

Quantitative Analysis of Infrared Images
For Early Breast Cancer Detection

by

Jessica Phaik Sun Koay

A thesis submitted to

The Faculty of Graduate Studies and Research

In partial fulfilment of the requirements for the degree of

Master of Applied Science

in Electrical Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering

Department of Systems and Computer Engineering

Carleton University

Ottawa, Ontario, Canada

December 2004

©Jessica Phaik Sun Koay, 2004



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*

ISBN: 0-494-00748-6

Our file *Notre référence*

ISBN: 0-494-00748-6

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

To my parents, for always supporting me in my decisions.

To my two sisters, for being my best friends.

And

To Dean, for endless reasons.

Acknowledgements

I would like to begin by thanking my thesis supervisor, Dr. Monique Frize for her ongoing guidance and support, as well as giving me the opportunity to be part of this very exciting research group. Her enthusiasm and energy is truly inspiring!

Thank you as well to Christophe Herry for valuable suggestions and answering my many, many questions.

I would also like to thank

Dr. Frize for providing the images from Moncton, and Dr. R. Roberge for providing expertise, patient base, and clinical results for these images.

Dr. Keyserlingk from Ville Marie Breast and Oncology Center for providing images from Montreal.

Mr. Thomas DiCicco from Infrared Science Corporation for providing the images from New York.

And finally, a big thank you to Dean Melvin for proof-reading my paper. Surely it was very exciting!

Abstract

This thesis presents a quantitative method to analyze and evaluate infrared images to improve early breast cancer detection. Researchers have demonstrated that breast thermography, or infrared imaging, has great potential in early cancer detection and prognosis indication. However, manual examinations of the thermograms for abnormal hyperthermia and hyper-vascularity related to tumour growth can be highly subjective, tedious and challenging if the differences between the normal and abnormal images are subtle. To address the problem, this thesis proposes a method whereby statistical analysis and image processing techniques are applied to extract thermal features that may be relevant in discriminating the images. Specifically, algorithms were developed to extract quantitative parameters to represent the vascular patterns of the thermal images. An Artificial Neural Network is then employed to predict clinical outcomes based on the extracted parameters and to identify the most relevant set of features for image classification.

Contents

<i>Acknowledgements</i>	<i>iv</i>
<i>Abstract</i>	<i>v</i>
<i>Contents</i>	<i>vi</i>
<i>List of Figures</i>	<i>ix</i>
<i>List of Tables</i>	<i>x</i>
1 Introduction	1
1.1 Motivation	1
1.2 Thesis Objective and Problem Statement	3
1.3 Thesis Layout	5
2 Background Information	6
2.1 Sensitivity and Specificity of Clinical Trials	7
2.2 Mammography	9
2.2.1 Sensitivity and Specificity of Mammography	9
2.2.2 Effectiveness of Mammography in Reducing Breast Cancer Mortality	10
2.2.3 Risks of Mammography	11
2.3 Imaging Modalities	12
2.4 Medical Thermography	14
2.4.1 Historical Background	15
2.4.2 Physics of Medical Thermography	16
2.4.3 Breast Thermography	20
2.4.4 Imaging Technology in Medical Thermography	24
2.5 Artificial Neural Networks (ANN)	27
2.5.1 Training the ANN	28

2.5.2	ANN for Pattern Classifications	31
2.5.3	Medical Applications for Pattern Classification	32
3	<i>Review of Existing Methods and Approaches</i>	33
3.1	Analysis of Infrared Images	34
3.1.1	Spatial Characteristics	35
3.1.2	Temporal Characteristics	39
3.1.3	Image Segmentation Techniques	41
3.2	Artificial Neural Networks (ANN)	43
3.2.1	ANN in breast cancer detection and diagnosis	43
3.2.2	ANNs in Breast Thermography	44
3.3	Critical Analysis of Existing Techniques and Approaches	46
4	<i>Infrared Image Analysis for Breast Cancer Detection</i>	49
4.1	Problem Statement	49
4.2	Research Data	50
4.2.1	Moncton Images	51
4.2.2	Montreal Images	52
4.2.3	New York Images	54
4.3	Software Tools	55
4.4	Image Pre-processing	55
4.4.1	Image Conversion, Cropping and Resizing	56
4.4.2	Image Enhancement	58
4.5	Image Processing	60
4.5.1	Identifying Region of Interest	60
4.6	Feature Extraction	62
4.6.1	Statistical Analysis	62
4.6.2	Quadtree decomposition analysis	64

4.6.2.1	Quadtree Decomposition Implementation in MATLAB	65
4.6.2.2	Algorithm for Extracting Quadtree Decomposition Parameter	67
4.6.2.3	Threshold Selection	69
4.6.3	Edge detection analysis	69
4.6.3.1	Algorithm to Extract Edge Detection Ratio	70
4.6.3.2	Selection of Edge Detection Thresholds	71
4.7	Feature Analysis	72
4.7.1	Architecture of ANN	72
4.7.1.1	Backpropagation Neural Networks	72
4.7.1.2	Selection of Input Nodes	74
4.7.1.3	Selection of Hidden Nodes	75
4.7.2	Performance Measures	76
5	Results and Discussions	78
5.1	Input parameters	79
5.1.1	Statistical Parameters	79
5.2	QDP and EDR as indicators of vascular asymmetries	81
5.2.1	Performance of QDP	81
5.2.2	Performance of EDR	83
5.2.3	Discussions on Performance of QDP and EDR	84
5.3	ANN Analyses Results	85
5.3.1	Performance of the Various Network Configurations	85
5.3.2	Predictive Power of the ANN	88
5.3.3	Discussions on the Performance of the ANN	91
6	Conclusions	97
6.1	Concluding Remarks and Summary of Contributions	97
6.2	Future Work	99
	References	102
	Appendix	117

List of Figures

Figure 1-1 Canadian Female Cancer Incidence Percentage, 2004.....	2
Figure 1-2 Canadian Female Cancer Mortality Percentage, 2004	2
Figure 2-1 A Single Layer Network.....	29
Figure 2-2 A Fully Connected With 2 Hidden Layers and 1 Output Nodes.....	29
Figure 4-2 A Sample Image Obtained from ISC, New York.....	57
Figure 4-3 A Sample Montreal Image (Before and After Resizing)	57
Figure 4-4 Image Histogram of a New York Image Before and After Intensity Adjustment.....	58
Figure 4-5 A Sample New York Image Before and After Intensity Adjustment.....	59
Figure 4-6 Image Histograms of a Montreal Image Before and After Intensity Adjustment	59
Figure 4-7 A Sample Montreal Image Before and After Intensity Adjustment.....	60
Figure 4-8 Segmented Breast Regions (showing an abnormal sign in the right UOQ).....	62
Figure 4-9 Percentage of Breast Tumour Occurrence by Quadrants	63
Figure 4-10 Two Pictures with the Same Statistics.....	64
Figure 4-11 Quadtree Decomposition Analysis with Two Images and Two Thresholds.....	66
Figure 4-12 Quadtree Decomposition Analysis of a Breast Image.....	68
Figure 4-13 Edge Detection Analysis for Parameter Extraction.....	70
Figure 4-14 A Three-layer Backpropagation Network.....	72
Figure 4-15 Tan-Sigmoid Transfer Function.....	74
Figure 5-1 Breast Quadrants Numbering.....	80
Figure 5-2 Typical Training Sessions of the RP and LM Algorithms.....	87
Figure 5-3 Interpretations of column headings for Table 5-6.....	89
Figure 5-4 Comparisons of Clinical Diagnoses and ANN predictions for Selected Images	94
Figure 5-5 ANN Predictions for Some Representative Images	96

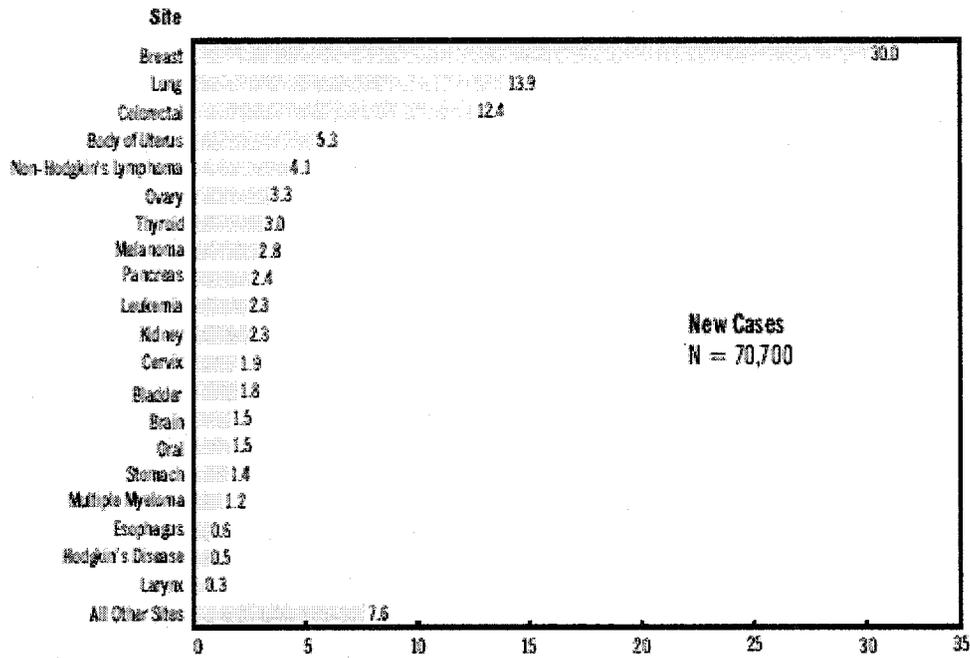
List of Tables

Table 4-1 Ville Marie Infrared (IR) Grading Scale.....	53
Table 5-1 Non-parametric Bivariate Correlation Analysis Results Summary.....	79
Table 5-2 Results from the Quadtree Decomposition Analysis.....	82
Table 5-3 Results from the Edge Detection Analysis.....	83
Table 5-4 Training Parameters for the RP and LM Training Algorithms.....	85
Table 5-5 Results of ANN Analysis with Various Network Configurations.....	86
Table 5-6 Summary Results for the Best-Performed Networks.....	91

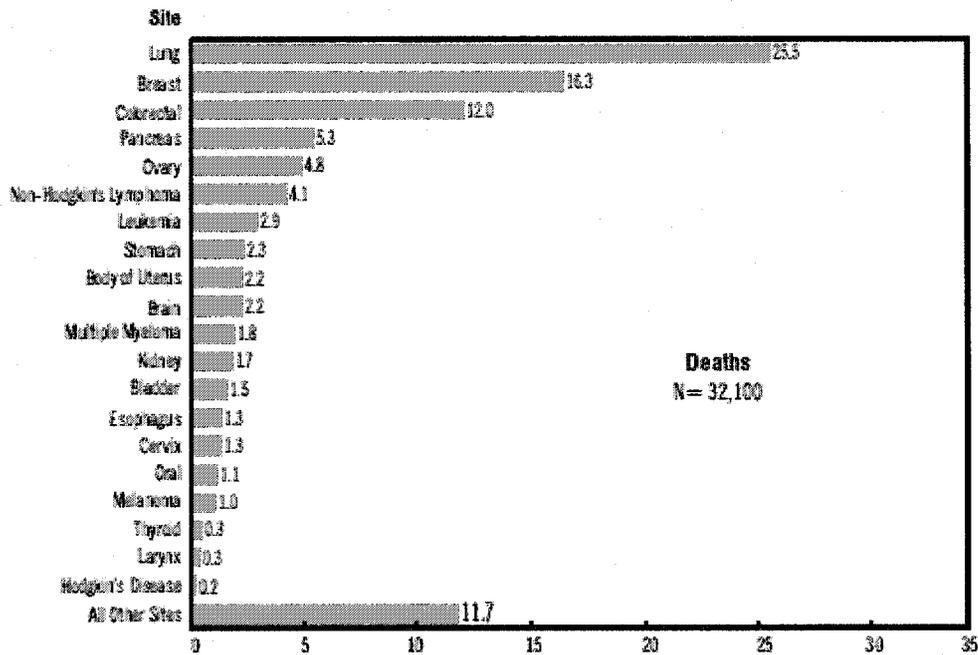
1 Introduction

1.1 Motivation

Breast cancer has been known for decades to be the most common type of cancer among women. Statistics compiled by government agencies show that almost one in three new Canadian female cancer cases is breast cancer. In 2004, Canada is expected to see 21,200 more new cases and a total of 5,200 deaths among women aged 20 and older [63]. Based on mortality rate and incidence rate data obtained from the National Center for Health Statistics (NCHS) and the National Cancer Institute Surveillance respectively, the American Cancer Society estimates 40,110 breast cancer deaths and 215,990 new cases of invasive breast cancer among the American female population for 2004 [3].



¹Figure 1-1 Canadian Female Cancer Incidence Percentage, 2004



²Figure 1-2 Canadian Female Cancer Mortality Percentage, 2004

¹ **Note:** Incidence figures exclude an estimated 76,000 new cases of non-melanoma (basal and squamous) skin cancer.
Source: Surveillance and Risk Assessment Division, CCDPC, Health Canada

² **Source:** Surveillance and Risk Assessment Division, CCDPC, Health Canada

Despite the high incidence rate, the breast cancer mortality rate has actually been decreasing since the late eighties. The 2004 estimated number of new cases and deaths for selected cancers in Canada are shown in Figure 1-1 and Figure 1-2 respectively. Most medical experts and researchers credit the encouraging results to early breast cancer detection as well as better cancer treatment. Decades of research and technology advances have produced more reliable screening techniques and procedures. A combination of breast self examinations (BSE), clinical breast examinations (CSE), imaging techniques such as mammography, ultrasonography and magnetic resonance imaging (MRI) are utilized to facilitate early breast cancer detection. The goal is to detect the cancer before tumorous cells become invasive, or metastasis has not yet started. More treatment options are available during early stages of the cancer and thus early detection generally increases the survival rate. Although the imaging techniques employed today, with the most widely used being mammography, are quite effective for early breast cancer detection, researchers around the world continue to study ways to improve the levels of sensitivity and specificity of the screening processes. Some have suggested new techniques to analyze mammograms while others look for alternative imaging techniques to complement the current technology.

1.2 Thesis Objective and Problem Statement

The application of thermography or infrared imaging in early breast cancer detection and risk prediction has been widely studied since the sixties. Results from the studies vary significantly. Some researchers have demonstrated that breast thermography used either independently or in conjunction with mammography, has great potential in

early cancer detection. On the contrary, different groups of researchers obtained unsatisfactory clinical results in breast cancer detection with thermography. The inconsistencies are partly due to non-standardized equipment and procedures practised during image acquisition, complicated processing and analyzing of the thermograms and most importantly, lack of standard protocols to guide the final prediction. Manual thermogram evaluation is highly subjective, and possibly inaccurate due to the limitation of untrained eyes to recognize the most subtle differences between the normal and abnormal images.

In this thesis, quantitative analysis techniques will be developed and employed to assess breast images objectively and reliably. The ultimate objective of the thesis is to determine if thermography is an effective method for early breast cancer detection. Various statistical and non-statistical parameters representing thermal characteristics of the infrared images will be extracted from the image. With the help of advanced computer technology to automate the procedures, Artificial Neural Networks (ANNs), widely used in medical decision support tools, can be employed with different sets of parameters and configurations for clinical outcome predictions. It is not known at the time of data extraction which parameters will be most discriminative for differentiating the normal from the abnormal images. The ANN will also be employed to select a minimal set of parameters with the best predicting ability. Once the combination that provides the best performance is obtained, sensitivity and specificity values can be used to evaluate the overall effectiveness of the method.

1.3 Thesis Layout

This section describes briefly the organization of the thesis. Chapter 2 presents background information on the problem being addressed. Various imaging techniques used for breast cancer detection will be briefly discussed. This is followed by an overview consisting of history, technical information and potential advantages of medical thermography applied in early breast cancer detection.

Chapter 3 covers the literature review that presents the current state of the art in fields and topics related to the thesis. In particular, techniques used in image segmentation and thermographic feature extraction will be discussed. The latter part of the chapter reviews ANN application in breast cancer detection and diagnosis.

Chapter 4 describes in greater detail the problem statement of the thesis. Additional information will be provided on the breast images used in the experiment. This is followed by a comprehensive explanation of the methodology used in the experiments including data processing techniques, feature extraction and neural network modelling.

Chapter 5 presents the result obtained in the experiments. The performance of the feature extraction techniques will be discussed. The overall results from the experiments will then be presented to formulate an answer to the problem statement of this thesis. This chapter will be concluded with a discussion of the results.

Chapter 6, the final chapter of the thesis will offer some concluding remarks. A summary of contributions to knowledge and a brief discussion of future work will complete the thesis presentation.

2 Background Information

With the goal of improving survival rates for breast cancer patients, scientists have spent decades researching more effective early breast cancer detection methods. Given that this thesis spans two broad subject areas, namely medicine and engineering, background information of relevant research conducted to date will introduce important concepts that build the foundation for this thesis. Section 2.1 provides a brief description of methods for evaluating the effectiveness of clinical trials. Section 2.2 and 2.3 describe current and actively researched technologies for early detection. Emphasis will be placed on mammography, as it remains the most recognized and widely used detection modality. Various aspects of medical thermography and clinical breast thermography will be presented in detail in section 2.4. Topics to be covered include physiology and pathology of breast cancer, basic theory and physics of medical thermography, historical overview,

current infrared imaging technology, and advantages and potential pitfalls of breast thermography. Finally, an overview of ANN concepts and applications is provided.

2.1 Sensitivity and Specificity of Clinical Trials

In the clinical environment, the effectiveness of a cancer detection method is commonly evaluated by its sensitivity and specificity. *Sensitivity* of the method is defined as the probability that a positive test result is obtained given that the individual tested actually has the disease. *Specificity* of the method is defined as the probability that the test result is negative for an individual who does not have the disease [67].

In the context of breast cancer screening, a positive test result could be an abnormal breast image. Sensitivity indicates the percentage of people who have positive test results among a group of people with breast cancer. A negative test result given to a person with breast cancer is denoted as a *False Negative* (FN). Conversely, a positive test result given to a person with breast cancer is denoted as a *True Positive* (TP). Specificity is the percentage of people who have normal breast images among those who do not have breast cancer. A positive test result given to a person without breast cancer is called *False Positive* (FP), and a negative test result given to such a person is called *True Negative* (TN). With TP, TN, FP, and FN defined, the *Positive Predictive Value* (PPV), the percentage of positive results which are actually true, and the *Negative Predictive Value* (NPV), the percentage of negative results which are actually true can be formulated as follows:

$$PPV = \frac{TP}{TP + FP} \qquad NPV = \frac{TN}{TN + FN}$$

A detection imaging modality with low sensitivity leads to frequent misses in detecting any abnormalities in the breasts at an early stage. Since fatal breast cancer is caused by cancerous cells that have spread into lymphatic nodes and further into distant organs over time, treatments for breast cancer are usually more effective at the early stages of breast cancer when the malignant tumour can be removed locally or the metastatic spread can be blocked by cancer therapies. Therefore, it can be concluded that improving the sensitivity of a detection method is an important step in reducing the breast cancer mortality rate.

On the other hand, while low specificity (high occurrence of false positives) does not cause direct bodily harm, it does bring unnecessary anxiety and distress to women, and may lead to unnecessary invasive diagnostic procedures. In most cases, women are called back to obtain more mammograms for different views of the breasts and magnification of the suspicious “object”. Some of these women will further be recommended for diagnostic evaluations such as ultrasound and biopsies. These follow-up evaluations increase the cost of operating breast screening clinics substantially; one study estimated that 33% of screening cost was spent evaluating false positive results [59] [20]. In addition, women are unnecessarily exposed to ionizing radiation during follow-up mammography that could otherwise be avoided. There are several types of biopsy procedures used for diagnostics. Most biopsy procedures are invasive which causes discomfort in addition to the possibility of surgical complications such as pneumothorax in some rare cases [9].

2.2 Mammography

In its simplest definition, mammography is the utilization of ionizing radiation to create images of the breast, and therefore mammograms are X-ray pictures of breast tissue [80]. The terms “mammography” and “mammogram” are often used interchangeably. Before the first specialized mammography machine was developed in 1966, standard X-ray machines were used to obtain images of the breast. However, as breasts consist of soft tissues with low-contrast internal structure, insufficient level of detail could be captured to guarantee a reliable diagnosis. Today, mammography equipment is designed to render images with the high spatial resolution and high contrast required for breast cancer screening and diagnosis while minimizing the patient’s radiation exposure. To realize this goal, high contrast and high speed radiographic film is used. In addition, the breasts are compressed to less than 5cm between two plastic plates to reduce the object-film distance and to produce sharper images [78]. The compression is also necessary to produce more uniform thickness of the breasts for better image quality, and spreading of the tissue through compression decreases overlapping of the breast structures for more accurate mammogram interpretation.

2.2.1 Sensitivity and Specificity of Mammography

Sensitivity and specificity levels of mammography are affected by a number factors such as imaging technology and procedures which determine the image quality, interpretation skill of the radiologists, size of the lesion, age of the patient, density of the breasts and whether the patient is undergoing hormonal therapy. Biologic factors such as the presence of extensive parenchymal densities (dense breast tissue), tumours of lobular histology (lobular carcinomas in situ), and small tumours could also lead to failure of

breast cancer detection by mammography [60]. While image quality and interpretation accuracy are human factors that can be controlled reasonably well, other factors have been widely investigated to determine their effects on the sensitivity and specificity of mammography.

A 1998 research study conducted in New Mexico with a database consisting of 183,134 mammograms has shown that the sensitivity of mammography is 81% for women aged 64 or older, 78% for women between 50 and 64 years old, 77% for women between 40-49 years old and only 54% for women younger than 40 years old [73]. In the same study, researchers also found that the sensitivity is 74% versus 81% for estrogen replacement therapy users and non-users respectively, and 68% versus 85% for dense breasts and low density breasts respectively. More recently, a similar study conducted with a pooled mammography registry consisting of 329,495 women investigated the individual and combined effects of age, breast density and hormone replacement therapy (HRT). The researchers concluded that the accuracy of screening mammography decreased with higher breast density and younger age. Specificity for women from all age groups who used HRT is 91.7%. The levels vary slightly for women who did not use HRT where it is for 91.4% for women aged 40-44 and 94.4% for women aged 80-89 [11].

2.2.2 Effectiveness of Mammography in Reducing Breast Cancer Mortality

Since the introduction of mammography for breast cancer detection in the 1960s, numerous controlled trials around the world have been performed to investigate its effectiveness to detect breast cancer at an early stage, and therefore reduce the mortality rate of the cancer. Among these, several large randomized controlled trials conducted in

the United States, Canada, Sweden and Edinburgh are frequently quoted in the literature. In total, almost half a million women aged between 40 and 74 participated. These meta-analyses vary in the number of participants, age group of the women screened, screening modalities, and screening intervals.

Primic-Zakelj constructed a table adapted from the Report of the International Workshop for Breast Cancer Screening [23] to clearly show the characteristics of each trial [70]. Results from these trials were also summarized in the report by Fletcher et al [23]. It can be concluded that breast cancer screening is more effective in reducing the death rate among women aged 50 to 69 than among women aged 40 to 49. In fact, breast cancer mortality is not significantly reduced by breast cancer screening among women aged 40 to 49 whereas the reported mortality reduction due to mammography screening varies from 21% to 37% among women aged 50 to 69.

Recently, debates regarding the effectiveness of mammography arose when Gotzsche and Olsen challenged the accountability of the meta-analyses. The researchers pointed out the inconsistencies and baseline imbalances among these randomized trials, and further concluded that mammography screening is unjustifiable [33]. Despite the questions raised regarding the effectiveness of mammography in reducing breast cancer mortality rate, the method is still by far the most common practise for breast cancer screening, and generally accepted as the 'golden standard'. Other detection methods currently in use complement mammography for more accurate diagnosis.

2.2.3 Risks of Mammography

The risk of ionizing radiation used in obtaining mammograms is of concern among the medical community. The risk of radiation induced cancer is constantly

weighed against the potential benefit of breast cancer screening. Before technology advancement in mammography equipment reduced the radiation dosage required for good quality images, some medical experts had in fact discouraged women of all ages to attend regular mammographic screening. Today, with the low radiation dosage used, controversy remains as to when women should be exposed to the radiation initially and what a safe interval for exposure is. In other words, there are still no definite answers to at what age a woman should begin regular mammographic breast screening and when she should go back for subsequent screenings.

A National Research Council (NRC) committee analyzing the biological effects of ionizing radiation proposed a model to evaluate the relative risks of cancer [15]. In this model, it was shown that the age-at-exposure is strongly associated with the subsequent risk of breast cancer following exposure for both the incidence and mortality rates of breast cancer. The relative risks of ionizing radiation are highest among younger women and decrease with increasing age at exposure. The carcinogenic potential of radiation is quite small for women over the age of 40.

2.3 Imaging Modalities

During the sixties and seventies, due to the potential carcinogenic radiation used in mammography, researchers began searching for other methods of detecting breast cancer. While radiation risk is not the main driving force these days, this type of research remains active. Researchers are finding ways to improve the overall effectiveness and risk profile of breast cancer detection and diagnosis. Some of the imaging modalities currently in use or investigated include ultrasound or sonography, MRI, positron emission tomography (PET), and electric impedance scanning (EIS).

Ultrasound is most commonly employed in diagnosis to analyze the composition of a non palpable lesion detected by mammography. It is used to identify a fluid filled cyst from a solid tumour. However, due to its inability to differentiate between a benign and a malignant tumour and to detect microcalcifications, an important early sign of potential breast carcinoma, it is not suitable for early breast cancer detection.

MRI plays a complementary role to mammography. As the efficiency of mammography is lower for screening women with breast implants and denser breasts, researchers are studying the suitability of MRI for screening these women. Screening young women at higher risk of breast cancer is also being investigated as MRI is radiation free. In general, it is more commonly used in diagnostics to evaluate the extent of the malignancy once cancer is diagnosed and to investigate palpable lesions that cannot be detected by conventional mammography. However, due to expensive setup costs and the inability to detect microcalcifications and certain lesions not identifiable on cross-sectional images, MRI has a very low efficacy for screening breast cancer.

Both PET and EIS are not used extensively on patients but are actively being studied in clinical trials. In PET, the patient is injected with a small amount of sugar and radioactive material. As cancerous cells absorb sugar faster than normal tissue, radioactive sugar accumulates around the tumour. A scanner is then used to detect any radiation emitted from the tumour and the location is computed. PET is not commonly used in early cancer detection. Rather, it is said to be most useful in evaluating recurrent breast cancers and the extent of metastatic spread, as well as monitoring the patients' response to cancer therapy [21][75].

EIS is employed in clinical settings to locate a tumour in an abnormal area indicated by mammograms. The method is based on the fact that breast tumours conduct electricity better than normal breast tissue. A small current is transmitted through the body and a scanning probe placed over the breast is used to generate measurements of the electric impedance. Researchers claim that the use of EIS can reduce the number of biopsies required to determine the malignancy of a breast mass [61]. However, EIS is not used for screening and is not recommended for patients with an implanted electrical device such as a pacemaker.

In summary, advancements in science and technology bring medical society more imaging techniques that have great potential to improve breast cancer detection or diagnosis accuracy. However, no single method fulfills all the challenges. Scientists are still actively searching for non-invasive, safe, reliable and effective technology for breast cancer management. This is especially true in the field of early breast cancer detection where mammography despite some drawbacks and risks, is still the only method widely available.

2.4 Medical Thermography

Thermography measures temperature variations on the surface of the body by detecting heat emitted by regions of the body and subsequently transforming these signals into images for analysis. Medical thermography refers to the utilization of the technique to detect, diagnose or monitor diseases or illnesses in the clinical environment. In this section, some historical and theoretical background of thermography will be presented. A brief overview of thermography in various clinical applications will follow, with

emphasis given to breast cancer detection and risk prediction. Imaging technology employed in thermography will also be discussed.

2.4.1 Historical Background

Medical thermography has a rich history. Early Greek physicians are said to have diagnosed tumours beneath the skin by applying wet mud to certain parts of the body, and observing the area that dried out in the shortest time. Centuries later, around 400BC, the “Father of Medicine”, Hippocrates, recommended physicians examine the patient for any part of the body that was hotter or colder than the rest as disease would be present there [8]. Hippocrates used his hands to measure the relative hot and cold spots. When Galileo invented a thermoscope in the sixteenth century, a more quantitative approach could be taken to measure heat.

Scales were soon added to the thermoscope, a somewhat cumbersome apparatus, and the first thermometer was invented. Although thermometers were very useful for measuring temperature, they were not suitable for taking surface temperatures of the body, as surface temperature is correlated with the ambient temperature and skin dissipates heat in all directions. A breakthrough came when Sir William Herschel’s experiment with glass filters of different colours discovered infrared radiation. He observed that the amount of heat associated with light passing through a filter was related to the colour of the filter. Thermometer readings were highest when placed at the far end of the red spectrum, the infrared zone. The scientific community at first reacted sceptically to the discovery of infrared. Nevertheless, when Sir John F. W. Herschel, son of Sir William Herschel, successfully recorded the heating rays on a specially prepared

paper in 1840, and further evidence was provided by other scientists within the next few decades, the debate was finally put to rest.

Recent technical advances in infrared imaging could be traced back to World War II when the ability of infrared cameras to take pictures in total darkness found important applications in the military. The infrared scanner was also invented during this time. Compared to infrared cameras available on the market today, these infrared scanners were slow and the resolution of the images was poor. However, they provided medical experts a valuable tool to measure the surface temperature of the human body.

Clinical thermography was said to have started in Montreal. With the infrared technology available in the 1950s, Dr. Ray Lawson began to study the implications of surface temperature to diagnose breast cancer. Two papers entitled “Implications of surface temperature in the diagnosis of breast cancer” and “Thermography – a new tool in the investigation of breast lesions” were published in 1956 and 1957 respectively [56] [57]. In the next two decades after the significant work of Dr. Lawson, clinical thermography saw a great interest among researchers. Potential applications of clinical thermography include detection of stroke, diagnosis of carpal tunnel syndrome, pain management, cardiovascular disease monitoring, and breast cancer detection [74][6][45][34].

2.4.2 Physics of Medical Thermography

Since Sir William Herschel’s discovery, knowledge of infrared radiation and electromagnetic waves in general has improved tremendously. The natural phenomenon of heat, energy transfer, and temperature variations are explained more thoroughly by theories and experiments.

Heat is the energy that arises from molecular motion within matter, and is one of the many forms of energy that exist in the universe. It can be transferred from one object to another in a few different ways. The sun transfers heat energy to the earth through infrared radiation. The Stefan-Boltzmann law is expressed in analytical form in equation (2-1).

$$E = \sigma T^4 \quad (2-1)$$

Where E is the total emissive power ($\text{W}\cdot\text{M}^2$) of a black body heated to temperature T (K), and σ is a proportionality constant $5.668 \times 10^{-8} \text{ W}\cdot\text{M}^2\cdot\text{K}^{-4}$. According to the law, all objects heated above absolute zero (0 K or $-273.15 \text{ }^\circ\text{C}$) emit energy in the form of electromagnetic waves at a rate proportional to the fourth power of the object's temperature [68].

Electromagnetic waves occupy a vast band on the frequency scale. Infrared radiation with wavelength ranges roughly from $0.8 \text{ }\mu\text{m}$ to $1000 \text{ }\mu\text{m}$ lies just beyond the red end of the visible light spectrum. The term spectral emissive power is often used to refer to the variation of power radiated at different wavelengths. The wavelength at which maximum heat energy is radiated varies with the temperature of the object. This observation can be derived from Wien's displacement law, shown here analytically in equation (2-2), where λ is the wavelength and T is the absolute temperature of the object. Wien's displacement law was developed by W. Wien in an attempt to explain measured spectral emissive power.

$$\lambda_1.T_1 = \lambda_2.T_2 \quad (2-2)$$

From Wien's displacement law, it can be inferred that as the temperature of the object increases, the wavelength at which maximum radiation is emitted becomes shorter,

and thus approaches the visible light range. The red glow of a piece of metal heated to a very high temperature is an observation of this phenomenon. While the laws established by Stefan, Boltzmann and Wien were invaluable in explaining most experimental data, their analysis based on the classical theory of electromagnetic wave has its limitations. Finally in the early twentieth century, Planck who did his analysis based on the quantum theory unveiled by Einstein placed the final piece of the puzzle for radiation energy with his equation expressed in equation (2-3).

$$E_{\lambda} = \frac{2\pi hc^2}{\lambda^5 (e^{\frac{A}{\lambda T}} - 1)}, \quad A = \frac{hc}{k} \quad (2-3)$$

Where E_{λ} is the spectral emissive power ($\text{W}\cdot\text{M}^{-2}\cdot\text{Micron}^{-1}$) at the wavelength λ , absolute temperature T (K), h is the Planck's constant 6.624×10^{-34} Joule-Seconds, c is the velocity of light and k is the Boltzmann's constant 1.38×10^{-23} $\text{J}\cdot\text{K}^{-1}$.

From a basic understanding of the physics behind infrared radiation, follows a focus on the biology behind thermal radiation of the human body. From the thermal control perspective, the body can be roughly divided into two zones, the shell (which is basically the skin), and the core. As homeotherms, the human core temperature is kept fairly constant at 37 °C. This core temperature is vital in maintaining the functioning of human organs and biological processes, and is kept stable by the thermoregulatory system of the body. The autonomic nervous system consists of the hypothalamus in the brain and thermoreceptors located mainly in the hypothalamus, spinal cord, the mucosa of the airways and the skin. One third of the thermoreceptors are located in the skin. The hypothalamus can be seen as the central processing unit, and acts like a negative

feedback circuit with feedback provided by the temperature sensing thermoreceptors [50]. Data provided by the thermoreceptors is constantly being monitored against an internal set point. When the body temperature drops, heat is generated by metabolism as well as muscular activity such as shivering and conserved through mechanism such as vasoconstriction. When the body temperature rises, heat loss is attained through perspiration, vasodilatation, and thermal radiation. Almost 50% of body heat is lost through the skin. In general, heat transfers within the human body and between the human body and the surroundings occur in four different ways: conduction, convection, radiation and evaporation [14] [22].

In contrast to the core temperature, skin or surface temperature fluctuates with ambient temperature. As long as there is a temperature gradient, heat is transferred by infrared radiation between the skin and the surrounding. By definition, a black body is a body which completely absorbs all incidence radiation, and emits the maximum amount of radiation energy as specified by Planck's Law. Emissivity (ϵ) is a parameter used to indicate the ability of a body to emit heat radiation. Emissivity of all other bodies is given relative to the emissivity of a black body, which is set as 1. Although infrared radiation emitted from the surface of a human body ranges from 3 – 50 μm , maximal emission occurs around 8 to 10 μm with a peak at 9.3 μm [22]. Furthermore, experimental results have shown that emissivity of the skin varies significantly for wavelengths in the range of 3-5 μm . However, the skin approximates a thermal black body very well for wavelengths longer than 6 μm . At this range, ϵ is approximately 0.98. Therefore, equipment designed to measure skin temperature should focus on detecting wavelengths within the maximal range to ensure more accurate and reliable results. In

addition, any lotion or ointment applied to the skin can also affect its emissivity. Hence, techniques and protocols practiced during the surface temperature recording should also take into careful consideration human and environmental factors that could change the emissivity of the skin. When these external factors, including the temperature, air flow and humidity of the recording room are carefully controlled, thermal patterns observed on the infrared images can be related to the thermo-physiological processes within the body [71].

2.4.3 Breast Thermography

As mentioned earlier, the study of thermography in the detection of breast disease began as early as the late fifties. The most common application lies in the early detection of breast cancer employing techniques of symmetry analysis between contralateral breasts. The breast images are investigated for an increase in temperature level and a change in the overall thermal pattern caused by metabolism and vascular changes of the underlying breast tissue. Although the metabolism rate and vascularity of the tissue fluctuates significantly with various physiological processes, it has been shown that normal physiological processes such as thalamic thermal control, pregnancy and menstrual cycle as well as external stimulants such as alcohol and cigarettes bring about symmetrical changes to the breasts. However, unilateral pathological disease such as breast carcinoma can be detected from asymmetrical thermal images. Although the observed abnormality in the images varies due to the different breast cancer pathologies, an increase in temperature level and/or a change in superficial vascular patterns can be observed. Hyperthermia around the tumour is thought to be caused by two mechanisms:

1) Higher metabolic rate of the cancerous cells and 2) angiogenesis of the tumour. Angiogenesis refers to the formation of the new blood vessels around the tumour.

A growing tumour has a higher metabolic rate and therefore transfers more heat to the surrounding tissue. A study conducted by Gautherie and Gros in the Laboratory of Biomedical Thermology in Strasbourg, France involving approximately 58,000 patients in the span of 12 years found that venous blood leaving the tumour has a higher temperature compared to the arterial blood and thus can be considered a heat source for the surrounding tissue [29]. The temperature and thermal conductivity were higher for the cancerous breast when compared to the contralateral healthy breast due to an increase in blood flow. This is consistent with the physiological concept that increased blood flow, higher metabolic rate and greater heat production have a cause-effect relationship.

Gautherie and Gros demonstrated that breast thermography can also be useful in cancer risk prediction [30]. Thermograms were classified into one of five stages - Th I to Th V, according to increasing probability of cancer. Of the 58,000 women examined, 1563 women were placed in Th III, the level where thermograms are suspicious but inconclusive. Approximately 90% of women classified in either category Th IV or Th V as well as 18% in Th III had histologically confirmed cancers during their first visit. Among the 1563 women from Th III, 784 were diagnosed "normal" by other conventional means such as mammography or physical examinations. Within five years, 38% of these women developed histologically confirmed malignancy. Thus, women with thermograms categorized in Th III have a higher risk of developing cancer and should be monitored more frequently. Furthermore, these researchers were able to show that the growth rate (specifically doubling time) of the tumour is well correlated to its metabolic

rate and therefore heat production. A greater amount of heat is produced when the doubling time for tumour growth is shorter. This observation can be useful in determining the therapeutic strategy, as “hotter” cancerous cells may need more aggressive treatment.

In angiogenesis, new blood vessels are formed to bring adequate blood supply once tumours grow beyond 1-2 mm³. Greater vasculature in turn facilitates rapid growth of the tumour and increases the probability of dissemination by entering the circulation system [18]. Angiogenesis of the tumour has been widely investigated for primary cancers and further into the stage of metastatic spread and found to be a significant prognosis indicator for breast cancer [58][13]. Microvessel density (MVD) is used to measure the extent of the angiogenesis. A high MVD has been shown to correspond to lower disease-free survival. Infrared cameras can capture the formation of new vessels that lead to neovascularization (vascular hot spots). In a clinical research conducted by Gamagami and co-workers involving 530 patients, hyperthermia and hypervascularity could be observed in 86% of non-palpable cancer, 15% of which were not detectable on the mammograms [28]. The observations were made even before the appearance of a cancerous mass on the mammograms. The result of this investigation is consistent with previous studies by Gautherie and Gros in which 38% of women had an abnormal thermogram as early as five years before malignancy was confirmed.

Breast thermography which is non-invasive, painless, and relatively inexpensive is still experiencing scepticism among healthcare professionals. Some believe that the sensitivity and the specificity of the method are too low to be used as a screening tool [82]. Others criticize the method as unable to identify the exact location of the tumour

for biopsy once an abnormality is detected. Thermography reflects the physiologic parameters of the tumour whereas mammography images the anatomic parameters.

Even clinical trials have demonstrated varying results over the last few decades. Some researchers such as Gautherie *et al.*, and Isard *et al.* demonstrated that patients with poor thermographic prognosis factors had shorter survival time [30][48], but Sterns *et al.* found that thermography was not an independent prognosis indicator [76]. The inconsistencies are partly due to image acquisition equipment and procedures [41], complicated processing and analysis of the thermograms and most importantly, lack of standard protocols to guide the final prediction. Evaluation of the thermograms was often done manually and thus affected by human subjectivity. With advanced technology found in today's image acquisition equipment, computer hardware and sophisticated imaging processing software, some of the inconsistencies can be overcome by more controlled image acquisition protocols, as well as faster and more intelligent image analysis algorithms. Evaluation of the thermograms can be performed more objectively by employing quantitative analysis and image classification software.

Increasingly, medical thermography researchers believe that infrared imaging would be a valuable imaging modality to complement mammography in breast cancer screening. Since infrared images do not map the breasts anatomically, the results are independent of the composition of the breasts. Therefore, the method is just as effective for young women, women with dense breasts or breast implants. It has been shown that this multi-modality approach, which takes advantage of a functional modality (IR imaging) and structural modality (mammography) as a first line detection method can improve the overall sensitivity of breast cancer screening [53]. Perhaps even more

importantly, infrared imaging's ability to detect metabolic and vascular changes at an early stage of tumour genesis, before the morphological changes can be detected by mammography, may prompt an earlier detection during the cancer growth and thus improved survival rate [2].

2.4.4 Imaging Technology in Medical Thermography

Different technologies to quantify surface temperature have been attempted since the discovery of infrared radiation emitted by a thermal body. An early technology known as evaporagraph was used in the 1950s, but was soon replaced due to poor quality and resolution. Generally, all thermal imaging systems can be categorized as contact or non-contact systems.

The most common contact thermal imaging technology is known as the liquid-crystal contact thermography. Liquid crystals are organic substances which, at a specific temperature, exist in the mesophase, the phase between solid and liquid. Liquid crystals have the mobility of a liquid with several degrees of freedom and demonstrate the optical properties of a crystal. When spread out in a thin film, the geometric property of a crystalline structure determines the spectral intensities of the light reflected. The reflected wavelengths give the crystal its distinct colour. Typical liquid crystals used in medical thermography go through colour transitions of red, yellow, green, blue and violet when the crystals are warmed. The liquid crystals used to measure skin temperature are encapsulated in a thin film and usually mounted with a black carrier layer. The film is spread out on the area of the skin to be measured and colour change is observed. Cool air can be applied to the skin, and the film can be reapplied to obtain a dynamic thermal and

vascular image. Different mixes of crystals are used to achieve the temperature spans in which the encapsulated liquid crystal (ELC) film can measure.

Despite giving an overall picture of the cool and hot spots of the skin, liquid-crystal thermography is not commonly used these days due to several limitations and downfalls. Due to the nature of contact thermography, the film inevitably affects the skin temperature and influences the heat exchange between the skin and the surrounding. The temperature resolution is limited by the overlapping of wavelengths reflected by adjacent crystals and the optical properties of liquid-crystal used in practical applications. The spatial resolution and the temporal resolution are limited by the ELC film and the carrier film it is attached with. In addition, inconvenient calibration methods severely restrict the ability of the device to measure the absolute temperature of a surface.

The imaging systems widely employed in research and clinical applications are non-contact infrared cameras that are able to capture very small amounts of radiant power from the source and convert it to electrical signals. Cameras utilize either a linear array or more efficient focal plane (two-dimensional) array of detectors to detect the incidence radiant flux. Infrared detectors are generally categorized by their spectral response to the radiant energy, the speed of the response, and their detectivity, a parameter indicating the signal-to-noise ratio per watt of incident power.

There are two different types of infrared detectors: thermal detectors and photon detectors. In thermal detectors, the temperature of the sensing element increases with the absorption of incidence energy and in turn triggers a change in the electrical output signals. As the thermal characteristic of the sensing element is not wavelength dependent, the output signal in response to the incident radiance can be considered constant over a

wide range of wavelengths. A variety of thermal detectors known as a microbolometer employs bolometer film whose electrical resistance changes when its elements on the focal plane absorb radiant energy. The rate of change in the electrical resistance with respect to the power dissipated is related to the rate of change with respect to temperature by the thermal impedance [51]. When coupled with the bias current, this rate of change characterizes the responsivity of the detector. Other types of thermal detectors being researched for suitability in commercial infrared cameras are thermopiles and pyroelectrics. A thermopile, consisting of connected thermocouples, detects temperature change by utilizing the thermo-electric property of a thermocouple junction (two metals connected in a loop) which generates a voltage proportional to the difference of temperature between the two metals. Pyroelectrics are ferroelectric materials that create an electric potential as a result of a change in temperature due to migration of charges in the material. As it detects the rate of temperature change instead of the temperature change, it is not as susceptible to the constant background radiation [16] [81].

Compared to the thermal detectors, photon detectors generally have faster response times and better detectivity, but are more expensive. However, until the recent technical advances in thermal detectors, photon detectors such as photovoltaic and photoconductive detectors were the only kind of detectors with spectral response, speed of response and detectivity suitable for medical thermographic equipment. In photovoltaic detectors, the incident radiant energy is converted into a change in voltage when the photon flux excites the electronic energy states of the semiconductor material. Since the energy of a photon is dependent on the wavelength as shown in equation (2-3), the output signals are therefore proportional to the incident wavelength. In

photoconductive detectors, free charge carriers are released when the incidence photon flux is absorbed. The increase in conductance caused by these free charge carriers can then be measured by applying a bias voltage across the photoconductive material [44]. Most photon detectors operate at cryogenic temperature to increase the signal to noise performance. As a result, a cooling mechanism such as liquid nitrogen is required.

Recent technical advances in infrared sensors and systems coupled with computer power for image processing have greatly improved the thermal, spatial and temporal resolution achievable as well as the spectral ranges measurable. For example, a commercially available, FDA approved thermal imaging system with a thermal resolution of 0.0125 °C can produce a 1024x768 image with 12-bit representation of the pixel value at a speed of 15 frames per second. Reducing the collection interval between thermal images has enabled real-time imaging and improved accuracy in dynamic thermographic analysis. Furthermore, cost reduction introduced by innovations in sensor technology has also made high performance systems more available to the medical and research community. A good example of the new infrared sensor technology is the uncooled focal plane array [7]. Focal plane arrays are comprised of microbridge - readout circuit pairs. A temperature change in the microbridges will trigger a change in current through the readout unit. An uncooled infrared array weighs only a few grams and provides very high thermal sensitivity and spatial resolution. As cryogenic cooling is not required, cameras containing thermoelectric sensors are much more compact and reliable.

2.5 Artificial Neural Networks (ANN)

The human brain is the most powerful and complex part of a human body. It performs in seconds a task that takes the smartest machine hours to process. Over the

years, neurobiologists have tried to model the structures and functions of the human nervous system. The human brain, comprised of simple neurons, is the central processing unit of the nervous system constantly receiving, processing, and storing information, making decisions, and initiating further actions. Structurally, the human brain can be seen as a network of neurons interconnected by elementary units called synapses.

An artificial neural network is a network which models the structure and mimics the function of the brain. ANNs are therefore said to be a form of artificial intelligence. Resembling the brain, ANNs are also structurally built on neurons (nodes) and synaptic weights. An ANN acts as a parallel distributed processor comprised of simple processing units (neurons) that have the ability to acquire and store knowledge through a learning process using interneuron connection strength (synaptic weights) [40]. The parallel and nonlinear nature of the ANN makes ANNs a powerful tool in modelling biomedical systems and solving engineering problems. More details about the learning processes will be presented in the first subsection “Training the ANN”. The second subsection discusses a specific problem related to modelling biomedical systems: “Pattern Classification Systems in Medical Applications.”

2.5.1 Training the ANN

An ANN is commonly applied to model an unknown relationship between system variables (inputs) and the observed responses (outputs). Since the relationship is not known in advance, the ANN must be able to learn from input-output pairs provided. The learning process is performed by training the network. The goal of training the network is to obtain a system that generates appropriate outputs based on input parameters provided.

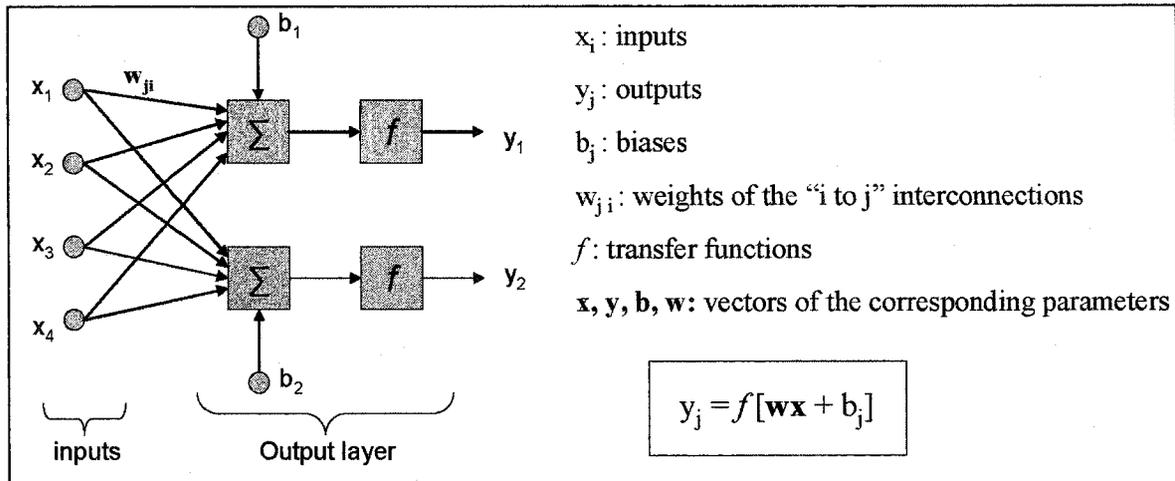


Figure 2-1 A Single Layer Network

The simplest neural network consisting of input nodes and an output layer with output nodes is shown in Figure 2-1. In this example, there are four input nodes and two output nodes. The output of the network is simply a function f of the sum of the weighted inputs. The function f is called an activation function or a transfer function, and its main function is to limit the magnitude of the output to some threshold values, for example -1 and 1. If the system to be modelled is more complex, additional layers of neurons, commonly known as hidden layers, are often added between the input nodes and the output layer. In Figure 2-2, an example of a fully connected network with five input nodes, two hidden layers and one output node is illustrated. Summation operators and transfer functions are not shown explicitly.

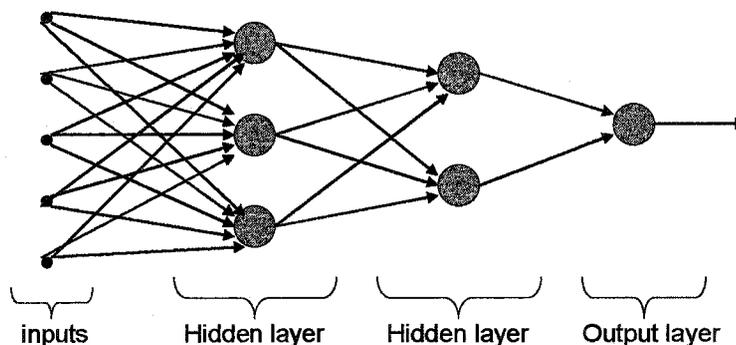


Figure 2-2 A Fully Connected With 2 Hidden Layers and 1 Output Nodes

During ANN training, weights are adjusted to minimize the difference between the ANN predictions and the desired outputs supplied for training. This difference is commonly referred as the training error. The method used to search for a solution (weight assignments) that has a minimum error is specified by a predefined learning algorithm. The search process will be reiterated until a certain stopping criterion or a performance goal is met. A performance goal specifies the upper limit for the training errors. Once a solution is found, the network is said to have converged.

Once the interconnection weights are selected during the training procedure with the provided input-output pairs, the ANN should be able to generate reasonable output for input that was not included in the training sets. This post-condition is commonly known as generalization. A well-designed neural network must have the capability to generalize. The ability of the ANN to generalize depends on the number of input datasets used to train the network, and whether they are a good representation of all the possible inputs. The “optimal” number of input datasets to be used for training in turn depends on the total number of nodes in the network. A general guideline is that approximately ten input datasets are required for each variable in the network. If the network is trained with too few input datasets, it may not be sufficiently trained to model the system. However, if the ANN is trained with too many hidden nodes, the network may be overfitted or overtrained. An overfitted network has “memorized” the training sets and lost the ability to generalize when a new set of inputs is presented. Hence, modelling a system with ANNs must take into consideration the number of training sets, the architecture of the neural network (number of nodes), and the learning algorithm selected for the training.

2.5.2 ANN for Pattern Classifications

One successful application of ANNs has been for system identification and pattern classification problems. Pattern recognition is sometimes used interchangeably with pattern classification. Features are properties of items to be analyzed to assist in classification. An important step in the design process of an ANN is feature extraction where the input variables are selected [46]. Pattern classification analyses in medical environment often involve predicting clinical outcomes in two types of situations. In the first situation, information to be analyzed may include personal data such as age, medical history and other signs and symptoms recorded by the doctors and nurses for a person with a particular medical condition. In the second case, the information to be analyzed consists of features extracted from medical images recorded to detect or diagnose a specific disease. In either case, identifying the most relevant set of features that is useful in predicting clinical outcomes is an important step in designing a medical decision support tool.

Once the input variables to the ANN are identified, the architecture of the ANN can be designed taking into consideration the number of cases available for training and validating the network. Validating datasets are used to obtain the performance measures of the trained network. In practice, training sets and validating sets are chosen randomly from the same database of input vectors and their corresponding outputs. Once the ANN is trained, the output it generates for the validating sets will be compared to the desired outputs to evaluate the effectiveness of the ANN by calculating performance measures such as sensitivity, specificity, and the total accuracy.

2.5.3 Medical Applications for Pattern Classification

In clinical settings, solving pattern classification problems with ANNs has proven valuable as a medical decision support tool in the areas of prevention, detection, diagnosis and intervention. In medical applications, due to the difficulties in collecting data or incomplete medical records, the number of cases available for training and validating the network can be quite limited. Hence, although a large set of features can be extracted, only a subset can be used as input variables to the ANN as more variables require a larger network (more nodes) to analyze. The subset can be chosen by consulting medical experts on the relevance of each feature. On the other hand, if the number of medical cases is abundant, all features extracted can be employed as ANN input variables. By studying the weights ANN assigned to these input variables, the most relevant features in classifying the medical cases can be identified.

3 Review of Existing Methods and Approaches

Significant improvements have been made in various areas of image acquisition and data analysis for the application of breast thermography. Advanced technology in the design and manufacture of infrared cameras bring images that have better spatial, temporal and temperature resolutions by orders of magnitude compared to the first generation of thermographic cameras. With the availability of powerful computer hardware and software, large amount of data collected can be analyzed more quickly and thoroughly with complex algorithms.

This chapter will present the state of the art with respect to breast thermography and image analysis. A review of recent publications will be presented with each section covering an important concept relevant to the thesis. The first section, infrared image analysis, presents state of the art techniques to analyse the spatial characteristics of still

images, as well as techniques to analyse the temporal characteristics of image sequences. The second section describes the application of ANN to various breast cancer detection scenarios with emphasises on ANN application to breast thermography. Finally, limitations in the state of the art that form the motivation for this thesis are discussed.

3.1 Analysis of Infrared Images

In the past, capturing breast thermograms was not vastly different from an informal photo shoot. A clinician, technician or scientist would take “pictures” manually, possibly from different angles to obtain more complete views of the objects. Each of the images could then be studied at a later time. Today, during a typical breast imaging session, dozens of images may be taken within a few minutes to produce an image sequence. Most researchers employ moving cameras or mounted mirrors to obtain lateral views of the breasts. A sequence of images comprises single image frames which can be analyzed individually.

Primary image breast thermography analysis techniques involve extracting thermal parameters and then studying the spatial characteristics of the infrared images. Different approaches and decision making strategies are used to assess the normality or abnormality of the breast images. As computers became more readily available in the early eighties, researchers began to utilize computer-assisted analysis for faster and more thorough data processing. The introduction of digital infrared imaging a decade later allowed infrared images to be processed and analysed using computers thus improving data analysis efficiency. For example, digital image processing tools can enhance an image to remove noise and improve contrast, and computer programs can extract relevant

image features for outcome prediction. In addition, manual thermogram evaluations suffer the limitation of subjective assessments. Therefore, significant effort has been expended to automate image processing, analysis and evaluation.

3.1.1 Spatial Characteristics

Spatial characteristics research encapsulates selecting an appropriate set of features from still thermal images and designing subsequent analyzing algorithms. The numerical values extracted must represent the visual heat pattern and be discriminative in nature in order to separate normal and abnormal thermograms. The performance of the quantitative analysis is evaluated based on the accuracy of outcome prediction. An automated approach to cancer detection should equal or exceed the accuracy of manual assessment, and approach the success rate of mammography. One of the earlier attempts was published by Goin and Haberman in 1983 [32]. A questionnaire to assess the frequency of occurrence and magnitude of the features was given to expert thermographers in order to identify a list of diagnostic features. The list was then further reduced with statistical techniques. The results of their study signified that five important diagnostic features were number of veins, occurrence of hot segments, Upper Outer Quadrant (UOQ) veins, Lower Outer Quadrant (LOQ) veins and whether there are any hot spots in the breast regions. The first four features characterize the vascularity and the fifth, the temperature distributions. The difference in values between the left and right breast features are further processed and used to classify the thermograms.

To obtain the spatial characteristics of a thermal image, one of the most commonly employed techniques is statistical analysis. A set of statistical parameters include mean temperature, standard deviation, skewness, and kurtosis for a defined

region of interest (ROI). The defined region may be the entire left or right breast, breast quadrants, a uniquely identified area unit or isolated hyperthermic spots. Depending on the methodology, the ROI may be of different shapes or sizes. The strategies and procedures exercised to identify ROI will be covered in the section describing image segmentation techniques. Of all the statistical parameters, mean temperature is the most common and thoroughly exploited for obvious reasons. It may give an indication of abnormal localized or regional hyperthermia when compared to the symmetrical region on the contralateral breast or the rest of the ipsilateral breast [27] [42] [52] [43].

While mean temperature gives a good overall assessment of the thermal characteristics, important spatial features and magnitude differences may be lost due to the averaging operations, especially if only mean temperatures of the whole breast are compared. Thus in addition to temperature differences calculated for whole breasts, Head and Lipari also proposed two other methods based on temperature difference of the breast quadrants. In one method, they assigned a score to each quadrant of the breasts based on the temperature difference between the contralateral breasts. A higher score is given if the temperature difference is greater. An index used to predict abnormal heat pattern is acquired by adding the scores from all quadrants. The second method involved taking the absolute value of the sum of all temperature differences between breast quadrants. After comparing the predictions with diagnoses given by an expert thermographer, they found that the second method with a normal cut-off temperature of 1°C for the quadrant differences showed more promising results [42]. However, when this technique was applied by Frize et al. to their sample of nine images, they demonstrated that the results could be improved by raising the threshold of normalcy

from 1°C to 1.5°C. With the threshold of 1°C suggested by Head and Lipari, they found two false positives on their sample of nine patients. However, when the threshold was increased to 1.5°C, no false negatives and false positives were found [26].

Although mean temperature has been shown to be a significant feature of the thermal images for asymmetry analysis, other features or derivations of features have been investigated for suitability. Qi and colleagues suggested the Bézier histogram technique [35]. In this algorithm, Bézier splines were employed to smooth image histograms that were previously computed to illustrate intensity (temperature) distribution of the heat pattern. Asymmetry analysis between the contralateral breasts was then performed by comparing the feature curves acquired by taking first and second order derivatives of the Bézier histograms. However, this technique was highly sensitive to noise. A different technique based on K-means clustering was proposed the next year [36]. After every pixel was properly labelled to a specific cluster, the mean temperature of each cluster was calculated. By studying the cluster mean distributions between the contralateral breasts, asymmetric abnormalities if any, could be identified if present. In the year 2002, the research group presented another asymmetry analysis algorithm investigating statistical parameters extracted from the thermal images [55]. The statistical parameters extracted included the four moments: mean, variance, skewness, and kurtosis, as well as entropy, joint entropy, and peak pixel intensity of the correlated image. These statistical values were not used directly in the asymmetry analysis. Instead, bilateral ratio, a number derived from feature values from the right and left breasts were calculated and compared among the cancerous images and non-cancerous images. From

the small samples of images analyzed, they observed that higher order statistics are more effective in measuring asymmetry.

Besides statistical analysis, a different approach for feature extraction is wavelet transformation of the infrared images. Wavelet analysis is a widely used image processing technique where images are subdivided into smaller images of different scales according to the level of detail each contains. This process is achieved by filtering the image with low pass filters and high pass filters iteratively. A low pass filter applied to infrared images generates sub-images showing information about the global temperature distribution and the temperature level (energy of the signal). A high pass filter operation produces a representation of the temperature gradient of the image [49]. These unique global and local representations of the image details can be compared between the contralateral breasts in heat pattern asymmetry analysis.

A relatively recent method of analysis in infrared imaging for breast cancer detection is the concept of thermal texture mapping (TTM) first introduced by researchers in China. The system utilizes a simple heat transfer model with heat source analogous to an electrical circuit with battery. Resistors with varying resistance in the circuit are the equivalent of body parts with different heat resistance rates in the thermodynamic systems. By assuming that surface temperature follows the Gaussian distribution, and finding the half power point of the heat emanation, either manually or automatically, the approximate depth of the heat source can be identified [37][38]. TTM is a relatively recent development in the analysis of infrared images. Ongoing research is exploring its potential and possible contributions to the medial infrared imaging field. In a study published in 2004 that compared the performance of TTM in breast cancer detection to

that of mammography and ultrasound, the research group found that TTM findings were well correlated to the pathological findings, and better than the other two modalities [83].

3.1.2 Temporal Characteristics

Temporal characteristics of infrared images are generated by the technique of dynamic thermography. Dynamic thermography refers to the capture and analysis of a series of infrared images. Temporal characteristics of the images are recorded by studying images sequentially and fluctuations in the heat pattern are analyzed for abnormalities. With this technique, caution must be taken to isolate artefacts caused by movement of the chest during inhalation and exhalation.

Traditionally, temporal characteristics of the surface temperature are analyzed by applying the “cold stress” method. The most convenient cold stress method is by circulating cold air around the chest area to induce cooling of the breast tissues. During the recovery (warming) period, a series of images are recorded for later analysis. The inspiration for the cold stress method is the observation that invasive malignant tumours with high thermodynamic activities override the regular thermoregulation system. The abnormal hot spots and vascular patterns seen on the thermograms cool down at a slower rate and warm up at a faster rate after cooling down compared to the surrounding healthy breast tissue.

Quantitative analysis of abnormal thermodynamic behaviour relies on parameters extracted from a specific image frame such as the one taken immediately after cooling is terminated and a series of subtraction images. Subtraction images are obtained by taking the temperature difference between two consecutive images. For example in one study, one of the parameters for discrimination analysis is derived by dividing the mean value of

the upper 30% of a subtraction image histogram into the overall mean temperature of the breast [31]. When dynamic thermography is included during an imaging session, the diagnostics criteria usually consists of both diagnostic features from dynamic thermography as well as those from static analysis such as asymmetry hot spots or vascular patterns, as these techniques may complement each other and as a result, improve overall detection accuracy [66].

A very different approach to dynamic thermography analysis was taken by Anbar and colleagues. In Dynamic Area Telethermography (DAT), temporal characteristics of the breast thermal pattern are fully explored by taking the Fast Fourier Transform (FFT) of the infrared images, and extracting the relevant features in the frequency domain [4]. Anbar proposed that abnormal heat patterns observed for cancerous breasts is a result of excessive vasodilation (widening of the blood vessels) induced by production of nanomolar quantities of Nitric Oxide (NO) at the neoplastic lesions. This extravascular NO overrides or attenuates the perfusion modulation controlled by autonomic neuronal stimulations. Blood perfusion has a complex modulation frequency spectrum caused by intricate autonomic processes that regulate the normal amount of intravascular NO (in picomolar quantities) and blood flow. DAT is applied to study quantitatively the effect of excessive NO on the perfusion modulation. Parameters extracted from the frequency analysis are selected to easily identify any amplitude differences at specific frequencies of the FFT spectra when the cancerous breast and non-cancerous breasts are compared. FFT spectra are obtained for both the temperature and the spatial thermal homogeneity, a derivation of temperature variance.

Each series of images comprised of 1024 images acquired at a rate of 100 images per second simultaneously from the frontal, medial and lateral views while patients were asked to hold their breath for 11 seconds. After reducing thousands of observations from a series of images into one diagnostic parameter that could be used to differentiate abnormal and normal thermal images, the method was evaluated by calculating the sensitivity and specificity of the method. The result showed that DAT can be an effective method in distinguishing cancerous and non-cancerous breasts. It is worth pointing out that the most discriminative diagnostic parameter is selected for each series regardless of the imaging view. However, the specific diagnostic parameters obtained vary significantly among the three imaging views indicating that the NO effects on the blood perfusion may be perceived quite differently depending on the view of image collection.

3.1.3 Image Segmentation Techniques

In this section, some image segmentation techniques employed by different research groups are described. As mentioned previously, selection of ROI is a crucial step in any computerized image analysis. From the most basic requirement of identifying the breast region on the thermal image, to the more algorithm dependent segmentation of sub-areas to compare between the contralateral breasts, image segmentation strategies may be a significant factor in generating experimental results. The task of manually identifying breast regions is straightforward, though it may be a tedious task in a study with a large database of images. However, to automate the process by computer algorithms, a simple visual task can be challenging due to the wide variations of breast sizes and shapes.

A relatively simple approach to this problem is by fitting a rectangle around the breast region after an edge detection operation. Each rectangle is further segmented into four rectangles of unequal sizes with the biggest rectangle fitted over the breast region with the least amount of distortion, and the smallest rectangle fitted over the region with the most distortion [27]. The varying amount of “distortion” is a combined result of capturing a 3-D object onto a 2-D image, and the natural shape of female breasts.

Another approach to automate the breast region identification is by applying the Hough Transform to extract the lower boundary curves and combine the result with the left and right body boundaries [36]. The left and right boundaries for the breast are obtained by first locating the right and left armpits. The Hough Transform is then used to detect two parabolic curves that may approximate the lower boundaries of the breasts. The intersection point of the parabolic curves is assigned to be the mid point between two breasts. Although this method works relatively well for images investigated by the authors, it may not be as effective for images with very different visual characteristics.

Since the mammogram reading uses the system of four quadrants to subdivide the breast region for localization, it seems to be a reasonable approach to take when segmenting the breast region for more accurate comparison between the contralateral breasts. Head and Lipari suggested a segmentation technique that divides the breast into four quadrants by utilizing reference points such as the nipples and the chin [42]. As this was the method followed for breast quadrant analysis for this thesis, the method will be described more thoroughly later in section 4.5.1.

3.2 Artificial Neural Networks (ANN)

ANNs have been applied extensively and successfully in the clinical environment for pattern recognition problems. Applications of ANNs in pattern recognition problems specific to breast cancer detection and diagnosis will be briefly discussed in this section.

3.2.1 ANN in breast cancer detection and diagnosis

The success of computer-aided analysis in clinical applications has encouraged breast cancer researchers to explore employing ANNs in widely different settings ranging from breast screening strategy to predicting response to cancer therapy [72] [10]. Active research in this field is focussed on mammography, MRI and ultrasound. In mammography, an important contribution of ANNs is in predicting malignancy of microcalcifications [62] [24]. Features selected for ANN analysis such as mean area, density and brightness of the clusters correspond to mammographic findings normally recorded by radiologists for malignancy evaluation. ANN analysis was found to effectively decrease false positive rates in mammography, and thereby, reduce the number of biopsies performed on benign lesions.

For MRI, as in mammography, ANNs are mainly used for breast carcinoma diagnosis aiming at reducing the number of benign lesions sent to biopsies. In 2001, Abdolmaleki *et al.* employed ANNs to extract quantitative parameters for an MRI time-intensity profile and subsequently used them as inputs to classify breast lesions [1]. A more recent study conducted in Sweden applied ANNs in their analysis of dynamic magnetic resonance images to select the most discriminative features from a set of morphologic and kinetic features considered in their investigation [79]. By receiver operating characteristics (ROC) curves, both studies were able to demonstrate that

sensitivity and specificity of malignancy predictions generated by the ANNs were comparable to that of a medical expert.

Researchers in ultrasound imaging have also explored the suitability of ANNs to diagnose breast lesions on sonograms. Drukker and colleagues performed a round robin analysis with a Bayesian neural network to discriminate actual lesions from false positives [17]. An automated approach to diagnose breast lesions based on sonographic texture analysis was developed by a group of scientist in Taiwan [12]. In the segmentation algorithm, radial lines were drawn outward from the approximate center of the ROI. A critical point belonging to the ROI boundary was identified by applying wavelet transformation to a moving window along the radial line, and searching for a point with maximum local variance in the wavelet sub-bands. Once the images were segmented, three feasible features were extracted and considered in the ANN analysis: variance contrast, autocorrelation contrast and distribution distortions of wavelet coefficients. The results of the experiment demonstrated the potential of combining image texture analysis and ANNs in assisting clinical decision making.

3.2.2 ANNs in Breast Thermography

Despite showing great potential, image analysis with ANNs has been investigated by relatively few researchers in breast thermography. Using data collected at the Singapore General Hospital, Ng and colleagues attempted the concurrent use of thermography and ANNs for the diagnosis of breast cancer [65]. Before the ANNs were used to evaluate the thermograms, the mean temperatures of all patients were calculated with statistical software. They found that average mean temperature of a breast

diagnosed with invasive breast carcinoma was 0.51°C higher than that of a breast with benign lesions and 0.85°C higher than that of a healthy breast.

In this study, statistical data extracted from the thermograms and physiological information gathered from the patients were input candidates to the ANNs. The statistical parameters considered were mean, median, mode, standard deviation and skewness. The physiological parameters included were age, family history, hormone replacement therapy, the presence of a palpable lump, and other relevant physiological information. Four sets of inputs containing either the statistical data or the physiological data, or a combination of both were investigated. Each network trained was said to have converged if the Root Mean Square (RMS) error between the ANN predictions and the target vectors fell within 0.15. The experimental results for various configurations were compared. The network trained with mean, median and modal of the temperature statistics had a higher sensitivity of 68.97%, but did not yield fast convergence. The network trained with the entire set of statistical data and physiological data, a total of 13 input parameters, was concluded to be the best network created. The network obtained a sensitivity of 61.54% and a specificity of 40%.

Recent research work published in September 2004 presented the preliminary results of employing ANNs to classify breast thermal images based on features extracted from wavelet transformation [49]. The feature extraction performed by wavelet analysis was briefly described previously in this thesis. Results of the ANN analysis are presented here. After the wavelet transformations, seven features extracted from the wavelet subbands were used as the inputs to the ANN. The architecture of the network employed one additional hidden layer which further transformed the features to two more separable

features for better classification by the output layer. Of the 40 images available for this study, 20 images with equal number of healthy and cancerous images were pre-selected for validation while the remaining images were used as the training set. For this preliminary study, classification of thermal images with ANNs provided a sensitivity between 80% and 100% depending on whether the image was of frontal or side position and a specificity of 93% for both positions.

3.3 Critical Analysis of Existing Techniques and Approaches

Research work in various areas of computerized analysis of breast thermography and ANN applications in breast cancer detection was presented. In the analysis of breast infrared images, strategies employed in feature extraction and image segmentation were reviewed.

To automate the analysis, image segmentation algorithms to identify ROI (breast regions) from the background were discussed. Qi suggested an automated approach using a Hough Transform that works relatively well for breast images with good contrast and well defined breast contours [36]. In this case, the lower breast contour was approximated by parabolic curves. The most common form of Hough Transform requires feature lines or curves to be expressed in parametric form for detection. Hence, if the breast contours do not resemble circles, ellipses, or simple parabolas, piecewise polynomial curves may be required to approximate the contour resulting in a highly complex algorithm.

Once the breast regions are identified, thermal patterns may be extracted and represented quantitatively. Some researchers calculated feature parameters for the entire breast region and compared them between contralateral breasts [65] [55]. However, to

facilitate more accurate comparison, it may be beneficial to further divide the ROI into smaller sub-regions. In doing so, more localized thermal patterns can be detected as well. Although dividing the ROI into rectangular segments fulfills this requirement and could be easily implemented, it does not map to the general shape of the breast [24]. A more intuitive approach was taken by segmenting the breasts into Upper Outer Quadrant (UOQ), Lower Outer Quadrant (LOQ), Upper Inner Quadrant (UIQ) and Lower Inner Quadrant (LIQ) using the nipples and chin if available as reference points [42].

In feature extraction, researchers have shown that both spatial characteristics and temporal characteristics could potentially be used to discriminate between normal and abnormal images in the asymmetry analysis. Among the statistical parameters, mean temperature and its derivation were explored extensively to quantitatively identify the possible presence of hot spots. However, one of the abnormal signs on thermal images is asymmetric vascular patterns, which may or may not be represented by difference in mean temperature. Haberman and Goin included a visual indication of vascularity difference in their analysis, but the ranking system they used was quite subjective [32]. More quantitative methods to represent the vascular patterns objectively are needed.

Dynamic Area Telethermography (DAT) and cold stress are two different methods designed to explore the temporal characteristics of the thermal patterns. Analysis of subtraction images and DAT techniques proposed by Anbar demonstrated the benefits of including temporal characteristics in the image analysis [4]. Although further research is required to establish its true value, DAT was demonstrated to potentially become the next new approach to more effective breast cancer detection. Unfortunately, conducting research on dynamic thermography has its challenges. Organized imaging

sessions designed specifically to facilitate dynamic thermography are required and thus may not be possible for some research environments. Furthermore, careful consideration must be given to prevent or isolate motion artefacts that may appear on the thermal images.

The potential of ANNs as a clinical decision making tool was reviewed for various imaging modalities but emphasis was given to breast thermography. In general, very few breast thermography research projects employed ANNs in their analysis of infrared images. Among the few, the research group of Jakubowska investigated applying wavelet transformation to 40 breast thermograms prior to the ANN analysis [49]. This technique showed some promising results, although additional samples are required for more extensive studies. Image analysis with ANNs was also investigated by Ng and colleagues with 200 patients from Singapore General Hospital [65]. In their method, thermal data and physiological information were used in different combinations as inputs to various ANNs. The best network tested gave a sensitivity of 61.54% and specificity of 40%. These levels of sensitivity and specificity are not sufficient to demonstrate that thermography can be considered a good method for early breast cancer detection. A more efficient method of analysis is required.

4 Infrared Image Analysis for Breast Cancer Detection

This chapter defines the problem statement of this thesis. Further subsections explain the methodology selected to solve the problem stated. Section 4.2 and section 4.3 describe the infrared images used for this thesis work and the software tool employed. Section 4.4 illustrates the image pre-processing procedures performed. The segmentation techniques applied to identify the ROI, the extraction of relevant parameters from the segmented images, and the ANN analysis performed with these parameters are explained in sections 4.5 to 4.7.

4.1 Problem Statement

This thesis tackles the problem of finding a quantitative method to analyze and evaluate breast infrared images to improve early breast cancer detection. The method's

effectiveness will be judged on whether sensitivity and specificity levels are equal to or better than that of mammography.

As mentioned in the literature review, statistical analysis of infrared images is widely conducted by researchers to extract image features. However, statistical parameters may not give a good numerical representation of the vascular patterns. This thesis investigates other potential quantitative parameters to identify vascular asymmetry between contralateral breasts.

Once representative parameters are generated, an analysis method is required to evaluate whether an image is normal or abnormal. This thesis evaluates the potential of ANNs to assess the images based on input parameters and to select the most relevant features in predicting clinical outcomes. A further objective was to minimize the number of input parameters required while maintaining the predictive power of the ANNs.

4.2 Research Data

Infrared images for this thesis work should ideally satisfy three conditions:

1. Clinical diagnosis for each image should be known, preferably verified by biopsy results or a different imaging modality in order to assess the accuracy of the ANN predictions.
2. All images should have the same origin to ensure that the quality and the characteristics of the images are comparable.
3. Sufficient number of images available for the ANN analysis to generate a more reliable conclusion.

At the time this study was conducted, in order to satisfy condition 1 and 3, infrared images of the patients had to be acquired from three different sources. A total of

49 images were available for the analysis. These images were taken by breast clinics in Moncton, Montreal, and New York. The names “Moncton images”, “Montreal images” and “New York images” will be used to address these groups of images. These three groups of images are described in detail in the following subsection.

4.2.1 Moncton Images

Thermal images from Moncton were recorded in 1984 by Dr. Monique Frize and her team in Moncton Hospital’s breast cancer screening clinic. The recording system consisted of the Agatronics Thermovision 680 Medical, a first generation thermographic camera and Agatronics OSCAR 780, an offline system for computer access and recording. To ensure the best results, data collection followed a rigid protocol summarized as follows: [69]

1. Tests were performed in the week following menstruation or between the 6th and 10th day of the cycle, up to the 13th day.
2. Patients were asked to refrain from alcohol, caffeine, pain medication lotions and smoking two hours before the test.
3. Chest area of the patient was cooled for approximately 10 minutes with a fan prior to the recording.
4. Room temperature was controlled at approximately 22°C and darkened during the test.

Procedures 3 and 4 were implemented to ensure optimum temperature contrast between the hot and cold areas on the body. Images taken were stored on a digital tape recorder for retrieval at a later time for analysis. Although thermal images from 86 patients were recorded at the breast screening clinic, only 19 images have known clinical

outcomes confirmed histologically and thus were included in this study. The clinical outcomes for the images were provided in four classifications: normal, benign, fibrocystic, and cancerous. However, due to the relatively small image database, and the fact that clinical diagnoses were reported in different ways by various centres where images were acquired, we grouped images into “normal” or “abnormal”. The 15 image outcomes previously classified as normal, benign or fibrocystic were considered “normal”, and the 4 image outcomes previously classified as cancerous were considered “abnormal”. The matrix for each image consisted of 128x128 pixels. Each pixel was represented by 8-bit value ranging from 0 to 255 in the grey scale.

For the 19 images used in this study, the actual temperature range of the thermographic camera was determined to be 10°C. Knowing the actual temperature range, the pixel values could be mapped to temperature values using the mean temperature value obtained from the camera as reference.

4.2.2 Montreal Images

Thirteen images were obtained from Dr. Keyserlingk at the Ville Marie Breast and Oncology Center in Montreal in 2002. The images were taken with a scanning-mirror optical system containing a mercury-cadmium-telluride detector (Bales Scientific, CA) with a spatial resolution of 600 optical lines [52]. The infrared images were recorded in a draft-free room with temperature controlled between 18°C and 20°C. Patients were requested to sit disrobed in the room for 5 minutes during the equilibrium period. For three hours prior to the imaging session, patients were asked to avoid alcohol, coffee, smoking, exercising, lotions and deodorant. Four images were recorded, each

from a different view namely the anterior, undersurface and two lateral views. Images generated were digitally stored on laser discs that could be retrieved for future analysis.

Abnormal Signs	
The following signs are considered abnormal signs, unless stable on serial imaging or due to known non-cancer causes.	
<ul style="list-style-type: none"> • Significant vascular asymmetry • Vascular anarchy consisting of unusual tortuous or serpiginous vessels that form clusters, loops, abnormal arborisation, or aberrant patterns. • A 1°C focal increase in temperature (ΔT) when compared to the contralateral site when associated with the area of clinical abnormality. • A 2°C focal ΔT versus the contralateral site. • A 3°C focal ΔT versus the rest of the ipsilateral breast when not present on the contralateral site. • Global breast ΔT of 1.5°C versus the contralateral breast. 	
Infrared Scale	
IR1	Absence of any vascular pattern to mild vascular symmetry
IR2	Significant but symmetrical vascular pattern to moderate vascular asymmetry, particularly if similar to prior imaging
IR3	One abnormal sign
IR4	Two abnormal sign
IR5	Three abnormal sign

Source: <http://www.villemariemed.com/Slideshow1/sld003.htm>

Table 4-1 Ville Marie Infrared (IR) Grading Scale

Images included in this study were taken from a collection of anterior view images with the pre-condition that they had been previously evaluated. For the Montreal images, actual outcomes comprise of assessments generated by a physician experienced in interpreting thermograms, mammograms, and sonograms. Each breast was graded on an infrared scale from IR1 to IR4 based on the presence of abnormal signs. For example, one patient may have left breast graded IR1 and right breast graded IR2. Table 4-1 shows the grading scale, which may also be found from the Ville Marie website. The

dimensions of the original images were 320x240 pixels. Each pixel has an 8-bit representation. The temperature range for the images is around 15°C.

4.2.3 New York Images

Infrared images from New York consisted of 17 images provided by Mr. DiCicco from Infrared Science Corporation (ISC). Images were collected at the ISC BreastScan IR™ station which consists of a digital infrared camera, a patient chair with infrared reflective side mirrors, a cool air source, and a video display where patients can see their images in real time [47]. Images used in this study were collected between year 2000 and 2004. Earlier images were recorded with the FLIR Systems Thermovision infrared camera whereas more recent images (2003 and 2004) were recorded with the FLIR system S-series camera. During the serial imaging session, temperature controlled air flow was directed at the breasts for four minutes while the camera recorded the surface cool down at a rate of 250 frames per second to generate an image sequence for analysis [5].

Static images provided for this study were extracted from the image sequences. Each image is 320x240 pixels with a thermal resolution of 0.05°C. The temperature range specified for each image is different and varies between 9.6°C and 16.6°C. Among the 17 patients, nine patients had cancer confirmed by biopsy or other diagnostic methods. The remaining eight patients were “normal”, but presented infrared images that had abnormal characteristics. To understand what “normal” patients with abnormal characteristics signify, the evaluation scheme used by BreastScan IR is briefly described. BreastScan IR employed a software algorithm that analyzed different parameters of the image sequences for abnormality and evaluated the risk based on the results obtained for

the various parameters. A cancerous patient and healthy patient should get a positive score and negative score respectively. The report also classified each measured sign of abnormality such as degree of asymmetry and temperature of focal hot spot into normal, equivocal, risk and abnormal. With respect to the eight “normal” patients in this study, it can be assumed that these patients were not diagnosed with cancer at the time images were collected, but their infrared images were given a positive score as several measured signs were classified to be at risk or abnormal. For this thesis, images from the nine patients with confirmed cancers were considered “abnormal” and images from the eight patients with only abnormal characteristics were considered “normal”.

4.3 Software Tools

Advanced computer technology has led to the creation of many powerful software packages to assist in data mining and analysis. For this thesis work, a software package that provides extensive image processing and ANN analysis tools is required. MATLAB was chosen as the software tool for this project for various reasons. Previously acquired experience and familiarity with the product was one of the main reasons. MATLAB also provides extensive data analysis functionality and embedded image processing tools essential to solve the problem. In addition, the help manuals, online technical support and newsgroups offered extra information for getting started and technical assistance.

4.4 Image Pre-processing

For the scope of this thesis, image pre-processing refers to the steps required to make images from multiple sources more compatible and hence improves the feasibility of analyzing the images collectively in the ANN analysis. During the preliminary

investigation phase, the group of 19 infrared images from Moncton were analyzed and evaluated with a backpropagation ANN. Although experimental results obtained were very encouraging and the work was published [54], drawing more reliable conclusions required a more thorough analysis with a greater set of images with their corresponding clinical results. Since the Moncton images were the smallest in size (128x128 pixels), the New York and Montreal images were cropped, resized and downsampled accordingly. In addition, raw data from New York needed to be converted into grayscale images with 8-bit representations, as this is the format of the Moncton and Montreal images. Figure 4-1 shows an example of original images from Moncton.

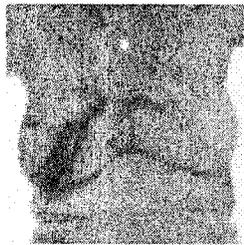


Figure 4-1 An example of Images collected at the Moncton Hospital's Breast Clinic

4.4.1 Image Conversion, Cropping and Resizing

Figure 4-2 shows one of the 17 images obtained from ISC in New York. The picture on the left of the figure is the original image displayed in colour. Each point on the original image is represented by double precision floating point numbers specifying the actual temperature values in two decimal points. The original image was cropped to isolate the chest region and then converted to an intensity image with 8-bit representation. The resulting image is shown to the right of the figure. Sizes of these images shown here are only approximations of their original dimensions.

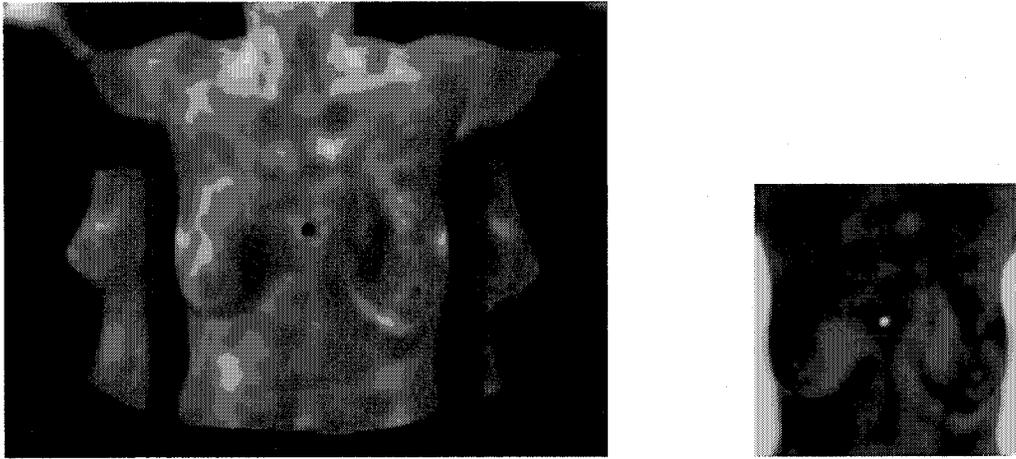


Figure 4-2 A Sample Image Obtained from ISC, New York.

As mentioned in section 4.2.2, images from Montreal used in this study were represented in a matrix of 320x240 pixels. With MATLAB toolbox function, the images were resized to 170x128 pixels. The resizing was done by bilinear interpolation, and since the resulting images would be smaller than the original images, a low pass filter was applied before the interpolation to reduce aliasing. The original ratio of the dimensions was kept to avoid any distortion of the images. An example of original images from Montreal and the corresponding resized image are shown in Figure 4-3. The dimensions of the images relative to each other are kept to illustrate the resizing operation.



Figure 4-3 A Sample Montreal Image (Before and After Resizing)

4.4.2 Image Enhancement

Poor contrast of an image may affect the analysis of infrared images as the ability to inspect small details and to detect subtle differences is compromised. The grayscale image in Figure 4-2 can be used to illustrate the point. In this case, the display of thermal patterns is not optimal mainly due to poor contrast of the image. This fact can easily be confirmed by inspecting the intensity histogram on the left in Figure 4-4.

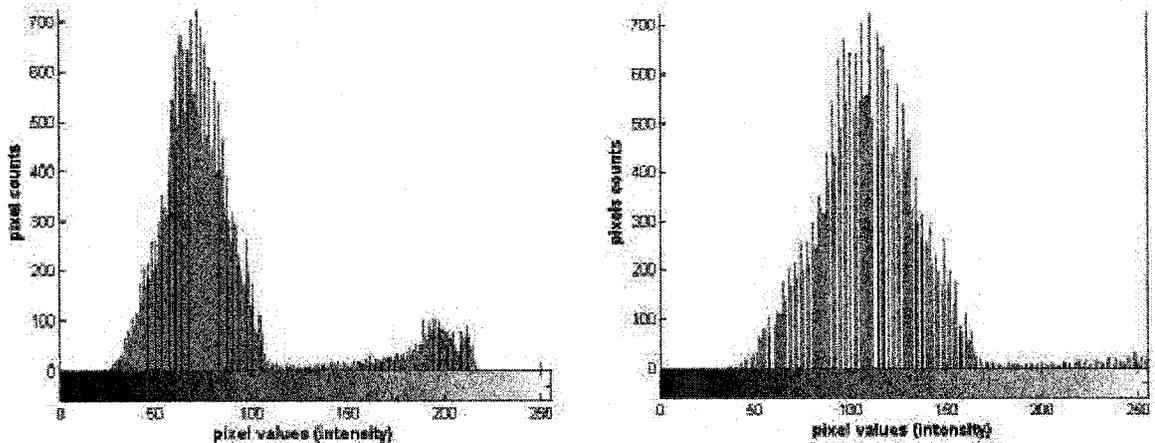


Figure 4-4 Image Histogram of a New York Image Before and After Intensity Adjustment

There are two distinct “groups” on the histogram of the pre-adjusted image. Pixels from the chest region form the group with a higher peak at a lower pixel value (lower intensity, darker on the image), while pixels from the image background form the second group. In fact, this characteristic of obvious intensity separation between the chest region and the background can be observed for all the New York images. These images can be enhanced by scaling the pixel distribution of the chest region to fill the entire intensity range, as illustrated in Figure 4-4. In this case, pixels with values between 0 and 165 were mapped to a new range of 0 and 255. Pixels that falls outside the region to be scaled (the input range) will be “clipped” at the maximum value of the output range, 255 in this case. After this operation, the chest region will have a much better contrast,

and the background will become completely white, since the pixel values are all clipped at 255. It should be emphasized that although the pixel distribution for the chest region is scaled to utilize the full dynamic range, the relative difference between pixel values remains the same. The original grayscale image from Figure 4-2 is replicated in Figure 4-5 together with the enhanced image to show the improvement. The white dot in the sternum area was an infrared target used for image recording placed by the New York team and does not reflect the actual body surface temperature.

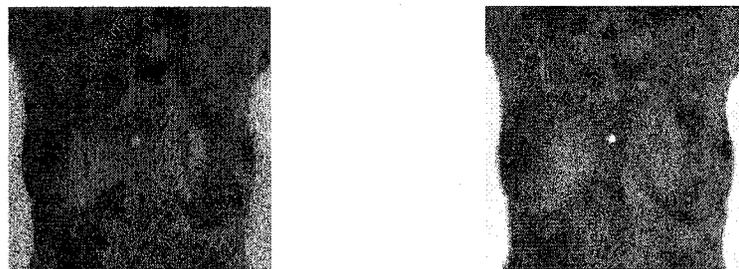


Figure 4-5 A Sample New York Image Before and After Intensity Adjustment

Similar intensity adjustment was required for the Montreal images. The histograms before and after the intensity adjustments are shown in Figure 4-6. A sample Montreal image before and after the adjustment is shown in Figure 4-7.

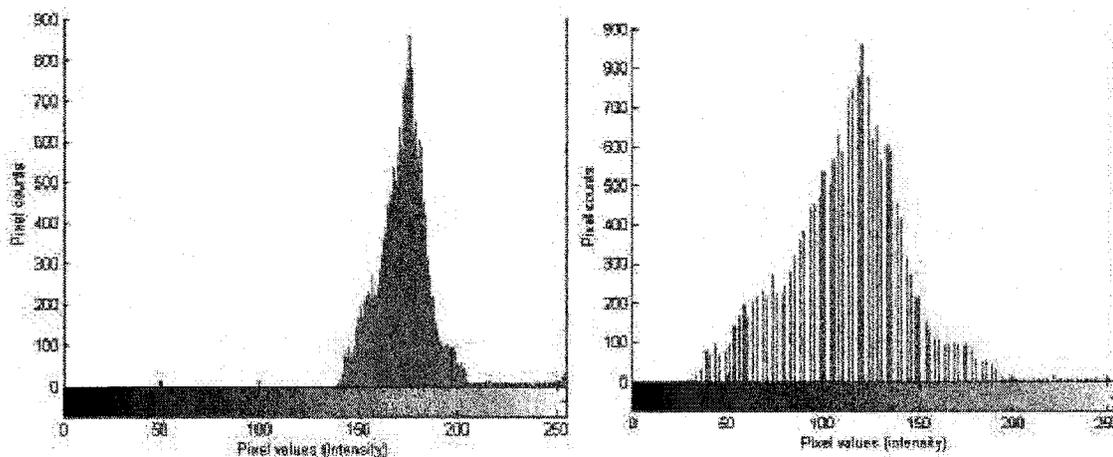


Figure 4-6 Image Histograms of a Montreal Image Before and After Intensity Adjustment

It can be observed that the contrast of the “before” image is lower than that of the New York image. Naturally one can expect that the pixels distribution for the chest region will be narrower, and thus the input range for the histogram “stretching” is also smaller. The output range is 0 to 255 as before.



Figure 4-7 A Sample Montreal Image Before and After Intensity Adjustment

4.5 Image Processing

Image processing of the infrared images involves identifying the ROI and dividing the ROI into quadrants.

4.5.1 Identifying Region of Interest

To automate the identification of the breast region on the image, an algorithm utilizing ellipse fitting of the breast region was tested. The ellipse fitting algorithm was originally presented in Halir and Flusser’s paper [39], and modified by Herry for the segmentation of the breast ROI [54]. In the algorithm, edge detection was first performed on the breast image. Based on the hypothesis that the centre of mass is most likely located within the breast region, the edges closer to the center of mass were considered more likely to belong to the breast area. Hence, a cut-off distance from the centre of mass was chosen to separate edges that are likely candidates of the breast region, and

those that were not. Finally, an ellipse was fitted to the smallest convex region that enclosed all the included edges.

Although for most images the ellipses generated by the algorithm approximate the breast regions quite well, it was not used in this thesis work for two reasons. Firstly, segmentation of the breast region was less than satisfactory for a number of images. Secondly, although some of the ellipses approximate the breast region quite well, they also include the “heat dam” just under the lower boundary of the breasts, and thus create an artefact of hyperthermic region on the surface of the breast.

Due to the fact that these images have been collected from multiple sources, image characteristics vary widely. Given this complexity and the small overall number of images available, manually segmenting the boundaries of the breasts ensured consistency of all images.

Once the breast region was identified, the breast was further divided to four quadrants using an approach suggested by Head and Lipari [42]. In this approach, the nipples, the chin, and the lowest contour point on each breast were used as reference points for further division. If the chin was not visible on the image, the uppermost pixel of the vertical symmetry line was used. Figure 4-8 clearly illustrated the lines drawn with the reference points to obtain the four quadrants. With this division, the quadrants on each side may be referred to with the terminology UOQ, UIQ, LOQ and LIQ, which is commonly used in mammography for positioning (see Figure 4-8). This method generated approximately symmetrical quadrants which are important for the asymmetry analysis described below.

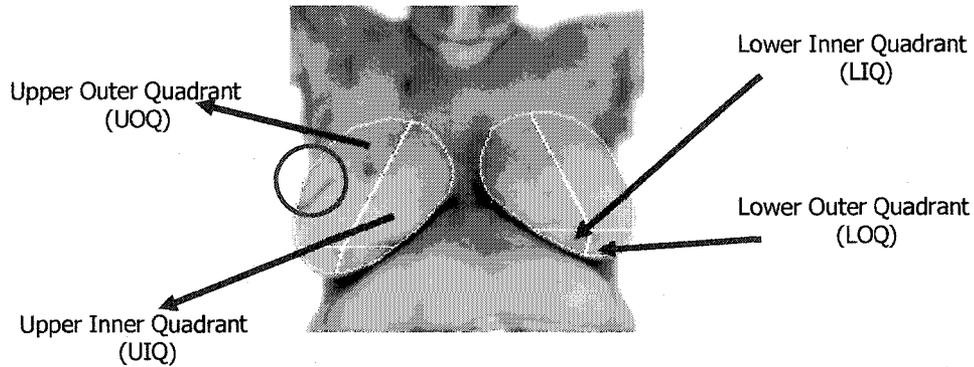


Figure 4-8 Segmented Breast Regions (showing an abnormal sign in the right UOQ)

4.6 Feature Extraction

In this section, methodologies for the extraction of quantitative parameters from infrared images are presented. Three separate techniques are developed to capture different thermal characteristics of the images.

4.6.1 Statistical Analysis

From the literature reviewed in Chapter 3, it is reasonable to say that statistical analysis has been applied extensively to study infrared images. First order statistics such as mean and median, second order statistics, either standard deviation or variance are the most commonly analyzed parameters. Higher order statistics such as skewness and kurtosis may be equally useful to describe asymmetrical thermal patterns and presence of hot spots. In this study, statistical analysis was performed on the four quadrants of each breast. For each quadrant, statistical parameters extracted included mean, median, minimum, maximum, interquartile range, standard deviation, variance, skewness, kurtosis, heat content and entropy.

Calculating the statistical parameters per quadrant enables more localized characterization of thermal pattern. The importance of this capability is illustrated in Figure 4-9 [77]. It can be observed that the percentage of tumour occurrence varies widely from quadrant to quadrant. Recent research reports that 48% to 54% of breast tumours were found in the UOQ of the breast [19].

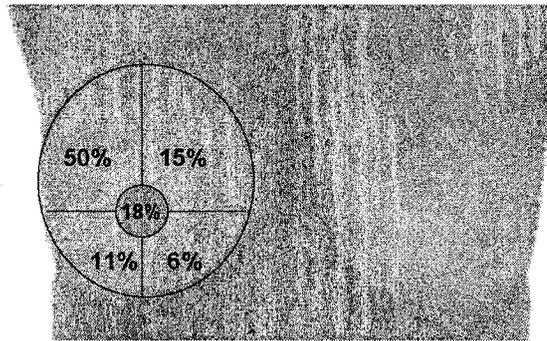


Figure 4-9 Percentage of Breast Tumour Occurrence by Quadrants

The most common parameter used to detect hyperthermia, the mean temperature, is not as discriminative when the temperature values average over a larger region instead of a more localised area. Considering the infrared images of one cancerous patient shown in Figure 4-8, a tumour was detected at the spot indicated by the red circle. It is unlikely that the “hot vein” would be detected if the mean temperature of the entire right breast is compared with the mean temperature of the left breast.

In evaluating thermograms, a visual inspection of the breast images usually involves a symmetry analysis of the contralateral breasts. To facilitate the quantitative asymmetry analysis, a set of *difference statistics* is computed by taking the difference of statistical values between the contralateral breasts. Hence, the difference statistics for

each image can be seen as an $m \times 4$ matrix, where m is the number of statistical parameters obtained for each quadrants.

4.6.2 Quadtree decomposition analysis

In the problem statement of the thesis, it was mentioned that one of the thesis objectives is to develop a quantitative parameter that could give a better representation of vascular asymmetry between contralateral breasts. A review of state of the art (see section 3.3) revealed that statistical analysis is the most widely applied technique to extract thermal features from the infrared images. However, statistical parameters do not provide a complete representation of symmetry analysis of the breast images. Two regions that look completely different may have the same statistics, as illustrated in Figure 4-10. This is due to the fact that relative locations of the pixels on the image are not taken into account during statistical analysis. Other quantitative parameters that can potentially be more representative of the thermal patterns asymmetry are required. This section of the thesis presents an algorithm to extract such a parameter through quadtree decomposition analysis.

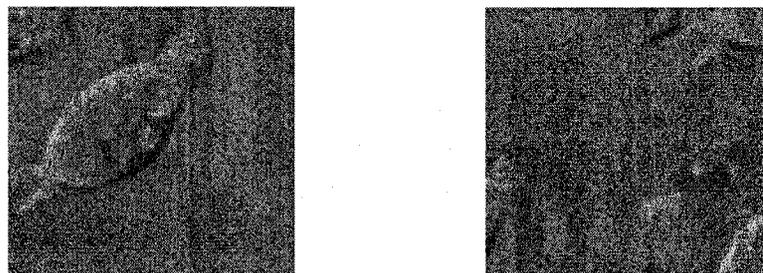


Figure 4-10 Two Pictures with the Same Statistics

4.6.2.1 Quadtree Decomposition Implementation in MATLAB

Quadtree decomposition analysis can be performed to assess the homogeneity of an image. It is commonly used in image compression algorithms to obtain information about the structure of the image. In Matlab, quadtree decomposition analysis is performed as follows:

1. A square image is divided into four equal-sized blocks.
2. Within each block, the algorithm checks if all pixels are within a specified dynamic range, the criterion for image homogeneity.
3. If the criterion is not met, the block is further divided into four equal-sized blocks.
4. Step 2 and 3 is iterated until all blocks meet the homogeneity criterion.

The criterion for image homogeneity requires that the difference between the minimum pixel values and the maximum pixel values must be smaller than a specified threshold. For a thermal intensity image with 8-bit representation, the criterion can be summarized as follows:

$$\text{Max}(\text{block}) - \text{Min}(\text{block}) < \text{threshold} * 255$$

The degree of homogeneity tested can be adjusted by changing the threshold of the test criterion. Since the threshold determines the dynamic range of the pixels allowed in a block, the greater the threshold, the less homogeneous the block will be (pixel values within a block will have a larger variation).

Because each block that does not meet the criterion will be further subdivided into four equal-sized blocks, the input image to the algorithm will be most efficiently analyzed if the dimensions of the image are a power of 2. For these images, the

dimensions of the smallest block possible will be 1x1, containing a single pixel. After the quadtree decomposition analysis, a fairly homogeneous image will have more large blocks and fewer small blocks in comparison to a “busy” image. The concept of quadtree decomposition analysis is shown in Figure 4-11.

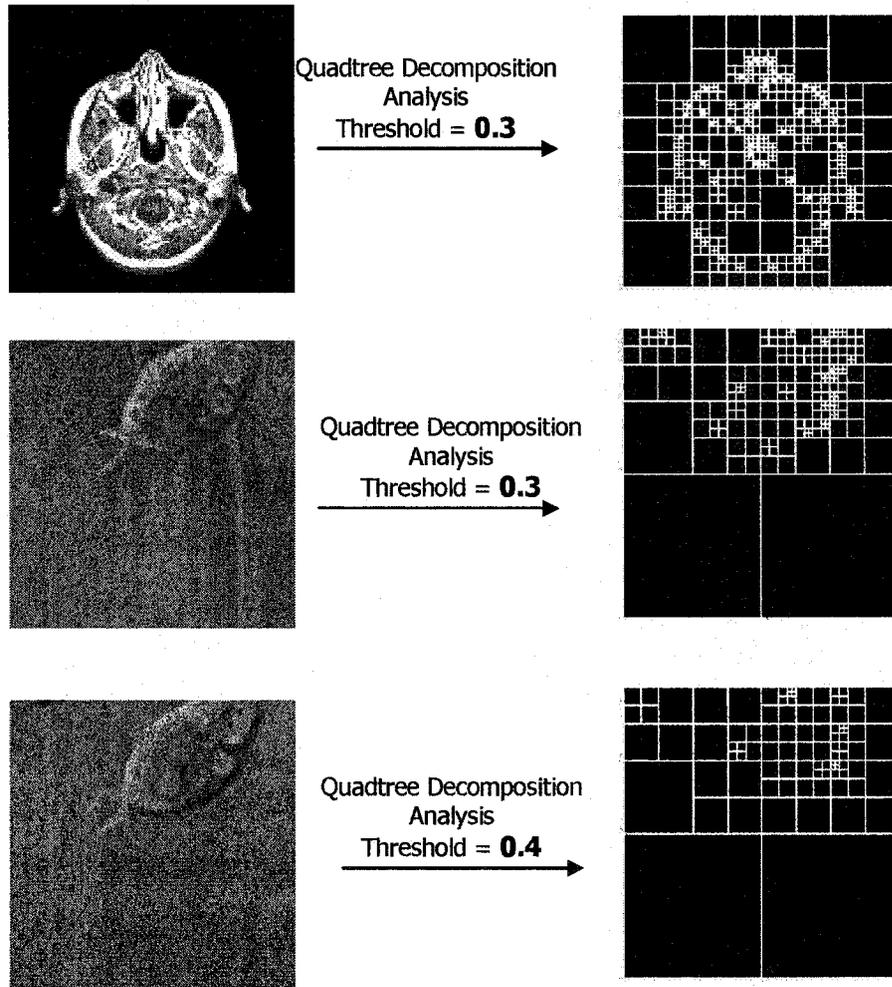


Figure 4-11 Quadtree Decomposition Analysis with Two Images and Two Thresholds

4.6.2.2 Algorithm for Extracting Quadtree Decomposition Parameter

The MATLAB image processing toolbox provides simple functions to perform quadtree decomposition operations. Utilizing these MATLAB functionalities, a MATLAB program was written to implement an algorithm to analyze the breast images with quadtree decomposition techniques. The algorithm designed and developed for this thesis work is summarized as follow:

1. The segmented breast regions were split into two separate images, consisting of the right and left breast.
2. Each image was padded so that the dimensions of the image were a power of two to maximize the number of divisions that can be performed.
3. Quadtree decomposition analysis was performed on each image individually.
4. For each image, the total count of blocks for each size was recorded.
5. The Quadtree Decomposition Parameter (QDP), which would be used as an input feature to the ANN, was obtained by summing the differences in total block counts between contralateral breasts, as shown in equation (4-1).

$$\mathbf{QDP} = \text{sum} \{ \text{diff} (2 \times 2 \text{ block}), \text{diff} (4 \times 4 \text{ block}), \text{diff} (8 \times 8 \text{ block}) \} \quad (4-1)$$

To facilitate more accurate comparison between the vascular patterns of the contralateral breasts, differences between the numbers for 1x1 blocks and blocks larger than 8x8 were not used. 1x1 blocks were excluded since they were made up of mainly the boundary pixels between the segmented breast regions and the padded background, and hence gave a poor representation of the inner vascular patterns. On the other hand, blocks larger than 16x16 represented mostly the padded background, and the more

homogeneous regions of the breast. Since vascular patterns were feature of interest in this analysis, the larger blocks were excluded as well in the QDP calculation.

Figure 4-12 shows a typical result from quadtree decomposition analysis for the breast image shown on the left of the figure. Squares drawn around the segmented region do not reflect the actual dimensions. The image block diagrams illustrate the result obtained from the quadtree decomposition analysis. In this diagram, the bigger squares map to the more homogeneous regions of the image whereas the smaller squares correspond to region showing significant temperature change. Two red circles were drawn in the figure to indicate the concentration of smaller squares around the “hot vein”.

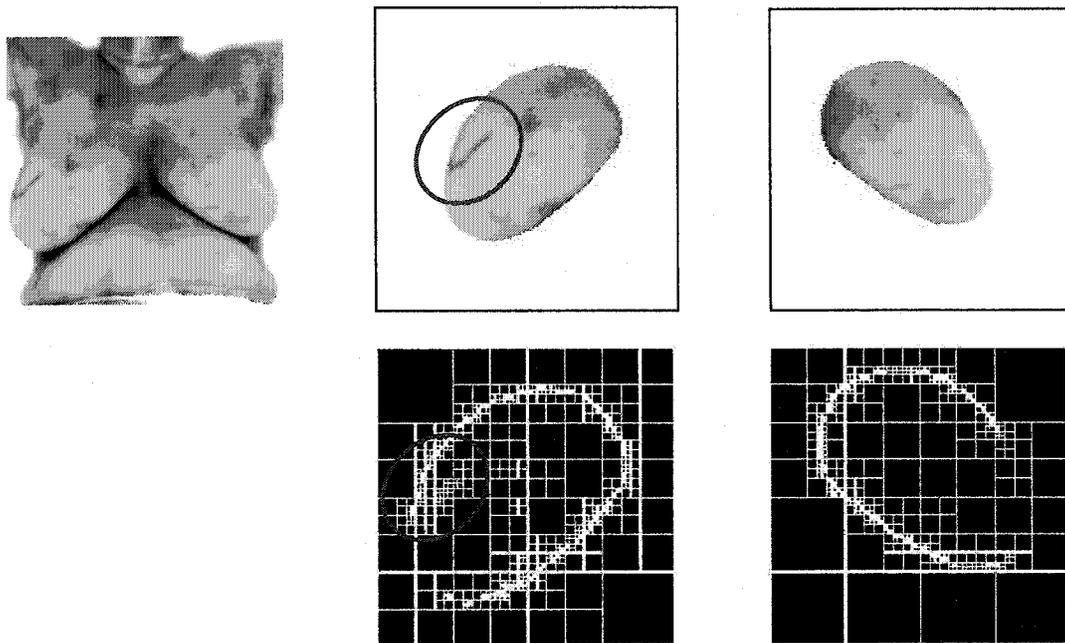


Figure 4-12 Quadtree Decomposition Analysis of a Breast Image

4.6.2.3 Threshold Selection

The parameter QDP is dependent on the threshold selected for the quadtree decomposition analysis. The selection of the value of a threshold is based on the contrast of an image, which was enhanced through histogram scaling described in section 4.4.2 and the level of vascular patterns visible on the thermal images, which in turn differs from patient to patient, and depends on the thermal resolution of the infrared camera. A pre-selected threshold did not perform well in capturing the vascularity difference of all the images in this study due to the varying characteristics of these images acquired from multiple sources. An approach that accounted for the dissimilarity between the image sets was required. For each image group, the resulting quadtree diagrams from different thresholds were inspected with a trial and error approach and an optimal threshold was determined by evaluating the performance of different thresholds in identifying abnormal patterns on the images. Thresholds selected for the Moncton, Montreal and New York images were 0.4, 0.3 and 0.35 respectively. With the optimal thresholds, QDP values were calculated as described in the previous section and included as an input feature of the ANN analysis.

4.6.3 Edge detection analysis

To complement quadtree decomposition analysis, an edge detection analysis was employed to extract an additional parameter that may be more indicative of vascular asymmetry than statistical parameters. A canny edge detector was used to detect any significant temperature gradient change on the surface of the breast. Thresholds for the edge detection can be adjusted so that only “strong” edges that might correspond to neovascularization (vascular hot spots) are detected.

4.6.3.1 Algorithm to Extract Edge Detection Ratio

The algorithm developed for feature extraction based on the edge detection technique is summarized in Figure 4-13. After edges of the segmented breast region are detected, edges corresponding to the boundaries of the segmented regions are removed since they do not represent the thermal patterns of the breast image. After the boundaries are removed, the number of edge pixels is counted for each image. An “edge detection” Ratio (EDR) was used as an input to the ANN, calculated with equation (4-2).

$$\text{EDR} = \frac{\text{Absolute (right breast edge pixels - left breast edge pixels)}}{\text{Average (right breast edge pixels, left breast edge pixels)}} \quad (4-2)$$

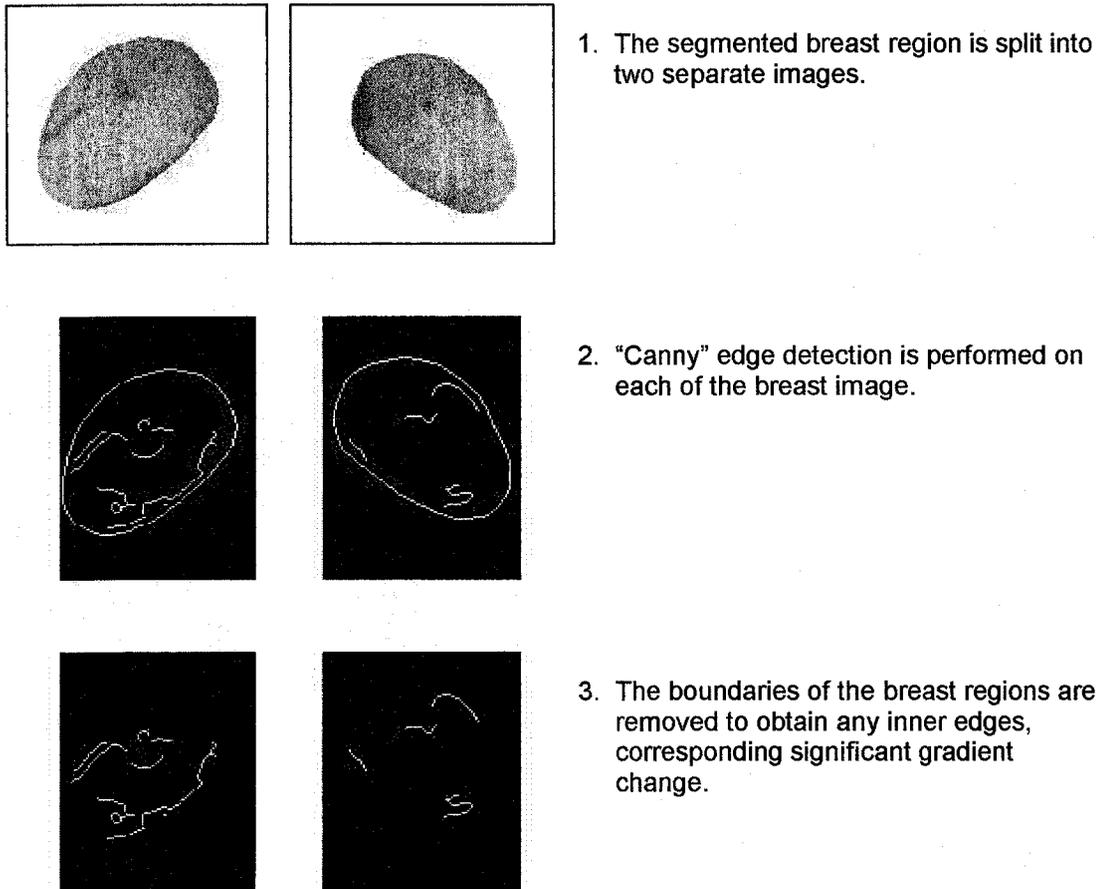


Figure 4-13 Edge Detection Analysis for Parameter Extraction

4.6.3.2 Selection of Edge Detection Thresholds

The Canny edge detection method uses two thresholds to detect strong and weak edges on the image. Implementing the Canny edge detection in Matlab, these thresholds may be automatically chosen or explicitly specified. To achieve a consistent comparison across all the images, it was desirable to specify a single set of high and low thresholds for all images. However, due to the issue with multiple sourced images, the thresholds had to be individually selected for each group of images.

In addition to the threshold selection, calculations of the EDR were also modified for images obtained in Moncton. In general, images collected in Moncton showed less vascularized thermal patterns, probably due to the lower thermal resolution of the camera used. The formula used to calculate EDR, shown in equation (4-3), was always 200% if the edge pixels count was zero for any one of the breasts. It was therefore decided that instead of dividing the difference (the numerator) with the average (denominator), the difference was used directly as an input parameter to the neural network. Considering that the pixels count was generally lower for the Moncton images compared to the Montreal and New York images, this modification was actually a feasible solution. After calculating several different sets of EDR values for different edge detection thresholds, the best-performed thresholds selected based on the ability to reflect actual clinical diagnoses were [0.12, 0.225] for the Moncton images and [0.06, 0.15] for both the Montreal and New York images. The two numbers in the brackets indicate the high and low thresholds that Canny edge detection technique uses to detect strong and weak edges.

4.7 Feature Analysis

This section presents a quantitative asymmetry analysis performed with parameters obtained from feature extraction techniques discussed in the previous section. The extracted parameters were used as input features for the neural network analysis. The ANN was trained and validated to predict clinical outcomes.

4.7.1 Architecture of ANN

Analyzing the breast thermograms and discriminating the abnormal from the normal images is a typical pattern classification problem. Various classes of ANNs with different learning algorithms and varying degrees of complexity have been applied to pattern classification problems. Among these network classes, Freeman and Skapura found the backpropagation (BP) network to be especially useful in addressing complex pattern recognition problems [25].

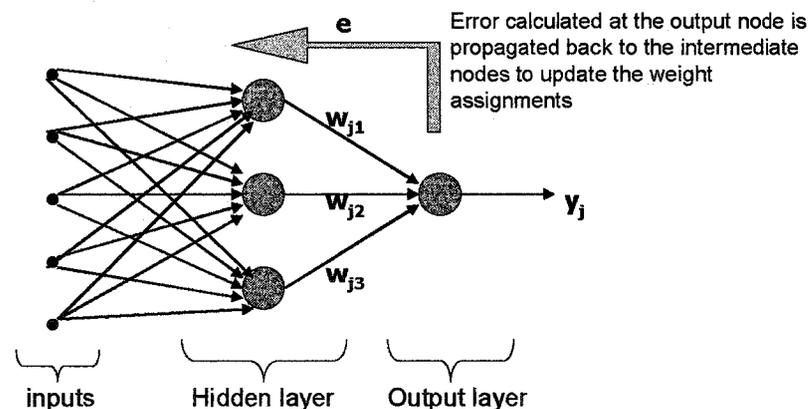


Figure 4-14 A Three-layer Backpropagation Network

4.7.1.1 Backpropagation Neural Networks

A BP neural network is a multilayer, fully connected, feedforward network that uses supervised learning to model a system, as shown in Figure 4-14. The main

characteristic of BP networks is that during each training cycle, errors calculated at the output nodes are propagated back through the network to update the connection weights. The basic learning algorithm of BP networks uses the gradient descent procedure where the steepest descent on the error “surface” is taken to find the minimum. It is possible for the search to stop at a local minimum instead of the global minimum, and therefore causing the network not to converge.

The BP neural network was implemented with MATLAB Neural Network Toolbox built in functions. In MATLAB, a set of predefined training algorithms is developed for BP networks. Each training algorithm has different characteristics in speed of convergence and memory storage requirement. The backpropagation training algorithms employed in the analysis were the Resilient Backpropagation and the Levenberg-Marquardt (LM). The Resilient Backpropagation (RP) algorithm was described as the fastest algorithm for pattern recognition problems [64]. The LM training algorithm was chosen as it was described as the training algorithm that yields the fastest convergence and lowest mean square error (MSE), especially for small to moderate-sized networks. It is also claimed to have the most efficient implementation in MATLAB.

For BP networks, the transfer function used must be fully differentiable due to the gradient (partial derivatives) descent search method. In this study, the transfer function used for all the neurons was the tan-sigmoid function as shown in Figure 4-15. All inputs and outputs were normalized to be in the range of -1 and 1.

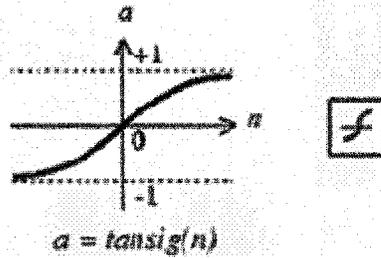


Figure 4-15 Tan-Sigmoid Transfer Function³

4.7.1.2 Selection of Input Nodes

Due to the fairly small sample size available for this analysis, the number of network nodes must be limited to a size that could be supported by the training data. A general rule of thumb requires five to ten training datasets for each node in the network [46]. There are two approaches that can be taken to achieve the goal of limiting network nodes. The first approach is to reduce the network connectivity by reducing the number of hidden layers. The second approach is to reduce the number of input nodes, or in other words, the input features to be analyzed. The input vector to the neural network consisted of features extracted from the infrared images. The size of the input vector may be reduced in a number of ways. The method employed in this thesis work was eliminating features which are highly correlated. The correlation coefficients between any two features can be obtained by performing correlation analysis.

As discussed in sections 4.6.1, 4.6.2, and 4.6.3, eleven statistical parameters and two potential thermal vascularity indicators were extracted from the images. Due to the limited number of images available for analysis, only a subset of these parameters can be used as input features. Correlation analysis was performed in SPSS statistical software to

³ Figure is reproduced from Mathworks website.
<http://www.mathworks.com/access/helpdesk/help/toolbox/nnet/backpr53.html>

investigate the relative correlation among the parameters. Highly correlated parameters explain the same variations within the data set and thus are redundant. Results obtained from the bivariate correlation analysis between the statistical parameters can be found in the Appendix. The correlation tables were used as a reference when the selection of input features was made.

4.7.1.3 Selection of Hidden Nodes

Once the input features to be analyzed were selected, the optimal number of nodes in the hidden layer was obtained through a trial and error process. These guidelines were followed to ensure that the network was sufficiently trained but not overfitted to the training sets.

1. If training does not meet the performance goal or does not converge, the number of hidden nodes is increased.
2. If training meets the performance goal but the output generated is not consistent, the number of input parameters is decreased.
3. If training meets the performance goal and the output generated for the same image is consistent, the number of hidden nodes is decreased to obtain the smallest network achievable.

Once the network architecture that achieved the best performance during the training sessions was established, the network was tested with the validation datasets to obtain the overall efficiency of the ANN.

4.7.2 Performance Measures

In clinical settings, the effectiveness of a method is often evaluated by three performance measures: sensitivity, specificity and the accuracy in predicting clinical diagnoses. The accuracy is determined by the correct classification rate (CCR) when ANNs are employed for outcome predictions. In addition, the positive predictive value (PPV) and the negative predictive value (NPV) may be used as well. These numbers are obtained during the validation stage of the ANN analysis after the network has been properly trained.

Due to the limited number of infrared images available for ANN analysis, an iterative process was developed to obtain the optimal number of hidden nodes to calculate the performance measures. To determine the optimal number of hidden nodes following the guidelines described in the previous section, this iterative process was initially executed to obtain an overall convergence rate for the network architecture. If the convergence rate was too low, the network was not sufficiently trained, probably due to insufficient number of hidden nodes. If the convergence rate was 100% or close to 100%, there was a high risk of network overfitting. Hence, in this methodology, a convergence rate of approximately 75% for the network was taken to be optimal. This convergence rate was chosen as a compromise between a sufficiently trained network and a network whose size was limited by the small set of training data available in this analysis.

Once the number of hidden nodes was determined, a second iterative process was performed to train and validate the network. The iterative training and validating procedure is summarized as follow:

1. From the 49 datasets, randomly choose 12 datasets to be used for validation, with the remaining 37 datasets to be used for training.
2. Verify that the training sets are a good representation of the entire sampling population by checking the outputs of the images to ensure that the proportions of the normal and abnormal images included are approximately the same as the sampling population. If the condition is not satisfied, go back to step 1.
3. Train the network with the training datasets. If the network does not converge during the training, go back to step 1. As mentioned earlier, the target convergence rate for the network was 75%.
4. Simulate the network with the validating datasets to generate outputs (predicted clinical outcomes)
5. Compare the ANN outputs for the validating datasets with their corresponding actual clinical diagnosis to obtain the sensitivity, specificity, PPV, NPV and CCR for the network.
6. Reiterate Step 1 to Step 5.

The procedure is similar to k-fold cross validation but each fold is an independently chosen subset of the sample. The performance measures used to evaluate the effectiveness of the method were the average values calculated from results obtained in all the iterations. These values of sensitivity, specificity, PPV, and NPV are compared to that of mammography to determine whether breast thermography is an effective method for early breast cancer detection.

5 Results and Discussions

This chapter presents the results obtained from various analyses described in the methodology. Results from the correlation analysis are presented and the derivation of statistical input parameters to the ANN is explained. Quantitative parameters extracted in the quadtree decomposition and edge detection analyses are discussed, and their abilities to classify the images based on thermal vascularity are examined. Finally, results obtained in the ANN analyses with various network configurations and input parameters are presented. The chapter concludes with a discussion of the results.

5.1 Input parameters

5.1.1 Statistical Parameters

Due to the limited number of images available for the ANN analysis, the size of the input vector was reduced to restrict the number of hidden nodes required in the neural network. In selecting a subset of the statistical parameters, correlation analysis was performed for the difference statistics obtained from each quadrant. Although results differed slightly from quadrant to quadrant, they were fairly consistent and gave a good indication of the overall correlations among the parameters. The results are summarized and shown in Table 5-1. More detailed results are included in the Appendix.

The correlation analysis results showed that mean, standard deviation and kurtosis are the least correlated parameters. Skewness was found to be highly correlated with mean in two of the quadrants and with kurtosis for the other two quadrants. It was therefore not included for the ANN analysis. As a result, mean, standard deviation and kurtosis formed the subset of statistical parameters to be used as ANN input parameters. Before these parameters were used in the neural network analysis, they were adjusted according to the original temperature range of each image, and the intensity scaling used during the image enhancement.

Non-parametric Bivariate Correlation Analysis	
Mean, Median, Standard Deviation, Interquartile Range, Skewness, Kurtosis, Minimum, Maximum, Entropy and Heat Content.	
	Highly Correlated With
Mean	Median, Minimum, Heat Content, <i>Skewness</i>
Standard Deviation	Interquartile Range, Entropy
Kurtosis	Maximum, <i>Skewness</i>

Table 5-1 Non-parametric Bivariate Correlation Analysis Results Summary

Once the statistical parameters to be used in the ANN analysis were determined, four sets of numbers from the four quadrants were computed based on the following formulas to obtain the input parameters: mean, standard deviation and kurtosis.

Mean:

$$\frac{\text{abs}(Q1_mean_diff) + \text{abs}(Q2_mean_diff) + \text{abs}(Q3_mean_diff) + \text{abs}(Q4_mean_diff)}{4}$$

Std_dev:

$$\frac{\text{sqr}(Q1_std_diff) + \text{sqr}(Q2_std_diff) + \text{sqr}(Q3_std_diff) + \text{sqr}(Q4_std_diff)}{4}$$

Kurtosis:

$$\frac{\text{abs}(Q1_kurt_diff) + \text{abs}(Q2_kurt_diff) + \text{abs}(Q3_kurt_diff) + \text{abs}(Q4_kurt_diff)}{4}$$

Q1_mean_diff - mean difference calculated for quadrant 1 between the contralateral breasts. This naming convention is followed for all the variables. Figure 5-1 shows the numbering of each quadrant.

“abs” - absolute value of the variable

“sqr” - square of the variable

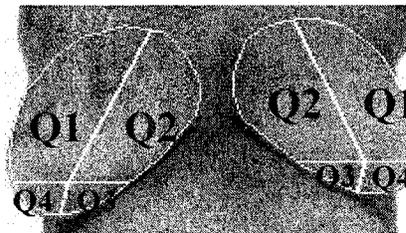


Figure 5-1 Breast Quadrants Numbering

5.2 QDP and EDR as indicators of vascular asymmetries

As described in sections 4.6.2 and 4.6.3, algorithms for quadtree decomposition analysis and edge detection analysis were developed to extract quantitative parameters QDP and EDR that may better indicate vascular asymmetry between contralateral breasts which otherwise were not indicated by statistical parameters. In this section, the ability of these parameters to differentiate the abnormal images from the normal images is presented.

5.2.1 Performance of QDP

Table 5-2 shows the results from the quadtree decomposition analysis. The second column shows the actual diagnosis of each image. Rows containing images with positive clinical diagnoses are shaded in grey. The images are sorted by QDP values in ascending order to show the correlation between positive clinical diagnoses and high QDP values.

In order to evaluate the performance of QDP in classifying images, measures of accuracy were calculated for a pattern classification method using the QDP as the only discriminating feature. On the table, a red line was drawn between the QDP values of 33 and 34. If these values were used as thresholds to classify the images into normal and abnormal, the correct classification rate (CCR), sensitivity and specificity for this analysis would be 0.82, 0.71 and 0.86 respectively.

Image	Diagnosis	QDP
BT2_1107	0	1
BT2_1152	0	3
52702	0	6
NY008	0	6
NY004	1	8
52102	0	9
BT1_1127	0	11
51702	0	13
BT3_1100	0	14
52202	0	14
BT2_1028	0	15
BT3_1030	0	15
BT1_1102	0	16
BT2_1132	0	17
NY009	0	17
Test8	0	18
51802	0	19
BT1_1132	0	22
BT2_1102	0	23
NY002	1	23
BT2_1127	0	24
51602	0	24
52002	0	25
Test92	0	25
BT2_1082	1	26

Image	Diagnosis	QDP
BT2_1112	0	26
51402	0	26
52302	0	26
BT1_1025	0	27
NY003	1	27
NY001	0	28
Test6	0	28
BT1_1107	0	32
51902	0	33
NY006	1	34
Test7	0	34
52502	0	35
NY011	1	35
BT3_1015	1	37
BT3_1090	1	38
51502	1	42
Test5	0	45
NY012	1	48
NY010	1	49
51302	0	57
NY005	1	68
BT2_1008	0	74
BT3_1070	1	100
NY007	1	125

Table 5-2 Results from the Quadtree Decomposition Analysis

$$\begin{aligned}
 \text{CCR} &= \text{numbers classified correctly} / \text{total number of images} \\
 &= 40 / 49 \\
 &= \mathbf{0.82}
 \end{aligned}$$

$$\begin{aligned}
 \text{Sensitivity} &= \text{numbers classified correctly as positive} / \text{numbers of true positives} \\
 &= 10/14 \\
 &= \mathbf{0.71}
 \end{aligned}$$

$$\begin{aligned}
 \text{Specificity} &= \text{numbers classified correctly as negative} / \text{numbers of true negatives} \\
 &= 30/35 \\
 &= \mathbf{0.86}
 \end{aligned}$$

5.2.2 Performance of EDR

Table 5-3 shows a similar table to present the results from the edge detection analysis. The images were sorted by EDR values in ascending order.

Image	Diagnosis	EDR
BT1_1025	0	0
BT1_1127	0	0
BT2_1008	0	0
BT2_1028	0	0
BT2_1112	0	0
BT2_1127	0	0
BT2_1132	0	0
Test6	0	1
51402	0	2
52202	0	2
52002	0	3
BT1_1107	0	4
52702	0	5
Test8	0	6
BT1_1102	0	7
52302	0	10
BT3_1015	1	10
NY001	0	10
51802	0	13
BT2_1152	0	15
BT2_1102	0	16
NY004	1	17
NY006	1	18
51302	0	20
BT1_1132	0	21

Image	Diagnosis	EDR
BT3_1100	0	21
BT2_1107	0	26
BT2_1082	1	27
NY005	1	30
NY002	1	42
Test92	0	46
NY009	0	51
Test5	0	53
NY011	1	54
NY007	1	55
NY008	0	58
52502	0	61
51702	0	72
NY012	1	82
51602	0	86
Test7	0	91
BT3_1030	0	96
NY003	1	103
52102	0	104
BT3_1070	1	107
BT3_1090	1	111
NY010	1	122
51902	0	139
51502	1	153

Table 5-3 Results from the Edge Detection Analysis

$$\begin{aligned} \text{CCR} &= 35 / 49 \\ &= \mathbf{0.71} \end{aligned}$$

$$\begin{aligned} \text{Sensitivity} &= 11 / 14 \\ &= \mathbf{0.79} \end{aligned}$$

$$\begin{aligned} \text{Specificity} &= 24 / 35 \\ &= \mathbf{0.68} \end{aligned}$$

To calculate the measures of accuracy for EDR, a red line was drawn between EDR values of 26 and 27 to separate images obtained from cancerous patients and images obtained from healthy patients. The same formulas as shown in the previous section were used. For EDR used as the only discriminating feature, the CCR, sensitivity and specificity of the method were 0.71, 0.79 and 0.68 respectively.

5.2.3 Discussions on Performance of QDP and EDR

Based on measures of accuracy calculated, QDP appeared to be a more effective feature in classifying images on the basis of thermal vascular asymmetry. Although the sensitivity of the EDR method was slightly higher, results obtained for EDR were overall less satisfactory. Although the algorithm implemented to extract EDR values from the images was relatively simple, it was also more sensitive to image noise and different characteristics of the multiple sourced images due to the nature of edge detection techniques. If the technique was to be used, a more sophisticated approach over counting edge pixels will be required to overcome the problems stated above.

On the other hand, the sensitivity, specificity and CCR measurements obtained in quadtree decomposition analysis demonstrated that QDP was quite effective as a quantitative indicator of thermal vascular asymmetries in classifying thermal images. To increase the potential of QDP as a discriminating feature, more improvements are needed to increase the sensitivity of the feature in analyzing less vascular but yet cancerous images.

5.3 ANN Analyses Results

Quantitative analysis of the image features was performed with the ANN using different network configurations and input parameters. This section summarizes the significant experimental results extracted from the various configurations tested.

5.3.1 Performance of the Various Network Configurations

As mentioned in the methodology, two training algorithms provided in the MATLAB Neural Network toolbox were investigated in this study. They were the Resilient Backpropagation (RP) algorithm and the Levenberg-Marquardt (LM) algorithm. All network trainings began with five input parameters. Appropriate adjustments were then made to the number of hidden nodes and the number of input parameters based on results obtained. The training parameters common to both training algorithms are shown in Table 5-4.

Performance goal	0.1
Learning Rate	0.3
Minimum Gradient	$1e^{-9}$
Maximum Epoch	1000

Table 5-4 Training Parameters for the RP and LM Training Algorithms

The *performance goal* is the mean square error between the network generated outputs and the desired outputs provided with the training datasets. The *learning rate* is a multiplier which determines the magnitude of change in network weights and biases during each training cycles. With a low learning rate, the algorithm takes smaller steps when updating the weights and biases, and thus may take longer to converge. However, a learning rate that is set too high may result in an unstable network. The *minimum gradient* is a stopping criterion used to discontinue the training cycles if the gradient of

the error function falls below a specified value. When the gradient decreases to a very small value, change in network weights and biases will not be significant enough to produce any appreciable effect on the overall network performance. An *epoch* is a single pass through all training inputs and target vectors.

Selected results obtained from the ANN analysis with various configurations are summarized in Table 5-5. In addition to five input parameters, different combinations of three and four input parameters were also tested. Only the more effective combinations with better convergence and correct classification rate (CCR) are shown in Table 5-5. The numbers shown in the columns of convergence rate and CCR were averages calculated with numbers obtained from numerous training iterations. The total number of training iterations performed typically ranged from one to two hundred. For each iteration, the network was trained with 37 randomly chosen training input vectors and validated with the 12 remaining input vectors.

Parameters	# of Parameters	# of Hidden Nodes	Convergence Rate		CCR	
			RP	LM	RP	LM
Mean, Std_dev, Kurtosis, QDP, EDR	5	8	0.88	0.84	0.68	0.67
	5	7	0.81	0.74	0.69	0.68
	5	6	0.76	0.77	0.68	0.68
Mean, Kurtosis, QDP, EDR	4	8	0.78	0.69	0.72	0.70
	4	7	0.76	0.70	0.73	0.70
	4	6	0.63	0.6	0.70	0.70
Mean, QDP, EDR	3	7	0.73		0.72	
	3	6	0.66		0.74	

Table 5-5 Results of ANN Analysis with Various Network Configurations

Figure 5-2 shows typical learning curves of the ANN implementing RP algorithms and LM algorithms. The two images on the left show the network converged

to eventually meet the performance goal, whereas the images on the right show networks that failed to converge. In general, the RP algorithm converged much slower than the LM algorithm. The RP networks that failed to converge most likely stopped training at the maximum epoch whereas the LM network stopped at the minimum gradient. Although networks trained with the LM algorithms yield faster convergence, the network CCR was slightly lower than networks trained with the RP algorithm. This observation was as expected since the RP training algorithm is described to have the best performance in pattern classification problems while the LM algorithm is generally more efficient in function approximation problems than in pattern recognition problems [64].

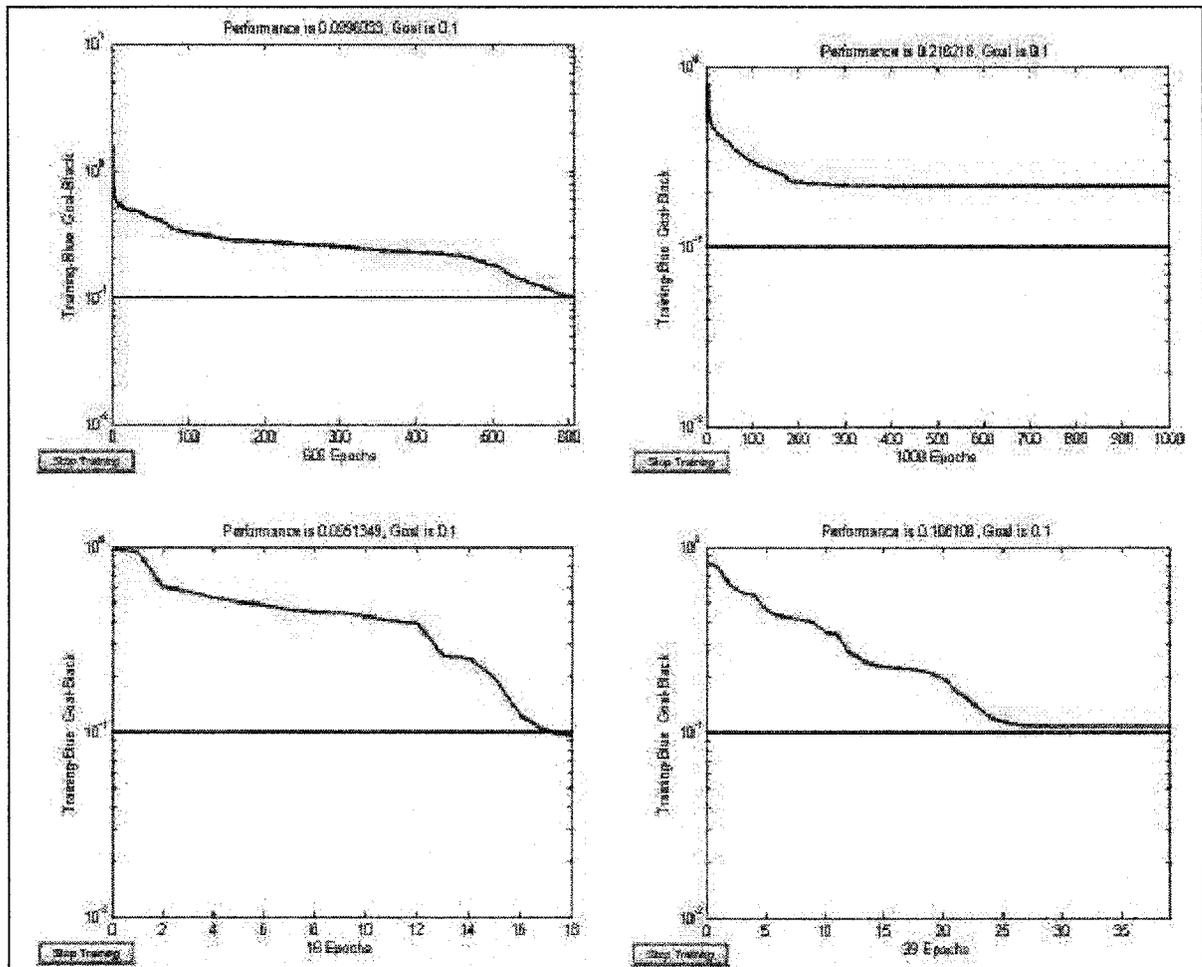


Figure 5-2 Typical Training Sessions of the RP and LM Algorithms

When the performances of networks trained with different numbers of input parameters were compared, networks trained with five input parameters converged more frequently although the outputs generated were less accurate. A probable explanation for this observation is that although networks with five input parameters had more information to process in order to achieve faster convergence, the added features were not effective in discriminating the images and hence did not improve the accuracy of predictions. Since one of the goals of the ANN analysis in this study was to minimize the number of input parameters required to establish an equally effective network, further analysis of network performance will focus on ANNs trained with three and four input features which demonstrated higher predictive power. Table 5-5 shows that among the networks trained with three input features, the network trained with mean, QDP and EDR achieved the highest CCR of 0.74. Similarly, among the networks trained with four input features the network trained with mean, kurtosis, QDP and EDR achieved the highest CCR of 0.73.

5.3.2 Predictive Power of the ANN

In clinical settings, predictive power of the ANN is determined by several measures of accuracy: sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and the CCR of outcome predictions. This section examines various performance measures achieved by the best performed networks in this experiment to determine their respective predictive power. Performance of networks with three input parameters and four input parameters are compared.

As mentioned in the previous section, the network trained with four input parameters, namely mean, kurtosis, QDP and EDR, and seven hidden nodes has a CCR of

0.73, while the network trained with three input parameters, mean, QDP and EDR, and six hidden nodes has a CCR of 0.74. ANN outputs obtained for these two networks are summarized in Table 5-6 for comparison. In Table 5-6, all images with positive clinical diagnosis are colour coded in red. The table entries shaded in blue show images that were predicted incorrectly by the ANN more than half of the time. Column headings for the CCR results of each individual image are shown shaded in grey. An interpretation for the column headings is provided in Figure 5-3.

Descriptions of column headings for Table 5-6

clinical diag. – corresponding clinical diagnoses for the images.

pred. positive - number of times an image was predicted as “positive” by the ANN.

test image - number of times the image was randomly chosen as a test image during the training iterations.

CCR -
$$\frac{\text{total number of correct ANN predictions}}{\text{test image (defined above)}}$$

Example

The image *BT2_1127*, which was clinically diagnosed to be *normal*, was randomly chosen to be a test image for the *6x1* network in 27 iterations. In 1 out of the 27 iterations, the ANN classified the image as *abnormal* resulting in a *CCR of 95%*.

Note

See Figure 5-4 and section 5.3.3 for a discussion of images 51902 and NY004, two images poorly predicted by the ANN.

Figure 5-3 Interpretations of column headings for Table 5-6

Network Architecture	7 x 1
Input Parameters	4
Training Algorithm	RP
Sensitivity	0.57
Specificity	0.80
Positive Pred. Val.	0.57
Negative Pred. Val.	0.83
CCR	0.73

Network Architecture	6 x 1
Input Parameters	3
Training Algorithm	RP
Sensitivity	0.60
Specificity	0.79
Positive Pred. Val.	0.58
Negative Pred. Val.	0.84
CCR	0.74

Image	clinical diag.	pred. positive	test image	CCR
BT1_11020	0	0	23	1.00
BT2_10280	0	0	26	1.00
BT2_11120	0	0	25	1.00
BT2_11320	0	0	24	1.00
BT2_11520	0	0	25	1.00
51402	0	0	25	1.00
51702	0	0	24	1.00
51802	0	0	17	1.00
52002	0	0	31	1.00
52202	0	0	27	1.00
52302	0	0	19	1.00
52702	0	0	25	1.00
NY001	0	0	24	1.00
BT2_11070	1	1	29	0.97
BT3_11000	1	1	23	0.96
BT1_11320	1	1	21	0.95
BT2_11270	1	1	21	0.95
BT1_11070	1	1	20	0.95
BT1_10250	1	1	19	0.95
BT3_10300	2	2	26	0.92
52102	0	2	22	0.91
51602	0	2	14	0.86
NY008	0	5	31	0.84
51302	0	4	18	0.78
NY009	0	8	28	0.71
Test8	0	6	21	0.71
BT1_11270	6	6	16	0.63
Test92	0	12	26	0.54
52502	0	12	25	0.52

Image	clinical diag.	pred. positive	test image	CCR
BT1_10250	0	0	25	1.00
BT1_11020	0	0	18	1.00
BT1_11070	0	0	23	1.00
BT2_10280	0	0	29	1.00
BT2_11020	0	0	26	1.00
BT2_11120	0	0	23	1.00
BT2_11320	0	0	26	1.00
BT2_11520	0	0	24	1.00
51402	0	0	25	1.00
51802	0	0	23	1.00
52002	0	0	31	1.00
52202	0	0	24	1.00
52302	0	0	20	1.00
52702	0	0	20	1.00
NY001	0	0	28	1.00
BT2_11270	1	1	27	0.96
51702	0	1	26	0.96
BT1_11270	1	1	22	0.95
BT1_11320	1	1	22	0.95
NY008	0	1	17	0.94
BT3_10300	2	2	26	0.92
BT3_11000	2	2	26	0.92
BT2_11070	2	2	24	0.92
52102	0	4	30	0.87
51302	0	4	20	0.80
NY009	0	6	26	0.77
51602	0	10	28	0.64
BT2_10080	8	8	18	0.56
Test8	0	10	22	0.55

Test6	0	11	20	0.45	52502	0	12	22	0.45
BT2_1102	0	14	24	0.42	Test92	0	12	22	0.45
BT2_1008	0	11	17	0.35	Test7	0	16	21	0.24
Test7	0	19	22	0.14	Test6	0	22	25	0.12
Test5	0	24	26	0.08	Test5	0	24	25	0.04
51902	0	21	21	0.00	51902	0	37	37	0.00
BT3_1070	1	31	31	1.00	BT3_1070	1	28	28	1.00
NY007	1	18	20	0.90	NY007	1	26	26	1.00
NY010	1	17	19	0.89	NY005	1	26	28	0.93
NY012	1	23	27	0.85	NY010	1	31	35	0.89
BT3_1015	1	13	17	0.76	51502	1	20	27	0.74
51502	1	24	32	0.75	NY011	1	16	23	0.70
NY005	1	15	24	0.63	NY006	1	13	24	0.54
BT2_1082	1	9	18	0.50	BT3_1090	1	10	19	0.53
BT3_1090	1	12	24	0.50	NY002	1	10	19	0.53
NY002	1	10	20	0.50	NY012	1	14	28	0.50
NY006	1	5	24	0.21	BT3_1015	1	8	19	0.42
NY011	1	3	22	0.14	BT2_1082	1	5	25	0.20
NY003	1	1	24	0.04	NY003	1	2	26	0.08
NY004	1	0	21	0.00	NY004	1	0	22	0.00

Table 5-6 Summary Results for the Best-Performed Networks

5.3.3 Discussions on the Performance of the ANN

Measures of accuracy obtained for the two networks shows that the abilities of the networks to predict clinical outcomes are quite comparable. It may therefore be concluded that adding the kurtosis as a fourth input parameter did not improve the ANN predictive power significantly in this study. However, a more conclusive statement regarding the relevance of kurtosis can only be made in a larger study where more training datasets are available since the 7x1 network results obtained in this analysis may be generated by an inadequately trained network.

In regards to the performance of the ANN to predict clinical outcomes, the networks were more accurate in evaluating normal thermograms and less effective in recognizing abnormal thermograms. Overall performance of the network is limited by the small set of training samples in this study. As a rule of thumb to obtain a reliable network, five to ten input vectors are required for the weight training of each node in the network, including the input nodes [46]. In the case of this study, the smaller 6x1 network with three input parameters, six hidden nodes, and one output node requires at least 50 images to train. However, since 12 images were set aside for validation, only 37 images were available for training the network. In addition, since the images were obtained from different sources and thus varied widely in their characteristics, more images are required for the network to sufficiently learn the image features for the classification task.

Low sensitivity values and PPV, which means a high number of false negatives in the predicted outcomes, are partly attributed to the relatively small sample of cancerous patients in the analysis. Hence, the networks were not sufficiently trained to recognize features that may be used to differentiate the abnormal images from the normal images, especially since different pathologies of breast cancer manifest themselves differently on breast thermograms, and therefore require a larger and more complete training database. Furthermore, as the ANNs in this study were not trained to differentiate between benign cases and healthy cases, some abnormal thermal characteristics of images with benign lesions were classified as normal although the same characteristics may be present on images with malignant lesions.

As recalled from the description of New York images in section 4.2.3, there were actually no “normal” patients included among the 17 patients in the group. Nine patients were diagnosed with cancer and the remaining eight patients all had images that showed some degrees of abnormalities in their breast infrared images. In other words, the latter eight images that were given a “healthy” diagnosis may actually share some abnormal characteristics with the cancerous images such as highly vascularized thermal patterns. If the ANNs were trained to classify these eight images as normal, the ability to recognise an abnormal image that shared some common abnormal signs with the “normal” images would be severely affected, as the image may be falsely classified as negative.

In addition, the ability of the trained network to classify the images into cancerous and healthy images may also be affected by inconsistent diagnoses of thermal images. As the images were obtained from multiple sources, two images that have similar thermal characteristics may be diagnosed as cancerous in one clinic but normal in the other.

Diagnoses for images obtained from Montreal were not consistent with the rest of the samples. For the Montreal images, each breast was given a score on the infrared scale based on the presence of abnormal signs. Hence, the “clinical diagnosis” of these images was actually not healthy or cancerous, but a set of two numbers such as [1, 2], [2, 2] and [1, 4]. To classify these numbered images as healthy and cancerous, it was decided that images scoring IR2 or less for both breast were considered healthy while images scoring IR3, IR4 or IR5 in any one of the breasts were considered cancerous. Therefore, an image with [2, 2] diagnosis was classified as normal while an image with [1, 4] was classified as abnormal. With this sorting scheme, 12 out of 13 images were healthy, although eight of these had a score of IR2. A score of IR2, as shown in Table 4-1, was

given to images with significant but symmetrical vascular patterns, or images with moderate vascular asymmetry. To illustrate the inconsistency among clinical diagnoses given for this study, some sample images and their corresponding diagnosis are shown in Figure 5-4. The image “NY004” shown at the top right corner was obtained from a cancerous patient but displayed fairly symmetrical vascular patterns. As a result, the ANN consistently predicted the image to be normal. On the other hand, the image “51902” at the bottom left corner showed a significant temperature difference between the contralateral breasts, just as image “NY007”. As a result, the ANN predicted both images to be “abnormal”. The clinical diagnosis for “NY007” was in fact cancerous. However, as “51902” was acquired from Montreal and had a clinical outcome of [2, 2], its clinical diagnosis was set to normal.

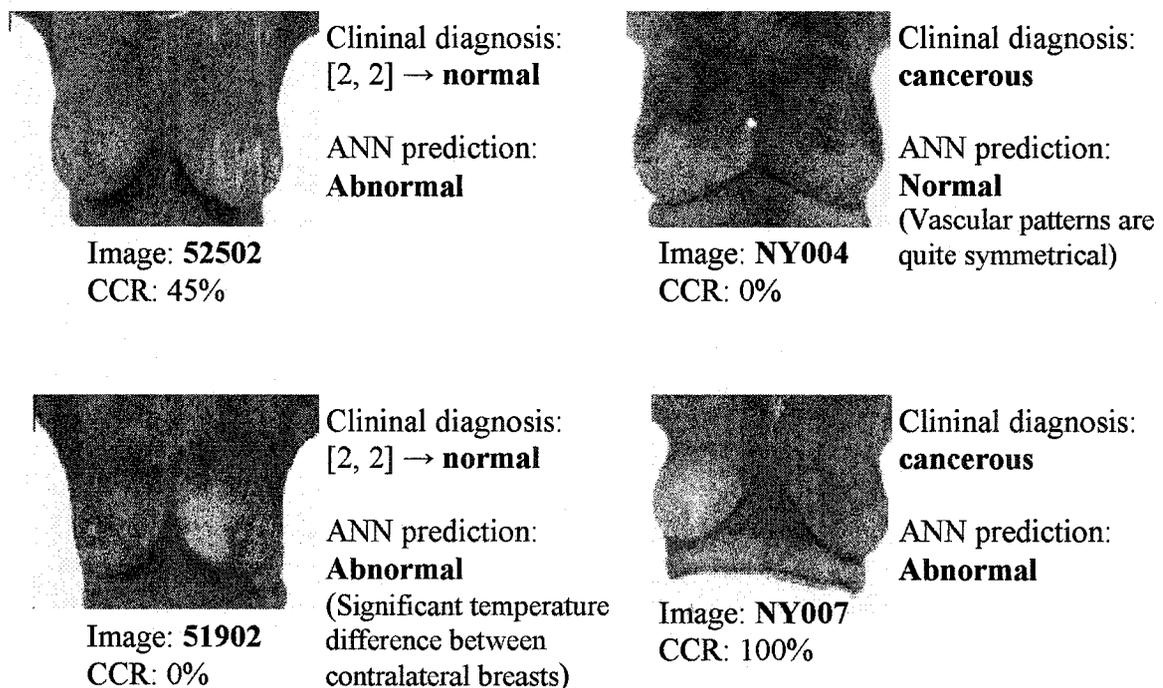
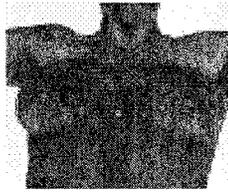


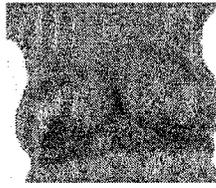
Figure 5-4 Comparisons of Clinical Diagnoses and ANN predictions for Selected Images

Although the performance of ANN analysis in predicting clinical outcomes of breast thermograms in this study was not sufficient to prove that breast thermography is more effective in detecting early breast cancer than mammography, it demonstrated the potential of quantitative analysis for objective breast thermogram evaluations. Furthermore, previous research had demonstrated that thermography was able to detect a breast cancer as early as five years before the malignancy was confirmed [30]. Since images obtained from Montreal and New York were acquired within the past four years, some of the false positive results obtained among these images may actually be cancerous patients that cannot be confirmed at this time by a different detection modality. Some examples of correct ANN predictions are shown in Figure 5-5. The numbers for CCR are taken from Table 5-6.

However, to fully assess the ability of ANNs to predict true outcomes based on quantitative parameters extracted from the images, a larger image database is required. More cancerous images covering different breast cancer pathologies are also essential so that the ANNs learn to recognize a variety of abnormal signs associated with breast cancer.



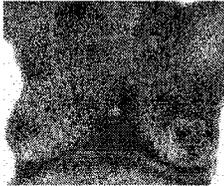
Diag: Normal
CCR: 100%



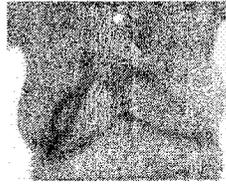
Diag: Normal
CCR: 92-97%



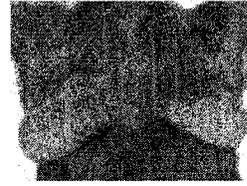
Diag: Normal [2, 2]
CCR: 96-100%



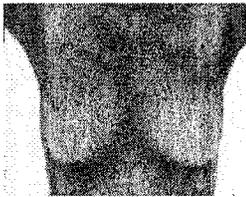
Diag: Abnormal
CCR: 93%
(With 6x1 network)



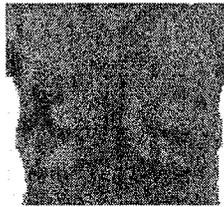
Diag: Abnormal
CCR: 100%



Diag: Abnormal
CCR: 89%



Diag: Normal [1, 1]
CCR: 100%



Diag: Normal
CCR: 100%



Diag: Normal [2, 2]
CCR: 100%

Figure 5-5 ANN Predictions for Some Representative Images

6 Conclusions

6.1 Concluding Remarks and Summary of Contributions

This thesis developed a quantitative method to analyze and evaluate breast infrared images. The method employed statistical analysis and image processing techniques to extract relevant features of the thermal patterns on the breast images. To facilitate objective evaluations of these images, the Artificial Neural Network was employed to analyze the image features extracted and to predict clinical outcomes.

With this method of analyzing breast thermograms quantitatively and objectively, sensitivity and specificity obtained in the experiments were 60% and 79% respectively. These levels of sensitivity and specificity were compared to similar measures of performance of mammography to evaluate the effectiveness of infrared imaging for early breast cancer detection. As stated in section 2.2.1, sensitivity of mammography ranges from 54% for women 40 years old or younger, and 81% for women 64 years old or older.

Specificity of mammography ranges insignificantly from 91.4% to 94.4% for women from all age groups. Although the sensitivity level obtained with the ANN predictions was slightly higher for young women, overall the sensitivity and specificity levels were not superior to those of mammography. Therefore, the results obtained in this thesis were not able to show that breast thermography is more effective than mammography is in detecting early breast cancer. The reason for the method's inadequacy could be partly attributed to the limited and inconsistent image database available.

Vascular patterns observed on breast thermograms can often provide as much information regarding the abnormality of the images as the global and local temperature difference between contralateral breasts. However, they are often not taken into consideration by researchers performing statistical analysis of infrared images. Two algorithms were developed in this study to extract quantitative parameters from the infrared images that may give a good numerical representation of the vascular patterns. The EDR parameter extracted with edge detection analysis was more sensitive to image noise and different thermal resolutions of the multiple-sourced images and was not reliable in discriminating the images. The second algorithm extracted a quantitative parameter, QDP, by first performing quadtree decomposition analysis on the images followed by asymmetry analysis between the contralateral breasts. The QDP parameter was found to be quite effective in classifying the breast thermograms, and demonstrated potential as a discriminating feature based on thermal vascular asymmetry complementing the widely employed statistical features.

In this study, ANNs were also employed to identify the most relevant features in predicting outcomes of breast thermograms. Due to the small image database available

for ANN analysis, correlation analysis was first performed on the statistical parameters to reduce the number of input parameters. Among these input parameters, the combinations of the mean temperature, the QDP and the EDR generated the highest CCR in the ANN analyses. Hence, these would be the minimum inputs required to effectively classify the breast thermograms used in this thesis.

In conclusion, the thesis presented a quantitative method to facilitate computerised and objective assessment of the breast thermograms for potential early breast cancer detection. Different technical aspects of the process in image acquisition, image segmentation, feature extraction, and image analysis were investigated. Although the quantitative analysis of infrared images employed in this thesis was not able to demonstrate the effectiveness of breast thermography, with a larger database and expansion of future work, the method may still demonstrate infrared imaging to be an effective method for early breast cancer detection.

6.2 Future Work

To improve the effectiveness of infrared imaging in early breast cancer detection, several aspects of image acquisition, image processing and image analysis can be considered for future work.

Building a larger image database should be the starting point. To avoid inconsistencies and to reduce complexities for the image processing and analyzing work during the preliminary stages, the breast infrared images should be acquired from the same equipment, or have very similar characteristics. Once an effective method is developed, images from variety of sources should be included in the evaluation to improve the robustness of the method.

With a larger database of image/diagnosis pairs, ANN analysis could be performed with more input features and larger network configurations with less risk of overfitting. The potential of *fractal dimension* for thermal texture analysis or a measure of vascularity may be explored. Transforming the images with wavelet analysis to obtain local and global details of the images may reveal other relevant features in classifying breast thermograms. To ensure more accurate symmetry analysis between contralateral breasts, image registration to re-align the breasts may be considered.

In addition, other segmentation techniques to identify the breast regions (ROI) could be explored and evaluated. An automated image segmentation routine will be required for more efficient ROI identifications, especially for a larger image database. If acquired images are noisy, denoising algorithms may improve the accuracy of the outcome predictions.

In this thesis, ANN analysis with a backpropagation neural network was performed with a fixed learning rate. More efficient training may be obtained by using training algorithms that employ adaptive learning rates. For comparative analysis, different types of neural networks may be experimented with. If the ANN is trained with an image database that has insufficient cancerous cases, re-sampling images from patients with a positive diagnosis may be considered to improve the ANN's predictability.

Finally, assessments of breast thermograms in this study were performed based on analysis of images taken from the front view only. A more effective method may be obtained if images taken from the lateral views were also analysed. Furthermore, although the ANN model developed evaluates the health of the breasts, and classifies the images as normal or abnormal, it does not provide information regarding possible

location of the tumour on abnormal images. A more sophisticated method utilizing more image features, as well as images taken from different camera views would be required to localize the tumour.

References

1. Abdolmaleki P, Buadu LD, Naderimansh H, "Feature extraction and classification of breast cancer on dynamic magnetic resonance imaging using artificial neural network," *Cancer letters*, 171(2):183-91, Oct 10, 2001.
2. American Cancer Society, "Mammograms and Other Breast Imaging Procedures,"
http://www.cancer.org/docroot/CRI/content/CRI_2_6X_Mammography_and_othe_r_Breast_Imaging_Procedures_5.asp?sitearea=CRI
3. American Cancer Society, *Cancer Facts and Figures 2004*.
4. Anbar M, Milescu L, Naumov A, Brown C, Button T, Carly C, AlDulaimi K, "Detection of cancerous breasts by dynamic area telethermometry," *Engineering in Medicine and Biology Magazine*, IEEE, 20(5): 80-91, Sept-Oct, 2001.
5. Arena F, Barone C, DiCicco T, "Use of digital infrared imaging in enhanced breast cancer detection and monitoring of the clinical response to treatment," In *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2, pp. 1129–1132, Cancun, Mexico, 2002.

6. Aubry-Frize M, Quartey GRC, Evans H, Lapalme D, "The thermographic detection of pain," In *Proceedings of the 3rd Canadian Clinical Engineering Conference*, pp. 82–83, Saskatoon, 1981.
7. Balcerak RS, Lupo J, "Advances in infrared sensor technology and systems," In *Proceedings of the Second Joint EMBS/BMES Conference*, Houston, Texas. Vol. 2, pages 1124–1125, Oct 23-26 2002.
8. Barnes RB, "The early history of thermography," *Medical Thermography, Theory and Clinical Application*, Brentwood Publishing Corporation, Los Angeles, California, 1976.
9. Bates T, Davidson T, Mansel RE, "Litigation for pneumothorax as a complication of fine-needle aspiration of the breast," *The British journal of surgery*, 89(2):134-7, Feb 2002.
10. Burke HB, Hoang A, Iglehart JD, Marks JR, "Predicting response to adjuvant and radiation therapy in patients with early stage breast carcinoma," *Cancer*, 82(5):874-7, Mar 1, 1998.

11. Carney P, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, et al, "Individual and combined effects of age, breast density and hormone replacement therapy use on the accuracy of mammography," *Annals of Internal Medicine*, 138: 168-75, 2003.
12. Chen DR, Chang RF, Kuo WJ, Chen MC, Huang YL, "Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks," *Ultrasound in medicine & biology*, 28(10):1301-10, Oct 2002.
13. Chu JS, Huang CS, Chang KJ, "The prognostic significance of tumor angiogenesis in Taiwanese patients with invasive ductal breast carcinomas," *Cancer Letter*, 134(1): 7-14, 1998.
14. Clark RP, "Human skin temperature and its relevance in physiology and clinical assessment," *Recent advances in medical thermology*, New York: Plenum, 1984.
15. Committee on the Biological Effects of Ionizing Radiation (BEIR V), National Research Council, "Health Effects of Exposure to Low Levels of Ionizing Radiation: BEIR V, 1990," The National Academies Press.
<http://www.nap.edu/catalog/1224.html>
16. Donati S, *Photodetectors: Devices, Circuits, and Applications*. Prentice-Hall PTR, Upper Saddle River, NJ, 1999.

17. Drukker K, Giger ML, Horsch K, Kupinski MA, Vyborny CJ, Mendelson EB. "Computerized lesion detection on breast ultrasound," *Medical physics*, 29(7):1438-46, Jul 2002.
18. Ellis LM, Fidler IJ, "Angiogenesis and breast cancer metastasis," *The Lancet*, August, 346(8972): 388-390, 1995.
19. Ellsworth DL, Ellsworth RE, Love B, Deyarmin B, Lubert SM, Mittal V, Hooke JA, Shriver CD, "Outer Breast Quadrants Demonstrate Increased Levels of Genomic Instability," *Annals of Surgical Oncology*, 11(9): 861-868, September 2004.
20. Elmore JG, Barton MB, Mocerri VM, Polk S, Arena PJ, Fletcher SW, "Ten-year risk of false positive screening mammograms and clinical breast examinations," *The New England journal of medicine*, 338(16):1089-96, Apr 16, 1998.
21. Eubank WB, Mankoff DA, "Current and future uses of positron emission tomography in breast cancer imaging," *Seminars in Nuclear Medicine*, 34(3):224-40, Jul 2004.
22. Flesch U, "Physics of skin-surface temperature," *Thermological Methods*, pp 21-33. VCH Verlagsgesellschaft mbh, Weinheim, Germany, 1985.

23. Fletcher SW, Black W, Harris R, Rimer BK and Shapiro S, "Report of the International Workshop on Screening for Breast Cancer," *Journal of the National Cancer Institute*, 85(20):1644-56, Oct 20, 1993.
24. Floyd CE Jr, Lo JY, Yun AJ, Sullivan DC, Kornguth PJ, "Prediction of breast cancer malignancy using an artificial neural network," *Cancer*, 74(11):2944-8, Dec 1, 1994.
25. Freeman JA, Skapura DM, *Neural Networks: Algorithms, Applications, and Programming Techniques*. Addison-Wesley Publishing Company, 1992.
26. Frize M, Herry C, Roberge R, "Processing of thermal images to detect Breast Cancer: A comparison with previous work," In *Proceedings of the Second Joint EMBS/BMES Conference, Houston, Texas*. vol. 2, pp.1159–1160, Oct 23-26 2002.
27. Frize M, Herry CL, and Scales N, "Processing thermal images to detect breast cancer and assess pain," In *Proceedings of 4th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine*, pp. 234-237, April 24-26, 2003.

28. Gamagami P, Silverstein MJ, Waisman JR, "Infrared imaging in breast cancer,"
In *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 677-680, Chicago, IL, 1997.
29. Gautherie M, "Thermopathology of breast cancer: measurement and analysis of in vivo temperature and blood flow," *Annals New York Academy Sciences*, 335:383-415, 1980.
30. Gautherie M, Gros CM, "Breast thermography and cancer risk prediction," *Cancer*, 45: 51-6, 1980.
31. Geser HM, Bosiger P, Stucki D, Landolt C, "Computer-assisted dynamic breast thermography," *Thermology*, 2:538-544, 1987.
32. Goin JE, Haberman JD, "Automated breast cancer detection by thermography: Performance goal and diagnostic feature identification," *Pattern Recognition*, 16(2):125-129, 1983.
33. Gotzsche PC and Olsen O, "Is screening for breast cancer with mammography justifiable?" *Lancet*, 355(9198):129-34, Jan 8, 2000.
34. Gros C, Gautherie M, Bourjat P, Vrousos C, "Thermography of Breast Diseases," *Bibliotheca Radiological*, 5:68-81, 1969.

35. Hairong Qi, Snyder WE, Head JF, Elliott RL, "Detecting breast cancer from infrared images by asymmetry analysis," In *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1227–1228, Chicago, IL, 2000.
36. Hairong Qi, Head JF, "Asymmetry analysis using automatic segmentation and classification for breast cancer detection in thermograms," In *Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2866–2869, Istanbul, Turkey, 2001.
37. Hairong Qi, Kuruganti PT, Zhongqi Liu, "Early detection of breast cancer using thermal texture maps," In *Proceedings of 2002 IEEE International Symposium on Biomedical Imaging*, pp. 309-312, 2002.
38. Hairong Qi, Zhongqi Liu, Chen Wang, "Breast cancer identification through shape analysis in thermal texture maps," In *Proceedings of the Second Joint EMBS/BMES Conference*, Houston, Texas. vol. 2, pp.1129–1130, Oct 23-26 2002.
39. Halir R, Flusser J, "Numerically stable direct Least Squares Fitting of Ellipses," WSCG'98, pp. 125-132, 1998.

40. Haykin S, *Neural Networks: A comprehensive foundation*. Prentice Hall, Upper Saddle River, New Jersey, 1999.
41. Head JF, Elliott RL, "Thermography. Its relation to pathologic characteristics, vascularity, proliferation rate, and survival of patients with invasive ductal carcinoma of the breast," *Cancer*, 79(1):186-8, Jan 1, 1997.
42. Head JF, Lipari CA, Wang F, Elliott RL, "Image analysis of digitized infrared images of the breasts from a first generation infrared imaging system," In *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 681-684, Chicago, IL, 1997.
43. Head JF, Lipari CA, Elliott RL, "Determination of mean temperatures of normal whole breast and breast quadrants by infrared imaging and image analysis," In *Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2823-2825, Istanbul, Turkey, 2001.
44. Herry C, "Quantitative assessment of pain through clinical digital infrared imaging," M.A.Sc. thesis. Department of System and Computer Engineering, Carleton University. Ottawa, Canada. 2002.

45. Herry CL, Frize M, "Quantitative assessment of pain-related thermal dysfunction through clinical digital infrared thermal imaging," *Biomedical Engineering Online*, 3(1):19, Jun 28, 2004.
46. Hudson DL, Cohen ME, *Neural Networks and Artificial Intelligence for Biomedical Engineering*. IEEE Press, New York, NY, 2000.
47. Infrared Sciences Corporation, "BreastScan IR,"
<http://www.infraredsciences.com/pages/4/index.htm>
48. Isard HJ, Sweitzer CJ, Edelstein GR, "Breast thermography. A prognostic indicator for breast cancer survival," *Cancer*, 62(3):484-8, Aug 1, 1988.
49. Jakubowska T, Wiecek B, Wysocki M, Drews-Peszynski C, Strzelecki M, "Classification of breast thermal images using artificial neural networks," In *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1155-1158, San Francisco, CA, 2004.
50. Jones BF, "A reappraisal of the use of infrared thermal image analysis in medicine," *IEEE Transactions on Medical Imaging*, 17(6):1019-1027, 1998.

51. Jones BF, Plassmann P, "Digital infrared thermal imaging of human skin," *Engineering in Medicine and Biology Magazine*, IEEE, 21(6): 41-48, Nov-Dec 2002.
52. Keyserlingk JR, Ahlgren PD, Yu E, Belliveau N, and Yassa M, "Functional infrared imaging of the breast," *Engineering in Medicine and Biology Magazine*, IEEE, 19(3): 30-41, 2000.
53. Keyserlingk JR, Ahlgren PD, Yassa M and Belliveau N, "Overview of functional infrared imaging as part of a multi-imaging strategy for breast cancer detection and therapeutic monitoring," In *Proceedings of the Second Joint EMBS/BMES Conference*, Houston, Texas, vol.2, pp. 1126–1128, Oct 23-26 2002.
54. Koay J, Herry C, Frize M, "Analysis of breast thermography with an Artificial Neural Network," In *Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1159–1162, San Francisco, CA, 2004.
55. Kuruganti PT, Hairong Qi, "Asymmetry analysis in breast cancer detection using thermal infrared images," In *Proceedings of the Second Joint EMBS/BMES Conference*, Houston, Texas. vol. 2, pp. 1155–1156, Oct 23-26 2002.

56. Lawson R, "Implications of surface temperature in the diagnosis of breast cancer," *Canadian Medical Association Journal*, 75: 309, August 5, 1956.
57. Lawson R, "Thermography; a new tool in the investigation of breast lesions," *Canadian Services Medical Journal*, 8(8):517-24, Sep 1957.
58. Leek RD, "The prognostic role of angiogenesis in breast cancer," *Anticancer Research*, 21(6B):4325-31, Nov-Dec 2001.
59. Lidbrink E, Elfving J, Frisell J, Jonsson E, "Neglected aspects of false positive findings of mammography in breast cancer screening: analysis of false positive cases from the Stockholm trial," *BMJ (Clinical research ed.)*, 312(7026):273-6, Feb 3, 1996.
60. Ma L, Fishell E, Wright B, Hanna W, Allan S, Boyd N F, "Case-control study of factors associated with failure to detect breast cancer by mammography," *Journal Of The National Cancer Institute*, 84(10):781-785, May 20, 1992.
61. Malich A, Fritsch T, Anderson R, Boehm T, Freesmeyer MG, Fleck M, Kaiser WA, "Electrical impedance scanning for classifying suspicious breast lesions: first results," *European radiology*, 10(10):1555-61, 2000.

62. Markopoulos C, Kouskos E, Koufopoulos K, Kyriakou V, Gogas J, "Use of artificial neural networks (computer analysis) in the diagnosis of microcalcifications on mammography," *European journal of radiology*, 39(1):60-5, Jul 2001.
63. National Cancer Institute of Canada, *Canadian Cancer Statistics 2004*, Toronto, Canada, 2004. ISSN 0835-2976, April 2004.
64. Neural Network Toolbox (version 4.0) User's Guide, The Math Works, Natick, MA, 2003, pp. 5-32 - 5-50.
65. Ng EY, Fok SC, Peh YC, Ng FC, Sim LS, "Computerized detection of breast cancer with artificial intelligence and thermograms," *Journal of medical engineering & technology*, 26(4):152-7, Jul-Aug, 2002.
66. Ohashi Y, Uchida I, "Applying dynamic thermography in the diagnosis of breast cancer," *Engineering in Medicine and Biology Magazine*, IEEE, 19(3): 42-51, May-June, 2000.
67. Pagano M and Gauvreau K, *Principles of Biostatistics*. Duxbury Thompson Learning, 2000.

68. Poole DO, "The basic physics of clinical thermography," *Medical Thermography, Theory and Clinical Applications*. Brentwood Publishing Corporation, Los Angeles, California, 1976.
69. Poussart, M Frize, R Roberge, "The re-evaluation of thermography in breast cancer detection by new image enhancement techniques," *Research Report*, Terry Fox Foundation, Fredericton, NB. Canada, 1988.
70. Primic-Zakelj M, "Screening mammography for early detection of breast cancer," *Annals of Oncology: official journal of the European Society for Medical Oncology / ESMO*, 10 Suppl 6:121-7, 1999.
71. Ring EFJ, "Standardization of thermal imaging in medicine," *Thermal Assessment in breast health*, Ed. Gautherie M, Albert E, Keith L. MTP Press Ltd, Lancaster, England, pp. 29-36, 1984.
72. Ronco AL, "Use of artificial neural networks in modeling associations of discriminant factors: towards an intelligent selective breast cancer screening," *Artificial intelligence in medicine*, 16(3):299-309, Jul 1999.

73. Rosenberg RD, Hunt WC, Williamson MR, Gilliland FD, Wiest PW, Kelsey CA, Key CR, Linver MN, "Effects of age, breast density, ethnicity, and estrogen replacement therapy on screening mammographic sensitivity and cancer stage at diagnosis: review of 183,134 screening mammograms," *Radiology*, 209(2):511-8, Nov 1998.
74. Schulte BP, Bomhof MA, Aarts NJ, "Facial thermography in the diagnosis of cerebrovascular disease and in evaluation of carotid endarterectomy," *Clinical Neurology and Neurosurgery*, 78(2):118-30, 1975.
75. Smith IC, Gilbert FJ, "Role of positron emission tomography in the management of breast cancer," *Breast*, 8(6):303-10, Dec 1999.
76. Sterns EE, Zee B, SenGupta S, Saunders FW, "Thermography: Its relation to pathologic characteristics, vascularity, proliferation rate, and survival of patients with invasive ductal carcinoma of the breast," *Cancer*, 77(7): 1324-8, 1996.
77. Stony Brook State University of New York, "Breast Health Knowledge Base", <http://www.stonybrookhospital.com/BALDWIN/kindsof.htm>
78. Suri JS, Setarehdan SK, and Singh S, *Advanced Algorithmic Approaches to Medical Image Segmentation. State of the Art Application in Cardiology, Neurology, Mammography and Pathology*. Springer-Verlag London Ltd, 2002.

79. Szabo BK, Wiberg MK, Bone B, Aspelin P, "Application of artificial neural networks to the analysis of dynamic MR imaging features of the breast," *European radiology*, 14(7):1217-25, Jul 2004.
80. U.S. National Library of Medicine and National Institute of Health, MedlinePlus Medical Encyclopedia.
<http://www.nlm.nih.gov/medlineplus/encyclopedia.html>
81. Webb S, Ed., *The physics of medical imaging*. Philadelphia : Hilger, 1988.
82. Williams KL, Phillips BH, Jones PA, Beaman SA, Fleming PJ, "Thermography in screening for breast cancer," *J Epidemiol Community Health*, 44(2): 112-3, 1990.
83. Yune Yuan, Qing Wang, Yan Tan, Li Rong, Song San Tai, Li Ji Ye, Zhongqi Liu, "Analysis of breast diseases examination with thermal texture mapping, mammography and ultrasound," In *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1166-1169, San Francisco, CA, 2004.

Appendix

Quadrant 1 Nonparametric Correlations												
		mean	median	std_dev	interquart	skew	kurtosis	min	max	entropy	heat	
Spearman's rho	mean	Correlation Coefficient	1.000	0.981	-0.047	-0.074	-0.502	0.033	0.475	0.113	0.013	0.350
		Sig. (2-tailed)		0.000	0.750	0.613	0.000	0.824	0.001	0.440	0.930	0.014
		N	49	49	49	49	49	49	49	49	49	49
	median	Correlation Coefficient	0.981	1.000	-0.010	-0.066	-0.516	0.054	0.461	0.147	0.021	0.325
		Sig. (2-tailed)	0.000		0.946	0.651	0.000	0.712	0.001	0.314	0.887	0.023
		N	49	49	49	49	49	49	49	49	49	49
	std_dev	Correlation Coefficient	-0.047	-0.010	1.000	0.850	0.111	-0.247	-0.237	0.277	0.896	-0.214
		Sig. (2-tailed)	0.750	0.946		0.000	0.446	0.087	0.102	0.054	0.000	0.140
		N	49	49	49	49	49	49	49	49	49	49
	interquartile	Correlation Coefficient	-0.074	-0.066	0.850	1.000	0.184	-0.557	-0.041	0.072	0.873	-0.157
		Sig. (2-tailed)	0.613	0.651	0.000		0.206	0.000	0.782	0.622	0.000	0.281
		N	49	49	49	49	49	49	49	49	49	49
	skewness	Correlation Coefficient	-0.502	-0.516	0.111	0.184	1.000	0.011	0.012	0.299	0.056	-0.167
		Sig. (2-tailed)	0.000	0.000	0.446	0.206		0.942	0.934	0.037	0.701	0.251
		N	49	49	49	49	49	49	49	49	49	49
	kurtosis	Correlation Coefficient	0.033	0.054	-0.247	-0.557	0.011	1.000	-0.288	0.478	-0.450	-0.153
		Sig. (2-tailed)	0.824	0.712	0.087	0.000	0.942		0.045	0.001	0.001	0.294
		N	49	49	49	49	49	49	49	49	49	49
	min	Correlation Coefficient	0.475	0.461	-0.237	-0.041	0.012	-0.288	1.000	-0.072	-0.137	0.214
		Sig. (2-tailed)	0.001	0.001	0.102	0.782	0.934	0.045		0.622	0.347	0.140
		N	49	49	49	49	49	49	49	49	49	49
	max	Correlation Coefficient	0.113	0.147	0.277	0.072	0.299	0.478	-0.072	1.000	0.204	-0.075
		Sig. (2-tailed)	0.440	0.314	0.054	0.622	0.037	0.001	0.622		0.160	0.609
		N	49	49	49	49	49	49	49	49	49	49
entropy	Correlation Coefficient	0.013	0.021	0.896	0.873	0.056	-0.450	-0.137	0.204	1.000	0.013	
	Sig. (2-tailed)	0.930	0.887	0.000	0.000	0.701	0.001	0.347	0.160		0.931	
	N	49	49	49	49	49	49	49	49	49	49	
heat	Correlation Coefficient	0.350	0.325	-0.214	-0.157	-0.167	-0.153	0.214	-0.075	0.013	1.000	
	Sig. (2-tailed)	0.014	0.023	0.140	0.281	0.251	0.294	0.140	0.609	0.931		
	N	49	49	49	49	49	49	49	49	49	49	

Quadrant 2 Nonparametric Correlations												
			mean	median	std dev	interquart	skew	kurtosis	min	max	entropy	heat
Spearman's rho	mean	Correlation Coefficient	1.000	0.963	-0.297	-0.336	-0.169	0.268	0.685	0.432	-0.294	0.623
		Sig. (2-tailed)		0.000	0.039	0.018	0.245	0.062	0.000	0.002	0.040	0.000
		N	49	49	49	49	49	49	49	49	49	49
	median	Correlation Coefficient	0.963	1.000	-0.261	-0.292	-0.259	0.172	0.654	0.321	-0.263	0.548
		Sig. (2-tailed)	0.000		0.070	0.042	0.073	0.238	0.000	0.025	0.067	0.000
		N	49	49	49	49	49	49	49	49	49	49
	std_dev	Correlation Coefficient	-0.297	-0.261	1.000	0.890	0.053	-0.408	-0.588	0.288	0.945	-0.032
		Sig. (2-tailed)	0.039	0.070		0.000	0.715	0.004	0.000	0.045	0.000	0.825
		N	49	49	49	49	49	49	49	49	49	49
	interquartile	Correlation Coefficient	-0.336	-0.292	0.890	1.000	0.184	-0.673	-0.459	0.166	0.879	-0.063
		Sig. (2-tailed)	0.018	0.042	0.000		0.205	0.000	0.001	0.254	0.000	0.666
		N	49	49	49	49	49	49	49	49	49	49
	skewness	Correlation Coefficient	-0.169	-0.259	0.053	0.184	1.000	-0.185	0.163	0.135	0.072	-0.166
		Sig. (2-tailed)	0.245	0.073	0.715	0.205		0.203	0.262	0.356	0.621	0.253
		N	49	49	49	49	49	49	49	49	49	49
	kurtosis	Correlation Coefficient	0.268	0.172	-0.408	-0.673	-0.185	1.000	0.033	0.181	-0.489	0.140
		Sig. (2-tailed)	0.062	0.238	0.004	0.000	0.203		0.824	0.212	0.000	0.339
		N	49	49	49	49	49	49	49	49	49	49
	min	Correlation Coefficient	0.685	0.654	-0.588	-0.459	0.163	0.033	1.000	0.095	-0.531	0.296
		Sig. (2-tailed)	0.000	0.000	0.000	0.001	0.262	0.824		0.517	0.000	0.039
		N	49	49	49	49	49	49	49	49	49	49
	max	Correlation Coefficient	0.432	0.321	0.288	0.166	0.135	0.181	0.095	1.000	0.249	0.431
		Sig. (2-tailed)	0.002	0.025	0.045	0.254	0.356	0.212	0.517		0.084	0.002
		N	49	49	49	49	49	49	49	49	49	49
entropy	Correlation Coefficient	-0.294	-0.263	0.945	0.879	0.072	-0.489	-0.531	0.249	1.000	-0.071	
	Sig. (2-tailed)	0.040	0.067	0.000	0.000	0.621	0.000	0.000	0.084		0.627	
	N	49	49	49	49	49	49	49	49	49	49	49
heat	Correlation Coefficient	0.623	0.548	-0.032	-0.063	-0.166	0.140	0.296	0.431	-0.071	1.000	
	Sig. (2-tailed)	0.000	0.000	0.825	0.666	0.253	0.339	0.039	0.002	0.627		
	N	49	49	49	49	49	49	49	49	49	49	49

			Quadrant 3 Nonparametric Correlations									
			mean	median	std_dev	interquart	skew	kurtosis	min	max	entropy	heat
Spearman's rho	mean	Correlation Coefficient	1.000	0.989	0.196	0.420	-0.546	-0.433	0.766	0.474	0.296	0.565
		Sig. (2-tailed)		0.000	0.176	0.003	0.000	0.002	0.000	0.001	0.039	0.000
		N	49	49	49	49	49	49	49	49	49	49
	median	Correlation Coefficient	0.989	1.000	0.156	0.414	-0.574	-0.412	0.749	0.445	0.278	0.586
		Sig. (2-tailed)	0.000		0.285	0.003	0.000	0.003	0.000	0.001	0.053	0.000
		N	49	49	49	49	49	49	49	49	49	49
	std_dev	Correlation Coefficient	0.196	0.156	1.000	0.722	0.066	-0.199	-0.055	0.637	0.695	0.177
		Sig. (2-tailed)	0.176	0.285		0.000	0.653	0.171	0.708	0.000	0.000	0.223
		N	49	49	49	49	49	49	49	49	49	49
	interquartile	Correlation Coefficient	0.420	0.414	0.722	1.000	-0.298	-0.536	0.095	0.401	0.765	0.259
		Sig. (2-tailed)	0.003	0.003	0.000		0.037	0.000	0.517	0.004	0.000	0.073
		N	49	49	49	49	49	49	49	49	49	49
	skewness	Correlation Coefficient	-0.546	-0.574	0.066	-0.298	1.000	0.637	-0.139	0.176	-0.339	-0.410
		Sig. (2-tailed)	0.000	0.000	0.653	0.037		0.000	0.341	0.226	0.017	0.003
		N	49	49	49	49	49	49	49	49	49	49
	kurtosis	Correlation Coefficient	-0.433	-0.412	-0.199	-0.536	0.637	1.000	-0.222	0.150	-0.536	-0.393
		Sig. (2-tailed)	0.002	0.003	0.171	0.000	0.000		0.126	0.303	0.000	0.005
		N	49	49	49	49	49	49	49	49	49	49
	min	Correlation Coefficient	0.766	0.749	-0.055	0.095	-0.139	-0.222	1.000	0.441	-0.069	0.345
		Sig. (2-tailed)	0.000	0.000	0.708	0.517	0.341	0.126		0.002	0.640	0.015
		N	49	49	49	49	49	49	49	49	49	49
	max	Correlation Coefficient	0.474	0.445	0.637	0.401	0.176	0.150	0.441	1.000	0.265	0.149
		Sig. (2-tailed)	0.001	0.001	0.000	0.004	0.226	0.303	0.002		0.066	0.306
		N	49	49	49	49	49	49	49	49	49	49
entropy	Correlation Coefficient	0.296	0.278	0.695	0.765	-0.339	-0.536	-0.069	0.265	1.000	0.297	
	Sig. (2-tailed)	0.039	0.053	0.000	0.000	0.017	0.000	0.640	0.066		0.039	
	N	49	49	49	49	49	49	49	49	49	49	49
heat	Correlation Coefficient	0.565	0.586	0.177	0.259	-0.410	-0.393	0.345	0.149	0.297	1.000	
	Sig. (2-tailed)	0.000	0.000	0.223	0.073	0.003	0.005	0.015	0.306	0.039		
	N	49	49	49	49	49	49	49	49	49	49	49

			Quadrant 4 Nonparametric Correlations									
			mean	median	std_dev	interquart	skew	kurtosis	min	max	entropy	heat
Spearman's rho	mean	Correlation Coefficient	1.000	0.964	0.292	0.342	-0.206	-0.039	0.607	0.438	0.309	0.599
		Sig. (2-tailed)		0.000	0.042	0.016	0.156	0.788	0.000	0.002	0.031	0.000
		N	49	49	49	49	49	49	49	49	49	49
	median	Correlation Coefficient	0.964	1.000	0.233	0.304	-0.300	-0.053	0.546	0.390	0.241	0.571
		Sig. (2-tailed)	0.000		0.107	0.034	0.036	0.720	0.000	0.006	0.095	0.000
		N	49	49	49	49	49	49	49	49	49	49
	std_dev	Correlation Coefficient	0.292	0.233	1.000	0.817	0.158	0.197	-0.254	0.657	0.807	0.183
		Sig. (2-tailed)	0.042	0.107		0.000	0.278	0.175	0.078	0.000	0.000	0.208
		N	49	49	49	49	49	49	49	49	49	49
	interquartile	Correlation Coefficient	0.342	0.304	0.817	1.000	-0.014	-0.150	-0.068	0.422	0.827	0.264
		Sig. (2-tailed)	0.016	0.034	0.000		0.924	0.304	0.644	0.003	0.000	0.067
		N	49	49	49	49	49	49	49	49	49	49
	skewness	Correlation Coefficient	-0.206	-0.300	0.158	-0.014	1.000	0.493	0.026	0.434	0.088	-0.145
		Sig. (2-tailed)	0.156	0.036	0.278	0.924		0.000	0.860	0.002	0.550	0.319
		N	49	49	49	49	49	49	49	49	49	49
	kurtosis	Correlation Coefficient	-0.039	-0.053	0.197	-0.150	0.493	1.000	-0.243	0.629	-0.028	-0.048
		Sig. (2-tailed)	0.788	0.720	0.175	0.304	0.000		0.092	0.000	0.848	0.743
		N	49	49	49	49	49	49	49	49	49	49
	min	Correlation Coefficient	0.607	0.546	-0.254	-0.068	0.026	-0.243	1.000	0.039	-0.158	0.322
		Sig. (2-tailed)	0.000	0.000	0.078	0.644	0.860	0.092		0.790	0.279	0.024
		N	49	49	49	49	49	49	49	49	49	49
	max	Correlation Coefficient	0.438	0.390	0.657	0.422	0.434	0.629	0.039	1.000	0.530	0.323
		Sig. (2-tailed)	0.002	0.006	0.000	0.003	0.002	0.000	0.790		0.000	0.024
		N	49	49	49	49	49	49	49	49	49	49
entropy	Correlation Coefficient	0.309	0.241	0.807	0.827	0.088	-0.028	-0.158	0.530	1.000	0.406	
	Sig. (2-tailed)	0.031	0.095	0.000	0.000	0.550	0.848	0.279	0.000		0.004	
	N	49	49	49	49	49	49	49	49	49	49	
heat	Correlation Coefficient	0.599	0.571	0.183	0.264	-0.145	-0.048	0.322	0.323	0.406	1.000	
	Sig. (2-tailed)	0.000	0.000	0.208	0.067	0.319	0.743	0.024	0.024	0.004		
	N	49	49	49	49	49	49	49	49	49	49	