Malignancy Detection Performance Using Excised Breast Tumor Margin Spectroscopic Data and an Optimal Decision Fusion Based Approach

by

Sarah Magdy Oraby

Department of Electrical & Computer Engineering
Duke University

Date:______________
Approved:

Loren Nolte, Supervisor

Gary Ybarra

Nirmala Ramanujam

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Electrical & Computer Engineering in the Graduate School of Duke University

2010
ABSTRACT

Malignancy Detection Performance Using Excised Breast Tumor Margin Spectroscopic Data and an Optimal Decision Fusion Based Approach

by

Sarah Oraby

Department of Electrical & Computer Engineering
Duke University

Date:_______________________
Approved:

___________________________
Dr. Loren Nolte, Supervisor

___________________________
Dr. Nirmala Ramanujam

___________________________
Dr. Gary Ybarra

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Electrical and Computer Engineering in the Graduate School of Duke University

2010
Abstract

Approximately 20-70% of women with breast cancer who choose to undergo breast-conserving surgery (BCS) need to return to the operating table for re-excision [2]. Now devices utilizing optical spectroscopy are emerging as a new platform for intraoperative tumor margin assessment. This study aims to evaluate an optimal decision fusion approach for malignancy detection of measured spectroscopic data from a first-generation optical visible spectral imaging platform that can image the molecular composition of breast tumor margins by implementing a Monte Carlo method with measured diffuse reflectance [2, 3]. The device measures the diffuse reflectance across 450-600 nm. After implementing the Monte Carlo algorithm the absorption and scattering spectra is derived and is used to provide insight on different optical properties present in the tissue mass [3]. Although the extracted optical properties may provide insight on the biological composition of a specimen, it may not be ideal for malignancy detection. Demographic factors may also affect a women’s normal tissue breast composition, which makes malignancy detection more complicated. This optimal decision fusion approach implements the basic decision fusion methodology on acquired spectroscopic data to evaluate the effect on malignancy detection for different extracted optical parameters. The results of this automated and systematic approach indicate that a performance of 90% sensitivity and 68% specificity can be achieved with this approach for the diffuse reflectance spectrum, which outperforms the extracted optical properties. However, when only considering post-menopausal patients, the absorption spectrum can yield a sensitivity of 90% and specificity of 82% and has the best performance of all other features for this demographic group.
Dedication

I dedicate this thesis and the completion of my Master’s at Duke University to my husband, my family, my professors, my fellowship advisor and fellow peers. Their unyielding support, knowledge, and love played a major role in pushing me through this enormous feat.
Content

Abstract ................................................................................................................................. iv
List of Tables ....................................................................................................................... ix
List of Figures .................................................................................................................... xi
Acknowledgements ........................................................................................................ xiv
1 Introduction ................................................................................................................. 1
   1.1 Background ............................................................................................................. 1
      1.1.1 Breast Cancer Disease .................................................................................... 2
      1.1.2 Breast Cancer Detection ............................................................................... 3
      1.1.3 Breast Cancer Treatment .............................................................................. 4
   1.2 Diffuse Reflectance Spectroscopy ......................................................................... 5
   1.3 Motivation ............................................................................................................. 7
2 Methods ....................................................................................................................... 7
   2.1 Data ..................................................................................................................... 7
      2.1.1 Tissue Optical Properties .............................................................................. 8
   2.2 Decision Fusion .................................................................................................... 9
      2.2.1 Likelihood Ratio ............................................................................................. 11
      2.2.2 Fusing Feature Level Binary Decisions ......................................................... 12
      2.2.3 Detection Performance Metrics .................................................................. 13
   2.3 The Optimal Decision Fusion Approach ............................................................. 13
   2.4 Determining Feature Significance ...................................................................... 15
   2.5 Data Evaluation ................................................................................................. 16
      2.4.1 Spectral Data ................................................................................................. 16
      2.4.2 The Effect of Tissue Type on Specificity ....................................................... 17
2.4.3 Decision Fusion Validation with Data Subsets ................................................. 17

3 Results & Discussion ............................................................................................. 17

3.1 Extracted Optical Properties ............................................................................. 19

3.1.1 Stage One: Individual Wavelength Classifier ................................................. 20

3.1.2 Stage Two: Pair-Wise Decision Fusion ......................................................... 21

3.1.3 Stage Three: Group-Wise Decision Fusion .................................................... 24

3.1.4 Demographic Groups ..................................................................................... 25

3.1.4.1 Post-Menopausal ....................................................................................... 25

3.1.4.2 Chemotherapy Naïve ............................................................................... 29

3.1.5 Tissue Type Analysis ..................................................................................... 33

3.1.6 Summary ....................................................................................................... 35

3.2 Measured Spectral Data ..................................................................................... 35

3.2.1 Stage One: Individual Wavelength Classifier ................................................. 36

3.3 Absorption ........................................................................................................ 36

3.3.1 Stage One: Individual Wavelength Classifier Performance ................................ 37

3.3.2 Stage Two: Pair-Wise Decision Fusion Performance ...................................... 38

3.3.3 Stage Three: Group-Wise Decision Fusion ..................................................... 41

3.3.4 Demographic Stratification ........................................................................... 42

3.3.4.1 Post-Menopausal ....................................................................................... 43

3.3.4.2 Chemotherapy Naïve ............................................................................... 46

3.3.4.3 Summary .................................................................................................... 47

3.3.5 Tissue Type Analysis ..................................................................................... 48

3.4 Scattering ........................................................................................................... 49

3.4.1 Stage One: Individual Wavelength Classifier Performance .......................... 49

3.4.2 Stage Two: Pair-Wise Decision Fusion ......................................................... 50
3.4.3 Stage Three: Group-Wise Decision Fusion ............................................................ 52
3.4.4 Demographic Status ............................................................................................. 53
  3.4.4.1 Post-Menopausal .......................................................................................... 54
  3.4.4.2 Chemotherapy Naïve .................................................................................. 56
3.4.5 Tissue Type Analysis ........................................................................................... 58
3.5 Diffuse Reflectance ................................................................................................. 59
  3.5.1 Stage One: Individual Wavelength Classifier Performance............................ 59
  3.5.2 Stage Two: Pair-Wise Decision Fusion ............................................................. 60
  3.5.3 Stage Three: Group-Wise Decision Fusion ....................................................... 62
  3.5.4 Demographic Status ......................................................................................... 64
    3.5.4.1 Post-Menopausal ...................................................................................... 65
    3.5.4.2 Chemotherapy Naïve ................................................................................ 67
4 Conclusions ................................................................................................................. 71
  4.1 Feature Selection .................................................................................................... 71
Appendix A: Threshold Value Assessment ...................................................................... 73
Appendix B: Validation Box Plots .................................................................................. 76
Appendix C: Stage Three ROC Curves ........................................................................... 86
References ...................................................................................................................... 89
List of Tables

Table 1: Descriptions of the different data subsets used in this study ................................ 16
Table 2: AUC achieved for each pair of optical properties .................................................. 23
Table 3: The following table lists the achieved sensitivity and specificity by the decision fusion of certain pairs of extracted optical properties .................................................. 23
Table 4: The following table lists the AUC achieved for the decision fusion of all possible pairs of extracted optical properties for the post-menopausal group ........................................... 26
Table 5: Sensitivity and Specificity from the group-wise decision fusion curve for the post-menopausal data set .................................................................................................................. 28
Table 6: The following table lists the achieved AUC for each pair of extracted optical properties for the chemo-naive data set ........................................................................................................... 31
Table 7: The Sensitivity and Specificity achieved from the group-wise decision fusion classifier of the chemotherapy naive data set ........................................................................................................... 32
Table 8: Summary of the performance for each tissue type decision fusion curve ........... 34
Table 9: The highest AUC and corresponding wavelength pairs from the absorption spectrum .................................................................................................................................................. 40
Table 10: The following table lists the wavelengths with the highest average AUC in descending order .................................................................................................................................................. 40
Table 11: Summary of the ROC performance achieved from fusing all ten highest of the best performing wavelengths from stage I ........................................................................................................... 42
Table 12: The ROC performance achieved from fusing the five best performing wavelengths from stage I .................................................................................................................................................. 42
Table 13: The sensitivity and specificity for absorption spectrum stratified by post-menopausal sites for 10 wavelength fusion ........................................................................................................... 44
Table 14: The AUC and its corresponding wavelength pair .................................................. 51
Table 15: The following table lists the wavelengths with the highest average AUC in descending order .................................................................................................................................................. 52
Table 16: Summary of the ROC performance achieved from fusing the ten (left) and the upper 5 (right) best performing wavelengths from stage I for the Scattering Spectrum ................................................ 52
Table 17: Top Performing Pairs from Stage Two of the Scattering Post-menopausal data set .................................................................................................................................................. 55
Table 18: Wavelength Pairs with the highest AUC from the Chemotherapy-Naive Scattering Spectral data

Table 19: The AUC of the highest performing wavelength pairs

Table 20: The following table lists the wavelengths with the highest average AUC in descending order

Table 21: Summary of the ROC performance achieved from fusing the ten (left) and the upper 5 (right) best performing wavelengths from stage I for the Scattering Spectrum

Table 22: Highest AUC for wavelength pairs of diffuse reflectance post-menopausal data set

Table 23: Detection sensitivity and specificity for the post-menopausal data set for Diffuse Reflectance for 10 wavelengths

Table 0-1: The following Table lists the threshold for the selected classifier chosen in Stage I for each extracted optical property. The measured value will be classified as malignant if it meets the criteria described by the inequality
List of Figures

Figure 1: Image adapted from American Breast Cancer Society ........................................... 3

Figure 2: Optimized Detector Selection from [25]................................................................. 10

Figure 3: The Decision Fusion Scheme adapted from [1] where Z is the feature, Y1 is operating point selected from the ROC curve of Z and X is the binary decision. D is then the final fused ROC curve....................................................................................................... 14

Figure 4: Group-Wise Decision Fusion approach adapted from [1].................................... 15

Figure 5: This figure illustrates the AUC achieved for the entire data set, the post-menopausal data set and the chemotherapy naïve data set. .................................................. 21

Figure 6: A graphical representation of the AUC achieved from implementing decision fusion on each pair of extracted optical properties..................................................... 22

Figure 7: The ROC curve achieved from implementing decision fusion with all seven extracted optical properties ................................................................................................... 24

Figure 8: The AUC for each pair of extracted optical properties ........................................ 26

Figure 9: The group-wise ROC decision fusion curve for all seven extracted optical properties for the post-menopausal data set................................................................. 28

Figure 10: Pair-wise performance for the decision fusion of each pair of extracted optical properties for the chemo-naïve group ............................................................................. 30

Figure 11: The group-wise decision fusion ROC curve for the chemotherapy naive data set............................................................................................................................... 32

Figure 12: The ROC curves attained from the decision fusion of all extracted optical properties excluding specific tissue types. .............................................................................. 34

Figure 13: The following figure illustrates the AUC achieved by the assigned classifier determined by each individual wavelength by either the measured Diffuse Reflectance spectrum, Scattering spectrum or Absorption spectrum. .................................................. 36

Figure 14: The figure above illustrates the AUC achieved by the assigned classifier at each wavelength for the measured absorption. The lower figure illustrates the difference in the average measured absorption at each wavelength between normal and malignant tissue. ............................................................................................................. 38

Figure 15: This figure visualizes the AUC after decision fusion with every combination of wavelength pairs. ........................................................................................................ 39
Figure 16: This figure presents the absorption fusion curves plotted against the fused extracted optical properties curve. ................................................................. 41

Figure 17: The AUC achieved at each wavelength of the Absorption spectrum .......... 43

Figure 18: The AUC resulting from the decision fusion of all combinations of wavelength pairs from the absorption spectrum stratified by post-menopausal status ............... 44

Figure 19: ROC curves of the group-wise decision fusion classifier for both the 10 and 5 wavelengths plotted with the ROC curve for the extracted optical properties. All of which represent the post-menopausal data set. ................................................................. 45

Figure 20: The AUC resulting from the decision fusion of all combinations of two wavelength pairs from the absorption spectrum stratified by chemotherapy. .......... 46

Figure 21: The stage-three ROC curve performance for the absorption spectrum plotted with the stage three ROC curve for the extracted optical properties. ......................... 47

Figure 22: Stage three ROC curves for the absorption data excluding specific tissue types. ....................................................................................................................... 48

Figure 23: This figure illustrates the AUC achieved by the assigned classifier at each wavelength for the measured scattering. The lower figure illustrates the difference in the average measured absorption at each wavelength between normal and malignant tissue. ................................................................. 50

Figure 24: This figure visualizes the achieved AUC for each wavelength pair from the scattering spectrum ............................................................................................................................................. 51

Figure 25: This figure compares the 10, 5 and extracted optical property group-wise fusion ROC curves. ................................................................................................................. 53

Figure 26: The AUC achieved for each wavelength for the data sets stratified by post-menopausal sites, chemotherapy naïve sites and the entire data set. ................................. 54

Figure 27: The AUC resulting from the decision fusion of all combinations of two wavelength pairs from the scattering spectrum stratified by post-menopausal status.... 55

Figure 28: The following figure is a plot of the group-wise decision fusion curves for both the extracted optical properties and scattering spectrum stratified by post-menopausal status ................................................................................................................. 56

Figure 29: The AUC resulting from the decision fusion of all combinations of two wavelength pairs from the scattering spectrum stratified by chemotherapy naïve. ....... 57

Figure 30: Significant Wavelengths for Scattering Stratified by chemo-naïve ............ 58
Figure 31: ROC curve generated by excluding fibro-adipose, fibro-glandular, and adipose tissue types plotted with the ROC curve generated from including all tissue types.

Figure 32: Summary of the AUC for each wavelength independently for the diffuse reflectance (top). The bottom figure illustrates the average measured diffuse reflectance for the normal tissue and malignant tissue.

Figure 33: A graphical representation of the pair-wise decision fusion for all pairs of the Diffuse Reflectance spectra.

Figure 34: This figure visualizes the achieved AUC after decision fusion with 2 wavelengths. This figure is a collection of every combination of wavelength pairs.

Figure 35: The figure illustrates the AUC for each wavelength evaluated individually for the entire data set, the data set stratified by post-menopausal patients, and the data set stratified by patients who have not undergone chemotherapy.

Figure 36: The AUC resulting from the decision fusion of all combinations of two wavelength pairs from the diffuse reflectance spectrum stratified by post-menopausal status.

Figure 37: The following figure is a plot of the group-wise decision fusion curves for both the extracted optical properties and diffuse reflectance spectrum stratified by post-menopausal status.

Figure 38: The AUC resulting from the decision fusion of all combinations of two wavelength pairs from the diffuse reflectance spectrum stratified by chemotherapy.

Figure 39: Significant Wavelengths for Diffuse Reflectance Stratified by chemo.

Figure 40: The ROC curves for stage three of each of the tissue excluded tissue data sets.

Figure 0-A: The change in threshold values for the classifier of each extracted optical property for each demographic status.

Figure 0-A

Figure 0-B

Figure 0-C

Figure 0-D
Acknowledgements

I would like to acknowledge Dr. Loren Nolte for his continual support as an academic advisor and professor. I would also like to acknowledge Dr. Gay Ybarra for his commitment to my success as a student and women in completing a higher degree in electrical and computer engineering. I also want to acknowledge Dr. Nirmala Ramanujam for her generosity in extending her resources and welcoming me as a member of her team during the span of this research.
1 Introduction

Breast cancer treatment techniques spanning from diagnosis, surgery and radiation therapies continue to become more sophisticated as the disease and its implications on breast tissue are better understood through these advancements. Despite these advancements, once a breast tumor is located, successful removal of the tumor mass is still heavily dependent on biopsies by histopathology in post-screening procedures, requiring approximately 20-70% of women to re-operate because of recurrence [2-4]. A new frontier of medical imaging is emerging to tackle this deficiency in the operating room. Among these imaging modalities is the use of diffuse reflectance spectroscopy for the diagnosis of breast cancer. Diffuse reflectance spectroscopy in the ultraviolet-visible (UV-VIS) wavelength range combined with a Monte Carlo-based inverse model can measure the tissue absorption and scattering, which in turn can be used to extract the intrinsic physiological and structural properties of a tissue [2, 5, 6]. By applying previous knowledge of breast tissue changes due to cancer, tissue composition can be predicted and characterized. My interests in this problem involve building on the diffuse reflectance spectroscopy model by adding a statistical method that can identify specific features in the visible light spectrum that may improve tissue detection and provide insight for the physiological and structural properties of the tissue sample.

1.1 Background

The introduction section of this paper will provide the basic understanding of breast cancer diagnosis and treatment followed by a brief review of current diffuse reflectance spectroscopy methods in the pipeline.

Breast cancer is the development of a malignant tumor, originating in breast tissue. Although mainly affected are women, men may also develop breast cancer.
Research by the Breast Health Global Initiative estimated that more than 1.1 million women worldwide are newly diagnosed with breast cancer a year, and an estimated of all women diagnosed with breast cancer in the last 5 years, 4.4 women have survived [7, 8]. With approximately 192,370 new cases of invasive breast cancer in the U.S. in 2009, breast cancer leads as the most common type of cancer among women in the US [8]. Approximately 40,170 died from breast cancer in 2009, making it the second most deadly form of cancer after lung cancer [9]. Earlier detection and improved treatment options has decreased breast cancer death rates to less than 1 in 35 [8].

1.1.1 Breast Cancer Disease

The female breast tissue is comprised of several types of tissue types including fatty and connective tissue, known as adipose and fibro-Adipose respectively, blood vessels, and lymph vessels, known as fibro-glandular. There are also structures that create and transport breast milk, which include the lobules and ducts. Breast cancer commonly spreads through the lymph system in the breast; these forms of malignant tissue have very similar properties to benign fibro-glandular tissue. Breast cancer can enter a lymphatic vessel, small veins that carry lymph fluid from the breast, and grow inside lymph nodes, small clusters of immune cells connected by lymphatic vessels. Several of the lymph vessels of the breast lead to lymph nodes under the arm, known as axillary nodes. When axillary nodes develop cancer cells they cause the node to swell. Cancer can spread to other regions of the body through these lymph vessels. Benign breast tumors may form and are generally caused by fibrocystic changes [8].
Breast cancer is characterized by its location in the breast. These types of cancer include carcinoma, adenocarcinoma, carcinoma in situ, invasive carcinoma and sarcoma. Identifying these types of cancers can assist in detection, localization and treatment. More specifically breast cancer can come in these forms, ductal carcinoma in situ (DCIS), lobular carcinoma in situ (LCIS), invasive lobular carcinoma (ILC) and inflammatory breast cancer (IBC) [8]. This study focuses primarily on DCIS and IDC.

1.1.2 Breast Cancer Detection

Currently several methods are in place to screen for cancer and identify key factors such as tumor size and location. Among these methods are x-ray mammography, ultrasound and palpation. These screening procedures alone do not accurately identify malignant tissue [10].

X-ray mammography and other regular screening programs from early detection has improved breast cancer mortality [11]. However mammography alone proves insufficient as 60-90% of suspicious lesions detected by mammography are proven benign upon biopsy [10, 11]. This is due to mammography’s low sensitivity and low
specificity, which results in patient trauma, time delay and high medical costs [12]. As a result, mammograms for high-risk patients are often supplemented with MRI screenings. Several prospective studies of women at high risk for the development of breast cancer have recently demonstrated that MRI is more sensitive than mammography for the detection of breast cancer [13].

1.1.3 Breast Cancer Treatment

Treatment for breast cancer can vary greatly from patient to patient. The entire course of breast cancer treatment may include several approaches including surgical removal procedures followed by either chemotherapy and/or radiation therapy [14]. The surgical procedures implemented for breast cancer treatment involve lumpectomy, mastectomy, or removal of axillary lymph nodes. Breast conserving surgery (BCS) is a type of lumpectomy procedure also known as a partial mastectomy.

A procedure that involves BCS followed by radiation therapy is known as Breast Conserving Therapy (BCT). Currently, approximately 50-75% of breast cancer patients receive BCT [15]. Mammography successfully detected ipsilateral breast tumor recurrence, predominantly as calcifications or masses, after breast conserving surgery with radiation therapy for DCIS in 97% of cases [16]. Although BCT and mastectomy have similar long-term survival rates, the importance of a completely excised tumor is emphasized by the 25% chance of mortality for a woman who develops a local recurrence after BCT [8].

Since 1982, BCS has increased from 1% to 25% in 2002. BCS has also replaced TM as the standard approach for localized ductal carcinoma in situ (DCIS) [8]. Successful BCS provides better cosmetic results and better quality of life than TM [15]. Patients diagnosed with breast cancer through a mammographic screening program tend to have early disease with smaller primary cancers making them particularly suited to BCS [8].
Safe breast conservation requires clear histological margins to minimize the risk of local recurrence [15]. Clear margins are not always obtained during initial surgery, necessitating re-excision to achieve adequate margins. BCS is becoming a more preferred alternative to breast cancer treatment and with the main causes for reoperation being due to margin control and lack of on-site pathological inspection, there is a greater need for intra-operative margin assessment tools.

Once removed, the excised tumor mass will be oriented as a 6-sided cube with an anterior, posterior, medial, lateral, inferior and superior margin. This specimen is then sent to pathology for both site and marginal evaluation that determines whether malignant cells are found within a 2mm margin. The marginal evaluation of the specimen will determine whether a patient is advised to have a re-excision. Published studies confirm that the pathologic margin status is the most important factor in determining the risk of recurrence after BCS. [17]. In this study, a positive margin indicates that malignant cells were found within 2mm of the margin and a negative margin if there are no malignant cells within a 2mm of the margin.

Intra-operative breast margin analysis is offered in some hospitals, such as Pathologic touch preparation cytology and frozen section analysis. However, these methods have several limitations, including the need for an on-site trained pathologist during operation [2]. Due to the heterogeneous nature of normal breast tissue, detecting cancerous tissue on margins is very difficult [18].

1.2 Diffuse Reflectance Spectroscopy

The use of optical spectroscopy has shown significant promises in overcoming the limitations of breast cancer screening, particularly for the evaluation of excised tumors during BCS [19]. Several research groups have determined significance between
the absorption and scattering, diffuse reflectance and Raman spectroscopy measured across small sample areas of a specimen margin.

Diffuse reflectance spectroscopy uses primarily light-tissue interactions of absorptions and scattering followed by spectral imaging to characterize differences in tissue composition of excised breast specimen margin [3]. Correctly identifying the absorption and scattering interactions in a tissue samples poses a challenge with diffuse reflectance measurements and as a result there are several analytical and numerical models developed that use fiber-optic based diffuse reflectance measurements from the ultra-violet to visible (UV-VIS) wavelengths to extract the scattering and absorption coefficients [20]. These methods include analytical approximations such as the diffusion equation, a modified form of the Monte Carlo model, and empirical models, such as neural networks [21].

For this study, an inverse Monte Carlo method was implemented with the diffuse reflectance spectroscopy. The margin data for the excised tumor masses were extracted using quantitative diffuse reflectance spectroscopy [3]. The probe for the spectroscopy device sampled 2 x 4 cm sections of the sample with a sensing depth range of 0.5 – 2.2 mm for 450-600 nm. The measured diffuse reflectance spectrum was then analyzed using a Monte Carlo model developed by the Tissue Optical Spectroscopy Lab (TOpS Lab) is directed by Professor Nimmi Ramanujam at Duke University to extract the optical features, which is used to determine the scattering and absorption contrast of the sample [3]. Currently a decision-tree model is applied to differentiate positive from negative margins for a data set of 55 margins from 48 patients yielding a sensitivity and specificity of 80% and 70% respectively optical assessment of tumor resection margins [3].
1.3 Motivation

My motivation stemmed from a business-marketing course that introduced to me the optical spectroscopic tumor margin assessment device. As a result of understanding the devices role in the current market and breast conserving surgery situation from the business perspective, I became very interested in the solution to these problems offered by the medical device. By conducting this research I hope to contribute to overcoming the limitations of intra-operative tumor margin assessment by providing a method that can diagnose and provide further insight on beast tissue composition and the implications on it by breast cancer.

2 Methods

The methodology for this research consists of an inter-disciplinary approach that applies an essentially statistical optimization approach known as decision fusion to a fundamentally biomedical imaging technique known as diffuse reflectance spectroscopy for the detection of malignant tissue on an excised breast tumor.

2.1 Data

The data for this study was collected by an ex-vivo diffuse optical spectroscopy method used to evaluate partial mastectomy specimens in patients undergoing surgery for breast malignancies. The Institutional Review Board at Duke University approved this method and more details about the device and collection are detailed in previous publications [22, 23]. The excised tumor mass data was collected for 104 patients, for each patient anywhere from 6-10 sites along the specimen margin and later diagnosed by histological evaluation by a pathologist, resulting in a total of 854 sites. However, for a variety of reasons described in [2] not all sites were considered for this study, and the final data set retained 595 normal samples and 38 malignant samples. For each patient
and their corresponding sites, different aspects of demographical information were recorded. The recorded demographic information includes extracted optical properties, menopausal status, chemotherapy status, therapy treatment history, etc. Menopausal status was differentiated as post-menopausal if they had either a bilateral salpingoopherectomy or lack of a menstrual cycle for greater than 1 year. A total of 45 optical sites were excluded from menopausal analyses due to an undefined menopausal status. The extracted optical properties for this study are detailed in Table 2. The diffuse reflectance, absorption and scattering data were measured at every 2.5 nm starting at 450 nm and ending at 600 nm, yielding spectra of 61 wavelengths in the visible light spectrum [2].

2.1.1 Tissue Optical Properties

Optical properties of the measured tissue were extracted from the diffuse reflectance spectra using a scalable Monte Carlo inverse model that has been previously published by the ToPS group at Duke University [24, 25]. The concentrations of β-carotene, oxygenated hemoglobin, de-oxygenated hemoglobin, and Lymphazurin were extracted using this model. Other parameters such as hemoglobin saturation and total hemoglobin (THb) were derived from the concentrations of oxygenated and de-oxygenated hemoglobin. The wavelength-dependent reduced scattering coefficient, μ_\text{s'}(λ) (scattering), was extracted from the model and the mean value, <μ_\text{s'}>, was calculated to describe the overall scattering properties of the probed tissue and likewise for absorption coefficient. The spectral data, extracted optical properties and demographic information will be referred to as features. Malignancy classification for all data samples was determined by pathology.
2.2 Decision Fusion

Due to the heterogeneous nature of the spectroscopic data a multi-stage decision fusion approach posed as a prime candidate for this type of classification problem, as it has shown to be effective for heterogeneous data sources over several different classification tasks [26]. Decision theory states that the likelihood ratio is the optimum detector for a binary signal detection problem under the assumption that the true probability density function under each hypothesis is known [27].

The binary hypothesis-testing problem assumes

\[ H_0: \text{Tissue is normal (not malignant)} \]
\[ H_1: \text{Tissue is malignant.} \]

The likelihood ratio can be defined as,

\[
\lambda(X) = \frac{P(X|H_1)}{P(X|H_0)}
\]

(1)

where the probability of the data X given that the tissue is malignant over the probability of the data X given that the tissue is normal. Decision-fusion theory introduces a procedure that combines local binary decisions optimally to determine the presence or absence of a signal [1, 26, 27]. A two-stage approach has been used before for decision fusion on similar applications[26, 28-30]. This approach is based on the two stage approach but incorporates a third stage that uses feedback from stage two to suggest significant features from the data set.

Stage one looks at each feature individually and assigns to it a local detector. In order to identify the optimum detector for stage one a series of threshold values are used.
as classifiers along the entire distribution of the data. Depending on the feature, if the data value lies above or below this threshold it will be classified by a binary decision value of either 1 or 0 (malignant or non-malignant). These different threshold values make up the operating points of the receiving operator characteristics (ROC) curve for that feature. For each threshold value, there is a specific false positive and false negative probability. Each individual feature will produce a unique ROC curve with different area under curve (AUC). The AUC serves as a metric to quantify the performance of the feature as a whole. However, at stage one the optimum detector is selected based on the particular operating point of the ROC curve where the difference between the Pd and Pf values are minimized. By selecting this as the local detector for this feature a sensitivity and specificity can be calculated.

Stage two assigned the optimum local detector to each feature and put every combination of feature detector pairs through the decision fusion processor. As a result of this procedure an optimum detector (based on the likelihood ratio calculated by a

Figure 2: Optimized Detector Selection from [25]
decision fusion processor) for each pair was determined along with a corresponding ROC curve. Each pair will have a unique detector and ROC curve and their significance was determined based on the AUC and corresponding sensitivity and specificity.

Stage three takes the features that performed best at the pair-wise level and runs a group of these features into the decision fusion processor. At this stage the final detector is the decision fusion of the local detectors for all the features involved. As in the previous stages, ROC curves and AUC was used to quantify performance. Stage three is designed to fuse the binary decisions for the features that are determined most significant from stage two to create a final, optimized classifier.

One of the most significant advantages of using decision fusion is that the high dimensionality and complexity of the problem is greatly reduced by considering each feature one at a time. Decision fusion is also known to be robust in heterogeneous data, not overly sensitive to the likelihood ratio threshold values and can handle missing data values[26].

2.2.1 Likelihood Ratio

The decision fusion classifier assumes the likelihood ratio as a testing criterion that tests the likelihood ratio against a given threshold. Where the decision, $u$ is declared given:

$$u = \begin{cases} 
1 \text{ if } \lambda \geq \tau \\
0 \text{ if } \lambda \leq \tau 
\end{cases}$$

(2)

where the decision ($u$) is either 1 (tissue is malignant) or 0 (tissue is not malignant) and $\tau$ is an arbitrary threshold value that spans the range of the distribution.
of data. Decision Theory can be applied to breast cancer detection where the hypothesis is that the null case is that there is no malignancy, and the alternative hypothesis is that malignancy is present. Previous work shows that the likelihood ratio can be an excellent classifier for breast cancer mass lesion data [1].

2.2.2 Fusing Feature Level Binary Decisions

With any of the detectors, the binary decisions were the probability of correctly identifying malignancy, known as $P_d$ where $P_d = P(u = 1 \mid H_1)$ and incorrectly identifying malignancy, $P_f$, where this probability is $P_f = P(u=1 \mid H_0)$. Conversely, there was also a probability of not classifying malignant tissue, $P(u=0 \mid H_1) = 1 - P_d$ or classifying normal tissue as malignant, $1 - P_f = P(u = 0 \mid H_0)$. At stage one the local classifier can be expressed as

$$
\lambda_{\text{decision}} = \begin{cases} 
\frac{P_d}{P_f} & \text{if } u = 1 \\
\frac{1 - P_d}{1 - P_f} & \text{if } u = 0 
\end{cases}
$$

(3)

At stage two these $P_d$ and $P_f$ values from equation 3 are implemented in the decision fusion likelihood ratio where $i$ indicates the features involved in decision fusion, and $u$ is the decision. The cumulative product of the likelihood ratios for each decision are taken for each feature and can be expressed as equation (4).

$$
\lambda_{\text{fused}}(u_1, \ldots, u_p) = \prod_{u=1}^{p} \frac{P_{d_i}}{P_{f_i}} \prod_{u=0}^{p} \frac{1 - P_{d_i}}{1 - P_{f_i}}
$$

(4)
2.2.3 Detection Performance Metrics

 Receivers operating characteristic (ROC) curves were used to assess the detection performance of the classifiers for each stage. At any (operating) point along the ROC there is a corresponding classifier and threshold \(i^{th}\) local decision). The \(P_d\) and \(P_f\) values can be extracted at each operating point to determine sensitivity (\(P_d\)) and specificity (1-\(P_f\)). At stage one the operating points were selected to find a point where the difference between \(P_d\) and \(P_f\) values were minimized, by determining the operating point closest to the coordinate \((0,1)\), this coordinate indicates perfect sensitivity and specificity, on the ROC plot. The selection of the likelihood ratio threshold values in stage two is important to maximize performance of the fused classifier in stage three [26].

 For stage two and three the \(P_d\) and \(P_f\) values extracted from the ROC curve can be expressed as the following:

\[
P_{d_{\text{fusion}}}(j) = \sum_{i=j}^{p} P(\lambda_{f_{\text{usedi}}} | H_1) j = 0,\ldots, p
\]  

(5)

\[
P_{f_{\text{fusion}}}(j) = \sum_{i=j}^{p} P(\lambda_{f_{\text{usedi}}} | H_0) j = 0,\ldots, p
\]

(6)

The threshold value for fusion determines the operating point on the ROC curve. By varying the value of the threshold, these \(P_{d_{\text{fusion}}}\) and \(P_{f_{\text{fusion}}}\) values trace the entire decision-fusion ROC curve [26].

2.3 The Optimal Decision Fusion Approach

 A systematic and automated approach was derived using decision fusion in order to determine and extract the optical spectroscopic features that yield the optimal detection performance.
Stage one involves creating ROC curves based on a simple threshold classifier for each feature (either an individual extracted optical property or wavelength) locally. Once the ROC curves were created, an optimal operating point was selected that provided the best compromise between sensitivity and specificity.

Stage two uses the PD and PF probabilities from the thresholds selected from stage one’s ROC curves for each feature and implements a pair-wise decision fusion to determine the best detector for the pair of interest. Secondary ROC curves were then generated based on the extracted decision fusion probabilities. The AUC from the ROC curves of these pairs were then assessed in order to identify best performance. These metrics involve the AUC and also the extracted Pd and Pf values for the selected operating point from the ROC curve.

![Decision Fusion Scheme](image)

**Figure 3:** The Decision Fusion Scheme adapted from [1] where Z is the feature, Y1 is operating point selected from the ROC curve of Z and X is the binary decision. D is then the final fused ROC curve.
The wavelength pairs selected in stage two fused together in a group-wise decision fusion in stage three. Stage three is essentially equivalent to stage two except the decisions of both 5 and 10 wavelengths were fused together.

2.4 Determining Feature Significance

Given the high emphasis of sensitivity and specificity in the clinical setting, the sensitivity was assessed at three different values, \( P_d = 0.9, 0.8, \) and 0.7 (sensitivity at 90\%, 80\% and 70\%). The ROC curve can provide more inference than just the sensitivity and specificity (1-Pf) of a given classifier. Another metric based on the ROC curve used for this method includes the area under the curve (AUC). In this study, the AUC metric was used to assess feature performance at stage-two, when decision fusion was applied pair-wise across all combinations of two features. The average AUC for a feature and all combinations of that feature with another feature was calculated (equation 7) and used as a metric to assess best feature performance [26].
2.5 Data Evaluation

In order to determine relative feature performance and significance, the data was assessed and broken into different subsets according to spectral information, extracted optical properties or demographic information. The different subsets are described in Table 1.

\[ \sum_{j=450}^{600} E \left( AUC_{j,450:600} \right) \]

(7)

Table 1: Descriptions of the different data subsets used in this study

<table>
<thead>
<tr>
<th>Data Subset</th>
<th>Sample Sizes</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td>Measured and normalized Diffuse Reflectance across visible light spectrum</td>
</tr>
<tr>
<td>B</td>
<td>Non-malignant: 595 Malignant: 38</td>
<td>Absorption Coefficients across visible light spectrum</td>
</tr>
<tr>
<td>C</td>
<td>Non-malignant: 535 Malignant: 30</td>
<td>Scattering Coefficients across visible light spectrum</td>
</tr>
<tr>
<td>D</td>
<td>Non-malignant: 426 Malignant: 30</td>
<td>Extracted Optical Properties, the mean absorption, mean scattering, absorption and scattering coefficients at 450,500, 550, 600 nm.</td>
</tr>
<tr>
<td>E</td>
<td>Non-malignant: 477 Malignant: 34</td>
<td>Data stratified by women who have not undergone Chemotherapy</td>
</tr>
<tr>
<td>F</td>
<td>Non-malignant: 535 Malignant: 30</td>
<td>Data stratified by women who have reached menopause.</td>
</tr>
<tr>
<td>G</td>
<td>Non-malignant: 271 Malignant:</td>
<td>Data set excluding sites composed of fibro-adipose tissue</td>
</tr>
<tr>
<td>H</td>
<td>Non-malignant:</td>
<td>Data set excluding sites composed of adipose tissue</td>
</tr>
<tr>
<td>I</td>
<td>Non-malignant: 571 Malignant:</td>
<td>Data set excluding sites composed of fibro-glandular tissue</td>
</tr>
</tbody>
</table>

2.4.1 Spectral Data

For the spectral data sets, the wavelength was the feature of interest and decision fusion was used as a tool to determine each wavelengths’ significance for malignancy
detection. Based on the decision fusion technique explained above, wavelength performance is relative to its specific data subset. In this study, the ten wavelengths with the greatest AUC averages in stage two were used in stage three.

2.4.2 The Effect of Tissue Type on Specificity

As an investigation to identify whether some normal tissue types may be easier to classify against malignant types, data sites comprised of these tissue types were each removed from the entire data set separately in order to quantify the effect on detection. These tissue types include fibro-adipose, fibro-glandular and adipose. Each spectrum and extracted optical properties were assessed through the decision fusion method excluding these tissue types.

2.4.3 Decision Fusion Validation with Data Subsets

For data subsets E and F, validation was performed to verify whether decision fusion performance was dependent on subset size. A type of validation was used to test 100 sample sets matching the size and detection sets of data subset E and F that involved collecting bootstrap sample sets. The bootstrap sample sets were data points randomly selected without replacement from the entire data. Performance was then assessed by looking at the box plot metrics associated with performance at Pd = 0.9 and average AUC. The validation procedure was used to determine whether decision fusion methods performance is dependent on the data set size or the actual content of the data.

3 Results & Discussion

The results section will cover the performance achieved at each stage in the detection approach for each feature and of the data sets previously described. The results section will be structured by each of the stages in the decision fusion approach.
Stage one of the approach involves creating an ROC curve for each individual wavelength or extracted optical property based on the distributions of normal and malignant sites. Each point on the ROC curve (known as an operating point) has a unique true positive and false negative probability resulting from a particular threshold value from within the range of the data set as the classifier. An ROC curve is comprised of many of these operating points. The systematic approach will select an optimized operating point from the ROC curve that is where the difference between the false positive and true positive is minimized. The reliability of this operating point selection approach is reviewed later in this section. The operating points selected in stage one for each wavelength or extracted optical property is carried onto stage two.

Stage two of the approach involves taking the selected operating points of any two wavelengths (for spectral data) or any two extracted optical properties and implementing decision fusion. Decision fusion takes these two operating points into a processor that determines an optimized classifier. In stage two, the interactions between two features can be assessed by the performance metrics of the ROC curve derived from the optimized decision fusion classifier. Features that show high contributions in performance (based on ROC curve metrics) in this stage are then carried to stage three. The main goal of stage two is to identify the key features from among the different data sets that provide higher detection performance.

Stage three of the approach examines up to ten of the best performing features from the data set of interest. The achieved detection performance from fusing the best ten performing wavelengths from any one of the spectral data sets will be compared to the decision fusion detection performance of the seven of the extracted optical properties. The goal of stage three is to evaluate the detection performance achieved by using the features selected by the approach. The second goal of stage three is to compare
the performance of the spectral data against the extracted optical properties in order to re-evaluate whether the extracted optical properties are a better tool for cancer detection or possibly suggest whether there is still an advantage to using specific wavelengths from the measured spectral data.

The initial sections will cover the extracted optical properties performances in stage I, II and III of the decision fusion approach. The three stage feature extraction optimization approach for classifying malignancy will be implemented on the entire data set, but also against different data sets that stratify the data into different demographic categories or by tissue types in order to further investigate the effect these factors have on detection performance.

The following section will review the performance achieved when stratifying the extracted optical property data by post-menopausal patients and chemo-naive patients (patients who have not had chemotherapy). The last section will then review the performance achieved when the fibro-adipose, fibro-glandular and adipose tissue sites are removed from the data set.

The next chapter section will review the measured spectrum and their performance given different data set stratifications.

### 3.1 Extracted Optical Properties

This section will review the detection performance for the extracted optical properties. These extracted optical properties refer to saturated hemoglobin (HB Sat), total hemoglobin (Total HB), beta-carotene (ß-carotene), de-oxygenated hemoglobin (DeOxy), oxygenated hemoglobin (Oxy), mean scattering (mean $\mu_s$), and mean absorption (mean $\mu_a$). Average scattering and average absorption refer to the value
obtained from taking the average across the entire measured scattering and absorption spectrum, relatively.

The extracted optical properties are extracted from the measured diffuse reflectance, scattering and absorption spectra by means of the inverse Monte Carlo method described in the background sections. For the purposes of bio-photonics and optical inference, the biological relevance of the extracted optical properties are believed to not only provide information with regard to tissue composition but also in turn serve as an improved classifier for malignancy detection [4]. Due to this reason, optical research has put a particular focus on these specific extracted optical properties for breast cancer detection and research [3].

3.1.1 Stage One: Individual Wavelength Classifier

As described previously, stage one finds the individual detection potential for each feature separately. This section will review the detection performance obtained from each extracted optical property.

The figure below summarizes the AUC achieved from stage one for each extracted optical property. The solid line indicates the AUC achieved for each optical property across all site data. The varied dotted lines indicate the AUC achieved for each optical property in the data sets stratified by post-menopausal sites and chemotherapy naïve sites.
Figure 5: This figure illustrates the AUC achieved for the entire data set, the post-menopausal data set and the chemotherapy naïve data set.

Across all the extracted optical properties, the AUC ranged from 0.5 to 0.67. With respect to the data set including all site data, the deoxygenated hemoglobin and mean $\mu s$ (scattering) stood out as the properties with the best ROC curves, as they had the highest AUC. This suggests that mean scattering provides the best contrast between malignant and non-malignant tissue. On the contrary, $\beta$-carotene had the lowest performing ROC curve with an AUC of merely 0.5. This suggests the $\beta$-carotene is not an ideal property for breast-malignancy detection.

3.1.2 Stage Two: Pair-Wise Decision Fusion

At this stage only two local detectors are considered for decision fusion at a time, therefore decision fusion was used to determine a new optimized classifier for every combination of pairs of extracted optical properties. The figure below is a graphical representation of the AUC as a function of the extracted optical properties.
representation of the AUC for the ROC curves achieved for each pair of extracted optical properties for all site data.

Figure 6: A graphical representation of the AUC achieved from implementing decision fusion on each pair of extracted optical properties.

The pairs resulting in higher AUC are redder in shade and the pairs resulting in a lower AUC are bluer in shade. All pairs including mean scattering distinctly resulted in higher AUC than any other extracted optical property, particularly the oxygenated hemoglobin and mean scattering pair. This suggests that mean scattering plays an influential role in providing high performance. Mean scattering performed best in stage one, with an AUC of approximately 0.64, however when paired with oxygenated hemoglobin, the AUC is greater than 0.68. Likewise, as β-carotene individually performed lowest in stage one, all pairs including β-carotene resulted in the ROC curves with the lowest AUC. This suggests that β-carotene detracts from detection performance and may not be an important property for breast cancer detection. The table below lists the exact AUC achieved for each pair of extracted optical properties.
Table 2: AUC achieved for each pair of optical properties

<table>
<thead>
<tr>
<th></th>
<th>β-Carotene</th>
<th>De-Oxy HB</th>
<th>Oxy HB</th>
<th>Total HB</th>
<th>Mean Abs</th>
<th>Mean Scat</th>
</tr>
</thead>
<tbody>
<tr>
<td>HB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturated</td>
<td>0.578</td>
<td>0.616</td>
<td>0.629</td>
<td>0.614</td>
<td>0.615</td>
<td><strong>0.654</strong></td>
</tr>
<tr>
<td>β-Carotene</td>
<td>0.608</td>
<td><strong>0.644</strong></td>
<td>0.621</td>
<td>0.606</td>
<td>0.606</td>
<td><strong>0.646</strong></td>
</tr>
<tr>
<td>De-Oxy HB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxy HB</td>
<td></td>
<td></td>
<td></td>
<td>0.643</td>
<td>0.643</td>
<td><strong>0.675</strong></td>
</tr>
<tr>
<td>Total HB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.629</td>
<td><strong>0.667</strong></td>
</tr>
<tr>
<td>Mean Abs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The higher AUC pairs range from 0.643 to 0.675. Overall pairs including oxygenated hemoglobin and/or mean scattering yielded the highest AUC, suggesting that these may be the most important properties for breast cancer detection. Although AUC is a strong indicator of classifier performance, the sensitivity and specificity that can be achieved can provide a more comprehensive understanding of performance potential. The table below summarizes the performance at the operating point on the ROC curve that corresponds to 80% sensitivity achieved from the decision fusion of the pairs with the highest AUC.

Table 3: The following table lists the achieved sensitivity and specificity by the decision fusion of certain pairs of extracted optical properties

<table>
<thead>
<tr>
<th>Sens, Spec</th>
<th>Oxy HB</th>
<th>Mean Scat</th>
</tr>
</thead>
<tbody>
<tr>
<td>De-Oxy HB</td>
<td>80%, 32%</td>
<td>80%, 30%</td>
</tr>
<tr>
<td>Oxy HB</td>
<td>80%, 35%</td>
<td></td>
</tr>
<tr>
<td>Total HB</td>
<td>80%, 32%</td>
<td></td>
</tr>
<tr>
<td>Mean Abs</td>
<td>80%, 37%</td>
<td></td>
</tr>
</tbody>
</table>

Although the mean scattering and oxygenated hemoglobin pair had the highest AUC, its specificity is second to the mean scattering and mean absorption pair, which yields the best specificity.
3.1.3 Stage Three: Group-Wise Decision Fusion

Stage three takes the classifiers for all the extracted properties and calculates a robust classifier using the decision fusion processor, unlike stage two which only fuses two different classifiers. The figure below plots the group-wise decision fusion curve of all extracted optical properties along with the individual ROC curves from stage one in order to see the improvement that decision fusion has on performance.

![ROC Decision Fusion of Extracted Optical Properties](image)

**Figure 7: The ROC curve achieved from implementing decision fusion with all seven extracted optical properties**

The group wise decision fusion classifier yields an ROC curve with an AUC of 0.6926, which is greater than even the best pair from stage two, which was 0.675. At 80% sensitivity a specificity of 44% can be achieved with the group-wise decision fusion classifier. At 90% sensitivity a specificity of 27% is attained, and at 70% sensitivity a specificity of 56% is attained.
3.1.4 Demographic Groups

The effects of specific patient conditions such as chemotherapy or menopause may play a role in changing the overall tissue composition of the breast. For the purposes of malignancy detection these characteristics can be exploited in order to expose any significant differences in malignancy and normal tissue distributions that can result in improved detection. Due to limitations of the data set in this study, only chemotherapy naïve (chemo naïve) sites and post-menopausal sites will be isolated for examination through the three step decision fusion approach. In this section, the data will be stratified in two different ways. In order to examine the effects of post-menopause on tissue composition and detection the data set was stratified by only the post-menopausal sites (data points). Likewise, in order to examine the effects on detection for sites that have not undergone chemotherapy treatment the data set will be stratified by data points from patients who have not undergone chemotherapy.

3.1.4.1 Post- Menopausal

A significant portion of the site level data was from patients who are post-menopausal. There are several biological changes that result in women during and after menopause, especially in relation to the female sex organs. Menopause, also closely related to the factor of age, may have a unique effect on breast tissue. These unique variations in the breast tissue of post-menopausal women may be measured and observed with optical spectroscopy. This section will review the detection performance achieved by looking solely at post-menopausal tissue sites.

Figure 5 reviews the stage one performance achieved for each of the three data sets, including the stratified post-menopausal data set. Compared against the stage one performance of using the entire data set, the individual detection performance improved for every extracted optical property except saturated hemoglobin. This suggests that the
normal and malignant distributions for the post-menopausal sites are uniquely
differentiated.

Stage two was performed in the same manner except with the post-menopausal
stratified data set and the results are summarized below.

**Figure 8: The AUC for each pair of extracted optical properties.**

For the post-menopausal group, oxygenated hemoglobin and total hemoglobin
pairs have the greatest AUC. The oxygenated hemoglobin paired with total hemoglobin
pair performs the best. This is slightly different from the entire data group, where mean
scattering pairs performed best. The table below summarizes the AUC for each pair.

**Table 4: The following table lists the AUC achieved for the decision fusion of
all possible pairs of extracted optical properties for the post-menopausal group.**

<table>
<thead>
<tr>
<th></th>
<th>B-Carotene</th>
<th>De-Oxy HB</th>
<th>Oxy HB</th>
<th>Total HB</th>
<th>Mean Abs</th>
<th>Mean Scat</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hb Saturated</strong></td>
<td>0.597</td>
<td>0.631</td>
<td>0.686</td>
<td>0.677</td>
<td>0.644</td>
<td>0.651</td>
</tr>
<tr>
<td><strong>B-Carotene</strong></td>
<td>0.649</td>
<td>0.703</td>
<td>0.694</td>
<td>0.662</td>
<td>0.669</td>
<td></td>
</tr>
</tbody>
</table>
The range for AUC for pair-wise decision fusion given this data set was between 0.597 and 0.739, which is an improvement from the entire data set, whose pair-wise decision fusion pairs ranged from 0.578 to 0.675 (refer to table 2). The highest performing pair was oxygenated hemoglobin and total hemoglobin, while the second highest performing pair was oxygenated hemoglobin and mean scattering. For the oxygenated hemoglobin and total hemoglobin pair, a sensitivity of 80% and a specificity of 42% were achieved. This suggests that the presence of oxygenated hemoglobin play a greater role in post-menopausal patients than mean scattering and mean absorption which were found to be better detectors for the entire data set.

For the stage three group-wise decision fusion of the post-menopausal stratified data set, all seven extracted optical properties classifiers were fused by decision fusion. The figure below plots the fused ROC curve for this data set against the ROC curves for the classifier of each individual extracted optical property.
Figure 9: The group-wise ROC decision fusion curve for all seven extracted optical properties for the post-menopausal data set.

The table below summarizes the sensitivity and specificity that can be achieved from the group-wise decision fusion curve for post-menopausal sites.

Table 5: Sensitivity and Specificity from the group-wise decision fusion curve for the post-menopausal data set

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Specificity: Post Meno</th>
<th>Specificity: All Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>53%</td>
<td>27%</td>
</tr>
<tr>
<td>80%</td>
<td>70%</td>
<td>44%</td>
</tr>
<tr>
<td>70%</td>
<td>80%</td>
<td>56%</td>
</tr>
</tbody>
</table>

The AUC for the group-wise ROC curve reached 0.831, which is an improvement from the original data set whose group-wise decision fusion curve achieved an AUC of 0.693. Not only did the AUC increase, but the sensitivity and specificity improved for the post-menopausal group as well, with 90% sensitivity and 53% specificity where the original data set had a specificity of only 27% at 90% sensitivity. That marks a 26%
increase in specificity from the original data set to the post-menopausal stratified data set.

3.1.4.2 Chemotherapy Naïve

The chemo-naïve group was investigated through the three stage approach in search for trends and to examine and compare the detection potential for this unique group of patients.

Figure 2 reviews the stage one performance achieved for each of the three data sets, including the stratified post-menopausal data set. Compared against the stage one performance of using the entire data set, the individual detection performance improved for every extracted optical property except ß-carotene. This suggests that the normal and malignant distributions are unique for the chemo-naïve sites in such a way that performance improves. Although there is an improvement from the original data set, the performance of the chemo-naïve group was also higher than the post-menopausal group for the saturated hemoglobin, deoxygenated hemoglobin and mean scattering.

Pair-wise decision fusion was then implemented for each pair of extracted optical properties with the chemo-naïve group. The figure below summarizes the pair-wise performance at stage two.
Figure 10: Pair-wise performance for the decision fusion of each pair of extracted optical properties for the chemo-naïve group

Like the original data set, most pairs with mean scattering had the highest AUC, while the pairs with deoxygenated hemoglobin had the next best AUC. Pairs with β-carotene had the lowest AUC. The pair with the greatest AUC was deoxygenated hemoglobin and mean scattering, which is consistent with the pattern for this group. This suggests that deoxygenated hemoglobin and mean scattering may play a special role in malignancy detection performance for chemo-naïve patients. The table below lists all the AUC for each pair of extracted optical properties for the chemo-naïve group.
Table 6: The following table lists the achieved AUC for each pair of extracted optical properties for the chemo-naïve data set.

<table>
<thead>
<tr>
<th></th>
<th>B-Carotene</th>
<th>De-Oxy HB</th>
<th>Oxy HB</th>
<th>Total HB</th>
<th>Mean Abs</th>
<th>Mean Scat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hb Saturated</td>
<td>0.595</td>
<td>0.664</td>
<td>0.652</td>
<td>0.636</td>
<td>0.645</td>
<td>0.685</td>
</tr>
<tr>
<td>B-Carotene</td>
<td>0.631</td>
<td>0.621</td>
<td>0.604</td>
<td>0.613</td>
<td>0.655</td>
<td></td>
</tr>
<tr>
<td>De-Oxy HB</td>
<td>0.678</td>
<td>0.669</td>
<td>0.674</td>
<td>0.701</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxy HB</td>
<td>0.656</td>
<td>0.661</td>
<td>0.697</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total HB</td>
<td>0.649</td>
<td>0.689</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Abs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.693</td>
</tr>
</tbody>
</table>

The range of AUC for this group was from 0.595 to 0.701. This range is smaller than that of post-menopausal but still larger than the original data group for stage two. Again, the best pair, deoxygenated hemoglobin and mean scattering had the highest AUC of 0.701. This suggests that stratifying the data by chemo-naïve patients can in fact improve detection performance. The improvement may be significant enough to use the two extracted optical properties alone for an alternative detection method.

In stage three, the group-wise decision fusion of all the extracted optical properties for the chemotherapy naïve group had an AUC of 0.789, which is larger than the AUC of 0.6926 for the original non-stratified data set. The figure below shows the group-wise decision fusion curve for the chemo-naïve stratified data set.
Similar to the results mentioned previously in stage two, the results for group-wise decision fusion show an improvement from using the original non-stratified data set. The sensitivity and specificity achieved using the group-wise classifier is summarized below.

Table 7: The Sensitivity and Specificity achieved from the group-wise decision fusion classifier of the chemotherapy naïve data set.

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Specificity: Chemo-Naïve</th>
<th>Specificity: All Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>43%</td>
<td>27%</td>
</tr>
<tr>
<td>80%</td>
<td>62%</td>
<td>44%</td>
</tr>
<tr>
<td>70%</td>
<td>73%</td>
<td>56%</td>
</tr>
</tbody>
</table>

The specificity improves by almost 60% from the original data set. This suggests that stratifying the data by chemo-naïve patients will enhance performance.
3.1.5 Tissue Type Analysis

Decision fusion can serve as a tool to determine classifiers for breast cancer tissue but can also be used as an investigative tool to study the effects of different factors, such as demographic status or tissue type on malignancy detection. This section will simply illustrate how decision fusion methods, such as the one in this paper can be used to analyze the effect of specific tissue types on malignancy detection. There are several types of tissues in any given excised breast tumor mass [2]. In this study there were as many as 10 different tissue types found in the samples. However, some research shows that the normal tissue types known as fibro-adipose and fibro-glandular have very similar optical finger-prints as the malignant tissue types known as DCIS and DC [3]. It may be possible that specificity may be adversely affected by these tissue type similarities. In an attempt to measure and quantify the potential affects on detection, decision fusion was implemented on data sets excluding either fibro-glandular, fibro-adipose, adipose, or all three of these tissue types. The figure below is the stage-three decision fusion curves for the extracted optical properties. Each curve corresponds to a data set that excludes a specific tissue type. The “All Tissue Types” refers to the original decision fusion curve in which no tissue types have been removed.
According to the figure, it appears that removing either fibro-adipose or fibro-glandular tissue sites from the data will just slightly improve the ROC curve. However, removing adipose tissue sites, because those are the majority tissue sample, the performance actually decreases, which suggests that adipose is easily classifiable with the decision fusion model. The table below reviews the extracted sensitivity and specificity for each curve in the figure above.

Table 8: Summary of the performance for each tissue type decision fusion curve.

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Fibro-Adipose</th>
<th>Fibro-Glandular</th>
<th>Adipose</th>
<th>All Tissues</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>37%</td>
<td>36%</td>
<td>27%</td>
<td>37%</td>
</tr>
<tr>
<td>80%</td>
<td>55%</td>
<td>54%</td>
<td>43%</td>
<td>55%</td>
</tr>
<tr>
<td>70%</td>
<td>67%</td>
<td>67%</td>
<td>56%</td>
<td>67%</td>
</tr>
<tr>
<td>Total Normal Sites</td>
<td>535</td>
<td>571</td>
<td>271</td>
<td>585</td>
</tr>
</tbody>
</table>

The sensitivity and specificity seems to be unaffected by excluding fibro-adipose or fibro-glandular tissue types. Removing adipose samples reduces performance, which is expected since the data set shrinks from 585 normal sites to 271 normal sites.
3.1.6 Summary  
Mean scattering and oxygenated hemoglobin were identified as the features that decision fusion can best exploit for detection. As a pair, mean scattering and oxygenated hemoglobin achieved a sensitivity of 80% and specificity of 35%. However, the optical property pair that has the highest specificity (37%) was mean scattering and mean absorption, which had the second highest AUC after mean scattering and oxygenated hemoglobin. By using decision fusion with all seven extracted optical properties, sensitivity improved to 90% with 37% specificity, or at 80% sensitivity a specificity of 55% was achieved. Specificity does not seem to be affected by the tissue types of fibro-adipose or fibro-glandular, however further investigation involving tissue type is possible for the remaining seven.

3.2 Measured Spectral Data  
Before optical properties can be extracted, optical spectroscopy begins by measuring the diffuse reflectance of a tissue sample. From there the inverse Monte Carlo model can be used to estimate the scattering and absorption that occurs within the tissue sample. By recognizing the scattering and absorption patterns of different molecules, the biological constituents of a tissue can be determined. However, it is possible that information is lost during this conversion. Although the raw spectral data does not directly provide a biological correspondence to the tissue it may still preserve different aspects of the tissue sample that may improve detection. The first goal of this study is to examine the detection potential of each of these measured spectral data, for the same data sites used for the extracted optical properties, and determine the performance strengths for each. The second goal of this study was to compare this performance against the performance achieved by using the extracted optical properties and determine if in fact the extracted optical properties are a better alternative to detection or
if the spectral data has more detection potential. The performance for the spectral data will be determined through the same three step approach used in the previous sections. As in the previous section the data sets were partitioned by demographic status and tissue types in order to explore their affects on the detection performance for each spectra.

3.2.1 Stage One: Individual Wavelength Classifier

This section summarizes the AUC achieved by the ROC curves created for each wavelength individual classifier against the different measured spectra.

![AUC as a function of Wavelength for each Spectrum](image)

Figure 13: The following figure illustrates the AUC achieved by the assigned classifier determined by each individual wavelength by either the measured Diffuse Reflectance spectrum, Scattering spectrum or Absorption spectrum.

3.3 Absorption

The following section will review the results obtained for the performance of the absorption spectrum in stage I, II and III of the decision fusion approach. The next section will review the performance achieved when stratifying the absorption spectrum data by post-menopausal patients and chemo-naïve patients. The last section will then review the performance achieved when the fibro-adipose, fibro-glandular and adipose tissue sites are removed from the data set.
3.3.1 Stage One: Individual Wavelength Classifier Performance

At this stage an ROC curve was created for each wavelength from the absorption spectrum independently in order to determine the most optimum detector specific to the normal and malignant distributions at each wavelength. Because light is absorbed differently at each wavelength, it is important to examine the measured absorption at each wavelength as particular wavelengths may possess a specific absorption pattern that is advantageous for malignancy detection. The AUC of the ROC curves for each wavelength are summarized in the upper plot of figure 11. The lower plot of figure 11 shows the average absorption for all the malignant cases and the average absorption for the normal cases.
Figure 14: The figure above illustrates the AUC achieved by the assigned classifier at each wavelength for the measured absorption. The lower figure illustrates the difference in the average measured absorption at each wavelength between normal and malignant tissue.

The greatest AUC was achieved at the longer wavelengths, between 580 and 600 nm, where the lowest AUC were achieved at the shorter wavelengths, between 455 and 500 nm. The measured absorption spectrum was also plotted in the bottom figure, showing the average measured absorption for normal tissue sites and the average measured absorption for the malignant sites.

3.3.2 Stage Two: Pair-Wise Decision Fusion Performance

As described in the previous section, decision fusion can exploit unexpected interactions that may occur between different combinations of features. In order to examine the combined decision potential of each wavelength, decision fusion was
implemented for every combination of wavelength pairs and the metrics of their resulting ROC curves were compared to find trends. The figure below illustrates AUC achieved by fusing the decisions of every combination of wavelength pairs. The combinations with the greatest AUC appear red while the pairs with lower AUC are blue.

Figure 15: This figure visualizes the AUC after decision fusion with every combination of wavelength pairs.

By visual inspection there are a few wavelengths that stand out as having a higher AUC. These wavelengths are summarized in the table below. There are a total of 60 different pairs for each wavelength, in order to determine the overall wavelength pair performance, the average AUC was taken for each of these pairs for each wavelength. The wavelengths with the highest averages were taken to the next stage for group-wise
decision fusion, and are considered the more influential wavelengths. The table below lists the highest performing wavelengths in descending order.

The table below summarizes the pairs from the pair-wise decision fusion whose ROC curves yielded the highest AUC.

Table 9: The highest AUC and corresponding wavelength pairs from the absorption spectrum.

<table>
<thead>
<tr>
<th>AUC</th>
<th>Wavelength (nm) Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6563</td>
<td>592.5 585</td>
</tr>
<tr>
<td>0.6564</td>
<td>470 592.5</td>
</tr>
<tr>
<td>0.6571</td>
<td>592.5 600</td>
</tr>
<tr>
<td>0.6573</td>
<td>500 592.5</td>
</tr>
<tr>
<td>0.6584</td>
<td>527.5 592.5</td>
</tr>
<tr>
<td>0.6585</td>
<td>450 592.5</td>
</tr>
<tr>
<td>0.6590</td>
<td>505 592.5</td>
</tr>
<tr>
<td>0.6597</td>
<td>582.5 592.5</td>
</tr>
<tr>
<td>0.6620</td>
<td>502.5 592.5</td>
</tr>
</tbody>
</table>

The table above summarizes the performance for the wavelength pairs that yield the best ROC curves in stage two. The pairs with the 592.5 nm wavelength consistently had high AUC.

Table 10: The following table lists the wavelengths with the highest average AUC in descending order

<table>
<thead>
<tr>
<th>Top 10 Wavelengths (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>592.5 502.5 582.5 505 527.5</td>
</tr>
<tr>
<td>450 585 500 600 495</td>
</tr>
</tbody>
</table>

The wavelengths in the table above consistently performed best in the pair-wise decision fusion. Given their higher AUC, they are assumed to have the most detection potential and are used in the next stage.
3.3.3 Stage Three: Group-Wise Decision Fusion

The figure below illustrates the ROC curves for the performance potential for the group-wise decision fusion of one, using the top 5 wavelengths, and to 10 wavelengths.

![ROC Curves](image)

**Figure 16:** This figure presents the absorption fusion curves plotted against the fused extracted optical properties curve.

Both of these group-wise fusion curves are plotted and compared with the performance achieved from fusing the 7 extracted optical properties. By using only 5 wavelengths from the absorption spectrum performs just as well as using all the information from the seven extracted optical properties. And by fusing all 10 of the high performing wavelengths from the absorption spectrum completely surpasses the detection performance of the extracted optical properties.

The table below reviews the specific performance metrics achieved from both the 10-wavelength and 5-wavelength group-wise decision fusion.
By using the highest performing wavelengths from stage two, a sensitivity and specificity of 90% and 50% respectively was achieved. Likewise, using only the 5 highest performing wavelengths from stage two rendered a sensitivity and specificity of 90% and 36% respectively. This is a 23% (10 wavelength group-wise) and 9% (5 wavelength group-wise) improvement in specificity from the group-wise detection potential of the extracted optical properties. These results suggest that using the absorption spectrum may provide better detection performance than with the extracted optical properties.

### 3.3.4 Demographic Stratification

As in the previous section, stratifying the data set by different demographic factors may improve detection performance. This section will review the performance in each stage of the decision fusion approach for the absorption data stratified by first post-menopausal and then by chemotherapy naïve data sites. The figure below summarizes the AUC achieved at stage one for the absorption spectrum for each data set.
Figure 17: The AUC achieved at each wavelength of the Absorption spectrum.

Although generally following the same shape, the AUC of the post-menopausal stratified data set are greater than the other two groups. While the chemotherapy naïve stratified data set did worse than the original all data set.

3.3.4.1 Post-Menopausal

The figure below summarizes the pair-wise decision fusion results for every combination of wavelength pairs.
Figure 18: The AUC resulting from the decision fusion of all combinations of wavelength pairs from the absorption spectrum stratified by post-menopausal status.

When stratified by post-menopausal sites, the greatest pair-wise AUC achieved is approximately 0.75, which is better than even the best pair from the original entire absorption data set of 0.66. In general, the pairs involving the wavelengths 450-460 nm and 500-515 nm resulted in the greatest AUC. The absorption spectrum out performed the extracted optical properties, whose post-menopausal data set had a maximum AUC of 0.739.

Table 13: The sensitivity and specificity for absorption spectrum stratified by post-menopausal sites for 10 wavelength fusion.

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>82%</td>
<td>0.94079</td>
</tr>
<tr>
<td>80%</td>
<td>92%</td>
<td></td>
</tr>
<tr>
<td>70%</td>
<td>95%</td>
<td></td>
</tr>
</tbody>
</table>

In stage-three group-wise decision fusion, the wavelengths that performed best as pairs were used in towards the fusion for the group-wise ROC curve. The figure
below summarizes the ROC curves for the absorption spectrum stratified by post-menopausal patients against the performance of the group-wise ROC curve for the extracted optical properties.

Figure 19: ROC curves of the group-wise decision fusion classifier for both the 10 and 5 wavelengths plotted with the ROC curve for the extracted optical properties. All of which represent the post-menopausal data set.

The ROC curve performance of the 5-wavelength group outperformed that of the extracted optical properties. The 10 wavelength group-wise ROC curve reached a specificity of 82% at 90% sensitivity.
3.3.4.2 Chemotherapy Naïve

The figure below summarizes the pair-wise decision fusion results for every combination of wavelength pairs.

**Figure 20:** The AUC resulting from the decision fusion of all combinations of two wavelength pairs from the absorption spectrum stratified by chemotherapy.

The pair-wise performance for the chemotherapy-naïve data set had lower AUC. However, all pairs with the 592.5 or 595 nm wavelength considerably outperformed the other pairs.

In stage-three group-wise decision fusion, the wavelengths that performed best as pairs were used in the fusion for the group-wise ROC curve. The figure below summarizes the ROC curves for the absorption spectrum stratified by chemotherapy.
naïve patients against the performance of the group-wise ROC curve for the extracted optical properties.

Figure 21: The stage-three ROC curve performance for the absorption spectrum plotted with the stage three ROC curve for the extracted optical properties.

For the chemotherapy naïve group, the 5-wavelength group classifier had lower performance than the extracted optical properties. This stratification did not do as well as post-menopausal stratification. At 90% sensitivity only a specificity of 52% was achieved. This suggests that using chemotherapy naïve status does not contribute to the detection performance power.

3.3.4.3 Summary

Across all the data groups (post-menopausal, chemotherapy naïve, and all data) the data set stratified by post-menopausal sites performed better in all stages, with a specificity of 82% and 90% sensitivity.
3.3.5 Tissue Type Analysis

As mentioned in the previous section, there are certain normal tissue types that are optically similar to the cancerous tissue types. To determine the effects of these tissue types on specificity, decision fusion identifies differences in the performance of data sets excluding different tissue types. This section is not intended to be an in depth examination but more of an example highlighting some of the advantages decision fusion can have in quantifying differences in detection of tissue types.

In the figure below, the ROC curves generated from the decision fusion of the 4 different data sets are below. The first data set excludes any sites that are of the fibro-adipose tissue types, the second data set excludes any fibro-glandular tissue types, the third excludes any adipose tissue types and the final data set is the original, all inclusive data set.

Figure 22: Stage three ROC curves for the absorption data excluding specific tissue types.
The figure above reveals that removing any of the three tissue types improves the detection performance of the absorption spectrum, however removing the fibro-glandular tissue type provides the greatest improvement. This suggests that it is the fibro-glandular tissue that is most frequently classified incorrectly.

3.4 Scattering

This section will review the results obtained for the performance of the scattering spectrum in stage I, II and III of the decision fusion approach. The following section will review the performance achieved when stratifying the scattering spectrum data by post-menopausal patients and chemo-naive patients. The last section will then review the performance achieved when the fibro-adipose, fibro-glandular and adipose tissue sites are removed from the data set.

3.4.1 Stage One: Individual Wavelength Classifier Performance

At this stage, an ROC curve was created for each wavelength from the scattering spectrum independently in order to determine the most optimum classifier specific to the $H_1$ and $H_0$ distributions at that wavelength. The AUC of the ROC curves for each wavelength are summarized in the figure below.
Figure 23: This figure illustrates the AUC achieved by the assigned classifier at each wavelength for the measured scattering. The lower figure illustrates the difference in the average measured absorption at each wavelength between normal and malignant tissue.

The greatest AUC was achieved at the longer wavelengths, between 540 and 600 nm, where the lowest AUC were achieved at the shorter wavelengths, between 450 and 520 nm. The measured scattering spectrum was also plotted in the bottom figure, showing the average measured scattering for normal tissue sites and the average measured scattering for the malignant sites.

3.4.2 Stage Two: Pair-Wise Decision Fusion

The following section reviews the performance at the stage-two level. The figure below summarizes the AUC for each pair of wavelengths.
Figure 24: This figure visualizes the achieved AUC for each wavelength pair from the scattering spectrum.

While there are several wavelengths that when paired perform relatively well, 560 nm is the most distinct. The tables below highlight the pairs that yielded the greatest AUC.

Table 14: The AUC and its corresponding wavelength pair.

<table>
<thead>
<tr>
<th>AUC</th>
<th>Wavelength Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6968</td>
<td>452.5</td>
</tr>
<tr>
<td>0.6970</td>
<td>452.5</td>
</tr>
<tr>
<td>0.6970</td>
<td>560</td>
</tr>
<tr>
<td>0.6972</td>
<td>535</td>
</tr>
<tr>
<td>0.6972</td>
<td>452.5</td>
</tr>
<tr>
<td>0.6974</td>
<td>560</td>
</tr>
<tr>
<td>0.6976</td>
<td>520</td>
</tr>
<tr>
<td>0.6978</td>
<td>577.5</td>
</tr>
<tr>
<td>0.6978</td>
<td>560</td>
</tr>
<tr>
<td>0.6982</td>
<td>452.5</td>
</tr>
</tbody>
</table>
The highest AUC achieved at stage one was 0.6982, from the pair of 452.5 and 560 nm. The average AUC for all pairs of each wavelength were taken and the results are summarized below.

Table 15: The following table lists the wavelengths with the highest average AUC in descending order.

<table>
<thead>
<tr>
<th>Wavelengths (nm) With Largest Average AUC</th>
<th>560</th>
<th>450</th>
<th>452.5</th>
<th>457.5</th>
<th>462.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>570</td>
<td>572.5</td>
<td>575</td>
<td>577.5</td>
<td>455</td>
<td></td>
</tr>
</tbody>
</table>

There is some overlap between the wavelengths that had the best pair-wise AUC and the wavelengths that had the highest average AUC across all pairs. The wavelengths from the table above were used in the next stage for group-wise fusion.

3.4.3 Stage Three: Group-Wise Decision Fusion

The following tables summarize the performance achieved from fusing all 10 of these wavelengths compared to just using the 5 most influential wavelengths. Both of these group-wise fusion curves were then compared to the performance achieved from fusing the 7 extracted optical properties.

Table 16: Summary of the ROC performance achieved from fusing the ten (left) and the upper 5 (right) best performing wavelengths from stage I for the Scattering Spectrum.

<table>
<thead>
<tr>
<th>Decision Fusion With 10 Wavelengths Performance</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90%</td>
<td>65%</td>
<td>0.88748</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>82%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70%</td>
<td>88%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Decision Fusion With 5 Wavelengths Performance</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90%</td>
<td>44%</td>
<td>0.8035</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>61%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70%</td>
<td>73%</td>
<td></td>
</tr>
</tbody>
</table>

With the scattering spectrum in group-wise decision fusion a classifier with 90% sensitivity and 65% specificity was achieved given the top 10 wavelengths. Even at the 5-
wavelength group at 90% sensitivity, 44% specificity is achieved. The specificity for both 10 and 5-wavelength group-wise decision fusion shows a 38% and 17% improvement from the extracted optical properties, respectively. Both the 10-wavelength and 5-wavelength group-wise decision fusion resulted in classifiers that outperformed the classifiers from the group-wise extracted optical properties. The table below summarizes the ROC for the scattering spectrum against the extracted optical properties.

<table>
<thead>
<tr>
<th>ROC Decision Fusion Curves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area Under Curve = 0.80235</td>
</tr>
<tr>
<td>Performance: [0.9 0.44 0.8 0.61 0.7 0.73]</td>
</tr>
<tr>
<td>Scattering Fusion</td>
</tr>
<tr>
<td>numPos: 38 numNeg: 595</td>
</tr>
</tbody>
</table>

![Graph](image)

**Figure 25:** This figure compares the 10, 5 and extracted optical property group-wise fusion ROC curves.

Both scattering ROC curves have a higher AUC and better sensitivity and specificity against the extracted optical properties.

### 3.4.4 Demographic Status

The detection results of the three-stage approach implemented on stratified sub-data sets are reviewed below. The stage one results are summarized in the figure below.
Figure 26: The AUC achieved for each wavelength for the data sets stratified by post-menopausal sites, chemotherapy naïve sites and the entire data set.

Stratifying the data sets based on these two demographic statuses improve the detection potential, as the AUC increases by 0.02. Stratifying the data by post-menopausal sites yields the greatest AUC and ROC performance.

3.4.4.1 Post-Menopausal

The stage two results for the post-menopausal data set are illustrated below in figure 22.
Figure 27: The AUC resulting from the decision fusion of all combinations of two wavelength pairs from the scattering spectrum stratified by post-menopausal status.

The pairs that consistently produce the largest AUC involve wavelengths 450, 452.5, and 565-577.5 nm. The table below reviews the top performing pairs from stage two.

Table 17: Top Performing Pairs from Stage Two of the Scattering Post-menopausal data set.

<table>
<thead>
<tr>
<th>AUC</th>
<th>Wavelength Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7248</td>
<td>510 575</td>
</tr>
<tr>
<td>0.7248</td>
<td>577.5 600</td>
</tr>
<tr>
<td>0.7252</td>
<td>525 575</td>
</tr>
<tr>
<td>0.7253</td>
<td>557.5 575</td>
</tr>
<tr>
<td>0.7254</td>
<td>562.5 577.5</td>
</tr>
<tr>
<td>0.7258</td>
<td>575 517.5</td>
</tr>
<tr>
<td>0.7258</td>
<td>517.5 575</td>
</tr>
<tr>
<td>0.7259</td>
<td>575 600</td>
</tr>
<tr>
<td>0.7264</td>
<td>562.5 575</td>
</tr>
<tr>
<td>0.7269</td>
<td>575 577.5</td>
</tr>
</tbody>
</table>
The greatest AUC achieved in stage two was 0.7269, with wavelengths 575nm and 577.5 nm. These AUC are higher than those of the extracted optical properties and higher than full scattering data set.

Figure 28: The following figure is a plot of the group-wise decision fusion curves for both the extracted optical properties and scattering spectrum stratified by post-menopausal status.

Figure 23 shows that the performance achieved by stratifying the data by post-menopausal gives about the same performance for the scattering spectrum (using 5 wavelengths) as for the extracted optical properties. Across the scattering data sets, the AUC of the post-menopausal data set is greater than the full data set.

3.4.4.2 Chemotherapy Naïve

The stage two results for the chemotherapy-naïve data set are illustrated below in figure 24.
Figure 29: The AUC resulting from the decision fusion of all combinations of two wavelength pairs from the scattering spectrum stratified by chemotherapy naïve.

The pairs that consistently produce the largest AUC involve wavelengths 450 – 465 nm. The table below reviews the top performing pairs from stage two.

Table 18: Wavelength Pairs with the highest AUC from the Chemotherapy-Naïve Scattering Spectral data.

<table>
<thead>
<tr>
<th>AUC</th>
<th>Wavelength Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7237</td>
<td>592.5 465</td>
</tr>
<tr>
<td>0.7240</td>
<td>465 570</td>
</tr>
<tr>
<td>0.7244</td>
<td>465 522.5</td>
</tr>
<tr>
<td>0.7245</td>
<td>465 550</td>
</tr>
<tr>
<td>0.7247</td>
<td>465 585</td>
</tr>
<tr>
<td>0.7248</td>
<td>465 527.5</td>
</tr>
<tr>
<td>0.7252</td>
<td>462.5 465</td>
</tr>
<tr>
<td>0.7254</td>
<td>465 515</td>
</tr>
<tr>
<td>0.7257</td>
<td>460 465</td>
</tr>
</tbody>
</table>
The highest AUC for stage II was 0.7257, which again shows improved detection performance from the full data set. The wavelength at 465 nm consistently performs well with the decision fusion approach.

![Figure 30: Significant Wavelengths for Scattering Stratified by chemo-naïve.](image)

In stage three of the decision fusion process, again it shows that stratifying the data by a specific demographic type does improve detection, as the ROC curves have a greater AUC and sensitivity and specificity. And again, the same pattern shows that even using just 5 of the optimal wavelengths from the spectrum can produce a better ROC curve than that of the extracted optical properties.

### 3.4.5 Tissue Type Analysis
Figure 31: ROC curve generated by excluding fibro-adipose, fibro-glandular, and adipose tissue types plotted with the ROC curve generated from including all tissue types.

3.5 Diffuse Reflectance

This section will review the results obtained for the performance of the diffuse reflectance spectrum in stage I, II and III of the decision fusion approach. The following section will review the performance achieved when stratifying the diffuse reflectance spectrum data by post-menopausal patients and chemo-naive patients. The last section will then review the performance achieved when the fibro-adipose, fibro-glandular and adipose tissue sites are removed from the data set.

3.5.1 Stage One: Individual Wavelength Classifier Performance

At this stage, an ROC curve was created for each wavelength from the diffuse reflectance spectrum independently in order to determine the most optimum classifier specific to the H₁ (malignant tissue) and H₀ (normal tissue) distributions at that wavelength. The AUC of the ROC curves for each wavelength are summarized in the figure below.
Figure 32: Summary of the AUC for each wavelength independently for the diffuse reflectance (top). The bottom figure illustrates the average measured diffuse reflectance for the normal tissue and malignant tissue.

The greatest AUC was achieved at the shorter wavelengths, at 450 nm and between 480 and 500 nm, where the lowest AUC were achieved intermittently along the diffuse reflectance spectrum at 460 nm, 500 nm, 520 nm, and 590 nm. The measured diffuse reflectance spectrum was also plotted in the bottom figure, showing the average measured scattering for normal tissue sites and the average measured diffuse reflectance for the malignant sites.

3.5.2 Stage Two: Pair-Wise Decision Fusion

The figure below is a graphical representation of the AUC achieved by every decision fusion wavelength pair. The combinations shown in red have the greater AUC.
Since the majority of the pairs had similar AUC, it is important to note that the wavelength pairs that consistently performed worst involved wavelengths 450-470 nm, 517.5-527.5 nm and 585-600 nm. The highest AUC and their corresponding pairs are listed below.

Figure 33: A graphical representation of the pair-wise decision fusion for all pairs of the Diffuse Reflectance spectra.
Table 19: The AUC of the highest performing wavelength pairs.

<table>
<thead>
<tr>
<th>AUC</th>
<th>Wavelength Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7044</td>
<td>552.5 555</td>
</tr>
<tr>
<td>0.7048</td>
<td>552.5 577.5</td>
</tr>
<tr>
<td>0.7048</td>
<td>552.5 557.5</td>
</tr>
<tr>
<td>0.7053</td>
<td>487.5 552.5</td>
</tr>
<tr>
<td>0.7059</td>
<td>552.5 492.5</td>
</tr>
<tr>
<td>0.7059</td>
<td>492.5 552.5</td>
</tr>
<tr>
<td>0.7062</td>
<td>552.5 547.5</td>
</tr>
<tr>
<td>0.7062</td>
<td>547.5 552.5</td>
</tr>
<tr>
<td>0.7070</td>
<td>552.5 575</td>
</tr>
<tr>
<td>0.7097</td>
<td>550 552.5</td>
</tr>
</tbody>
</table>

The highest AUC is 0.7097, the best pair-wise detection performance of both the spectral data and extracted optical properties. The wavelength pairs that achieved this AUC were 550 and 552.5 nm.

3.5.3 Stage Three: Group-Wise Decision Fusion

The tables 19 and 20 summarize the performance achieved from fusing all 10 of these wavelengths compared to just using the 5 most influential wavelengths. Both of these group-wise fusion curves were then compared to the performance achieved from fusing the 7 extracted optical properties.
Figure 34: This figure visualizes the achieved AUC after decision fusion with 2 wavelengths. This figure is a collection of every combination of wavelength pairs.

Even with just 5 of the selected wavelengths from the extracted optical properties, the detection performance is better than with all 7 of the extracted optical properties. Likewise, fusing 10 wavelengths provides very good detection performance.

Table 20: The following table lists the wavelengths with the highest average AUC in descending order.

<table>
<thead>
<tr>
<th>Wavelength (nm) with Greatest Average AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>552.5</td>
</tr>
<tr>
<td>557.5</td>
</tr>
</tbody>
</table>

Table 21: Summary of the ROC performance achieved from fusing the ten (left) and the upper 5 (right) best performing wavelengths from stage I for the Scattering Spectrum.

<table>
<thead>
<tr>
<th>Decision Fusion With 10 Wavelengths Performance</th>
<th>Decision Fusion With 5 Wavelengths Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>90%</td>
<td>68%</td>
</tr>
<tr>
<td>80%</td>
<td>81%</td>
</tr>
<tr>
<td>70%</td>
<td>89%</td>
</tr>
</tbody>
</table>
Overall detection performance for diffuse reflectance showed superior detection potential, as the sensitivity and specificity pairs were best out of all features.

3.5.4 Demographic Status

The following section will review the detection performance achieved by the optimal decision fusion approach for the stratified data sets against the performance of using the entire data set. There will also be a comparison against the performance of those data sets for the extracted optical properties.

Figure 26 shows the stage one performance for the three different data sets.

![AUC as a function of Diffuse Reflectance Wavelength](image)

Figure 35: The figure illustrates the AUC for each wavelength evaluated individually for the entire data set, the data set stratified by post-menopausal patients, and the data set stratified by patients who have not undergone chemotherapy.

The AUC for the post-menopausal data set performed slightly better in stage one with higher AUC. However, the chemotherapy naïve data set performed slightly worse with smaller AUC.
3.5.4.1 Post-Menopausal

The following section reviews the detection performance for the decision fusion of wavelength pairs.

![Figure 36](image.png)

Figure 36: The AUC resulting from the decision fusion of all combinations of two wavelength pairs from the diffuse reflectance spectrum stratified by post-menopausal status.

In stage two for the diffuse reflectance post-menopausal data set, the wavelengths pairs that consistently performed well correspond almost completely with the wavelengths that performed best in stage one. The table below lists the AUC for the best performing wavelength pairs.
Table 22: Highest AUC for wavelength pairs of diffuse reflectance post-menopausal data set.

<table>
<thead>
<tr>
<th>AUC</th>
<th>Wavelength Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7237</td>
<td>592.5 465</td>
</tr>
<tr>
<td>0.7240</td>
<td>465 570</td>
</tr>
<tr>
<td>0.7244</td>
<td>465 522.5</td>
</tr>
<tr>
<td>0.7245</td>
<td>465 550</td>
</tr>
<tr>
<td>0.7247</td>
<td>465 585</td>
</tr>
<tr>
<td>0.7248</td>
<td>465 482.5</td>
</tr>
<tr>
<td>0.7249</td>
<td>465 527.5</td>
</tr>
<tr>
<td>0.7252</td>
<td>462.5 465</td>
</tr>
<tr>
<td>0.7254</td>
<td>465 515</td>
</tr>
<tr>
<td>0.7257</td>
<td>460 465</td>
</tr>
</tbody>
</table>

The highest AUC was 0.7257, which shows an improvement of 0.02 from the non-stratified data set. The table below summarizes the detection sensitivity and specificity for using the top 10 wavelengths from stage II.

Table 23: Detection sensitivity and specificity for the post-menopausal data set for Diffuse Reflectance for 10 wavelengths.

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>73%</td>
<td>0.9178</td>
</tr>
<tr>
<td>80%</td>
<td>86%</td>
<td></td>
</tr>
<tr>
<td>70%</td>
<td>92%</td>
<td></td>
</tr>
</tbody>
</table>

The results for stage III are summarized in the figure below.
Figure 37: The following figure is a plot of the group-wise decision fusion curves for both the extracted optical properties and diffuse reflectance spectrum stratified by post-menopausal status.

The 5 wavelength group wise fusion curve did approximately as well as the extracted optical properties. However, using 10 wavelengths still shows improved detection performance.

3.5.4.2 Chemotherapy Naïve

The following section reviews stage two and three of the optimal decision fusion approach for the chemotherapy naïve data set.
Figure 38: The AUC resulting from the decision fusion of all combinations of two wavelength pairs from the diffuse reflectance spectrum stratified by chemotherapy.

The wavelength pairs that performed best were 450-452.5 nm and 530-555 nm, which is a slightly different span of wavelengths from the original data set. The table below summarizes pair detection performance.
Figure 39: Significant Wavelengths for Diffuse Reflectance Stratified by chemo

Similar to post-menopausal, there was no significant difference between utilizing the 5 wavelengths from stage two and the 7 extracted optical properties. However, the 10 wavelength fusion curve outperformed the two on the graph, however did not yield results as promising as the post-menopausal group.

3.5.5 Tissue Type Analysis

The following section will briefly summarize the effect on detection performance that excluding specific tissue types had on specificity.
The figure above illustrates that removing either tissue set does not improve detection performance. This plot includes an extra ROC curve, the exclusive set is the set of data that excludes all three tissue types.
4 Conclusions

4.1 Feature Selection

Aspects of the diffuse reflectance, absorption and scattering spectra are each intertwined in the extraction of optical properties for both the forward and inverse models for diffuse reflectance [20]. Because of this relationship, unique features existing in each spectra may be overlooked or degraded after extraction. As illustrated in the results section, each spectrum was analyzed with decision fusion to assess each one's own ability to classify the tissue and identify wavelengths along the spectra that had the most influence on data detection.

The decision fusion approach is considered optimal because at stage one an optical local detector is chosen and carried to stage two where the features with the best decision fusion detectors was then carried to stage three that finds the best detector with a reduced number of features. This approach provides a completely automated and systematic method for identifying the most optimized classifiers and up to the 10 features that show the highest detection performance. Once the performance was determined, the different spectra were compared against each other and the extracted optical properties to determine which format of optical measurements serves as the best for malignancy detection.

The initial objective was to identify which features can provide the best detection performance. Between the extracted optical properties, scattering, absorption and diffuse reflectance spectra, the diffuse reflectance spectrum outperformed all other features in all three stages of the approach with a final performance of 90% sensitivity and 68% specificity. Although the extracted optical properties provide insight on the biological constituents of the tissue sample it is not ideal for malignancy detection. The
scattering spectrum had relatively good performance as well, with 90% sensitivity and 65% specificity.

The second objective was to observe the effects on detection when stratifying the data set by a specific demographic group. The post-menopausal data set yielded the highest detection performance for all data sets. The post-menopausal data set of the absorption spectrum had the best performance of all features, with a sensitivity of 90% and a specificity of 82%. This suggests that there is a unique effect on tissue absorption for post-menopausal women that allows better differentiation between normal and malignant tissues.

This study shows how decision fusion can be used to improve the detection of malignancy in breast tumor margins given a large variety of parameters. However future work needs to be done in order to further explore the benefits of the decision fusion tool for extended tissue detection analysis. By removing certain tissue types from the data set and implementing decision fusion the specificity improvement was observed for fibro-adipose, fibro-glandular and adipose tissues. Because specificity slightly increased more when fibro-adipose tissues were removed it suggests that off the three tissue types in this study, fibro-adipose is most difficult to classify correctly against malignant tissue.
Appendix A: Threshold Value Assessment

Currently, some detection algorithms for breast cancer tissue involve determining a threshold value of some sort for an optical property. Similarly, this approach provides an automated and systematic approach to indirectly identifying the optimized threshold values for each optical property as a classifier. Understanding these threshold values may serve as a platform to understanding general breast tissue composition for normal and malignant tissue.

The table below lists the threshold values corresponding to the selected operating point on the ROC curve determined at the stage one. These threshold values have been extracted for three data sets, the all inclusive site data set, the chemo naïve stratified data set and the post-menopausal stratified data set. It is useful to compare the change in the optimized threshold value between demographic factors as they may suggest a trend in the difference between tissue that has or has not been exposed to these specific factors.

Table 0-1: The following Table lists the threshold for the selected classifier chosen in Stage I for each extracted optical property. The measured value will be classified as malignant if it meets the criteria described by the inequality.

<table>
<thead>
<tr>
<th></th>
<th>HB Sat</th>
<th>B-Carotene</th>
<th>De-Oxy HB</th>
<th>Oxy HB</th>
<th>Total HB</th>
<th>Mean Absorption</th>
<th>Mean Scattering</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Data</td>
<td>&lt;88.51</td>
<td>&lt;15.35</td>
<td>≥5.88</td>
<td>≥33.12</td>
<td>≥40.41</td>
<td>≥4.70</td>
<td>≥8.64</td>
</tr>
<tr>
<td>Chemo Naïve</td>
<td>&lt;92.72</td>
<td>≥18.38</td>
<td>≥2.37</td>
<td>≥33.27</td>
<td>≥40.24</td>
<td>≥4.71</td>
<td>≥7.95</td>
</tr>
<tr>
<td>Post-Menopausal</td>
<td>&lt;89.61</td>
<td>&lt;15.45</td>
<td>≥5.95</td>
<td>≥38.01</td>
<td>≥40.37</td>
<td>≥4.69</td>
<td>≥7.59</td>
</tr>
</tbody>
</table>

These values can be compared against selected threshold values for other algorithms used for this application. Generally, the classifier can indicate at what concentration of a given optical property constitutes cancerous tissue. The either positive or negative change in thresholds across demographic status provides an
indication of how general tissue composition changes given malignancy and how it may effect detection. The figure below illustrates the difference between the threshold for the classifier for the optical property for all the data and the threshold for the classifier of either demographic status. In the figure below, if the marker falls on the dotted line, it indicates no difference between the entire data set and the stratified data set. If the marker falls below the dotted line, it indicates that the threshold value decreases as a result of stratification. If the marker falls above the dotted line it indicates that the threshold value increases as a result of stratification.

Figure 0-A: The change in threshold values for the classifier of each extracted optical property for each demographic status

For saturated hemoglobin in both demographic factors the threshold value decreases, however it decreases most for chemo naive sites. This suggests that there is a lower concentration of saturated hemoglobin in the malignant breast tissue of women who are chemo –naive, the trend is similar for post-menopausal patients however not to the same extent.
Post-menopausal sites seem to have no effect on the concentration of \( \beta \)-carotene in malignant tissues, however chemo naïve sites seem to have an odd and dramatic affect and may suggest that chemo naïve patients have much higher concentration of \( \beta \)-carotene.

The differences in deoxygenated hemoglobin could argue that chemo naïve tissue has a lower concentration of \( \beta \)-carotene.

Lastly, for oxygenated hemoglobin, the threshold value increases slightly for post-menopausal sites. This suggests that the malignant tissues for post-menopausal women have higher oxygenated hemoglobin.
Appendix B: Validation Box Plots

The following plots show the validation results for the post-menopausal data sets for diffuse reflectance, scattering and absorption. The post-menopausal status for the absorption spectrum had the best validation results.
Figure 0-2

Validation Results Absorption Spectra Stratified By Menopausal Status 1

Performance Validation for Operating Points Metric

Performance Validation for AUC Metric

Figure 0-3

Validation Results Diffuse Reflectance Stratified By Menopausal Status 1

Performance Validation for Operating Points Metric

Performance Validation for AUC Metric
The following ROC curves are the corresponding group-wise fusion curves for each post-menopausal stratified data set.
Figure 0-7

The following figures show the validation results for the chemotherapy naïve data sets for scattering, diffuse reflectance and absorption.
Figure 0-8

Figure 0-9
Figure 0-10
ROC Decision Fusion of Wavelengths

Area Under Curve = 0.94079

Performance [0.9 0.82; 0.8 0.92; 0.7 0.95]

Absorption Stratified by Post-Menopausal
numPos: 30   numNeg: 426

PD (Sensitivity)

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

PF (1 - Specificity)

502.5 507.5 457.5 505 452.5 450 500 510 512.5 470

Decision Fusion
502.5 nm
507.5 nm
457.5 nm
505 nm
452.5 nm
450 nm
500 nm
510 nm
512.5 nm
470 nm
ROC Decision Fusion of Wavelengths

Area Under Curve = 0.91178
Performance [0.9 0.73; 0.8 0.86; 0.7 0.92]

Diffuse Reflectance stratified by Post-Menopausal
numPos: 34 numNeg: 422

- Decision Fusion
- 495 nm
- 552.5 nm
- 492.5 nm
- 550 nm
- 547.5 nm
- 490 nm
- 487.5 nm
- 490 nm
- 580 nm
- 575 nm
- 540 nm
ROC Decision Fusion of Wavelengths

Area Under Curve = 0.94079
Performance[0.9 0.82; 0.8 0.92; 0.7 0.95]
Absorption Stratified by Post-Menopausal
numPos: 30 numNeg: 426

PD (Sensitivity) vs PF (1 - Specificity)

- Decision Fusion
- 502.5 nm
- 507.5 nm
- 457.5 nm
- 505 nm
- 452.5 nm
- 450 nm
- 500 nm
- 510 nm
- 512.5 nm
- 470 nm
Appendix C: Stage Three ROC Curves
The follow figures show the results obtained from including only saturated hemoglobin, β-carotene, de-oxygenated hemoglobin, oxygenated hemoglobin and total hemoglobin as the extracted optical properties in the stratified data sets to remove any bias from using a type of raw spectral data source.
Figure 0-A

Figure 0-B
Figure 0-C

Figure 0-D
References


2682.


