Micro-Anatomical Quantitative Imaging Towards Enabling Automated Diagnosis of Thick Tissues at the Point of Care

by

Jenna Lynne Hook Mueller

Department of Biomedical Engineering
Duke University

Date:_______________________
Approved:

___________________________
Nimmi Ramanujam, Supervisor

___________________________
Mark W. Dewhirst

___________________________
G. Allan Johnson

___________________________
David G. Kirsch

___________________________
Greg M. Palmer

___________________________
Rebecca M. Willett

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biomedical Engineering in the Graduate School of Duke University

2015
ABSTRACT

Micro-Anatomical Quantitative Imaging Towards Enabling Automated Diagnosis of Thick Tissues at the Point of Care

by

Jenna Lynne Hook Mueller

Department of Biomedical Engineering
Duke University

Date:_______________________
Approved:

___________________________
Nimmi Ramanujam, Supervisor

___________________________
Mark W. Dewhirst

___________________________
G. Allan Johnson

___________________________
David G. Kirsch

___________________________
Greg M. Palmer

___________________________
Rebecca M. Willett

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biomedical Engineering in the Graduate School of Duke University

2015
Abstract

Histopathology is the clinical standard for tissue diagnosis. However, histopathology has several limitations including that it requires tissue processing, which can take 30 minutes or more, and requires a highly trained pathologist to diagnose the tissue. Additionally, the diagnosis is qualitative, and the lack of quantitation leads to possible observer-specific diagnosis. Taken together, it is difficult to diagnose tissue at the point of care using histopathology.

Several clinical situations could benefit from more rapid and automated histological processing, which could reduce the time and the number of steps required between obtaining a fresh tissue specimen and rendering a diagnosis. For example, there is need for rapid detection of residual cancer on the surface of tumor resection specimens during excisional surgeries, which is known as intraoperative tumor margin assessment. Additionally, rapid assessment of biopsy specimens at the point-of-care could enable clinicians to confirm that a suspicious lesion is successfully sampled, thus preventing an unnecessary repeat biopsy procedure. Rapid and low cost histological processing could also be potentially useful in settings lacking the human resources and equipment necessary to perform standard histologic assessment. Lastly, automated interpretation of tissue samples could potentially reduce inter-observer error, particularly in the diagnosis of borderline lesions.
To address these needs, high quality microscopic images of the tissue must be obtained in rapid timeframes, in order for a pathologic assessment to be useful for guiding the intervention. Optical microscopy is a powerful technique to obtain high-resolution images of tissue morphology in real-time at the point of care, without the need for tissue processing. In particular, a number of groups have combined fluorescence microscopy with vital fluorescent stains to visualize micro-anatomical features of thick (i.e. unsectioned or unprocessed) tissue. However, robust methods for segmentation and quantitative analysis of heterogeneous images are essential to enable automated diagnosis. Thus, the goal of this work was to obtain high resolution imaging of tissue morphology through employing fluorescence microscopy and vital fluorescent stains and to develop a quantitative strategy to segment and quantify tissue features in heterogeneous images, such as nuclei and the surrounding stroma, which will enable automated diagnosis of thick tissues.

To achieve these goals, three specific aims were proposed. The first aim was to develop an image processing method that can differentiate nuclei from background tissue heterogeneity and enable automated diagnosis of thick tissue at the point of care. A computational technique called sparse component analysis (SCA) was adapted to isolate features of interest, such as nuclei, from the background. SCA has been used previously in the image processing community for image compression, enhancement, and restoration, but has never been applied to separate distinct tissue types in a
heterogeneous image. In combination with a high resolution fluorescence microendoscope (HRME) and a contrast agent acriflavine, the utility of this technique was demonstrated through imaging preclinical sarcoma tumor margins. Acriflavine localizes to the nuclei of cells where it reversibly associates with RNA and DNA. Additionally, acriflavine shows some affinity for collagen and muscle. SCA was adapted to isolate acriflavine positive features or APFs (which correspond to RNA and DNA) from background tissue heterogeneity. The circle transform (CT) was applied to the SCA output to quantify the size and density of overlapping APFs. The sensitivity of the SCA+CT approach to variations in APF size, density and background heterogeneity was demonstrated through simulations. Specifically, SCA+CT achieved the lowest errors for higher contrast ratios and larger APF sizes. When applied to tissue images of excised sarcoma margins, SCA+CT correctly isolated APFs and showed consistently increased density in tumor and tumor + muscle images compared to images containing muscle. Next, variables were quantified from images of resected primary sarcomas and used to optimize a multivariate model. The sensitivity and specificity for differentiating positive from negative ex vivo resected tumor margins was 82% and 75%. The utility of this approach was further tested by imaging the in vivo tumor cavities from 34 mice after resection of a sarcoma with local recurrence as a bench mark. When applied prospectively to images from the tumor cavity, the sensitivity and specificity for differentiating local recurrence was 78% and 82%. The results indicate that SCA+CT can
accurately delineate APFs in heterogeneous tissue, which is essential to enable automated and rapid surveillance of tissue pathology.

Two primary challenges were identified in the work in aim 1. First, while SCA can be used to isolate features, such as APFs, from heterogeneous images, its performance is limited by the contrast between APFs and the background. Second, while it is feasible to create mosaics by scanning a sarcoma tumor bed in a mouse, which is on the order of 3-7 mm in any one dimension, it is not feasible to evaluate an entire human surgical margin. Thus, improvements to the microscopic imaging system were made to (1) improve image contrast through rejecting out-of-focus background fluorescence and to (2) increase the field of view (FOV) while maintaining the sub-cellular resolution needed for delineation of nuclei. To address these challenges, a technique called structured illumination microscopy (SIM) was employed in which the entire FOV is illuminated with a defined spatial pattern rather than scanning a focal spot, such as in confocal microscopy.

Thus, the second aim was to improve image contrast and increase the FOV through employing wide-field, non-contact structured illumination microscopy and optimize the segmentation algorithm for new imaging modality. Both image contrast and FOV were increased through the development of a wide-field fluorescence SIM system. Clear improvement in image contrast was seen in structured illumination images compared to uniform illumination images. Additionally, the FOV is over 13X
larger than the fluorescence microendoscope used in aim 1. Initial segmentation results 
of SIM images revealed that SCA is unable to segment large numbers of APFs in the 
tumor images. Because the FOV of the SIM system is over 13X larger than the FOV of the 
fluorescence microendoscope, dense collections of APFs commonly seen in tumor 
images could no longer be sparsely represented, and the fundamental sparsity 
assumption associated with SCA was no longer met. Thus, an algorithm called 
maximally stable extremal regions (MSER) was investigated as an alternative approach 
for APF segmentation in SIM images. MSER was able to accurately segment large 
numbers of APFs in SIM images of tumor tissue. In addition to optimizing MSER for 
SIM image segmentation, an optimal frequency of the illumination pattern used in SIM 
was carefully selected because the image signal to noise ratio (SNR) is dependent on the 
grid frequency. A grid frequency of 31.7 mm\(^{-1}\) led to the highest SNR and lowest 
percent error associated with MSER segmentation.

Once MSER was optimized for SIM image segmentation and the optimal grid 
frequency was selected, a quantitative model was developed to diagnose mouse 
sarcoma tumor margins that were imaged \textit{ex vivo} with SIM. Tumor margins were 
stained with acridine orange (AO) in aim 2 because AO was found to stain the sarcoma 
tissue more brightly than acriflavine. Both acriflavine and AO are intravital dyes, which 
have been shown to stain nuclei, skeletal muscle, and collagenous stroma. A tissue-type 
classification model was developed to differentiate localized regions (75x75 µm) of
tumor from skeletal muscle and adipose tissue based on the MSER segmentation output. Specifically, a logistic regression model was used to classify each localized region. The logistic regression model yielded an output in terms of probability (0-100%) that tumor was located within each 75x75 µm region. The model performance was tested using a receiver operator characteristic (ROC) curve analysis that revealed 77% sensitivity and 81% specificity. For margin classification, the whole margin image was divided into localized regions and this tissue-type classification model was applied. In a subset of 6 margins (3 negative, 3 positive), it was shown that with a tumor probability threshold of 50%, 8% of all regions from negative margins exceeded this threshold, while over 17% of all regions exceeded the threshold in the positive margins. Thus, 8% of regions in negative margins were considered false positives. These false positive regions are likely due to the high density of APFs present in normal tissues, which clearly demonstrates a challenge in implementing this automatic algorithm based on AO staining alone.

Thus, the third aim was to improve the specificity of the diagnostic model through leveraging other sources of contrast. Modifications were made to the SIM system to enable fluorescence imaging at a variety of wavelengths. Specifically, the SIM system was modified to enabling imaging of red fluorescent protein (RFP) expressing sarcomas, which were used to delineate the location of tumor cells within each image. Initial analysis of AO stained panels confirmed that there was room for improvement in tumor detection, particularly in regards to false positive regions that were negative for
RFP. One approach for improving the specificity of the diagnostic model was to investigate using a fluorophore that was more specific to staining tumor. Specifically, tetracycline was selected because it appeared to specifically stain freshly excised tumor tissue in a matter of minutes, and was non-toxic and stable in solution. Results indicated that tetracycline staining has promise for increasing the specificity of tumor detection in SIM images of a preclinical sarcoma model and further investigation is warranted.

In conclusion, this work presents the development of a combination of tools that is capable of automated segmentation and quantification of micro-anatomical images of thick tissue. When compared to the fluorescence microendoscope, wide-field multispectral fluorescence SIM imaging provided improved image contrast, a larger FOV with comparable resolution, and the ability to image a variety of fluorophores. MSER was an appropriate and rapid approach to segment dense collections of APFs from wide-field SIM images. Variables that reflect the morphology of the tissue, such as the density, size, and shape of nuclei and nucleoli, can be used to automatically diagnose SIM images. The clinical utility of SIM imaging and MSER segmentation to detect microscopic residual disease has been demonstrated by imaging excised preclinical sarcoma margins. Ultimately, this work demonstrates that fluorescence imaging of tissue micro-anatomy combined with a specialized algorithm for delineation and quantification of features is a means for rapid, non-destructive and automated detection
of microscopic disease, which could improve cancer management in a variety of clinical scenarios.
Dedication

This work is dedicated to all who have lost their battle to cancer and to all who have survived.
# Contents

Abstract ........................................................................................................ iv

Contents ........................................................................................................ xiii

List of Tables .................................................................................................. xviii

List of Figures ................................................................................................ xix

Symbols and Abbreviations .......................................................................... xxix

Acknowledgements ........................................................................................ xxxii

1. Background and Significance .................................................................... 1

1.1 Limitations of Histopathology ................................................................. 1

1.2 Clinical Implications ............................................................................... 2

1.2.1 Intraoperative Tumor Margin Assessment ............................................ 2

1.2.2 Rapid Assessment of Biopsy Specimens ............................................. 5

1.2.3 Low Cost Histological Processing for Rapid Tissue Diagnosis in Low Resource Settings ................................................................. 7

1.2.4 Automated Interpretation of Tissue Samples Could Reduce Inter-Observer Error ..................................................................................... 8

1.3 Potential Biophotonic Solutions ............................................................... 9

1.3.1 Optical Microscopy ........................................................................... 11

1.4 Need for Automated Diagnosis ............................................................... 14

1.4.1 Tissue Heterogeneity Presents a Challenge for Quantification .......... 15

1.4.2 Current Techniques for Nuclei Segmentation .................................... 15

1.5 Specific Aims and Organization of Chapters .......................................... 19
2. Automated Diagnosis.................................................................................................................. 21
  2.1 Rationale .................................................................................................................................. 21
    2.1.1 Sparse Component Analysis (SCA) .................................................................................. 21
  2.2 Methods ................................................................................................................................... 24
    2.2.1 Ethics Statement .............................................................................................................. 24
    2.2.2 Mice and Sarcoma Generation ......................................................................................... 24
    2.2.3 Imaging System and Contrast Agent ............................................................................... 24
    2.2.4 Ex Vivo Imaging of Excised Tissue Margins .................................................................. 25
    2.2.5 In Vivo Imaging of the Resected Tumor Cavity ............................................................... 27
    2.2.6 Generation of Simulated Sarcoma Images ...................................................................... 28
    2.2.7 Sparse Component Analysis and Circle Transform (SCA+CT) ...................................... 29
    2.2.8 Regularization Parameter Selection ............................................................................... 32
    2.2.9 Calculation of AFP Variables ......................................................................................... 32
    2.2.10 Model Development with Ex Vivo Data Set .................................................................. 33
    2.2.11 Application of Optimized Models to the In Vivo Data Set ............................................ 34
  2.3 Results ..................................................................................................................................... 34
    2.3.1 Illustration of Challenges Associated with APF Segmentation in Heterogeneous Tissue ......................................................................................................................... 34
    2.3.2 Sensitivity of SCA+CT to Background Heterogeneity, APF Size and Density .......... 36
    2.3.3 Trends Observed in APF Size and Density Quantified Using SCA+CT ................. 40
    2.3.4 Optimization of SCA+CT and Logistic Regression on Excised Tumor Margins .......... 44

xvi
2.3.5 Prediction of Local Recurrence Using an Optimized Algorithm Based on SCA+CT and Logistic Regression.................................................................52

2.4 Discussion........................................................................................................59

3. Improve Contrast..................................................................................................65

3.1 Rationale ............................................................................................................65

3.1.1 Contrast Affects Image Processing Performance........................................65

3.1.2 Current Techniques for Background Rejection.........................................65

3.2 Development of SIM System for Contrast Improvement.............................67

3.2.1 Methods .......................................................................................................67

3.2.1.1 SIM Theory............................................................................................67

3.2.1.2 SIM System and Contrast Agent .........................................................68

3.2.1.3 Tissue Phantom to Demonstrate Contrast Improvement .....................71

3.2.1.4 Ex Vivo Imaging of Excised Tissue Margins........................................72

3.2.1.5 Applying SCA to SIM images ...............................................................73

3.2.2 Results .........................................................................................................73

3.2.2.1 SIM Contrast Improvement Demonstrated in Tissue Phantoms ............73

3.2.2.2 SIM Contrast Improvement Demonstrated in Ex Vivo Images...............74

3.2.2.3 Applying SCA to Representative SIM images .....................................77

3.3 Optimizing SIM Image Segmentation.............................................................79

3.3.1 Methods .......................................................................................................80

3.3.1.1 Imaging Systems ....................................................................................80

3.3.1.2 Tissue Mimicking Phantom Preparation and Imaging .......................83
3.3.1.3 Signal to Background Ratio (SBR) Analysis .................................................. 84
3.3.1.4 Maximally Stable Extremal Regions (MSER) ............................................. 84
3.3.1.5 Percent Error Calculations ........................................................................... 87
3.3.1.6 Tissue Imaging and Analysis ....................................................................... 88
3.3.2 Results ............................................................................................................ 90
3.3.2.1 Applying MSER and SCA +CT Algorithms to Tissue Mimicking Phantom Images.................................................................................................................. 90
3.3.2.2 MSER Segmentation Analysis of Individual Beads Imaged with the HRME and SIM Microscopes ......................................................................................... 92
3.3.2.3 Comparing MSER Algorithm Performance to Optical Section Thickness. 93

3.4 Quantitative model development for SIM ............................................................ 99
3.4.1 Methods ........................................................................................................... 100
3.4.1.1 Margin Imaging Protocol ............................................................................. 100
3.4.1.2 Tissue Type Classification Model ................................................................. 103
3.4.1.3 Logistic Regression ...................................................................................... 104
3.4.2 Results ............................................................................................................ 106
3.4.2.1 Tissue Type Classification Model Results .................................................... 106
3.4.2.2 Application to Full Margin Images ............................................................... 111

3.5 Discussion ......................................................................................................... 115

4. Improve Specificity ............................................................................................. 121
4.1 Rationale ........................................................................................................... 121
4.1.1 Difficult to Identify Location of Normal Tissues ............................................ 121
4.1.2 Fluorophores for Multi-Spectral Imaging .................................................... 121
4.1.3 RFP Sarcoma Model Provides a Region Specific Diagnosis ........................................ 124

4.2 Methods .................................................................................................................. 125

4.2.1 Mice and Sarcoma Generation ............................................................................ 125

4.2.2 Modifications to SIM system .............................................................................. 126

4.2.3 Ex Vivo Imaging of Excised Sarcoma Margins .................................................... 127

4.3 Results .................................................................................................................. 129

4.3.1 RFP Sarcoma Model Provides a Region Specific Diagnosis .............................. 129

4.3.2 Can Multi-Spectral Properties of AO be Used to Improve Algorithm Specificity? .................................................................................................................. 132

4.3.3 Can Tetracycline Staining be Used to Improve Algorithm Specificity?........... 136

4.4 Discussion .............................................................................................................. 142

5. Conclusions and Future Directions ......................................................................... 147

5.1 Conclusions ........................................................................................................... 147

5.2 Future Work .......................................................................................................... 153

5.2.1 Tetracycline Preclinical Validation Study .......................................................... 154

5.2.2 SIM Technical Improvements ........................................................................... 158

5.2.3 Clinical Translation ............................................................................................ 159

5.2.3.1 Micro-Anatomical Quantitative Confocal Microscopy Towards Automated Assessment of Clinical Breast Tissues ................................................. 161

5.2.3.2 Combining SIM with Optical Spectral Imaging for Wide-Field Assessment of Breast Tumor Margins ........................................................... 190

References .................................................................................................................. 192

Biography .................................................................................................................... 209
List of Tables

Table 1: Biophotonic techniques that enable real-time visualization of tissue .................. 10
Table 2: Types of fluorescence microscopy studies .......................................................... 13
Table 3: Nuclear segmentation methods .............................................................................. 18
Table 4: Pathology and local recurrence endpoints for all margins .................................. 55
Table 5: Number of true positives, false negatives, true negatives, and false positives associated with the best performing model ................................................................. 58
Table 6: Properties of fluorescent microscopy systems ....................................................... 81
Table 7: Optical section thickness for each grid frequency .................................................. 83
Table 8: Model coefficients calculated using logistic regression for each combination of the three input variables ........................................................................................................ 109
Table 9: Fluorophore considerations .................................................................................... 123
Table 10: Summary of patients from which tissue specimens were acquired, sites analyzed, and histologic diagnoses ................................................................................................ 173
Table 11: Summary of top performing parameters for distinguishing between benign and malignant sites measured using algorithms .............................................................. 178
Table 12: Performance of model for classification of neoplasia, non-neoplasia, and individual histologic types of breast neoplasia ........................................................................ 183
List of Figures

Figure 1: Map of specific aims included in this work.......................................................... 20

Figure 2: Detailed schematic of the HRME imaging system.............................................. 25

Figure 3: Image preprocessing. An original image of a tumor + muscle site is shown in (a). (b) The image is first cropped to remove the rim of the fiber bundle. (c) Next a low pass Gaussian filter is applied to remove the fiber cores that are superimposed on the image. The image displayed in (c) is the input into the SCA algorithm................................. 29

Figure 4: Illustration of challenges associated with APF segmentation in heterogeneous tissue. The original image was created by blending different regions of a tumor (T), muscle (M), adipose (A), and tumor + muscle (T+M) image together. This image was put through SCA to yield the spatial, DCT, and curvelet outputs. Then, the spatial image was false colored green and laid on top of the curvelet image to yield the spatial curvelet overlay. This illustrates how APFs are spatially co-registered with some of the features isolated in the curvelet output. This information could be used diagnostically to indicate that these APFs are associated with benign adipose tissue, not with malignant tissue. .. 36

Figure 5: Image simulations used to evaluate SCA. Simulations of tumor (a), muscle (b), and tumor + muscle (c) were mimicked after observations seen in H&E sections (d-f). (g) To better understand how well SCA can capture APFs of various sizes and densities, a single simulated image was created that contained a variety of sizes and densities. The red box indicates the most commonly seen sizes and densities in our data. The outputs of SCA for a tumor + muscle simulation are shown (h). The colorbar shows the gray level intensities, which vary from 0 to 1. In this case, the curvelet image appears blank because no adipose is present in the original image. A zoomed in area of the upper right hand corner of the original image is shown. The same area in the spatial image is shown with CT applied to yield SCA+CT. Scale bar 50 µm......................................................... 38

Figure 6: SCA performance varies with APF size, APF density, and background contrast. (a) The image simulation with various sizes and densities of APFs was added to the muscle simulation with varying weights or ratios to create a tumor + muscle simulation. The ratios are reported as the (max tumor APF intensity/ max muscle intensity). The spatial output from SCA is shown in (b). The density was calculated by applying CT to the images in (b) to yield the images in (c). Red boxes indicate the most commonly seen APF size, APF density, and contrast ratio observed in our data. The percent error was
calculated and is shown as a contour plot in (d). The colorbar indicates the percent error (%). .................................................................39

Figure 7: SCA+CT applied to a panel of tissue section images illustrates trends in APF density. The specific locations where HRME images (Original, (b)) were collected are indicated by squares in the H&E section (H&E, (a)). APFs were isolated and quantified by using SCA+CT, which is illustrated in (d). For SCA+CT, APFs are overlaid onto the original image. APFs that are greater than 7 µm in diameter are false colored red, while those that are equal to or less than 7 µm are false colored green. The contrast of the original panel was adjusted in (c) to match the SCA+CT overlay in (d) in order to enable direct visual comparison. Scale bar 200 µm. The diameter (e), total density (f), density of the smaller APFs (green, (g)) and density of the larger APFs (red, (h)) were quantified for the 3 tumor images (images 3, 6, 9), 3 muscle images (images 1, 4, 7), and 3 tumor + muscle images (images 2, 5, 8). For diameter, the parenthetical values indicate the number of APFs whose diameter was quantified. For density, the parenthetical values indicate the number of images in which the density was calculated. In (e) the vertical dotted red line corresponds to 7 µm diameter, and the horizontal color bars show the mean and standard deviation for each variable.........................................................44

Figure 8: SCA+CT yields differences between ex vivo tissue types. Representative H&E images (column 1) of tumor, tumor + muscle (T+M), muscle, and adipose tissue and their corresponding HRME image (column 2) are shown (a)-(d) respectively. APFs were isolated and quantified by using SCA+CT and were overlaid onto the original image, which is illustrated in column 4. APFs that are greater than 7 µm in diameter are false colored red, while those that are equal to or less than 7 µm are false colored green. In column 3, the contrast of the original panel was adjusted to match the SCA+CT overlay in column 4. Scale bar 200 µm. The probability distribution functions (pdf) of the diameters from the four images are shown in (e). The parenthetical values indicate the number of APFs in each image which is synonymous to density, and the vertical dotted red line corresponds to a diameter of 7 µm. .................................................................46

Figure 9: APF variables calculated for a cohort of ex vivo sarcoma margin images. Density and mean diameter were calculated from 8 tumor, 6 tumor + muscle (T+M), and 13 muscle images. For density, boxplots were created for all APFs, the smaller features (green) and the larger features (red) are shown in (a), (c), and (d) respectively. A boxplot of the mean diameter of all APFs is shown in (b). P values calculated from Wilcoxon rank sums are shown in each boxplot.........................................................................48
Figure 10: Original H&E stained sections of sarcoma, inflammation, muscle, and adipose tissues are shown in (a). Scale bar 100 µm. Sarcoma, muscle and adipose nuclei are outlined in green. The inflammation image has two primary cell types – macrophages and lymphocytes. Macrophage nuclei are outlined in green and lymphocyte nuclei are outlined in yellow. The smaller features in the sarcoma and inflammation image, which correspond to sarcoma and macrophage nucleoli respectively, are outlined in blue. The average major and minor axis of the nuclei or nucleoli circled in each image are reported in (b).

Figure 11: Multivariate models developed based on ex vivo sarcoma margins. The receiver operator curves for all 2 variable and 3 variable combinations are shown in (a) - (d). Each plot contains curves associated with the original model as well as with cross-validation. The area under the curve (AUC) and the area under the curve associated with the cross-validation (Crossval AUC) are shown on each plot. The density of the smaller features (green), the density of the larger features (red), and the density of all APFs are referred to as Density (green), Density (red), and Density (both) respectively.

Figure 12: The application of an algorithm based on sparse component analysis (SCA), and circle transform (CT) applied to representative in vivo images from the tumor cavity after resection of the sarcoma. Overlays of a LR+/Path+ and LR-/Path- tumor cavities are shown in (a) and (b) respectively. The overlays were contrast-stretched in order to enable increased visibility of the false colored APFs. Scale bar is 200 µm. The probability distribution functions (pdf) of the APF diameters from the two tumor cavities are shown in (c). The parenthetical values indicate the number of APFs in each panel of images, and the vertical dotted red line corresponds to a diameter of 7 µm.

Figure 13: Multivariate models applied to the in vivo images from the tumor cavity predict local recurrence. The receiver operator characteristic curves associated with all of the models shown in Fig 11 are applied here. Each plot contains curves associated with all margins as well as with matching LR+/Path+ and LR-/Path- margins. The area under the curve for all margins (AUC all margins) and the area under the curve for matching LR+/Path+ and LR-/Path- margins (AUC matched LR/Path margins) are shown on each plot.

Figure 14. (a) Detailed schematic of the SIM imaging system. The λex peak was 480 nm and the λem peak was 520 nm. The spatial filter diameter was adjusted to allow only the 0 and +1 diffraction orders to pass. (b) and (c) show a comparison between the HRME FOV and the SIM FOV.
Figure 15: Images of a single layer of 10 µm fluorescent spheres embedded in PDMS and TiO$_2$. The reduced scattering coefficient of the phantom is approximately 10 cm$^{-1}$. The signal-to-background was calculated by taking the intensity of 5 manually selected spheres (indicated by red arrows) and dividing by the background ROI (indicated by the blue square). All scale bars are 100 µm.

Figure 16: Uniform and structured illumination images acquired from mouse tissue. (a) Images of skeletal muscle from mouse. Both longitudinal and cross-sectional muscle can be seen in the region of interest. (b) Image of tumor tissue from mouse sarcoma. Contrast enhancement is clearly seen in the sectioned images. Scale bars are 100 µm.

Figure 17: SCA applied to representative ex vivo SIM images. Representative sectioned images (column 1) of tumor, muscle, and adipose tissue are shown in (a)-(c) respectively. Zoomed in regions of interest (ROIs), which are outlined in red in column 1, are shown in column 2. APFs were isolated by using SCA and were overlaid onto the original ROIs, which is illustrated in column 3. All scale bars are 100 µm.

Figure 18: (a) Overview of the maximally stable extremal regions (MSER) algorithm. This algorithm, which is used in the automatic reconstruction of 3D scenes was selected as it was a more robust intensity thresholding approach. MSER tests all intensities thresholds within the image to ensure local maxima within various background intensities are not lost. (b) A detailed description and visual representation of the five variable parameters used in MSER, MinArea, MaxArea, Delta, MaxVariation, and MinDiversity.

Figure 19: Application of MSER and SCA+CT to HRME images. Representative HRME images of 1, 6, and 10 µm tissue mimicking phantoms are shown in (a) columns 1-3. Original images were analyzed with MSER and SCA+CT and overlays are shown in (b) and (c) respectively. Scale bar 200 µm. Boxplots of the approximated beads diameters obtained with MSER and SCA+CT are shown in (d) and (e) respectively. The red horizontal lines indicate the true bead sizes of 1, 6, and 10 µm.

Figure 20: Application of MSER and SCA+CT to SIM images. Representative SIM images of 1, 6, and 10 µm tissue mimicking phantoms are shown in (a) columns 1-3. Original images were analyzed with MSER and SCA+CT and overlays are shown in (b) and (c) respectively. Scale bar 200µm. Boxplots of the approximated beads diameters obtained with MSER and SCA+CT are shown in (d) and (e) respectively. The horizontal red lines indicate the true bead sizes of 1, 6, and 10 µm.
Figure 21: Intensity values associated with a line running through the center of one representative 1, 6, and 10 µm bead. Intensity plots for beads taken with the HRME and SIM microscopes are shown in (a) and (b) respectively. Images of the beads and the corresponding MSER overlay are shown on each plot. The edge of MSER segmentation for each bead is indicated by the vertical blue lines drawn on each plot. ..........................93

Figure 22: Relationship between MSER algorithm performance and optical section thickness. Average SBR values for 5 images of 10µm phantoms taken with the HRME and SIM system at different optical section thickness are shown as a boxplot in (a). A comparison of the average percent error associated with the MSER segmentation results and the optical section thickness is shown as a scatterplot in (b). ....................................94

Figure 23: Results demonstrating the impact of optical section thickness on quantification of APF density with MSER. Sectioned images of mouse sarcoma tissue at multiple illumination frequencies (a). The APFs isolated with MSER were false colored green and overlaid onto the original image (b). Mean density of APFs extracted from images mouse sarcoma tissue (c). The APF density extracted from the 50 µm thick frozen section images (N=9) was compared to density extracted from thick tissue imaging (N=5), both uniform and sectioned (84µm, 120µm, and 156µm) images. Scale bar 100µm.................................................................95

Figure 24: Flowchart demonstrating the imaging protocol immediately after the tumor margin was removed from the transgenic sarcoma model. After the procedure above was completed, the tissue sections were inked to identify the area imaged, stained with H&E and submitted to pathology. .................................................................102

Figure 25: Flow chart which demonstrates the quantitative analysis carried out on the site specific ROIs. The purpose of this analysis was to develop a predictive model to differentiate tumor tissue from muscle and adipose. A total of 30 images were used as a training dataset to develop the model. Each image was further divided into smaller bins to ensure very localized disease could be detected. Finally, three quantities were calculated for each bin, area fraction, average diameter, and average shape, and used as input variables for a logistic regression model.........................................................104

Figure 26: Application of MSER to both (a) structured and (b) uniform illumination representative images. It is evident that the structured illumination images provide a clear benefit with contrast enhancement which assists the MSER algorithm in accurately identifying APFs. (c) – (e) contain zoomed in images of tumor, muscle, and adipose tissue respectively. Zoomed in regions of H&E images are shown in column 1, structured illumination images are shown in column 2, and the MSER overlay is shown
in column 3. White and black arrows point to nuclei while red arrows point to features that were segmented with MSER but do not appear to correspond to nuclei. All scale bars are 100 µm.

Figure 27: Distribution functions of the three variables, area fraction, diameter, and shape. These include all 50x50 pixel bins from the training dataset, N=420 tumor bins and N=840 normal bins (420 muscle and 420 adipose). A Wilcoxon rank-sum test was used to compare the distributions and calculate the corresponding p-value. Based on the results, the tumor and normal distributions were statistically different for all three variables, so they were included as input to the logistic regression model.

Figure 28: The receiver operating characteristic (ROC) curves (top-left) were generated when logistic regression model was applied to the original training dataset of 30 site specific ROIs, 10 tumor sites vs. 20 normal sites (10 muscle and 10 adipose). The specific combinations of variables used as input to the logistic regression model are shown in the legend.

Figure 29: Representative images of a positive and negative margin. SIM images were analyzed with MSER and the segmented regions are highlighted in green. In addition, specific tissue types were identified by a pathologist and SIM images and labeled with T (tumor) and M (muscle). Scale bar 200 µm.

Figure 30: The results of the grid analysis for the positive and negative margin shown in the previous figure. The predicted tumor probability is represented spatially as heat maps. It is clear that the positive margin exhibits a much higher portion of bins with high tumor probability.

Figure 31: Detailed schematic of the multispectral SIM imaging system. Modifications are highlighted in red.

Figure 32: Excitation and emission spectra for AO1 and RFP are shown in green and yellow respectively. Excitation spectra are indicated by the dotted lines and emission spectra are indicated by solid lines.

Figure 33: Pure tumor and pure muscle tissues were imaged with a commercial microscope. Representative images of a pure tumor and muscle site are shown in rows 1 and 2 respectively. Images of RFP are shown in column 1. The tissue was then topically stained with AO and images were acquired of the same site (column 2). An overlay of the RFP and AO images are shown in column 3. RFP is false colored red and AO is false colored green for visualization purposes. Scale bar 100 µm.
Figure 34: An overlay of the RFP and AO images are shown in A. RFP is false colored red and AO is false colored green for visualization purposes. The zoomed in ROI is shown in B. Gray scale images of AO and RFP are shown separately in C and D respectively. Scale bar 100 µm.

Figure 35: A mosaic of AO and RFP images are shown in A and B respectively. RFP signal is outlined in red in B-D. MSER segmentation was applied to the panel in A to isolate AO positive features (APFs). APFs isolated with MSER were false colored green and overlaid onto the original images. The MSER overlay is shown in C. The logistic regression model was applied to the MSER panel. The predicted tumor probability (which varies from 0 to 100%) is represented spatially as a heat map in D. Scale bar 200 µm.

Figure 36: Multi-spectral properties of AO. (a) An tissue section image of tumor + muscle imaged with all emission wavelengths. (b) An overlay of 530 (green) and 630 nm (red) emission. (c) Corresponding H&E stained section. Scale bar 200 µm.

Figure 37: Mosaics of AO1, RFP, MSER overlays, and probability heat maps are shown in A-D respectively. E shows the AO2 signal that was detected. Areas that contained AO2 signal are outlined in blue. F contains the corrected probability heat map in which AO2 positive areas were set to zero. Scale bar 200 µm.

Figure 38: Comparison of AO1 and AO2 corrected probability heat maps. The percentage of bins in both positive and negative regions that have probabilities > 50% (or 0.5) were tabulated for AO1 heat maps (n = 5 mice). These variables are referred to as AO1+ for positive regions and AO1- for negative regions. AO2 was also imaged for 3 out of the 5 mice, and the percentage of bins in both positive and negative regions that have probabilities > 0.5 were tabulated for the AO2 heat maps. These variables are referred to as AO2+ for positive regions and AO2- for negative regions. A boxplot of these variables is shown here.

Figure 39: Representative images of tetracycline stained tissue sections. The corresponding H&E image is shown in A. The tumor section was outlined in B (the lower half of the panel contains tumor). The mosaic of tetracycline images is shown in C. Fluorescent positive features (FPFs) were isolated with MSER, false colored green, and overlaid onto the original tetracycline images. The MSER overlay is shown in D. Both C and D also show the tumor section outlined in red (the lower portion of the panel). E and F contain zoomed in images of tumor and muscle tissue respectively. Zoomed in regions of H&E images are shown in column 1, structured illumination images are
shown in column 2, and the MSER overlay is shown in column 3. All scale bars 200 µm.

Figure 40: Comparison of representative AO and tetracycline stained tissue sections. The corresponding H&E section is shown in A. The tumor region, which is in the lower left quadrant of the panel is outlined in B. The AO panel is shown in C and the tetracycline panel is shown in D. The tumor regions in C and D are outlined in red. All scale bars 200 µm.

Figure 41: Comparison of AO and tetracycline stained tumor and muscle regions of interest (ROIs). Representative tumor and muscle ROIs are shown in A and B respectively. Scale bar 100 µm. The tumor to muscle ratios were quantified from 3 sets of AO and tetracycline (tetra) stained ROIs. A boxplot of the tumor to muscle ratios is shown in C.

Figure 42: Comparison of AO and tetracycline tumor probability heat maps. The corresponding H&E section is shown in A. The tumor regions are outlined in B. The MSER overlay and probability heat map for the AO panel is shown in C and E respectively, while the MSER overlay and probability heat map for the tetracycline panel is shown in D and F respectively. The tumor regions in C-F are outlined in red. All scale bars 200 µm.

Figure 43: Comparison of AO and tetracycline tumor probability heat maps. The percentage of bins in both positive and negative regions that have probabilities > 50% (or 0.5) were tabulated for AO heat maps (n = 3 tissue sections). These variables are referred to as AO+ for positive regions and AO- for negative regions. Similarly, the percentage of bins in both positive and negative regions that have probabilities > 0.5 were tabulated for tetracycline heat maps. These variables are referred to as Tetra+ for positive regions and Tetra- for negative regions. A boxplot of these variables is shown here.

Figure 44: Illustration of RFP crosstalk in the blue channel. Two excised RFP margins were imaged. The first margin, which is shown in (a) and (b), was stained with tetracycline. The second margin, which is shown in (c) and (d) was not stained in order to demonstrate crosstalk in the blue channel (d). Scale bar 200 µm.

Figure 45: Spectral power density of the Fianium supercontinuum laser used in the SIM system. The particular model used in this work is the Fianium SC400-2, which is shown in red.
Figure 46: Excitation and emission spectra for AO and tetracycline (tetra) are shown in green and blue respectively. Excitation spectra are indicated by the dotted lines and emission spectra are indicated by solid lines. Both a dilute and concentrated solutions of tetracycline were measured to acquire the excitation spectra.

Figure 47: Algorithms for APF (A-D) and ductal (E-L) segmentation. APF segmentation: A: Raw image acquired from confocal fluorescence microscope with 750 x 750 µm² field of view. B: Region of interest selected in confocal fluorescence image with 75 x 75 µm² field of view. C: The maximally stable extremal regions (MSER) algorithm applies thresholds from 0 to 255 to B. D: At each threshold, the MSER algorithm identifies APFs as connected components and selects “maximally stable” components with the lowest size variation. Ductal segmentation: E: Raw image acquired from confocal fluorescence microscope with 750 x 750 µm² field of view. F: Wiener low pass filter and adaptive histogram equalization applied to E. G: The algorithm converts E to a binary image using an interactive threshold tool. H: Objects below range of APF area are removed and then user selects a region of interest (ROI) around ducts with an interactive polygon selection tool. I: The algorithm fills boundaries of ducts identified in H to segment the outer boundaries of the duct. J: The algorithm selects the complement of H to segment the inner boundaries of the duct (lumen). K: Duct wall width is measured by selecting the shortest distance from the outer to the inner duct boundaries (red lines). L: Ellipses are fitted to outer and inner duct boundaries. E-J: scale bar is 100 µm. K,L: scale bar is 25 µm.

Figure 48: Representative raw confocal fluorescence images of adipose tissue, fibrous tissue, lobules, invasive ductal carcinoma, and invasive lobular carcinoma are shown in A through E, respectively. F-J: APFs segmented by identifying maximally stable extremal regions (MSER) are false colored green and overlaid onto the raw confocal fluorescence image. K-O: Histologic slides with H&E staining show similar histology to confocal images in A-E. Slides were prepared with the same specimens from which confocal images were acquired. Scale bar is 100 µm.

Figure 49: Representative confocal images of normal, non-hyperplastic ducts (A), hyperplastic ducts (B), and ductal carcinoma in situ (C) analyzed with the APF segmentation algorithm (middle row) and with the ductal segmentation algorithm (bottom row). D-F: APFs segmented by identifying maximally stable extremal regions (MSER) are false colored green and overlaid onto the raw confocal fluorescence image. G-I: Breast ducts segmented using the ductal segmentation algorithm. J-L: Histologic slides with H&E staining show similar histology to confocal images in A-E. Slides were prepared with the same specimens from which confocal images were acquired.
prepared with the same specimens from which confocal images were acquired. Scale bar is 100 µm.

Figure 50: Mean value of parameters used to separate malignant from benign sites. APF parameters calculated with the APF segmentation algorithm are shown for all adipose, fibrous, lobules, invasive ductal carcinoma (IDC), and invasive lobular carcinoma (ILC) sites; A: standard deviation of IND; B: area fraction; C: range of IND. Ductal parameters calculated with the duct-based segmentation algorithm are shown for all normal, non-hyperplastic ducts, hyperplastic ducts (Hyperplasia), and ductal carcinoma in situ (DCIS); D: number of lumens; E: minor dimension of outer ellipse; F: area of outer ellipse. The number of sites represented in each box is represented by n. Significant differences between mean values of parameters measured at benign and malignant sites are indicated by asterisks (*).

Figure 51: Classification tree automatically generated when all APF and duct data was used. Duct- and APF-based parameters selected by classification regression tree analysis to optimize separation between benign and malignant sites. Bar graphs show the diagnoses of sites sorted into each classification category.

Figure 52: Representative images of sites with lowest classification accuracy in the decision tree model. A-D: invasive lobular carcinoma and lobules in confocal fluorescence images. E-H: APFs segmented by identifying maximally stable extremal regions (MSER) are false colored green and overlaid onto the raw confocal fluorescence image. I-L: Histologic slides with H&E staining show similar histology to confocal images in A-E. Slides were prepared with the same specimens from which confocal images were acquired. A, E: a true positive invasive lobular carcinoma (ILC) site; B, F: false negative ILC site; C, G: true negative lobules; and D, H: false positive lobules. Scale bar is 100 µm.
Symbols and Abbreviations

AF – area fraction

ALND – axillary lymph node dissection

AO – acridine orange

AO+ – acridine orange signal in positive tumor regions

AO- – acridine orange signal in negative normal tissue regions

AO1 – first emission peak of acridine orange

AO2 – second emission peak of acridine orange

APF – acriflavine or acridine orange positive feature

AUC – area under the curve

BCS – breast conserving surgery

CART – classification and regression tree

CCD – charge-coupled device

CT – circle transform

DCIS – ductal carcinoma in situ

DCT – discrete cosine transform

DRE – digital rectal exam

FM – fluorescence microendoscopy

FOV – field of view
FWHM – full-width half-maximum

GPSR – gradient projection for sparse reconstruction

H&E – hematoxylin and eosin

HRME – high resolution microendoscope

IDC – invasive ductal carcinoma

ILC – invasive lobular carcinoma

IND – internuclear distance

LED – light emitting diode

LMICs – low and middle income countries

LR+/Path+ – locally recurred and pathologically positive

LR+/Path- – locally recurred and pathologically negative

LR+/Path N/A – locally recurred and discordant pathology

LR-/Path- – recurrence free and pathologically negative

LR-/Path+ – recurrence free and pathologically positive

LR-/Path N/A – recurrence free and discordant pathology

MaxArea – maximum area

MaxVariation – maximum variation

MinArea – minimum area

MinDiversity – minimum diversity

MSER – maximally stable extremal regions
OCT – optical coherence tomography
OSI – optical spectral imaging
PDMS – polydimethylsiloxane
PBS – phosphate buffered saline
PSA – prostate-specific antigen
RFP – red fluorescent protein
ROC – receiver operator characteristic curve
ROI – region of interest
SBR – signal to background ratio
SCA – sparse component analysis
SCA+CT – sparse component analysis followed by the circle transform
SIM – structured illumination microscopy
SLM – spatial light modulator
SLNB – sentinel lymph node biopsy
SNR – signal to noise ratio
StdIND – standard deviation of internuclear distance
tetra – tetracycline
Tetra+ – tetracycline signal in positive tumor regions
Tetra- – tetracycline signal in negative normal tissue regions
UV-VIS – ultraviolet-visible

xxxi
Acknowledgements

First I’d like to thank my parents, Lynne and Wes Hook, for raising me in an environment that encouraged curiosity, dedication, independence, academic achievement, creativity, music, art, math and science. Without that foundation, none of this would have been achievable. I’d also like to thank my incredible siblings, James and Leah Hook, for joining me on many adventures (both real and imagined in our game room) and making me into a more well-rounded person. Next I’d like to thank my best friend and partner in life Collin Mueller. Thank you for so many evenings of encouragement on our porch and all of the amazing adventures we have had in the past ten years. To the rest of the Hook, Falke, and Mueller family, thank you for being amazing grandparents, aunts and uncles, cousins, and in-laws, and for taking an interest in my many nerdy exploits. To my many amazing friends, Jenny Becker, Elise Bugh, friends from childhood, Westlake High School, Rice University, Duke BME, Duke Divinity, the Duke Wesley Fellowship, Refuge Home Church, Six Plates, and American Village, thank you for your support, encouragement, and laughter. Thank you for teaching me to see the bigger picture and to enjoy the journey.

Next I’d like to thank the many amazing teachers and professors who have encouraged me along the way. To the late Mr. Furtado, thank you for teaching me to love biology, for sharing your passion for life, and for setting me on this path. To Mr. Mark Misage, thank you for literally and metaphorically shocking me with physics
lessons. You are a true example of how a teacher can make a remarkable difference in a student’s life and create a love of learning. To Dr. Ann Saterbak, thank you for encouraging me to stay in bioengineering, and for demonstrating how science, spirituality, and family can be passionately intertwined. To Dr. Maria Oden and Dr. Rebecca Richards-Kortum, thank you for seamlessly integrating global health and bioengineering and for providing incredibly enriching experiences to undergraduates. You have inspired so many students, and I am grateful to count myself as one of them. Next, I’d like to thank my Ph.D. advisor, Dr. Nimmi Ramanujam, for her scientific guidance, and for always pushing me to find the answer and to be a better scientist. Thank you for teaching me to strive for excellence, to be unafraid to ask big questions, and for encouraging me to be a part of the Global Women’s Health Technologies Center, whose mission resonates with me on so many levels. I’d also like to thank my collaborators and committee members, Dr. David Kirsch and Dr. Rebecca Willett, for their scientific guidance and for all of the time I’ve spent with them learning about medicine and image processing. To my other committee members, Dr. Mark Dewhirst, Dr. Allan Johnson, and Dr. Greg Palmer, I am grateful for your unbiased perspective and excellent feedback. To our other outstanding collaborators, Dr. Joseph Geradts, Dr. Diana Cardona, and Dr. Alaatin Erkanli, thank you for our insightful discussions and for taking the time to teach me about pathology and statistics. I’d also like to thank Dr.
Quincy Brown and Dr. Christine Mulvey for their incredible mentorship and life advice at many crucial junctures during my graduate school career.

Finally, I’d like to thank the great cast of characters from (and close to) the TOpS lab. To those graduate students who have gone before me, Dr. Torre Bydlon, Dr. Henry Fu, Dr. Justin Lo, Dr. Stephanie Kennedy, Dr. Stacy Millon, and Dr. Vivide Chang, thank you for answering my many questions and for helping me navigate the challenges of graduate school. To the current graduate students, Amy Frees, Fangyao Hu, Brandon Nichols, Chris Lam, Brian Crouch, Dan Salo, Mercy Asiedu, and Rob Morehard, thank you for your help with running experiments, analyzing data, and making the lab an entertaining and vibrant place to work. To Rhea Chitalia, thank you for being an incredible Pratt Fellow, for conducting countless experiments with me, for listening to my music, and for laughing at my jokes. To Marlee Krieger and Nada Hussain, thank you for helping me schedule countless meetings, navigate the IRB and audits, and for running the lab behind the scenes. To Jennifer Gallagher and Liz Wildermann, thank you for helping me navigate clinical studies and for spending many hours with me in cramped locations around the Duke ASC and Cancer Center. Lastly, to Dr. Jeff Mito, Melodi Whitley, and Dr. Zac Harmany, thank you for being incredible collaborators. I have enjoyed tasks that I never expected to because of your presence.
1. Background and Significance

1.1 Limitations of Histopathology

Histopathology is the clinical standard for tissue diagnosis. Pathologists examine high resolution images of small volumes of fixed, sectioned, and stained tissue [1]. When diagnosing cancer in particular, pathologists look for changes in tissue morphology including changes in nuclei and surrounding tissue. Nuclear changes that may indicate the presence of cancer include pleomorphism, which is variation in nuclear size and shape, increased nuclear-to-cytoplasmic ratio, increased nuclear density, decreased organization of nuclei, and increased mitotic rate [2, 3]. Changes in the surrounding tissue include the presence of reactive stroma, which is composed of connective tissue, blood vessels, macrophages, lymphocytes, and the presence of progressive infiltration, which involves the invasion and destruction of surrounding tissue [2, 3]

While histopathology is the gold standard, limitations include that it requires tissue processing, including sectioning and staining the tissue, which can take 30 minutes or more, and requires a highly trained pathologist to diagnose the tissue. Additionally, the diagnosis is qualitative and the lack of quantitation leads to possible observer specific diagnosis. Taken together, it is difficult to diagnose tissue at the point of care using histopathology.
1.2 Clinical Implications

Several clinical situations could benefit from more rapid and automated histological processing, which could reduce the time and the number of steps required between obtaining a fresh tissue specimen and rendering a diagnosis. For example, there is need for rapid detection of residual cancer on the surface of tumor resection specimens during excisional surgeries, which is known as intraoperative tumor margin assessment [4]. Additionally, rapid assessment of biopsy specimens at the point-of-care could enable clinicians to confirm that a suspicious lesion is successfully sampled, preventing an unnecessary repeat biopsy procedure. Rapid and low cost histological processing could also be potentially useful in settings lacking the human resources and equipment necessary to perform standard histologic assessment, which can be a challenge in many parts of the world [5]. Lastly, the diagnostic criteria used to classify suspicious lesions are qualitative and subjective, and inter-observer discordance has been shown to be a significant challenge in the diagnosis of borderline lesions [6]. Thus, automated interpretation of tissue samples could potentially reduce inter-observer error. Each of these clinical implications will be discussed in more detail below.

1.2.1 Intraoperative Tumor Margin Assessment

Most people with cancer will have some type of surgical resection, the goal of which is to achieve negative surgical margins, which are associated with lower rates of tumor recurrence [7]. For example, the goal of resection of soft tissue sarcomas located in
the extremity is to completely excise the tumor while preserving limb function. Typically, the completeness of the excision is determined by pathologic assessment several days after surgery. Positive margin status, which is indicated by the presence of tumor cells at the edge of the resected specimen, has been reported to correlate with local recurrence, development of metastasis, and overall survival [8-10]. In the absence of adjuvant radiation therapy, microscopic disease left at the surgical site causes local recurrence in up to 31% of sarcoma patients [9, 11].

Another example of tumor margin assessment is breast conserving surgery (BCS), which involves removal of malignant tissue with a surrounding margin of normal breast tissue. Post-operative histopathologic assessment of the resected specimen is the current gold standard by which microscopic residual tumor in the margin is detected. Post-operative positive pathologic margin status is an important predictor of local recurrence of an invasive or in-situ cancer after BCS [12, 13]. Re-excision surgery is performed if residual cancer is found within 2 mm of the surface of the excised lumpectomy specimen, in order to reduce the risk of recurrence [14]. Recently, meta-analyses by the Early Breast Cancer Trialists group showed that one death is averted for every four women in which a local recurrence is avoided. A review of the literature indicates that as many as 20-70% of BCS patients undergo re-excision surgery because the cancer was incompletely removed during the first BCS [4, 15-19]. This represents an enormous physical burden to the patient (increasing her chances for surgical
complications and/or eventual cancer-related mortality) and financial burden to the health care system (effectively doubling the cost of treatment for this group of patients). By 2015, it is expected that the number of patients undergoing BCS will rise from approximately 200,000 to more than 270,000 per year in the U.S., at an annual growth rate of 5.5% [20]. With no industry standard to prevent re-excision, it is expected that there will be a concomitant rise in the number of re-excision surgeries.

No tools for intra-operative margin assessment have been widely accepted across cancers. Intra-operative frozen section analysis and touch prep cytology are used to assess surgical margins at the time of first surgery at a few select high-volume centers with dedicated resources and personnel. However, these techniques have not been widely adopted because of (1) the need for pathologists in the operating room, (2) prolonged time during surgery for specimen processing (20-40 minutes), (3) technical challenges associated with processing fatty tissues, and (4) suboptimal sensitivity of gross pathologic evaluation to guide where to cut the sample for microscopic margin assessment (since the entire specimen cannot be practically sampled). Standard of care may involve techniques to macroscopically access margin status, such as intra-operative radiography, but generally the microscopic margin status is unknown until the specimen has been processed by pathology several days after surgery. Taken together, there exists an opportunity to improve intraoperative assessment of tumor margins.
1.2.2 Rapid Assessment of Biopsy Specimens

Rapid assessment of biopsy specimens at the point of care could enable clinicians to confirm that a suspicious lesion is successfully sampled, preventing an unnecessary repeat biopsy procedure. For example, the management of prostate cancer could benefit from rapid assessment of biopsy specimens. The current standard of care is the prostate-specific antigen (PSA) test, which assesses the concentration of PSA in the blood. PSA tests are usually accompanied by a digital rectal exam (DRE), which relies on physician palpation. If prostate specific antigen (PSA) and DRE tests are positive, diagnosis is confirmed by a biopsy of the prostate with 12 core needle samples taken from randomized locations. Prostate biopsies are usually performed blind or by using transrectal ultrasound to ensure that the prostate alone is sampled as opposed to other nearby tissues. Prostate biopsies are performed on 1 million men annually in the United States resulting in approximately 30% of these men being diagnosed with cancer [21]. Of the 1 million biopsies, approximately 700,000 come back negative; however, these men may still harbor prostate cancer [21]. Only 1 in 4 prostate biopsies are positive for prostate cancer, but due to the reportedly high false negative rate of biopsy results, which can be as high as 50%, repeat biopsies are commonly performed in men with elevated PSA test results [22, 23]. In the standard of care, 43% of men are referred to have 1 repeated biopsy, 44% of which have a second biopsy, and 43% of these have a third biopsy [24]. Taken together, current clinical practice misses many cancers, yet still subjects a large
patient population to unnecessary biopsy procedures. Even with frozen section analysis, it is not practical to freeze, stain, and section a large number of core needle biopsies (between 6 and 20 for prostate biopsy) within the timeframe of a typical biopsy procedure in the clinic. Thus there is a clinical need for a better method to detect prostate cancer at the point of care, which could be aided with rapid assessment of biopsy specimens to confirm that a suspicious lesion is successfully sampled.

Sentinel lymph node biopsy (SLNB) represents another clinical procedure in which rapid assessment of biopsy specimens could improve patient outcomes. Specifically, the presence or absence of axillary lymph node involvement represents one of the most important prognostic indicators of long-term patient outcome for breast cancer [25]. Similarly, accurate axillary staging is important for directing the appropriate selection of adjuvant therapies for breast cancer and guiding surgical management. SLNB is typically performed prior to axillary lymph node dissection (ALND) because SLNB is less morbid than ALND [25]. Consequently, patients with negative SLNB results can avoid ALND complications, such as upper extremity edema [25]. However, SLNB is associated with a false negative rate of 5 to 10% [26-28]. This high false negative rate is of clinical concern. Therefore, new approaches that can decrease the false negative rate associated with SLNB, such as rapid assessment of biopsy specimens to confirm that sentinel lymph nodes were successfully sampled, could improve breast cancer management.
1.2.3 Low Cost Histological Processing for Rapid Tissue Diagnosis in Low Resource Settings

Rapid and low cost histological processing could also be potentially useful in settings lacking the human resources and equipment necessary to perform standard histologic assessment [5]. For example, intermediate diagnostic biopsy is typically not performed between cancer screening and treatment in low and middle income countries (LMICs) due to the need for multiple visits (there is patient attrition with every clinic visit that is needed) and lack of resources. The number of pathologists in LMICs is small, even as a percentage of the total medical workforce. For example, there are only 15 pathologists in the entire country of Tanzania, which translates to 1 pathologist per 2.5 million people [29]. Although pathology should play a key role in the determination of appropriate care for the patient, its role is not well recognized by key stakeholders, such as administrators and politicians [29]. The lack of resources allocated to pathology leads inevitably to a large gap in health care for many patients. Correct management of the patient, even when available, is not administered because of the lack of pathologic information [29]. Surgery may be performed without the benefit of preoperative or postoperative pathologic confirmation of the diagnosis, let alone identification of important prognostic information [29]. To address this unmet clinical need in LMICs, technologies that enable rapid, automated, low cost histological processing could be placed in the hands of health care workers, such as nurses or community health care workers.
1.2.4 Automated Interpretation of Tissue Samples Could Reduce Inter-Observer Error

The diagnostic criteria used to classify suspicious lesions are qualitative and subjective, and inter-observer discordance has been shown to be a significant challenge in the diagnosis of borderline lesions [6]. For example, breast cancer diagnosis is an intricate process, which requires tissue procurement, rigorous tissue preparation and histologic assessment whether it is in the context of core needle biopsy diagnosis or surgical excision. Fixed tissue samples are processed after harvesting and are evaluated for presence and type of malignant breast tissue based on standardized histologic criteria [30-33], which employ cytological and qualitative architectural features. Breast tumors that are diagnosed as malignant in nature are graded using different types of grading systems to categorize the tumors into groups to reflect their biology of progression. One of the most widely-used grading systems was developed by Bloom and Richardson in 1957, which used only qualitative criteria to evaluate breast lesions [31]. In 1991, Elston and Ellis published the Nottingham modification to the Bloom and Richardson grading system, which incorporated semi-quantitative criteria to evaluate tubule formation, nuclear pleomorphism, and mitotic count [33]. Extensive research has been done to evaluate the rate of inter- and intra-observer discordance using these grading systems for histologic assessment of fixed breast tissue. While some studies have shown that inter-observer agreement is high in the majority of cases [34], other studies have shown that subjective criteria can lead to inter-observer variation for
margin assessment and poor reproducibility in evaluation of borderline and in situ lesions [6, 35-39]. The availability of techniques that use quantitative criteria that can be applied without subjecting the tissue to processing can overcome the subjectivity of interpretation and may reduce the inter- and intra-observer variability in the histological evaluation of breast tissue [40].

1.3 Potential Biophotonic Solutions

To address the unmet clinical needs described in section 1.2, a number of groups have published on the use of various biophotonic solutions to enable rapid visualization and quantification of tissue features. A summary of these various biophotonic techniques is included in Table 1. Several groups have published on the use of optical spectroscopy, in particular diffuse reflectance, fluorescence and/or Raman spectroscopy for the diagnosis of breast cancer [41, 42]. These groups have demonstrated that the primary sources of intrinsic contrast in breast cancer are alterations in cell density, fat and collagen content as well as tissue vascularity. Our group has developed a quantitative diffuse reflectance spectral imaging device to exploit these intrinsic sources of optical contrast for the imaging of breast tumor margins [41, 43, 44]. In an initial cohort of 48 patients, the technology had a sensitivity and specificity of 80% and 67%, respectively, for detection of residual cancer on the margins of excised specimens [44]. In a later cohort of 92 margins from 72 patients, the technology had a sensitivity and specificity of 89% and 70%, respectively [45]. Nguyen et al. [46] showed that imaging
with optical coherence tomography (OCT) provides an alternate approach to exploiting scattering contrast associated with increased cell density for the detection of microscopic disease in breast tumor margins \textit{ex vivo}, obtaining a sensitivity and specificity of 100% and 82%, respectively, in 20 patients. A trade-off between spectral imaging and OCT imaging is that spectral imaging exploits multiple sources of intrinsic contrast while OCT has inherently much higher spatial resolution. While these clinical studies show promise for \textit{ex vivo} margin assessment of the breast, one of the inherent challenges with exploiting intrinsic sources of optical contrast is that many of the features exploited for cancer detection are also sensitive to normal heterogeneity [47]. For example, fat and collagen content as well as tissue vascularity show variance within the normal breast depending on whether the tissue is predominantly fibroglandular, fibroadipose or adipose. The prevalence of these constituents is inextricably linked to age, breast density, and menopausal status [48, 49] as well as tissue changes associated with chemotherapy [50]. As seen, there is an important need to identify complementary approaches that are sensitive to hallmarks of carcinogenesis but are fairly insensitive to the variations associated with normal tissue heterogeneity.

\textbf{Table 1: Biophotonic techniques that enable real-time visualization of tissue}

<table>
<thead>
<tr>
<th>Optical Technique</th>
<th>Primary contrast</th>
<th>Depth (mm)</th>
<th>Resolution (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photoacoustic tomography</td>
<td>Absorption</td>
<td>3-50</td>
<td>45-700</td>
</tr>
<tr>
<td>Near-infrared spectroscopy</td>
<td>Absorption and elastic scattering</td>
<td>1-5</td>
<td>500-2000</td>
</tr>
<tr>
<td>Raman Spectroscopy</td>
<td>Inelastic scattering</td>
<td>1-3</td>
<td>500-2000</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------</td>
<td>-----</td>
<td>---------</td>
</tr>
<tr>
<td>UV-VIS Diffuse reflectance spectroscopy</td>
<td>Absorption and elastic scattering</td>
<td>0.5-2</td>
<td>500-2000</td>
</tr>
<tr>
<td>Fluorescence Spectroscopy</td>
<td>Fluorescence and elastic scattering</td>
<td>0.5-2</td>
<td>500-2000</td>
</tr>
<tr>
<td>Optical coherence tomography</td>
<td>Elastic scattering</td>
<td>1-2</td>
<td>1-10</td>
</tr>
<tr>
<td>Optical microscopy</td>
<td>Fluorescence and elastic scattering</td>
<td>0.2</td>
<td>0.3-3</td>
</tr>
</tbody>
</table>

**1.3.1 Optical Microscopy**

One strategy for accurate assessment of tumor margins is to adopt approaches that exploit features that are already leveraged in traditional histopathology, such as nuclear size and density, and relationships between tumor cells and surrounding support structures (collagen, fat). Optical microscopy is a powerful technique to obtain high-resolution images of tissue nuclei and stroma in real-time at the point-of-care, without the need for tissue processing. Various optical microscopy techniques including reflectance and fluorescence [51-53], confocal [54, 55], and optical coherence tomography [56, 57] have been used to exploit intrinsic sources of contrast in thick (i.e. unsectioned or unprocessed) tissues. The term thick tissue implies that the tissue is optically thick (i.e. if light is shown on the top side of the tissue, no light can be detected on the bottom side) and that it has not been sectioned with a microtome or cryostat. Additionally, fluorescence microscopy has been combined with vital fluorescent stains to visualize micro-anatomical features. Stains such as acridine orange (AO) [58-60], acriflavine [61,
and DAPI [63] have been used to visualize micro-anatomical features in skin [58], breast [63], ovarian [60], oral [61], and esophageal [62] cancers.

Specifically, a recent study by Farkas et al [63], employed UV-excited DAPI to stain cell DNA \textit{ex vivo} in rat mammary tumor xenografts and human breast tissue microarrays. The authors demonstrated that they could detect rodent tumors with greater than 95% sensitivity and specificity using automated calculation of nuclear features. Gareau et al have published a series of papers employing confocal reflectance and fluorescence microscopy in combination with acridine orange (AO) staining of nuclei for tumor margin assessment of skin cancer \textit{ex vivo} [64-69]. This approach exploited subjective human observation of the images to detect basal cell carcinomas with 97% sensitivity and 89% specificity. Gmitro’s group has published a number of studies on confocal laser scanning microendoscopy of AO stained tissue [70-73]. In an \textit{ex vivo} ovarian study, they showed that automated classification algorithms leveraging texture could diagnose ovarian cancer with better than 95% sensitivity and 90% specificity [70]. They have developed a mobile confocal endoscopy system, and shown images of microanatomical and nuclear detail from the ovary \textit{in vivo} and \textit{ex vivo} using AO as a contrast agent (under IND approval) [71]. Richards-Kortum et al have developed a high resolution fluorescence microendoscope (HRME) and used it in conjunction with acriflavine (similar to AO) to image dysplasia and early cancer in the esophagus [74, 75]. Esophageal images collected \textit{ex vivo} were analyzed both subjectively
by expert clinicians yielding a sensitivity and specificity of 87% and 61%, and quantitatively through the application of algorithms that focused on first order statistics and spatial frequency content, yielding a sensitivity and specificity of 87% and 85% [76]. Yaroslavsky et al. developed a technique for fluorescence polarization imaging of the nonmelanoma cancers stained using antibiotics from the tetracycline family [77]. Average intensity values of the fluorescence polarization images were able to accurately identify the location, size, and shape of lesions in approximately 90% of cases [77]. Wirth et al. stained freshly excised human gliomas with methylene blue and imaged the samples with confocal fluorescence microscopy [78]. Images were compared side by side to histology to investigate staining trends. Lastly, Zhang et al. aimed to assess the diagnostic accuracy and interobserver agreement crystal violet staining in predicting invasion depth of early colorectal cancers [79]. The accuracy for crystal violet staining for real time diagnosis was 93% and the kappa value for inter-observer agreement among 3 endoscopists was 0.70 [25]. A summary of these various fluorescence microscopy studies is included in Table 2.

### Table 2: Types of fluorescence microscopy studies

<table>
<thead>
<tr>
<th>Group</th>
<th>Contrast agent</th>
<th>Constituents stained</th>
<th>Technology</th>
<th>Context</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farkas</td>
<td>DAPI</td>
<td>DNA</td>
<td>Fluorescence microscopy</td>
<td>Ex vivo rat mammary tumor xenografts</td>
<td>Watershed algorithm</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(nuclear size, count)</td>
</tr>
<tr>
<td>Gareau</td>
<td>Acridine orange</td>
<td>DNA, RNA</td>
<td>Confocal fluorescence microscopy</td>
<td>Skin excisions from Mohs</td>
<td>Human-observer</td>
</tr>
</tbody>
</table>
1.4 Need for Automated Diagnosis

While optical microscopy is well suited to enable visualization of tissue morphology at the point of care, robust methods for segmentation and quantitative analysis are essential to enable automated diagnosis. As seen in Table 2, a majority of the groups use human observer analysis or first order statistics that rely on intensity values. Like histopathology, human observer analysis is qualitative and the lack of quantitation leads to inter-observer error. Quantifying intensity information may be appropriate if the background patterns and intensities do not vary greatly between images or patients. However, quantifying intensity information has limited utility in heterogeneous images in which background intensities vary greatly.
1.4.1 Tissue Heterogeneity Presents a Challenge for Quantification

There are three important criteria that have to be considered in the selection of an appropriate image analysis strategy. (1) If the background patterns and intensities vary greatly between images (i.e. if images are heterogeneous), will a method still be able to isolate features of interest, such as tumor nuclei? (2) Can the method resolve overlapping nuclei when attempting to characterize nuclear size or density? (3) Does the method require human intervention and supervision, thus introducing subjective bias and complexity into the analysis?

1.4.2 Current Techniques for Nuclei Segmentation

Many approaches for cell or cell nuclei segmentation exist. A summary of the advantages and disadvantages of commonly used approaches for nuclear segmentation in microscopy is included in Table 3. Global thresholding approaches work well when cell nuclei do not overlap and background intensities are evenly distributed, and its use in isolating cell nuclei is well established in the literature [60, 61, 63, 80]. However, it is also broadly recognized that global thresholding has many shortcomings. Specifically, global thresholding has limited utility in heterogeneous images in which background intensities vary greatly. While global thresholding takes intensity information into account, it does not incorporate geometric information, such as the expected sizes or shapes of nuclei. Thus, in an effort to take geometry into account, many groups have developed techniques that combine global and local image information, such as
adaptive window thresholding, or local maxima detection [80-86], active contours [87-90], watershed segmentation [86, 91, 92], high pass filtering [93], and the circle transform [94].

In adaptive window thresholding or local maxima detection, regions or windows of the image are examined separately and the nuclei within each region are identified based on intensity information through either finding maximum intensities or applying a threshold [80-86]. For heterogeneous images, the window size and threshold within each window should ideally vary across images and patients in order to effectively segment nuclei which are surrounded by various structures, such as muscle, adipose tissue, or other types of connective tissue. Tuning so many parameters on an image-by-image basis quickly can become unmanageable and introduce subjective bias into the quantification of nuclear size and density.

Active contours, such as snakes, find the boundaries of a feature by minimizing an “energy” function associated with the current contour that measures the contour’s curvature and enclosed area [87-90]. However, choosing or defining the energy function can be a complex process, and segmentation results are highly sensitive to this choice. Additionally, active contours require human intervention and supervision through manually guiding the outlining of features or selecting a pixel in the interior of each structure (e.g. tumor nucleus) to be extracted. In images that contain large collections of nuclei, this segmentation approach can quickly become unwieldy. Furthermore, due to
the complexity of this computational technique, it is difficult to know when an optimal solution has been achieved.

In watershed segmentation methods, an image is partitioned into regions separated by watershed lines. While watershed segmentation can identify overlapping nuclei, it is vulnerable to a well-recognized phenomenon called over-segmentation, in which homogeneous regions are segmented into multiple different regions erroneously [86, 91, 92]. These effects can be somewhat mitigated via an involved parameter tuning process requiring significant human intervention, or through variations to the watershed transform, such as viscous watershed [95]; generally, this is an area of active ongoing research. Furthermore, there are many regions in an image that are segmented via watershed methods that are not meaningful for our purpose of isolating nuclei within heterogeneous images. For example in heterogeneous images, the improper segmentation of background elements such as adipose or connective tissue can mistakenly be identified as nuclei, leading to incorrect quantitation of nuclear size and density.

High pass filtering is a technique that is commonly used to isolate edges in images, and can be used to segment small features such as nuclei [93]. While high pass filtering is simple to implement and easy to tune, it is highly sensitive to noise present in an image.
Lastly, the circle transform can be used to detect approximately circular objects of a specified range of radii within an image [94, 96]. While this technique is simple and can identify overlapping nuclei, it assumes that objects are approximately circular and it is sensitive to small variation in background intensity. Because of this sensitivity to small variations in the background, the circle transform, in isolation, is a suboptimal approach for heterogeneous images.

**Table 3: Nuclear segmentation methods**

<table>
<thead>
<tr>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global thresholding</td>
<td>Simple, easy to tune</td>
<td>Requires uniform background intensity</td>
</tr>
<tr>
<td>Adaptive thresholding</td>
<td>Simple</td>
<td>Requires varying window size across image and adjusting threshold within each window</td>
</tr>
<tr>
<td>Active contours</td>
<td>Can find object outlines in complex images</td>
<td>Requires defining complex energy function and human intervention and supervision</td>
</tr>
<tr>
<td>Watershed segmentation</td>
<td>Can identify overlapping nuclei</td>
<td>Results in over segmentation</td>
</tr>
<tr>
<td>High pass filter</td>
<td>Simple, easy to tune</td>
<td>Sensitive to noise</td>
</tr>
<tr>
<td>Circle transform</td>
<td>Simple, can identify overlapping nuclei</td>
<td>Sensitive to small variations in background intensity</td>
</tr>
</tbody>
</table>

Despite the diversity of approaches, segmentation of cells and cell nuclei remains a challenge due to the complexity of images that have varying levels of contrast and non-uniform background heterogeneity, as well as overlapping nuclear features. Thus,
the goal of this work is to obtain high resolution images of tissue morphology through employing fluorescence microscopy and vital fluorescent stains and to develop a quantitative strategy to segment and quantify tissue features, such as nuclei and the surrounding stroma, which will enable automated diagnosis of thick tissues.

### 1.5 Specific Aims and Organization of Chapters

My first aim was to develop an image processing algorithm that can segment nuclei from background tissue heterogeneity and enable automated diagnosis of thick tissue at the point of care. However, results from aim one indicated that low image contrast degraded the ability of the image processing algorithm to segment nuclei. Thus, my second aim was to improve image contrast through employing a technique called structured illumination microscopy and optimize the algorithm for the new imaging system. Results from aim two indicated that the initial contrast agent used to stain the tissue yielded a high false positive rate. Therefore, my third aim was to improve the specificity of the algorithm through leveraging other sources of contrast. A map of the specific aims included in this work is shown in **Fig 1**. Aims 1, 2, and 3 are discussed in detail in Chapters 2, 3, and 4 respectively. Together this work yields an optimized system that is capable of automated segmentation and quantification of micro-anatomical thick tissue images.
Figure 1: Map of specific aims included in this work.
2. Automated Diagnosis

The aim of chapter 2 was to develop an image processing algorithm that can segment nuclei from background tissue heterogeneity and enable automated diagnosis of thick tissue at the point of care. Chapter 2 describes a methodology that systematically evaluates the potential of an image processing approach or combination of approaches, for a specific biomedical problem. The image processing approach is sparse component analysis followed by the circle transform, and the specific indication is the ability to isolate nuclei from heterogeneous tumor margin images.

2.1 Rationale

2.1.1 Sparse Component Analysis (SCA)

A computational technique called sparse component analysis (SCA) was adapted to isolate features of interest, such as nuclei, from background tissue heterogeneity. SCA has been used previously in the image processing community for image compression, enhancement, and restoration, but has never been applied in order to separate distinct tissue types in a heterogeneous image [96]. More specifically, SCA is used here to decompose monochrome images of fluorescently-stained microanatomy into mathematically-discrete components. In combination with a high resolution fluorescence microendoscope (HRME) and the contrast agent, acriflavine [97, 98] the utility of this technique in the visualization of tissue histology in tumor margins is demonstrated. Acriflavine localizes in the nuclei of cells where it reversibly associates
with RNA and DNA [99]. Additionally, acriflavine shows some affinity for collagen and muscle [98].

When performing decomposition, first, SCA [96] is used to separate small acriflavine positive features (which correspond to RNA and DNA) from fibrous components, and adipose components. While acriflavine positive features (APFs) roughly correspond to nuclei, in some cases, RNA and DNA are concentrated within the nucleoli of neoplastic cells; therefore, we refer to these acriflavine positive features as APFs throughout this work. Second, the circle transform (CT) [94] is applied to the deconstructed image to quantify the size and density of overlapping features of interest, in this case, APFs as a means to identify the presence of residual disease in a tumor margin. While the CT is sensitive to small variations in the background, this effect is mitigated by first using SCA to remove the background. The rationale for selecting SCA is that it can segment different types of structures (APFs, muscle, and adipose tissue) in complex heterogeneous images. The CT was chosen to quantify the size and density of APFs because it can distinguish overlapping circular features and is easy to tune. It should be noted that the combination of SCA+CT is not the only solution to this complex problem; however, it is a well-justified approach to analyzing images from heterogeneous tissues and certainly could be adapted to include other methods if they can benefit the overall approach. Unlike image processing techniques which rely solely on intensity information (and are thus susceptible to calibration errors), SCA
incorporates geometric information through the property of sparsity. This leads to a highly flexible approach that requires tuning a very small number of parameters, can resolve overlapping APFs, and does not require human intervention or supervision. Additionally, this technique does not discard image content but rather retains all of the image information inherent in the image to preserve spatial relationships between tissue types, which are essential for proper interpretation of the images.

In this chapter, we demonstrate that SCA+CT can accurately isolate and quantify information within an image with a single stain both on the excised margin and more importantly, on the intact tumor bed. The sensitivity of the SCA+CT approach to variations in APF size, density and background heterogeneity is demonstrated through simulations. The quantitative attributes of this strategy are demonstrated by comparing the output of monochrome fluorescence images of tissue sections to that derived from histology. The clinical utility of this approach to detect residual disease is examined by imaging excised margins as well as the tumor bed in vivo in a cohort of mice after surgical resection of a sarcoma [100, 101]. Specifically, a logistic regression model was optimized on excised tissue margins and prospectively applied to the panel of images obtained from the tumor cavity of 34 mice. After surgery, the mice were monitored for local recurrence, and the results from the diagnostic model were compared to local recurrence endpoints.
2.2 Methods

2.2.1 Ethics Statement

This study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Duke University Institutional Animal Care and Use Committee (Protocol Number: A134-10-05). All surgery was performed under isoflurane gas anesthesia, and all efforts were made to minimize suffering.

2.2.2 Mice and Sarcoma Generation

The generation of the temporally- and spatially-restricted genetically engineered mouse model of sarcoma was performed as described by Kirsch et al [100]. Briefly, mice were anesthetized using isoflurane and soft tissue sarcomas were generated by intramuscular injection of a calcium phosphate precipitate of Ad-Cre (Gene Transfer Vector Core, University of Iowa) in to the proximal portion of the medial or lateral gastrocnemius muscle [100]. Mice were on a mixed 129 SvJae/C57/Bl6 background for these studies. Tumors were excised as described by Mito et al [102]. Mouse genotypes used to generate sarcomas included \( LSL-K\text{ras}^{G12D/+};Trp53^{\text{Flex/Flex}} \) [100] and \( Braf^{C/}\text{et/+};Trp53^{\text{Flex/Flex}} \) [103].

2.2.3 Imaging System and Contrast Agent

A high resolution fluorescence microendoscope (HRME) device that has previously been described in detail [61] was used to collect images of acriflavine stained
tissues. The system contained a 455 nm light emitting diode (Luxeon V Star, LXHL-LR5C), excitation filter (Semrock, FF01-452/45-25) dichroic mirror (Chroma 485 DCLP), emission filter (Semrock, FF01-550/88-25), CCD camera (Point Grey Research, GRAS-14S5), and coherent fiber bundle (Sumitomo, IGN-08/30). The fiber bundle was composed of 30,000 fibers giving a circular field of view of approximately 750 μm in diameter. The resolution of the system was 4.4 μm. A detailed schematic of the HRME system is shown in Fig 2. For both ex vivo and in vivo studies, images were produced by placing the fiber bundle in contact with the acriflavine stained tissue surface. Acriflavine (0.01% w/v, Sigma-Aldrich, [97, 98]) dissolved in phosphate buffered saline (PBS) was topically applied to all ex vivo and in vivo specimens immediately prior to imaging.

![Figure 2: Detailed schematic of the HRME imaging system.](image)

### 2.2.4 Ex Vivo Imaging of Excised Tissue Margins

Seven mice were euthanized immediately prior to surgical tumor resection. Within ten minutes of euthanasia, the tumor was excised from the leg. Seven tumors
were excised, six of which were imaged directly (bulk tissue imaging) and one of which was used for serial tissue sectioning (tissue section imaging).

For tissue section imaging, the excised tissue was flash frozen in liquid nitrogen, imbedded in optimal cutting temperature compound (Tissue-Tek), serially sectioned, and mounted on glass slides. Alternating 50 μm and 5 μm sections were cut with a Leica cryostat with 1-2 sections discarded between to allow for cryostat adjustment. 3-5 drops (0.15 – 0.25 mLs) of acriflavine was topically applied to the 50 μm sections, and after 30 seconds, the tissue sections were raster-scanned with the fiber probe in 1 mm increments to create mosaics. The alternating 5 μm sections were submitted for standard hematoxylin and eosin (H&E) staining.

For bulk tissue imaging, the tissue was laid flat and 3-5 drops (0.15 – 0.25 mLs) of acriflavine were topically applied. 30 seconds following the application of acriflavine, the distal end of the fiber bundle was placed in contact with the tissue and images were acquired from several discrete sites on the tissue (3 to 5 sites per specimen). Each imaged site was inked with a 1 mm dot to facilitate pathologic co-registration. The tissue was fixed, paraffin-embedded and sectioned. En face sections of inked regions were taken below the inked surface. The H&E stained slides were reviewed separately by three pathologists. For each H&E slide, the tissue was diagnosed as tumor, muscle, adipose, or any combination thereof. Only sites for which the diagnosis was concordant between the two pathologists (n= 27 of 33 sites) were included in subsequent analysis.
2.2.5 *In Vivo* Imaging of the Resected Tumor Cavity

A total of 34 mice were included in the *in vivo* study. Tumors located on the lower hind limbs of mice were surgically excised. 3 to 5 drops of acriflavine were topically applied to the tumor cavity, and the fiber probe was placed in contact with the tissue. The probe was raster scanned in 1 mm increments in order to create a mosaic of the *in vivo* margin. Mosaics varied from 3 x 3 images up to 5 x 4 or 4 x 5 images depending on the size of the tumor cavity. After imaging, the excised margin that mirrored the *in vivo* tumor cavity was inked and submitted for H&E processing. The excised margins were sectioned tangentially (*en face*) and the three most superficial sections were given separately to two pathologists for diagnosis. If there were tumor cells present in any of the sections, the margin was diagnosed as positive (Path+). If there were no tumor cells, the margin was diagnosed as negative (Path-). Only margins for which there was a concordant positive or negative pathology diagnosis between the two pathologists were labeled as Path+ or Path-. Mice with discordant pathology were labeled as Path N/A. Additionally, mice were followed for local recurrence for up to 200 days. The 200 day mark was selected because it was approximately twice the length of the latest recurrence in an initial cohort of mice [102]. If a palpable mass could be detected within 200 days, then the tumor locally recurred. If no palpable mass was detected after 200 days, then the tumor did not locally recur. If the mouse did not survive for 200 days post-surgery, then no local recurrence endpoint was achieved.
2.2.6 Generation of Simulated Sarcoma Images

Tumor, muscle, and tumor + muscle tissue images were simulated because these tissue types were most commonly observed in sarcoma margins. All simulated APFs were drawn the same way using MATLAB (2009a, Mathworks Inc., Natick, MA). Specifically, random locations were selected within each region. Then the fspecial command was used to create a disk of a specified radius at each location. The disk was then blurred with a Gaussian filter that had a standard deviation of 1.1 pixels. The blurring step was done to simulate the gradual falloff in intensity seen in APFs imaged experimentally. To select a biologically appropriate range of values for the APF sizes in the tumor simulation, a pathologist was consulted. Typically, APFs in a sarcoma mouse model range from 5 - 15 µm. Therefore, to fully cover the biologically observed range tumor APFs expected with our system (in which there were 1.36 pixels per µm), APFs were simulated with diameters that ranged from 4 up to 18 pixels (3 to 13 µm). Densities were simulated from 60-900 APFs/ 0.25 mm². This range was selected to ensure that a single APF could be detected (at the lowest end) all the way up to a dense collection of APFs in which all APFs were touching (at the highest end).

Muscle was simulated as longitudinal fibers, and tumor + muscle was simulated as tumor APFs on top of the longitudinal fibers, which was the weighted addition of the two components. Specifically, for the tumor + muscle simulations the APF phantom was multiplied by a weighting factor and added to the muscle phantom. The ratios between
tumor APFs and the underlying muscle were chosen based on observations from experimental images and were varied from 1.2-1.8.

2.2.7 Sparse Component Analysis and Circle Transform (SCA+CT)

All processing of the images was performed using MATLAB (2009a, Mathworks Inc., Natick, MA). Images were preprocessed to remove the rim of the fiber bundle and the fiber core pattern superimposed onto each image. Images were cropped to retain only the central portion containing an image of the tissue sample, which resulted in a field of view of 445 x 526 µm. Next, a low-pass Gaussian filter with a full-width half-maximum (FWHM) size of 4.7 µm was applied to remove the high-frequency fiber bundle pattern. This FWHM was selected such that it approximately matched the bandwidth of the Nyquist cutoff frequency for the fiber bundle pattern, equal to half of the distance between fiber cores [104]. An illustration of the image preprocessing steps is shown in Fig 3.

![Image Preprocessing](image)

**Figure 3:** Image preprocessing. An original image of a tumor + muscle site is shown in (a). (b) The image is first cropped to remove the rim of the fiber bundle. (c) Next a low pass Gaussian filter is applied to remove the fiber cores that are superimposed on the image. The image displayed in (c) is the input into the SCA algorithm.
Tissue components (APFs, muscle fibers, and the outline of adipose cells) were separated computationally using a sparse component analysis (SCA) method. Let \( y \) denote the preprocessed image data, modeled as

\[
y = x_{\text{APFs}} + x_{\text{muscle}} + x_{\text{adipose}} + w \tag{1}
\]

where \( x_{\text{APFs}}, x_{\text{muscle}}, \) and \( x_{\text{adipose}} \) denote the true APFs, muscle, and adipose components respectively, and \( w \) accounts for noise and small deviations from the model. The key assumption was that each tissue component has a different “sparsifying” basis or dictionary in which the expansion coefficients were nearly all zero, with only a few large coefficients. (For instance, an image of muscle fibers was relatively smooth, so it could be accurately approximated using a superposition of a small number of Fourier basis functions.) If the sparsifying dictionaries were sufficiently dissimilar, then the sparsity could be exploited to uniquely identify the different tissue components.

The pixel basis was used for the APFs dictionary to capture the small and spatially isolated APFs. The discrete cosine transform (DCT), a variant of the Fourier transform, basis was used to describe muscle components with periodic fiber structures. Specifically the DCT was performed on the entire image (not in blocks) to capture globally smooth muscle features. Mathematically, \( x_{\text{muscle}} = F \theta_{\text{muscle}} \), where \( F \) was a matrix representation of the DCT, and \( \theta_{\text{muscle}} \) was a vector of the DCT coefficients; most elements of \( \theta_{\text{muscle}} \) were zero. The curvelet dictionary was used to represent the curvilinear outlines of adipose cells [105]. Adipose tissue can be described as localized
piecewise smooth features, and therefore curvelets are well suited to capture adipose features. Specifically, curvelets, which are similar to wavelets, have dictionary elements corresponding to different scales and locations throughout an image and is relatively dissimilar to both the pixel and DCT bases [105]. Let $C$ denote the curvelet transform matrix so that $x_{\text{adipose}} = C \theta_{\text{adipose}}$ where $\theta_{\text{adipose}}$ was a sparse curvelet coefficient vector.

The sparse coefficient vectors were estimated by solving a regularized least-squares inversion [96]:

$$
\left( \hat{x}_{\text{APFs}}, \hat{\theta}_{\text{muscle}}, \hat{\theta}_{\text{adipose}} \right) = \underset{x_{\text{APFs}}, \theta_{\text{muscle}}, \theta_{\text{adipose}}}{\arg\min} \left\| y - \left( x_{\text{APFs}} + F \theta_{\text{muscle}} + C \theta_{\text{adipose}} \right) \right\|_2^2 + \tau_{\text{APFs}} \left\| x_{\text{APFs}} \right\|_1 + \tau_{\text{muscle}} \left\| \theta_{\text{muscle}} \right\|_1 + \tau_{\text{adipose}} \left\| \theta_{\text{adipose}} \right\|_1
$$

(2)

where $\arg\min \frac{\min f(x)}{x}$ returned the value of $x$ (the argument) that minimized $f$, $\left\| x \right\|_2^2 = \sum x_i^2$ (the $\ell_2$ norm) and $\left\| x \right\|_1 = \sum |x_i|$ (the $\ell_1$ norm), where $x_i$ was the $i$th component of $x$. The $\ell_2$ term in Eqn. 2 ensured that the approximation $\hat{x}_{\text{approx}}$ was a good fit to the observed data, while the $\ell_1$ terms promoted sparsity in the variables $x_{\text{APFs}}, \theta_{\text{muscle}},$ and $\theta_{\text{adipose}}$. The regularization parameters $\tau_{\text{APFs}}, \tau_{\text{muscle}},$ and $\tau_{\text{adipose}}$ were positive weights that controlled the balance between data fidelity and sparsity in the reconstruction. These parameters were selected using an empirical method described below. To solve this minimization problem, the Gradient Projection for Sparse
Reconstruction (GPSR) algorithm was used [106]. After SCA was applied to isolate APFs, the APF size and density were quantified by computing the circle transform [94] on $\hat{x}_{APFs}$ to detect approximately circular objects (i.e., APFs). This methodology, which is referred to as SCA+CT, was applied to analyze all images in chapter 2.

2.2.8 Regularization Parameter Selection

Four representative experimental images (one tumor, one muscle, one adipose, and one tumor + muscle image) were used to systematically select the best set of regularization parameters. Each image was analyzed with sixty-four different combinations of $\tau_{APFs}$, $\tau_{muscle}$, and $\tau_{adipose}$. For each triplet of parameters, a series of questions were asked of two independent reviewers: (1) does $\hat{x}_{APFs}$ capture the APFs present across the 4 representative images, (2) does $\hat{x}_{adipose}$ capture the curved features in the 4 images (such as the outline of adipose cells), (3) does $\hat{x}_{muscle}$ capture the low frequency features in the 4 images (such as the periodic longitudinal muscle fibers)? The reviewers determined that parameters $\tau_{APFs} = 0.10$, $\tau_{muscle} = 1.0$, and $\tau_{adipose} = 1.0$ gave strong performance across tissue types; this set was applied to all images in chapter 2.

2.2.9 Calculation of AFP Variables

AFP variables were designed to capture features that pathologists typically use to distinguish between normal and diseased tissue. Diseased features typically include increased nuclear density with aneuploidy and pleomorphism (the variation in size and
shape of nuclei) [2, 3]. Specifically AFP variables include density, which is the number of AFPs in a specified area, and diameter, which is defined as the diameter given by the output of CT.

2.2.10 Model Development with Ex Vivo Data Set

In order to develop a model to distinguish between positive and negative in vivo margins, the ex vivo data set was used to examine trends corresponding to the pathology diagnosis. First, Wilcoxon rank sums (non-parametric, two-tailed, alpha = 0.05) were used to determine whether quantitative image parameters were significantly different between positive and negative images. A significance level of $p < 0.05$ was considered to reject the null hypothesis for all analyses. The $n$ for each statistical analysis is listed in the respective figure legend. Next, a multivariate variable-selection analysis based on logistic regression in SAS programming environment was carried out in which all combinations of variables were initially considered for the ex vivo data set. Multivariate logistic regression models yielded receiver operator characteristic curves and the area under the curve for each variable-selection iteration. The area under the curve associated with each model was recorded and tabulated. Additionally, the cross-validated probabilities for each image were determined in SAS using leave one out cross-validation and then used to construct a receiver operator characteristic curve, which was built with a web-based tool [107]. A cut point on the receiver operator characteristic
curve was selected based on the objective function \( F = (1 - \text{sensitivity})^2 + (1 - \text{specificity})^2 \), which is minimized at the optimal cut point.

2.2.11 Application of Optimized Models to the In Vivo Data Set

The models that yielded the highest area under the curve for the ex vivo data set were directly applied to in vivo tumor cavity for which a local recurrence endpoint was obtained. For the in vivo tumor cavities, APF variables were calculated for the entire panel. The model that yielded the highest area under the curve for the in vivo data set was selected. A cut point on the receiver operator characteristic curve was selected based on the same objective function listed in the previous section.

2.3 Results

2.3.1 Illustration of Challenges Associated with APF Segmentation in Heterogeneous Tissue

First to illustrate the challenges associated with isolating APFs in heterogeneous tissue, an image was created with four different tissue types present, including tumor, muscle, adipose, and tumor + muscle tissues. Due to the small field of view of the HRME (0.63 mm\(^2\)) it was not possible to capture all four tissue types in a single image. Therefore, in order to illuminate the challenges associated with isolating APFs from different tissue types, the original image shown in Fig 4 was created by blending different regions of a tumor, muscle, adipose, and tumor + muscle image together. As can be seen, there are inherent challenges in segmenting APFs from each of these regions with a single approach. Specifically, the goal is to isolate the large amount of APFs from
the tumor and tumor + muscle regions without capturing any of the background and to isolate the APFs from the adipose region that are located on the edges of the adipocytes. To illustrate the ability of SCA to meet these goals, the original image was put through SCA to yield the spatial, DCT, and curvelet outputs. From the spatial output we can see that APFs are correctly isolated throughout the tumor and tumor + muscle regions, as well as in the adipose region. Lastly, SCA does not discard image content but rather retains all of the image information inherent in the image. For example, the curvilinear structures present in the original image, particularly in the adipose region, are isolated in the curvelet output. This information can be used to properly interpret images. To illustrate this point the APFs isolated in the spatial output were false colored green and then overlaid onto the curvelet output. As can be seen, the APFs in the adipose region are spatially co-registered with the outline of the adipocytes. This information could be used diagnostically to indicate that these APFs are associated with benign adipose tissue, not with malignant tissue.
Figure 4: Illustration of challenges associated with APF segmentation in heterogeneous tissue. The original image was created by blending different regions of a tumor (T), muscle (M), adipose (A), and tumor + muscle (T+M) image together. This image was put through SCA to yield the spatial, DCT, and curvelet outputs. Then, the spatial image was false colored green and laid on top of the curvelet image to yield to spatial curvelet overlay. This illustrates how APFs are spatially co-registered with some of the features isolated in the curvelet output. This information could be used diagnostically to indicate that these APFs are associated with benign adipose tissue, not with malignant tissue.

2.3.2 Sensitivity of SCA+CT to Background Heterogeneity, APF Size and Density

In order to determine the accuracy and sensitivity of SCA+CT to quantify the size and density of APFs, a series of simulated images were generated in which the size and density of APFs could be varied in the presence of background tissue heterogeneity. Tumor was simulated as randomly dispersed circular APFs ($x_{APFs}$, Fig 5a), muscle was
simulated as longitudinal fibers (\( x_{\text{muscle}} \), Fig 5b), and tumor + muscle was simulated as APFs on top of the longitudinal fibers, which was the weighted addition of the tumor and muscle components (Fig 5c). These simulations were based on characteristics observed in H&E stained histological sections of the same tissue (Fig 5d-f). Specifically, APFs were simulated with diameters ranging from 4-15 µm (4-18 pixels) and densities from 60-900 APFs/0.25 mm². These values were selected based on the biologically expected ranges. The simulated APFs in Fig 5g were added to the simulated muscle image to yield the tumor + muscle simulated image \( y \) (original, Fig 5h) which was the input image into SCA. The outputs of SCA include the spatial \( \hat{x}_{\text{APFs}} \), DCT \( \hat{x}_{\text{muscle}} \), and curvelet \( \hat{x}_{\text{adipose}} \) component estimates, whose summation was denoted as the approximation \( \hat{x}_{\text{approx}} \) of the original image. Next, CT was applied to \( \hat{x}_{\text{APFs}} \) to isolate overlapping APFs for accurate quantitation of size and density. A zoomed in area in the upper right hand corner of the spatial image \( \hat{x}_{\text{APFs}} \) is shown with CT applied to yield SCA+CT (Fig 5h). All subsequent images were analyzed with SCA+CT.
Figure 5: Image simulations used to evaluate SCA. Simulations of tumor (a), muscle (b), and tumor + muscle (c) were mimicked after observations seen in H&E sections (d-f). (g) To better understand how well SCA can capture APFs of various sizes and densities, a single simulated image was created that contained a variety of sizes and densities. The red box indicates the most commonly seen sizes and densities in our data. The outputs of SCA for a tumor + muscle simulation are shown (h). The colorbar shows the gray level intensities, which vary from 0 to 1. In this case, the curvelet image appears blank because no adipose is present in the original image. A zoomed in area of the upper right hand corner of the original image is shown. The same area in the spatial image is shown with CT applied to yield SCA+CT. Scale bar 50 µm.

A set of simulated images with varying contrast ratios was used to examine the influence of background heterogeneity on the ability of SCA+CT to quantify the size and density of APFs (Fig 6a). APFs from each image were isolated with SCA (Fig 6b) and
quantified with CT (Fig 6c). The percent error was calculated by comparing the densities calculated from the simulated images in Fig 6c to the original densities in the tumor simulation, and can be seen in Fig 6d. As seen, the overall error decreases as contrast increases. The highest errors are observed on the left hand side of each image, which corresponds to the area with the smallest APFs. Because CT is based on local gradients, it is unable to distinguish APFs that are smaller than 5 pixels in diameter.

Figure 6: SCA performance varies with APF size, APF density, and background contrast. (a) The image simulation with various sizes and densities of APFs was added to the muscle simulation with varying weights or ratios to create a tumor + muscle simulation. The ratios are reported as the (max tumor APF intensity/ max muscle intensity). The spatial output from SCA is shown in (b). The density was calculated
by applying CT to the images in (b) to yield the images in (c). Red boxes indicate the most commonly seen APF size, APF density, and contrast ratio observed in our data. The percent error was calculated and is shown as a contour plot in (d). The colorbar indicates the percent error (%).

2.3.3 Trends Observed in APF Size and Density Quantified Using SCA+CT

Images captured from frozen tissue sections mounted on glass slides were used to facilitate more direct comparison of the SCA+CT to hematoxylin and eosin (H&E) stained micrographs. For this experiment, the excised tissue was frozen, sectioned serially, and mounted on glass slides. Alternating 50 µm thick and 5 µm thick sections were cut. The 50 µm thick section was imaged with the HRME, and the proceeding 5 µm thick section was stained with H&E and imaged with a standard transmission microscope. This protocol was followed so that optimal images could be acquired of both the H&E stained tissue and the acriflavine stained tissue. Fig 7b contains a panel of representative HRME images and the corresponding H&E en face section micrograph (Fig 7a). The approximate locations where FM images were collected are indicated by squares in the H&E section. The tissue types in the H&E section generally correspond with the tissue types observed in the FM images; however features do not match up perfectly because these panels were acquired from different tissue slices. APFs in the panel of images were isolated with SCA+CT (Fig 7d) as described previously. The SCA+CT panel (Fig 7d) is shown as an overlay in which APFs isolated through SCA+CT are overlaid onto the original image. APFs that are greater than 7 µm in diameter are false colored red, while those that are equal to or less than 7 µm are false colored green.
The cutoff of '7 µm' was chosen because two distinct populations were observed in the histogram of APF diameters shown in Fig 7e. Specifically, one population was located at approximately 5 µm and the other was located at approximately 10 µm, while 7 µm appeared to be located at the dip between these two populations. This two-color scheme is used throughout the remainder of the manuscript to highlight the bimodal distribution observed in Fig 7e. The images with false colored APFs were contrast-stretched in order to enable increased visibility of the false colored APFs. The contrast of the original panel was adjusted in Fig 7c to match the SCA+CT overlay in Fig 7d in order to enable direct visual comparison. As seen, SCA+CT isolates APFs that visually correspond to the locations of nuclei observed in the H&E micrograph (gold standard).

Next, the APF size and density for images in the panel that contained mostly tumor tissue (3, 6, 9), mostly muscle tissue (1, 4, 7), and a mixture of tumor + muscle tissue (2, 5, 8) were quantified using SCA+CT (Fig 7e, f). The diameter results are plotted as histograms or probability distribution functions (pdfs) to show the small differences seen in APF size between groups. The vertical dotted red line corresponds to the 7 µm diameter cutoff, and the horizontal color bars show the mean and standard deviation for each variable. The density results are shown as boxplots. For all boxplots shown in this work, the center line corresponds to the median and the edges of the box correspond to the 25th and 75th percentiles. The whiskers correspond to the most extreme data points not considered outliers, and outliers are plotted individually. Because there are only 3
images in each group, no statistical analysis was performed; however, a general
decreasing trend in APF density from tumor to tumor + muscle to muscle tissue can be
observed. This trend is observed for both smaller APFs (green) and larger APFs (red).
However, there appears to be larger differences between tumor and muscle as well as
tumor + muscle and muscle for the smaller APFs (green).
Figure 7: SCA+CT applied to a panel of tissue section images illustrates trends in APF density. The specific locations where HRME images (Original, (b)) were collected are indicated by squares in the H&E section (H&E, (a)). APFs were isolated and quantified by using SCA+CT, which is illustrated in (d). For SCA+CT, APFs are overlaid onto the original image. APFs that are greater than 7 µm in diameter are false colored red, while those that are equal to or less than 7 µm are false colored green. The contrast of the original panel was adjusted in (c) to match the SCA+CT overlay in (d) in order to enable direct visual comparison. Scale bar 200 µm. The diameter (e), total density (f), density of the smaller APFs (green, (g)) and density of the larger APFs (red, (h)) were quantified for the 3 tumor images (images 3, 6, 9), 3 muscle images (images 1, 4, 7), and 3 tumor + muscle images (images 2, 5, 8). For diameter, the parenthetical values indicate the number of APFs whose diameter was quantified. For density, the parenthetical values indicate the number of images in which the density was calculated. In (e) the vertical dotted red line corresponds to 7 µm diameter, and the horizontal color bars show the mean and standard deviation for each variable.

2.3.4 Optimization of SCA+CT and Logistic Regression on Excised Tumor Margins

SCA+CT was applied to ex vivo FM images from the margins of freshly excised tumors, and representative images are shown in Fig 8a-d. Fig 8a contains an H&E micrograph (column 1) of tumor and the corresponding FM image (column 2). The APFs from these images were isolated using SCA+CT, which can be seen in column 4. The images in column 4 were contrast-stretched in order to enable increased visibility of the false colored APFs. The contrast of the original panel was adjusted in column 3 to match the SCA+CT overlay in column 4 in order to enable direct visual comparison. The SCA+CT processed tumor image shows a dense, disorganized collection of APFs, which is characteristic of malignant tissue. Fig 8b-d contain images of tumor + muscle, muscle, and adipose respectively. The SCA+CT processed tumor + muscle image contains a slightly less dense collection of APFs than is seen in the tumor image, which is
characteristic of residual tumor, while the SCA+CT processed muscle image contains few APFs, as is characteristic of muscle or fibrous tissue. In adipose tissue, APFs are located at the periphery of the adipocytes. In the adipose image in Fig 8d, SCA+CT correctly isolated APFs that are located on the border of the adipocytes. A probability density function of the differences observed in APF size is shown in Fig 8e. Similar to Fig 7e, two distinct populations can be seen in the histogram of APF diameters in Fig 8e.
Figure 8: SCA+CT yields differences between ex vivo tissue types. Representative H&E images (column 1) of tumor, tumor + muscle (T+M), muscle, and adipose tissue and their corresponding HRME image (column 2) are shown (a)-(d) respectively. APFs were isolated and quantified by using SCA+CT and were overlaid onto the original image, which is illustrated in column 4. APFs that are greater than 7 \( \mu m \) in diameter are false colored red, while those that are equal to or less than 7 \( \mu m \)
are false colored green. In column 3, the contrast of the original panel was adjusted to match the SCA+CT overlay in column 4. Scale bar 200 µm. The probability distribution functions (pdf) of the diameters from the four images are shown in (e). The parenthetical values indicate the number of APFs in each image which is synonymous to density, and the vertical dotted red line corresponds to a diameter of 7 µm.

**Fig 9a-c** show boxplots of density for all APFs, as well as the smaller and larger APFs calculated from tumor (n=8), muscle (n=13), and T+M (n=6) images. As expected, each density boxplot shows a decreasing trend from tumor to T+M to muscle. The density of the smaller APFs (green) yields the most significant differences between malignant and benign images (p = 0.0016). **Fig 9d** shows significantly smaller mean diameters for tumor and T+M compared to muscle (p = 0.021).
Figure 9: APF variables calculated for a cohort of ex vivo sarcoma margin images. Density and mean diameter were calculated from 8 tumor, 6 tumor + muscle (T+M), and 13 muscle images. For density, boxplots were created for all APFs, the smaller features (green) and the larger features (red) are shown in (a), (c), and (d) respectively. A boxplot of the mean diameter of all APFs is shown in (b). P values calculated from Wilcoxon rank sums are shown in each boxplot.

As seen in Fig 9b, there were lower values of mean diameter for positive ex vivo margin images than for negative ex vivo images. This result is most likely due to the acriflavine stained nucleic acids being highly concentrated in the nucleoli within sarcoma nuclei, which reduces the average diameter. Conversely, the nucleic acids are more diffuse within muscle and adipose nuclei. Additionally, sarcoma cells are often
interspersed with inflammatory cells, such as macrophages and lymphocytes, whose nuclei are typically smaller than sarcoma cell nuclei, which also reduce the average diameter. To further investigate this trend, a separate analysis was conducted on H&E stained slides in which both the major and minor axis of 10 nuclei were quantified by a pathologist (Fig 10). Results in Fig 10 indicate that sarcoma nuclei are the largest, while sarcoma nucleoli, macrophage nucleoli, and lymphocyte nuclei are much smaller and consistent with the size distributions seen in Fig 7 and Fig 8. Additionally, SCA+CT is designed to detect objects that are approximately circular; therefore, if an APF is more ellipsoidal, the minor axis is likely being measured.
Figure 10: Original H&E stained sections of sarcoma, inflammation, muscle, and adipose tissues are shown in (a). Scale bar 100 µm. Sarcoma, muscle and adipose nuclei are outlined in green. The inflammation image has two primary cell types – macrophages and lymphocytes. Macrophage nuclei are outlined in green and lymphocyte nuclei are outlined in yellow. The smaller features in the sarcoma and
inflammation image, which correspond to sarcoma and macrophage nucleoli respectively, are outlined in blue. The average major and minor axis of the nuclei or nucleoli circled in each image are reported in (b).

Multivariate models were constructed using combinations of variables shown in Fig 9. Each density variable was paired with the mean diameter variable, and the receiver operator characteristic curves achieved with each of these models are shown in Fig 11. All combinations of variables performed comparably on the training set. The set of variables which had the smallest difference between the training and cross-validation set were a combination of the density of the smaller APFs (green) and mean diameter. The optimal cutpoint on the density of the smaller APFs (green) + mean diameter curve in Fig 11b yielded a sensitivity of 82% and a specificity of 75%.
Figure 11: Multivariate models developed based on *ex vivo* sarcoma margins. The receiver operator curves for all 2 variable and 3 variable combinations are shown in (a) - (d). Each plot contains curves associated with the original model as well as with cross-validation. The area under the curve (AUC) and the area under the curve associated with the cross-validation (Crossval AUC) are shown on each plot. The density of the smaller features (green), the density of the larger features (red), and the density of all APFs are referred to as Density (green), Density (red), and Density (both) respectively.

2.3.5 Prediction of Local Recurrence Using an Optimized Algorithm Based on SCA+CT and Logistic Regression

The SCA+CT overlays of a representative LR+/Path+ and LR-/Path- margin are shown in **Fig 12a** and **Fig 12b** respectively. The APF diameters from the two margins
were quantified using SCA+CT and are shown in Fig 12c. As seen, the distribution of the LR-/Path- margin is slightly shifted to the right, which is similar to the trends seen in Fig 7 and Fig 8. Additionally, large differences in density are present between the two margins; specifically, the LR+/Path+ margin contains 617 total APFs (123 APFs/mm$^2$) while the LR-/Path- margin contains 141 total APFs (47 APFs/mm$^2$).
Figure 12: The application of an algorithm based on sparse component analysis (SCA), and circle transform (CT) applied to representative in vivo images from the tumor cavity after resection of the sarcoma. Overlays of a LR+/Path+ and LR-/Path- tumor cavities are shown in (a) and (b) respectively. The overlays were contrast-stretched in order to enable increased visibility of the false colored APFs. Scale bar is 200 µm. The probability distribution functions (pdf) of the APF diameters from the two tumor cavities are shown in (c). The parenthetical values indicate the number of APFs in each panel of images, and the vertical dotted red line corresponds to a diameter of 7 µm.

Local recurrence and pathology endpoints for the 34 mouse study are listed in Table 4. Local recurrence endpoints were achieved for 26 out of 34 mice; the other 8 mice died in post-operative period. Out of the 26 mice where local recurrence could be scored, 9 locally recurred and 17 did not locally recur. Local recurrence and pathology endpoints were not always concordant. Out of the 26 mice that achieved local recurrence endpoints, only 12 (approximately 46%) had matched pathology and local recurrence endpoints—2 mice were LR+/Path+ and 10 were LR-/Path-. For 9 of the 26 mice, the endpoints did not match—5 mice were LR+/Path- and 4 mice were LR-/Path+. The remaining 5 mice that achieved a local recurrence endpoint had discordant pathological diagnosis between the two pathologists. Because of the high degree of discordance, local recurrence was used as the primary endpoint to compare the imaging results.

Table 4: Pathology and local recurrence endpoints for all margins

<table>
<thead>
<tr>
<th></th>
<th>Locally recurred (LR+)</th>
<th>Recurrent free (LR-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathologically positive (Path+)</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Pathologically negative (Path-)</td>
<td>5</td>
<td>10</td>
</tr>
</tbody>
</table>
Next, the multivariate models combining SCA+CT and logistic regression were applied to the in vivo local recurrence data set. The receiver operator characteristic curves achieved with each of the models shown in Fig 11 are shown in Fig 13. Density (green) + diameter and density (both) + diameter are the best performing two variable models for the in vivo local recurrence data set with an area under the curve for all margins = 0.81 and 0.82, respectively (Fig 11a, b). While the three variable model also achieved an area under the curve = 0.82, no improvement in performance was obtained through adding an additional variable.
Figure 13: Multivariate models applied to the *in vivo* images from the tumor cavity predict local recurrence. The receiver operator characteristic curves associated with all of the models shown in Fig 11 are applied here. Each plot contains curves associated with all margins as well as with matching LR+/Path+ and LR-/Path- margins. The area under the curve for all margins (AUC all margins) and the area under the curve for matching LR+/Path+ and LR-/Path- margins (AUC matched LR/Path margins) are shown on each plot.

The optimal cut point on the density (green) + diameter curve in Fig 13b yielded 7 true positives, 2 false negatives, 14 true negatives, and 3 false positives, resulting in a sensitivity of 78% and a specificity of 82%. For comparison, if pathology was used to predict local recurrence in this data set, it would achieve a sensitivity of 29% and a
specificity of 71%. The number of true positives, false negatives, true negatives, and false positives associated with density (green) + diameter that fell within each category is shown in Table 5. Interestingly, for the 2 (100%) false negatives and 1 out of the 3 (33%) false positives, the imaging correlated with the pathology assessment, but not with local recurrence (LR+/Path- and LR-/Path+). For the other 2 false positives – one had matched pathology (LR-/Path-) and the other had a discordant pathological diagnosis (LR-/Path N/A). For 3 out of the 7 (43%) true positives, the pathology did not match local recurrence (LR+/Path-). In addition, in 4 out of the 14 (29%) true negatives, the pathology did not match local recurrence (LR-/Path+). For the margins that had matched local recurrence and pathology, 100% of LR+/Path+ margins were correctly classified as true positives by imaging and 90% of LR-/Path- margins were correctly classified as true negatives by imaging.

Table 5: Number of true positives, false negatives, true negatives, and false positives associated with the best performing model

<table>
<thead>
<tr>
<th>Category</th>
<th>Total</th>
<th>True positives</th>
<th>False negatives</th>
<th>True negatives</th>
<th>False positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR+/Path+</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LR+/Path-</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LR+/Path N/A</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LR-/Path-</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>LR-/Path+</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>LR-/Path N/A</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Abbreviations: LR+/Path+, locally recurred and pathologically positive. LR+/Path-, locally recurred and pathologically negative. LR+/Path N/A, locally recurred and discordant pathology. LR-/Path-, recurrence free and pathologically negative. LR-/Path+, recurrence free and pathologically positive. LR-/Path N/A recurrence free and discordant pathology.
2.4 Discussion

Fluorescent microscopy imaging of acriflavine stained tissue combined with an algorithm that leverages sparse component analysis and circle transform (SCA+CT) provides a rapid, non-destructive and automated strategy for quantitative pathology of thick tissues with non-uniform background heterogeneity. A strength of this method is that it is particularly well suited for applications to different organ sites given that (1) it leverages the micro-anatomical changes in pathological tissue, similar to pathology and (2) it can be applied to highly heterogeneous tissues consisting of multiple tissue types. Moreover, no additional optimization of SCA+CT and logistic regression was required for the algorithm to be effective in analyzing images from the in vivo tumor cavity after resection of the sarcomas, suggesting that the features identified by our technology are independent of whether it is applied to excised or intact tissues. This combination of approaches provides a powerful alternative to complicated and time-intensive immunohistochemistry techniques, which require fixing, sectioning, and staining.

SCA+CT achieved the lowest errors for higher contrast ratios and larger APF sizes. Fig 6 reveals that circle transform (CT) may have a lower limit to the size of APFs that it is able to detect. More specifically, CT has difficulty identifying APFs whose diameter is less than 5 pixels. While tissue simulations provided insight into the performance of SCA+CT and how it varies with APF size, density, and contrast, the shape of the APFs was not varied, which could bias the output of the circle transform.
While SCA+CT is optimized in a sarcoma margin model with a fluorescence microendoscope, the proposed strategy for performing in situ quantitative pathology may lend itself to other types of morphological imaging. Specifically, in chapter 2 the impact of contrast is investigated in Fig 6 to essentially emulate different contrast ranges that would be true of different microscopy techniques. SCA+CT can isolate APFs even at ratios of 1.2 (max tumor APF intensity / max background intensity); however, errors associated with APF density increase as contrast decreases. As mentioned previously, Fig 6 reveals that circle transform (CT) may have a lower limit to the size of APFs that it is able to detect (less than 5 pixels). While this is potentially beneficial in that CT will not pick up small artifacts or aberrations, it could also miss APFs if the pixel size is too large (and the resolution is too low). Taken together, the analytical methodology described here may be generalized to different imaging techniques and staining approaches as long as certain conditions are met. Specifically, conditions include situations in which images have components or sources that can be represented sparsely in a given basis (for example APFs in the pixel basis, muscle in a frequency basis) and the elements of one basis cannot be sparsely represented in any of the other bases. Additionally, while standard Gaussian noise models are assumed in the development of the SCA+CT algorithm, the fundamental ideas can be easily extended to photon noise models common in fluorescence, confocal, and multiphoton microscopy [96, 108-110].
SCA+CT correctly isolated APFs in *ex vivo* images and showed consistently increased density in tumor and tumor + muscle images compared to images containing muscle (Fig 7, 8). In particular, the combination of the density of the smaller APFs (green) and mean diameter achieved the best performance in distinguishing between positive and negative *ex vivo* images (Fig 11) and distinguishing between *in vivo* tumor cavities that locally recurred and cavities that did not locally recur (Fig 13). As expected, there are higher values of density for images of positive *ex vivo* margins and *in vivo* tumor cavities than for images of negative *ex vivo* margins and *in vivo* tumor cavities. Conversely, there were lower values of mean diameter for positive *ex vivo* margins and *in vivo* tumor cavities than for negative *ex vivo* margins and *in vivo* tumor cavities. This is most likely due to the acriflavine stained nucleic acids being highly concentrated in the nucleoli within sarcoma nuclei, which reduces the average diameter. Based on the distributions seen in Fig 7, 8, and 12, and the H&E results in Fig 10, the smaller APFs primarily correspond to sarcoma nucleoli, macrophage nucleoli, and lymphocyte cell nuclei and the larger APFs primarily correspond to muscle and adipose nuclei. In summary, the trends from the H&E analysis in Fig 10 are reflected in the fluorescence microscopy data—the average diameter of the sarcoma APFs is smaller than the average diameter of muscle and adipose APFs. Lastly, the fact that the best model determined from the *ex vivo* data set—density (green) + diameter yielded the highest performance when applied to the *in vivo* data set shows consistency between *ex vivo* and *in vivo*
imaging, which suggests that *ex vivo* imaging can be a good surrogate for *in vivo* imaging when *in vivo* imaging is not clinically feasible.

Similar to human soft tissue sarcomas [9, 11], we found that local recurrence and pathology of the resected margins for mouse sarcomas do not always match. Therefore, this study compared imaging results to local recurrence and pathology separately. In a separate cohort of 25 mice that had resection of a primary sarcoma on which only histological margin assessment was performed (i.e. no imaging), no significant difference in the rate of local recurrence was observed for mice with either a negative or positive margin diagnosis [111]. Possible reasons for the discrepancy between margin status and local recurrence include the limitations and potential errors associated with the pathologic assessment. For example, the ink used to delineate the excised tissue margin can run into crevices where it does not belong, either of which could yield incorrect results. Moreover, it is possible that residual tumor cells (positive margin) are eliminated by the post-surgical inflammatory response or cautery effect on the tumor, thus preventing growth of a local recurrence. Additionally, tumor deposits are not necessarily continuous and can migrate resulting in ‘skip lesions’, which may not be detected. Lastly, *en face/tangential* sectioning of the excised tissue margin was chosen for this study to assess the superficial area of the excised tumor that mirrored the *in vivo* tumor cavity and imaged surface; however, perpendicular sectioning is traditionally done in clinical settings, which may ultimately better correlate with local recurrence.
Due to these possible sources of error associated with pathology and the ultimate clinical importance, local recurrence was used as the endpoint on which to compare our imaging technique. Importantly, when local recurrence and pathology endpoints matched, there was an additional degree of concordance with imaging: 100% of LR+/Path+ margins were correctly classified as true positives by imaging and 90% of LR-/Path- margins were correctly classified as true negatives by imaging. This increased performance may correlate with the tumor volume left behind—specifically for LR+/Path+ margins, there may be a large amount of tumor present on the excised margin as well as in the in vivo tumor cavity.

The primary challenge with this technique is the field of view (FOV). While it is feasible to create mosaics by scanning a sarcoma tumor bed in a mouse, which is on the order of 3-7 mm in any one dimension, it is not feasible to evaluate an entire human surgical margin. We foresee that high resolution anatomical imaging could be combined with wide-field devices, such as a device that we have utilized to detect protease-activated fluorescent imaging agents [102, 112]. Specifically, the wide-field imaging device has a 9.0 x 6.6 mm (59.4 mm²) FOV and has been used to intraoperatively detect residual tumor after wide resections of soft tissue sarcoma in 9 dogs [102, 112]. Additionally, improvements to the microscopic imaging system could also be made to increase the FOV while maintaining the sub-cellular resolution needed for delineation of nucleoli. One approach is to employ a technique called structured illumination
microscopy (SIM), where the entire FOV is illuminated with a defined spatial pattern rather than scanning a focal spot, such as in confocal microscopy [113]. Towards this end, we have developed a wide-field fluorescence SIM system with a FOV of 2.06×1.56 mm (3.21 mm²) and sub-cellular resolution (4.4 µm), which will be described in detail in chapter 3 [114].

In conclusion, topical acriflavine staining and high resolution microscopy combined with appropriate strategies for segmentation and selection of APFs for automated detection of microscopic residual sarcoma in vivo. This combination of technologies is particularly useful in deconstructing images of heterogeneous tissues.
3. Improve Contrast

The aim of chapter 3 was to improve image contrast and increase the FOV through employing wide-field, non-contact structured illumination microscopy and optimize the algorithm from chapter 2 for the new imaging modality.

3.1 Rationale

3.1.1 Contrast Affects Image Processing Performance

While SCA can be used to isolate features, such as APFs, from heterogeneous images, its performance is limited by the contrast between nuclei and the background. Fig 6 illustrates how the overall error associated with SCA+CT increases as contrast decreases. One way in which image contrast can be improved is through rejecting out-of-focus background fluorescence.

3.1.2 Current Techniques for Background Rejection

Many microscopy techniques for background rejection exist, including confocal microscopy [115], two-photon microscopy [115], light sheet microscopy [116], and structured illumination microscopy (SIM) [117]. Confocal microscopy reduces the amount of scattered light detected by isolating signal from the focus of a scanning laser beam. While confocal microscopy can achieve high spatial resolution (<1 μm), the need for beam scanning limits the volume of tissue that can be surveyed in a given amount of time and ultimately limits its translatability to the clinic. Two-photon microscopy works in a similar way to confocal microscopy, but uses longer wavelength pulsed laser light
capable of inducing a non-linear fluorescence excitation at only the very focus of the laser beam. While this technique enables fluorescence imaging at extended depths, it also requires beam scanning. In light sheet fluorescence microscopy, samples are illuminated from the side by a thin sheet of light positioned exactly at the focal plane of the objective, and images are gathered by a camera. This approach is fast as it does not rely on beam scanning; however, it requires relatively transparent low scattering samples. An alternative approach to reject background fluorescence is using structured illumination microscopy (SIM), where the entire field of view is illuminated with a defined spatial pattern and scanning of a focal spot is not required [118]. Other than the use of patterned illumination, the illumination and collection geometry is identical to that of conventional fluorescence microscopy, so a standard CCD may be used for detection. SIM has been shown to perform equivalently (and at times better) than confocal microscopy with respect to optical sectioning and SNR, particularly in superficial tissues [119-121]. However, SIM has the added advantage of full-field illumination and non-descanned detection, thus lowering the complexity compared to confocal scanning systems, and increasing the speed with which microscopy of large tissue areas or fields of view (FOV) can be performed.

Thus in chapter 3, both image contrast and the FOV are increased through employing structured illumination microscopy (SIM). First, the development of a SIM system for thick tissue imaging is discussed. Next the segmentation algorithm for SIM
images is optimized through a series of experiments. In particular, the amount of improvement in image contrast can be controlled through changing the grid frequency that is projected onto the tissue. Thus, the relationship between contrast and algorithm performance is determined through imaging tissue mimicking phantoms and tumor tissue and varying the frequency of the grid pattern used for structured illumination. Lastly, a quantitative model is developed to diagnose mouse sarcoma tumor margins that were imaged ex vivo with structured illumination microscopy.

3.2 Development of SIM System for Contrast Improvement

3.2.1 Methods

3.2.1.1 SIM Theory

As only the in-focus fluorescence from the object is modulated by the sinusoidal component of the illumination pattern, a proper demodulation method can be applied to extract only this information. The most straightforward and commonly used algorithm relies on square law detection to extract and demodulate the in-focus component.

\[
I_{\text{sectioned}} = \sqrt{(I_1 - I_2)^2 + (I_1 - I_3)^2 + (I_2 - I_3)^2} \tag{3}
\]

As the equation shows, the image, which contains only information from the focal plane, is calculated by acquiring three separate images \((I_1, I_2, \text{ and } I_3)\) which differ only in phase shift \((0, 2\pi/3, 4\pi/3)\). The resulting sectioned image thus only contains the modulated portion of the object, which corresponds to the plane of focus.
3.2.1.2 SIM System and Contrast Agent

A SIM system that has previously been described in detail [122] was used to collect images of acridine orange (AO) stained tissues. AO, like acriflavine, is an intravital dye that stains RNA and DNA, skeletal muscle, and collagenous stroma [59, 60, 122]. In chapter 2 the abbreviation APFs was used to describe acriflavine positive features. In chapter 3 APFs are similarly used to describe acridine orange positive features. A detailed schematic of the system is shown in Fig 14. A broadband super continuum laser (Fianium SC400-2) was used to provide illumination for fluorophore excitation. The output from the laser traveled through a bandpass filter centered at 480 nm with a bandpass of 20 nm, which corresponded to the excitation peak of AO. The filtered beam was then passed through a 6X beam expander and a polarizing beam splitter that redirected the light toward an LCoS spatial light modulator (SLM) display chip (Holoeye LCR-720). Following reflection off the SLM chip, the light traveled through a series of 4 lenses, the last of which was the microscope objective (Nikon 4x E Plan Fluor, NA = 0.1). An iris placed one focal length after the first lens was used to spatially filter the diffraction orders created by the sinusoidal pattern on the SLM. The iris was aligned on the optical axis to only pass the 0 and +1 diffraction orders. Allowing these two diffraction orders to pass yielded a sinusoidal pattern at the sample plane, free of undesired higher frequency harmonics, which produce sectioning artifacts if allowed to pass to the sample. The resulting fluorescence generated by the illumination pattern
incident on the sample was collected by the objective, filtered by the dichroic (505nm) and emission filter (peak 520nm), and imaged onto the CCD (LaVision Imager 3 QE) using a 200 mm focal length tube lens (Nikon MXA20696). The system has a measured resolution of 4.4 μm and measured field of view (FOV) of 2.06×1.56 mm (3.21 mm²) [114]. Thus, the FOV of the SIM system is over 13X larger than the FOV of the HRME used in chapter 2 (which had a FOV of 445 x 526 μm or 0.23 mm²). A comparison between the HRME FOV and SIM FOV is illustrated in Fig 14. The illumination pattern used in this section (section 3.2) had a frequency of 31.7 mm⁻¹ at the sample plane which corresponded to an optical section thickness of 120 μm [114]. Additionally frequencies are investigated in section 3.3.
A

\[ \lambda_{\text{ex}} \quad \lambda_{\text{em}} \]

SLM

Spatial Filter

\[ f = 150 \text{ mm} \quad f = 150 \text{ mm} \quad f = 50 \text{ mm} \]

6x Beam Expander

Ex Filter

Fianium Laser

CCD

Em Filter

Dichroic Mirror

Sample

HRME

FOV: 445 x 526 \( \mu \text{m} \) (0.23 mm\(^2\))

Resolution: 4.4 \( \mu \text{m} \)

SIM

FOV: 1.56 x 2.06 mm (3.21 mm\(^2\))

Resolution: 4.4 \( \mu \text{m} \)
Figure 14. (a) Detailed schematic of the SIM imaging system. The $\lambda_{ex}$ peak was 480 nm and the $\lambda_{em}$ peak was 520 nm. The spatial filter diameter was adjusted to allow only the 0 and +1 diffraction orders to pass. (b) and (c) show a comparison between the HRME FOV and the SIM FOV.

3.2.1.3 Tissue Phantom to Demonstrate Contrast Improvement

A tissue phantom was constructed in order to simulate the type of environment seen in thick tissue samples stained with acridine orange. The phantom consisted of fluorescence spheres (Polysciences, Fluoresbrite YG Microspheres) and TiO$_2$ (Sigma, T8141) in a polydimethylsiloxane (PDMS) sample (Dow Corning, Slygard 184). The phantom was constructed in a petri dish with a cover glass window on the bottom (Mattek, P35G-0-14-C). The phantom contained a layer of 10 µm diameter fluorescent spheres dried on the cover glass to generate an optically thin layer of fluorescence (simulating the superficial layer of acridine orange in tissue). The size of the fluorescent spheres was chosen to simulate the size of targets the system was designed to detect (APFs). A 1 cm layer of PDMS and TiO$_2$ was added behind the fluorescent layer. A quantity of 2.25 grams of TiO$_2$ per gram of uncured PDMS was added to the PDMS to generate a scattering coefficient of 10 cm$^{-1}$, which is a commonly measured reduced scattering coefficient in soft tissues [47]. The TiO$_2$ was thoroughly mixed with the PDMS prior to curing. Finally the phantom was placed in a vacuum chamber to draw out the residual air bubbles from the mixing process and to effectively cure the PDMS.

In conventional fluorescence microscopy, the entire field of view is uniformly illuminated with the excitation light. Thus in order to compare SIM images to
conventional microscopy, the phantom was imaged both with structured illumination and uniform illumination. This resulted in a sectioned image, which was calculated by applying equation 3 above, and a uniform image. Each image was normalized by dividing each pixel by the maximum intensity within the image. A signal to background contrast ratio was calculated in both images by manually selecting regions of interest (ROIs) for 5 fluorescent spheres and dividing by a background region of interest (ROI).

3.2.1.4 *Ex Vivo* Imaging of Excised Tissue Margins

A protocol similar to chapter 2 (section 2.2.4) was used to image excised tissue margins from the murine sarcoma model described previously. Briefly, mice were euthanized and tumor was excised from the leg. The tissue was laid flat and 3-5 drops (0.15 – 0.25 mLs) of acridine orange (0.01% w/v, Sigma-Aldrich) dissolved in water was topically applied. After ~30 seconds, the tissue was thoroughly rinsed with a phosphate-buffered saline solution to remove any excess contrast agent. A cover glass was placed over the stained tumor tissue to create a flat surface and images were acquired from several sites on the tissue. Each site was imaged with uniform illumination and with structured illumination using a frequency of 31.7 mm⁻¹. The tissue was inked to facilitate pathologic co-registration. The tissue was fixed, paraffin-embedded and sectioned. For each H&E slide, the tissue was diagnosed by a pathologist.
3.2.1.5 Applying SCA to SIM images

In order to apply SCA to SIM images, which have a different resolution, FOV, and pixel density than the images previously acquired with fluorescence microendoscope, the regularization parameters needed to be re-tuned. The same protocol described in section 2.2.8 was used to select the optimal set of regularization parameters. Briefly, four representative SIM images (one tumor, one muscle, one adipose, and one tumor + muscle image) were used to systematically select the best set of regularization parameters. Each image was analyzed with sixty-four different combinations of \( \tau_{\text{nuclei}} \), \( \tau_{\text{muscle}} \), and \( \tau_{\text{adipose}} \). It was determined that \( \tau_{\text{nuclei}} = 0.10 \), \( \tau_{\text{muscle}} = 1.0 \), and \( \tau_{\text{adipose}} = 1.0 \) gave strong performance across tissue types; this set was applied to all SIM images in chapter 3.

3.2.2 Results

3.2.2.1 SIM Contrast Improvement Demonstrated in Tissue Phantoms

Images were first taken on the phantoms consisting of single layer of 10 µm diameter fluorescence spheres. The spheres were embedded in PDMS and TiO\(_2\) giving the surrounding medium a biologically relevant reduced scattering coefficient value (\( \mu_s' = 10 \text{ cm}^{-1} \)) to introduce background fluorescence. Each image was normalized by dividing each pixel by the maximum intensity within the image. The images shown in Fig 15 clearly demonstrate the improvement seen in structured illumination compared to uniform illumination. The contrast ratio was calculated directly in both
images by manually selecting ROIs for 5 spheres and dividing by a background ROI. The contrast ratio was averaged over these five spheres, which showed a significant quantitative improvement, 889±58% greater in the sectioned image over the uniform image.

![Image of uniform and structured illumination](image)

**Figure 15:** Images of a single layer of 10 μm fluorescent spheres embedded in PDMS and TiO₂. The reduced scattering coefficient of the phantom is approximately 10 cm⁻¹. The signal-to-background was calculated by taking the intensity of 5 manually selected spheres (indicated by red arrows) and dividing by the background ROI (indicated by the blue square). All scale bars are 100 μm.

### 3.2.2.2 SIM Contrast Improvement Demonstrated in *Ex Vivo* Images

To demonstrate the feasibility of imaging of thick tissue, freshly excised sarcoma tumor tissue was stained with AO and imaged using the SIM system. Representative images (uniform and sectioned) of mouse skeletal muscle are shown in **Fig 16a** and sarcoma tumor in **Fig 16b**. Because these are thick tissue samples, a large amount of
scattered background fluorescence is visible in the uniform illumination images. As a result, the contrast enhancement provided by structured illumination improves visualization of individual muscle fibers in normal tissue and individual APFs in tumor tissue.
Figure 16: Uniform and structured illumination images acquired from mouse tissue. (a) Images of skeletal muscle from mouse. Both longitudinal and cross-sectional muscle can be seen in the region of interest. (b) Image of tumor tissue from
mouse sarcoma. Contrast enhancement is clearly seen in the sectioned images. Scale bars are 100 µm.

3.2.2.3 Applying SCA to Representative SIM images

SCA was applied to ex vivo SIM images from the margins of freshly excised tumors, and representative images are shown in Fig 17a-c. Each row contains the original sectioned image, a zoomed in region of interest (ROI), and an SCA overlay of the ROI. Specifically, APFs were isolated with SCA, false colored green and overlaid onto the original ROI. CT was not applied to quantify the APF size and density because previously it was discovered in chapter 2 that CT is only able to pick up APFs that are greater than 5 pixels in diameter. For images acquired with the HRME, most APFs were greater than 5 pixels; however, the pixel density is different for the SIM system. While the HRME has a pixel density of approximately 0.735 µm/pixel, the SIM system has a pixel density of 1.5 µm/pixel. Therefore, CT is not an optimal approach to calculate APF size and density in SIM images. To address this limitation, a new approach was applied to SIM images. Specifically, a threshold value of 0.01 was applied to the spatial output from SCA, and any pixel that was greater than 0.01 was considered part of an APF and false colored green in the SCA overlay. Fig 17 reveals that very few APFs are segmented in SIM images with SCA. This is particularly noticeable in the tumor image in Fig 17a which contains very few APFs.
Figure 17: SCA applied to representative *ex vivo* SIM images. Representative sectioned images (column 1) of tumor, muscle, and adipose tissue are shown in (a)-(c) respectively. Zoomed in regions of interest (ROIs), which are outlined in red in column 1, are shown in column 2. APFs were isolated by using SCA and were overlaid onto the original ROIs, which is illustrated in column 3. All scale bars are 100 µm.

As seen, the tissue phantoms images shown in Fig 15 and *ex vivo* images shown in Fig 16 clearly demonstrate the improvement seen in structured illumination compared to uniform illumination. However, initial segmentation results with SCA in Fig 17 do not look promising. In particular, SCA is unable to segment large numbers of...
APFs in the tumor image in Fig 17a. Previously it was stated that the SCA could be extended to other microscopes if certain conditions were met, including that components (such as APFs) could be represented sparsely in a given basis (such as the pixel basis). Based on the FOV in the SIM system, which is over 13X larger than the FOV in the HRME, dense collections of APFs commonly seen in tumor images may no longer be able to be sparsely represented in the pixel basis. Thus, the sparsity condition mentioned above may no longer be met. Thus, further investigation is warranted in order to optimize image segmentation and analysis for the SIM system.

### 3.3 Optimizing SIM Image Segmentation

Initial results from section 3.2 indicate that SCA may not be the optimal approach for segmenting APFs from SIM images. Therefore, an additional approach for image segmentation was investigated. As mentioned in section 2.1.2, many approaches for nuclei or cell segmentation exist; however, nuclei segmentation remains a challenge due to the complexity of images that have varying levels of contrast and non-uniform background heterogeneity.

In 2004 Matas described a technique for detecting regions (i.e. APFs) in an image that remain stable over a range of threshold values called maximally stable extremal regions (MSER) [123]. The MSER method has important characteristics that are useful for the segmentation of objects (i.e. APFs) from complex (i.e. heterogeneous) images. In particular, MSER is not sensitive to pixel intensity changes and non-uniformities in
background intensities because it is only dependent on pixel intensities within maximally stable extremal regions [124]. The feasibility for MSER for detecting cultured cells was recently demonstrated by Arteta et al [125]. Additionally, MSER has been used to detect cells in phase-contrast images [124].

The objective of section 3.3 is to compare SCA and MSER segmentation algorithms with the goal of optimizing SIM image segmentation. Specifically, tissue mimicking phantoms that contained fluorescent beads of various sizes were imaged with both the HRME and SIM system. In previous work, our group illustrated that the frequency of the illumination pattern used in SIM must be carefully selected because the image signal to noise ratio is dependent on the grid frequency [114]. Therefore, images of tissue mimicking phantoms and fluorescently stained tumor sections were also acquired at different grid frequencies with the SIM system in order to select an optimal frequency.

3.3.1 Methods

3.3.1.1 Imaging Systems

Two different microscopy systems were used for section 3.3, including the high resolution microendoscope (HRME) previously described in detail in section 2.23, and the structured illumination microscopy (SIM) system previously described in section 3.2.1.2. Table 6 contains a summary of the different characteristics of each of these systems. While the FOV of the HRME fiber bundle was approximately 750 µm in
diameter; images were cropped to remove the rim of the fiber bundle, which resulted in a FOV of 445 x 526 µm.

**Table 6: Properties of fluorescent microscopy systems**

<table>
<thead>
<tr>
<th></th>
<th>HRME</th>
<th>SIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective</td>
<td>10x</td>
<td>4x</td>
</tr>
<tr>
<td>Resolution</td>
<td>4.4µm</td>
<td>4.4µm</td>
</tr>
<tr>
<td>Number of pixels</td>
<td>606 x 716</td>
<td>1040 x 1376</td>
</tr>
<tr>
<td>Field of view</td>
<td>445 x 526µm (0.23 mm²)</td>
<td>1560 x 2064µm (3.21mm²)</td>
</tr>
<tr>
<td>µm/pixel</td>
<td>0.735</td>
<td>1.50</td>
</tr>
</tbody>
</table>

While the configuration of the SIM system was the same as illustrated in Fig 14, one small adjustment was made in order to enable switching between different grid frequencies. Previously, an iris was placed one focal length after the first lens to spatially filter the diffraction orders created by the sinusoidal pattern on the SLM and only allow the 0 and +1 diffraction orders to pass to the sample plane. This prevented sectioning artifacts to pass to the sample. When changing between frequencies, the distance between the 0 and +1 diffraction orders varies; therefore, the iris was replaced with an adjustable mechanical slit (Thorlabs VA100) in order to enable reproducible switching between frequencies.

Because of the finite pixel number and size of the SLM, a limited number of discrete frequencies could be produced for structured illumination. Using the experimentally measured FOV, the value of each discrete spatial frequency (in mm⁻¹) at the sample plane (for each pattern generated by the SLM) was determined by imaging
the fluorescence from a uniform calibration test slide and calculating the Fourier
Transform of the resulting image. The absolute spatial frequency values of the various
sinusoidal patterns included 47.7 mm\(^{-1}\), 31.7 mm\(^{-1}\), and 24.1 mm\(^{-1}\).

The optical section thicknesses that correspond to each of these frequencies were
estimated empirically using methodology previously described by Fu et al [114]. Briefly,
a tissue mimicking phantom was translated axially toward the objective and images
were acquired at each 10 µm step. The mean fluorescence intensity from a region of
interest (ROI) within the sectioned image was plotted against axial depth to determine
the section thickness of the system. In order to verify the experimental optical section
thickness measurements, the results were also compared to predicted theoretical results.
Specifically, the defocus of the structured illumination pattern has been previously
shown to match the Stokseth empirical approximation of the optical transfer function
[126, 127]. Fu et al [122] showed that the Stokseth empirical approximation matches the
experimental optical section thickness measurements. The optical section thicknesses
that correspond to each frequency are shown in Table 7. As seen, the axial resolution
can be tuned by varying pattern frequency, while the lateral resolution of 4.4 µm is
symmetric over the field of view and determined by the optics within the SIM system.
Table 7: Optical section thickness for each grid frequency

<table>
<thead>
<tr>
<th>Absolute frequency at sample plane (mm⁻¹)</th>
<th>Optical section thickness estimate (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>47.7</td>
<td>84</td>
</tr>
<tr>
<td>31.7</td>
<td>120</td>
</tr>
<tr>
<td>24.1</td>
<td>156</td>
</tr>
</tbody>
</table>

3.3.1.2 Tissue Mimicking Phantom Preparation and Imaging

The phantoms used in this study were created to model tissue that has been stained with the contrast agent acridine orange (AO). Details on the preparation of the tissue mimicking phantom used in this study have previously been described in detail elsewhere [114]. Briefly, the phantoms were created by homogenously mixing fluorescent spheres (Polysciences, Fluoresbrite YG Microspheres) of varying sizes, with titanium dioxide (TiO₂) and polydimethylsiloxane (PDMS). Phantoms were create by mixing 2.25 grams of TiO₂ per gram of PDMS as described in a prior study [128]. This resulted in phantoms with a reduced scattering coefficient (µ′ₚ) of 10 cm⁻¹, which is the µ′ₚ of most soft tissues [129, 130]. Three set of phantoms were created using 1, 6, and 10 µm spheres respectively. Each phantom was placed in a vacuum chamber in order to remove residual air bubbles created during mixing.

The tissue mimicking phantoms were placed at the focal plane of the HRME and SIM systems. Five unique sites were chosen in order to maximize the coverage on each phantom resulting in five images of each phantom. Each phantom (all sphere sizes) was imaged with the two microscopy systems.
3.3.1.3 Signal to Background Ratio (SBR) Analysis

The signal to background ratio (SBR) was determined in MATLAB by selecting three signal regions (from fluorescent beads) and three nearby background regions of equal area from images of 10µm scatterer phantoms. The average SBR value for each image was found by calculating the ratio of each signal value to the corresponding background value and determining the average. If the SBR went to infinity due to a small background value in the denominator, the SBR value was set to 255, which is the highest pixel intensity value for our 8-bit images.

3.3.1.4 Maximally Stable Extremal Regions (MSER)

The first computational technique SCA+CT has previously been described in detail in sections 2.2.7, 2.2.8, and 3.2.1.5. The second computation technique, called maximally stable extremal regions (MSER) was also applied to all images. MSER has previously been used within the image processing community to reconstruct 3D scenes [131]. A flow chart with an overview of the MSER algorithm is shown in Fig 18. MSER utilizes basic intensity thresholding; however, no global or optimal threshold is sought, rather all thresholds are tested and the stability of the isolated connected components (i.e. APFs) are evaluated. More specifically, all possible thresholds are applied to an image and all sets of connected components (adjacent pixels with intensity values that exceed the current threshold) that are isolated with each threshold are stored. This yields a data structure in which the area of each connected component is stored as a function of
the intensity threshold. Finally, the intensity thresholds that correspond to a local minimum in the rate of change of the area function for each connected component are selected as thresholds producing MSER. In section 3.3, MSER has been adapted to isolate fluorescent beads from tissue phantoms and APFs from tumor images.

Figure 18: (a) Overview of the maximally stable extremal regions (MSER) algorithm. This algorithm, which is used in the automatic reconstruction of 3D scenes was selected as it was a more robust intensity thresholding approach. MSER tests all intensities thresholds within the image to ensure local maxima within various background intensities are not lost. (b) A detailed description and visual representation of the five variable parameters used in MSER, MinArea, MaxArea, Delta, MaxVariation, and MinDiversity.
Mathematical notation for MSER is included below. Assume a given pixel, \( p(x,y) \), has a specific location \((x,y)\) in the image. At a given threshold \( T_i \) where intensity of \( p(x,y) > T_i \), a region of connected components can be identified which contain \( p(x,y) \). That region is \( CC_{p(x,y)}(T_i) \).

\[
V(T_i) = \frac{\text{size}[CC_{p(x,y)}(T_i - 0.5\Delta)] - \text{size}[CC_{p(x,y)}(T_i + 0.5\Delta)]}{\text{size}[CC_{p(x,y)}(T_i)]}
\]

Delta is used in the calculation of variation, which is calculated over all threshold values, \( T_i \) for a given pixel. A given threshold will yield a stable region if its calculated variation is a local minimum, or in other words:

\[
V(T_i) < V(T_{i-1}) \text{ AND } V(T_i) < V(T_{i+1})
\]

For the region to be considered maximally stable, then it must satisfy the following:

\[
V(T_i) < \text{MaxVariation} \text{ AND } \text{MinArea} < \text{size}[CC_{p(x,y)}(T_i)] < \text{MaxArea}
\]

The final parameter is MinDiversity which is used to prune smaller maximally stable regions that are embedded in other maximally stable regions. If two thresholds, \( T_1 \) and \( T_2 \), both yield a maximally stable region containing pixel \( p(x,y) \), then the diversity is calculated with the following:

\[
D(p_{(x,y)}) = \frac{\text{size}[CC_{p(x,y)}(T_2)] - \text{size}[CC_{p(x,y)}(T_1)]}{\text{size}[CC_{p(x,y)}(T_1)]}
\]
If the calculated diversity is less than MinDiversity, than \( T_1 \) is selected as the threshold. Otherwise, \( T_2 \) is selected.

In order to apply MSER specifically to our images, five tuning parameters associated with MSER, which are detailed in Fig 18, needed to be selected. The two most straightforward parameters are MinArea and MaxArea, which are related to the expected size of the connected components (i.e. fluorescent beads). In order to cover the range of bead sizes present in the tissue mimicking phantoms, the minimum diameter was set to 1\( \mu \)m and the maximum diameter was set to 12\( \mu \)m. The next set of parameters was related to the intensity thresholds and included MaxVariation, MinDiversity, and Delta. These intensity parameters were systematically tuned through the range of values on 3 representative phantom images containing 1, 6, and 10 \( \mu \)m fluorescent beads. Ultimately, MaxVariation was determined to be 10, MinDiversity was determined to be 0.5, and Delta was determined to be 10. These parameters were robust and consistent across all imaging systems.

3.3.1.5 Percent Error Calculations

After isolating beads using either SCA+CT or MSER, the average diameter of all beads within an image was determined. The average diameter values calculated from five images of each phantom were used to create boxplots in Figures 1-3. The percent error in the algorithm approximations was determined by comparing the measured or
approximated bead size to the actual bead size (1 µm, 6 µm, and 10 µm) through using the following formula:

\[
\text{Percent error} = \frac{\text{Exact} - \text{Approximated}}{\text{Approximated}} \times 100
\]

### 3.3.1.6 Tissue Imaging and Analysis

The tumors used in this study were generated using methods previously described by Kirsch et al [132]. Tumors were excised as described by Mito et al [102]. Five mice were euthanized immediately prior to surgical tumor resection. Within ten minutes of euthanasia, the tumor was excised from the leg. Five tumors were excised, two of which were imaged directly (thick tissue imaging) and three of which was used for serial tissue sectioning (tissue section imaging).

For tissue section imaging, the excised tissue was flash frozen using liquid nitrogen. The tissue was then imbedded in optical cutting temperature compound (Tissue-Tek) and serially sectioned with a Leica cryostat, with alternating 50µm and 5 µm sections being cut. The 5 µm sections were submitted for hematoxylin and eosin (H&E) staining while the 50µm sections were mounted onto a glass slide and stained by topically applying 3-5 drops of AO (0.01%). Each slide was then washed with 3-5 drops of PBS. Excess liquid was drained off of the slide. A total of nine tumor images were acquired from three animals.
For thick tissue imaging, the tissue was laid flat and 3-5 drops (0.15 – 0.25 mLs) of AO were topically applied. After ~30 seconds, the tissue was thoroughly rinsed with a phosphate-buffered saline solution to remove any excess contrast agent. A cover glass was placed over the stained tumor tissue to create a flat surface and images were acquired from several sites on the tissue (2 to 3 sites per specimen). Each site was imaged with uniform and structured illumination using three different optical section thicknesses (84µm, 120µm, and 156µm). The tissue was inked to facilitate pathologic co-registration. The tissue was fixed, paraffin-embedded and sectioned. For each H&E slide, the tissue was diagnosed by a pathologist. A total of five tumor images were acquired from two animals.

MSER was first applied to the tissue section images in order to separate APFs from fibrous components, and adipose components. The goal was to determine the expected APF density observed in sarcoma tumor tissue. Nine tissue section images were processed using MSER to isolate and count the number of APFs per mm², which was then used as a benchmark when imaging thick tissues. Five thick tissue sites from freshly excised sarcoma tumors were imaged at multiple frequencies. The APF density from the thick tissue images at each optical section thickness was compared to the benchmark APF density established from tissue section imaging using a Wilcoxon rank-sum test. For all tests, a p value of less than 0.05 was considered to reject the null hypothesis.
3.3.2 Results

3.3.2.1 Applying MSER and SCA +CT Algorithms to Tissue Mimicking Phantom Images

Representative images and analysis of tissue mimicking phantoms taken with the HRME and SIM system can be seen in Fig 19 and Fig 20 respectively. In both figures, original images of the 1, 6, and 10 µm phantoms can be seen in A. The images were processed using both MSER and SCA+CT and overlays are shown in B and C respectively. For the overlay, segmented beads were false colored green and overlaid onto the original image. As seen, SCA+CT was able to more accurately segment fluorescent beads from the tissue mimicking phantoms in images acquired with the HRME while MSER was able to more accurately segment beads in images acquired with the SIM system.
Figure 19: Application of MSER and SCA+CT to HRME images.
Representative HRME images of 1, 6, and 10 µm tissue mimicking phantoms are shown in (a) columns 1-3. Original images were analyzed with MSER and SCA+CT and overlays are shown in (b) and (c) respectively. Scale bar 200 µm. Boxplots of the approximated beads diameters obtained with MSER and SCA+CT are shown in (d) and (e) respectively. The red horizontal lines indicate the true bead sizes of 1, 6, and 10 µm.
Figure 20: Application of MSER and SCA+CT to SIM images. Representative SIM images of 1, 6, and 10 µm tissue mimicking phantoms are shown in (a) columns 1-3. Original images were analyzed with MSER and SCA+CT and overlays are shown in (b) and (c) respectively. Scale bar 200µm. Boxplots of the approximated beads diameters obtained with MSER and SCA+CT are shown in (d) and (e) respectively. The horizontal red lines indicate the true bead sizes of 1, 6, and 10 µm.

3.3.2.2 MSER Segmentation Analysis of Individual Beads Imaged with the HRME and SIM Microscopes

In order to further understand why MSER works for SIM images but not for HRME images, one bead from a representative image of the 1µm, 6µm, and 10µm phantoms acquired with each imaging system was isolated using the imcrop function in MATLAB. Plots of the intensity values associated with a line running through the center of the representative 1µm, 6µm, and 10µm beads are shown in Fig 21 columns 1-3.
respectively. The intensity plots of beads imaged with the HRME and SIM microscopes can be seen in Fig 21a-b. Images of the beads and the corresponding MSER overlay can be seen on each intensity plot.

Figure 21: Intensity values associated with a line running through the center of one representative 1, 6, and 10 µm bead. Intensity plots for beads taken with the HRME and SIM microscopes are shown in (a) and (b) respectively. Images of the beads and the corresponding MSER overlay are shown on each plot. The edge of MSER segmentation for each bead is indicated by the vertical blue lines drawn on each plot.

3.3.2.3 Comparing MSER Algorithm Performance to Optical Section Thickness

Average SBR values were calculated in 5 images of 10 µm phantoms taken with the HRME and SIM system at different grid frequencies. The grid frequencies (47.7 mm⁻¹, 31.7 mm⁻¹, and 24.1 mm⁻¹) correspond to optical section thickness (84 µm, 120 µm, and 156 µm). Average SBR values for each optical section thickness are shown as a boxplot.
in Fig 22. A comparison of average percent error associated with the MSER segmentation results and optical section thicknesses can be seen in Fig 22b.

![Figure 22](image)

**Figure 22: Relationship between MSER algorithm performance and optical section thickness.** Average SBR values for 5 images of 10µm phantoms taken with the HRME and SIM system at different optical section thickness are shown as a boxplot in (a). A comparison of the average percent error associated with the MSER segmentation results and the optical section thickness is shown as a scatterplot in (b).

In order to examine if trends established in phantoms held up in tissue images, MSER was applied to quantify APF density in tissue section images (N=9) and thick tissue images (N=5), each of which were imaged with both uniform illumination and structured illumination at different frequencies. The corresponding images are shown in Fig 23a, which qualitatively show the impact of optical section thickness on isolating APFs. The corresponding MSER overlays for these images are shown in Fig 23b. A comparison of the average APF density of the uniform and sectioned images acquired at three optical section thicknesses to the mean nuclear density from the 50 µm frozen section images are shown in Fig 23c. The p-values were calculated using a Wilcoxon
rank-sum test, and revealed that the uniform images were significantly different (p<0.01) than the 50 µm tissue slice images and the sectioned images.

Figure 23: Results demonstrating the impact of optical section thickness on quantification of APF density with MSER. Sectioned images of mouse sarcoma tissue at multiple illumination frequencies (a). The APFs isolated with MSER were false colored green and overlaid onto the original image (b). Mean density of APFs extracted from images mouse sarcoma tissue (c). The APF density extracted from the 50 µm thick frozen section images (N=9) was compared to density extracted from thick tissue imaging (N=5), both uniform and sectioned (84µm, 120µm, and 156µm) images. Scale bar 100µm.

The objective of section 3.3 is to compare SCA and MSER segmentation algorithms with the goal of optimizing SIM image segmentation. While SCA+CT was able to more accurately segment fluorescent beads from the tissue mimicking phantoms
in images acquired with the HRME, MSER was able to more accurately segment beads in images acquired with both the SIM system. Lastly, the grid frequency in the structured illumination microscope was varied to investigate MSER dependence on SBR. Specifically, the optical section thickness and therefore SBR varied as the grid frequency was changed. As SBR increased, MSER accuracy increased. Images of fluorescently stained tumor sections were also acquired at different grid frequencies with the structured illumination microscope to examine if trends held up in tissue images. Results indicated that a grid frequency of 31.7 mm$^{-1}$ yielded the highest density of APFs.

While SCA+CT worked well for segmenting beads in HRME images, it was less accurate at segmenting beads in SIM images. Specifically, SCA+CT accurately segmented 6 and 10 µm beads in HRME images. However, SCA+CT was unable to accurately segment 1 µm beads because they were below the limit of detection in the HRME system. Additionally, SCA+CT was able to accurately segment some of the 10 µm beads in SIM images; however, it significantly overestimated the size of 6 µm beads most likely because the circle transform (CT) could not detect beads that are smaller than 5 pixels in diameter. For the HRME and SIM system, 5 pixels corresponded to 3.7 µm and 7.5 µm respectively. Thus, SCA+CT could only pick up beads or clumps of beads that are larger than 7.5 µm, which led to systematic overestimation of the 6 µm bead size in SIM images. While SCA+CT was able to accurately segment some 10 µm beads in the SIM images, it only segmented a small fraction of the total beads present in each image.
This is particularly noticeable when the SCA+CT overlays were compared to the MSER overlays in Fig 20. As mentioned previously, the key assumption of SCA was that different image components (such as fluorescent beads or APFs) have a different basis (such as the pixel basis) in which they could be represented sparsely. Based on the FOV of the SIM system, which is over 13X larger than the FOV in the HRME, dense collections of fluorescent beads or APFs commonly seen in tumor images could no longer be sparsely represented in the pixel basis. Thus, the sparsity condition was no longer met for the SIM system. Consequently, MSER appeared to be a more appropriate approach for segmentation of SIM images.

As seen in Fig 20 and Fig 22, MSER worked well in systems with some optical sectioning and higher SBR. Specifically, MSER was able to accurately segment 6 and 10 µm beads in SIM images. However, MSER was unable to accurately segment 1 µm beads because they were below the limit of detection in the SIM system. MSER overestimated the size of fluorescent beads in HRME images, which could be explained through close examination of Fig 21. Specifically, MSER was based on local intensity thresholding and compared intensity values of neighboring pixels. All possible thresholds were applied to an image and all sets of connected components (adjacent pixels with intensity values that exceed the current threshold) that are isolated with each threshold were stored. This yielded a data structure in which the area of each connected component was stored as a function of the intensity threshold. Finally, the intensity
thresholds that corresponded to a local minimum in the rate of change of the area function for each connected component were selected as thresholds producing MSER. **Fig 21** shows where the MSER boundaries were located relative to the change in intensity across a fluorescent bead. For the HRME, there was no clear distinction between the bead boundary and the background intensity; therefore, MSER overestimated the true bead size. For the SIM system, the rate of change of the area function was steeper between the bead boundary and the background, which led to accurate segmentation of fluorescent beads with MSER.

In order to further investigate how optical sectioning and SBR affect MSER segmentation, the grid frequency in the structured illumination microscope was varied. **Fig 22** showed that there was a linear relationship between MSER segmentation error and SBR. Specifically, as SBR increased, percent error associated with MSER segmentation decreased. However, the relationship between SBR and grid frequency was not linear. In particular, the highest grid frequency (47.7 mm⁻¹), which corresponded to the thinnest optical section (84 µm), did not yield the highest SBR. Conversely, both 31.7 and 24.1 mm⁻¹ led to higher SBRs, with 31.7 mm⁻¹ yielding the highest average SBR. Increasing the grid frequency led to thinner optical sectioning thicknesses, more background rejection, and higher image contrast. However, it has been previously demonstrated that increasing the grid frequency also led to a decrease in modulation depth, which represented the amplitude of the sinusoidal illumination pattern.
transferred onto the sample [114]. This is important because signal to noise ratio (SNR) was directly linked to modulation depth, and a decrease in modulation depth led to lower SNR [114]. Thus, a tradeoff between optical section thickness and modulation depth must be reached when selecting an optimal grid frequency. To examine if the trends in Fig 22 held up in tissue, images of fluorescently stained tumor sections were also acquired with structured illumination at different grid frequencies. First, the density of APFs isolated with MSER was underestimated in uniform images compared to sectioned images of thick tissue, which illustrates how rejecting background fluorescence through optical sectioning can greatly improve contrast and the performance of MSER. Additionally, Fig 23b indicates that the thinnest optical section did not yield the highest density, which is likely due to the decrease in modulation depth and SNR at higher frequencies. Specifically, a grid frequency of 31.7 mm⁻¹ yielded the highest density of APFs. In summary, 31.7 mm⁻¹ led to the highest SBR, lowest percent error in phantom images, and highest density of APFs in tissue images. Together these results indicated that 31.7 mm⁻¹ was the optimal grid frequency for the SIM system described in this work.

3.4 Quantitative model development for SIM

Results from section 3.3 indicated that MSER segmentation was well suited to isolate APFs in SIM images. Additionally, a grid frequency of 31.7 mm⁻¹ yielded optimal MSER segmentation results. The goal of section 3.4 was to demonstrate the clinical
utility of SIM imaging and MSER segmentation for the detection of residual disease on a preclinical model sarcoma model. A tissue-type classification model was developed to differentiate localized regions (75x75 µm) of tumor from skeletal muscle and adipose tissue based on the MSER segmentation output. A logistic regression model was used for classification of each localized region. For margin classification, the whole margin image was divided into localized regions and the tissue-type classification model was applied to each region.

3.4.1 Methods

3.4.1.1 Margin Imaging Protocol

To simulate clinical tumor growth and disease progression, a transgenic sarcoma model was selected as a primary test bed for margin classification [132]. Mice with conditional mutations in oncogenic K-ras or Braf and p53 were injected intramuscularly in their hind leg with an adenovirus expressing Cre recombinase to induce primary sarcomas. The spontaneous nature of this model accurately represented a clinical tumor margin compared to a tumor xenograft that would tend to develop a fibrous encapsulation.

After injection of the adenovirus, the sarcoma tumor was allowed to develop and grow to 500-700 mm³ (60-90 days) and then surgically removed from the animal. To simulate clinical treatment, the mouse was anesthetized with isofluorane for the duration of the surgical procedure. During the procedure, the entire leg, including the
tumor mass, was amputated to expose the relevant margin. The excised tumor margin was immediately prepared for imaging by topically applying a 0.01% aqueous solution of AO and then rinsed with 0.1M phosphate-buffered solution (PBS). To obtain the highest quality images, a cover glass was placed on top of the tumor margin to create a flat focal plane for imaging.

A brief overview of the imaging procedure is shown in Fig 24. The sample was placed on a 3-axis translation stage under the imaging objective for the purposes of focusing and lateral translation. Due to the large relative size of the margins (~3x5 mm) relative to the field of view (FOV) of the SIM system, multiple images were acquired to cover entire margin. Beginning at the first image location, the sample was translated to each subsequent image location using a micrometer to ensure the sample was moved an equal distance between image sites. A small amount of overlap between adjacent images was allowed to ensure the entire margin was imaged. The sample was moved 1.7 mm in the x-axis and 2.0 mm in the y-axis, with a typical margin fully imaged using 4-6 sites.
To correlate the imaging results obtained with the SIM system to pathology, a procedure was carefully followed immediately after imaging. Tattoo ink was painted over the region of the tissue that was imaged, which allowed the pathologist to identify the corresponding margin after fixing and slicing the tissue. After applying ink, the tissue was immediately frozen in optimal cutting temperature gel using liquid nitrogen. A microtome was used to slice 5 µm thick tissue sections (cross-section to the tissue face painted with ink) spaced 500 µm apart. The tissue sections were then fixed in formalin and then stained with H&E. A pathologist reviewed all slides from a given margin to classify it as a positive or negative margin. The exact number of H&E slides varied based on the physical size of the margin. The criterion used to determine a positive margin was to identify the presence of any tumor cells directly in contact with ink.
3.4.1.2 Tissue Type Classification Model

Additionally, we developed and optimized a predictive model to distinguish between images of distinct tissue types. A total of 30 images (10 of each tissue type) were manually selected as a training dataset for constructing this site level-predictive model. Each image was a 350x300 pixel region of interest (ROI) from the original dataset of 23 margins. This ROI size was chosen because it was roughly equivalent to the FOV of a 10x objective, a typical magnification used by pathologists when studying a suspicious region.

One additional step was taken to further divide each image into smaller elements and each image was broken into 42 evenly spaced 50x50 pixel bins. This procedure is shown visually in Fig 25. The rationale for this step was if classification was carried out on the 350x300 pixel ROIs, it is likely that small focal areas of tumor would not be detected. A bin size of 50x50 pixels was chosen because this corresponded to the diameter of a single skeletal muscle fiber or a single adipose cell. After this step, the true number of measurements in the training dataset was N=1260 bins (30 images with 42 bins per image) with N = 420 tumor bins and N = 420 muscle bins, and N = 420 adipose).
Figure 25: Flow chart which demonstrates the quantitative analysis carried out on the site specific ROIs. The purpose of this analysis was to develop a predictive model to differentiate tumor tissue from muscle and adipose. A total of 30 images were used as a training dataset to develop the model. Each image was further divided into smaller bins to ensure very localized disease could be detected. Finally, three quantities were calculated for each bin, area fraction, average diameter, and average shape, and used as input variables for a logistic regression model.

Using the segmentation output from the MSER algorithm, three different variables were calculated for each bin, area fraction, average diameter, and average shape (perimeter/area).

3.4.1.3 Logistic Regression

Once the segmented area fraction, average diameter, and average shape were quantified for each bin, the goal was to construct a predictive classification model to differentiate tumor bins from normal bins. While the training dataset contained three distinct tissue types, a binary logistic regression model was chosen for the classification algorithm. To reduce the three group training dataset to a binary dataset, the adipose and muscle tissue were classified together as normal tissue sites.

Briefly, logistic regression is statistical classification model used to predict a binary response given specific variables from a data point. In this case, the data points
corresponded to each 50x50 pixel bin (N = 1260) and the binary response corresponded to known classification (1 = tumor, 0 = normal). The logistic regression model was implemented using the statistical software package R. The general form of a 3 variable logistic regression model is as follows:

\[ g(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \]

Where \( g(x) \) is the model predictor value which is a fitted logistic function, \( \beta_0 \) represents the model coefficients, and \( X_1, X_2, X_3 \) are the variable values. In our case, \( X_1, X_2, X_3 \) are the quantified variables from each bin which where:

- \( X_1 \) – Area Fraction
- \( X_2 \) – Average Diameter – average value of all segmented regions within the grid element
- \( X_3 \) – Average Shape – average value of all segmented regions within the grid element

For each bin, these variables, along with the diagnosis (1 for tumor, 0 for muscle or adipose), were input into the R logistic regression code which estimated the coefficients \( \beta_0, \beta_1, \beta_2, \beta_3 \). The model predictor value, \( g(x) \), is then calculated for each bin, where a higher value corresponds to a higher likelihood of presence of tumor. Furthermore, the probability (between 0 and 100%) a given bin is tumor can be calculated from \( g(x) \) using the following equation:

\[
Probability \, (\%) \, of \, tumor = \frac{1}{1 + e^{-g(x)}}
\]
3.4.2 Results

3.4.2.1 Tissue Type Classification Model Results

As expected, the margin-level images clearly contained a mix of various tissue types including tumor, muscle, fat, blood vessels, nerves, and more. The goal was not necessarily to separately identify every tissue type within a margin, but simply to identify any presence of tumor tissue. This would lead to the conclusion that the presence of tumor is positive and conversely, the absence of tumor would be negative.

Three representative ROIs of tumor, muscle, and adipose are shown again in Fig 26 as overlays. Specifically APFs were segmented with MSER, false colored green, and overlaid onto the original image. Both the structured and uniform illumination images are included to demonstrate the importance of contrast enhancement using SIM.

Zoomed in structured illumination images of tumor, muscle, and adipose tissue are also included in Fig 26 to show where MSER segments APFs. As seen, most APFs approximately correspond to the location of nuclei or nucleoli; however, occasionally MSER segments features particularly on the edges of muscle fibers of adipocytes that may not correspond to nuclei.
Figure 26: Application of MSER to both (a) structured and (b) uniform illumination representative images. It is evident that the structured illumination images provide a clear benefit with contrast enhancement which assists the MSER algorithm in accurately identifying APFs. (c) – (e) contain zoomed in images of tumor, muscle, and adipose tissue respectively. Zoomed in regions of H&E images are shown in column 1, structured illumination images are shown in column 2, and the MSER overlay is shown in column 3. White and black arrows point to nuclei while red arrows point to features that were segmented with MSER but do not appear to correspond to nuclei. All scale bars are 100 µm.

Following the procedure in the methods section, the training dataset of 30 tissue specific images was subdivided into N=1260 50x50 pixel bins. The area fraction, average diameter, and average shape were calculated for all bins based on the MSER segmentation output. These calculated values of all bins from the training dataset are displayed as boxplots and shown in Fig 27.

Figure 27: Distribution functions of the three variables, area fraction, diameter, and shape. These include all 50x50 pixel bins from the training dataset, N=420 tumor bins and N=840 normal bins (420 muscle and 420 adipose). A Wilcoxon rank-sum test was used to compare the distributions and calculate the corresponding p-value. Based on the results, the tumor and normal distributions were statistically different for all three variables, so they were included as input to the logistic regression model.

The data from the 50x50 pixel bins were input to the logistic regression to estimate coefficients for $\beta_0$, $\beta_1$, $\beta_2$, $\beta_3$. First, the classification model was constructed using all three of input variables. Three additional classification models were
constructed using two input variables, where one variable was left out of each of the
iterations. The estimated logistic regression coefficients of each model are shown in
Table 8.

<table>
<thead>
<tr>
<th></th>
<th>Area Fraction</th>
<th>Area Fraction</th>
<th>Area Fraction</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average Diameter</td>
<td>Average Diameter</td>
<td>Average Shape</td>
<td>Average Shape</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>2.659</td>
<td>1.374</td>
<td>-4.660</td>
<td>1.3587</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>44.074</td>
<td>43.818</td>
<td>30.12</td>
<td>N/A</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-0.962</td>
<td>-0.870</td>
<td>N/A</td>
<td>-0.2114</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>-0.8652</td>
<td>N/A</td>
<td>12.09</td>
<td>-5.1622</td>
</tr>
</tbody>
</table>

Using these estimated coefficients, the predictor value $g(x)$ was calculated for
each bin. Then using this predictor value, $g(x)$, as the binary classification parameter, a
receiver operating characteristic (ROC) curve was generated by varying the classification
threshold and is shown on the top-left axis of Fig 28. For comparison, three other ROC
curves are also shown that were generated from a logistic regression models built using
only two of three variables. In addition, the area under the curve (AUC) for each ROC
curve was calculated and displayed in the legend. The AUC provided an appropriate
measure for the performance of each model.
Figure 28: The receiver operating characteristic (ROC) curves (top-left) were generated when logistic regression model was applied to the original training dataset of 30 site specific ROIs, 10 tumor sites vs. 20 normal sites (10 muscle and 10 adipose). The specific combinations of variables used as input to the logistic regression model are shown in the legend.

As shown by the ROC curve, the model constructed using all three variables performed exceptionally well (AUC=0.90). Though this was expected since the model was simply applied to its training dataset. To test the unbiased accuracy of the site level model, a separate validation dataset was selected. This validation dataset contained 25 additional images (10 tumor, 10 muscle, 5 adipose) that were separate from the training dataset. Only 5 adipose images were selected for this validation dataset due to the smaller number of adipose sites found in the original dataset of 23 margins. As with the training dataset, each image in the validation dataset was divided into the 50x50 pixel bins and the model predictor value $g(x)$ was calculated for each bin. The ROC curves for the validation dataset are shown in top-right plot of Fig 28. These ROC curves demonstrate that the model constructed with only two-variables (area fraction and average diameter) performs almost identically compared to the three-variable model.
(area fraction, average diameter, and average shape). Therefore, to reduce the complexity, only calculation of area fraction and average diameter were required when the grid analysis approach was applied in the remainder of this study.

3.4.2.2 Application to Full Margin Images

The final step of the analysis was to demonstrate the feasibility of applying the site level predictive model to representative positive and negative full margin images. In order to fully visualize each margin, multi-image mosaics were generated by placing each image site in its proper physical location, while also accounting for the overlap among adjacent images. Selected examples of a positive and negative margin are shown in Fig 29. Following the procedure of the site level images, MSER was then applied to the margin images to segment features. The segmentation results of the sectioned images are shown in with the MSER-segmented regions false-colored in green. For each margin, the classification of an area as positive or negative was provided by the pathologist using the post-operative H&E sections. In addition, a pathologist also inspected the SIM images of the representative margins, while blinded to her original diagnosis using the H&E sections. Using the SIM images alone, she identified specific tissue types such as T (tumor), M (muscle), or A (adipose). The positive margin example in Fig 29 contain tumor in the bottom right of the panel while the remainder of the panel contain muscle. The negative margin example in Fig 29 contains muscle throughout the panel.
Figure 29: Representative images of a positive and negative margin. SIM images were analyzed with MSER and the segmented regions are highlighted in green. In addition, specific tissue types were identified by a pathologist and SIM images and labeled with T (tumor) and M (muscle). Scale bar 200 µm.
Finally, the segmented images from the whole margins were divided into the smaller 50x50 pixel bins and used as input for the site level predictive model. The area fraction, average diameter, and average shape were calculated for each bin within the margin and the predictor value \( g(x) \) was determined. The corresponding analyses are shown in Fig 30, where the tumor probability is represented as a heat map for each margin. From this visualization, it is clear that a larger portion of bins in the positive margins correspond to a high tumor probability. This is an expected result since residual tumor remaining on the positive margin would cause many of the bins to display a high probability. When this analysis was extended to a subset of 6 margins (3 negative, 3 positive), it was shown that at a tumor probability threshold of 50% only 8% of all regions from a negative margins exceeded this threshold, while over 17% of all regions exceeded the threshold in the positive margins.
Figure 30: The results of the grid analysis for the positive and negative margin shown in the previous figure. The predicted tumor probability is represented spatially as heat maps. It is clear that the positive margin exhibits a much higher portion of bins with high tumor probability.
3.5 Discussion

The aim of chapter 3 was to improve image contrast and increase the FOV through employing wide-field, non-contact structured illumination microscopy and optimize the algorithm for the new imaging modality. First, the development of a SIM system for thick tissue imaging was discussed in section 3.2. Clear improvement in image contrast was seen in structured illumination images compared to uniform illumination images. In particular, the tissue phantoms images shown in Fig 15 and ex vivo images shown in Fig 16 clearly demonstrate the improvement seen in structured illumination compared to uniform illumination. However, initial segmentation results of SIM images in Fig 17 revealed that SCA was unable to segment large numbers of APFs in the tumor images, which represented a high false negative rate (in other words features were not segmented with SCA when they should have been). Because the FOV of the SIM system is over 13X larger than the FOV of the HRME, dense collections of APFs commonly seen in tumor images could no longer be sparsely represented in the pixel basis. Thus, SCA was a suboptimal approach for segmenting APFs in SIM images because the sparsity constraint was no longer satisfied. Therefore, MSER was investigated as an alternative approach for image segmentation in SIM images. As seen in section 3.3, MSER was able to accurately segment 6 and 10 µm fluorescent beads (which are approximately the size of APFs commonly seen in our sarcoma model) in SIM images of tissue mimicking phantoms. In previous work, our
group illustrated that the frequency of the illumination pattern used in SIM must be carefully selected because the image signal to noise ratio is dependent on the grid frequency [114]. Therefore, images of tissue mimicking phantoms and fluorescently stained tumor sections were also acquired at different grid frequencies with the SIM system in order to select an optimal frequency. As seen in section 3.2, a grid frequency of 31.7 mm⁻¹ led to the highest SBR, lowest percent error associated with MSER segmentation in phantom images, and highest density of APFs segmented with MSER in tissue images. Together, these results indicated that 31.7 mm⁻¹ was the optimal grid frequency for the SIM system described in this work.

Lastly, a quantitative model was developed to diagnose mouse sarcoma tumor margins that were imaged ex vivo with structured illumination microscopy. Due to the relatively large per frame field of view (>3 mm²), a typical SIM image from the sarcoma margin typically contains more than a single type of tissue. Given the variety of tissue types and the respective differences in staining intensity, a simple global threshold was not an appropriate option. MSER was a robust algorithm with the potential to identify APFs in complex tissue types. The approach allowed the intensity threshold to be selected locally, which was ideal due to the wide range of nuclear intensity relative to the background. Zoomed in structured illumination images of tumor, muscle, and adipose tissue in Fig 26 revealed that most APFs segmented with MSER approximately correspond to the location of nuclei or nucleoli. For example, most APFs in muscle or
adipose were located at the edge of a muscle fiber or adipocyte respectively. However, occasionally MSER segmented features that may not correspond to nuclei. For example, small variations within a muscle fiber could mistakenly be identified as an APF, which was a potential source of false positive error (in other words a feature was segmented when it should not have been). Additionally, MSER occasionally segments oddly shaped APFs, which could represent several nuclei or nucleoli that were mistakenly grouped together, which could lead to an overestimation of APF size or an underestimation of APF density. Despite the potential sources of error associated with MSER segmentation, significant differences in the area fraction, diameter, and shape of APFs were seen between malignant and benign tissues in Fig 27.

Once MSER was optimized for SIM image segmentation (section 3.3), the focus was to develop a predictive model to distinguish and classify different tissue types using “pure” tissue types for characterization and model construction. Area fraction, average diameter, and average shape were selected as parameters since they were quantifiable metrics of pleomorphic changes that pathologists use to identify tumor tissue in H&E. Based on the ROC curves, the model was able to classify tumor and normal tissue with 77% sensitivity and 81% specificity. For an unbiased measure of the model performance, it was applied to a separate validation dataset that resulted in 73% sensitivity and 80% specificity.
Once the site-level grid analysis method was applied to the representative positive and negative margin, the tumor probability heat map provided a visual representation that roughly corresponded to the pathology diagnosis. Specifically, the tissue regions labeled by the pathologist as tumor corresponded to a high tumor probability. Conversely, the muscle regions in the negative margin displayed a low tumor probability. However, there were still many regions that had a high tumor probability even though the pathologist identified them as normal muscle tissue. This could be seen both within the negative margin and the muscle regions of the positive margin. When the analysis was applied to a subset of 6 margins (3 negative, 3 positive), 8% of all regions from a negative margins exceeded a tumor probability threshold of 50%. Thus, 8% of regions in negative margins were considered false positives. These false positive regions are likely due to the high density of APFs present in normal tissues, which clearly demonstrated the challenge in implementing this automatic algorithm based on AO staining alone.

One potential solution to improve the specificity of the algorithm was to incorporate multispectral information that can aid in delineating where normal tissue was located. An additional feature of AO that has not been explored was its spectral-dependent features. AO is known to be a metachromatic dye—in monomeric form its fluorescence emission peak is 525 nm; however in aggregated form the emission peak is shifted to 590-630 nm [98]. It has been reported that AO tends to aggregate between
myofibrils in skeletal muscle [98]. Thus, imaging the two emission peaks of AO could aid in identifying specifically where muscle is located within an image. This approach was further investigated in chapter 4.

Another limitation of this study was that while we acquired a margin level diagnosis, it was difficult to determine exactly where tumor was located within individual images. Specifically, a pathologist reviewed all H&E slides from a given margin to classify it as a positive or negative margin. In addition, a pathologist also inspected the SIM images of the representative margins, and through using the SIM images alone, she identified specific tissue types within each image, such as tumor and muscle. However, this region specific diagnosis could not be confirmed with H&E because H&E slides were prepared by slicing tangentially into the tumor while our SIM imaging was completed by imaging the superficial en face tissue. Thus H&E and SIM images could not be spatially co-registered to determine the exact location of tumor within the SIM images. One potential solution was to use red fluorescent protein (RFP) expressing sarcomas that could be imaged at the same sites that AO was imaged in order to know exactly where tumor was located. This approach was further investigated in chapter 4.

Despite the limitations, SIM imaging combined with MSER for APF segmentation has the potential to be a powerful tool for distinguishing tumor from muscle tissue. Staining tissue with AO was a straightforward procedure that required
only topical application and no specialized tissue or contrast agent preparation.

Additionally, the high fluorescence yield of AO generated more than enough signal for the SIM system to detect. This allowed the resected tumor to be rapidly imaged after excision (<15 min), which is a critical requirement for potential intraoperative use.
4. Improve Specificity

Results from chapter 3 indicated that the initial contrast agent used to stain the tissue yielded a high false positive rate. Therefore, the aim of chapter 4 was to improve the specificity of the algorithm through leveraging other sources of contrast.

4.1 Rationale

4.1.1 Difficult to Identify Location of Normal Tissues

While MSER was able to segment APFs from heterogeneous SIM images, it was also important to place APFs within the context of the entire histological landscape [2, 3]. For example, context information could be used to indicate whether an APF was associated with benign tissue such as stroma, adipose, collagen, or muscle, or with malignant tissue. MSER was originally designed to separate APFs from muscle, and adipose tissue; however, it was difficult to identify where normal tissue, such as muscle was located with a single monochrome image. Specifically, Fig 30 illustrated that many normal tissue regions had a high probability of being malignant (according to our diagnostic model developed in section 3.4) even though the pathologist identified them as normal muscle tissue. Therefore, chapter 4 sought to improve the specificity of the diagnostic model through leveraging other sources of contrast.

4.1.2 Fluorophores for Multi-Spectral Imaging

Acriflavine and acridine orange (AO) have previously been used to visualize micro-anatomical features in skin [58], ovarian [60], oral [61], and esophageal [62]
cancers. Both acriflavine and AO are intravital dyes, which have been shown to stain nuclei, skeletal muscle, and collagenous stroma [98]. Additionally acriflavine and AO have very similar spectral characteristics—acriflavine has an excitation and emission peak of approximately 450 nm and 505 nm respectively and AO has an excitation and emission peak of approximately 480 nm and 525 nm respectively [97]. Acriflavine was used in chapter 2 in conjunction with the HRME, and AO was used primarily in chapter 3 with the SIM system because it was experimentally found to stain the sarcoma tissue more brightly.

For multi-spectral imaging, either additional emission spectra from AO can be collected or additional fluorophores could be used to stain different tissue components. AO is known to be a metachromatic dye—in monomeric form it has its fluorescence emission maximum at 525 nm, while in aggregated forms its peak emission is shifted to 590-630 nm [98]. It has been reported that AO tends to aggregate between myofibrils in skeletal muscle [98]. Thus, imaging the two emission peaks of AO could aid in identifying where muscle is located within an image. Either AO or other fluorophores could be incorporated into the imaging protocol in order to acquire multi-spectral image sets.

Another approach for improving the specificity of the diagnostic model was to investigate using a fluorophore that was more specific to staining tumor. However, in order for a fluorophore to be appropriate for micro-anatomical quantitative imaging at
the point of care, several criteria had to be met. Criteria to consider are listed in Table 9. Specifically, criteria included the specificity of the stain, toxicity, vitality, photobleaching, stability, brightness, and the spectral profile.

<table>
<thead>
<tr>
<th>Question</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>What does the fluorophore stain? How specific is it?</td>
<td>Specific for one type of tissue</td>
</tr>
<tr>
<td>Is it non-toxic?</td>
<td>Non-toxic or requires high dose for toxicity</td>
</tr>
<tr>
<td>At what dose/concentration is it toxic?</td>
<td></td>
</tr>
<tr>
<td>Is it vital? Can it stain fresh tissue? How quickly?</td>
<td>Vital, stains fresh tissue in a few minutes</td>
</tr>
<tr>
<td>How easily does it photobleach?</td>
<td>Little to no photobleaching, stable</td>
</tr>
<tr>
<td>Is it stable in solution?</td>
<td></td>
</tr>
<tr>
<td>How brightly does it stain?</td>
<td>Can be visualized with SIM system</td>
</tr>
<tr>
<td>What is the quantum yield?</td>
<td></td>
</tr>
<tr>
<td>What is the spectral profile? Does it overlap with other stains?</td>
<td>Can be visualized with SIM system, no overlap if simultaneous imaging is required</td>
</tr>
</tbody>
</table>

One potential set of fluorophores that met several of the criteria listed above were the antibiotics from the tetracycline family, which have been commonly used to treat infection and inflammatory disorders. Since 1957, when Rall et al [133] had noted fluorescence in breast tumors following tetracycline therapy, multiple studies have been conducted that utilized this phenomenon as an aid to diagnostics [77, 78, 134-136] and therapy [135, 137, 138] of the different tumor types. In particular, Yaroslavsky et al developed a technique for fluorescence polarization imaging of excised non-melanoma skin cancers stained using tetracycline [77]. Wirth et al employed demeclocycline, which
is an antibiotic from the tetracycline family, combined with fluorescence confocal microscopy to enhance the optical contrast of excised gliomas [78]. These studies indicated that tetracycline appeared to specifically stain freshly excised tumor tissue in a matter of minutes. Additionally, tetracycline has previously been used as an antibiotic, is non-toxic and appears to be stable in solution. Thus, tetracycline was an appropriate candidate for micro-anatomical quantitative imaging at the point of care, and was therefore investigated in chapter 4 as an alternative to AO staining.

4.1.3 RFP Sarcoma Model Provides a Region Specific Diagnosis

One limitation of the study described in section 3.4 was that while a margin level diagnosis was acquired for each mouse, it was difficult to determine exactly where tumor was located within the image panel. In particular, the region specific diagnosis described in section 3.4.2.3 could not be confirmed with H&E because H&E slides were prepared by slicing tangentially into the tumor while our SIM imaging was completed by imaging the superficial en face tissue. To address this limitation, red fluorescent protein (RFP) expressing sarcomas were used in this study to delineate the location of tumor cells within each image.

The purpose of chapter 4 was to improve the specificity of the algorithm through leveraging other sources of contrast. First modifications were made to the structured illumination microscopy (SIM) system to enable fluorescence imaging at a variety of wavelengths. Next, RFP expressing sarcoma were imaged and used to indicate where
tumor is located within image panels. Two potential approaches were investigated for improving the specificity of the diagnostic model. The first approach was imaging the two emission peaks of AO, which could aid in identifying where muscle is located within an image. The second approach was to investigate using tetracycline, which may be more specific to staining tumor.

4.2 Methods

4.2.1 Mice and Sarcoma Generation

The generation of the temporally- and spatially-restricted genetically engineered mouse model of sarcoma was performed as described by Kirsch et al [100]. Briefly, mice were anesthetized using isoflurane and soft tissue sarcomas were generated by intramuscular injection of a calcium phosphate precipitate of Ad-Cre (Gene Transfer Vector Core, University of Iowa) in to the proximal portion of the medial or lateral gastrocnemius muscle [100]. Mice were on a mixed 129 SvJae/C57/Bl6 background for these studies. Mouse genotypes used to generate sarcomas were LSL-KrasG12D+/+;Trp53Flox/Flox [100] and LSL-tdTomato [139]. LSL-tdTomato were obtained from Jackson Laboratories (Bar Harbor, ME). After administration of Ad-Cre, mice formed tdTomato-expressing soft tissue sarcomas at the site of injection. Tumors were excised as described by Mito et al [102].
4.2.2 Modifications to SIM system

In order to obtain spectral information from the SIM system, the excitation and emission filters were replaced with manual filter wheels (Thorlabs CFW6), and the dichroic mirror was replaced with a manual dichroic filter wheel (Thorlabs CDFW5). The modifications to the SIM system are shown in Fig 31.

Figure 31: Detailed schematic of the multispectral SIM imaging system. Modifications are highlighted in red.

The blue channel contained a 480/30 nm bandpass excitation filter, 505 nm dichroic mirror, and 515 nm long pass emission filter. The blue channel was used to image both the first peak of AO (AO1) and tetracycline since the excitation and emission spectra of tetracycline appeared to be similar to AO1 [140]. A second 650/12 nm bandpass emission filter was used to capture the second peak of AO (AO2). The green
channel, which was used to image RFP, contained a 555/25 nm bandpass excitation filter, 562 nm dichroic mirror, and no emission filter (in order to maximize the signal acquired from RFP). These filters were chosen in order to maximize the fluorescence signal acquired with the SIM system. The excitation and emission spectra for AO and RFP are shown in Fig 32.

4.2.3 Ex Vivo Imaging of Excised Sarcoma Margins

A protocol similar to chapter 2 and 3 was used to image excised sarcoma margins. Briefly, mice were anesthetized and RFP expressing sarcomas were excised from the leg. Five tumors were excised, all of which were imaged directly (thick tissue imaging) and three of which were subsequently prepared for serial tissue sectioning (tissue section imaging).
For thick tissue imaging, the excised tumor margin was immediately prepared for imaging by topically applying a 0.01% aqueous solution of AO and then rinsing with 0.1M phosphate-buffered solution (PBS). To obtain the highest quality images, a cover glass was placed on top of the tumor margin to create a flat focal plane for imaging. Multiple images were acquired to cover the entire margin. Beginning at the first image location, the sample was translated to each subsequent image location using a micrometer to ensure the sample was moved an equal distance between image sites. A small amount of overlap between adjacent images was allowed to ensure the entire margin was imaged. The sample was moved 1.7 mm in the x-axis and 2.0 mm in the y-axis, with a typical margin fully imaged using 4-6 sites. Each site was imaged with the blue and green channels in order to acquire mosaics of both AO and RFP.

For tissue section imaging, the excised tissue was flash frozen using liquid nitrogen, imbedded in optical cutting temperature compound (Tissue-Tek) and serially sectioned into alternating 50µm and 5 µm sections, and mounted on a glass slide. The 50µm sections were stained by topically applying 3-5 drops of either AO (0.01% w/v) or tetracycline hydrochloride (Sigma-Aldrich, T3383, 100 mg/mL). Each stained slide was then washed with 3-5 drops of PBS. Excess liquid was drained off of the slide. Mosaics of each tissue section were acquired by translating the sample with a micrometer (as described in the above paragraph). Tissue sections were only imaged with the blue channel, since RFP signal could no longer be detected after tissue processing. The
alternating 5 µm sections were fixed in formalin, stained with H&E, and imaged using a standard bright-field microscope (Leica DM5500 B microscope).

4.3 Results

4.3.1 RFP Sarcoma Model Provides a Region Specific Diagnosis

Red fluorescent protein (RFP) expressing sarcomas were used in this study to delineate the location of tumor cells within each image. To determine if RFP could be detected, excised tissue was first imaged with a commercial upright fluorescent microscope system (Leica DM5500 B microscope). Representative images of pure tumor and pure muscle tissue are shown in Fig 33. As seen RFP signal could be detected in pure tumor tissue, but not in muscle.

Figure 33: Pure tumor and pure muscle tissues were imaged with a commercial microscope. Representative images of a pure tumor and muscle site are shown in rows 1 and 2 respectively. Images of RFP are shown in column 1. The tissue was then topically stained with AO and images were acquired of the same site (column 2). An
overlay of the RFP and AO images are shown in column 3. RFP is false colored red and AO is false colored green for visualization purposes. Scale bar 100 µm.

To determine the specific location of AO versus RFP signal within individual sarcoma cells, a zoomed in region of interest (ROI) was investigated. The ROI is shown in Fig 34. As seen, AO primarily stained nuclei while RFP was expressed in the cytoplasm of sarcoma cells. Thus RFP could be used to delineate where sarcoma cells were located, but not individual nuclei.

![Image](image.png)

**Figure 34:** An overlay of the RFP and AO images are shown in A. RFP is false colored red and AO is false colored green for visualization purposes. The zoomed in ROI is shown in B. Gray scale images of AO and RFP are shown separately in C and D respectively. Scale bar 100 µm.

Next, an RFP tumor was excised, stained with AO, and imaged with the SIM system. Mosaics of AO and RFP images are shown in Fig 35. The RFP signal was
outlined in red, which indicated that tumor was located in the upper left hand portion of the panel. The maximally stable extremal regions (MSER) algorithm described in section 3.3.1.4 was applied to the AO panel to isolate and quantify AO positive features (APFs). Then, the logistic regression model described in section 3.4.2.2 was applied to the MSER output to generate a predicted tumor probability heat map. Specifically, the logistic regression model included area fraction (AF) and average diameter of APFs. The tumor probability heat map, which varies from 0 to 100%, could be directly compared to the RFP signal outlined in red. While many of the 50 x 50 pixel grid elements in the tumor region have high tumor probabilities, there were still many normal tissue regions that have high tumor probabilities. Specifically, at a tumor probability threshold of 50%, 35% of tumor regions and 11% of normal tissue regions in the representative mosaic below exceeded this threshold. Thus, 11% of normal tissue regions were considered false positives. These false positive regions were likely due to the high density of APFs present in normal tissues, which clearly demonstrated the challenge of implementing this automatic algorithm based on AO staining alone.
Figure 35: A mosaic of AO and RFP images are shown in A and B respectively. RFP signal is outlined in red in B-D. MSER segmentation was applied to the panel in A to isolate AO positive features (APFs). APFs isolated with MSER were false colored green and overlaid onto the original images. The MSER overlay is shown in C. The logistic regression model was applied to the MSER panel. The predicted tumor probability (which varies from 0 to 100%) is represented spatially as a heat map in D. Scale bar 200 µm.

4.3.2 Can Multi-Spectral Properties of AO be Used to Improve Algorithm Specificity?

One approach for improving the specificity of the diagnostic model was to image the two emission peaks of AO, which could aid in identifying where muscle was located within an image. A tissue section containing both tumor and muscle was imaged with all emission wavelengths as well as discrete spectral bins through using different emission filters. An overlay of the monomeric emission maximum (530 nm) and the aggregated emission maximum (630 nm) is shown in Fig 36. As seen, the skeletal muscle
region in the upper right hand corner primarily fluoresced at the longer red wavelength, while the tumor region in the left half of the image primarily fluoresced at the shorter green wavelength.

**Figure 36.** Multi-spectral properties of AO. (a) An tissue section image of tumor + muscle imaged with all emission wavelengths. (b) An overlay of 530 (green) and 630 nm (red) emission. (c) Corresponding H&E stained section. Scale bar 200 µm.

To further investigate if imaging the two emission peaks of AO was a robust approach for improving the specificity of the algorithm, a cohort of RFP tumors were excised, stained with AO, and imaged with the SIM system. Both emission peaks of AO, which are referred to as AO1 and AO2, were imaged. A representative margin is shown in **Fig 37.** Specifically, the same analysis illustrated above was applied to this margin; however, the images of AO2 were also included. Areas that contained AO2 signal were outlined in blue. A corrected probability heat map is also shown below in which AO2 positive areas were set to zero, since AO2 likely corresponded to muscle locations. For the AO1 probability heat map in **Fig 37d,** 31% of bins in positive regions have probabilities > 50%, while only 6% of bins in negative regions have probabilities > 50%. For the AO2 corrected probability map in **Fig 37f,** the percentage of bins in negative
regions that have probabilities > 50% decreased to 1%, which shows an increase in specificity. However, the percentage of bins in positive regions that have probabilities > 50% also decreased to 16%, which shows a decrease in sensitivity. Thus, initial results indicated that AO2 may not be specific to muscle or normal tissue.

Figure 37: Mosaics of AO1, RFP, MSER overlays, and probability heat maps are shown in A-D respectively. E shows the AO2 signal that was detected. Areas that contained AO2 signal are outlined in blue. F contains the corrected probability heat map in which AO2 positive areas were set to zero. Scale bar 200 µm.

The same analysis illustrated above was applied to a cohort of 5 excised RFP margins. Specifically the percentage of bins in both positive and negative regions that
have probabilities > 50% were tabulated for AO1 probability heat maps. These variables are referred to as AO1+ for positive regions and AO1- for negative regions. AO2 was also imaged for 3 out of the 5 mice and the percentage of bins in both positive and negative regions that have probabilities > 50% were tabulated for the AO2 corrected probability maps. Similarly, these variables are referred to as AO2+ for positive regions and AO2- for negative regions. A boxplot of these percentages is shown in Fig 38. As seen, AO2- was lower than AO1-, which indicated an increase in specificity. However, AO2+ was also lower than AO1+, which indicated a decrease in sensitivity. Thus, AO2 was not specific to muscle or normal tissue and was not a robust approach for improving the specificity of the algorithm.

Figure 38: Comparison of AO1 and AO2 corrected probability heat maps. The percentage of bins in both positive and negative regions that have probabilities > 50% (or 0.5) were tabulated for AO1 heat maps (n = 5 mice). These variables are referred to as AO1+ for positive regions and AO1- for negative regions. AO2 was also imaged for 3 out of the 5 mice, and the percentage of bins in both positive and negative regions that have probabilities > 0.5 were tabulated for the AO2 heat maps. These variables
are referred to as AO2+ for positive regions and AO2- for negative regions. A boxplot of these variables is shown here.

4.3.3 Can Tetracycline Staining be Used to Improve Algorithm Specificity?

Another approach for improving the specificity of the diagnostic model was to investigate using a fluorophore that was more specific to staining tumor. Specifically, tetracycline was selected because it appeared to specifically stain freshly excised tumor tissue in a matter of minutes, and was non-toxic and stable in solution. To initially assess tetracycline staining, tissue sections were stained with tetracycline hydrochloride and imaged with the AO channel of the SIM system. The corresponding H&E slides were imaged with a bright field microscope. Images of a representative tissue section are shown in Fig 39. The tumor section of the H&E section was traced in Matlab and used to indicate where tumor is located within the corresponding tetracycline images. Fluorescent positive features (FPFs) were isolated with MSER, false colored green, and overlaid onto the original tetracycline images. As seen, tetracycline appeared bright in tumor regions, but not in muscle. Additionally, a higher density of FPFs was isolated within the tumor region compared to the muscle region. Zoomed in regions of tumor and muscle tissue are also shown in Fig 39 for enhanced visualization. Overall tetracycline images correlated well with the H&E images.
Figure 39: Representative images of tetracycline stained tissue sections. The corresponding H&E image is shown in A. The tumor section was outlined in B (the lower half of the panel contains tumor). The mosaic of tetracycline images is shown in C. Fluorescent positive features (FPFs) were isolated with MSER, false colored green, and overlaid onto the original tetracycline images. The MSER overlay is shown in D. Both C and D also show the tumor section outlined in red (the lower portion of the
panel). E and F contain zoomed in images of tumor and muscle tissue respectively. Zoomed in regions of H&E images are shown in column 1, structured illumination images are shown in column 2, and the MSER overlay is shown in column 3. All scale bars 200 µm.

Fig 39 revealed that tetracycline did not stain muscle fibers. To illustrate the difference in AO versus tetracycline staining of muscle, serial tissue sections were stained with AO and tetracycline. Images of a representative set of serial sections are shown in Fig 40. The H&E section indicated that tumor was primarily located in the lower left quadrant of the panel while muscle was located in the upper right quadrant. As seen, the muscle region appeared bright in the AO panel, but appeared dim in the tetracycline panel.
Figure 40: Comparison of representative AO and tetracycline stained tissue sections. The corresponding H&E section is shown in A. The tumor region, which is in the lower left quadrant of the panel is outlined in B. The AO panel is shown in C and the tetracycline panel is shown in D. The tumor regions in C and D are outlined in red. All scale bars 200 µm.

To quantify the difference in AO versus tetracycline staining, 3 representative tumor and muscle regions of interest (ROIs) were selected from 3 different sets of AO and tetracycline stained tissue sections. For each set of tumor and muscle ROIs, the ratio of the intensities present in the tumor to muscle ROIs was calculated. Representative ROIs as well as a boxplot of the tumor to muscle ratios are shown in Fig 41. As seen, the brightness of AO stained tumor and muscle ROIs was comparable; thus, the tumor to
muscle ratios are close to 1. However, tetracycline stained tumor more brightly than muscle and achieved an average tumor to muscle ratio of approximately 1.6.

![Image: Comparison of AO and tetracycline stained tumor and muscle regions of interest (ROIs). Representative tumor and muscle ROIs are shown in A and B respectively. Scale bar 100 µm. The tumor to muscle ratios were quantified from 3 sets of AO and tetracycline (tetra) stained ROIs. A boxplot of the tumor to muscle ratios is shown in C.](image)

**Figure 41:** Comparison of AO and tetracycline stained tumor and muscle regions of interest (ROIs). Representative tumor and muscle ROIs are shown in A and B respectively. Scale bar 100 µm. The tumor to muscle ratios were quantified from 3 sets of AO and tetracycline (tetra) stained ROIs. A boxplot of the tumor to muscle ratios is shown in C.

**Fig 42** further demonstrates the differences in AO and tetracycline staining. Two serial tissue sections were stained with AO and tetracycline. The H&E section indicated that there was a river of muscle running down the left hand side of the panel. The tumor probability heat maps for the AO and tetracycline panels are shown in **Fig 42e** and **Fig 42f** respectively. As seen, less of the muscle region in the tetracycline panel had high tumor probabilities, indicating that tetracycline yielded higher specificity.
Figure 42: Comparison of AO and tetracycline tumor probability heat maps. The corresponding H&E section is shown in A. The tumor regions are outlined in B. The MSER overlay and probability heat map for the AO panel is shown in C and E respectively, while the MSER overlay and probability heat map for the tetracycline panel is shown in D and F respectively. The tumor regions in C-F are outlined in red. All scale bars 200 µm.

The same analysis illustrated above was applied to a cohort of 3 sets of AO and tetracycline stained serial tissue sections. Specifically the percentage of bins in both positive and negative regions that have probabilities > 50% were tabulated for AO probability heat maps. These variables are referred to as AO+ for positive regions and AO- for negative regions. Similarly, percentage of bins in both positive and negative regions that have probabilities > 50% were tabulated for the tetracycline probability
maps. Similarly, these variables are referred to as Tetra+ for positive regions and Tetra- for negative regions. A boxplot of these percentages is shown in Fig 43. As seen, Tetra+ and AO+ had similar sensitivities; however, Tetra- was lower than AO-, which indicated an increase in specificity. Thus, tetracycline shows promise for increasing the specificity of tumor detection in SIM images of a preclinical sarcoma model.

Figure 43: Comparison of AO and tetracycline tumor probability heat maps. The percentage of bins in both positive and negative regions that have probabilities > 50% (or 0.5) were tabulated for AO heat maps (n = 3 tissue sections). These variables are referred to as AO+ for positive regions and AO- for negative regions. Similarly, the percentage of bins in both positive and negative regions that have probabilities > 0.5 were tabulated for tetracycline heat maps. These variables are referred to as Tetra+ for positive regions and Tetra- for negative regions. A boxplot of these variables is shown here.

4.4 Discussion

The aim of chapter 4 was to improve the specificity of the algorithm through leveraging other sources of contrast. First, modifications were made to the structured illumination microscopy (SIM) system to enable fluorescence imaging at a variety of
wavelengths. Specifically, the SIM system was modified to enabling imaging of red fluorescent protein (RFP) expressing sarcomas. RFP tumors were used in this study to delineate the location of tumor cells within each image. Results from section 4.3.1 indicated that RFP can be visualized with the multi-spectral SIM system and used to indicate where tumor is located within image panels. Initial analysis of AO stained panels confirmed that there was room for improvement in tumor detection, particularly in regards to false positive regions that were negative for RFP.

One potential approach for improving the specificity of the diagnostic model was to image the two emission peaks of AO, which could aid in identifying where muscle was located within an image. However, results from section 4.3.2 indicate that while the second peak of AO (AO2) led to an increase in the specificity it also led to a decrease in sensitivity. Thus, AO2 is not specific to normal tissue, such as muscle, and therefore was not an appropriate approach for improving tumor detection.

Another approach for improving the specificity of the diagnostic model was to investigate using a fluorophore that was more specific to staining tumor. Specifically, tetracycline was selected because it appears to specifically stain freshly excised tumor tissue in a matter of minutes, and was non-toxic and stable in solution. Tissue sections were stained with tetracycline and imaged with the SIM system. As seen in section 4.3.3, tetracycline images correlated well with the H&E images. Additionally, tetracycline appeared bright in tumor regions, but not in muscle. To quantify the difference in AO
and tetracycline staining, representative tumor and muscle regions of interest (ROIs) were selected from AO and tetracycline stained tissue sections. For each set of tumor and muscle ROIs, the ratio of the intensities present in the tumor to muscle ROIs was calculated. Quantification revealed that the brightness of AO stained tumor and muscle ROIs was comparable and consequently the tumor to muscle ratios were close to 1. Conversely, tetracycline stained tumor more brightly than muscle and achieved tumor to muscle ratios around 1.6. Further comparison between AO and tetracycline stained serial tissue sections revealed that AO and tetracycline achieved similar sensitivities; however, tetracycline achieved better specificity. Thus, tetracycline shows promise for increasing the specificity of tumor detection in SIM images of a preclinical sarcoma model and further investigation is warranted.

A significant strength of this study was that RFP expressing sarcomas enabled us to determine where positivity is located within each margin. Additionally, RFP images enabled us to confirm that improvements could be made in the specificity of the algorithm, but that AO2 was not a robust approach since it led to degradation in the sensitivity of the algorithm. Serial tissue section imaging enabled comparison between both AO and tetracycline stained image and H&E images. Initially experiments with tetracycline showed promise for increasing the specificity of tumor detection, and thus further experiments are warranted.
While this study illustrates that tetracycline may improve the specificity of the tumor detection, additional work is needed with a larger independent data set to validate the feasibility and reproducibility of these parameters. Future studies could focus on validating tetracycline staining and analysis in thick tissues, such as the excised RFP tumor margins. However, tetracycline was not as bright as AO, and longer integration times (~5 second) were required to capture images of tetracycline compared to AO (~1 second). One approach to address the weak signal acquired from tetracycline is increase the optical power in the blue channel. In order to increase the excitation power, a new source could be explored since the Fianium laser source has low power density in the blue wavelengths. Laser diodes or LEDs can provide a feasible option cost significantly less than the Fianium. More details on the tetracycline preclinical validation study and technical improvements for the SIM system are included in sections 5.2.1 and 5.2.2.

Lastly, this study was conducted in a preclinical sarcoma model in order to assess whether a combination of imaging tools could detect residual disease in resected tumor cavities. Moving forward, our goal is to expand this work to clinical tumor margin assessment in humans. Towards this end, our group has is focusing on two different projects. The first project is validating that morphological staining, fluorescence microscopy, and automated segmentation of features enables rapid visualization and quantitative interpretation of clinical breast tissue morphology. The second project is to
combine high resolution morphological imaging with wide-field devices, such as our group’s spectral imaging platform [141]. More details on these projects are included in sections 5.2.3.1 and 5.2.3.2.
5. Conclusions and Future Directions

5.1 Conclusions

The primary goal of this dissertation was to design a system capable of automated segmentation and quantification of micro-anatomical thick tissue images. This system enables more rapid and automated morphological imaging, which could benefit several different clinical situations. For example, automated morphological imaging could be used to rapidly detect residual cancer on the surface of tumor resection specimens during excisional surgeries [4]. Additionally, rapid assessment of biopsy specimens at the point of care could enable clinicians to confirm that a suspicious lesion is successfully sampled, preventing an unnecessary repeat biopsy procedure. Rapid and low cost morphological imaging could also be potentially useful in settings lacking the human resources and equipment necessary to perform standard histologic assessment [5]. Lastly, automated interpretation of tissue samples could potentially reduce inter-observer error, particularly in the diagnosis of borderline lesions [6].

To address these unmet clinical needs, a number of groups have published on the use of various biophotonic solutions, including optical spectroscopy [41, 42], optical coherence tomography [46], and various optical microscopy techniques [62, 142] to enable rapid visualization and quantification of tissue features. This work focuses on fluorescence optical microscopy combined with topical stains that highlight tissue morphology because these technologies can be used to obtain high-resolution images of
tissue nuclei and stroma without tissue processing. While many group have developed fluorescence microscopy techniques that enable visualization of tissue morphology, robust methods for segmentation and quantitative analysis are essential to enable automated diagnosis. As discussed in section 1.4, a majority of these groups use human observer analysis or first order statistics that rely on intensity values. While these approaches are appropriate for tissues that have little variation in background intensity, tissue heterogeneity can present a challenge for quantification. Thus, the goal of this work was to maintain high resolution imaging of tissue morphology through employing fluorescence microscopy and vital fluorescent stains, but to also develop a quantitative strategy to segment and quantify tissue features, such as nuclei and the surrounding stroma, which will enable automated diagnosis of thick tissues.

Specifically, my first aim was to develop an image processing method that can differentiate nuclei and nucleoli (which are collectively referred to as acriflavine positive features or APFs) from background tissue heterogeneity and enable automated diagnosis of thick tissue at the point of care. Sparse component analysis (SCA) was adapted to isolate APFs from background tissue heterogeneity. The circle transform (CT) was applied to the SCA output to quantify the size and density of overlapping APFs. The sensitivity of the SCA+CT approach to variations in APF size, density and background heterogeneity was demonstrated through simulations. Specifically, SCA+CT achieved the lowest errors for higher contrast ratios and larger APF sizes. The clinical
utility of this approach was examined by imaging acriflavine stained tumor margins with a fluorescence microendoscope. SCA+CT correctly isolated APFs in \textit{ex vivo} images and showed consistently increased density in tumor and tumor + muscle images compared to images containing muscle. A logistic regression model was optimized on excised tissue margins and prospectively applied to the panel of images obtained from the \textit{in vivo} tumor cavity of 34 mice. The sensitivity and specificity for differentiating positive from negative excised tumor margins was 82\% and 75\%. When applied prospectively to images from the \textit{in vivo} tumor cavity, the sensitivity and specificity for differentiating local recurrence was 78\% and 82\%. The results in chapter 2 indicated that SCA+CT could accurately delineate APFs in heterogeneous tissue, which is essential to enable automated and rapid surveillance of tissue pathology.

There are two primary challenges that were identified in the work in chapter 2. First, while SCA could be used to isolate features, such as APFs, from heterogeneous images, its performance was limited by the contrast between APFs and the background. Second, while it was feasible to create mosaics by scanning a sarcoma tumor bed in a mouse, which is on the order of 3-7 mm in any one dimension, it is not feasible to evaluate an entire human surgical margin. Thus, improvements to the microscopic imaging system were made to (1) improve image contrast through rejecting out-of-focus background fluorescence and to (2) increase the FOV while maintaining the sub-cellular resolution needed for delineation of nuclei and nucleoli. To address these challenges, a
technique called structured illumination microscopy (SIM) was employed in which the entire FOV is illuminated with a defined spatial pattern rather than scanning a focal spot, such as in confocal microscopy.

In chapter 3, image contrast and FOV were increased through the development of a wide-field fluorescence SIM system. Clear improvement in image contrast was seen in structured illumination images compared to uniform illumination images. Additionally, the FOV is over 13X larger than the fluorescence microendoscope used in chapter 2.

Next the segmentation algorithm for SIM images was optimized through a series of experiments. Segmentation results of SIM images with SCA revealed that SCA was unable to segment large numbers of APFs in the tumor images. Because the FOV of the SIM system was over 13X larger than the FOV of the HRME, dense collections of APFs commonly seen in tumor images could no longer be sparsely represented in the pixel basis. Thus, SCA was a suboptimal approach for segmenting APFs in SIM images because the sparsity constraint was no longer satisfied. Therefore, a technique called maximally stable extremal regions (MSER) was investigated as an alternative approach for image segmentation in SIM images. MSER was able to accurately segment 6 and 10 μm fluorescent beads, which were approximately the size of APFs commonly seen in the sarcoma model, in SIM images of tissue mimicking phantoms. In previous work, our group illustrated that the frequency of the illumination pattern used in SIM must be
carefully selected because the image signal to noise ratio is dependent on the grid frequency [114]. Therefore, images of tissue mimicking phantoms and fluorescently stained tumor sections were also acquired at different grid frequencies with the SIM system in order to select an optimal frequency. A grid frequency of 31.7 mm\(^{-1}\) led to the highest signal to background ratio, lowest percent error associated with MSER segmentation in phantom images, and highest density of APFs segmented with MSER in tissue images.

Once MSER was optimized for SIM image segmentation and the optimal grid frequency was selected, a quantitative model was developed to diagnose mouse sarcoma tumor margins that were imaged \textit{ex vivo} with SIM. Images of “pure” tissue types were used for characterization and model construction. The model was able to classify pure tumor and normal tissue with 77% sensitivity and 81% specificity. Next the model was applied to a subset of positive and negative margins. Initial results indicated that with a tumor probability threshold of 50%, 8% of all regions from negative margins exceeded this threshold, while 17% of all regions exceeded the threshold in the positive margins. Thus, 8% of regions in negative margins were considered false positives. These false positive regions were likely due to the high density of APFs present in normal tissues, which clearly demonstrated the challenge in implementing the algorithm based on AO staining alone.
Therefore, chapter 4 sought to improve the specificity of the diagnostic model through leveraging other sources of contrast. Modifications were made to the structured illumination microscopy (SIM) system to enable fluorescence imaging at a variety of wavelengths. Specifically, the SIM system was modified to enabling imaging of red fluorescent protein (RFP) expressing sarcomas, which were used to delineate the location of tumor cells within each image. Analysis of AO stained panels indicated that there was room for improvement in tumor detection, particularly in regards to false positive regions that were negative for RFP. One approach for improving the specificity of the diagnostic model was to investigate using a fluorophore that was more specific to staining tumor. Specifically, tetracycline was selected because it appeared to specifically stain freshly excised tumor tissue in a matter of minutes, and was non-toxic and stable in solution. Results indicated that tetracycline staining showed promise for increasing the specificity of tumor detection in SIM images of a preclinical sarcoma model and further investigation is warranted.

In summary, this work presented the development of a combination of tools that were capable of automated segmentation and quantification of micro-anatomical thick tissue images. When compared to the fluorescence microendoscope used in chapter 2, wide-field multispectral fluorescence SIM imaging provided improved image contrast, a larger FOV with comparable resolution, and the ability to image a variety of fluorophores. MSER was an appropriate and rapid approach to segment dense
collections of fluorescent positive features (FPFs) from wide-field SIM images. Variables that reflect the morphology of the tissue, such as the density, size, and shape of nuclei and nucleoli, could be used to automatically diagnose SIM images. The clinical utility of SIM imaging and MSER segmentation to detect microscopic residual disease was demonstrated by imaging excised preclinical sarcoma margins. This combination of technologies was particularly useful in deconstructing images of heterogeneous tissues. Ultimately, this work demonstrated that fluorescence imaging of tissue micro-anatomy combined with a specialized algorithm for delineation and quantification of features was a means for rapid, non-destructive and automated detection of microscopic disease, which could improve cancer management in a variety of clinical scenarios.

5.2 Future Work

Future work is primarily focused on two studies, which include (1) validating tetracycline staining and analysis in thick tissues, which may require optimizing the excitation wavelengths used in the SIM system and (2) clinical translation of tetracycline staining, SIM imaging, and MSER analysis to human specimens. The tetracycline preclinical validation study is discussed in section 5.2.1. The SIM technical improvements, which includes optimizing the excitation wavelength for tetracycline, are further discussed in section 5.2.2. Clinical translation to human breast specimens is discussed in section 5.2.3.
5.2.1 Tetracycline Preclinical Validation Study

While chapter 4 illustrated several ways to improve the specificity of the tumor detection, additional work is needed with a larger independent data set to validate the feasibility and reproducibility of these parameters. Results from chapter 4 indicated that tetracycline staining correlated well with pathology. In particular, tumor appeared to be well delineated in areas that correspond to sarcoma in matching H&E sections. Future studies could focus on validating tetracycline staining and analysis in thick tissues, such as the excised RFP tumor margins used in chapter 4. However, tetracycline is not as bright as AO, and longer integration times (~5 second) were required to capture images of tetracycline compared to AO (~1 second). Additionally, since RFP is extremely bright, a high degree of crosstalk was captured in the blue channel, which was used to image both AO and tetracycline. Since AO is bright, RFP crosstalk was insignificant when imaging AO. However, when imaging tetracycline, RFP crosstalk remained a dominant signal in the blue channel. Two representative RFP sarcoma margins were imaged to demonstrate this limitation and are shown in Fig 44. The first RFP margin was stained with tetracycline and imaged with both channels. The second RFP margin was not stained and was simply imaged with both channels in order to demonstrate crosstalk.
Figure 44: Illustration of RFP crosstalk in the blue channel. Two excised RFP margins were imaged. The first margin, which is shown in (a) and (b), was stained with tetracycline. The second margin, which is shown in (c) and (d) was not stained in order to demonstrate crosstalk in the blue channel (d). Scale bar 200 µm.

One approach to address the weak signal acquired from tetracycline is increase the optical power in the blue channel. Originally, the Fianium supercontinuum laser was selected because a wide range of wavelengths were available to image potential fluorophores. Specifically, the wavelengths produced by the Fianium range from 400 – 2500 nm; however, the peak wavelength (with the highest spectral power density) is around 1100 nm. Conversely, the power density is much lower in the blue wavelengths
(400-500 nm). A graph of the spectral power density versus wavelength is shown in Fig 45. The particular model included in the SIM system described in this work is the Fianium SC400-2. As seen, the power density at 480 nm, which was the center wavelength used to excite both tetracycline and AO, is less than 1 mW/nm. In order to increase the excitation power, a new source must be explored. Laser diodes or LEDs can provide a feasible option cost significantly less than the Fianium laser source.

Figure 45: Spectral power density of the Fianium supercontinuum laser used in the SIM system. The particular model used in this work is the Fianium SC400-2, which is shown in red.

The blue channel was initially used to image tetracycline because the excitation and emission spectra appeared to be similar to AO. Specifically, Pautke et al. used a 458 nm argon laser to excite all tetracycline derivatives investigated in their study, and found the peak emission for tetracycline was located around 529 nm [140]. However, conflicting reports of the peak excitation of tetracycline were found in the literature. While Pautke et al. used 458 nm excitation, others used 390 nm [77] or 402 nm [78] to
excite tetracycline derivatives. To measure the peak excitation and emission of
tetracycline and AO, fluorescence spectra were acquired with a standard fluorometer
and are shown in Fig 46. As seen, the peak excitation and emission of AO were
approximately 470 nm and 550 nm respectively. Conversely, the peak excitation and
emission of tetracycline were approximately 355 nm and 530 nm respectively. While the
blue channel (which contained a 480/30 nm bandpass excitation filter and 515 nm long
pass emission filter) was well suited to excite and capture the emission of AO, it was not
ideal for tetracycline imaging. In particular, lower excitation wavelengths could be used
to acquire more fluorescence signal from tetracycline. However, the excitation spectra
from the concentrated solution of tetracycline (100 mg/mL), which was used in the
studies in chapter 4, revealed that some excitation was achieved between 450 and 500
nm. Additionally, the emission spectra from tetracycline confirmed that a 515 long pass
emission filter was appropriate for tetracycline imaging. Thus, while some fluorescence
signal from tetracycline was captured by the blue channel, the weak signal could be
improved by using a more powerful source, such as a deep ultraviolet LED with a peak
wavelength centered around 355 nm.
Figure 46: Excitation and emission spectra for AO and tetracycline (tetra) are shown in green and blue respectively. Excitation spectra are indicated by the dotted lines and emission spectra are indicated by solid lines. Both a dilute and concentrated solutions of tetracycline were measured to acquire the excitation spectra.

Lastly, the multispectral SIM system and sarcoma model is an excellent test bed to assess how additional contrast agents stain thick tissues. Both yellow fluorescent protein (YFP) and RFP-expressing sarcomas are readily available [102]. Thus depending on the excitation and emission wavelengths of the fluorophore, either YFP or RFP sarcomas could be imaged to indicate where tumor is located within an image and can be directly compared to the fluorophore imaging results.

5.2.2 SIM Technical Improvements

A number of technical improvements can be made to increase the viability and efficacy of the SIM system. As mentioned previously, increasing the power density, particularly in the deep ultraviolet wavelengths, could enable better visualization of tetracycline. Additionally, increasing the excitation power could decrease the image
acquisition time. In its current form, the image acquisition time for AO stained sarcoma tissue is 1-2 seconds. This is compounded by the fact that SIM requires three images to be captured at each location. For this technique to be truly scalable to larger human surgical margins, the acquisition time must be reduced. Thus, new sources, such as laser diodes or LEDs must be explored.

In order to enable imaging larger samples in clinically relevant time windows, video-rate structured illumination microscopy for high-throughput imaging of large tissue areas could be explored. Brown et al recently reported the development of a next-generation Rapid Optical Sectioning Specimen Scanner (ROS3) based on incoherent structured illumination microscopy, which was specifically designed for rapid, high-area throughput fluorescence microscopy of intact surgical and biopsy specimens [143]. Specifically, they used the latest high-speed digital SLM and sCMOS camera technology to achieve high-resolution, optically sectioned images of fluorescent samples at area-throughput rates of up to 4.4 cm²/min with 1.3 µm lateral resolution [143]. Similar improvements could be made to the SIM system described here in order to enable imaging of large clinical specimens.

**5.2.3 Clinical Translation**

As discussed previously, several clinical situations could benefit from more rapid histological processing, which could reduce the time and the number of steps required between obtaining a fresh tissue specimen and rendering a diagnosis. The clinical
implications of this work were described in detail in section 1.2. For example, there is need for rapid detection of residual cancer on the surface of tumor resection specimens, which are known as positive tumor margins [4]. Current pathology techniques to determine tumor margin status are too slow and labor intensive to allow comprehensive analysis of large surgical specimens. More rapid assessment of tumor margins could prevent cancer from being left behind in the patient during surgical procedures.

The studies described in this work were conducted in a preclinical sarcoma model in order to assess whether a combination of imaging tools could detect residual disease in resected tumor cavities. Moving forward, the goal is to expand this work to breast tumor margin assessment in humans. Towards this end, our group has is focusing on two different projects. The first project is validating that morphological staining, fluorescence microscopy, and automated segmentation of features enables rapid visualization and quantitative interpretation of breast tissue morphology. Thus, images of freshly excised breast tissue specimens were acquired using confocal fluorescence microscopy and acriflavine as a topical stain. Computerized algorithms to segment and quantify nuclear and ductal parameters were developed to that characterize breast architectural features. Details of this work are included in section 5.2.3.1. The second project is to combine high resolution morphological imaging with wide-field devices, such as our group’s spectral imaging platform [141]. While it is feasible to create mosaics by scanning a tumor bed in a mouse, which is on the order of 3–7 mm in any one
dimension, it is not feasible to evaluate an entire human surgical margin, in which a single breast resection margin areas can cover up to 40 cm² [130]. Thus we foresee combining SIM imaging with a portable optical spectral imaging (OSI) system for wide-field imaging of tumor margins. Further details are included in section 5.2.3.2.

5.2.3.1 Micro-Anatomical Quantitative Confocal Microscopy Towards Automated Assessment of Clinical Breast Tissues

In order to characterize quantitative criteria to classify breast architecture, several studies have described segmentation algorithms based on nuclear [144-148] and ductal [149-151] morphometry in images of fixed tissue stained with hematoxylin and eosin (H&E) staining. Additionally, some recent studies evaluated nuclear morphometric parameters using wide-field fluorescence microscopy [152] and micro-optical computed tomography [153] to acquire images of breast tissue. Specifically, wide field fluorescence microscopy combined with watershed segmentation to quantify nuclei found that area fraction could distinguish between tumor and normal regions in excised rat mammary tissue with 97% accuracy [152]. Micro-optical computed tomography and nuclear morphometry was used to compare variations between human breast cell lines and found that nuclear volumes increased from normal to metastatic breast cells and that nuclei of abnormal cells contained more nucleoli [153].

The idea of establishing quantitative criteria on fixed tissue can be taken one step further to be applied directly to intact specimens using other imaging modalities, which can obviate the need for extensive tissue processing. Several studies have already
described the feasibility of imaging breast tissue with confocal microscopy in a clinical setting, [154-159]. Schiffhauer and colleagues showed that confocal reflectance microscopy could be used to image benign and malignant breast features and provide visual similarity to H&E micrographs [156]. Abeytunge and colleagues demonstrated that confocal fluorescence microscopy can be used to rapidly acquire images of fresh tissue specimens between 1-2.5 cm² in size [156]. Kortum and colleagues recently showed that confocal fluorescence microscopy yields images with sufficient detail to identify benign and malignant breast architecture in freshly-excised tissue [154]. In another recent study, Kortum et al. demonstrated that confocal fluorescence images can be used to estimate percent tumor cellularity in core needle biopsy specimens and can indicate the adequacy of procured tissue for diagnosis and ancillary molecular and immunophenotypic studies [155].

The goal of this work is to combine both quantitative image processing techniques with optical microscopy of intact breast tissue specimens for interpretation of breast tissue at the point of care. The benefits of this approach are minimal tissue processing, rapid diagnosis, and quantitative criteria that could potentially reduce the subjectivity with intra- and inter-observer variation in the evaluation of breast histology. In this study, we combine clinical confocal microscopy with a computerized image processing algorithm to quantify both nuclear and ductal morphology of breast tissue; we develop an algorithm using these parameters to classify breast tissue as benign or
malignant. Although previous studies have described evaluation of breast architecture in histologic images [144-151], these studies only considered either nuclear or ductal parameters. We show that combining both yields improved diagnostic performance, particularly in the diagnosis of invasive ductal cancer (IDC) and ductal carcinoma in situ (DCIS). The APF and ductal parameters described in this study could potentially be used for objective categorization of breast lesions.

5.2.3.1.1 Methods

5.2.3.1.1 Breast Tissue Acquisition and Preparation

Fresh human breast tissue specimens were acquired through a protocol approved by The University of Texas MD Anderson Cancer Center and Rice University Institutional Review Boards. Fresh breast tissue was acquired from patients undergoing surgery to excise a clinically abnormal lesion. The procedure for tissue preparation has been described previously [154]. In brief, two tissue specimens - one grossly abnormal and one grossly normal in appearance were acquired from each patient for image acquisition and evaluation; each specimen measured approximately 15 x 15 mm² in size, with thickness varying from 2-7 mm. Within 30 minutes of surgical excision, breast tissue specimens were stained for one minute in a solution of 0.01% acriflavine in 1X phosphate buffered saline (PBS). Acriflavine is a nuclear contrast agent [160, 161], which has been used to stain breast tissue, oral mucosa, Barrett’s esophagus, cervical tissue, and sarcoma in previous studies [76, 154, 155, 162-166]. Following topical application of acriflavine, specimens were washed with 1X PBS and then immediately imaged.
5.2.3.1.2 Image Acquisition and Evaluation
Confocal fluorescence images were acquired from multiple sites within each specimen using a multi-wavelength scanning confocal microscope (Vivascope 2500®, Caliber Imaging and Diagnostics) as described previously [154, 155, 167]. Following topical application of acriflavine and the PBS wash, each tissue specimen was positioned on the microscope stage and imaged using 2.1 ± 0.4 mW power at 488 nm laser excitation, and the fluorescence was detected in a band pass of 550 ± 44 nm with a 30x water immersion lens. At these settings, the lateral and axial resolution was 1.0 µm and 5.0 µm, respectively, in the center of the 750 x 750 µm² field of view. A 12 x 12 mm² composite image was created for both sides of each tissue specimen. To create the composite image, images were acquired from contiguous sites in a grid pattern (maximum area 12.2 x 12.2 mm²) over the surface of the specimen at an approximate depth of 20 µm. Following image acquisition, specimens were kept moist in 1X PBS and were submitted for routine histologic preparation and fixation. Samples were stained with H&E and fixed on microscope slides for histologic assessment.

A board-certified, breast pathologist (author S. Krishnamurthy) viewed composite confocal images and fixed tissue specimens stained with H&E using a conventional light microscope to identify sites that corresponded to the same approximate location in the specimen based on similar image morphology. Specifically we selected in-focus confocal microscope fields of view that contain representative examples of characteristic benign and malignant breast features. Thus, at each site, a
corresponding pair of confocal and H&E images were available from a 750 µm x 750 µm field of view. At each site, the H&E images of fixed tissue specimens were used as a reference standard to identify breast architectural features that should be present in corresponding confocal images [30, 154]. Benign breast features identified in reference H&E images included adipose and fibrous tissue, lobules, non-hyperplastic ducts, and ductal hyperplasia. Malignant breast features identified in reference H&E images included: ductal carcinoma in situ (DCIS), invasive ductal carcinoma (IDC), and invasive lobular carcinoma (ILC).

5.2.3.1.3 Segmentation and Connected Components Algorithms for Identifying APFs
MSER has previously been described in section 3.3.1.4. In order to apply MSER to the confocal images of breast tissue, the five tuning parameters associated with MSER were re-selected. The first two parameters, which included the minimum area (MinArea) and maximum area (MaxArea) of the connected components, are related to the expected size of APFs. These parameters were selected based on the biologically expected range of nuclear diameters. Specifically, other groups have found nuclear volume to range from approximately 200 to 1500 µm$^3$, which corresponds to 7 to 14 µm in diameter [153]. Therefore, MaxArea was set to 500 pixels, which corresponds to 19 µm in diameter, which is larger than the expected nuclear size for our images. MinArea was set to 35 pixels, which corresponds to 5 µm in diameter, which is smaller than the expected nuclear size for our images. The next set of parameters is related to the intensity thresholds and includes maximum variation (MaxVariation), minimum diversity
(MinDiversity), and Delta. These intensity parameters were systematically tuned through applying a range of values to representative images in order to select the best value for each parameter. Specifically, one input parameter was varied over a wide range while other input parameters were held constant. For each iteration, the area fraction (AF) from representative images of tumor and normal tissue was calculated and overlays of the features isolated with that particular setting were displayed. The values that led to the largest differences in AF between tumor and normal tissues, while isolating features that approximately corresponded to nuclei or nucleoli, which we refer to as acriflavine positive features (APFs), were selected. Specifically, MaxVariation was set equal to 2.5, MinDiversity to 0.5, and Delta to 6. These parameter values are in terms of relative intensity, which for our 8 bit images ranges from 0 to 255.

After APFs were isolated with MSER, a connected components algorithm was applied in order to calculate parameters such as the density and diameter of APFs. In the connected components algorithm, all touching or connected pixels are assumed to belong to the same APF. Parameters include APF density (the number of APFs in a unit area), area fraction (the total APF area divided by the total area), minimum inter-APF distance (the distance from an APF center to the next closest APF center), and APF diameter (the length of the major axis of each APF). APF density and area fraction (AF) represent scalar variables – only one value is returned for each image, while the minimum inter-APF distance (IND) and APF diameter represent vector variables – a
value is calculated for each APF in the image. In order to consolidate the vector variables into a scalar value, several summary statistics were evaluated, including mean, median, mode, interquartile range, and standard deviation.

5.2.3.1.1.4 Ductal segmentation algorithm and quantification of ductal parameters
An algorithm was developed to measure ductal parameters, which segments non-hyperplastic ducts, ductal hyperplasia, and DCIS lesions based on the intensity of acriflavine staining. An illustration of both the APF and ductal segmentation is shown in Fig 47. To reduce noise and increase image contrast, a Wiener lowpass filter was first applied followed by contrast-limited adaptive histogram equalization (CLAHE). Images were converted from grayscale to binary using a user-defined threshold based on relative intensity. The mean threshold used to segment ducts was 107 ± 27 (range: 52-168) on a scale of 0 to 255 for 8 bit images. It was not possible to select a universal threshold, because in order to accurately segment ducts from surrounding tissue, it is necessary to isolate both nuclei in the duct walls and inter-nuclear space between them. The relative intensity of these features differed between images due to the variation in illumination power used for image acquisition and the variation in acriflavine staining. Areas smaller than the upper threshold for APFs (approximately equivalent to 280 µm² or 500 pixels, with a diameter of 19 µm [168]) were removed to avoid segmenting individual APFs outside of the duct walls. Individual ducts were manually segmented using a user-defined polygon selection tool to define architectural features corresponding to breast ducts. After application of the ductal segmentation algorithm,
the binary confocal image showed the segmented duct walls and the outer and inner boundaries of the duct used to measure ductal parameters.
Figure 47: Algorithms for APF (A-D) and ductal (E-L) segmentation. APF segmentation: A: Raw image acquired from confocal fluorescence microscope with 750 x 750 µm² field of view. B: Region of interest selected in confocal fluorescence image with 75 x 75 µm² field of view. C: The maximally stable extremal regions (MSER) algorithm applies thresholds from 0 to 255 to B. D: At each threshold, the
MSER algorithm identifies APFs as connected components and selects “maximally stable” components with the lowest size variation. Ductal segmentation: E: Raw image acquired from confocal fluorescence microscope with 750 x 750 μm² field of view. F: Wiener low pass filter and adaptive histogram equalization applied to E. G: The algorithm converts E to a binary image using an interactive threshold tool. H: Objects below range of APF area are removed and then user selects a region of interest (ROI) around ducts with an interactive polygon selection tool. I: The algorithm fills boundaries of ducts identified in H to segment the outer boundaries of the duct. J: The algorithm selects the complement of H to segment the inner boundaries of the duct (lumen). K: Duct wall width is measured by selecting the shortest distance from the outer to the inner duct boundaries (red lines). L: Ellipses are fitted to outer and inner duct boundaries. E-J: scale bar is 100 μm. K,L: scale bar is 25 μm.

Following segmentation of ducts, a number of ductal parameters were measured based on the properties of the inner and outer duct boundaries. The outer boundary defines the outer edge of the duct wall and the inner boundary defines the inner edge of the duct wall; the lumen. The width of the duct wall was measured at every pixel on the outer edge of the duct wall. This was done by finding the shortest distance between every point on the outer boundary and the nearest point on the inner boundary. Duct wall width was measured for each non-hyperplastic duct, ductal hyperplasia, and DCIS lesion and the vector of values were summarized by calculating the mean, median, mode, interquartile range, and standard deviation. Other scalar parameters measured include the area of the duct wall, area of the lumen, area of an ellipse approximating the duct wall, area of an ellipse approximating the lumen, lengths of the major and minor axes for the duct and the lumen, solidity of the duct and the lumen, and eccentricity of the duct and the lumen.
5.2.3.1.5 Statistical analysis and model building

APF parameters were calculated for all sites (n = 259) and ductal parameters were calculated for all sites that contained ducts (n=50), and the diagnostic performance of each image parameter was individually assessed by determining the classification accuracy. Two-class linear discriminant analysis was performed to classify malignant from benign breast architectural features based on each individual APF or ductal parameter; receiver operator characteristic (ROC) curves were constructed and area under the curve (AUC) was calculated for each ROC curve. Sensitivity and specificity values were determined at the optimal cutpoint. Parameters were sorted by accuracy for classification of neoplasia based on AUC values. Boxplots were created for the parameters with the highest AUCs. A Student t-test for samples with unequal variances was used to identify statistically significant differences between mean parameter values measured in benign and malignant tissues. This analysis was performed to evaluate individual APF and ductal parameters to incorporate into a classification model.

Next we sought to develop a multivariate model to yield optimal separation between benign and malignant tissues. Towards that end, all 33 APF and ductal variables were used as input for a classification and regression tree (CART) function in Matlab. Decision trees were constructed using the automated Matlab function classregtree, which selects parameters and cutpoints that lead to the optimal classification of benign and malignant breast architectural features. Decision trees were pruned to prevent a single APF or ductal from being used at more than one node within
the tree. Pruning was also performed to prevent the number of categories for classification of malignant breast features from exceeding 3: the number of malignant tissue types (IDC, ILC, and DCIS). After construction, decision tree nodes were pruned by finding the next higher node whose decision point led to two categories, one with a majority of neoplastic sites, and one with a majority of benign sites. A custom leave one out cross-validation algorithm was also developed in order to calculate the cross-validated sensitivity and specificity. Specifically, 258 of the 259 data points were used to build a CART model, which contained the same two variables at the first and second decision points as is seen in Figure 5. Specifically the standard deviation of IND (StdIND) was the first decision point and the number of lumens was the second decision point. However, with each iteration of leave one out cross-validation, the cutoff value of StdIND could vary. The cutoff value associated with the number of lumens (number of lumens >1) was held constant because biologically normal ducts are expected to only contain a single lumen; therefore, this was considered to be the optimal and only logical cutoff value and therefore was held constant. Then the model was applied to the remaining data point, which was classified as either benign or malignant. This process was repeated for all 259 data points, and the calculated diagnosis for each image was compared to the known diagnosis in order to calculate sensitivity and specificity for the cross-validated model. The performance of the decision tree was characterized by computing sensitivity and specificity for classification of malignant breast architectural
features. Additionally, sensitivity and specificity were calculated for each individual histologic type of malignant tissue in order to determine the relative classification accuracy for IDC, ILC, and DCIS sites. For example, in order to calculate sensitivity for IDC, true positives were defined as IDC sites that had been classified as malignant by the decision tree, and false negatives were defined as IDC sites that had been classified as benign. Specificity was calculated by defining true negatives as benign sites that were correctly classified in the decision tree and false positives were defined as benign sites that were incorrectly classified. An ROC curve was constructed for the decision tree model. All sites were sorted in order of ascending StdIND value and then sensitivity and specificity for classification of neoplasia were calculated at every StdIND value. The cutoff value for number of lumens was held constant at 1 lumen because biologically normal ducts are expected to only contain a single lumen. AUC was calculated based on the resulting ROC curve.

### 5.2.3.1.2 Results

A total of 259 sites from 36 patients were identified in composite confocal fluorescence images. A summary of patients, sites, and diagnoses are included in Table 10. In total there were 179 benign sites, which included adipose tissue, fibrous tissue, lobules, and benign ducts, and 80 malignant sites, which included DCIS, IDC, and ILC.

**Table 10: Summary of patients from which tissue specimens were acquired, sites analyzed, and histologic diagnoses**

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Patients</th>
<th>Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>Row 1</td>
<td>Row 2</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>Adipose tissue</td>
<td>18</td>
<td>42</td>
</tr>
<tr>
<td>Fibrous tissue</td>
<td>16</td>
<td>31</td>
</tr>
<tr>
<td>Lobules</td>
<td>12</td>
<td>82</td>
</tr>
<tr>
<td>Non-hyperplastic ducts</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>Hyperplastic ducts</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td><strong>Malignant</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ductal carcinoma in situ (DCIS)</td>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td>Invasive ductal carcinoma (IDC)</td>
<td>15</td>
<td>37</td>
</tr>
<tr>
<td>Invasive lobular carcinoma (ILC)</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>36</td>
<td>259</td>
</tr>
</tbody>
</table>

Fig 48 shows representative confocal images of sites without ducts acquired by confocal fluorescence microscopy in row 1 and APFs isolated with MSER at those sites in row 2. Row 3 shows sites in the corresponding histologic slide with H&E staining that have similar histology to the confocal sites. APFs were false-colored green and overlaid onto the original images for visualization. As seen, APFs are isolated at the periphery of adipose cells and are dispersed throughout the fibrous tissue image. Denser clusters of APFs are isolated in and around lobules. APFs are the densest at sites with malignant tissue, including IDC and ILC.
Figure 48: Representative raw confocal fluorescence images of adipose tissue, fibrous tissue, lobules, invasive ductal carcinoma, and invasive lobular carcinoma are shown in A through E, respectively. F-J: APFs segmented by identifying maximally stable extremal regions (MSER) are false colored green and overlaid onto the raw confocal fluorescence image. K-O: Histologic slides with H&E staining show similar histology to confocal images in A-E. Slides were prepared with the same specimens from which confocal images were acquired. Scale bar is 100 µm.

Fig 49 shows representative images of breast ducts acquired with confocal fluorescence microscopy in row 1, APFs that were isolated at sites with breast ducts using MSER in row 2, and ducts that were segmented with the ductal segmentation algorithm in row 3. Row 4 of Figure 3 shows sites in the corresponding histologic slide with H&E staining that have similar histology to the confocal sites. APF density in and around the ducts increases from the non-hyperplastic duct, to the hyperplastic duct, to DCIS. However, relatively few APFs are successfully isolated using MSER within the
non-hyperplastic and hyperplastic ducts, which is most likely due to the fact that the borders of individual APFs are difficult to visually discern in confocal fluorescence images. The images of sites isolated with the ductal segmentation algorithm show well-defined lumens in both the non-hyperplastic duct and hyperplastic duct. Conversely, the image of DCIS shows bridges of cells crossing the lumen to create a cribriform pattern with several lumens.
Figure 49: Representative confocal images of normal, non-hyperplastic ducts (A), hyperplastic ducts (B), and ductal carcinoma *in situ* (C) analyzed with the APF segmentation algorithm (middle row) and with the ductal segmentation algorithm (bottom row). D-F: APFs segmented by identifying maximally stable extremal regions (MSER) are false colored green and overlaid onto the raw confocal fluorescence image. G-I: Breast ducts segmented using the ductal segmentation algorithm. J-L: Histologic slides with H&E staining show similar histology to confocal images in A-
E. Slides were prepared with the same specimens from which confocal images were acquired. Scale bar is 100 µm.

The parameters that yielded the highest performance for distinguishing between benign and malignant sites are shown in Table 11. We evaluated the performance of APF parameters for classification of benign and malignant features in all sites and in sub-groups of sites that did or did not contain ducts to determine the groups for which APF parameters had the highest classification accuracy. We only evaluated the classification accuracy of ductal parameters at sites that contained ducts. APF parameters measured at non-duct sites achieve higher performance (AUC = 0.93) than APF parameters measured at duct sites (AUC = 0.69). Conversely, ductal parameters achieve higher performance (AUC = 0.92) than APF parameters for classification of duct sites (AUC = 0.69). These findings suggest that a combination of APF parameters measured at non-duct sites and ductal parameters measured at duct sites may yield improved separation between all benign and malignant sites.

Table 11: Summary of top performing parameters for distinguishing between benign and malignant sites measured using algorithms

<table>
<thead>
<tr>
<th>Group</th>
<th>Performance metric</th>
<th>Standard deviation of IND</th>
<th>Area fraction</th>
<th>Range of IND</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Classification by APF</td>
<td>AUC</td>
<td>0.87</td>
<td>0.86</td>
<td>0.87</td>
</tr>
</tbody>
</table>
### B. Classification by APF

<table>
<thead>
<tr>
<th>Parameter – Non-duct Sites</th>
<th>AUC</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.93</td>
<td>0.92</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>0.92</td>
<td>0.80</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>0.91</td>
<td>0.87</td>
<td>0.88</td>
</tr>
</tbody>
</table>

### C. Classification by APF

<table>
<thead>
<tr>
<th>Parameter – Duct Sites</th>
<th>AUC</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.68</td>
<td>0.72</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>0.65</td>
<td>0.65</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>0.70</td>
<td>0.70</td>
<td>0.96</td>
</tr>
</tbody>
</table>

### D. Classification by Duct

<table>
<thead>
<tr>
<th>Parameter – Duct Sites</th>
<th>AUC</th>
<th>Number of lumens</th>
<th>Minor dimension of outer ellipse</th>
<th>Area of outer ellipse</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.92</td>
<td>0.83</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.88</td>
<td>0.73</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.88</td>
<td>0.79</td>
<td>0.75</td>
<td></td>
</tr>
</tbody>
</table>

Boxplots showing the mean and interquartile range of the top three performing APF parameters are shown in **Fig 50**. Both Std IND and Range IND decrease from adipose to fibrous to lobules to ILC to IDC, while AF increases from adipose to fibrous to lobules to ILC to IDC. This trend suggests that the number of clusters of APFs increases from adipose tissue, which has the fewest, to IDC, which has the greatest number of clusters of APFs. All comparisons between benign (adipose, fibrous, lobules) and malignant (IDC, ILC) sites were significant. Similarly, AF increases from adipose to fibrous to lobules to IDC, which suggests increasing APF density.

Boxplots showing the mean and interquartile range of the top three performing ductal parameters for duct sites are also shown in **Fig 50**. DCIS lesions have a significantly higher number of lumens than hyperplastic and non-hyperplastic ducts (p < 0.001), which is consistent with the cribriform pattern that occurs when abnormally high cellular proliferation causes the luminal space to be filled with epithelial cells. The minor dimension of the outer ellipse approximating the duct is significantly smaller in

---

179
normal, non-hyperplastic ducts than in DCIS lesions (p < 0.001). There is no significant difference in the minor dimension between ellipses approximating hyperplastic ducts and DCIS lesions. The area of the outer ellipse approximating duct area was significantly smaller in normal, non-hyperplastic ducts than in DCIS lesions (p < 0.001). There is no significant difference between the average area of outer ellipses approximating hyperplastic ducts and DCIS lesions.

Figure 50: Mean value of parameters used to separate malignant from benign sites. APF parameters calculated with the APF segmentation algorithm are shown for all adipose, fibrous, lobules, invasive ductal carcinoma (IDC), and invasive lobular carcinoma (ILC) sites; A: standard deviation of IND; B: area fraction; C: range of IND. Ductal parameters calculated with the duct-based segmentation algorithm are shown for all normal, non-hyperplastic ducts, hyperplastic ducts (Hyperplasia), and ductal carcinoma in situ (DCIS); D: number of lumens; E: minor dimension of outer ellipse;
F: area of outer ellipse. The number of sites represented in each box is represented by n. Significant differences between mean values of parameters measured at benign and malignant sites are indicated by asterisks (*).

All 33 APF and ductal parameters were used as input for a classification and regression tree (CART) algorithm to automate selection of parameters to discriminate benign and malignant sites. The CART algorithm was pruned to remove redundancies and over-fitting to the data set. The classification tree generated through this process is shown in Fig 51. Std IND with a cutoff value of 6.83 µm is the first decision point selected for classification by the decision tree, followed by number of lumens with a cutoff value of 1. Std IND < 6.83 µm separates out 52 true positives composed of IDC, DCIS, and ILC sites and 9 false positives composed of fibrous and lobule sites. The remaining sites enter the second node – Number of lumens > 1– which separates out 13 true positive DCIS sites and 3 false positive hyperplasia and normal duct sites. The remaining sites are classified as benign and are composed of 167 true negative adipose, fibrous, lobule, normal duct, and hyperplasia sites and 15 false positive IDC, DCIS, and ILC sites. Overall, the model achieved a sensitivity and specificity of 81% and 93% respectively, corresponding to an area under the curve of 0.93 and 90% overall classification accuracy, as shown in Table 12. If the model is evaluated based on classification of individual histologic types of neoplasia, 92% of IDC sites and 96% of DCIS sites were classified correctly. However, the model correctly classified only 35% of ILC sites. Additionally, leave one out cross-validation was performed, which yielded a
cross-validated sensitivity of 75% and specificity of 93%. Specifically, cross-validation resulted in a 6% drop in sensitivity (from 81% to 75%) due to the fact that 5 additional IDC images were incorrectly classified during cross-validation. When each of these 5 cases were left out of the original cohort of data used to form the model (in other words during the leave one out cross-validation exercise), the cutoff value associated with StdIND dropped. This resulted in each of the 5 cases being classified as a false negative. For the remainder of the images, the cutoff value associated with StdIND remained the same as it is in Figure 5, resulting in the same specificity of 93%.

Figure 51: Classification tree automatically generated when all APF and duct data was used. Duct- and APF-based parameters selected by classification regression
tree analysis to optimize separation between benign and malignant sites. Bar graphs show the diagnoses of sites sorted into each classification category.

Table 12: Performance of model for classification of neoplasia, non-neoplasia, and individual histologic types of breast neoplasia

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification Tree Model</td>
<td>81% (65/80)</td>
<td>93% (167/179)</td>
</tr>
<tr>
<td>Cross-validated Model</td>
<td>75% (60/80)</td>
<td>93% (167/179)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Correctly Classified</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Sites</td>
<td>90% (232/259)</td>
</tr>
<tr>
<td>DCIS</td>
<td>96% (25/26)</td>
</tr>
<tr>
<td>IDC</td>
<td>92% (34/37)</td>
</tr>
<tr>
<td>ILC</td>
<td>35% (6/17)</td>
</tr>
</tbody>
</table>

As seen in the histograms in Fig 51, ILC sites account for the largest number of false negative (n = 11 out of 17 sites) while lobule sites account for the largest number of false positives (n = 8 out of 82 sites). Fig 52 shows representative confocal images of a true positive ILC, false negative ILC, true negative lobules, and false positive lobules sites in row 1 and APFs isolated with MSER at those sites in row 2. Row 3 shows sites in the corresponding histologic slide with H&E staining that have similar histology to the confocal sites. As seen, there are large differences in the density and clustering of APFs between the true positive ILC site and true negative lobules site. In comparison, the false negative ILC site in has relatively few APFs, which appear to be predominately clustered in the upper left region of the image. Conversely, the false positive lobules site contains more APFs than Fig 52g, particularly in stromal tissue located in between lobules.
Figure 52: Representative images of sites with lowest classification accuracy in the decision tree model. A-D: invasive lobular carcinoma and lobules in confocal fluorescence images. E-H: APFs segmented by identifying maximally stable extremal regions (MSER) are false colored green and overlaid onto the raw confocal fluorescence image. I-L: Histologic slides with H&E staining show similar histology to confocal images in A-E. Slides were prepared with the same specimens from which confocal images were acquired. A, E: a true positive invasive lobular carcinoma (ILC) site; B, F: false negative ILC site; C, G: true negative lobules; and D, H: false positive lobules. Scale bar is 100 µm.

5.2.3.1.3 Discussion

In this study, we performed quantitative analysis of breast histology in confocal fluorescence images by designing algorithms to segment and measure APF and ductal parameters. We combined APF and ductal parameters to develop a classification tree
model to classify malignant from benign changes in the breast parenchyma with 81% sensitivity and 93% specificity, which corresponded to an AUC of 0.93 and an overall accuracy of 90%. The cross-validated model classified the same sites with 75% sensitivity, 93% specificity, and 88% overall accuracy.

Several groups have used automated morphometric evaluation of nuclei in H&E stained sections of breast tissue [144-147], cytological smears of breast tissue [148], and fluorescence microscopy images of mouse tissue [152, 166] to classify benign and malignant breast features. While these groups demonstrate that quantitative nuclear parameters can be used to classify benign and malignant breast features, some lesions are more difficult to distinguish. For example, Rajesh et al used automated nuclear morphometry to classify ILC, IDC, and borderline lesions [146]. While significant differences were found between parameters measured for ILC and IDC, no significant difference was found between parameters measured for ILC and benign borderline lesions [146]. We found similar results to the other studies – namely that ILC is difficult to distinguish from non-neoplasia based on APF features alone. Additionally, several studies have demonstrated the feasibility of computerized image analysis to distinguish between non-hyperplastic ducts, hyperplastic ducts, and DCIS. Mayr et al used computerized image analysis to quantify ductal parameters in H&E-stained slides of breast biopsies and found that the most significant parameters for differentiation between normal ducts and DCIS were duct mean diameter and the presence of necrosis.
Anderson et al. used a computerized segmentation algorithm to measure parameters of ductal hyperplasia and DCIS in tissue sections stained with the antibody cocktail AE 1/3, and showed that the highest classification accuracy for DCIS was achieved by combining parameters of ducts and lumina [150]. The findings from our work agree with previous studies, which showed that quantitative ductal parameters can be used to classify benign and malignant ducts [149, 150].

The strengths of our study are that we demonstrate that APF and ductal parameters can be measured in confocal fluorescence images of clinical samples acquired at the point of care. We perform quantitative analysis of breast tissue architecture without requiring tissue fixation, cutting, and staining and achieve comparable classification accuracy to studies that performed computerized analysis on fixed breast tissue stained with H&E. The model classified IDC and DCIS with greater than 90% accuracy using parameters that were based on the morphological characteristics of each malignant tissue type. Specifically, IDC was classified with 92% accuracy using standard deviation of IND as a parameter, which identifies dense clusters of APFs. DCIS was classified with 96% accuracy based on the presence of more than one lumen, which is consistent with the cribriform pattern. Overall we achieve high performance (AUC = 0.93) on a large number of sites (n = 259).

There are several limitations associated with this study. While our initial data set contains a large number of sites (n = 259), the data was acquired at a single center (The
University of Texas M.D. Anderson Cancer Center), and some individual categories, such as ILC contain relatively few sites (n = 17); therefore, additional work is needed with a large, independent data set composed of data from more than one center to validate the feasibility and reproducibility of these parameters. However, our initial results indicate that leave one out cross-validation of the CART model yields similar performance to the original model suggesting that our model may generalize to an independent data set. Another limitation to the study are the large variances observed for the APF parameters that reflect the high degree of heterogeneity in APF size and spacing in benign and malignant breast epithelia. Changes in APF area and spacing in breast epithelia occur frequently in both ductal and lobular nuclei and can be due to a number of clinical features, including sexual maturity, pregnancy, menopausal status, use of hormonal contraceptives, and presence of mammary carcinoma [169]. The presence of heterogeneity in APF area and distribution within benign breast tissue is a potential source of variance for the nuclear parameters measured in this study. Similarly, IDC and ILC typically contain irregular nuclear sizes and an irregular distribution of nuclei [31, 170], which is another potential source of variance within APF parameters. In addition, the algorithm designed for ductal segmentation uses an interactive threshold to convert images from grayscale to binary and a user-defined selection tool to isolate ducts from surrounding nuclei. The ductal segmentation process is a potential source of variability between users, particularly for parameters that could
be impacted by a user’s visual assessment of the duct wall boundaries, such as duct wall width. However, the ductal parameter that was ultimately selected for the decision tree model was the number of lumens, which is unlikely to vary at the decision point (number of lumens greater than 1) based on slight variations to the threshold value or by excluding surrounding nuclei. This is because it is readily apparent if a duct has one or more lumens based on visual assessment, however the segmentation algorithm could assist in identifying ducts with more than 1 lumen. Lastly, examination of the breakdown of false negatives and false positives reveals that our algorithm does most poorly at distinguishing ILC and lobule sites. Specifically 65% (n = 11 out of 17 sites) of ILC sites and 10% of lobule sites (n = 8 out of 82 sites) are incorrectly classified. Figure 6 reveals that there are differences in quantity and clustering of APFs between the true positive and false negative ILC sites. In particular, APFs in the false negative ILC site appear to be predominately located in the upper left region of the image, suggesting that only the upper left region of the image contains ILC while the remainder of the image may contain other benign tissue. Therefore, the fraction of the image that consists of a malignant tissue type may correlate with the likelihood that it is correctly classified as a true positive site. Conversely, the false positive lobules site contains more APFs than the true negative lobule site, particularly in stromal tissue located in between lobules. This indicates that the stromal tissue that lobules or other features are embedded within may lead to incorrect classification as a false positive site. In future work, additional
parameters are needed in order to classify lobules as benign and ILC as malignant with greater accuracy.

It is to be noted that while confocal microscopy provides high resolution high quality images, currently its cost, footprint, and maintenance requirements limit the ability to translate this imaging platform to routine usage in patient care. Specifically, the need for beam scanning in confocal microscopy limits the volume of tissue that can be surveyed in a given amount of time. However, structured illumination microscopy (SIM) has the added advantage of full-field illumination and non-descanned detection, which lowers the complexity and cost and increases the speed with which microscopy of large tissue areas can be performed. Thus, in future work the algorithms described here could be combined with SIM imaging of breast tissue to enable automated diagnosis of large tissue areas.

In conclusion, quantitative APF and ductal parameters were measured in confocal fluorescence images of fresh breast tissue and used to develop a classification algorithm that distinguishes between 259 benign and malignant sites with an accuracy of 88%. The APF and ductal parameters described in this study could be used to develop criteria to automate breast lesion diagnosis for immediate evaluation of fresh tissue at the point of care obviating the need for extensive tissue preparation. Quantitative diagnostic criteria developed on fluorescence confocal images in section 5.2.3.1.1 could
be applied to SIM images of breast tissue in the future in order to enable automated assessment of breast tumor margins.

5.2.3.2 Combining SIM with Optical Spectral Imaging for Wide-Field Assessment of Breast Tumor Margins

While it is feasible to use the SIM system to quickly survey the entire tumor resection cavity in a mouse, it is not feasible to image an entire human breast margin in a clinically relevant time window. Thus we foresee combining SIM imaging with a portable optical spectral imaging (OSI) system for wide-field imaging of tumor margins.

In particular our group has developed an OSI system for wide-field imaging of tumor margins in patients undergoing breast-conserving surgery (BCS). The system has a coverage area of 4x4 cm and 1 mm resolution. Initial results of OSI imaging indicate that the variables β-carotene and the reduced scattering coefficient <μs'> reliably distinguish adipose and malignant tissue, but not dense glandular and malignant tissue. Thus, SIM micro-anatomical imaging capabilities could be integrated into the current OSI system such that hot spots identified by wide field imaging with the OSI could be examined at higher resolution with micro-anatomical imaging. The integration of the two systems would enable wide-field imaging with OSI and improved specificity with SIM imaging.

Future work could focus on system integration. Specifically, this will require (1) a common optical path for both systems, (2) co-registration of the images from the two systems and (3) the integration of the image processing algorithms for the two systems.
into one unit. A series of *in vitro* phantom investigations could establish the performance metrics of the integrated device, specifically, coverage area, resolution, speed, contrast and reproducibility. Additionally, phantom investigations could be used to demonstrate software capability that allows for manipulation of key variables in real time.

Ultimately, the integrated device could be taken to the clinic to demonstrate co-registration of OSI and SIM imaging and feature identification from the images that reflect margin positivity.
References


94. 


159. Abeytunge, S., et al., Confocal microscopy with strip mosaicing for rapid imaging over large areas of excised tissue. Journal of Biomedical Optics, 2013. 18(6).


Biography

Jenna Mueller was born August 1, 1986 in Austin, Texas. She received her B.S. degree in Bioengineering from Rice University in May 2009 with a minor in Global Health Technologies. She joined Dr. Nimmi Ramanujam’s tissue optical spectroscopy lab at Duke University in August 2009 and completed her M.S. degree in Biomedical Engineering in September 2013 and her Ph.D. in Biomedical Engineering in September 2015.

Publications (*denotes that authors contributed equally)


Awards and Fellowships

- Bass Teaching Fellow, Bass Instructional Fellows Program, Duke University (2014)
- SPIE Optics and Photonics Education Scholarship, SPIE – The International Society for Optics and Photonics (2013)
- McCChesney Fellowship in Biomedical Engineering, Department of Biomedical Engineering, Duke University (2010 – 2011)
- National Science Foundation Graduate Research Fellowship Program, Honorable Mention, Rice University (2009)
- Best Global Health Technologies Engineering Design Project, Department of Bioengineering, Rice University (2009)
- Clinton Global Initiative University Outstanding Commitment Award, Awarded by former President Bill Clinton for diagnostic Lab-in-a-Backpack research, Tulane University (2008)