigmoid
Learning Rate	0.04
Momentum Rate	0.6
Maximum Number of Epochs	10000
Mean Squared Error goal	0.04

Table 4.5: Third Stage: Neural Networks Specification

Percentage Correct Classification on training dataset	75%
Percentage Incorrect Classification on training dataset	25%
Percentage Correct Classification on testing dataset	78%
Percentage Incorrect Classification on testing dataset	32%
The Mean Recognized of Training Data	80
Mean Train Accuracy	100%
The Mean Recognized of Testing Data	20
Mean Test Accuracy	100%

Table 4.6: Third Stage: Neural Networks Results

### 3. Third Stage:

Table 3 summarizes the Neural Network specifications for the this stage while table 3 shows the obtained results.

Figures 4.35 and ?? show the confusion matrices of the training and testing dataset obtained in this stage.

Figures 4.35 and 4.38 show the ROC plots of the training and testing dataset obtained in this stage

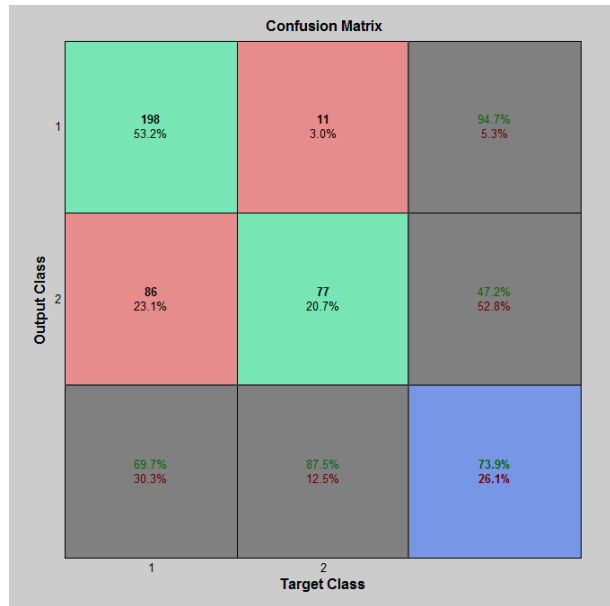


Figure 4.35: Third Stage: Confusion Matrix of Training Dataset

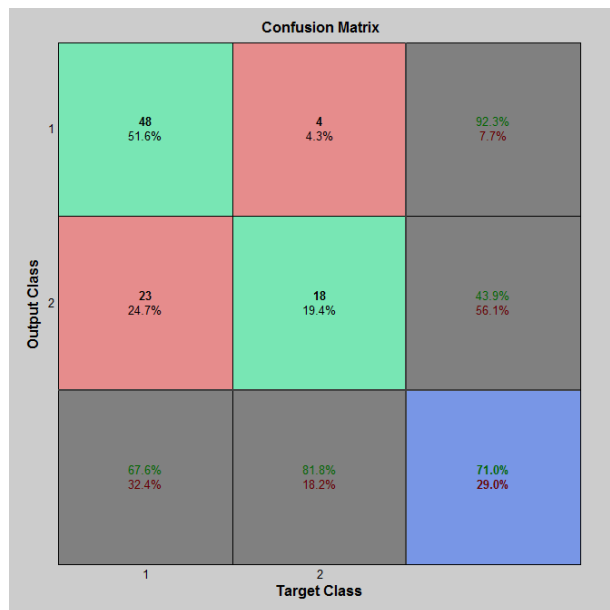


Figure 4.36: Third Stage: Confusion Matrix of Testing Dataset

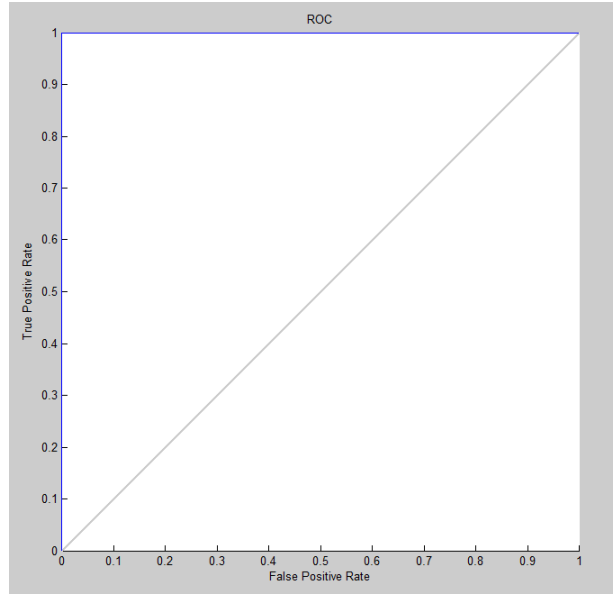


Figure 4.37: Third Stage: ROC of Training Dataset

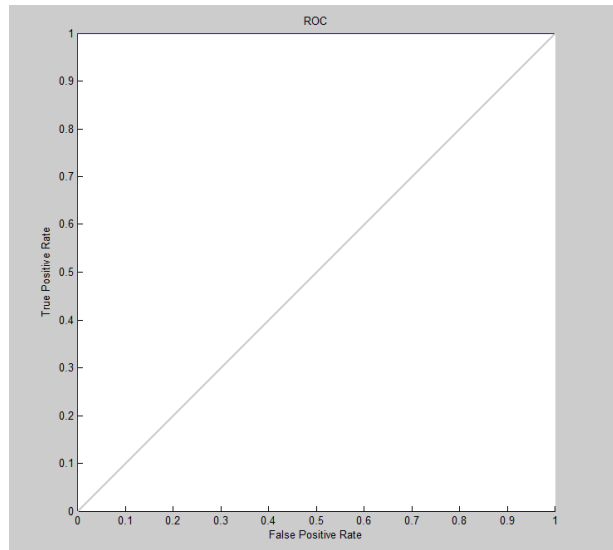


Figure 4.38: Third Stage: ROC of Testing Dataset

# Chapter 5

## Conclusion and Future Work Directions

This thesis has examined the rectangular cavity perturbation technique as it used for materials detection with the use of Artificial Neural Networks. The proposed method to determine a change in the material composition of a cavity resonator was able to classify different changes in the cavity material using changes in the resonant frequencies of the cavity and applying a Neural Network classifier. The performance of this method was experimentally tested for different material changes inside the cavity. The advantage of this method over other commonly used methods is in its ability to detect changes even when the material to be detected is not in the line of sight of the radiation source.

The 100% classification rate from human intervention suggests that future work should include optimization of the preprocessing algorithms and applying different implementations of Neural Networks. Other future work will include applying these techniques to more complex cavities, incorporating partial knowledge of the resonant frequency behaviour of the system into the model, and applying the same technique using two or more antennas in different locations around the cavity. Also, use multiple classifiers technique as a decision maker instead of only using one classifier. Future work could also include application to the biomedical field; for example, the detection of cancerous tissues.



# References

- [1] A.W. Ham and DH Cormack. Textbook of histology. *Ch*, 28:903–940, 1974.
- [2] AD Vyas, VA Rana, DH Gadani, and AN Prajapati. Cavity perturbation technique for complex permittivity measurement of dielectric materials at x-band microwave frequency. In *Recent Advances in Microwave Theory and Applications, 2008. MICROWAVE 2008. International Conference on*, pages 836–838. IEEE, 2008.
- [3] W. Che, Z. Wang, Y. Chang, and P. Russer. Permittivity measurement of biological materials with improved microwave cavity perturbation technique. In *Microwave Conference, 2008. EuMC 2008. 38th European*, pages 905–908. IEEE, 2008.
- [4] Q. Wang, X. Deng, M. Yang, Y. Fan, and W. Wang. Measuring glucose concentration by microwave cavity perturbation and dsp technology. In *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on*, volume 3, pages 943–946. IEEE, 2010.
- [5] A. Sklyuyev, M. Ciureanu, C. Akyel, P. Ciureanu, D. Menard, and A. Yelon. Measurement of complex permeability of ferromagnetic nanowires using cavity perturbation techniques. In *Electrical and Computer Engineering, 2006. CCECE'06. Canadian Conference on*, pages 1486–1489. IEEE, 2006.
- [6] ZH Wang, HHS Javadi, and AJ Epstein. Application of microwave cavity perturbation techniques in conducting polymers. In *Instrumentation and Measurement Technology Conference, 1991. IMTC-91. Conference Record., 8th IEEE*, pages 79–82. IEEE, 1991.
- [7] O. Ogunlade, Y. Chen, and P. Kosmas. Measurement of the complex permittivity of microbubbles using a cavity perturbation technique for contrast enhanced ultra-wideband breast cancer detection. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 6733–6736. IEEE, 2010.

- [8] R. Franklin and U.B. Halabe. Knowledge-based assistant for ultrasonic inspection in metals. In *Proceedings of SPIE*, volume 3203, page 44, 1997.
- [9] V.K. Devabhaktuni, M.C.E. Yagoub, Y. Fang, J. Xu, and Q.J. Zhang. Neural networks for microwave modeling: Model development issues and nonlinear modeling techniques. *International Journal of RF and Microwave Computer Aided Engineering*, 11:4–21, 2001.
- [10] PM Watson, KC Gupta, and RL Mahajan. Development of knowledge based artificial neural network models for microwave components. In *Microwave Symposium Digest, 1998 IEEE MTT-S International*, volume 1, pages 9–12. IEEE, 1998.
- [11] V.K. Devabhaktuni, M.C.E. Yagoub, and Q.J. Zhang. A robust algorithm for automatic development of neural-network models for microwave applications. *Microwave Theory and Techniques, IEEE Transactions on*, 49(12):2282–2291, 2001.
- [12] Z.Ž. Stanković, B. Milovanović, and N. Dončov. Neural model of microwave cylindrical cavity loaded with arbitrary-raised dielectric slab. *International Journal of RF and Microwave Computer-Aided Engineering*, 19(3):317–327, 2009.
- [13] F. Wang, V.K. Devabhaktuni, C. Xi, and Q.J. Zhang. Neural network structures and training algorithms for rf and microwave applications. *International Journal of RF and Microwave Computer-Aided Engineering*, 9(3):216–240, 1999.
- [14] Y. Kim, S. Keely, J. Ghosh, and H. Ling. Application of artificial neural networks to broadband antenna design based on a parametric frequency model. *Antennas and Propagation, IEEE Transactions on*, 55(3):669–674, 2007.
- [15] I. Ratner, H. Ali, and E.M. Petriu. Neural network simulation of a dielectric ring resonator antenna. *Journal of systems architecture*, 44(8):569–581, 1998.
- [16] H.J. Delgado, M.H. Thursby, and F.M. Ham. A novel neural network for the synthesis of antennas and microwave devices. *Neural Networks, IEEE Transactions on*, 16(6):1590–1600, 2005.
- [17] S. Selleri, S. Manetti, and G. Pelosi. Neural network applications in microwave device design. *International Journal of RF and Microwave Computer-Aided Engineering*, 12(1):90–97, 2002.
- [18] A. Patnaik and RK Mishra. Ann techniques in microwave engineering. *Microwave Magazine, IEEE*, 1(1):55–60, 2000.

- [19] Z. Su and L. Ye. An intelligent signal processing and pattern recognition technique for defect identification using an active sensor network. *Smart materials and structures*, 13:957, 2004.
- [20] L. Wang, L. Li, W. Qi, and H. Yang. Pattern recognition and size determination of internal wood defects based on wavelet neural networks. *Computers and Electronics in Agriculture*, 69(2):142–148, 2009.
- [21] Q.J. Zhang, K.C. Gupta, and V.K. Devabhaktuni. Artificial neural networks for rf and microwave design-from theory to practice. *Microwave Theory and Techniques, IEEE Transactions on*, 51(4):1339–1350, 2003.
- [22] F.M. Kahnert. Numerical methods in electromagnetic scattering theory. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 79:775–824, 2003.
- [23] Z.Z. Stankovic, B Milovanovic, and N Doncov. Neural model of microwave cylindrical cavity loaded with arbitrary-raised dielectric slab. *International Journal of RF and Microwave Computer-Aided Engineering*, 19:317–327, 2009.
- [24] D.A. Woten, J. Lusth, and M. El-Shenawee. Interpreting artificial neural networks for microwave detection of breast cancer. *Microwave and Wireless Components Letters, IEEE*, 17(12):825–827, 2007.
- [25] F.L. Penaranda Foix, J.M. Catala-Civera, A.J. Canos-Marin, and B. Garcia-Banos. Practical expression for frequency deviation due to insertion hole in cylindrical cavity. *IEEE Transactions on Instrumentation and Measurement*, 43:1338–1340, November 2007.
- [26] A.V. Farahani and A. Konrad. Fdtd calculation of cavity resonant frequencies in case of nonuniform internal magnetic field distribution. *IEEE Transactions on Magnetics*, 43:1517–1520, 2007.
- [27] D.A. Hill. *Electromagnetic Fields in Cavities: Deterministic and Statistical Theories*. Wiley IEEE Press, 2009.
- [28] U. Raveendranath and K. T. Mathew. New cavity perturbation technique for measuring complex permeability of ferrite materials. *Microwave and Optical Technology Letters*, pages 241–243, July 1998.
- [29] J. F. Ma, W. H. Yu, and R. Mittra. Detection of buried dielectric cavities using the finite-difference time-domain method in conjunction with signal processing techniques. *IEEE Transactions on Antennas and Propagation*, pages 1289–1294, September 2000.

- [30] S. Kharkovsky and R. Zoughi. Microwave and millimeter wave nondestructive testing and evaluation-overview and recent advances. *Instrumentation & Measurement Magazine, IEEE*, 10(2):26–38, 2007.
- [31] H.A. Bethe and J.S. Schwinger. Perturbation theory for cavities, 1943.
- [32] K. Folgerø. Bilinear calibration of coaxial transmission/reflection cells for permittivity measurement of low-loss liquids. *Measurement Science and Technology*, 7:1260, 1996.
- [33] V.V. Varadan, R.D. Hollinger, D.K. Ghodgaonkar, and V.K. Varadan. Free-space, broadband measurements of high-temperature, complex dielectric properties at microwave frequencies. *Instrumentation and Measurement, IEEE Transactions on*, 40(5):842–846, 1991.
- [34] U.C. Hasar. A fast and accurate amplitude-only transmission-reflection method for complex permittivity determination of lossy materials. *Microwave Theory and Techniques, IEEE Transactions on*, 56(9):2129–2135, 2008.
- [35] M.N. Afsar and K.J. Button. Millimeter-wave dielectric measurement of materials. *Proceedings of the IEEE*, 73(1):131–153, 1985.
- [36] R.F. Harrington. *Time Harmonic Electromagnetic Fields*. Wiley IEEE Press, 2001.
- [37] Ral Rojas. *Neural networks : a systematic introduction*. Springer-Verlag, Berlin ; New York, 1996.
- [38] K.J. Hunt, D. Sbarbaro, R. Żbikowski, and PJ Gawthrop. Neural networks for control systemsa survey. *Automatica*, 28(6):1083–1112, 1992.
- [39] H.X. Li and C.L.P. Chen. The equivalence between fuzzy logic systems and feedforward neural networks. *Neural Networks, IEEE Transactions on*, 11(2):356–365, 2000.
- [40] B. Igel'nik and N. Parikh. Kolmogorov's spline network. *Neural Networks, IEEE Transactions on*, 14(4):725–733, 2003.
- [41] L.H. Tsoukalas and R.E. Uhrig. *Fuzzy and neural approaches in engineering*. John Wiley & Sons, Inc., 1996.
- [42] Simon S. Haykin. *Neural networks : a comprehensive foundation*. New York : Macmillan ; Toronto : Maxwell Macmillan Canada, 1994.

- [43] J.C. Principe, N.R. Euliano, and W.C. Lefebvre. *Neural and adaptive systems: fundamentals through simulations*. Wiley New York, 2000.
- [44] A.K. Jain, J. Mao, and K.M. Mohiuddin. Artificial neural networks: A tutorial. *Computer*, 29(3):31–44, 1996.
- [45] R. Battiti. Using mutual information for selecting features in supervised neural net learning. *Neural Networks, IEEE Transactions on*, 5(4):537–550, 1994.
- [46] L.M. Belue, K.W. Bauer, et al. Determining input features for multilayer perceptrons. *Neurocomputing*, 7(2):111–121, 1995.
- [47] I.V. Tetko, A.E.P. Villa, and D.J. Livingstone. Neural network studies. 2. variable selection. *Journal of chemical information and computer sciences*, 36(4):794–803, 1996.
- [48] J. Utans, J. Moody, S. Rehfuss, and H. Siegelmann. Input variable selection for neural networks: Application to predicting the us business cycle. In *Computational Intelligence for Financial Engineering, 1995., Proceedings of the IEEE/IAFE 1995*, pages 118–122. IEEE, 1995.
- [49] C.Y. Suen. *Pattern Recognition*. Elsevier Inc., 2009.
- [50] G. Acciani, G. Brunetti, G. Fornarelli, and A. Giaquinto. Angular and axial evaluation of superficial defects on non-accessible pipes by wavelet transform and neural network-based classification. *Ultrasonics*, 50(1):13–25, 2010.
- [51] L.G. Esteban, F.G. Fernández, and P. de Palacios. Moisture prediction in pine timber: Application of an artificial neural network using non-destructive testing. *Computers & Structures*, 87(21):1360–1365, 2009.
- [52] T. D’Orazio, C. Guaragnella, M. Leo, and P. Spagnolo. Defect detection in aircraft composites by using a neural approach in the analysis of thermographic images. *NDT & E International*, 38(8):665–673, 2005.
- [53] F.M. Kahnert. Numerical methods in electromagnetic scattering theory. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 80:755–824, August 2002.
- [54] L. Merad, F.T. Bendimerad, S.M. Meriah, and S.A. Djennas. Neural networks for synthesis and optimization of antenna arrays. *RADIOENGINEERING-PRAGUE*, 16(1):23, 2007.

- [55] M. H. Beale, M. T. Hagan, and H. B. Demuth. *Neural Network Toolbox MATLAB User's Guide*. The MathWorks, Inc, 1992 to 2012.
- [56] R. Venkatesan and B. Balamurugan. A real-time hardware fault detector using an artificial neural network for distance protection. *Power Delivery, IEEE Transactions on*, 16(1):75–82, 2001.