12-1995

Computer-Aided Diagnosis of Mammographic Masses

William E. Polakowski

COMPUTER-AIDED DIAGNOSIS OF MAMMOGRAPHIC MASSES

THESIS

William E. Polakowski
Captain, USAF

AFIT/GEO/ENG/95D-02

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GEO/ENG/95D-02

COMPUTER-AIDED DIAGNOSIS OF MAMMOGRAPHIC MASSES

THESIS

William E. Polakowski
Captain, USAF

AFIT/GEO/ENG/95D-02

AFIT/GEO/ENG/95D-02

Computer-Aided Diagnosis of Mammographic Masses

THESIS

Presented to the Faculty of the Graduate School of Engineering

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Electro-Optical Engineering

William E. Polakowski, BS

Captain, USAF

December, 1995

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U. S. Government.

Acknowledgements

First, I would like to acknowledge God's hand in this whole endeavor, and my wife and family's incredible support while I have been at AFIT. God has given me the life and the perspective to accomplish incredible things while staying true to myself. My wife stepped in at home and kept our household running and really did neat activities with our kids, Ana and Gen to help them grow while allowing me time to study. I would also like to acknowledge the help of Dr Steve Rogers, my thesis advisor; and Major Dennis Ruck, Capt Rick Raines, and Major Jeff Hoffmeister, my committee members, for all their help and support. The success of this thesis was due to their efforts at educating me in the medical pattern recognition area. In addition, I appreciated the help of all those at the Wright Patterson AFB hospital for working so diligently to make the MDIS system work and the mammograms available so we could acquire our database. In addition, Mr Bruce Noren from Wright Laboratory's Avionics Directorate provided computer hardware to transfer the images to our system. And, finally, thanks goes to the breast cancer working group members who provided endless ideas and computer help: Dave Bramlage, Capt Ron Dauk, 1st Lt Drew McCandless, and Capt Amy Magnus. Thank you all!

William E. Polakowski

Table of Contents

List of Figures

Abstract

This thesis describes a new Air Force Institute of Technology (AFIT) system used to find regions of interest, specifically masses, in digitized mammograms. After finding these regions, the second contribution of this work was to identify malignant masses. This AFIT system achieves a sensitivity of 92 percent for segmentation of malignant masses and a classification accuracy of 100 percent for the segmented malignant masses. These results are from the AFIT, biopsy-proven database of 272 images (12 bit, 100μm) with 36 malignant mass, 53 benign mass, and 183 microcalcification or healthy images. Of the 53 biopsied benign cases, 74 percent were rejected or correctly classified by the algorithm as benign. The algorithm/architecture is based on the Model Based Vision (MBV) approach which has never been applied to breast cancer diagnosis. The Focus of Attention (segmentation) Module algorithm relies on a physiologically motivated Difference of Gaussians (DoG) bandpass frequency filter to highlight mass-like regions in the mammogram. These regions were then passed through size and texture tests to reduce the number of false regions from 8.4 to 1.8 per image. The segmented regions were indexed (a stage of the MBV architecture) as to their hypothesized class: large mass or medium mass. Size, shape, contrast, and Laws texture features were used to develop the Prediction Module's mass model. Statistical and derivative-based feature saliency techniques were used to determine the best features. Nine features were chosen to define the model. Using this model, the regions were then classified using a multilayer perceptron neural network architecture trained with an imbalanced training set weight update algorithm to achieve the above results.

Computer-Aided Diagnosis of Mammographic Masses

I. Introduction

Computer pattern recognition techniques and systems have successfully been applied to many military and non-military problems (2, 3, 4, 5, 6). Many of these pattern recognition systems have relied on human-visual-system-based pattern recognition principles and human-brain-modeled neural networks to quickly and accurately classify the patterns of interest with low false alarm rates (5, 6, 7, 8, 9, 10). This thesis transitions military pattern recognition techniques to hospital use. In this medical arena, breast cancer diagnosis using computer pattern recognition techniques is ready to transition to general hospital use (6). The following algorithm presents a new approach to implementing a Computer-Aided Diagnosis System (CADx) for breast cancer detection and diagnosis.

1.1 Breast Cancer

The National Cancer Institute estimated that in the United States in 1994 over 182,000 women were newly diagnosed with breast cancer, with over 46,000 deaths per year (11). Current estimates predict the rate will increase for the foreseeable future (10, 11, 12). The lifetime risk that a woman will develop breast cancer is 1 in 10 assuming the average life expectancy of 79 years, or it is 1 in 8 assuming longevity of 95 years (13). Breast cancer is the second leading cause of death from cancer (following lung cancer) for women (12). In addition to the trauma for the woman and her family, this places a huge strain on the radiologists, doctors, and the medical system in terms of the number of mammograms to diagnose, biopsies to perform, and if necessary, treatment that must be accomplished.

The difficulty of this diagnosis is increased by two factors. First, as in many medical imaging areas, normal tissue presents a very cluttered background to the radiologist. Considering that the vast majority of mammograms are benign, the radiologists have trouble seeing the low contrast cancers in the normal breast tissue. The second reason is that there are many other normal or benign structures in the breast that look very similar to

the two types of cancer found in the breast: Mass lesions and microcalcifications. Mass lesions are lumps of tumorous tissue, but they appear very similar to glands or dense portions of the breast, and they are frequently hidden inside of those regions. Malignant microcalcifications are groupings of tiny calcium deposits associated with breast cancer. Frequently there are numerous non-cancerous calcification deposits lining blood vessels or cysts or just scattered throughout the breast. Thus, due to the low contrast and many similar noncancerous structures of the breast, the cancerous regions are very difficult to diagnose (12, 14).

1.2 Traditional Breast Cancer Diagnosis

Radiologist diagnosis of breast cancer using X-ray film mammograms has allowed for more efficient diagnosis of breast cancers at an earlier stage of development than simply relying on locating masses by palpation during yearly breast exams. Silverberg et al., (15) state that mortality has been reduced by 30 to 35 percent for those women who participate in yearly mammogram screening. Yet, problems persist. The overworked radiologists misdiagnose 10 to 30 percent of the malignant cases; two-thirds of which were evident in the mammogram in retrospect (6). These are called false negative diagnoses which could be fatal. And, of the cases sent for surgical biopsy, only 10 to 20 percent are actually malignant (6). The remaining 80 to 90 percent are called false positives and they have the effect of causing unnecessary trauma to the patient and are a burden on the medical system. Having two radiologists read each mammogram has been suggested, but that would only add to radiologist workload and fatigue. A Computer Aided breast cancer Diagnosis (CADx) system used as an aid to the radiologist could help to ease the workload by helping to correctly diagnose the missed (20 percent) malignant cases and reduce the number of unnecessary (85 percent) surgical biopsies.

1.3 Computer-Aided Breast Cancer Diagnosis

The use of the CADx system then, is to check on the radiologist's diagnosis. After the radiologist makes their diagnosis, the films could be placed in the computer's digitizer, and then the computer's diagnosis could be reviewed on the spot or at a later date. A

complete CADx system automatically does all the steps from receiving a film mammogram to outputting the diagnosis and the cancer's location. The first step is image acquisition. Currently, CADx systems rely on a digitized X-ray film as the input to the system versus direct digital acquisition. While direct digital acquisition of the mammogram using stereotactic imaging techniques has been introduced for biopsies, only a small portion of the breast is imaged at any one time. For the true computer diagnosis system, a full mammogram is needed; so, in the second step of the process, the computer must automatically segment, or identify, Regions Of Interest (ROIs).

This initial segmentation of the image breaks the huge (1800 x 2400 element) digitized mammogram into smaller (i.e. 140 x 140 element) ROIs that are easier to work with. The segmentation done by the Focus of Attention Module must be sensitive enough not to miss any cancerous regions (to eliminate the false negative diagnoses), but it cannot overload the highly complex matcher/classifier algorithms with too many noncancerous regions. This would slow down the system and potentially result in too many false positive diagnoses, with a huge rise in unnecessary biopsies. Once the ROIs are found from this segmentation of the entire X-ray film, features are extracted, and classification algorithms are implemented on each ROI to determine malignancy.

These computationally intensive classification algorithms need to be able to identify both types of cancer: clusters of small microcalcifications and mass lesions. The algorithms use various 'features' in the ROIs as inputs. Features such as the size, shape, or intensity of the patterns in the ROIs, can be input as numbers into the classifier. For example, the higher the number of microcalcifications present in the ROI, the more likely that the ROI is cancerous. So, one input to a classifier could be the number of small pixel groupings in the ROI above a certain threshold intensity. Given the set of features, various classification techniques can be used for this problem, but one of the most efficient techniques involves neural networks. These networks are loosely patterned off the networks of neurons in human brain tissue (7). Information from the ROIs are fed into this neural network, and they efficiently provide the radiologist with the computer's analysis of the mammogram. To provide the radiologist with that 'second opinion', the CADx system will

identify each region of the breast that contains suspicious tissue and provide the indicated cancer diagnosis.

1.4 Statement of the Problem

Radiologists need help identifying difficult-to-see mass lesion cancers to decrease the number of cancers missed and to reduce the number of unnecessary biopsies of benign tissue. This thesis will develop algorithms to focus the radiologist's attention on suspect mass-like regions of the full mammogram. It will match those regions with model-based predictors of normal breast tissue, benign masses, and malignant masses to provide the radiologist with the most probable diagnosis of each region of the mammogram. It can be used as an in-situ diagnostic tool or as a diagnosis review tool at a later date.

1.5 Scope

This CADx system was proposed as a backup to or second look for the radiologist to review the potentially malignant areas in the mammogram. The algorithms that define the CADx system were written mostly in the Matlab software environment with a few C routines used for better efficiency. The Focus of Attention and Matching algorithms were used to detect and diagnose masses from a database of 300 radiologist diagnosed and/or pathology-truthed mammograms from the Wright-Patterson Air Force Base (WPAFB) Hospital. The mammograms were digitized to 12 bits of grayscale and 100μm resolution, and were cropped to 2048 x 1024 pixels.

1.6 Methodology

A Model Based Vision (MBV) architecture (16) is used to focus the radiologist's attention to indexed regions of a mammogram. The initial Focus of Attention module implemented a Difference of Gaussians (DoG) (17, 18) human-based visual system filter to identify potential ROIs. After dynamically thresholding the filtered image and rank ordering the ROIs with an area to perimeter ratio to reduce the number of false ROIs, the ROIs were indexed into mass size categories. Based on the indexing label of an ROI,

features were extracted and matched to predicted models of that type of tissue. The resulting hypothesis would then be presented to the radiologist.

1.7 Overview

The remainder of the thesis is structured as follows: Chapter II examines breast cancer diagnosis in more depth, discusses relevant research, and defines MBV. Chapter III discusses the WPAFB hospital database. Chapter IV discusses the methodology used in the CADx system. Chapter V presents the results of the Focus of Attention Module, the Indexing Module, the Prediction Module, and the Matching Module tested using the WPAFB database. Chapter VI discusses the conclusions regarding the usefulness of the system. The database description, the medical protocol, the image acquisition process, and the Matlab code are provided in the appendices.

II. Background

This chapter discusses the relevant contributions of other researchers in the areas of focusing a radiologist's attention to specific regions of a mammogram and in the area of feature selection and region classification. It also provides background into the Model Based Vision (MBV) approach to pattern recognition. The MBV approach consists of the Focus of Attention Module, the Indexing Module, the Prediction Module, and the Matching Module.

2.1 Breast Tissue

Numerous groups have demonstrated the feasibility of computers to analyze and classify different types of textures like those found in breast tissue (6). To understand these techniques, a discussion of breast structures and textures follows. Figures 1 and 2 illustrate various mammograms from the WPAFB database.

2.1.1 Normal Tissue. The parenchymal pattern of healthy breast tissue is a conglomeration of dense tissue, supporting ligaments, connective tissue, milk-producing glands and ducts, and calcium deposits. This variety of normal tissue types presents a highly-variable and well-structured image to the radiologist. Dense tissue, supporting ligaments, connective tissue and glands can by themselves or in concert obscure or mimic the appearance of malignant tumors. The tissue shown in Figure 1c illustrates this, since it appears very similar to the malignant mass shown in Figure 1b. In some literature, these natural tumor-like structures are labeled benign tumors. The remaining normal structures, ducts, and calcium deposits also appear on the X-ray mammogram in a manner similar to malignant groupings of microcalcifications. Complicating the diagnosis further is the overall decrease in density of the normal breast tissue with age. Thus, mammograms from younger women have much more structure and variability in gray-levels than those from older women (14).

2.1.2 Mass Tumors. Mass lesions are lumps of tumorous tissue, but they appear very similar to glands, cysts, or dense portions of the breast, and they are frequently

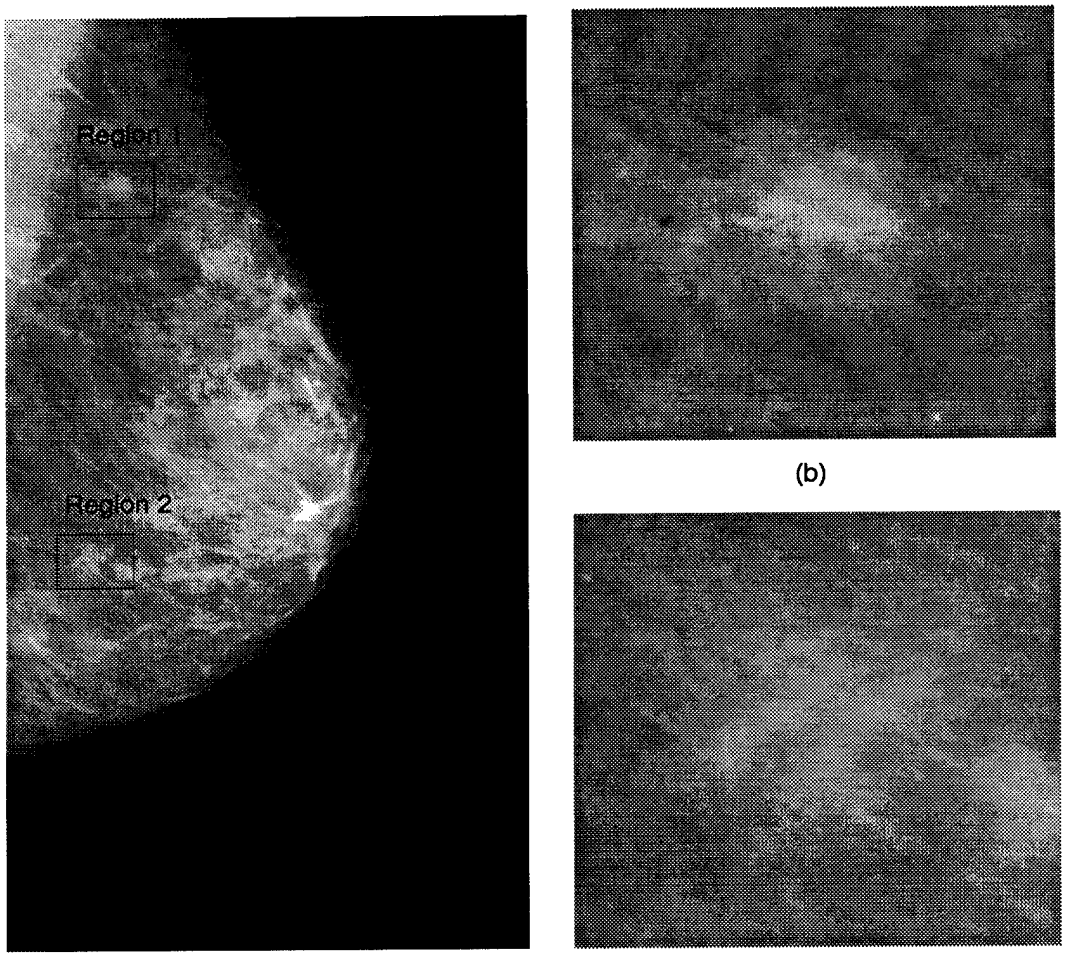


Figure 1. (a) Digitized Mammogram 1 with a Malignant Mass
(b) Malignant Mass in Region 1
(c) Normal Tissue in Region 2

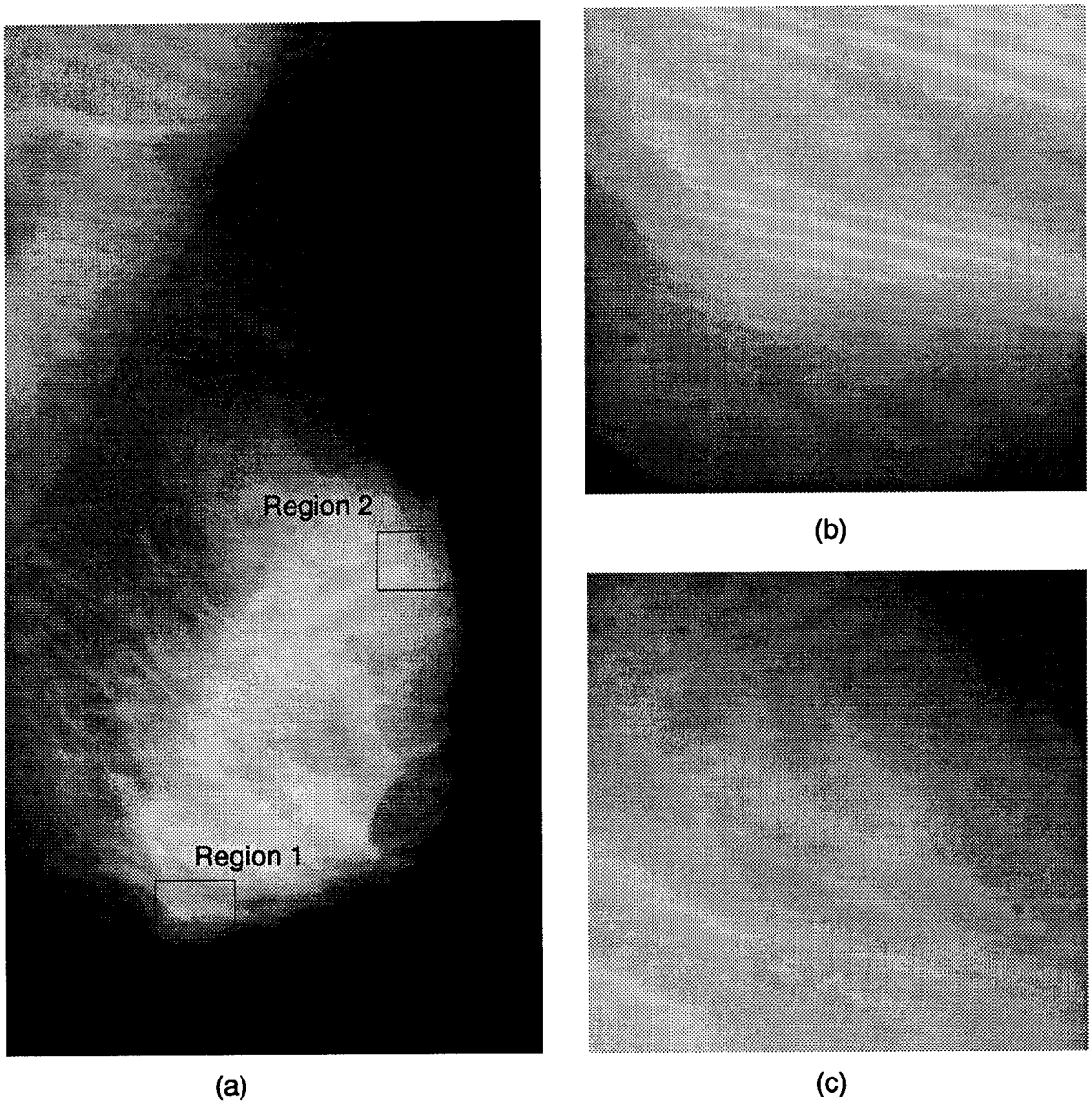


Figure 2. (a) Digitized Mammogram 2 with a Malignant Mass
(b) Malignant Mass in Region 1
(c) Normal Tissue in Region 2

hidden inside of those regions. Malignant masses typically are lower contrast with respect to the surrounding tissue than benign masses, but both are brighter than the surrounding tissue. Malignant masses have other characteristics that distinguish them from healthy tissue. Malignant masses occur in two basic categories: Stellate and circumscribed (12, 14). Radiologists, in their diagnosis, first determine the margination of the mass, whether it is well circumscribed by a fatty (darker) halo, or spiculated (radiating tendrils from a central mass) and poorly defined (14). A strong halo signature, with a well-defined border is usually benign, but the ill-defined, spiculated masses are more likely to be cancerous. Seventy to Eighty percent of breast cancers are of this type (14), but as Swann points out (19) none of these features are absolute. Beyond margination, the determination of malignancy is judged by the shape, size, and pattern or texture of the suspect tissue. These same features can be used in computer diagnosis (20). This thesis used shape, size, contrast, and texture features similar to the inherent techniques radiologist's use in their diagnosis (20).

2.1.3 Microcalcifications. Calcifications of some type are found in the majority of mammograms. Malignant microcalcifications are groupings of tiny calcium deposits that are associated with breast cancer, but appear very similar to non-cancerous calcification deposits lining blood vessels or cysts or just scattered throughout the breast that contains these malignant microcalcifications (14). Most calcifications are much brighter than the surrounding tissue, but their small size (100-300μm) makes them difficult to detect. They are usually distinguished from benign calcifications by their margination, number per volume, shape, size, and distribution (20).

All of these factors combine to make breast cancer detection and diagnosis very challenging.

2.2 Breast Cancer Detection and Diagnosis

Initially, the only way to detect breast cancer was during a breast exam when a palpable mass was detected. With the advent of Mammography, a radiologist could detect masses and even microcalcifications significantly prior to the cancer becoming palpable.

Technology has now advanced so computers are able to aid radiologists in their diagnosis of breast cancer, and in the future, the X-ray film digitization process will be replaced with purely digital mammogram acquisition and computer-aided diagnosis. Giger (6) provides an extensive review of all aspects of this CADx area.

2.3 Traditional: Radiologist Diagnosis of X-Ray Film

Radiologist diagnosis of breast cancer using X-ray film mammography is currently the most effective method for decreasing the severity of breast cancer. Radiologists' correct diagnosis of malignant cancers range from 70 to 90 percent, but to attain these high results they send 4 to 5 benign masses for biopsy for every malignant mass biopsied (6). These problems can be attributed to several factors including poor image quality, radiologist fatigue, and human oversight. A Computer Aided breast cancer Diagnosis (CADx) system used as an aid to the radiologist could help to ease the workload by helping to correctly diagnose the missed malignant cases and reduce the number of unnecessary surgical biopsies. The first step towards this full CADx system was to use computers to enhance the images so the radiologist could identify the differences between the cancerous and noncancerous regions more easily.

Much work has been done in this area to enhance mammograms for the radiologist (21, 22, 23, 24, 25, 1). Many techniques have used wavelets or multiresolution analysis to weight specific frequency decomposition scales. For example, Laine, et al., (21) implemented an approach for mammographic feature enhancement based on the image's multiresolution representation. These multiresolution coefficients from the dyadic, ϕ, and hexagonal wavelet transforms were modified by nonlinear operators and then used to reconstruct an enhanced image. This method enhanced the cancerous regions for easier detection. Yoshida et al., at the University of Chicago, has used the Least Asymmetric Daubechies' wavelets to enhance and classify microcalcifications (22). For enhancement, they modified the weights of certain wavelet coefficients to enhance microcalcifications and masses. Yoshida proceeded one step further (see Section 2.4.2) and developed a computer diagnosis algorithm to work on the enhanced image (22).

Another way to enhance the image is to reduce the noise effect. By using Gabor filters, cosine transforms, and other wavelet methods, the underlying 'noisy' texture of the image can be found and then subtracted from the original image (23). Dhawan, et al., used an adaptive filter based on the local contrast in the image (24, 25). It seemed to increase the contrast in the image without increasing the noise.

One final researcher, Lai, et al. (1), used median filtering and selective averaging to enhance the mammograms, but their enhancement attempts were strictly for computer diagnosis (Section 2.4.2). The trend is continuing to place the emphasis on computer diagnosis rather than just mammographic enhancement for radiologists.

2.4 CADx: Radiologist With Computer-Aided Diagnosis

Research into computer-aided breast cancer diagnosis has been going on since 1979 (26), with many individuals and groups contributing techniques to solve this problem. Since the properties of masses and microcalcifications are so different, the techniques used to detect them are very different too. Various techniques are described below for each area.

2.4.1 Current Research - Microcalcifications. Microcalcification detection and classification is a challenging job given the fact that the microcalcifications are generally only a few pixels in size for a 100μm resolution image and not even visible at lower resolutions. To complicate matters, the mammograms contain severe background noise that is comparable to the signature of the microcalcifications. Thus most of the literature has focused on classifying microcalcifications in hand-segmented ROIs rather than computer segmenting the image, because computer segmentation schemes generally result in a high number of false ROIs.

Recent work by Chitre and Dhawan used second-order gray level histogram-based features for microcalcification classification in 100 difficult-to-diagnose cases (160μm resolution). They used entropy, contrast, and angular second moment based features, among others, in a neural network architecture for a 73 percent true-positive rate and a 35 percent false-positive rate (10, 27, 28, 29, 30).

Capt Kocur's Computer-Aided Breast Cancer Diagnosis thesis work in 1994 (9, 31) built on Dhawan and Chitre's work but explored many different techniques for and aspects of microcalcification classification. Her thesis investigated feature extraction and image classification of hard-to-detect microcalcifications in 94 digitized mammograms (160μm resolution). She implemented a variety of features and feature saliency techniques within a neural network classification architecture. The angular second moment features based on the second-order gray-level histogram obtained 62 percent correct classification. The Karhunen-Loeve Transform features used to do an eigenvector coordinate transformation to obtain the eigenmass coefficients achieved 65 percent correct classification. Using biorthogonal wavelet features, the accuracy was increased from 74 percent to 88 percent after performing some feature selection and reduction techniques and after using neural network decision boundaries (9, 31).

As mentioned before, Yoshida et al., used the Least Asymmetric Daubechies' wavelets for microcalcification enhancement (22). He then used global and local thresholding, morphological erosion, and texture analysis on the enhanced image to achieve 85 percent correct classification but at a cost of 5 false positives per image (22).

Many other researchers have tried various approaches to detect and classify microcalcifications, and the best summary of these techniques is found in Giger's Computer-aided Diagnosis article (6). Current research at the Air Force Institute of Technology (32, 33) uses wavelets and morphological digital image processing to detect/diagnose microcalcifications.

2.4.2 Current Research - Masses. While, in general, the literature on microcalcifications focused on classification, the literature on masses includes the segmentation and classification aspects of the problem.

Brzakovic, et al., detected and classified large masses in 25 mammograms using multiresolution analysis combined with fuzzy pyramid linking for the segmentation step, and Bayes classifiers based on the shape and intensity characteristics of the masses for the classification step (34). Brzakovic trained on ten images and then tested using all 25 images. Of the 20 tumors present, they missed two malignant tumors and misclassified one

			-1	-1	-1			
		-1	-1	0	-1	-1		
	-1	-1	0	1	0	-1	-1	
-1	-1	0	1	1	1	0	-1	-1
-1	0	1	1	1	1	1	0	-1
-1	-1	0	1	1	1	0	-1	-1
	-1	-1	0	1	0	-1	-1	
		-1	-1	0	-1	-1		
			-1	-1	-1			

Figure 3. Lai's template for matching tumors of five pixels in diameter (1).

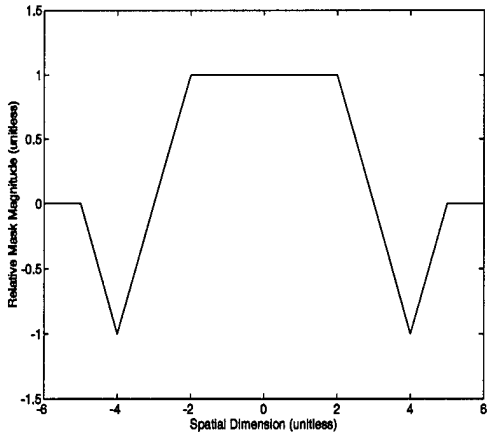


Figure 4. One dimensional plot of Lai's template.

benign tumor as malignant, but the full mammograms were only digitized at low resolution in 256x256 arrays.

Lai, et al., worked from selective median filtered images to attempt to segment and classify circumscribed masses (1). They used template matching to segment the image and histogram tests to classify the masses. An example of one of their 12 templates with a one dimensional plot of the template are shown in Figures 3 and 4. It relies on the three characteristics of tumors: brightness contrast, uniform density, and circular shape. On a database of 17 images they achieved 100 percent true-positive detection with 1.7 false-positives per image.

-1	0	2	0	-1
-4	0	8	0	-4
-6	0	12	0	-6
-4	0	8	0	-4
-1	0	2	0	-1

(a)

-1	-2	0	2	1
-4	-8	0	8	4
-6	-12	0	12	6
-4	-8	0	8	4
-1	-2	0	2	1

(b)

1	-4	6	-4	1
-4	16	-24	16	-4
6	-24	36	-24	6
-4	16	-24	16	-4
1	-4	6	-4	1

(c)

1	0	-2	0	1
2	0	-4	0	2
0	0	0	0	0
-2	0	4	0	-2
-1	0	2	0	-1

(d)

Figure 5. Kegelmeyer's four Laws kernels:(a) l5s5 (b) l5e5 (c) r5r5 (d) e5s5

Kegelmeyer used a local oriented edge analysis algorithm, Laws texture analysis, and a binary decision tree classification algorithm to detect and classify stellate lesions. The binary decision tree determines the probability of malignancy of each pixel based on the edge and texture analysis processes (35, 36, 37). The four Laws texture features Kegelmeyer used were derived from the convolution of the l5s5, l5e5, r5r5, and e5s5 kernels with the image. The Kernels are shown in Figure 5, but a more detailed description is in Section 4.3. He tested the algorithm on a portion of the University of South Florida database (200 μm resolution). Out of one hundred images, he used their twelve stellate lesion examples and their fifty normal images. The results were 100 percent probability of detection with only 0.27 false alarms per image (36). In a second test conducted to determine the increased performance of radiologists when shown the CADx results, the algorithm detected 66 of 68 spiculated lesions, correctly classified 82 percent of the spiculated lesions, and had 0.28 false positives per image on a database of 84 cases (4 views per case) at 240 μm resolution (37).

Wei et. al., from the University of Michigan used multiresolution texture analysis to differentiate masses from normal tissue (38). The texture features were derived from the spatial gray level dependencies (angular second moment) of the image and of the wavelet decomposition images. They constructed their database of 672 ROIs by hand-segmenting four ROIs from each of 168 images. Thus, each image had one tumor and three normal

tissue ROIs represented in the database. They achieved a 95 percent true-positive fraction with a false-positive fraction of 55 percent. Since their database consisted of one mass (malignant or benign) and three normal tissue ROIs from each mammogram, these results translate to over two out of three or four benign regions per mammogram being false-positives (38).

Giger and Yin, et. al., from the University of Chicago, have worked towards implementing a complete CADx system to identify suspect regions of mammograms for both microcalcifications and masses (6, 39, 40, 41, 42). Their technique for segmentation of masses is based on a comparison of the right and left mammograms of the same view (CC or MLO). Both mammographic views show a similar pattern and symmetry for normal parenchymal tissue, but for most cancer cases, the cancer only appears in one breast. Using an autonomous, non-linear, bi-lateral subtraction technique to eliminate most of the normal tissue results in unusual masses being highlighted in either breast. Radiologists regularly compare the opposing mammographic views in their diagnosis, and Giger and Yin have found this idea to be superior to their single image processing techniques (41). When tested on 77 patient cases (308 mammograms), they achieved 91 percent sensitivity with a false-positive rate of 6.5 per image (42). In earlier tests, on a smaller subset of the data (46 pairs of mammograms), they achieved 95 percent sensitivity with a false-positive rate of 3 per image (40). The autonomous classification scheme is based on the spiculation of the masses detected by the segmenter. They used a number of morphological and averaging steps to determine the area and boundary of the masses at the different steps in their morphological process. Comparing the features at each of these steps yielded a 97 percent true-positive rate and a 79 percent false positive rate on a database of 50 masses. Their full CADx system combines these mass algorithms with the microcalcification algorithms, and it is currently in clinical testing.

Figure 6 summarizes the mass lesion results of the authors mentioned above. Kegelmeyer's and Yin's second entry and Giger's entry are the most relevant results to this thesis.

Researcher	Database	False Alarms per Image	True Positive Fraction	False Positive Fraction	Citation
Brzakovic	25	0.8	0.8	na	(34)
Lai	17	1.7	1.0	na	(1)
Kegelmeyer 1	62	0.27	1.0	na	(36)
Kegelmeyer 2*	340	0.28	0.97	na	(37)
Wei	168 ROIs	>2	0.95	0.55	(38)
Yin 1	92	3	0.95	na	(40)
Yin 2*	308	6.5	0.91	na	(38)
Giger*	50 ROIs	na	0.95	0.79	(38)

Figure 6. Summary of the Results for the Detection of Masses in Mammograms. * The Most Relevant to This Work. na = not available/applicable.

2.5 The Model Based Vision CADx Process

The Model Based Vision (MBV) approach consists of the Focus of Attention Module (FOA), the Indexing Module, the Prediction Module, the Feature Selection Module and the Matching Module (16). The FOA module is similar to traditional segmentation, the Prediction Module and Feature Selection Module are similar to feature extraction/selection, and the Matching module is similar to traditional classification. Figure 7 shows the architecture.

2.5.1 Focus of Attention: Segmentation. The Focus of Attention (FOA) module identifies regions in an image that require more attention. It is an information reduction step that highlights regions in the mammogram that a radiologist would most likely spend more time on during their diagnosis. It boils down to the art of identifying the regions of interest in an image for the application of classifier algorithms. For example, in a digitized mammogram, the segmenter's job would be to determine if there were suspicious patterns or objects in the mammogram that required further study. The regions that contained the suspicious Regions Of Interest (ROIs) would be passed to the feature extraction algorithms. There are many ways to do this, but the point is that, in some way, an 1,800 x 2,400 element digital mammogram is boiled down to a few small ROIs (140 x 140 elements in this case) for closer study.

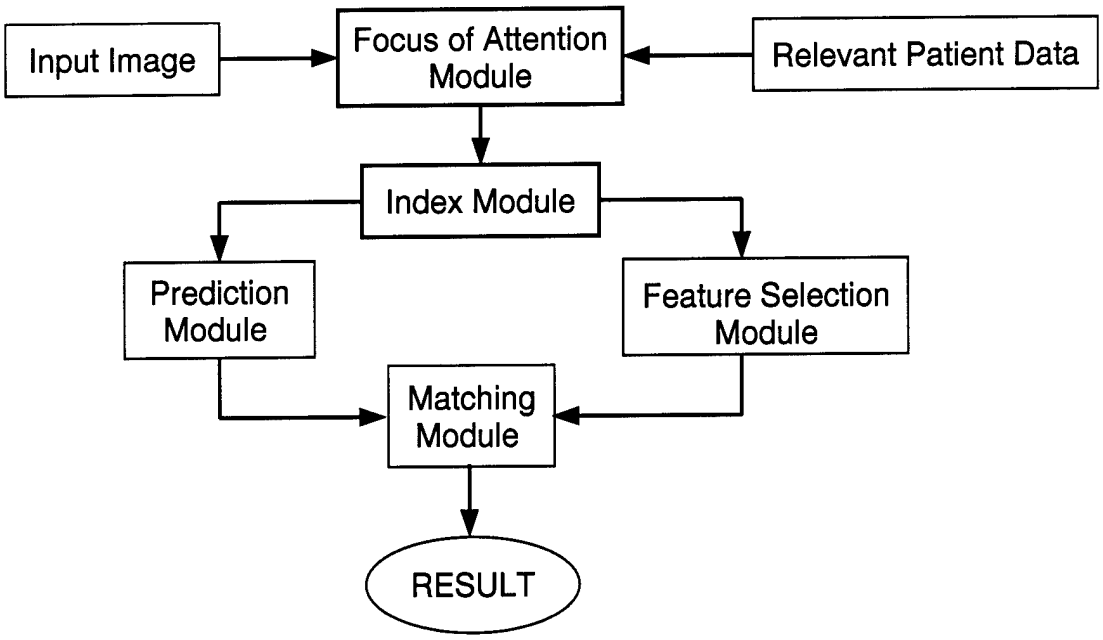


Figure 7. The Model-Based Vision Flow Diagram

2.6 Indexing

The Indexing module receives the ROIs from the FOA module and separates the ROIs into categories. It does this by applying various hypothesis tests on the ROIs to make a high level prediction of the type of object in an ROI. Example demarcations for hypothesis generation could be based on the size or shape of the object in the ROI. Continuing with the size-based hypothesis, the features extracted from and the matching model used for a small object could be different than those used for a large object. Thus the features, models, and hypothesis tests are tailored for specifically indexed ROIs.

2.7 Prediction / Feature Selection

The Prediction Module develops a model of the tissue types the indexer specifies; in this case, different sized malignant masses. It builds the model by extracting the best features that define the characteristics of a malignant mass. A subset of the available data, the 'training' set is used to develop this model, while the rest of the data is used to test the model. The job of the feature extractor then is to take the information present in an individual ROI and reduce it to a few pieces of discriminantly useful information that

can be fed into a matching or classification algorithm: in this case a neural network. For example, the ROI consists of 140 x 140 pixels with 12 bits of grayscale. An energy texture feature extractor finds the energy in the various images after they have been convolved with various texture kernels. Some features may be correlated with each other (redundant) or useless. Features whose class probability density functions (pdfs) having the least overlap will be the most discriminantly useful features. The goal is to find the fewest number of best features that allow the classifier to separate the data into the proper classes.

There are a number of reasons to reduce the number of features ultimately used for developing the model. The first is that the matching/classification will be much faster with fewer features. The second, and more important reason, is to ensure the model will generalize to work for all cases in the clinical environment. There are a number of rules used to determine the maximum number of features one should use. These rules are based on a neural network classification architecture, which is used in the Matching/Classification Module. Foley's rule (7) states that there needs to be three times the number of training samples in each class than there are features, while Uncle Bernie's rule (7) states that the total number of training samples should be ten times the number of connections used in the neural network architecture. Thus, depending on the amount of data, a certain maximum number of features needs to be found to develop the model.

There are a number of techniques available to determine these best features. The most basic is the probability of error metric. This technique picks the most relevant features by the overlap of their class conditional probability density functions (pdfs). The less overlap, the higher the discriminant power of the feature. The second method is called the f-ratio. For each feature, this method compares the distance between the means of the classes and their standard deviations. The larger the difference between the means and the smaller the standard deviations, the higher the f-ratio and the higher the feature's saliency. The equation from Parson's text (43) is shown below:

$$f - ratio = \frac{(\mu_{class1} - \mu_{class2})^2}{\sigma^2_{class1} + \sigma^2_{class2}} \quad (1)$$

A third method, developed by Steppe (44, 45) eliminates features one by one that have the least positive effect on the classification accuracy of the remaining features. The

exhaustive algorithm trains ten neural networks for each feature selection level, and maintains the overall classification results for each level. A certain subset of features will yield the best results for this method. Lee and Langrebe (46, 47) developed a fourth method which only uses the correctly classified samples output from a Gaussian or Parzen Window classifier. The decision boundary is approximated using the normal vector to the decision boundary along the vectors connecting the closest samples on either side of the decision boundary. In this manner the eigenvectors with the highest magnitude identify the most relevant features. The last technique, developed by Ruck (48, 49), examines the derivative of the output of a trained neural network with respect to each training sample's features. The features that have the biggest effect on the classification output will have the highest derivative. Each feature saliency method mentioned will yield similar results. Steppe seems to provide the most exhaustive results, but for faster analysis, the Ruck method is suitable for most applications.

While the Prediction Module is used to select the features for the models of indexed tissue types, the Feature Selection Module selects those same features for each sample. These features from each Module are then compared to determine how closely any given sample matches the models of the tissue types.

2.8 Matching / Classification

2.8.1 Traditional. Many classification algorithms or discriminant functions can be applied to the pattern recognition problem. For details on these algorithms, consult Duda and Hart's book (3) or Fukunaga's book (2) on pattern recognition. Bayesian classifiers minimize the probability of error for a given set of features. They provide the optimum solution since, they place the class decision boundary at the point of least overlap of the pdfs. But these classification algorithms are too computationally complex to use on the entire image. For example, the Gaussian classifier uses Gaussian distributions about each class over the entire feature space. Then the Bayesian discriminant function is calculated from the class means, variances, and covariances. Another type of classifier, the nonparametric KNN classifier, identifies the test sample features with the K nearest training sample's features and determines the class by voting on which class's features are closest.

Finally, neural networks can approximate the results of an optimum Bayes classifier by approximating the Bayes optimal discriminant function (50). Once the neural network is trained, the testing and classifying are potentially much quicker than for the conventional Bayes classifiers.

2.8.2 Model-Based Matching. Model-Based classification computes a probability that the test sample's features match the features generated from a model. Given a high enough correlation, the sample will be classified as a certain class with a stated probability. If the correlation of the features is not high enough, the hypothesis generated in the indexing step may be revised to try to match the sample to another class, or to change a parameter in the model to better match the sample. For example, given a tank template for an infrared image, the aspect or orientation of the tank hypothesized and then used to generate the model signature may not precisely match the actual tank's orientation. The hypothesis then might be changed and the model tweaked to better match the sample.

2.8.3 Imbalanced Training Sets. For this thesis, a modified multilayer perceptron neural network classifier was used to classify the samples versus, strictly matching them to the model. The concepts are similar, but the actual techniques are different. One method compares samples to a model, while the other classifies samples by which side of the learned decision boundary the samples fall upon. For the neural network case, it is important to note that in practice, the number of false ROIs and benign cases is far greater than the number of malignant cases. This results in imbalanced malignant and benign training sets. Since the network tries to reduce the overall mean square error for all samples, this means that it usually classifies the dominant class samples correctly at the expense of the smaller class's samples. In practice, correct malignant classification (the smaller class) is more critical than benign classification (the dominant class). Therefore, modifications to the standard neural network learning rule were used to reduce the impact of the imbalanced training set problem. The method used was developed by Anand, et al. (5). He trained the network in batch mode (sigmoidal activations) but with the results separated into the two classes. Then the bisector of the two error gradients was used to determine the weight update for each epoch. The effect is to force the mean square error of both classes to

zero. The network has a single clamped output node, but can have any number of hidden nodes. The only drawback is that the batch technique tends to get stuck in local minima, but by carefully selecting the initial starting point, the network will converge to a suitable minimum error. Even with the drawback, the training can be tailored to equally balance the two error gradients or weight one more heavily than the other with the result of much better classification versus the standard approach.

2.9 Background Summary

While microcalcification detection and classification are required for a complete CADx system, the focus of this research is on mass detection and classification. Referencing Figure 6, Kegelmeyer's second set of results, and Yin's second set of results are the most promising results for segmentation performance comparison. They were both done on large, representative databases. Contrasting Kegelmeyer's binary decision tree algorithm, and Yin's standard segmentation approach, this thesis uses the model based vision architecture to focus the radiologist's attention on mammographic regions, and in addition, it provides the most probable hypothesis.

III. Database

The 72 medical cases used for this study were obtained using the Medical Diagnostic Imaging Support (MDIS) system at the Wright-Patterson Air Force Base (WPAFB) Hospital located at WPAFB, Ohio. In most cases, all four screening views were digitized for a total of 284 images. They were digitized according to a signed medical protocol with the WPAFB hospital (Appendix B). The protocol was necessary to acquire the images. The MDIS system is described in Section 3.1 and the malignant and benign biopsied mass images used in the database are listed in Appendix A.

3.1 MDIS System

The MDIS system includes a film digitization and archival storage system in use at the WPAFB hospital. The most important MDIS module for this thesis was the Lumiscan 200 automatic laser film digitizer made by Lumisys. The automatic film handler holds up to 70 films from sizes 8"x10" to 14"x17". The resolution of the system varies from 100μm to 420μm based on the film size. (The mammograms were digitized at 100 μm.) Each pixel is assigned a value equal to 1000 times the film's Optical Density (OD). The digitizer's density capability ranges from 0 to 3.5 OD at 0.001 OD resolution. This translates to a possible grayscale range of 0 to 3500 for twelve bits of grayscale resolution. The density resolution and precision are linear functions. The data are permanently stored on magneto-optical platters at the hospital, which are easily accessed by the MDIS system. Thus, the MDIS system provides an optimal system for mammogram digitization and storage.

3.2 Database Management

At the 100μm resolution, each 8" x 10" image varied from 1500 to 1800 columns by 2400 to 2500 rows with 12 bit grayscale. To make the database uniform and manageable, and to protect the patients' privacy, the tissue areas were hand-cropped from the patient label portion of the film resulting in a 1024x2048 array. In most cases, no tissue shown on the X-ray film was lost, and if tissue had to be cropped out, it was taken from the chest wall side of the image. One hundred pixels corresponds to a centimeter on the X-ray film;

so, the images are 20.48cm by 10.24cm. The image files were approximately 4 MB each, defining a total of storage capacity requirement of 1.14 GB. Thus, only relevant images were retained on the hard drive while the non-biopsied images were stored on tape backup. For this thesis' purposes, the database was split into five groups of images:

(1) Malignant Mass Training Set	10 cases (18 images)
(2) Malignant Mass Testing Set	9 cases (18 images)
(3) Malignant Mass Evaluation Set	6 cases (12 images)
(4) Benign Mass Set	28 cases (53 images)
(5) Non-Mass Set	19 cases (183 images)

The Malignant Mass Training, Testing, and Evaluation Sets each contained a biopsy-proven malignant mass in each image. The Training Set was used to define and develop the algorithm, and the Testing Set was then used to evaluate and modify the algorithm. Finally, the Evaluation Set was used to project how well the algorithm generalized for all clinical tests. The Evaluation Set's role in this effort was very important in determining how well the algorithm worked on images the algorithm had not previously seen. Using the Evaluation Set is the only way to determine the extensibility of these techniques to the clinical environment; since given enough time, any algorithm can be fine-tuned to do well on the training data. The Benign Mass Set consisted of biopsy-referred masses whose diagnosis was benign. The Benign Set would test the algorithm's ability to reduce the number of unnecessary biopsies. The size distribution of masses for the four datasets is shown in Figure 8. The images in the Non-Mass Set (5) either contained benign microcalcifications, malignant microcalcifications, or were one of the other mammographic views of a breast that had not been biopsied.

3.3 Case Selection

Records were selected from the WPAFBH pathology record book by a trained radiologist, and all cases selected were biopsied with a pathological diagnosis of the tissue. The original radiologist's diagnosis and the pathology of the cases were recorded with a detailed

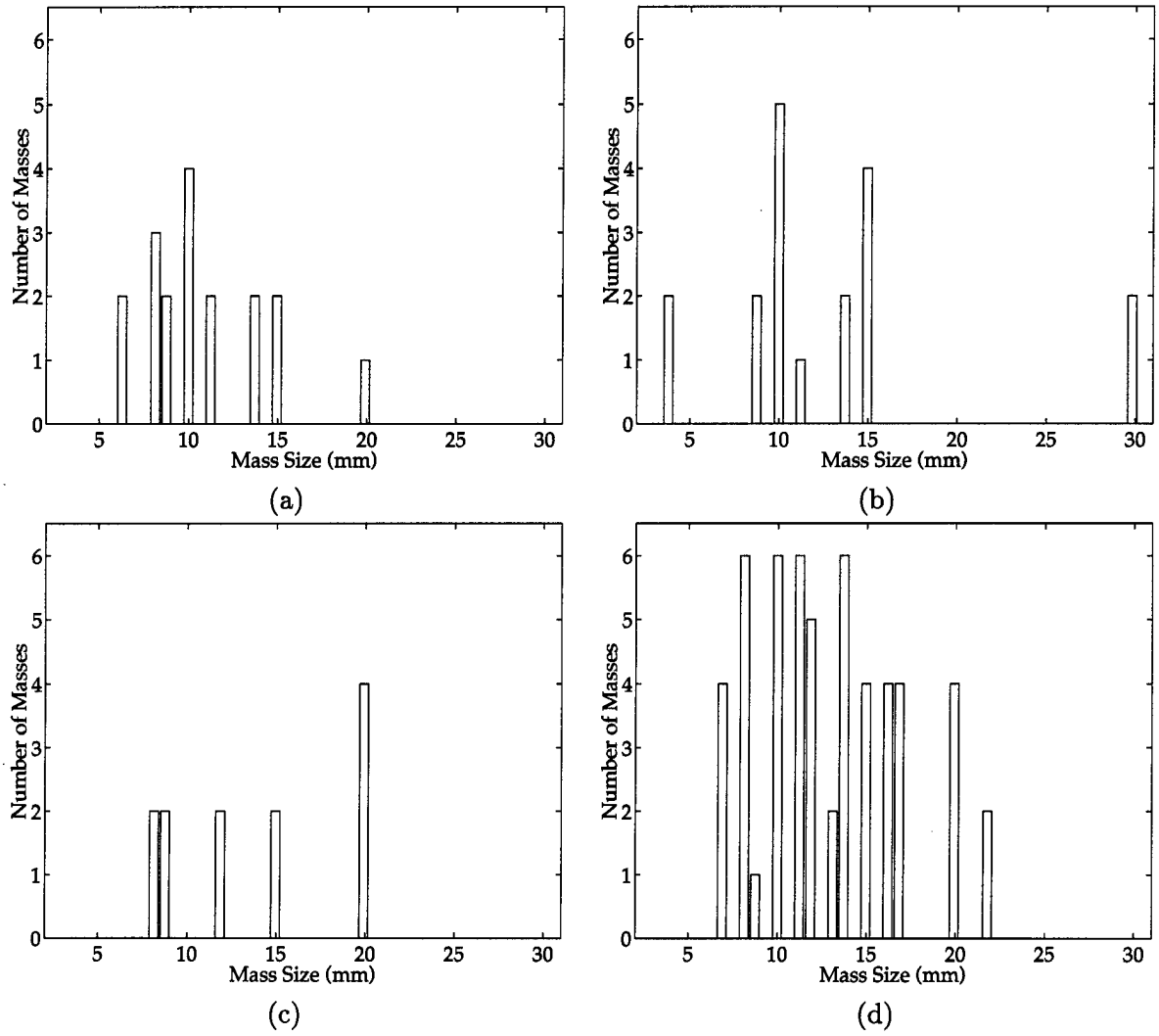


Figure 8. Mass Sizes for:
(a) Malignant Mass Training Set (b) Malignant Mass Testing Set
(c) Malignant Mass Evaluation Set (d) Benign Mass Set

description of the location and characteristics of the tissue biopsied. The diameter of the mass or the number of microcalcifications were recorded. All four mammographic views (right and left craniocaudal and right and left mediolateral oblique) from the screening session when the initial diagnosis was made were digitized.

The composition and number of cases present in a database can greatly effect the development of the algorithm and the experimental outcomes (51). Nishikawa, et al., used a database of 90 images and compared the performance of his classifier on the 'easy' cases versus the 'hard' cases. Results varied from 100 percent to 26 percent true-positive scores for the hard versus easy cases. In fact, while just switching 10 hard cases into the easy case subset, the 100 percent correct was reduced to 74 percent correct (51). Thus, no pruning of the records from the WPAFBH record book was done for this research. As many cases as possible were selected for this study to ensure a good cross-section of typical masses and microcalcification cases.

For this thesis, the entire database could not be used to develop the algorithms. Since the database contained both malignant and benign mass and microcalcification examples with the opposing mammographic views, and since the database was being acquired at the same time the methodology was being developed, the algorithms were developed on a limited subset of the entire database. The majority of the development of the algorithm in the next chapter was done using only the Malignant Mass Training Set which comprised just 6 percent of the entire database.

IV. Methodology

This chapter describes the Model Based Vision (MBV) techniques developed and the algorithms used to implement the Focus of Attention module's segmentation, the Indexing Module's labeling criteria, the Prediction Module's feature selection, and the Matching Module's classification of mass lesions in the mammograms.

4.1 Focus of Attention (Segmentation)

Masses of interest all have the general characteristic that they are of a higher brightness than the surrounding tissue, but in many cases the difference of gray levels separating a mass from the surrounding tissue is not large. In addition, they are generally not the brightest region in the image; networks of glands and ducts often are brighter. The difference between these mass regions and the other high intensity regions is that the networks of glands and ducts are often interlaced to present a large area of higher intensity. Thus, a frequency filter that would find smaller, distinct mass-like structures was developed.

The FOA Module process is shown in the flow diagram (Figure 9). Each step is described in the following sections.

4.1.1 Difference of Gaussians (DoG) Filter. There are many filtering techniques that can pass certain frequency ranges, but the Difference of Gaussians (DoG) (52) and Laplacian of Gaussian (LoG) (17) bandpass operations have been linked to the way humans preprocess an image (18, 53, 54). Since human diagnosis has been the best technique available, there is a good basis for modeling this approach. The DoG bandpass filter was the one chosen for this research since it is energy normalized and it has a broader frequency bandpass than the LoG filter. This allows it to respond to a wider range of mass sizes, but it also passes through more false ROIs that need to be dealt with in the next step in the MBV process.

The DoG filter is constructed by subtracting two Gaussians of different standard deviation, σ, and then taking the Fourier Transform of the image. Filtering an image with this result is analogous to convolving the DoG with the image. The DoG convolutional kernel and the analogous filter used for this research are shown in Figure 10. Figure 10a

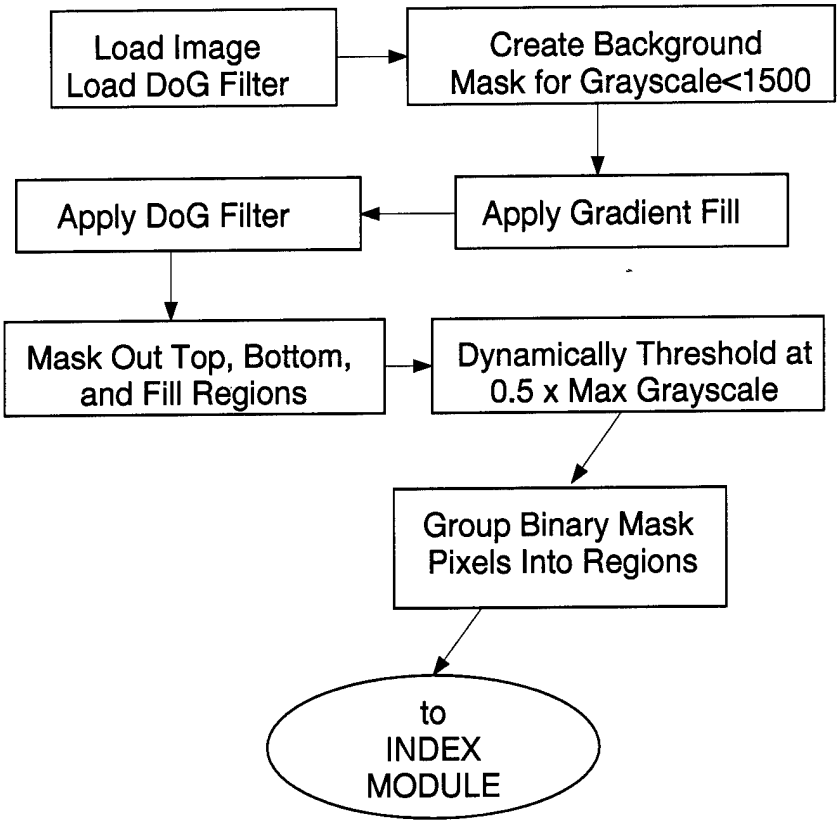


Figure 9. Focus of Attention Module Process

shows the spatial size of the convolution kernel, while Figure 10b shows the one-dimensional detail view of the kernel. The process as used in the research is similar to the matched filter implemented by Lai, et al., (1). The DoG kernel is a matched filter for structures of about 100 pixels (1cm) in diameter, similar to the tumors shown in Figures 1b and 2b. Notice how the DoG has the positive and negative aspects that Lai's template has in Figure 4, but the DoG is energy normalized and has a more gradual response.

The derivation for frequency response and the peak response follows. Using the Gaussian form shown in Equation 2 ensures that the DoG is energy normalized so that the energy above and below the axis is equal.

$$gauss(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \tag{2}$$

Subtracting two Gaussians of different σ yields:

$$DoG(x,y) = \frac{1}{2\pi\sigma_1^2} e^{\frac{-(x^2+y^2)}{2\sigma_1^2}} - \frac{1}{2\pi\sigma_2^2} e^{\frac{-(x^2+y^2)}{2\sigma_2^2}} \tag{3}$$

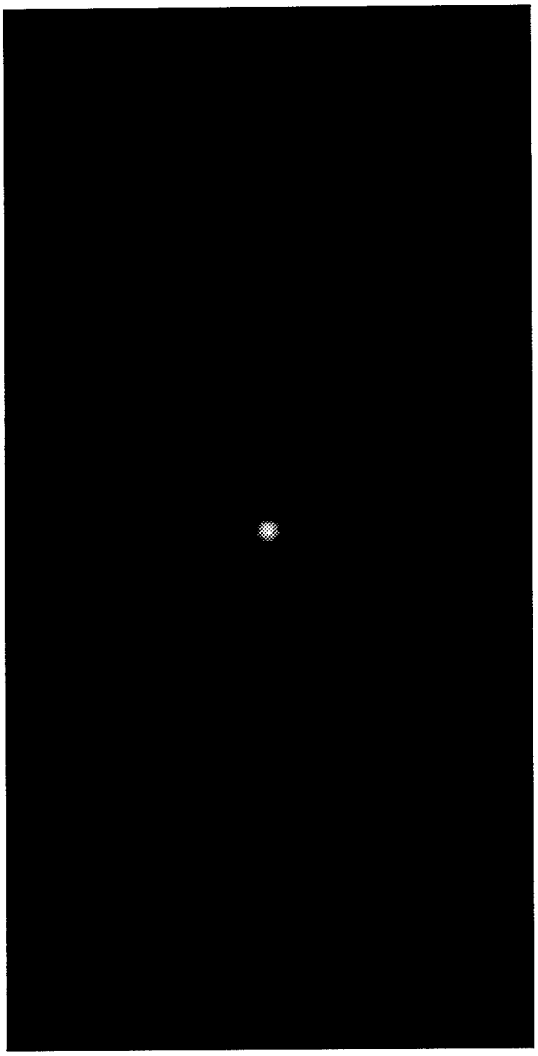
The 1D and 2D DoG plots for $\sigma_1 = 20$ and $\sigma_2 = 50$ are shown in Figure 10a and b. Using Gaskill's (55) definition of a Gaussian ($gauss(x) = e^{-\pi(\frac{x}{b})^2}$), and his Fourier transform relationship ($gaus(\frac{x}{b}) <==> |b|gaus(bf_x)$), the resulting DoG frequency filter becomes,

$$filter(f_x, f_y) = e^{-\sqrt{2\pi}\sigma_1^2(f_x^2+f_y^2)} - e^{-\sqrt{2\pi}\sigma_2^2(f_x^2+f_y^2)} \tag{4}$$

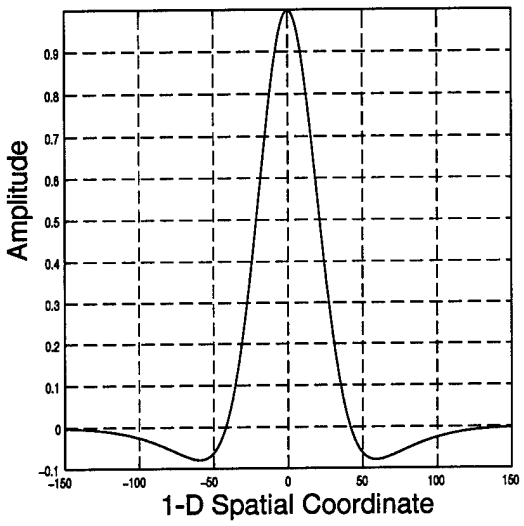
The filter is shown in Figure 10c. This view is the cutout from the entire frequency plot. The maximum spatial frequency is 1024 cycles or 512 cycles depending on the axis. Setting the derivative of Equation 4 equal to zero and solving for f_x or f_y yields,

$$f_x = \left(\frac{1}{\sqrt{2\pi}} \frac{\ln \frac{\sigma_2}{\sigma_1}}{(\sigma_2^2 - \sigma_1^2)}\right)^{\frac{1}{2}} \tag{5}$$

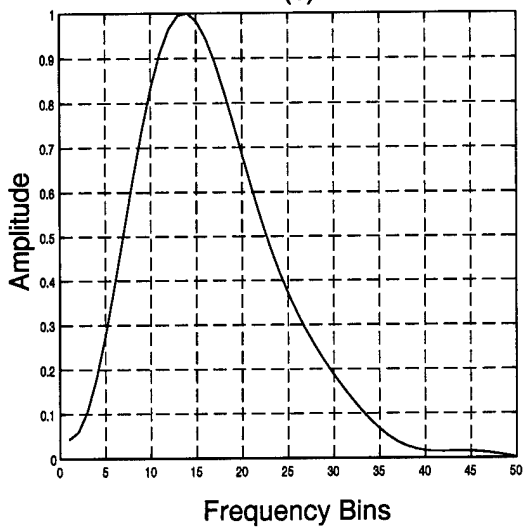
This equation has a similar form as found in reference (52). The difference lies in the definition of the Gaussian and its Fourier transform relationship. The peak response



(a)



(b)



(c)

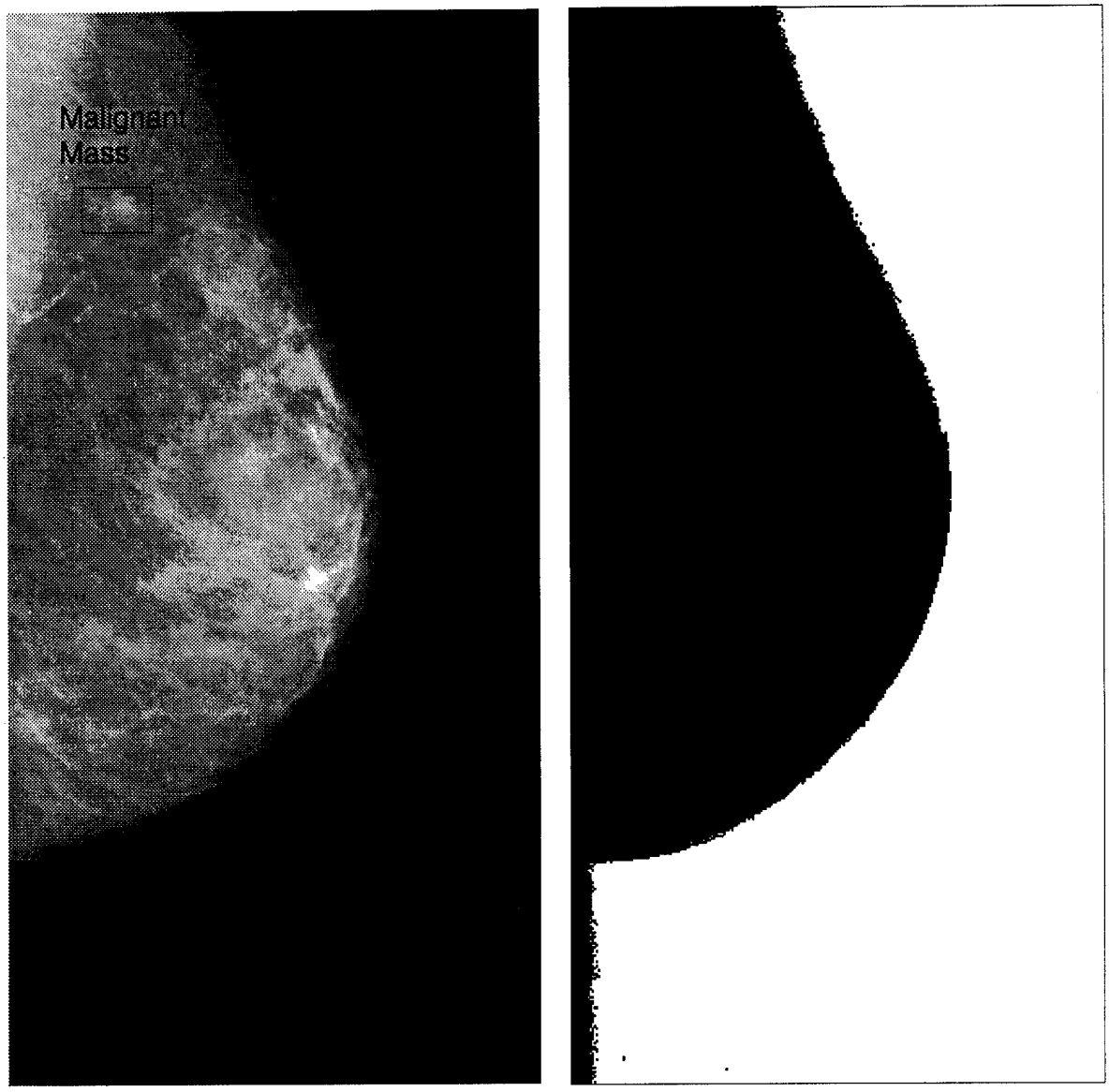
Figure 10. DoG Convolutional Kernel and Filter for $\sigma_1 = 20$ and $\sigma_2 = 50$
(a) Full Size 2D Kernel (b) 1D Detail View of Kernel
(c) 1D Detail View of DoG Filter

of the filter, f_x, for $\sigma_1 = 20$ and $\sigma_2 = 50$ from Equation 5 is 0.013 cycles per image which when multiplied by the maximum frequency bin represented in the matrix, 1024, yields 13.5 bin. Thus, the peak frequency response is between the 13th and 14th frequency bins along the row frequency axis. Since the matrix is not square (2048x1024) the peak response for the column frequency axis is 6.75 frequency bins from DC. For this research, the filter was constructed in a 512x512 array and then padded with zeros to the correct dimensions before Fourier transformation. The magnitude of the result was used.

4.1.2 Image Preparation for DoG Implementation. Since the DoG filter convolves the Kernel with the entire image, some preprocessing steps were taken to eliminate some unwanted artifacts. The three main sources of the artifacts were the transitions from the breast tissue to the dark background, the X-ray film markers used for film identification, and the edge effects due to the circular convolutional nature of the Fourier Transform. An additional unwanted outcome was the sensitivity to the grayscale changes between the chest wall and breast tissue, but that outcome was not as critical as the others.

4.1.2.1 Preprocessing. The first preprocessing step to remove the artifacts was a thresholding step used to identify breast tissue vs background X-ray pixels. Due to the calibration of the X-ray machines and the X-ray films used in this database, the brightest gray level that could, in general, be attributed to background and not breast tissue was 1500. Raising the threshold higher caused breast tissue that was in the interior of the breast to be included in the mask. Thus a mask was created of all pixels with a gray level < 1500 (see Figure 11b).

The next step was to eliminate or at least reduce the edge transition from the breast tissue to the background by filling in the background pixels with higher grayscale values comparable to the breast tissue's grayscale. Figure 13a-d shows the wide variety of horizontal transitions that occur in mammograms. As shown in the detail views of Figure 13a-d, plots e-h show that in general, the transitions occurred over 75 to 100 columns from the edge detected by the mask. In other words, the relevant breast tissue is approximately 75 to 100 columns to the left of the mask edge for any particular row (Figure 13e-h). This



(a) (b)

Figure 11. (a) Figure 1a Reproduced with Malignant Mass Identified
(b) Mask of Grayscale Pixel Values Less Than 1500

transition region includes the skin tissue of the mammogram which is not needed for the diagnosis.

To simplify the fill process, the replacement of background pixels was done row by row, left to right, with all mammograms oriented with the chest wall on the left of the image. So as the algorithm scanned a row of the mask, it would detect the first 'on' pixel in the mask and identify the pixel position 55 columns to the left as the 'edge pixel'. As discussed below, it would then fill in all pixels from the edge pixel all the way to the right edge of the image. This would have two major effects: It would reduce the grayscale transition without eliminating relevant edge features and structures, and it would eliminate the radiologist film markers on the image.

Two types of grayscale fills were attempted. The first technique filled each row with a constant gray level based on the edge pixel's value in the row. This reduced the breast to background edge effect, but caused a mismatch between the left and right edges of the mammogram. This artificial edge caused artifacts in the filtered image due to the periodic properties of the Fourier Transform. (Note: This artifact was even stronger prior to the fill.) The technique implemented to reduce this artifact was a gradient method. The difference in gray levels between the edge pixel and the first pixel in the row was found. The algorithm then implemented a linear gradient fill that caused the gray levels to match both end pixel grayscale values and decrease (or increase) linearly between them. Figure 12 shows the full filled-in image with specific rows identified. Figure 13 shows the plots of the rows shown in Figure 12 before application of the fill algorithm, and Figure 14 shows the row plots after the application of the fill algorithm. The grayscale transitions were dramatically reduced, while the relevant breast tissue was retained.

The DoG results for the unprocessed and preprocessed images are shown in Figure 15. The tumor region is located in the white box in the figure. (Reference Region 1 in Figure 1a to see the actual tumor.) Leaving the transition unchanged (no fill algorithm) caused tumors even relatively far from the edge to be obscured by the breast tissue to background transition and the image edge effect Figure 15a. Figure 15b shows the results of the DoG filter applied to the preprocessed, gradient-filled image. Although the fill algorithm generated some artifacts in the fill region, the edge transition artifact was greatly reduced

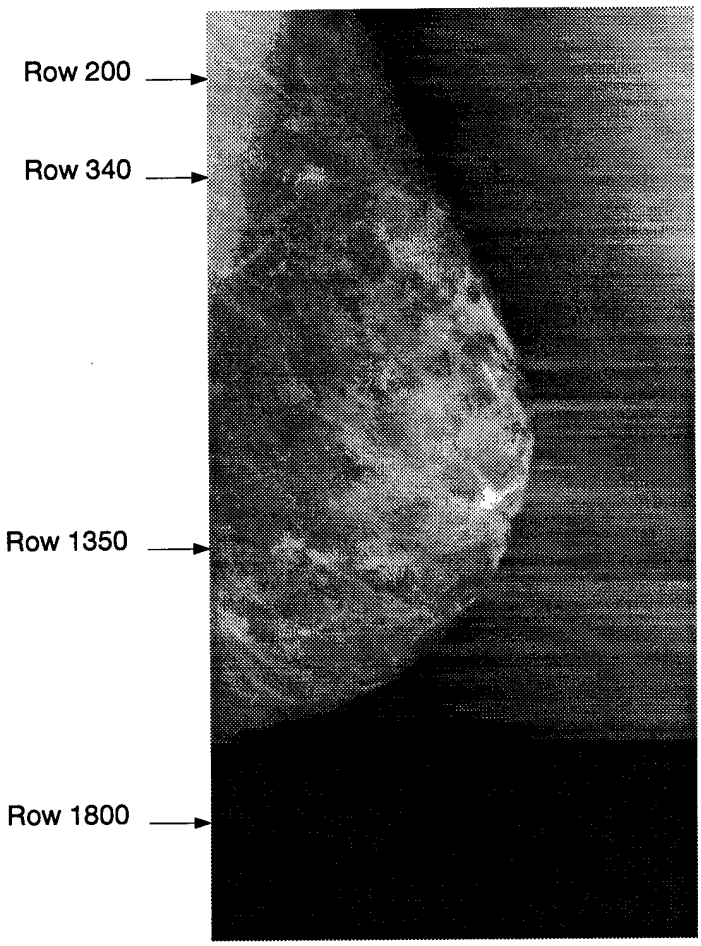


Figure 12. (a) Image From Figure 11a with Gradient Fill.

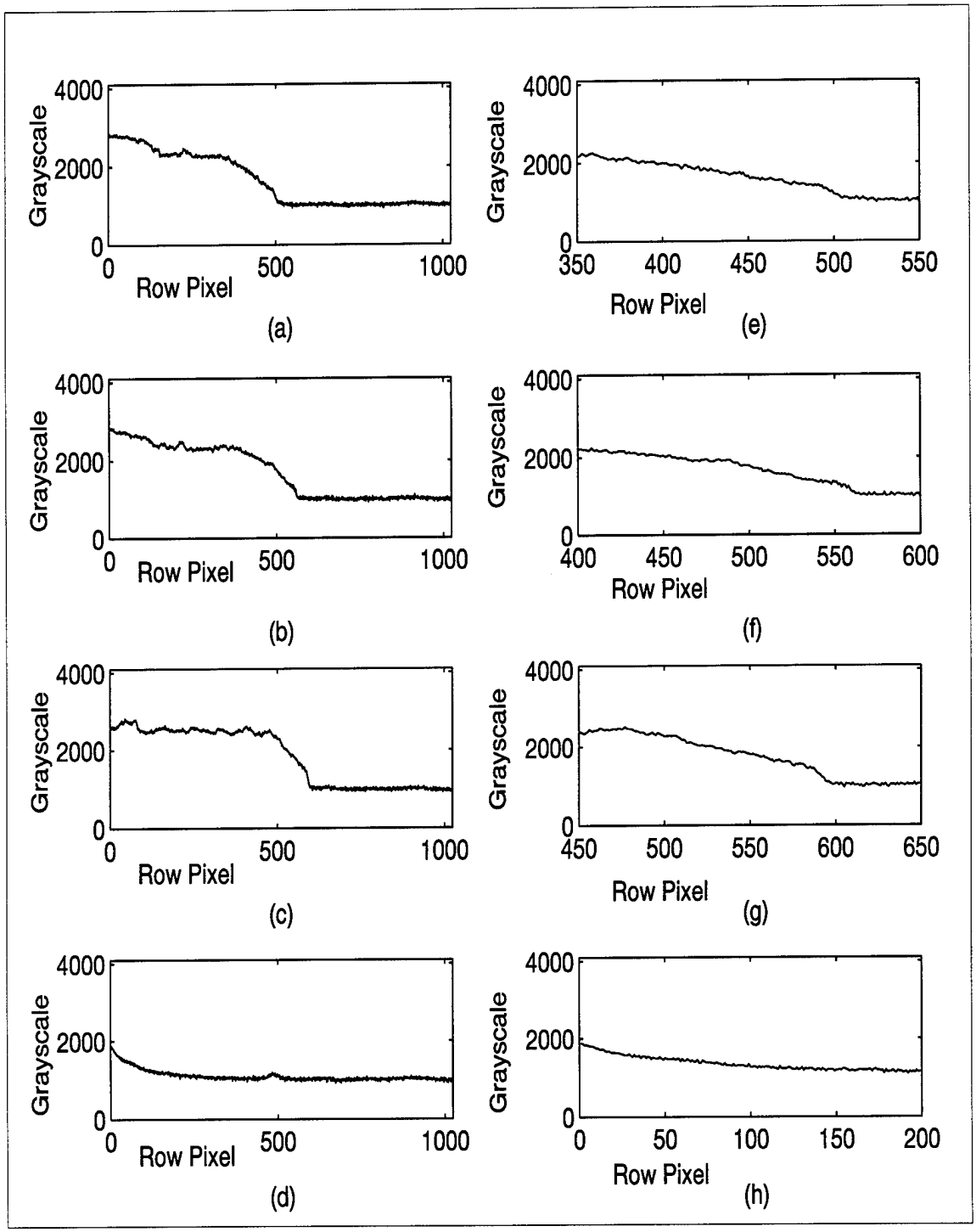


Figure 13. Row Grayscales in the Unfilled Image at the Points Shown in Figure 12
Full Row Plots (a) row 200 (b) row 340 (c) row 1350 (d) row 1800
Detail Views of the Transition Regions (e) - (h).

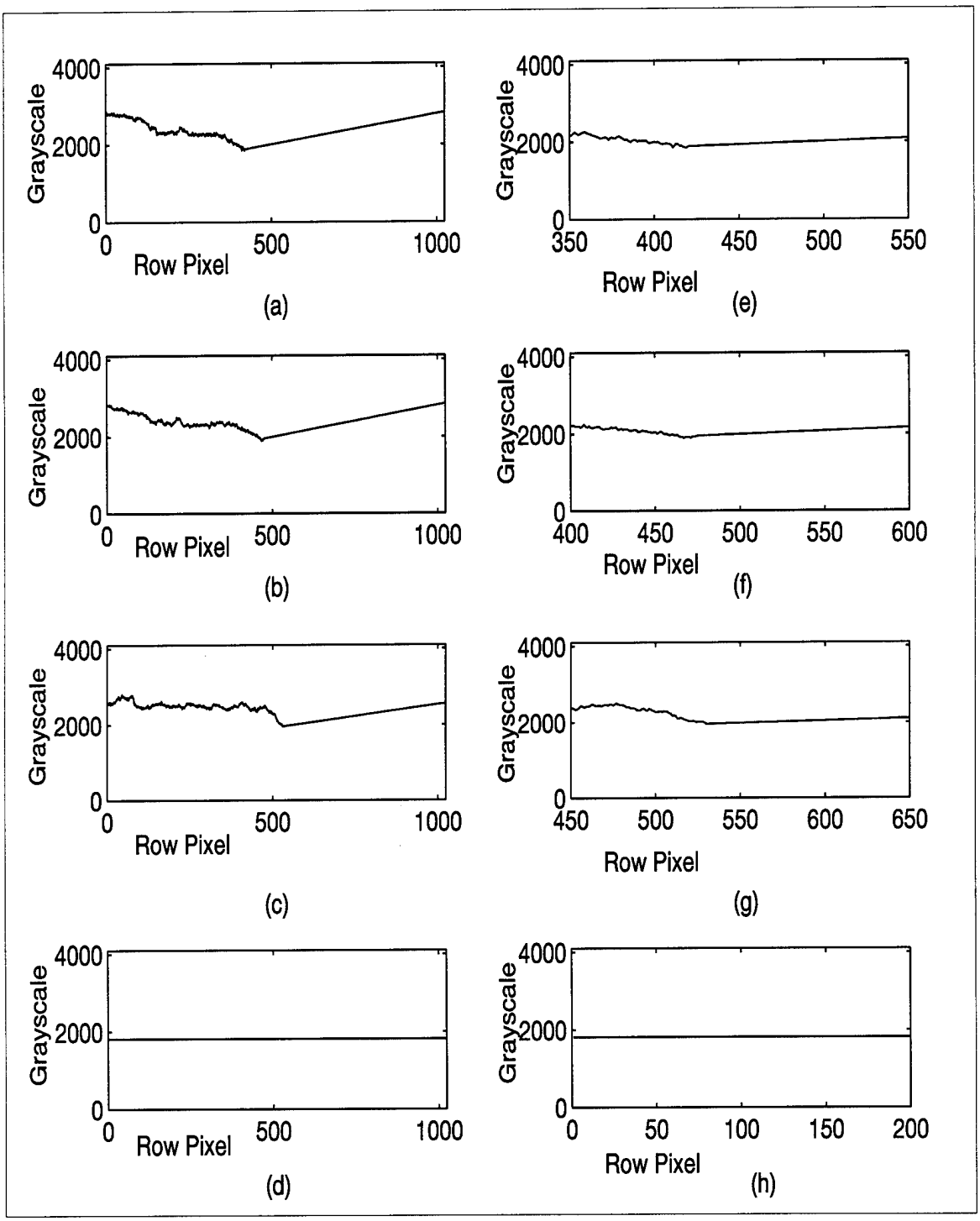
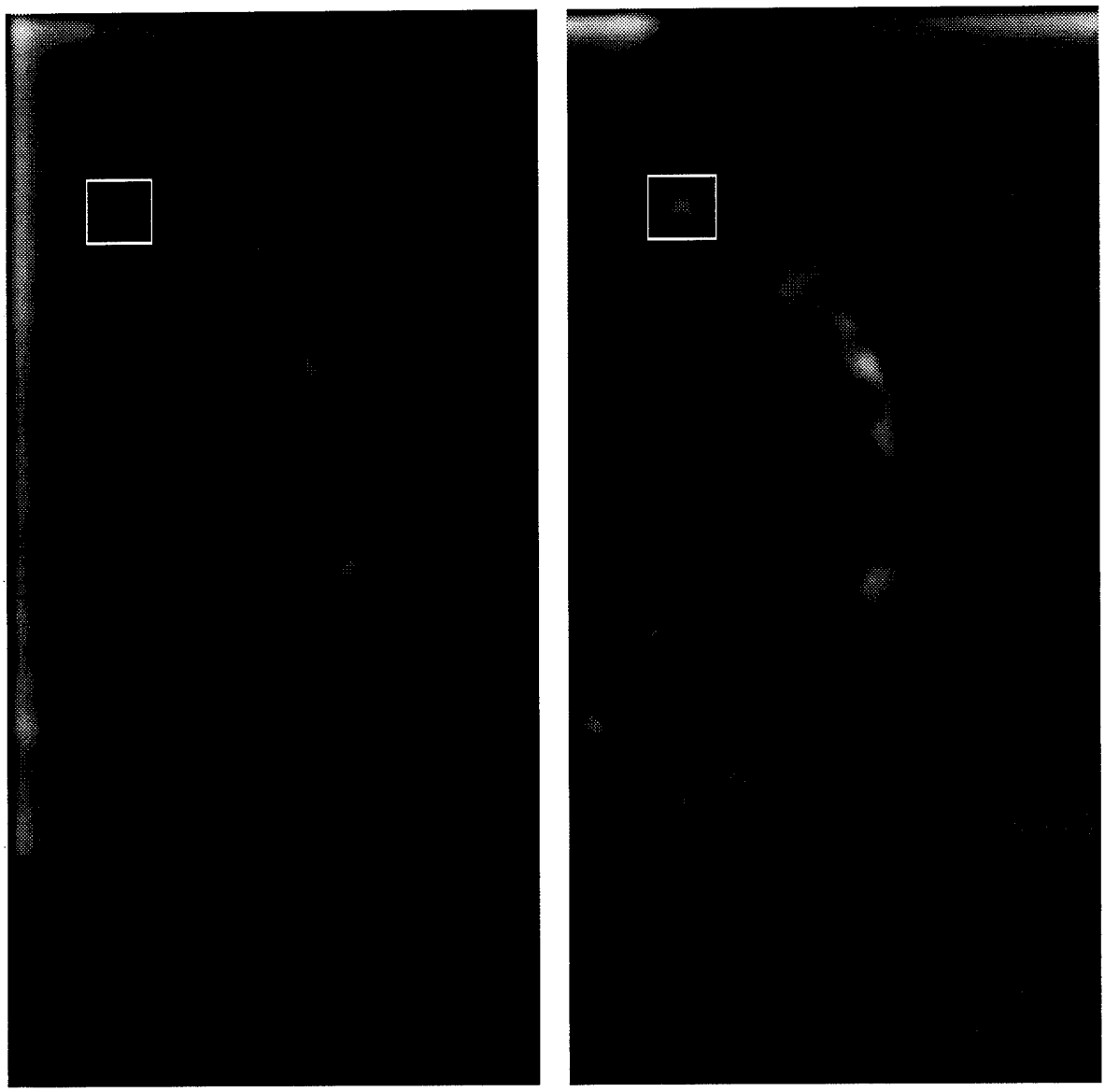


Figure 14. Row Grayscales at the Points Shown in Figure 12 After Application of the Fill Algorithm.
Full Row Plots (a) row 200 (b) row 340 (c) row 1350 (d) row 1800
Detail Views of the Transition Regions (e) - (h).

and the tumor is much more visible. In addition, the left and right edge effect was greatly diminished, so now the chest wall to breast tissue transition is more visible too. With some minimal post-processing, the remainder of the artifacts could be removed.

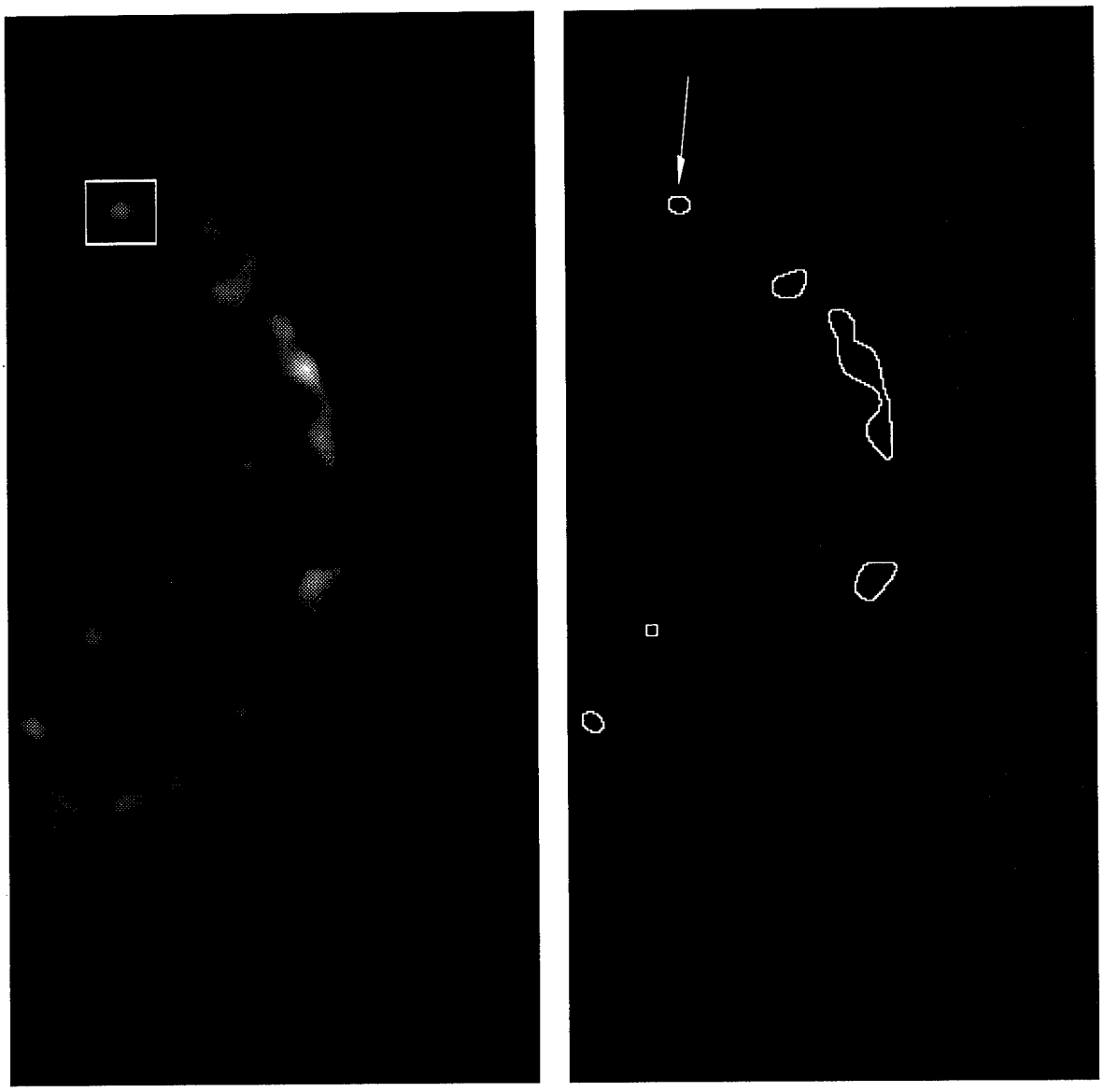
4.1.2.2 Postprocessing. Once the fill algorithm was implemented, and the DoG filter was passed over the image, the resulting image shown in Figure 15b still contained some artifacts of the process. Two candidate items for postprocessing are of note in this image. First, the DoG filter was sensitive to the vertical mismatch in gray levels from the top to the bottom of the image, and second, there were pseudo-tumors in the filled-in area of the image. In addition to some normal tissue, the real malignant tumor was highlighted in the filtered image (see the white box in the figure), but the gray level mismatch from the top to the bottom of the image dominates and masks the items of interest. In some images there was no mismatch, but in others, the real malignant tumor was completely obscured due to the high intensity of the top to bottom edge effect. No structures of interest were located within the top or bottom 100 rows (out of 2048 total rows), so these rows were masked out. In addition the threshold mask used in the preprocessing step (Figure 11b) was used to eliminate any pseudo-masses located in the fill region. These pseudo-masses were due to the vertical grayscale mismatches present from the horizontal gradient fill algorithm. The resulting, postprocessed image is shown in Figure 16a.

With the result in Figure 16a, the Focus of Attention Module does one more step before passing the filtered image to the Indexing Module. A dynamic threshold of the image was implemented to select the pixels with grayscale values greater than 50 percent of the maximum grayscale value in the postprocessed matrix. The binary image contained groupings of pixels that corresponded to the ROIs the FOA Module was tasked to identify. These ROIs in the binary image were passed to the Indexing Module for the next stage of analysis. Most of the binary regions selected corresponded to reasonable regions that one would like the FOA Module to select.



(a) (b)

Figure 15. (a) DoG Filter Results on the Unpreprocessed Image Shown in Figure 11a
(b) DoG Filter Results on the Preprocessed Image Shown in Figure 11a.



(a) (b)

Figure 16. (a) DoG Filter Results on the Post-processed Image Shown in Figure 11a
(b) The Binary Regions Selected. The Arrow Shows the Malignant Region.

4.2 Indexing

Using the binary image from the FOA Module and the original grayscale image, the Indexing Module did two tasks: It labeled each grouping of pixels in the binary image into a discreet set of categories, and it subjected each grouping to a set of tests to reduce the number of non-malignant ROIs. A flow diagram is shown in Figure 17. Each process is described in the following sections.

4.2.1 Labeling. The binary image passed to the Indexing Module from the FOA Module generally contained a number of groupings of pixels (ROIs). Each of these ROIs in the binary image were categorized into five indexes: Edge ROI, Small ROI, Medium ROI, Large ROI, and Extra Large ROI. The Edge ROI label was created to eliminate any small groupings of pixels close to the chest wall side of the mammogram. Any pixel groupings that extended less than 60 pixels into the image were labeled as Edge ROIs. Since most detectable masses are greater than 6.0mm in diameter and occur in the interior of the breast, all Edge ROIs were considered to be artifacts of the DoG process. All other ROIs were passed through a minimum bounding box area to perimeter ratio test to determine their label. The thresholds between the categories were box to perimeter ratios of 3.0, 36, and 50. The value of the ratio corresponds to the size and complexity of the ROI. Small ROIs (ratios < 3.0) and Extra Large ROIs (ratios > 50) were eliminated since no malignant masses fell in these ranges. Considering the two remaining index labels, the Large ROI index ($36 <$ ratios < 50) was designed to detect masses greater than 2 cm in diameter, while the Medium ROI was designed to detect masses from 0.5 cm to 2.0 cm in diameter. Most masses fell into the Medium ROI indexing label. Since, there were only two Large ROI examples, and since large masses are usually easily diagnosed, nothing further was done with the Large ROIs besides noting if the malignant mass was detected and noting the number of non-malignant ROIs detected. This 'ratio' test used to eliminate undesired ROIs and label the medium and large masses was the first of a number of tests used to reduce the number false ROIs. At this point the Medium ROIs were hypothesized to be 0.5 cm to 2.0 cm diameter masses (malignant or benign), or background tissue.

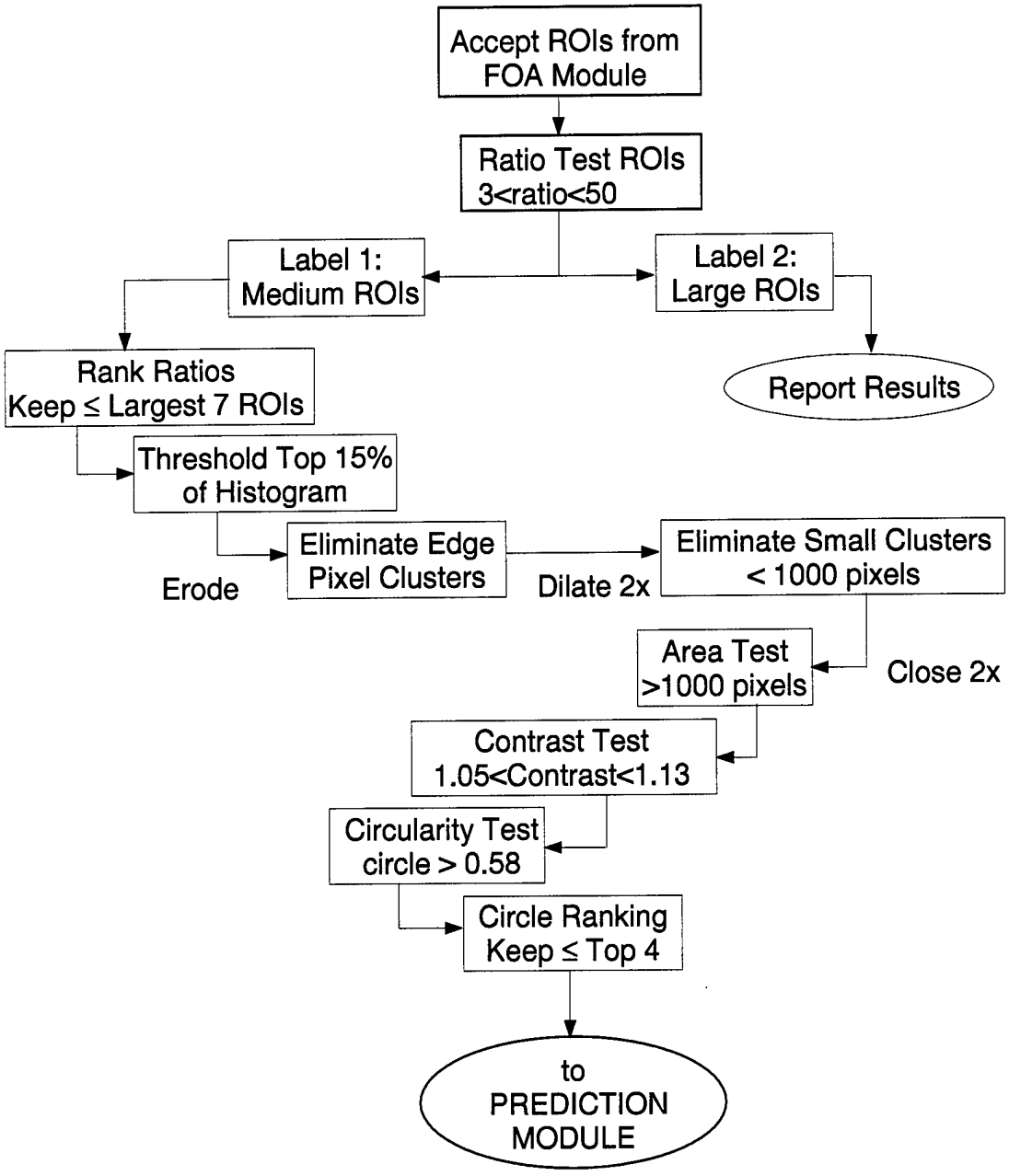


Figure 17. Index Module Flow Diagram

4.2.2 Testing. This section can best be described as a loop back to the FOA Module, but this time based on the indexed ROIs. Although features are selected, it is inherently not a feature selection process, but rather a reduction of false ROIs/segmentation process. The intent is to eliminate all background tissue ROIs, and be left with only malignant and benign masses. The following tests: ranked-ratio, area, contrast, and circularity were all implemented only on the Medium ROIs that passed the ratio test.

For the ranked-ratio test the ratios used in the labeling step were ranked in descending order. A maximum of seven ROIs could be retained. This criteria passed all the malignant ROIs while eliminating many of the false ROIs. The images had between 1 and 22 ROIs, but the malignant ROI, if found, always had a ratio that put it in the top seven ROIs. For the example shown in Figure 16b, this criteria had no effect.

The next three tests conducted are a variation of the false alarm reduction process used by Yin (42). Yin segmented a suspicious region from the image and then found the peak pixel value in that region. A region was morphologically grown from that point until the mask only encompassed pixels that were within 97 percent of the peak pixel's value. He then did a minimum/maximum area test, a circularity test, and a normalized contrast test. The morphological process and the contrast test used in this thesis are unique, but the area and circularity tests were very similar in idea to Yin's work. The tests are described below.

The tests were conducted on ROIs extracted from the original image. The 140x140 ROIs were extracted based upon the centers of the ROIs passed through the ranked-ratio test (Figure 18a). The tests required a binary mask representing the shape of the hypothesized mass in the ROI. To create the mask, the top 15 percent (3000) of the pixels in the image's histogram were used as a baseline (Figure 18b). All of the following morphological operations were done using the standard Matlab 3x3 square kernels. A morphological erode was implemented to eliminate any small pixel clusters. Since true masses should be centered in the ROI, any pixel groupings close to the image edge were set to zero (Figure 18c). The masses also typically had a large variation in grayscale values in their interiors, so two morphological dilations and a morphological close were used to

connect and fill close but disjoint regions. The 'on' portions of this mask (Figure 18d) provided the basis for the following tests and the feature extraction algorithm.

The three tests (area, contrast, and circularity) were each based on malignant mass vs background tissue characteristics. The area test first set any pixel groupings in the mask containing <1000 pixels to zero, and it then rejected any masks that contained <1000 pixels. Since the mask started with 3000 pixels, and perfectly circular 5mm to 10mm masses should ideally contain 2000 to 8000 pixels, this was a very conservative bound. Next, the contrast test assumed that the masses are of a higher mean grayscale value than the surrounding tissue, or they would not be discernible. Thus, it divided the mean grayscale value of the original image within the mask by the mean of the original image within the entire ROI. Only ROIs with contrast values from 1.05 to 1.13 were retained.

Since masses tended to be more circular than other structures in the breast, the final test was the circularity test. A circle of area equal to the area of the mask was created and positioned at the center of the 'on' pixels in the mask (Figure 19).

The area of overlap of the circle and the mask was divided by the area of the circle to produce a number ranging from zero to one. ROIs with two surviving regions in them or long narrow regions failed this test. The threshold for circularity was set at 0.58. One final aspect of circularity was used for the last false alarm reduction step. The remaining ROIs (up to seven) were ranked in descending order once again, and only the top four were retained. For those images that still contained seven mass-like ROIs, it was found that the malignant ROI was usually the first or second ranking ROI, and it was never past the fourth in line in the training or test sets. Although this technique (picking the top four ROIs) was not influenced by Wei's hand segmentation of four ROIs per image, it makes for a good comparison of the results.

Initially, each test was conducted on all ROIs, but the tests were much more effective when they were done in succession on only the ROIs that survived the previous tests. Thus the final binary mask for the example is shown in Figure 20. The best four ROIs were retained and sent on to the Matching Module. In almost all of the mammograms, the best four ROIs corresponded to regions that a radiologist should and most likely does

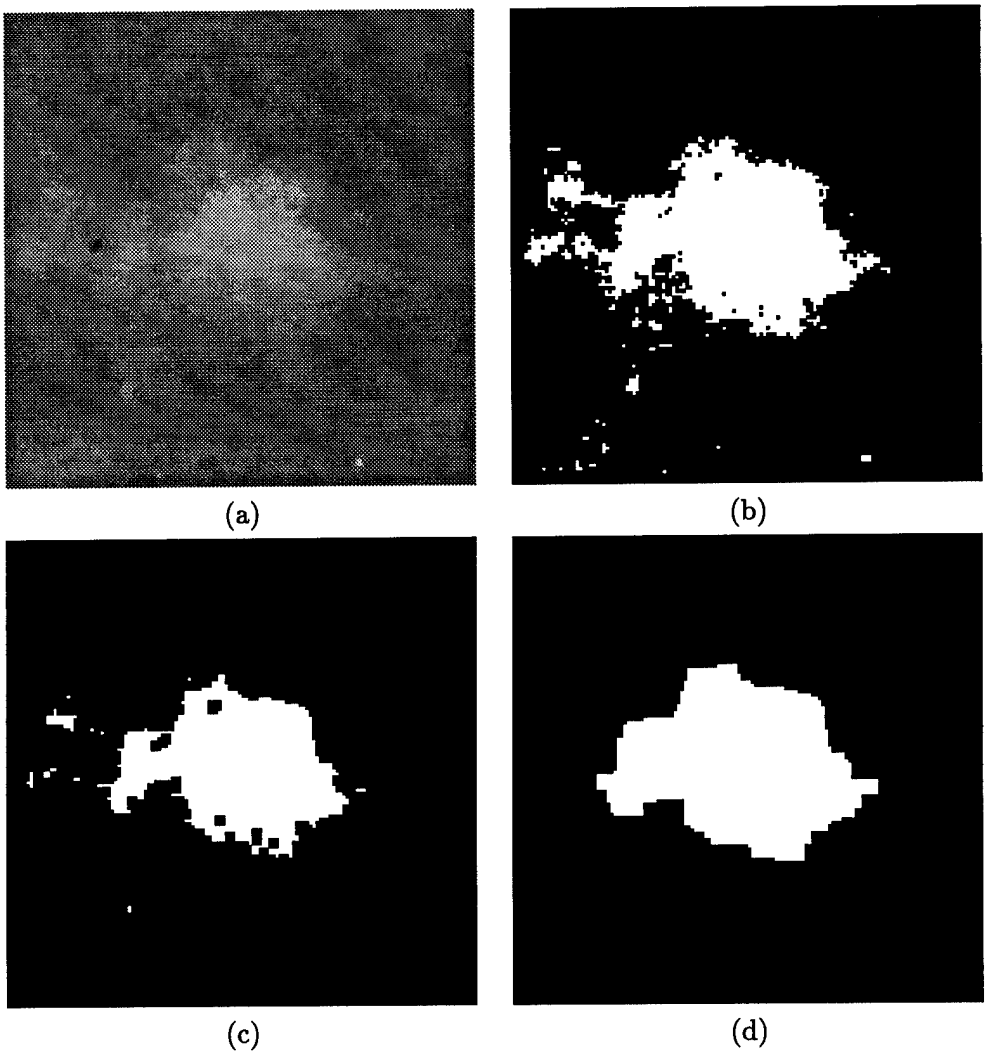


Figure 18. The ROI from Figure 1b and it's Morphological Masks
(a) Malignant Mass ROI (b) Top 15 Percent of the Histogram
(c) Erosion (d) Final Mask

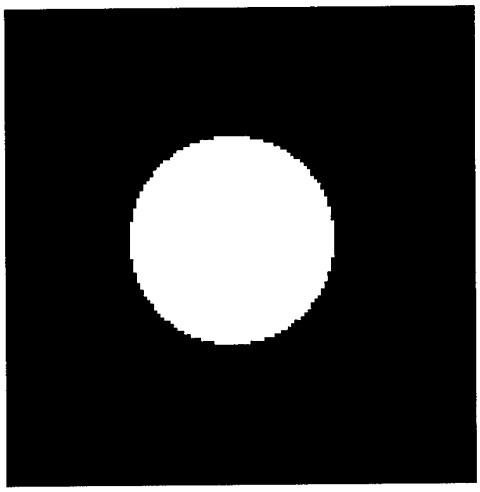


Figure 19. (a) Circular Mask for the ROI Mask Shown in Figure 18d.

scrutinize during their diagnosis. Thus, selecting up to four of the most mass-like regions in a mammogram, without even the Matching Module's computer diagnosis is a very useful tool for the radiologist.

Once the ROIs, if any, successfully passed through the Index Module's procedure, they were sent on to the Prediction/Feature Selection Module.

4.3 Prediction / Feature Selection

In the literature, the features that seemed most likely to be separable to the radiologist, and thus, to the CADx system, were texture, shape, and mass border transition features (20). Kegelmeyer (37) seemed to have the best results using the Laws texture features; so those were the first set of features chosen in the Prediction Module. In addition, the four shape and border features that were already calculated in the indexing step were used.

The Laws features are, in general, used for image segmentation, but Kegelmeyer (35, 36, 37) and Wei (38) have both applied them to classification of masses. Miller and Astley also used the Laws features to classify other types of breast tissue, such as, fatty and glandular tissue (56). The Laws features are derived from a set of five convolution kernels

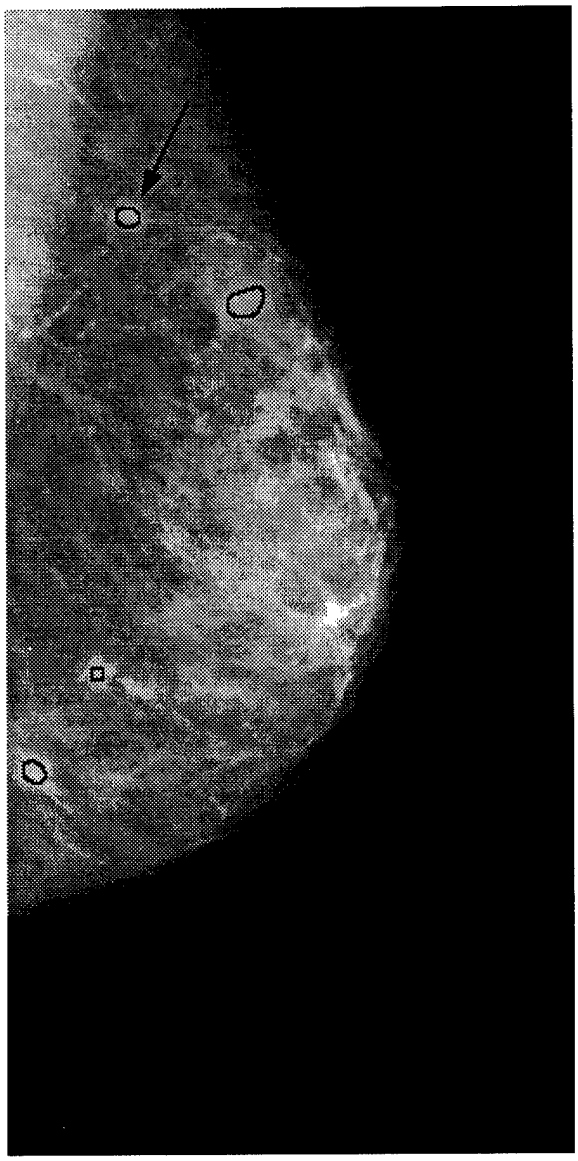


Figure 20. The final output of the Index Module. It shows the original example image with the suspicious regions outlined. The true malignant mass is identified by the arrow.

Kernel Name	Kernel Label	Kernel Values
Average	l	(1 4 6 4 1)
Spot	s	(-1 0 2 0 -1)
Edge	e	(1 -4 6 -4 1)
Ripple	r	(-1 -2 0 2 1)
Wave	w	(-1 2 0 -2 1)

Figure 21. Laws one-dimensional kernels

1	4	6	4	1
4	16	24	16	4
6	24	36	24	6
4	16	24	16	4
1	4	6	4	1

Figure 22. The Laws L5L5 two-dimensional kernel

that are applied to an image. Each of the kernels shown below responds to different local texture behavior.

For image texture analysis, these five kernels are convolved together to form twenty-five 5x5 kernels. These kernels are designated by l5l5, l5s5, ..., s5l5, s5s5, ..., and w5w5. Notice that the switched position of the labels corresponds to the transpose of the original kernel ($l5s5 = s5l5^T$). All twenty-five kernels were convolved with the ROIs. A number of techniques were used to determine how to select one number to represent the resultant energy in the convolved images. Kegelmeyer (35, 36, 37) used a 15x15 averaging filter to force a consensus between neighboring pixels, but he was classifying individual pixels in the raw image, not ROIs. The best feature set for classifying an entire ROI would rather be a number that contains the correlation of the ROI with the particular Laws kernel. An average pixel value in the convolved image should capture the desired correlation. However, since in most cases the true masses encompassed only a subset of the ROI, only the resultant pixel values within the Index Module's final mask were used. (The mask was dilated twice more to ensure most of the mass and the transition region was included.) This was done to keep the 'normal' tissue from diluting the mass texture results. Of these twenty-five Laws features and the ratio, mass area, contrast, and circularity Index Module features the best features needed to be found.

Since up to four regions could be retained per image, there were a possible thirty-six malignant ROIs and possibly three to four times that number of benign ROIs in the Malignant Mass Training and Test Sets. So, recalling Foley's Rule (number of class samples $>$ 3 times number of features) up to a maximum of about ten features could be used to develop the model. As discussed in Chapter II, there are many ways to select the 'best' features. Although each method may yield different 'best' features, they should be in some agreement. Since the Matching Module uses a neural network, the best method for feature selection would be a neural network technique.

The feature saliency technique chosen was the derivative-based technique using the imbalanced training set neural network architecture with sigmoidal activations. However, recalling the maximum number of features rule discussed in Section 2.7, even for feature saliency tests, there is a risk in using too many features. The results may not be useful. Recall also that the goal was to obtain about ten features, so more than ten features were needed so a selection could be made. The number chosen from which to make this down selection was twenty features. Since the f-ratio rule is the simplest of the rules, it was used to eliminate the nine least relevant features. Then the more sophisticated derivative-based method could be used on the remaining twenty features. Although, using twenty features for thirty-six malignant training samples still violates Foley's rule, the saliency results should retain most of their validity. To maintain a reasonable balance between Uncle Bernie's rule and a valid network architecture, no hidden nodes were used with one output node per Anand's specifications. Although this corresponded to a linear discriminant architecture, adding additional nodes could cause the data to be memorized, and thus invalidate the saliency results. This resulted in twenty-one network connections for a ratio of about five samples per connection. Thus, there is the risk that the network will memorize the data in the 10,000 epoch training run, but the features should still retain their ranking.

The final ten features then define the differences between the malignant masses and all other benign tissue. In essence these feature parameters form the model which is used for the Matching Module.

4.4 Matching / Classification

The Matching Module used the features selected from the Prediction Module and determined the best neural network architecture and coefficients for classification of all of the evaluation data. This evaluation data included all the ROIs from the 251 images not in the Malignant Mass Train or Test sets. The only malignant masses it contained were from the Malignant Mass Eval set. Thus, it accurately represents the true classification rate for the benign ROIs, but only had at best twelve malignant ROIs to determine the classification rate for the Malignant ROIs.

4.5 Summary of Methodology

This chapter has described the Model-Based Vision (MBV) process for detecting and classifying suspicious regions in digitized mammograms. The MBV Modules identified the key tasks to perform and the requirements for interfacing each of the pattern recognition concepts into one complex algorithm. The FOA Module identified the suspicious regions. The Index Module separated the regions into two labeled categories and reduced the number of false ROIs for the Medium Mass Category. The Prediction Module defined a number of features and then selected the best features from which it developed it's models for malignant and non-malignant tissue. The Matching Module used the best features and designed the best neural network architecture and parameters to correctly classify the malignant and benign regions. Each of these Modules involved a complex series of tasks and tests which the MBV process neatly structured into the appropriately grouped processes for a highly functional pattern recognition architecture.

V. Analysis

Receiver Operating Characteristic (ROC) Curves are used to report the results. These plots give a better picture of the performance of the algorithm because they include true-malignant, false-malignant, true-benign, and false-benign information. Their development, meaning, and use are described in more detail in Giger's article (6) and Metz's article (57).

The biopsied benign cases are included in this analysis separately to determine at what point in the process they are either ignored, or they are classified and presented to the radiologist as benign ROIs.

5.1 Focus of Attention

Recall, the Focus of Attention (FOA) module consisted of the steps shown in Figure 23.

The results for the FOA module are listed below, but the overall performance was 85 percent correct segmentation of malignant regions with 67 percent of the benign biopsied regions being retained and an average of 8.3 false ROIs per image being passed on to the next Module.

For analysis, the Malignant Mass Train and Test Set results are combined into the indexed classes. Figure 24a lists the results for the Medium ROIs and Figure 24b lists the results for the Large ROIs. Of the 36 malignant masses present, 31 of them were indexed as Medium ROIs with 241 Medium false ROIs. In addition, 2 of the medium malignant regions had two Medium ROIs associated with them. Two of the malignant masses were indexed as Large ROIs with a corresponding 14 Large false ROIs. Of the three ROIs that

1. Create a threshold mask for pixel values < 1500.
2. Gradient fill the background pixels.
3. Apply the DoG Filter.
4. Mask out the top, bottom, and fill regions.
5. Dynamically threshold at 0.5 times the maximum gray level.
6. Group the binary mask regions.

Figure 23. FOA Process

Data Set	Number of Images	Correct Regions Found	True Positive Fraction	Total False ROIs	False ROIs per Image
Malignant Mass Train and Test Set	36	31/34	0.91	241	6.7
Malignant Mass Eval Set	12	7/11	0.64	86	7.2
Benign Mass Set	53	36/54	0.67	414	7.8
Non-Mass Set	183	na	na	1517	8.3
Medium ROI Results	284	39/46	0.85	2258	8.0

(a)

Data Set	Number of Images	Correct Regions Found	True Positive Fraction	Total False ROIs	False ROIs per Image
Malignant Mass Train and Test Set	36	2/2	1.00	14	0.39
Malignant Mass Eval Set	12	1/1	1.00	1	0.08
Benign Mass Set	53	1/1	1.00	16	0.30
Non-Mass Set	183	na	na	78	0.43
Large ROI Results	284	3/3	1.00	109	0.38

(b)

Data Set	Number of Images	Malignant Regions Found	True Positive Fraction	Total False ROIs	False ROIs per Image
Combined Results	284	41/48	0.85	2367	8.3

(c)

Figure 24. Focus of Attention Module's Performance:
(a) Medium ROI Index Results (b) Large ROI Index Results
(c) Combined Results

were missed, two images (one case) contained a 4mm mass, and the other one was buried too deeply into a large dense region of the mammogram to be separated out in the CC view. However, it was detected in the MLO view. An image from each missed case is shown in Figure 26. Note that the algorithm was never designed to detect masses less than 7mm in diameter.

The two ROC curves in Figure 25a show the tradeoff for adjusting the ratios that defined what regions were sent to the Indexing Module or the tradeoff for the number of regions sent to the Indexing Module. For the ratio ROC curve, the lower threshold for the ratio was varied from zero to fifty. Only regions with ratios higher than the threshold were retained. The number of correct malignant ROIs retained vs the number of false ROIs retained were plotted for the ratio thresholds. The second ROC curve was done with the ratio thresholds set to retain all regions with ratios between three and fifty. Then with the regions ranked by ratio, successively fewer ROIs were retained. The number of correct ROIs retained within the top 'X' regions vs the number of false ROIs in the top 'X' regions define the second ROC curve. Although this test was done in the Indexing Module, the ratio was calculated in the FOA Module, and the ranked ratio test is really a loop back into the FOA Module. These curves apply to all thirty-six malignant images and both indexed classes. Recall, since the large masses are usually palpable and therefore easily diagnosed, no attempt was made to reduce the number of Large false ROIs. The final algorithm's parameters were set to retain the top seven ROIs with ratios between three and thirty-six plus all of the Large ROIs with ratios between thirty-six and fifty. These settings were used for the Malignant Eval Set, the Benign Mass Set, and all other images.

For the Malignant Mass Eval Test Set, the results are listed in Figure 24a for the Medium ROIs and Figure 24b for the Large ROIs. Of the 12 malignant masses present, 7 of them were indexed as Medium ROIs with 86 Medium false ROIs. One malignant mass was indexed as a Large ROI with a corresponding single Large ROI false alarm. Of the four masses that were missed, one image was very dark, one mass was ranked eighth, one was too close to the film edge, and one was really only discernible in the opposing view (the 8th ranked ROI image). The dark image mass could fairly easily be detected by lowering the mask threshold from 1500 to 1480 and the eighth-ranked mass could be detected by

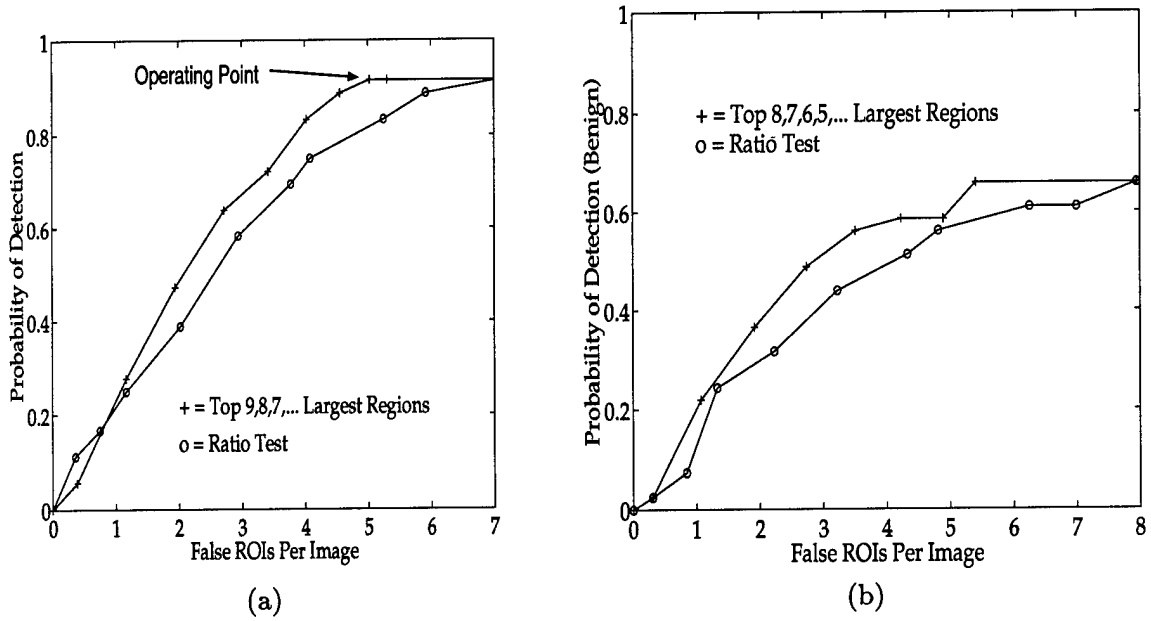
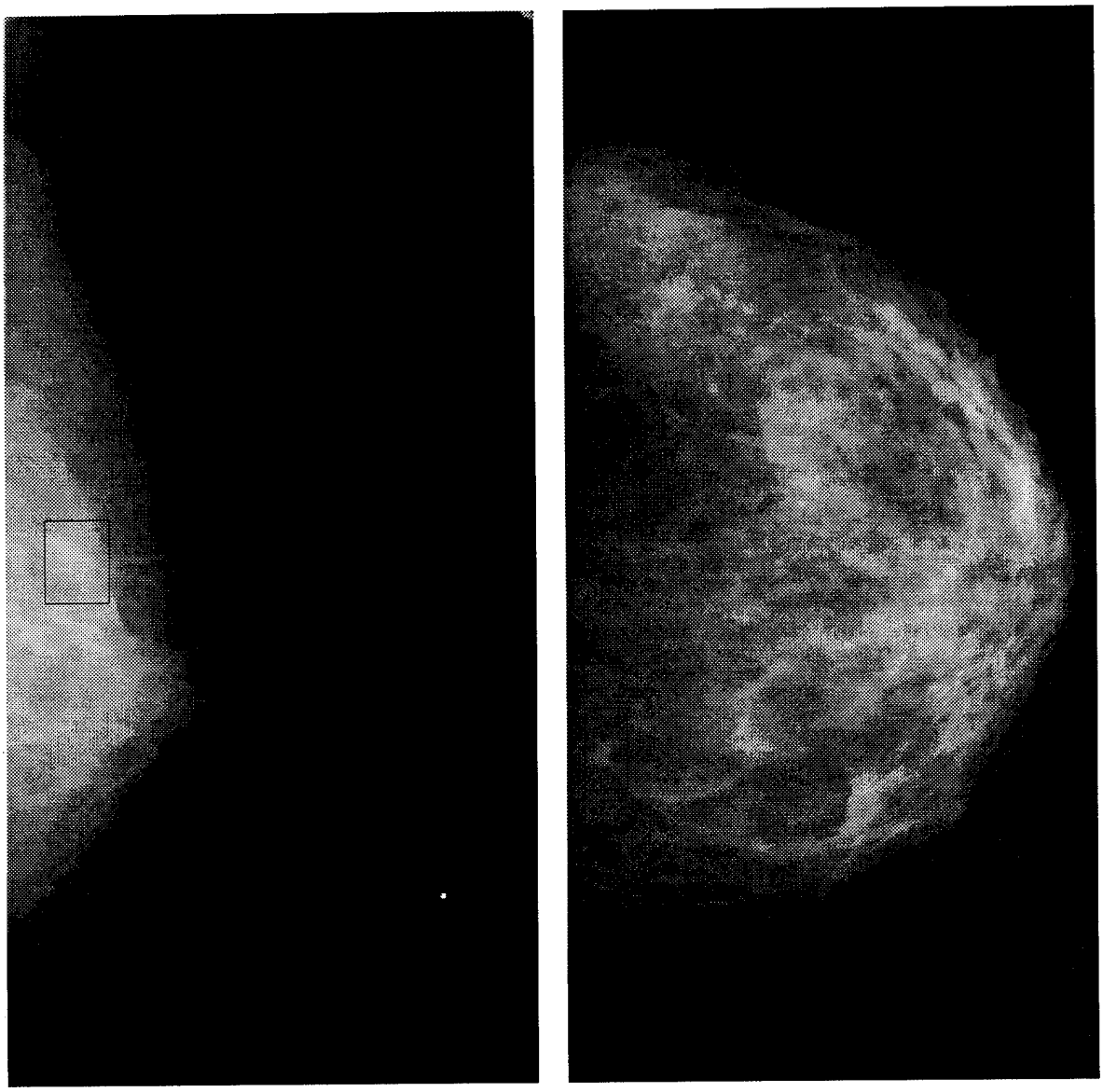


Figure 25. ROC Curves for the Focus of Attention Module
(a) Combined Train and Test Malignant Sets (b) Benign Set

passing one more ROI per image to the Indexing Module's morphological tests. Neither of these changes would increase the false alarm count much since the Index Module only passes the top four regions anyway.

For the Benign Mass Set, recall that images in this set contained a radiologist identified mass that was biopsied and proven to be benign. The intent of the FOA module is to either reject these masses outright, or pass them along to the rest of the MBV Modules to identify these ROIs as non-cancerous regions. Of the 53 biopsied benign masses present, 36 of them were indexed as Medium ROIs with 414 Medium false ROIs. One of the benign masses was indexed as a Large ROI with a corresponding 16 Large false ROIs. The ROC curve in Figure 25b shows the tradeoff for adjusting the ratios and the top rank ratio ROIs retained that defined what regions were sent to the Indexing Module. The same criterion used to make the malignant ROC curves were used to make the benign ROC curves.

Figure 24c contains the overall results for the Focus of Attention Module. It treats all non-malignant ROIs as false ROIs and includes both Indexed classes. Including the two missed Eval Set masses as detections, brings the true-positive fraction to 0.90.



(a) (b)

Figure 26. Two Example Malignant Images that Failed to Pass the FOA Module Process:
(a) The Box Shows the Missed Malignant Mass Within Dense Tissue.
(b) The Box Shows the Missed 4 mm Malignant Mass.

5.2 *Indexing*

Recall, the Indexing module consisted of the steps shown in Figure 27.

1. Ratio test the ROIs ($3 < \text{ratio} < 50$).
2. Index the ROIs, and only continue with the Medium Mass ROIs.
3. Rank the ROIs in descending order and keep ≤ 7 ROIs.
4. Threshold select the top 15 percent of the ROI's histogram.
5. Erode the binary mask histogram image.
6. Eliminate any edge pixel groupings.
7. Dilate twice.
8. Eliminate pixel groupings with < 1000 pixels.
9. Morphologically close twice.
10. Select ROIs with > 1000 pixels.
11. Select ROIs with contrasts between 1.05 and 1.13.
12. Select ROIs with a circularity coefficient > 0.58.
13. Select ≤ 4 ROIs per image based on the circularity test.

Figure 27. Index Module process

The purpose of the thirteen steps was to separate the ROIs into Large and Medium Indexes, and then reduce the number of false ROIs. Since, the large masses are easily diagnosed and since there were only two Large Malignant ROIs, the Large ROIs were not processed by the Index Module. So, steps 3 through 13 apply only to Medium ROIs.

Based on the results of the morphological operations and tests (steps 3 - 9), many of the masks contained few or no 'on' pixels in them. Many false ROIs occurring near the breast-background transition had the brightest pixels near the edge of the ROI, and were eliminated by these steps. Then, the area, contrast, and circularity tests each reduced the false ROI rate further. Figure 28 lists the final results for the different datasets. All 284 images were used and the results recorded, but only Medium ROIs were used in these tests. The overall results were 80 correct classification of malignant masses with 43 percent of the benign masses retained and 2.36 false ROIs per image.

For the Malignant Mass Train and Test Set results, of the 36 malignant masses present, once again 31 of them were indexed as Medium ROIs with 67 Medium false ROIs, and the two medium malignant regions still had the two Medium ROIs associated with them. Thus there were a total of 33 ROIs. The ROC curve in Figure 29a shows the tradeoff for adjusting the parameters for each of the tests. The area parameters were varied from

Data Set	Number of Images	Correct Regions Found	True Positive Fraction	Total False Alarms	False Alarms per Image
Malignant Mass Train and Test Set	36	31/33	0.94	67	1.9
Malignant Mass Eval Set	12	4/11	0.36	27	2.3
Benign Mass Set	53	23/54	0.43	111	2.1
Non-Mass Set	183	na	na	465	2.5
Combined Results	284	35/44	0.80	670	2.36

Figure 28. Indexing Module's Performance for Medium ROIs.

>1000 pixels to >3500 pixels, the contrast parameters were varied from >1.05 to >1.13, the circularity parameters were varied from >0.58 to >0.88, and the top seven through the top one circular ROIs were retained for the final ROC curve. The final parameters were shown in Figure 27 above.

For the Malignant Mass Eval Test set results, of the 11 Medium ROI malignant masses present, 5 were indexed as Medium ROIs with 27 Medium false ROIs. Only 7 Medium ROIs were actually passed to the Index Module, so 2 of the 7 were rejected by the morphological tests. One was touching the edge of the film and was thus eliminated by the morphological test, and the other had a contrast of 1.16. The edge ROI could be detected by changing the process to allow masses touching the left edge of the ROI to pass the first morphological step. The additional false ROIs passed would probably be rejected using the other tests. The other ROI could easily be included by eliminating the upper limit on the contrast test. There is really no reason for the upper limit, since only true masses should have that high of a contrast. Thus, the algorithm should be able to pick up two more malignant masses with a small increase in false ROIs.

For the Benign Mass Set results, of the 54 benign biopsied masses present in the 53 images, 23 of them were indexed as Medium ROIs with 111 Medium false ROIs. The ROC curve in Figure 29b shows the results for this data set. The parameters were adjusted in the same way as they were for the malignant ROC curve.

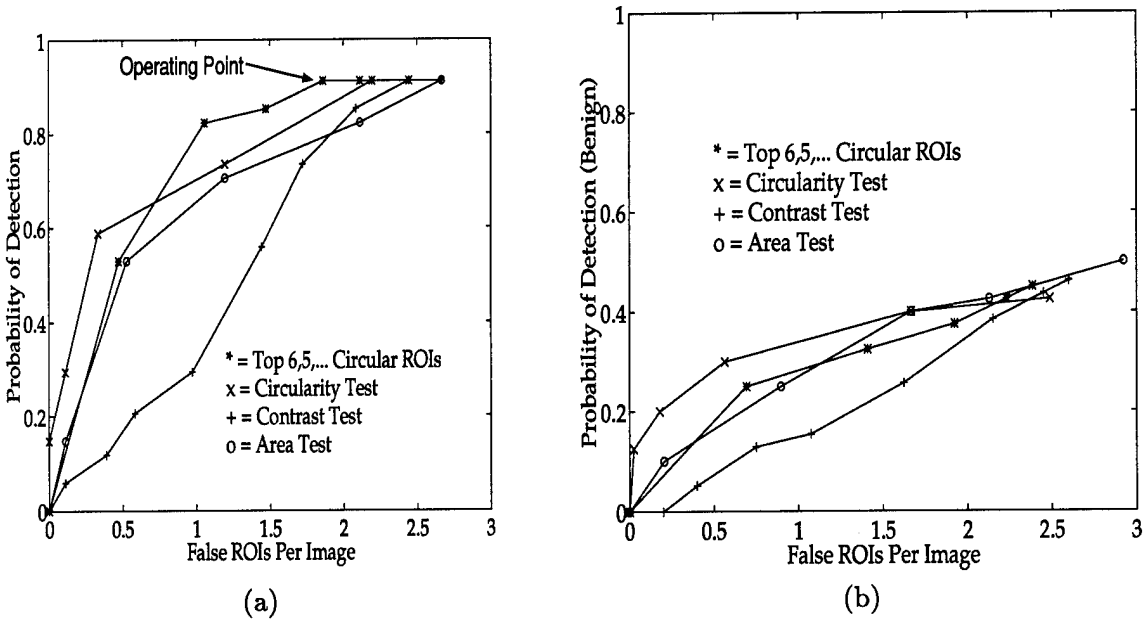


Figure 29. ROC Curves for the Indexing Module (Medium ROIs only)
(a) Combined Train and Test Malignant Sets (b) Benign Set

The bottom row of Figure 28 contains the overall results for the Indexing Module. It treats all non-malignant ROIs as false ROIs. With the changes proposed to pass the two Medium malignant masses in the Eval set, the true positive fraction would increase to 0.84. If the two additional images rejected by the FOA Module were successfully passed by the Index Module, the true positive fraction could reach 0.89. Passing these four Medium ROIs through the Index Module increases the Eval Set true positive rate to 0.72 which is much closer to the 0.94 rate achieved with the Train and Test sets.

5.3 Prediction / Feature Selection

The Prediction Module extracted the Laws features and the Index Module's four features to determine the best ones to use to develop the malignant mass and non-malignant tissue models. It used all of the ROIs from the Malignant Mass Train and Test sets and only the correctly segmented biopsied regions from the Benign Mass Set. The benign biopsied regions were included to ensure the model would be defined using regions that radiologists had identified as very close to being malignant. The other false ROIs from

Rank	f-ratio	Feature	Rank	f-ratio	Feature	Rank	f-ratio	Feature
1	0.1774	26-area	11	0.0591	7-s5s5	21	0.0476	22-w5s5
2	0.0855	4-l5e5	12	0.0574	21-w5l5	22	0.0474	5-l5w5
3	0.0850	19-e5e5	13	0.0574	20-e5w5	23	0.0472	24-w5e5
4	0.0825	27-circle	14	0.0558	17-e5s5	24	0.0436	15-r5w5
5	0.0688	23-w5r5	15	0.0539	9-s5e5	25	0.0414	12-r5s5
6	0.0627	10-s5w5	16	0.0527	29-ratio	26	0.0391	14-r5e5
7	0.0622	16-e5l5	17	0.0517	25-w5w5	27	0.0384	11-r5l5
8	0.0615	8-s5r5	18	0.0514	18-e5r5	28	0.0301	3-l5r5
9	0.0596	13-r5r5	19	0.0513	2-l5s5	29	0.0034	28-contrast
10	0.0592	6-s5l5	20	0.0502	1-l5l5			

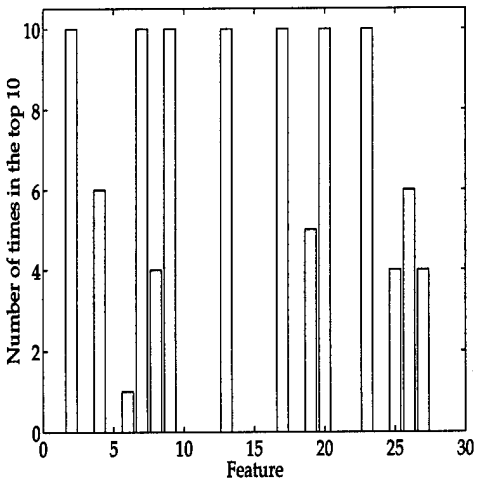
Figure 30. F-ratio ranking for the 29 features.

the Benign Mass set were not included to keep the imbalanced training set problem to a minimum. As it was, there were 117 samples: 33 malignant and 84 benign.

The twenty-five Laws features described in Section 4.3 and the ratio, area, contrast, and circularity features from the 117 ROIs were ranked according to their f-ratio. The ranking is shown in Figure 30.

Although the actual f-ratio values did not show a dramatic break point for good versus bad features, the area feature was much better than any of the other features. As for the Laws features, the spot (s) and edge (e) filters performed better than the others, and the ripple filters were the worst overall. The first two columns of features (20 in all) in Figure 30 were retained and used for the derivative-based feature saliency test.

The derivative-based feature saliency test was done using the top twenty f-ratio features using the imbalanced neural network algorithm with sigmoidal activations and one clamped output node. The clamp was set to 0.1. Since batch algorithms tend to get stuck in local minima of the error surface, tracking of the initial starting point on the error surface was monitored. This was done by evaluating the mean square error (mse) for both classes for the first epoch. These mse's ranged from 0.08 to 0.45 for either class. Since the desired outcome was to correctly classify malignants at the expense of incorrect benign classifications, the algorithm was restarted if the initial mse for the malignant class for epoch one was >0.20. Then to obtain reasonable statistics, ten networks with random initial weights were run and the results compared.



(a)

Derivative Rank	Feature Number	Kernel	F-ratio Rank
1	23	w5r5	5
2	13	r5r5	9
3	7	s5s5	11
4	9	s5e5	15
5	17	e5s5	14
6	2	l5s5	19
7	20	e5w5	13
8	4	l5e5	2
9	26	area	1

(b)

Figure 31. (a) A histogram of the number of times each feature appeared in the top ten ranking of features for ten independent neural network feature saliency trials. (b) The Feature Ranking

Figure 31 shows the occurance histogram depicting the number of times each feature appeared as one of the top ten features in each of the ten independent feature saliency trials. Thus, the best a feature could do was appear ten times in the histogram. Features 2, 7, 9, 13, 17, 20, and 23 each occurred ten times, with features 4 and 26 appearing six times. The top features were also listed by their average ranking for the ten trials.

The ten networks were designed to start at a position in weight space that kept the Mean Square Error (mse) after the first epoch for the malignant class samples < 0.20. This ensured the networks favored classifying malignant samples at the expense of benign samples. As shown in Figures 32, there was not much of a problem with the networks

	Malignant	Benign
mse Epoch 1	0.19±0.01	0.32±0.02
mse Epoch 10000	0.11±0.01	0.21±0.01

(a)

Actual/Classify	malignant	benign
malignant	30±1	3±1
benign	20±2	64±2

(b)

Figure 32. (a)Mean square error results per class for feature saliency trials.
(b) Feature saliency trial confusion matrix.

	Malignant	Benign
mse Epoch 1	0.15±0.04	0.39±0.07
mse Epoch 10000	0.09±0.04	0.23±0.06

(a)

Actual/Classify	malignant	benign
malignant	30±3	3±3
benign	30±11	54±11

(b)

Figure 33. (a) Mean square error results per class for network architecture trials.
(b) Network architecture trial confusion matrix.

getting stuck in local minimums for this architecture and feature set. The networks' average mse per class for the first and the ten thousandth epoch are shown as is the confusion matrix. The standard deviations are shown too.

It is interesting to note, that the four kernels Kegelmeyer picked from the literature as being the best ones (l5s5, l5e5, r5r5, and e5s5), also appeared in the top nine derivative-based saliency features. With the nine best features in hand, the model was ready to be used to classify all Medium ROIs.

5.4 Matching / Classification

The Classification module used the nine features found from the Prediction Module and determined the best neural network architecture and weights to use for the classification of all the Medium ROI samples. Two hidden nodes were chosen for the architecture. This yielded 23 connections for 117 training samples for an Uncle Bernie ratio of five. Of course, Foley's rule was satisfied since there were only 9 features and 33 malignant samples. Ten trials were run, but this time there was much greater variability in the results. Notice the standard deviations in Figure 33.

Of the ten trials, the weights for the best trial were chosen as the best model/classification architecture combination. These weights correctly classified all 37 malignant samples and 284 of the 670 benign samples. (Recall that two masses had two ROIs associated with them.) The classifier achieved a true positive rate of 1.0 and a false positive rate of 0.58 for all Medium ROIs from the 284 images. Of the 23 remaining benign biopsied ROIs, 8 of them were classified correctly as benign. Figure 34 shows an example final output to the radiologist with the computer's diagnosis. The computer diagnosis correctly identified the malignant mass and two of the three benign regions. Thus, there was one false alarm, and three correct diagnoses for this image.

5.5 Analysis Summary

Recalling all of the steps in the process: First the algorithms were developed and tested on the 36 images of the Train and Test sets. Then the algorithms were evaluated on all the other 248 images including 12 malignant mass images and 53 benign biopsied images. The final results for detecting and classifying both Large and Medium ROIs for all 284 images are shown in Figure 35. The results for all of the data, excluding the 12 Malignant Mass Eval set images, are shown in Figure 36. As discussed above, options are available to lift the Eval set performance up to the level of the Train and Test set values.

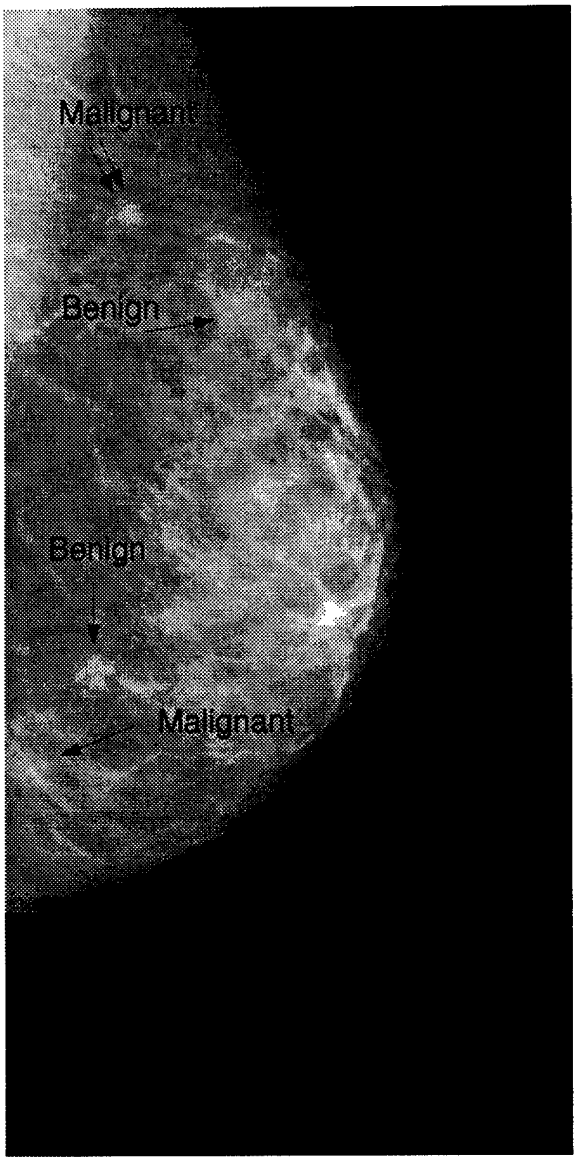


Figure 34. The final output to the radiologist. It shows the original example image with the suspicious regions and the computer's classification identified. The true malignant mass is identified by the double arrow. The other three regions are truly benign regions.

Data Set	Number of Images	Malignant Regions Found	True Positive Fraction	Total False Alarms	False Alarms per Image
FOA Module All Indexes	284	41/48	0.85	2367	8.3
Index Module Medium ROIs	284	35/44	0.80	670	2.36
Matching Module Medium ROIs	284	37/37	1.0	403	1.4
Final Results All ROIs	284	37/48	0.77	512	1.8

Figure 35. Final Results for the FOA Module, Index Module, and Matching Module using all data.

Data Set	Number of Images	Malignant Regions Found	True Positive Fraction	Total False Alarms	False Alarms per Image
FOA Module All Indexes	272	33/36	0.92	2280	8.4
Index Module Medium ROIs	272	31/33	0.94	643	2.36
Matching Module Medium ROIs	272	33/33	1.0	376	1.4
Final Results All ROIs	272	33/36	0.92	484	1.8

Figure 36. Final Results for the FOA Module, Index Module, and Matching Module using all but the 12 malignant Eval Set images.

VI. Conclusion

6.1 Summary

The Focus of Attention Module's DoG filter in conjunction with the Indexing Module's morphological operations and tests proved to be an effective tool for locating malignant ROI's in mammograms. Considering the 272 image database, the MBV algorithm detected 92 percent of the malignant ROIs with less than two false malignant ROIs per image. The Prediction/Feature Selection Module tested the Laws features that were most applicable for classification, and through derivative-based feature saliency testing, it found the best nine features. The classification results using these nine features were a true-positive rate of 1.0 and a false-positive rate of 0.58 for the Medium ROIs. The false-positive rate corresponded to an average of 1.8 false malignant ROIs per image for both Medium and Large ROIs combined.

These results compare very well to the most current and relevant results in the literature shown in Figure 37. Of the large database researchers, only Kegelmeyer has done both segmentation and classification in one system, but he restricted his database to spiculated masses only. This work's results are much better than Yin's in terms of false alarms, and are comparable to both Yin's and Kegelmeyer's true-positive fractions for the segmentation process. The classification results are better than Giger's in terms of the false-negative fraction, and are comparable to both Giger's and Wei's true-positive fractions. Thus, the MBV system, with it's ability to Focus, Index, Model, and Match, has good possibilities for implementation in a breast cancer diagnosis system.

The new contributions this thesis made were:

The first approved medical protocol with the Wright Patterson Air Force Base hospital was accomplished as a part of this thesis. The protocol was initiated to acquire the database necessary for this work.

The database generated as a part of this thesis is the largest, high resolution database of any of the current researchers' databases. It has a higher resolution than Kegelmeyer's database (2.4 times) and Giger's and Yin's database (4 times). This provides more in-

Researcher	Database	False Alarms per Image	True Positive Fraction	False Positive Fraction	Citation
Brzakovic	25	0.8	0.8	na	(34)
Lai	17	1.7	1.0	na	(1)
Kegelmeyer 1	62	0.27	1.0	na	(36)
Kegelmeyer 2	340	0.28	0.97	na	(37)
Wei	168 ROIs	>2	0.95	0.55	(38)
Yin 1	92	3	0.95	na	(40)
Yin 2	308	6.5	0.91	na	(38)
Giger	50 ROIs	na	0.95	0.79	(38)
Polakowski	272	1.8	0.92	0.58	

Figure 37. Comparison of DoG Results to Alternate Techniques for Detection of Masses in Mammograms.

formation for detection and diagnosis, and the detailed description of each biopsied case allows more insight into malignant vs benign tissue's characteristics.

This is the first application of the Model Based Vision (MBV) process to the detection and diagnosis of breast cancer masses. New ratio, area, and contrast features were developed for this thesis, and the MBV process is in place and operational.

This is the first application of a physiologically motivated Difference of Gaussians (Dog) filter to breast cancer detection. It models the best breast cancer detection system in use, the radiologist and their optical detector, the eye.

This is the first application of any feature saliency algorithm to the selection of mass-specific features. In most cases, researcher's used trial and error to determine the best features.

This is the first use of data partitioning to determine the extensibility of the algorithm to the clinical environment.

6.2 Conclusions

This work has shown that the Model Based Vision approach is well-matched to a CADx system for breast cancer diagnosis. Using the human-based perception Difference of Gaussian technique for focusing the radiologist's attention on a small number of regions of interest in a mammogram could greatly improve their diagnosis capability. But, by also

indexing and classifying those regions as malignant regions or benign regions, the CADx system can be an instrumental breast cancer diagnosis tool as a fail-safe or second reader for the radiologist. At 92 percent correct segmentation with less than 2 false malignant ROIs per image, this CADx system is ready to be tested in a clinical setting as a radiologist's aide.

Matlab Filename	AFIT Filename	WPAFBH Filename	Mass Diameter	Row Coord	Column Coord
mal1	af004	AFIT001	8mm	370	230
mal2	af006	AFIT004	20mm	1210	370
mal3	af009	AFIT006	14mm	1430	570
mal4	af011	AFIT006	14mm	1755	360
mal5	af014	AFIT009	8mm	1605	240
mal6	af016	AFIT009	8mm	1475	300
mal7	af033	AFIT014	9mm	464	310
mal8	af035	AFIT014	9mm	1040	420
mal9	af042	AFIT016	10mm	1128	129
mal10	af044	AFIT016	10mm	1540	100
mal11	af055	AFIT021	15mm	1320	220
mal12	af057	AFIT021	15mm	770	370
mal13	af067	AFIT028	11mm	1030	890
mal14	af069	AFIT028	11mm	914	140
mal15	af079	AFIT035	10mm	1564	704
mal16	af081	AFIT035	10mm	1114	400
mal17	af083	AFIT036	6mm	490	110
mal18	af085	AFIT036	6mm	1180	300

Table 1. Malignant Mass Training Set

Note: The mass center coordinates are referenced from the top left of each image after the image has been oriented with the chest wall on the left of the image.

Matlab Filename	AFIT Filename	WPAFBH Filename	Mass Diameter	Row Coord	Column Coord
mal19	afl10	AFIT050	10mm	750	340
mal20	afl23	AFIT056	10mm	540	230
mal21	afl25	AFIT056	10mm	870	250
mal22	afl04	AFIT049	10mm	530	120
mal23	afl49	AFIT062	11mm	1100	80
mal24	afl51	AFIT064	15mm	750	270
mal25	afl53	AFIT064	15mm	1450	330
mal26	afl72	AFIT073	4mm	1420	75
mal27	afl74	AFIT073	4mm	820	100
mal28	afl75	AFIT074	14mm	750	110
mal29	afl77	AFIT074	14mm	1090	140
mal30	afl89	AFIT083	15mm	490	100
mal31	afl90	AFIT083	30mm	775	460
mal32	afl91	AFIT083	15mm	550	360
mal33	afl92	AFIT083	30mm	870	770
mal34	afl93	AFIT090	9mm	1210	80
mal35	afl95	AFIT090	9mm	690	210
mal36	afl06	AFIT049	10mm	880	170

Table 2. Malignant Mass Test Set

Matlab Filename	AFIT Filename	WPAFBH Filename	Mass Diameter	Row Coord	Column Coord
eval1	af206	AFIT094	20mm	620	250
eval2	af208	AFIT094	20mm	850	540
eval3	af210	AFIT098	20mm	1010	280
eval4	af212	AFIT098	20mm	1570	130
eval5	af243	AFIT115	9mm	500	150
eval6	af245	AFIT115	9mm	930	420
eval7	af248	AFIT117	8mm	310	250
eval8	af250	AFIT117	8mm	770	500
eval9	af278	AFIT134	12mm	880	50
eval10	af280	AFIT134	12mm	760	440
eval11	af285	AFIT138	15mm	1540	110
eval12	af287	AFIT138	15mm	1470	100

Table 3. Malignant Mass Eval Set

Matlab Filename	AFIT Filename	WPAFBH Filename	Mass Diameter	Row Coord	Column Coord
ben1	af025	AFIT012	14mm	550	470
ben2	af027	AFIT012	14mm	1080	500
ben3	af030	AFIT013	16mm	580	580
ben4	af032	AFIT013	16mm	1030	730
ben5	af046	AFIT017	7mm	1300	730
ben6	af048	AFIT017	7mm	1400	580
ben7	af050	AFIT018	7mm	1480	130
ben8	af064	AFIT024	17mm	580	130
ben9	af066	AFIT024	17mm	800	300
ben10	af071	AFIT032	14mm	1180	350
ben11	af073	AFIT032	14mm	1430	300
ben12	af076	AFIT034	14mm	1280	730
ben13	af078	AFIT034	14mm	1730	680
ben14	af087	AFIT037	11mm	1230	760
ben15	af089	AFIT037	11mm	1680	640
ben16	af093	AFIT039	10mm	630	75
ben17	af095	AFIT042	8mm	410	10
ben17	af095	AFIT042	8mm	1180	440
ben18	af097	AFIT042	8mm	680	400
ben18	af097	AFIT042	8mm	1220	520
ben19	af116	AFIT054	20mm	1500	220
ben20	af118	AFIT054	20mm	1050	320
ben21	af132	AFIT058	22mm	1080	120
ben22	af134	AFIT058	22mm	860	460
ben23	af135	AFIT059	10mm	880	360
ben24	af137	AFIT059	10mm	900	680
ben25	af151	AFIT064	15mm	825	340
ben26	af153	AFIT064	15mm	1510	264
ben27	af155	AFIT065	15mm	740	124
ben28	af156	AFIT065	20mm	910	210
ben29	af157	AFIT065	15mm	870	480
ben30	af158	AFIT065	20mm	1050	550
ben31	af163	AFIT068	12mm	900	220
ben32	af165	AFIT068	12mm	1350	760
ben33	af167	AFIT071	10mm	1635	460
ben34	af169	AFIT071	10mm	1760	280
ben35	af180	AFIT075	16mm	1350	280
ben36	af182	AFIT075	16mm	1240	230
ben37	af183	AFIT078	12mm	525	920
ben38	af197	AFIT091	8mm	1230	170
ben39	af199	AFIT091	8mm	650	410

Table 4. Benign Mass Set Images 1 - 39

Matlab Filename	AFIT Filename	WPAFBH Filename	Mass Diameter	Row Coord	Column Coord
ben40	af213	AFIT099	11mm	1120	520
ben41	af215	AFIT099	11mm	1650	590
ben42	af217	AFIT101	10mm	1780	120
ben43	af228	AFIT110	13mm	1350	710
ben44	af230	AFIT110	13mm	1650	700
ben45	af235	AFIT112	12mm	870	650
ben46	af237	AFIT112	12mm	1630	750
ben47	af252	AFIT118	11mm	700	620
ben48	af254	AFIT118	11mm	1430	530
ben49	af256	AFIT120	17mm	950	620
ben50	af258	AFIT120	17mm	1250	480
ben51	af272	AFIT132	9mm	420	510
ben52	af274	AFIT133	7mm	1280	450
ben53	af276	AFIT133	7mm	1230	670

Table 5. Benign Mass Set Images 40 - 53

Appendix B. WPAFB Protocal Letter

EXEMPT PROTOCOL SUMMARY

TITLE: Computer-Aided Breast Cancer Diagnosis Using Digitized Film Mammograms

PRINCIPAL INVESTIGATOR: Steven K. Rogers, PhD
Professor of Electrical and Computer Engineering

FACILITY: The Air Force Institute of Technology (AFIT)

1. SUMMARY. We request the short term use of selected Wright-Patterson Air Force Base Hospital (WPAFB) film mammograms to digitize them for later computer analysis. The goal is to develop a Computer-Aided-Diagnosis (CADx) system to aid radiologists in accurately diagnosing Mammographic films. Privacy act regulations will be followed with patient names being covered during the digitization process to ensure privacy. The subject population is women who have had Mammographic screening at the WPAFB Hospital. Only mammograms from patients medical files will be used so there is no patient risk or any other risk involved.

a. AFIT has a breast cancer diagnosis group, led by Prof. Rogers, that is pursuing the use of computers to diagnose masses and microcalcifications in digitized mammograms as malignant or benign. We have been working in this area for two years, with close cooperation from Capt Jeffrey Hoffmeister, M.D. from Armstrong Laboratory's Crew Systems Directorate at WPAFB. The groups' objective is to transfer its thirty years worth of military image processing experience to medical CADx. We are continuing with last years highly successful results of 88% correct classification of difficult-to-diagnose microcalcification cases obtained from the University of Cincinnati. This year, there are four AFIT Master's Program students working on various CADx implementations. To increase the statistical validity of the results and robustness of the algorithms, we require a larger and higher resolution data base.

b. By combining the WPAFB Hospital medical files with the high resolution digital cameras at the WPAFB graphics shop, we can add to our current database and obtain the digital resolution required for our classification algorithms. The real benefit will be the increase in resolution from 160um currently to 9um for the WPAFB Hospital images. The resolution has a dramatic effect on the accuracy and types of classification techniques employed. Based on Capt Hoffmeister's estimation, we can increase the number of images over last year's

database from 95 (38 malignant, 57 benign) by approximately 50 malignant and greater than 50 benign images and also obtain the other contralateral mammograms for further classification accuracy. In addition, we hope to get associated data (such as age) to add those risk factor analyses to the CADx algorithms.

2. PROCEDURE. This procedure will be used to digitize five to ten cases at a time until all of the cases are exhausted. The procedure for digitizing the films follows:

a. Capt Hoffmeister reviews medical records and pulls cases with microcalcifications.

b. All mammograms are kept in the original folder. A naming system is used to track the individual films of interest: target film with microcalcification(s) and matching film of opposite breast.

c. The location of the region of interest is noted with the diagnosis.

d. Personal data is covered with a label, historical risk factor data is compiled, and the film is digitized using the Bldg 20 Area B camera facility.

e. Films are returned within two working days.

f. Digitized versions of the films are transferred to AFIT control for CADx testing.

3. MANHOURS. The bulk of the manhours required for this study derive from the six months of fulltime work from the four AFIT students pursuing their theses. There is minimal impact on hospital staff. The estimates are for digitizing ten cases at a time.

Med Center Personnel	
Medical record review (records clerks)	0.5 hrs
AFIT Personnel	
Medical record review (Capt Hoffmeister)	4.0 hrs
Risk factor annotation	2.0 hrs
Film annotation for tracking (Students)	2.0 hrs

Digitization (Students)	3.0 hrs
Total/10 cases	11.5 hrs
Total manhours for 100 cases =	115 hrs

Analysis (Thesis Work)

4 students full time for 6 months

2 faculty half time for 6 months

4. STATISTICS. The number of samples (malignant and benign) we have in our database will determine the number of features we can use to have statistically relevant results. Under certain restrictions on the training samples, Foley's Rule says that when there are three times as many training samples per class as there are features, the error rate on the training set is a good predictor of the error rate on an independent test set. With our current database limited to 38 malignant samples, using Foley's rule, we can have a maximum of about 12 features to use in our neural networks or other classification algorithms. By adding the WPAFB Hospital data (50 more samples of microcalcifications), we hope to increase our feature set to almost 30. Whatever the number of final images we obtain, any data will help in our classification success.

5. For further information and consultation, contact
Capt Bill Polakowski, 252-4476 or email wpolakow@afit.af.mil.

Steven K. Rogers, PhD
Professor of Electrical and
Computer Engineering

Appendix C. Digitization Procedure

C.1 Search the Pathology Follow-Up Book Review (PFUBR) :

NOTE : Jeff Hoffmeister, MD, completes this section.

a) Search for microcalcification cases, they may not be classified this way.

b) Fill out the PFUBR, but DO NOT ASSIGN A CASE NUMBER!

C.2 Pull Film Jackets for Patients Identified in the PFUBR :

NOTE : Jeff Hoffmeister, MD, completes this section

a) Look at the last 2 of the last 4 digits of the patient's SS# :
   i) 00-18 : upstairs
   ii) 19-99 : downstairs

b) Fill out an OUT form and place it in an OUT folder, then replace the film jacket with the OUT folder. The OUT form should contain :
   i) Patient's Name
   ii) Last four numbers of the patient's SS#
   iii) Date
   iv) Room number where the films are being reviewed

c) Bring PFUBR and the film jackets to a mammogram reviewer

C.3 Review Mammograms :

NOTE : Jeff Hoffmeister, MD, completes this section.

a) Find 4 films from the screening exam for which a lesion was identified for biopsy.
   NOTE: Make sure the most recent mammograms, "for biopsy", are used.

b) The group of four mammograms should contain one of each of the following :
   i) R-CC - Right Cranial-Caudal
   ii) L-CC - Left Cranial-Caudal
   iii) R-MLO - Right Medial-Lateral Oblique
   iv) L-MLO - Left Medial-Lateral Oblique

c) Fill out mammogram review and the mammogram diagrams.

d) The order the mammograms should stay in throughout this process is RCC, LCC, RMLO, and LMLO.

C.4 Digitize Mammograms :

a) Start up the Big Mac (Mac IIfx) found in the basement of the X-Ray filing room. There is a large folder in the center of the screen with three icons on it.

b) Select "Film Digitizer".

c) Login as a Registered User using HOFFMEISTERJ, hit **TAB**, then type the

password obtained from Jeff, and hit return.

d) Goto "Special" on the button bar on top of the screen, then select "Calibrate" from the pull-down menu.

NOTE : The digitizer only needs to be calibrated once per day.

e) Put six, 14" x 17" films into the digitizer, which is located to the right of the Big Mac. The films are located in the yellow box near the digitizer, usually leaning up against the side.

f) Hit "OK" and the calibration procedure will begin.

WHILE THE DIGITIZER IS CALIBRATING

g) Login into Lil Mac (Mac IIsi) located to the left of Big Mac. There should be a large folder in the center of the screen with two icons on it.

h) Select Paris.

i) Type in patient's name as **AFIT, TEST** in the box under where it says Patient's Name, then hit return.

ii) Click on big "Exams for a Patient" icon on the top of the Paris window.

iii) Under the list of exams for "AFIT, TEST", select an exam. It does not matter which exam is selected.

iv) Goto "Exams" on the button bar on top of the screen, then select "Duplicate" from the pull down menu.

v) Input "Mammography" in Requesting Ward/Clinic Box located in the center right portion of the screen (if necessary).

vi) Click on "New Exam" button at the bottom of the screen and pull the bar code slip from the machine on the left.

NOTES: Lil Mac will not be needed any further. Lil Mac will shut down on its own. A bar code slip is good for 16 hours.

AFTER BAR CODE PROCEDURE IS COMPLETE

i) Scan the bar code by holding the scanner, located between Big and Lil Mac, at a 45° angle,approximately 6-8 inches away from the bar code. Pull the trigger and listen for the beep from Big Mac.

h) Place mammography films into the digitizer. Make sure that the films are as clean as possible (ie: wipe off any grease pencil or other smudges).

i) Place in sticker side down. **IMPORTANT!! :** The LMLO view is on the bottom, then RMLO, LCC, and the RCC goes on top. This order is critical in the naming of the files.

ii) Place the long side of the films flush against the right side of the auto-feeder on the digitizer

iii) On Big MAC, click on "Digitize" in the Film Digitizer window, then "OK". The digitization process will then start.

iv) When all the films have been digitized, click "OK".

iv) Goto "File" on the button bar on top of the screen and select "Quit" from the pull down menu, then "Quit" on the Login window.

C.5 Getting Image Specs :

a) Check for filename and image size
   i) From the folder in the Main Window, choose the "LiteBox" icon
   ii) Goto the "Patient Name" box and type in AFIT, TEST
   iii) Select exam, it should be the exam with an "I" in the ONLINE column. The "I" means that the process is still active
   iv) Click on "Images" on top of the window, then "Display"
   NOTE : The process takes awhile. There is a display in the lower left hand corner that updates the number of images loaded
   v) Click on an image to select it
   vi) Goto "Image" on the top button bar and select "Image About"
   vii) Record filenames and image sizes, the size is in the middle of the window and the filename is at the bottom
   NOTE : When copying the filename DO NOT include the WSU prefix
b) Press the $\gg$ button in the lower right hand corner to view the next image. Repeat the same process for all images
c) To exit "LiteBox" select "File" from the top button bar and then "Quit"
   NOTE : Never save changes

C.6 Naming Images :

a) Recommended file naming convention :
   NOTE : The final filenames will be given after the images have been chopped
   AFITxxx.yz
   xxx = case number from the mammogram review
   y = 1: R-CC
   y = 2: L-CC
   y = 3: R-MLO
   y = 4: L-MLO
   z = M: Malignant
   z = B: Benign

C.7 Finishing Up Digitization :

a) Fill out the mammogram diagrams.
b) Place the films back into the Mammogram jacket inside the film jacket.
c) Return films to the filing room (DO NOT RE-FILE!).
d) Download data to a portable MAC Hard Drive. **NOTE :** Curt has fought for our privilege to download our files, so follow the directions carefully.
   i) Obtain the adaptor cable for the portable MAC drive from Curt.
   ii) Connect the portable MAC to the big MAC-see Curt if you have questions on this step. **NOTE :** Make sure the system is shut-down before connecting the portable MAC!

iii) Start system back up and goto "File" on the button bar, then "Go to Finder" **NOTE :** The password must be obtained from Curt.

iv) Place disk in portable MAC, and an icon labeled Polakowski will appear.

v) Erase disk by clicking on "Special", "Erase Disk", sometimes after erasing, the disk must be ejected and re-inserted for the correct directory display.

vi) Click on the File Copy Icon. **NOTE :** This procedure only works one time, after copying a file quit the window and reopen the File Copy Icon.

a) First prompt, drive where file is to be copied, type -2 for WSU:

b) Second prompt, filename, type the entire case-sensitive filename.

c) Third prompt, drive to be copied to, type -3 for Polakowski drive.

d) Fourth prompt, filename to be saved as, use the AFITxxx.yz format.

vii) When the disk is full, goto "Special", "Eject Disk", place another disk in the portable drive and repeat the process.

viii) Return the adaptor cable to Curt.

e) Return to AFIT and transfer files into our system.

i) Use Lab MAC II in Rm. 2011 and connect the portable MAC drive. **NOTE :** Make sure the system is shut down before connecting the portable drive.

ii) If Polakowski icon does not appear.

a) Click on Romulus icon to open the window.

b) Select "Utilities", then "Alliance Power Tools".

c) Look down the "Product" column for Beta 150 and select the corresponding ID **NOTE :** The icon may appear as soon as you open the "Appliance Power Tools" window.

d) Click "Mount", then "Quit".

iii) Open the Polakowski drive window.

iv) Click on the Apple icon in the upper left corner of the screen and scroll down to "TPC/Connect II-A".

v) Select "FTP" on the button bar, then "Connect".

vi) Change host name to barruc and login.

vii) Setting selection.

a) Choose Image data type.

b) Select the MacBinary box.

c) Under Options, choose Binary and unselect "Prompt for every file".

viii) Set the Directory to /home/pinna1/bdata/wpafbh

ix) Select all of the files and click "Copy".

NOTE : The disk sometimes does not eject like it should so you must shut-down the system and start all over.

C.8 Viewing the Images :

a) Login into any local machine and goto the directory where the image is located.

i) Goto Command Tool window, type :

≫ cabcd

..bdata ≫ cd wpafbh

ii) Transfer elements to the ZOO network, type :

≫ ftp {ZOO machine}

NOTE : Make sure you are in the correct directory before this step.

iii) Send files to the ZOO, type :

ftp> *put* {filename}

NOTE : Make sure you are in the correct directory before this step.

iv) Modify getmamo*.m file for your ZOO directory

a) type pwd to identify the directory, should be similar to :

/tmp_mnt/home/birds0/dbramlag is current directory

b) ftp≫ put getmamo*.m

v) quit out of the ZOO, then exit out of the directory.

vi) Rlogin onto Unicorn or Pegasus.

a) You must type the following :

≫setenv LD_LIBRARY_PATH/usr/openwin/lib

≫setenv DISPLAY {Machine Name}:0

vii) Goto the Console Window and type :

≫xhost +*unicorn* (or pegasus)

viii) Return to the Command Tool Window and type, matlab.

NOTE : Make sure you are in the correct directory before this step.

ix) wpafbh_files.txt is a listing of the image sizes.

x) Type, x=getmamo('AFIT.xxx',row size); to view the image.

Appendix D. Matlab Code

D.1 Difference of Gaussians Code (dog.m)

```
% Bill Polakowski
% This is a batch file to read in 36 mass malignant mamos
% It does a 20,50 dog with a 55 offset fill to id regions
% It keeps the largest 7 ratio regions and the 4 most
% circular regions
% The outputs are the ratios, centers, contrasts, circularity,
% and areas of the regions, and the regions.

%%%%%%%%%%%%%%%%%%%%%%%%  Initialization  %%%%%%%%%%%%%%%%%%%%%%

ratio = zeros(43,10);
roiarea = zeros(43,10);
roicount = 0;
con = zeros(43,10);
circle = zeros(43,10);
correct = zeros(1,43);
rbcount = 1;
rmcount = 1;
fa_regions = 0;
fa_ratio = 0;
fa_7 = 0;
two = zeros(43,1);
lab = zeros(43,50);

%%%%%%%%%%%%%%%%%%%%%%%%  Load DoG Filter  %%%%%%%%%%%%%%%%%%%%%%

p = '../bdata/';
filename=[p 'massfilter'];
fid=fopen(filename, 'r');
[fil,cnt]=fread(fid,[2048,1024],'float');
% Change from 1024 to 1124 for large images
fclose(fid);

%%%%%%%%%%%%%%%%%%%%%%%%  Process 18 Malignant Mammos  %%%%%%%%%

for index = 1:36
index

%%%%%%%%%%%%%%%%%%%%%%%%  Load Image  %%%%%%%%%%%%%%%%%%%%%%%%%%

filename=[p 'n' int2str(index)];
```

```
fid=fopen(filename, 'r');
[x,cnt]=fread(fid,[2048,1024],'ushort');
fclose(fid);
clear fid cnt filename
%%%%%%%%%%%%%%%%%%%%%%%%  Orient Image  %%%%%%%%%%%%%%%%%%%%%%%%

if sum(x(:,1)) < sum(x(:,1024))
   x=fliplr(x);
end            % (if sum...)
%x=x(:,1:1024);  %  Include if using 2048x1124 images
marker = [];
position = 1;
c = ones(1,1024);
mamo = x;

%%%%%%%%%%%%%%%%%%%%%%%%  Background Mask  %%%%%%%%%%%%%%%%%%%%%%

f = (x < 1500);

%%%%%%%%%%%%%%%%%%%%%%%%  Gradient Fill (Offset 55)  %%%%%%%%%%

for i = 1:2048
   col = find(f(i,:) == 1);
   if size(col,2) >= 5
      if col(5) < 56 & size(col) > 969  % Fill background rows
         j = sum(x(i,1:10)) / 10;
         if j < 1500
            j = 1500;
         end
         x(i,:) = c * j;
      end  % (if col(5)...)
      if col(5) > (position - 55)  % Crude tracking of breast edge
         if col(5) < 56            % Fill completely bacground rows
            j = sum(x(i,1:10)) / 10;
            if j < 1500
               j = 1500;
            end
            x(i,:) = c * j;
         else  %  Normal gradient fill
            if col(5) < 61
               col(5) = 61;
            end
            begin = sum(x(i,(col(5)-60):(col(5)-51))) / 10;
            final = sum(x(i,1:10)) / 10;
            slope = (final - begin)/(1079-col(5));
```

```
      count = 0;
      for k = col(5)-55:1024
         x(i,k) = x(i,col(5)-55) + count;
         count = count + slope;
      end  % for k
   end     % (if col(5) < 56)
   position = col(5);
else       %  Disjoint mask region
   if position < 56
      temp = 21;
   elseif position > 964
      temp = 964;
   else
      temp = position;
   end
   %  Breast Edge Tracking
   marker = find(f(i,temp-20:temp+60) == 1);
   s = size(marker,2);
   if s > 0
      marker = marker + temp - 21;
      if s < 5       %  New breast edge position
         marker = marker(s);
      else
         marker = marker(5);
      end
   else
      marker = position;
   end
   col = find(f(i,1:marker - 20) == 1);  % Disjoint mask region
   s = size(col,2);
      if col(s) > marker - 25    % Fill mostly background row
         j = sum(x(i,1:10)) / 10;
         if j < 1500
            j = 1500;
         end
         x(i,:) = c * j;
      else
         s = s - 5;
         if s > 5            %  Fill disjoint region
            if col(1) < 26
            slope = 0;
            col(1) = 21;
            x(i,col(1)-20) = sum(x(i,col(s)+16:col(s)+25)) / 10;
            else
            col(1) = 26;
```

```
                begin = sum(x(i,(col(1)-25):(col(1)-16))) / 10;
                final = sum(x(i,col(s)+16:col(s)+25)) / 10;
                col(1) = 21;
                slope = (final - begin)/(col(s) - col(1) + 40);
                end
                count = 0;
                for k = col(1)-20 : col(s) + 20
                   x(i,k) = x(i,col(1)-20) + count;
                   count = count + slope;
                end      % (for k = ...)
             end        % (if s > 5)
             %  Fill the rest of the masked row
             begin = sum(x(i,(marker-60):(marker-51))) / 10;
             final = sum(x(i,1:10)) / 10;
             slope = (final - begin)/(1079-marker);
             count = 0;
             for k = marker-55:1024
                x(i,k) = x(i,marker-55) + count;
                count = count + slope;
             end      % (for k = marker...)
          end         % (if col(s)...)
       position = marker;
    end               % (if col(5) >...)
 end                  % (if size(col,1) >= 5)
end                   % (for i = 1:2048)
fprintf(1,'fill complete \n');
clear s col position k begin final marker position c slope
%%%%%%%%%%%%%%%%%%%%%%%%  Do the Dog  %%%%%%%%%%%%%%%%%%%%%%%%

m = mean(mean(x));
x = x - m;
x = fil .* (fft2(fftshift(x)));
y = real(fftshift(ifft2(x)));
fprintf(1, 'fft Complete \n');
clear x
%%%%%%%%%%%%%%%%%%%%%%%%  Clear Top & Bottom Artifacts  %%%%%%%%

f(1:100,:) = ones(100,1024);
f(1949:2048,:) = ones(100,1024);
f(101:150,1:50) = ones(50,50);
f(1899:1948,1:50) = ones(50,50);
y=y .* (1 - f);
clear f
%%%%%%%%%%%%%%%%%%%%%%%%  Threshold Dog Image  %%%%%%%%%%%%%%%%%
```

```
m=max(max(y));
y=(y>0.5*m);

%%%%%%%%%%%%%%%%%%%%%%%%  Find Regions  %%%%%%%%%%%%%%%%%%%%%%%%

center = [];
centers = [];
regions = [];
area = find_cluster(y);
m = dilate(y,'dilate') - y;
edge = find_cluster(m);
regions = max(max(edge));
fa_regions = fa_regions + regions;
if regions ~= []
clear m y
%%%%%%%%%%%%%%%%%%%%%%%%  Ratio Test  %%%%%%%%%%%%%%%%%%%%%%%%%%

for i = 1 : regions;
   [edger, edgec] = find(edge == i);
   if min(edgec) == 1 & max(edgec) < 60
      lab(index,i) = 0;
   else
      box = (max(edger) - min(edger)) * (max(edgec) - min(edgec));
      mass_perimeter = size(edger,1);
      ratio(index,i) = box / mass_perimeter;
      if ratio(index,i) < 36 & ratio(index,i) > 3
         center(i,1:2) = [mean(edger) mean(edgec)];
         lab(index,i) = 1;
      elseif ratio(index,i) < 50 & ratio(index,i) > 36
         center(i,1:2) = [mean(edger) mean(edgec)]
         lab(index,i) = 2;
      else
         ratio(index,i) = 0;
      end   % (if ratio...)
   end      % (if min(edgec) == 1)
end         % (for i = 1 : regions)
fa_ratio = fa_ratio + sum(ratio(index,:)~=0);
clear edge edgec edger box
%%%%%%%%%%%%%%%%%%%%%%%%  Index Out Large Masses  %%%%%%%%%%%%%%

for i = 1 : regions
   if lab(index,i) == 2
      fprintf(1, 'Index 2:  Large Mass Detected \n');
      ratio(index,i) = 0;
      center(i,1:2)
```

```
      center(i,1:2) = zeros(1,2);
      two(index) = two(index) + 1;
   end
end
if two(index) ~= 0
   label2_false_alarms = two(index)
end

%%%%%%%%%%%%%%%%%%%%%%%%  Pick Top 7 Ratio Regions  %%%%%%%%%%

if center ~= []
temp = size(find(center(:,1) ~= 0),1);
if temp > 7
   temp = 7;
end
[temp1,position] = sort(ratio(index,:));
temp1 = fliplr(temp1);
all_ratios(index,1:size(temp1,2)) = temp1;
clear ratio
ratio(index,1:temp) = temp1(1:temp);
position = fliplr(position);
for i = 1 : temp
   centers(i,1) = center(position(i),1);
   centers(i,2) = center(position(i),2);
end % (for 1 = 1 : temp)
clear center
top_7_ratios = ratio(index,1:temp)
clear center top_7_ratios
centers = centers(1:temp,:);
centers
fprintf(1,'Region Identification Complete \n');

%%%%%%%%%%%%%%%%%%%%%%%%  Extract the Regions for Tests %%%%%%%%

temp = 0;
hits = size(centers,1);
fa_7 = fa_7 + hits;
for i = 1 : hits
   if centers(i,2) < 70
      roi = mamo((centers(i,1)-69):(centers(i,1)+70) , 1:140);
   elseif centers(i,2) > 954
      roi = mamo((centers(i,1)-69):(centers(i,1)+70) , 955:1024);
   else
      roi = mamo((centers(i,1)-69):(centers(i,1)+70) ,
                 (centers(i,2)-69):(centers(i,2)+70));
```

```
   end

%%%%%%%%%%%%%%%%%%%%%%%%  Histogram-Based Morphing  %%%%%%%%%%%%

   t = roi;
   [a,b]=hgram(t);
   a=flipud(a);
   b=flipud(b);
   j=1;
   count=0;
   position=0;
   while count <=3000
      count = count + b(j);
      j = j + 1;
   end
   t=(t>=a(j));
   t=erode(t,'erode');
   reg = find_cluster(t);
   number = max(max(reg));
   for j = 1 : number
      [r,c] = find(reg == j);
      if min(min(r))<=2 | min(min(c))<=2 | max(max(r))>=139 | max(max(c))>=139
         t = t - (reg==j);
      end
   end
   t=dilate(t,'dilate',2);
   reg = find_cluster(t);
   number = max(max(reg));
   for j = 1 : number
      [r,c] = find(reg == j);
      if size(r,1) < 1000
         t = t - (reg==j);
      end
   end
   t = bwmorph(t,'close',3);

%%%%%%%%%%%%%%%%%%%%%%%%  Morphed Area Test  %%%%%%%%%%%%%%%%%%

   roiarea(index,i) = sum(sum(t));
   if roiarea(index,i) < 1000 | roiarea(index,i) > 3500
      roiarea(index,i) = 0;
   else

%%%%%%%%%%%%%%%%%%%%%%%%  Morphed-Mask/Real ROI Contrast Test  %
```

```
      ave_mass = sum(sum(roi .* t)) / roiarea(index,i);
      ave_roi = mean(mean(roi));
      con(index,i) = ave_mass / ave_roi;
      if con(index,i) <= 1.05 | con(index,i) >= 1.13
         con(index,i) = 0;
      else

%%%%%%%%%%%%%%%%%%%%%%  Morphed Mask Circularity Test  %%%%%%%

         mask = zeros(140,140);
         radius = sqrt(roiarea(index,i) / pi);
         [r,c] = find(t == 1);
         r = round(mean(r));
         c = round(mean(c));
         for m = 1:140
         for n = 1:140
            if norm([r c] - [m n]) <= radius
               mask(m,n) = 1;
            end
         end
         end
         circle(index,i) = sum(sum(mask .* t)) / sum(sum(mask));
         if circle(index,i) <= 0.58
            circle(index,i) = 0;
         else
            temp = 1;
         end   % (if size...)
      end      % (if con...)
   end         % (if area...)
end            % (for i=1:hits(index))

%%%%%%%%%%%%%%%%%%%%%%  Pick Top 4 Circular Regions %%%%%%%%%%

if temp == 1
temp = size(find(circle(index,:) ~= 0),2);
if temp > 4
   temp = 4;
end
[temp1,position] = sort(circle(index,:));
temp1 = fliplr(temp1);
circles(index,1:temp) = temp1(1:temp);
position = fliplr(position);
for i = 1 : temp
   center(i,1) = centers(position(i),1);
   center(i,2) = centers(position(i),2);
```

```
   ratios(index,i) = fliplr(ratio(index,position(i)));
   cons(index,i) = fliplr(con(index,position(i)));
   roiareas(index,i) = fliplr(roiarea(index,position(i)));
end      % (for 1 = 1 : temp)
if temp < 4
   center(temp+1:4,1:2) = ones(4-temp,2);
end
area(1,1) = -1;
for i = 1 : 4
   id(i) = area(center(i,1),center(i,2));
end
area = area==id(1) | area==id(2) | area==id(3) | area==id(4);
area(1,1) = 0;
clear ratio centers
final_ratios = ratios(index,1:temp)
centers = center(1:temp,:);
centers

%%%%%%%%%%%%%%%%%%%%%%%%  ROI Extraction  %%%%%%%%%%%%%%%%%%%%%%

rcount = 1;
for i = 1 : temp
   if centers(i,2) < 70
      roi = mamo((centers(i,1)-69):(centers(i,1)+70) , 1:140);
   elseif centers(i,2) > 954
      roi = mamo((centers(i,1)-69):(centers(i,1)+70) , 955:1024);
   else
      roi = mamo((centers(i,1)-69):(centers(i,1)+70) ,
                 (centers(i,2)-69):(centers(i,2)+70));
   end
   filename = ['roi' int2str(index) '.' int2str(rcount) ];
   rcount = rcount + 1;
   filename = [p filename];
   fid=fopen(filename, 'w+');
   fwrite(fid,roi,'ushort');
   fclose(fid);
end      % (for i = 1 : temp)
else
   correct(index) = 0;
   area = 0;
   centers = [];
end                        % (if temp == 1)
   else
   correct(index) = 0;
   area = 0;
```

```
   centers = [];
end                     % (if center ~= [])
else
   correct(index) = 0;
   area = 0;
   centers = [];
end                     % (if regions ~= [])

[i,j,s] = find(area);
[m,n] = size(area);
area = sparse(i,j,s,m,n);

eval(['save image' int2str(index) ' area ratios centers correct
roiarea roiareas con cons circle circles all_ratios']);
clear s1 s2 s3 s4 s5 area i j
end      % (for index = 1:36)

fa_regions = fa_regions-sum(correct~=0)
fa_ratio = fa_ratio-sum(correct~=0)
fa_top_7 = fa_7-sum(correct~=0)
fa_area = sum(sum(roiarea~=0))-sum(correct~=0)
fa_contrast = sum(sum(con~=0))-sum(correct~=0)
fa_final = sum(sum(circles~=0))-sum(correct~=0)
label2_false_alarms = sum(two)
```

D.2 Laws Features Extraction Code (laws.m)

```
% Feeds all the rois through the Laws filters for the 36 images
% Adds area, circularity, contrast, and ratio features

l = [1 4 6 4 1];
s = [-1 0 2 0 -1];
r = [1 -4 6 -4 1];
e = [-1 -2 0 2 1];
w = [-1 2 0 -2 1];

l5l5 = l' * l;
l5s5 = l' * s;  % Kegelmeyer feature
l5r5 = l' * r;
l5e5 = l' * e;  % Kegelmeyer feature
l5w5 = l' * w;
s5l5 = s' * l;
s5s5 = s' * s;
s5r5 = s' * r;
s5e5 = s' * e;
s5w5 = s' * w;
r5l5 = r' * l;
r5s5 = r' * s;
r5r5 = r' * r;  % Kegelmeyer feature
r5e5 = r' * e;
r5w5 = r' * w;
e5l5 = e' * l;
e5s5 = e' * s;  % Kegelmeyer feature
e5r5 = e' * r;
e5e5 = e' * e;
e5w5 = e' * w;
w5l5 = w' * l;
w5s5 = w' * s;
w5r5 = w' * r;
w5e5 = w' * e;
w5w5 = w' * w;

countm = 1;
countb = 1;

for i = 1 : 36
i
   eval(['load image' int2str(i)])
   s = size(centers,1);
   area = full(area);
```

```
   count = 1;

%%%%%%%%%%%%%%%%%%%%%%%%  Extract ROI Mask  %%%%%%%%%%%%%%%%%%%

   for j = 1 : s
     if centers(j,2) < 70
        roi = area((centers(j,1)-67):(centers(j,1)+68) , 1:136);
        n = area(centers(j,1),centers(j,2));
     elseif centers(j,2) > 954
        roi = area((centers(j,1)-67):(centers(j,1)+68) , 889:1024);
        n = area(centers(j,1),centers(j,2));
     else
        roi = area((centers(j,1)-67):(centers(j,1)+68) ,
                   (centers(j,2)-67):(centers(j,2)+68));
        n = area(centers(j,1),centers(j,2));
     end
     roi = (roi==n);
     roi = dilate(roi,'dilate',2);
     a = sum(sum(roi));

%%%%%%%%%%%%%%%%%%%%%%%%  Load ROI  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%

     filename = ['roi' int2str(i)];
     end
     p='../bdata/';
     filename=[p filename];
     fid=fopen(filename, 'r');
     [x,cnt]=fread(fid,[140,140],'ushort');
     fclose(fid);
     m(countm,1) = sum(sum(abs(conv2(x,l515,'valid')) .* roi))/ a;
     m(countm,2) = sum(sum(abs(conv2(x,l5s5,'valid')) .* roi))/ a;
     m(countm,3) = sum(sum(abs(conv2(x,l5r5,'valid')) .* roi))/ a;
     m(countm,4) = sum(sum(abs(conv2(x,l5e5,'valid')) .* roi))/ a;
     m(countm,5) = sum(sum(abs(conv2(x,l5w5,'valid')) .* roi))/ a;
     m(countm,6) = sum(sum(abs(conv2(x,s515,'valid')) .* roi))/ a;
     m(countm,7) = sum(sum(abs(conv2(x,s5s5,'valid')) .* roi))/ a;
     m(countm,8) = sum(sum(abs(conv2(x,s5r5,'valid')) .* roi))/ a;
     m(countm,9) = sum(sum(abs(conv2(x,s5e5,'valid')) .* roi))/ a;
     m(countm,10) = sum(sum(abs(conv2(x,s5w5,'valid')) .* roi))/ a;
     m(countm,11) = sum(sum(abs(conv2(x,r515,'valid')) .* roi))/ a;
     m(countm,12) = sum(sum(abs(conv2(x,r5s5,'valid')) .* roi))/ a;
     m(countm,13) = sum(sum(abs(conv2(x,r5r5,'valid')) .* roi))/ a;
     m(countm,14) = sum(sum(abs(conv2(x,r5e5,'valid')) .* roi))/ a;
     m(countm,15) = sum(sum(abs(conv2(x,r5w5,'valid')) .* roi))/ a;
     m(countm,16) = sum(sum(abs(conv2(x,e515,'valid')) .* roi))/ a;
```

```
      m(countm,17) = sum(sum(abs(conv2(x,e5s5,'valid')) .* roi))/ a;
      m(countm,18) = sum(sum(abs(conv2(x,e5r5,'valid')) .* roi))/ a;
      m(countm,19) = sum(sum(abs(conv2(x,e5e5,'valid')) .* roi))/ a;
      m(countm,20) = sum(sum(abs(conv2(x,e5w5,'valid')) .* roi))/ a;
      m(countm,21) = sum(sum(abs(conv2(x,w5l5,'valid')) .* roi))/ a;
      m(countm,22) = sum(sum(abs(conv2(x,w5s5,'valid')) .* roi))/ a;
      m(countm,23) = sum(sum(abs(conv2(x,w5r5,'valid')) .* roi))/ a;
      m(countm,24) = sum(sum(abs(conv2(x,w5e5,'valid')) .* roi))/ a;
      m(countm,25) = sum(sum(abs(conv2(x,w5w5,'valid')) .* roi))/ a;
      m(countm,26) = roiareas(i,j);
      m(countm,27) = circles(i,j);
      m(countm,28) = cons(i,j);
      m(countm,29) = ratios(i,j);
      countm = countm + 1;
   end     % (for j=1:s)
end          % (for i=1:36)

lbest=[m(:,23) m(:,13) m(:,7) m(:,9) m(:,17) m(:,2) m(:,20) m(:,4) m(:,26)];
save laws6 lbest -ascii
```

D.3 Imbalanced Training Set Neural Network Code (imb.m)

```
%  Imbalanced Training Set Neural Network Code
%  single hidden layer, sigmoid activation function,
%  single output neural net, training in batch mode,
%  derivative-based feature saliency.
%  [err_c0,err_c1,W1,W2]=seltrn(data,HL,maxepochs,lr,clamp);
%  INPUT:
%    data:  1st col class, remaining cols features,
%           # of row = # of samples
%    HL:  number of desired hidden nodes
%    maxepochs:  maximum number of epochs to train
%    lr:  learning rate
%    clamp:  clamp output > 1-clamp to 1-clamp or  <clamp to clamp
%
%  OUTPUT:
%    err_c0:  error for class 0 for each epoch
%    err_c1:  error for class 1 for each epoch
%    W1:  final weights for input to hidden layer
%    W2:  final weights for hidden layer to output node
%
%This program will train a neural net for an imbalanced training set
%with two classes with a selectable number of hidden nodes and a
%single output node.

function [dzdx,err_c0,err_c1,W1,W2]=seltrn(data,HL,maxepochs,lr,clamp)

rand('seed',sum(100*clock));  %rand seed value

[n,I]=size(data);
I=I-1;

ave=mean(data(:,2:I+1));  % normalize data
dev=std(data(:,2:I+1));

average=ones(n,1) * ave;
sigma=ones(n,1) * dev;
data(:,2:I+1)=(data(:,2:I+1)-average)./sigma;
data=data';

epoch_err_c0 = 1;
while epoch_err_c0 > 0.20

%  initialize weights in the net
```

```
W1=rand(HL,I+1)-0.5;   %[HL by I+1]
W2=rand(1,HL+1)-0.5; %[1 by HL+1]

err_c0=[];
err_c1=[];
epoch=0;
dummy1=1;

while epoch<maxepochs

  %Initialize variables
  mse0=[];
  mse1=[];
  index=randperm(n);
  count0=1;
  count1=1;
  z1_c0=[];
  z1_c1=[];
  z2_c0=[];
  z2_c1=[];
  X_c0=[];
  X_c1=[];
  n0=0;
  n1=0;
  for i=1:n;
    d(i)=data(1,index(i));  %  desired output
    X(:,i)=[data(2:I+1,index(i)); 1]; %feature vector (I+1 by n)

%  compute activation fuctions

    z1(:,i)=1./(1+exp(-W1 * X(:,i)));  %hidden layer (HL by n)
    z2(1,i)=1./(1+exp(-W2 * [z1(:,i);1]));  %output layer (1 by n)

      %  clamp output values

    if z2(1,i)>(1-clamp)
      z2(1,i)=1-clamp;
    elseif z2(1,i)<clamp
      z2(1,i)=clamp;
    else
      z2(1,i)=z2(1,i);
    end

%  divide input, hidden and output layer results by class
```

```
    if d(i)==1
      X_c1=[X_c1 X(:,i)];
      z1_c1=[z1_c1 z1(:,i)];
      z2_c1=[z2_c1 z2(1,i)];
      n1=n1+1;
    else
      X_c0=[X_c0 X(:,i)];
      z1_c0=[z1_c0 z1(:,i)];
      z2_c0=[z2_c0 z2(1,i)];
      n0=n0+1;
    end
  end;  %  all train samples through the net

%  find first derivative of hidden and output layers

  dz1_c0=z1_c0.*(1-z1_c0);  %  derivative of hidden layer (HL by n0)
  dz2_c0=z2_c0.*(1-z2_c0);  %  derivative of output layer (1 by n0)
  dz1_c1=z1_c1.*(1-z1_c1);  %  derivative of hidden layer (HL by n1)
  dz2_c1=z2_c1.*(1-z2_c1);  %  derivative of output layer (1 by n1)
  dout_c0=dz2_c0 .* (clamp-z2_c0);  %  (1 by n0)
  temp_c0=W2' * dout_c0;  %  (HL+1 by n0)
  dhl_c0 = dz1_c0 .* temp_c0(1:HL,:);  %  (HL by n0)
  dout_c1=dz2_c1 .* (1-clamp-z2_c1);  %  (1 by n1)
  temp_c1=W2' * dout_c1;  %  (HL+1 by n1)
  dhl_c1 = dz1_c1 .* temp_c1(1:HL,:);  %  (HL by n1)

%  calculate gradients for each class

  GE_W1_c0=dhl_c0 * X_c0';
  GE_W2_c0=dout_c0 * [z1_c0;ones(1,n0)]';
  GE_W1_c1=dhl_c1 * X_c1';
  GE_W2_c1=dout_c1 * [z1_c1;ones(1,n1)]';

%  find unit vectors for each gradient

  unit_GE_W1_c0=GE_W1_c0/sqrt(sum(sum(GE_W1_c0.^2)));
  unit_GE_W1_c1=GE_W1_c1/sqrt(sum(sum(GE_W1_c1.^2)));
  unit_GE_W2_c0=GE_W2_c0/sqrt(sum(GE_W2_c0.^2));
  unit_GE_W2_c1=GE_W2_c1/sqrt(sum(GE_W2_c1.^2));

%  find bisecting angle between the class GE vectors

  ang_GE_W1=(unit_GE_W1_c0 + unit_GE_W1_c1)/2;
  ang_GE_W2=(unit_GE_W2_c0 + unit_GE_W2_c1)/2;
```

```
%  calculate magnitude of GE vectors

  mag_GE_W1=sqrt(sum(sum((GE_W1_c0 + GE_W1_c1).^2)));
  mag_GE_W2=sqrt(sum((GE_W2_c0 + GE_W2_c1).^2));

%  create new GE vectors

  GE_W1=mag_GE_W1*ang_GE_W1;
  GE_W2=mag_GE_W2*ang_GE_W2;

%  update weights with new backprop

  W1=W1+lr*GE_W1;
  W2=W2+lr*GE_W2;

%  calculate the mse for each class

  for i=1:n
    if d(i)==0
      mse0(count0)=(d(i)-z2(i))^2;
      count0=count0+1;
    else
      mse1(count1)=(d(i)-z2(i))^2;
      count1=count1+1;
    end
  end

% compute epoch error for each class

  epoch_err_c0=mean(mse0);
  epoch_err_c1=mean(mse1);
  if epoch_err_c0 >= 0.20
    break
  end
  err_c0=[err_c0 epoch_err_c0];
  err_c1=[err_c1 epoch_err_c1];
  epoch=epoch+1;
fprintf(1,' Epoch %d ... ',epoch);
fprintf(1, 'Average mse = %6.3f     %6.3f\n', epoch_err_c0, epoch_err_c1);
end % (while epoch<maxepochs)
end % (while epoch_err_c0 > 0.20)

%  Feature Saliency

dzdx=zeros(1,I);
```

```
for i=1:n

  z1 = 1 ./ (1 + exp(-W1 * X(:,i)));
  z2 = 1 ./ (1 + exp(-W2 * [z1; 1]));
  fprime1 = z1 .* (1-z1);
  fprime2 = z2 .* (1-z2);

  %  dzdx contains each feature's saliency for all training samples

  dzdx1 = abs((W1(:,1:I)' * (((W2(:,1:HL)' * fprime2) .* fprime1)))');
  dzdx = dzdx + dzdx1;

end  %  (for i=1:n)

dzdx=dzdx/max(dzdx);
dzdx

% Testing

confusion = zeros(2,2);
for i=1:n

  z1 = 1 ./ (1 + exp(-W1 * X(:,i)));
  z2 = 1 ./ (1 + exp(-W2 * [z1; 1]));
if z2>=0.5
  guess = 2;
else
  guess = 1;
end
  d=data(1,index(i));
  confusion(d+1, guess) = confusion(d+1, guess) + 1;

end  % (for i = 1:n)

confusion
classify=trace(confusion)/n;
classify
```

Bibliography

1. S.-M. Lai, X. Li, and W. Bischof, "On techniques for detecting circumscribed masses in mammograms," *IEEE Transactions on Medical Imaging*, vol. 8, pp. 377–386, Dec. 1989.
2. K. Fukunaga, *Introduction to Statistical Pattern Recognition.* Academic Press, second ed., 1990.
3. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis.* New York: John Wiley and Sons, 1973.
4. S. K. Rogers, D. W. Ruck, M. K. G. L. Tarr, and M. E. Oxley, "Artificial neural networks for automatic object recognition," *SPIE Institute Series on Automatic Oject Recognition*, pp. 231–243, Apr. 1990.
5. R. Anand, K. G. Mehrotta, C. K. Mohan, and S. Ranka, "An improved algorithm for neural network classification of imbalanced training sets," *IEEE Transactions on Neural Networks*, vol. 4, pp. 962–969, Nov 1993.
6. M. L. Giger, "Computer-aided diagnosis," *RSNA Categorical Course in Physics*, pp. 283–298, 1993.
7. S. K. Rogers and M. Kabrisky, *An Introduction to Biological and Artificial Neural Networks for Pattern Recognition.* SPIE, 1991.
8. S. K. Rogers, D. W. Ruck, and M. Kabrisky, "Artificial neural networks for early detection and diagnosis of cancer," *Cancer Letters*, vol. 77, pp. 79–83, Mar. 1994.
9. C. M. Kocur, S. K. Rogers, L. Myers, T. Burns, J. Hoffmeister, Bauer, and J. M. Steppe, "Neural network feature selection for breast cancer diagnosis," *IEEE Transactions on Engineering in Medicine and Biology*, Accepted Mar 1995 to appear in early 1996.
10. A. P. Dhawan, Y. Chitre, and M. Moskowitz, "Artificial neural network based classification of mammographic microcalcifications using image structure features," *SPIE*, vol. 1905, pp. 820–831, 1993.
11. J. H. Tanne, "Everything you need to know about breast cancer...but were afraid to ask," *New York [GNYC]*, vol. 26, pp. 52–62, Oct. 1993.
12. R. A. Smith, "Epidemiology of breast cancer," *RSNA Categorical Course in Physics:*, pp. 21–33, 1994.
13. "Breast cancer: New perspectives can replace unrealistic fears," Tech. Rep. ISSN 0741-6254, Mayo Foundation for Medical Education and Research, Rochester, MN, Oct. 1994.
14. E. S. de Paredas, "Radiographic breast anatomy: Radiologic signs of breast cancer," *RSNA Categorical Course in Physics:*, pp. 35–46, 1994.
15. E. Silverberg and J. Lubera, "Cancer statistics," *Cancer*, vol. 39, 1987.
16. "Model-driven automatic target recognition," tech. rep., Wright Laboratory, Wright-Patterson AFB OH, Oct. 1994.

17. R. J. Schalkoff, *Digital Image Processing and Computer Vision.* John Wiley and Sons Inc, 1989.

18. A. Huertas and G. Medioni, "Detection of intensity changes with subpixel accuracy using laplacian-gaussian masks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 5, pp. 651–664, 1986.

19. C. A. Swann, D. B. Kopans, F. C. Koerner, K. A. McCarthy, G. White, and D. A. Hall, "The halo sign and malignant breast lesions," *American Journal of Roentgenology*, vol. 149, pp. 1145–1147, Dec 1987.

20. Y. Wu, M. Giger, K. Doi, C. Metz, C. Vyborny, and R. Schmidt, "Artificial neural networks in mammography: Application to decision making in the diagnosis of breast cancer," *Radiology*, vol. 187, pp. 955–963, Sept. 1993.

21. A. F. Laine, S. Schuler, J. Fan, and W. Huda, "Mammographic feature enhancement by multiscale analysis," *IEEE Transactions on Medical Imaging*, vol. 13, pp. 725–740, Dec. 1994.

22. H. Yoshida, K. Doi, and R. M. Nishikawa, "Automated detection of clustered microcalcifications in digital mammograms using wavelet transform techniques," *SPIE Image Processing*, vol. 2167, pp. 868–886, 1994.

23. T. Chang and C. J. Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Transactions on Image Processing*, vol. 2, pp. 429–441, Oct. 1993.

24. A. P. Dhawan, G. Buelloni, and R. Gordon, "Enhancement of mammographic features by optimal adaptive neighborhood image processing," *IEEE Transactions on Medical Imaging*, vol. MI-5, no. 1, pp. 8–15, 1986.

25. A. P. Dhawan and E. L. Royer, "Mammographic feature enhancement by computerized image processing," *Computer Methods and Programs in Biomedicine*, vol. 27, pp. 23–35, 1988.

26. W. Spiesberger, "Mammogram inspection by computer," *IEEE Transactions on Biomedical Engineering*, vol. 26, pp. 213–219, 1979.

27. Y. Chitre, A. P. Dhawan, and M. Moskowitz, "Artificial neural network based classification of mammographic microcalcifications using image structure features," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 7, no. 6, pp. 1377–1401, 1993.

28. Y. Chitre, A. P. Dhawan, and M. Moskowitz, "Artificial neural network based classification of mammographic microcalcifications using image structure features," *IEEE Engineering in Medicine and Biology*, vol. 15, pp. 50–51, 1993.

29. A. P. Dhawan, Y. Chitre, M. Moskowitz, and E. Gruenstein, "Classification of mammographic microcalcification and structural features using an artificial neural network," *IEEE Engineering in Medicine and Biology*, vol. 13, no. 3, pp. 1105–1106, 1991.

30. A. P. Dhawan, "Computerized mammographic image analysis for reducing false positive rate for biopsy recommendation," *SPIE*, vol. 1905, pp. 540–541, 1993.

31. C. M. Kocur, "Computer aided breast cancer diagnosis," Master's thesis, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, 1994.

32. R. C. Dauk, "Computer-aided detection of microcalcifications in breast tissue," Master's thesis, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, 1995.

33. D. A. McCandless, "Detection of clustered micro-calcifications using wavelets," Master's thesis, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, 1995.

34. D. Brzakovic, X. M. Luo, and P. Brzakovic, "An approach to automated detection of tumors in mammograms," *IEEE Transactions on Medical Imaging*, vol. 9, pp. 233–241, Sept. 1990.

35. W. Kegelmeyer, "Computer detection of stellate lesions in mammograms," *Proceedings of the SPIE*, vol. 1660, pp. 445–453, 1992.

36. W. Kegelmeyer, "Evaluation of stellate lesion detection in a standard mammogram data set," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 7, pp. 1477–1492, Dec 1993.

37. W. Kegelmeyer, J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs, and M. L. Nipper, "Computer-aided mammographic screening for spiculated lesions," *Radiology*, vol. 191, pp. 331–337, May 1994.

38. D. Wei, H.-P. Chan, M. Helvie, B. Sahiner, N. Petrick, D. Adler, and M. Goodsitt, "Classification of mass and normal breast tissue on digital mammograms: Multiresolution texture analysis," *Medical Physics*, to appear in 1995.

39. M. L. Giger, F.-F. Yin, K. Doi, C. E. Metz, R. A. Schmidt, and C. J. Vyborny, "Investigation of methods for the computerized detection and analysis of mammographic masses," *SPIE*, vol. 1233, pp. 183–184, 1990.

40. F.-F. Yin, M. L. Giger, K. Doi, C. E. Metz, R. A. Schmidt, and C. J. Vyborny, "Computerized detection of masses in digital mammograms: Analysis of bilateral-subtraction images," *Medical Physics*, vol. 18, pp. 473–481, Sep 1991.

41. F.-F. Yin, M. L. Giger, K. Doi, R. A. Schmidt, and C. J. Vyborny, "Comparison of bilateral-subtraction and single image processing techniques in the computerized detection of mammographic masses," *Investigative Radiology*, vol. 28, pp. 473–481, Jun 1993.

42. F.-F. Yin, M. L. Giger, K. Doi, R. A. Schmidt, and C. J. Vyborny, "Computerized detection of masses in digital mammograms: Investigation of feature analysis techniques," *Journal of Digital Imaging*, vol. 7, no. 1, pp. 18–26, 1994.

43. T. Parsons, *Voice and Speech Processing.* McGraw-Hill Book Co, 1987.

44. J. M. Steppe, *Feature and Model Selection in Feedforward Neural Networks.* PhD thesis, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, 1994.

45. J. M. Steppe, K. W. B. Jr, and S. K. Rogers, "Integrated feature and architecture selection," *IEEE Interactions on Neural Networks*, accepted in Mar 1995 (to appear).

46. C. Lee and D. Langrebe, "Decision boundary feature extraction for nonparametric classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, pp. 433–444, Mar. 1993.

47. C. Lee and D. Langrebe, "Feature extraction based on decision boundaries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 388–400, Apr. 1993.

48. D. W. Ruck, *Characterization of Multilayer Perceptrons and their Application to Multisensor Automatic Target Detection.* PhD thesis, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, 1994.

49. D. W. Ruck, S. K. Rogers, and M. Kabrisky, "Feature selection using a multilayer perceptron," *Journal of Neural Network Computing*, vol. 2, pp. 40–48, Oct 1990.

50. D. W. Ruck, S. K. Rogers, M. Kabrisky, M. Oxley, and B. Suter, "The multilayer perceptron as an approximation to the bayes optimal discriminant function," *IEEE Transaction on Neural Networks*, vol. 1, pp. 296–298, Dec. 1990.

51. R. M. Nishikawa, M. Giger, K. Doi, C. Metz, F.-F. Yin, C. Vyborny, and R. Schmidt, "Effect of case selection on the performance of computer-aided detection schemes," *Medical Physics*, vol. 21, pp. 265–269, Feb. 1994.

52. B. Jahne, *Digital Image Processing Concepts, Algorithms and Scientific Applications.* Springer-Verlag, 1991.

53. D. Mar and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London*, vol. 207, pp. 187–217, 1980.

54. D. Marr, *Vision.* Freeman, 1982.

55. J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics.* John Wiley and Sons, 1978.

56. P. Miller and S. Astley, "Classification of breast tissue by texture analysis," *Image Vision Computing*, vol. 10, pp. 277–282, June 1992.

57. C. E. Metz, "Roc methodology in radiologic imaging," *Investigative Radiology*, vol. 21, pp. 720–733, Sept. 1986.

Vita

Capt Polakowski obtained a BSEE from the University of Arizona and a commission as a 2nd Lt in the United States Air Force in December of 1989. From October, 1990 through May 1994 he was an Electro-Optical Countermeasures Program Manager in the Electronic Warfare Division, Avionics Directorate, Wright Laboratory, Wright-Patterson AFB (WPAFB), OH. Currently, he has been assigned to the Air Force Institute of Technology (AFIT), WPAFB, OH. At AFIT he has been pursuing an MS in Electrical Engineering, specializing in electro-optical pattern recognition. As of January, 1996, he can be reached at the Air Force Information Warfare Center, Kelly AFB, TX.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1995	Master's Thesis

4. TITLE AND SUBTITLE
COMPUTER-AIDED DIAGNOSIS OF MAMMOGRAPHIC MASSES

5. FUNDING NUMBERS

6. AUTHOR(S)
William E. Polakowski
Captain, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology, WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GEO/ENG/95D-02

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
NA

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; Distribution Unlimited

12b. DISTRIBUTION CODE
A

13. ABSTRACT *(Maximum 200 words)*

A new Model-Based Vision algorithm was developed to find possibly cancerous regions of interest (ROIs) in digitized mammograms and to correctly identify the malignant masses. This work has shown a sensitivity of 92 percent for locating malignant ROIs. The database contained 272 images (12 bit, 100μm) with 36 malignant and 53 benign mass images. Of the 53 biopsied benign cases, 74 percent were correctly classified. The Focus of Attention (segmentation) Module algorithm used a physiologically motivated Difference of Gaussians (DoG) filter to highlight mass-like regions in the mammogram. The Index Module labeled the regions by their hypothesized class: large or medium mass. Then it used size, shape, and contrast tests to reduce the number of non-malignant regions from 8.4 to 2.8 per image. Size, shape, contrast, and Laws texture features were used to develop the Prediction Module's mass model. Statistical and derivative-based feature saliency techniques were used to determine the best features. Nine features were chosen to define the model. Using this model, the Matching Module classified the regions using a multilayer perceptron neural network architecture trained with an imbalanced training set weight update algorithm to obtain an overall classification accuracy of 100 percent for the segmented malignant masses with a false-positive rate of 1.8/image.

14. SUBJECT TERMS
Pattern Recognition, Breast Cancer, Medical Imaging, Neural Networks, Model Based Vision, Difference of Gaussians

15. NUMBER OF PAGES
113

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL