IMPROVING ULTRASOUND MICROVASCULAR IMAGING WITH SUPERHARMONIC IMAGING AND MACHINE LEARNING

Thomas M. Kierski

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biomedical Engineering.

Chapel Hill
2022

Approved by:

Devin Hubbard

Paul A. Dayton

Gianmarco F. Pinton

Andrea Giovannucci

Yueh Z. Lee

ABSTRACT

Thomas M. Kierski: Improving Ultrasound Microvascular Imaging with Superharmonic Imaging and Machine Learning
(Under the direction of Paul A. Dayton)

Biomedical ultrasound imaging devices are safe, portable, inexpensive, and produce high-resolution images of soft tissues in real time, making ultrasound the ideal modality for a variety of applications. While ultrasound is useful for noninvasively locating suspicious lesions in organs such as the breast and prostate, a tissue biopsy is required to make an accurate diagnosis in most cases. Most biopsies are benign, and the procedures are typically invasive and uncomfortable.

A tumor is composed of abnormal cells that divide uncontrollably. This rapid growth is accompanied by a considerable increase in angiogenesis, or the formation of new blood vessels. The normal balance of pro- and anti-angiogenic factors is disrupted, resulting in a dense and disorderly network of vessels. The effect is so pronounced that this phenomenon has been described as one of the "hallmarks of cancer", and many studies have suggested that increased angiogenesis precedes the appearance of solid tumors. Therefore, this biomarker is an attractive target for both detecting and treating cancer.

Recently, two technologies have been developed for imaging angiogenesis with ultrasound: acoustic angiography and ultrasound localization microscopy. Preclinical studies have demonstrated that these imaging methods can detect small tumors and monitor response to treatment by quantifying vascular features. However, a number of roadblocks remain for clinical translation of these methods. Specifically, they require an excellent signal to noise ratio, and extracting vascular features from images is labor intensive.

In this dissertation, we explore four new ideas with the aim of improving the existing methods for ultrasound microvascular imaging. First, we explore superharmonic imaging for ultrasound

localization microscopy, and show that this approach is robust to physiological motion and improves signal quality for small blood vessels. Then, we adapt this methodology to achieve super-resolution acoustic molecular imaging *in vivo* for the first time, paving the way for a new mode of quantitative cancer imaging. Afterwards, we apply deep learning to improve the detection of contrast agents for ultrasound localization microscopy, improving resolution in the presence of noise and image artifacts. Finally, we train convolutional neural networks to accurately detect tumors in acoustic angiography images in real time.

To my wife Kathleen, I wouldn't have made it this far without you.
Thank you for your love and support.

TABLE OF CONTENTS

LIST OF FIGURES xi

LIST OF TABLES xv

CHAPTER 1: INTRODUCTION TO ULTRASOUND MICROVASCULAR IMAGING 1

Cancer screening 1

Angiogenesis 2

Biomedical ultrasound imaging 3

Contrast-enhanced ultrasound 4

Microvascular ultrasound imaging 5

Acoustic angiography 6

Ultrasound localization microscopy 8

Quantifying vascular images 8

Dissertation scope 10

REFERENCES 10

CHAPTER 2: SUPERHARMONIC ULTRASOUND LOCALIZATION MICROSCOPY 15

Introduction 15

Materials and Methods 18

Experimental setup 18

ULM imaging scheme 20

Tube imaging in vitro 21

Flow study in vitro 21

Kidney imaging in vivo 22

Results 24

Tube imaging in vivo 24

Flow study in vitro 25

Kidney imaging in vivo 26

Discussion 26

Conclusion 34

REFERENCES 35

CHAPTER 3: ACOUSTIC MOLECULAR IMAGING BEYOND THE DIFFRACTION LIMIT 40

Introduction 40

Materials and Methods 42

Contrast Agent Preparation 42

Dual-frequency system description 43

Baseline resolution measurement for the hybrid dual-frequency array 44

Data collection for in vitro binding experiment 44

Animal care and in vivo data collection 45

Localization microscopy processing 47

Optical - Ultrasound Calibration 48

Vessel centerline extraction and analysis 50

Measuring the resolution of ULM images with Fourier Ring Correlation 51

Estimating the degree of vessel reconstruction 52

Results 52

In vitro studies 52

In vivo imaging 55

Discussion 56

Conclusion 60

REFERENCES 61

CHAPTER 4: DEEP LEARNING METHODS FOR THE DETECTION AND LOCALIZATION OF MICROBUBBLE CONTRAST AGENTS IN HIGH-CLUTTER ENVIRONMENTS 66

Introduction and Background 66

Methods 69

Simulation parameters 69

Generating reverberation clutter 71

Microbubble contratst agent simulations 72

Synthetic channel noise 73

Creating the synthetic training data 73

Model implementation and training 74

Model evaluation 76

Statistical analysis 78

Results 78

Discussion 79

Conclusion 82

REFERENCES 82

CHAPTER 5: DETECTING CANCER IN ULTRASOUND MICROVASCULAR IMAGES USING DEEP CONVOLUTIONAL NEURAL NETWORKS 86

Introduction 86

Materials and Methods 89

Contrast agent preparation 89

Animal Care and Tumor Model 90

In vivo imaging 90

Image preprocessing and augmentation 91

Model training and evaluation 94

Results 96

Discussion 97

Conclusion 100

REFERENCES 100

CHAPTER 6: CONCLUDING REMARKS 105

LIST OF FIGURES

Figure 1.1 - Tumor-associated angiogenesis results in a dense and chaotic vessel network 2

Figure 1.2 - Examples of B-mode and contrast-enhanced ultrasound imaging in a rodent kidney. 3

Figure 1.3 - Optical image of microbubble contrast agents 4

Figure 1.4 - Maximum intensity projections from three-dimensional B-mode and acoustic angiography images of a rodent abdomen 6

Figure 1.5 - The effects of tumors on the local vasculature are easily visible in acoustic angiography images 7

Figure 1.6 - Visualization of the Rayleigh criterion for $F\# = 1$ 9

Figure 2.1 - Dual-frequency transducer for plane wave imaging 19

Figure 2.2 - Data collection and processing for super harmonic ULM 20

Figure 2.3 - A comparison of SHI-ULM and AA in vitro 22

Figure 2.4 - Velocity maps of crossed 46 μm tubes in a water bath 24

Figure 2.5 - Maximum intensity projections of SVD-filtered and superharmonic images in vitro 25

Figure 2.6 - SNR vs. flow rate for dual-frequency and SVD-filtered images 27

Figure 2.7 - Example of super harmonic ultrasound localization microscopy applied to a rodent kidney with motion correction 28

Figure 2.8 - Selected vessels from rodent kidney 3-D dataset 30

Figure 2.9 - A comparison of ULM with and without motion correction 31

Figure 2.10 - Mapping of blood velocity in a rodent kidney 33

Figure 2.11 - 3-D ultrasound localization microscopy in a rodent kidney 34

Figure 3.1 - A flowchart outlining the steps for molecular ultrasound localization microscopy 45

Figure 3.2 - Optical comparison of control and biotinylated contrast agents in a microvessel phantom coated with avidin 46

Figure 3.3 - Example super-resolution molecular images from control and targeted trials in a microflow phantom 47

Figure 3.4 - Measuring the correlation between optical and dual-frequency ultrasound bubble counts 49

Figure 3.5 - Resolution measurements of dual-frequency transducer in superharmonic mode 50

Figure 3.6 - Super-resolution ultrasound molecular imaging *in vivo* 53

Figure 3.7 - Fourier ring correlation plot for ULM image of vasculature 54

Figure 3.8 - Tube-pairing algorithm using in vivo data. 56

Figure 3.9 - Vessel segmentation from tumor images allows for quantification of tortuosity metrics 57

Figure 3.10 - Estimated degrees of vessel reconstruction (DOR) for each tumor 59

Figure 4.1 - Sources of image degradation in ultrasound imaging in heterogeneous media 69

Figure 4.2 - Representative abdominal and skull speed of sound maps for Fullwave simulations 70

Figure 4.3 - Synthetic data for training localization models 73

Figure 4.4 - Schematic of the deep unfolded architecture used for ultrasound localization microscopy 74

Figure 4.5 - Matching algorithm for localization data 75

Figure 4.6 - Predictions on test data from the deep unfolded network 76

Figure 4.7 - Precision and recall vs. CTR for both localization methods on abdomen test dataset 77

Figure 4.8 - Distributions of the magnitudes of errors in the axial and lateral dimensions for both methods. 79

Figure 4.9 - Distributions of the axial and lateral localization errors from the P4-1 test dataset 80

Figure 4.10 - ULM images of an *in vitro* microflow phantom using the deep unfolded network and standard localization algorithm 80

Figure 5.1 - Maximum intensity projections of acoustic angiography images in a rodent model 89

Figure 5.2 - An overview of the data augmentation pipeline 91

Figure 5.3 - 2-D maximum intensity projections of the entire acoustic angiography dataset 92

Figure 5.4 - Best models for each architecture after the hyperparameter search 93

Figure 5.5 - A comparison of the 2-D and 3-D implementations of EfficientNet-B0 94

Figure 5.6 - Examining the effects of learning rate and batch size on EfficientNet-B0 performance 98

Figure 5.7 - GradCAM saliency maps for correctly classified tumor images 100

LIST OF TABLES

Table 3.1 - Summary of the ultrasound contrast agents used in the molecular imaging study 43

Table 4.1 - Acoustic properties of tissues for Fullwave simulations 71

Table 5.1 - Model performance summary 95

Table 5.2 - Hyperparameter search space 96

Table 5.3 - Data augmentation parameters 99

CHAPTER 1: INTRODUCTION TO ULTRASOUND MICROVASCULAR IMAGING[1]

1.1 Cancer screening

As of 2021, cancer is the 2nd leading cause of death in the United States, killing approximately 1,600 individuals per day [1]. The same review highlights the substantial progress against cancer in recent decades, estimating that roughly 3.2 million deaths have been avoided since 1991 primarily because of reductions in smoking and improved treatment regimens. However, the article concludes with the assertion that "progress is slowing or halting for cancers amenable to early detection through screening, such as breast cancer [and] prostate cancer," suggesting that we are approaching the limits of the current screening techniques for these diseases. Since prostate and breast cancer are the most common cancer types in men and women, respectively, there is a strong motivation to develop the next generation of screening technologies and further suppress the cancer death rate.

Biomedical imaging plays an important role in screening as well as monitoring the response to treatment over time. The most common cancer imaging modalities are x-ray, computed tomography (CT), magnetic resonance (MR) imaging, ultrasound (US), positron emission tomography (PET), and single-photon emission computed tomography (SPECT) [2]. When a neoplasm is detected via imaging or some other form of screening, a biopsy is often required to determine an accurate diagnosis. Because the conventional screening methods, including the medical imaging modalities mentioned earlier, lack the specificity to rule out malignancy in many cases, a significant percentage of biopsies lead to a negative result. For example, a recent study found that more than half of the women who undergo a biopsy following a mammogram do not have malignant tumors [3]. The same pattern was noted for men with prostatic lesions [4]. These extra biopsies

[1]Portions of this chapter appear as an article in Applied Physics Letters. The original citation is as follows: T. M. Kierski and P. A. Dayton, “Perspectives on high resolution microvascular imaging with contrast ultrasound,” *Applied Physics Letters*, vol. 116, no. 21, p. 210 501, 202.

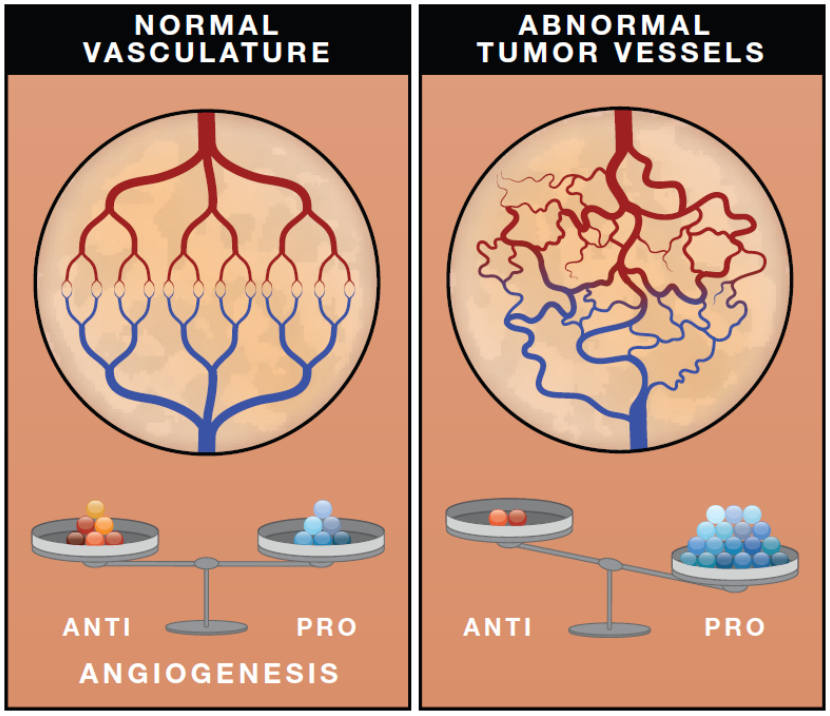

Figure 1.1: Tumor-associated angiogenesis results in a dense and chaotic vessel network. Figure reproduced with permission. Original citation: R. K. Jain and P. Carmeliet, "Snapshot: Tumor angiogenesis," *Cell*, vol. 149, no. 6, pp. 1408–1408, 2012.

impose a psychological burden on patients and a logistical and financial burden upon the health care system.

1.2 Angiogenesis

A well-known biomarker of cancer is angiogenesis and its associated molecular signaling [5], [6]. A tumor forms as the result of the prolonged and uncontrollable division of cancer cells, a process enabled in part by the disruption of normal angiogenic pathways (Fig. 1.1). Under normal circumstances, angiogenesis is confined to a relatively short time period, after which the process slows considerably. In contrast, a tumor must initiate a period of sustained local angiogenesis by shifting the balance of pro- and anti-angiogenic signaling to feed its growth [6]. The result is a network of densely-packed and tortuous blood vessels. Since this phenomenon is intrinsic to tumor development, it is regarded as one of the "hallmarks of cancer" [6]. Previous work has demonstrated that the recruitment of new blood vessels precedes the appearance of solid tumors [7], suggesting that angiogenesis is an attractive target for both the treatment and early

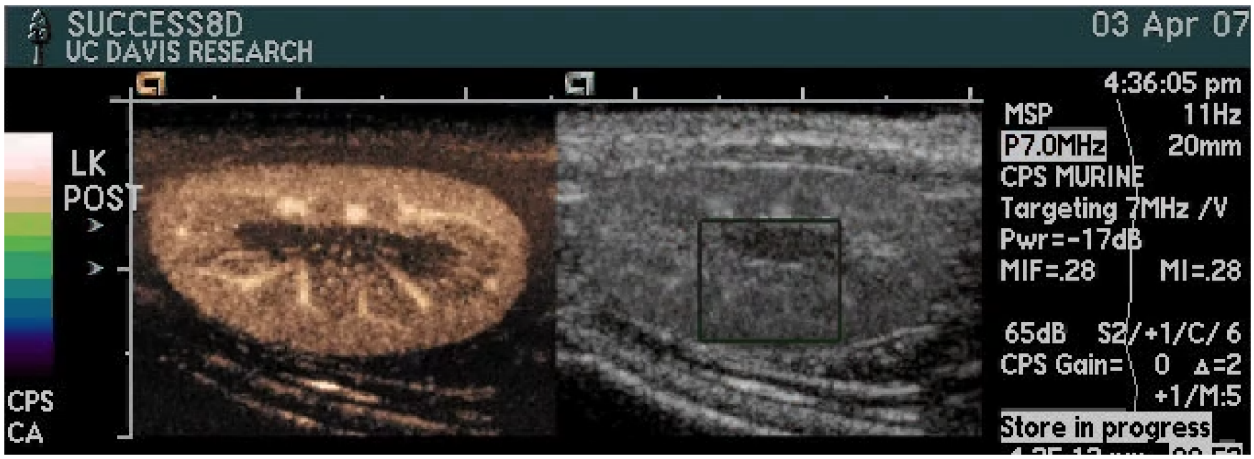

Figure 1.2: Examples of B-mode (*right*) and contrast-enhanced ultrasound imaging (*left*) in a rodent kidney. Image courtesy of Paul A. Dayton.

detection of cancer. Improving early detection is especially important, since it is associated with an improved five-year survival rate for virtually all cancer types [1].

1.3 Biomedical ultrasound imaging

Of the imaging modalities mentioned earlier, ultrasound is notable for its safety, accessibility, portability, and high spatial and temporal resolutions (<1 mm and >1,000 fps, respectively). Ultrasound images are created by transmitting a high-frequency acoustic wave into the target, recording the echoes, and beamforming the reflections into pictures of the underlying anatomy using the relationship between the time of flight and distance. These B-mode (short for "brightness-mode" images are often log-compressed and displayed in grayscale (Fig. 1.2). Ultrasonography has long been used by clinicians for investigating organs such as the kidney [8], heart [9], liver [10], spleen [11], prostate [12], thyroid [13], and others. In addition to interrogating soft tissues, it is also possible to noninvasively measure the mechanical properties of tissues, blood flow, and molecular signaling using elastography [13], Doppler imaging [14], and molecular imaging [15], [16], respectively. While B-mode and Doppler imaging are commonplace in many clinics, it is important to note their limitations. The spatial resolutions of these approaches are fundamentally diffraction-limited, which means that higher frequencies are necessary to image smaller targets. Many biological tissues have attenuative properties that grow exponentially as the frequency is increased [17]. Therefore, high-frequency imaging is generally restricted to shallow and superfi-

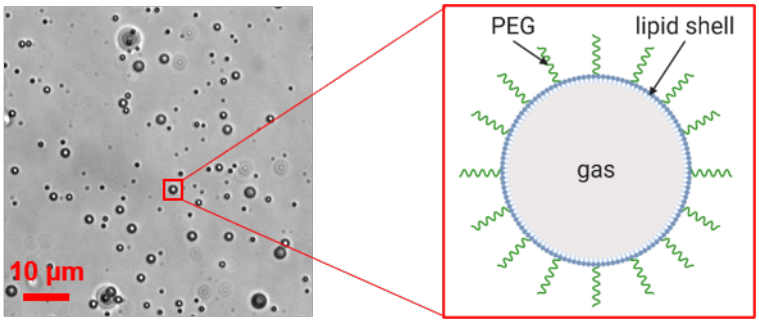

Figure 1.3: (*Left*) Optical image of microbubble contrast agents. (*Right*) Schematic of a single microbubble.

cial targets because of the poor signal to noise ratio (SNR) at depth, limiting the viability of using Doppler imaging for imaging the small blood vessels associated with cancer angiogenesis.

1.4 Contrast-enhanced ultrasound

It is possible to overcome low SNR in blood vessel imaging (*e.g.*, when the blood velocity is slow, as in a capillary) by introducing microbubble contrast agents (MCAs) [18], [19]. MCAs are small bubbles composed of a high molecular weight gas core and a lipid shell for stabilization (less common are shells made of proteins). Often, an additional polymer such as PEG is added to the shell to mitigate the immune response to the contrast agent. The typical diameter of an MCA ranges between 1 and 8 μm (Fig. 1.3), so it is able to traverse very small capillaries much like a red blood cell (roughly 8 μm in size). Microbubble contrast agents are excellent reflectors of sound because they have a much lower acoustic impedance than blood and most biological tissues. MCAs are administered intravenously and enhance the brightness of the tissues that they perfuse. This type of imaging is referred to as contrast-enhanced ultrasound (CEUS).

SNR can be further improved for CEUS by taking advantage of the nonlinear acoustic properties of microbubbles [18]. Since MCAs have a gas core, they are able to expand and contract in an acoustic field. When driven at or near their resonant frequency (a factor largely determined by the diameter of a given bubble), they respond with wide-bandwidth echoes with energy located at the harmonics of the transmitted frequency. Tissue also exhibits some degree of nonlinearity, but the effect is much less pronounced unless the acoustic pressure amplitude is quite large. Normally, a series of pulses with different phases and amplitudes are transmitted into the body, and

the echoes from each are recorded. These data are then scaled and summed in such a way that the linear content in the recordings is cancelled, while nonlinear frequencies remain. For example, a technique known as pulse inversion uses two transmissions that are 180° out of phase to suppress the majority of the signal from the tissue scatterers [20]. Another method, amplitude modulation, recovers the nonlinear component of the echoes by transmitting two pulses of different amplitudes (*e.g.*, half-amplitude and full-amplitude), scaling the radiofrequency data accordingly, and then taking the difference of the two signals prior to beamforming [21].

One application of CEUS is ultrasound molecular imaging (USMI), a technique that enables noninvasive mapping of the expression of molecular markers *in vivo*. This is accomplished by modifying the exterior of the standard microbubble architecture (Fig. 1.3) to include one or multiple targeting ligands [15], [16]. Acoustically, these modified bubbles are identical to the unmodified contrast agents described thus far and can be imaged with one of many different CEUS schemes. USMI is useful for both early detection of tumors [22] as well as quantifying response to treatment [23], [24] because it provides insight into the low-level molecular signaling that precedes macroscopic changes such as changes in tumor volume.

Another application of contrast-enhanced ultrasound is quantitative perfusion imaging, where the speed of wash-in is measured after administering a bolus of the contrast agent [25], [26] (Fig. 1.2). This is relevant for assessing the function of organs such as the kidney, or studying the effects of various drugs in preclinical studies. While the SNR of CEUS is increased compared to standard B-mode imaging, the resolution of CEUS is still limited by diffraction. As mentioned previously, resolution scales with frequency, so imaging smaller structures requires transducers with a higher center frequency. Therefore, most ultrasound systems cannot image blood vessels smaller than approximately 250 μm even with the addition of MCAs.

1.5 Microvascular ultrasound imaging

Recently, a variety of techniques for resolving very small blood vessels (≤ 100 μm) using ultrasound have been developed for applications such as cancer imaging. Some of these methods leverage the nonlinearity of the microbubble contrast agents while others make use of large

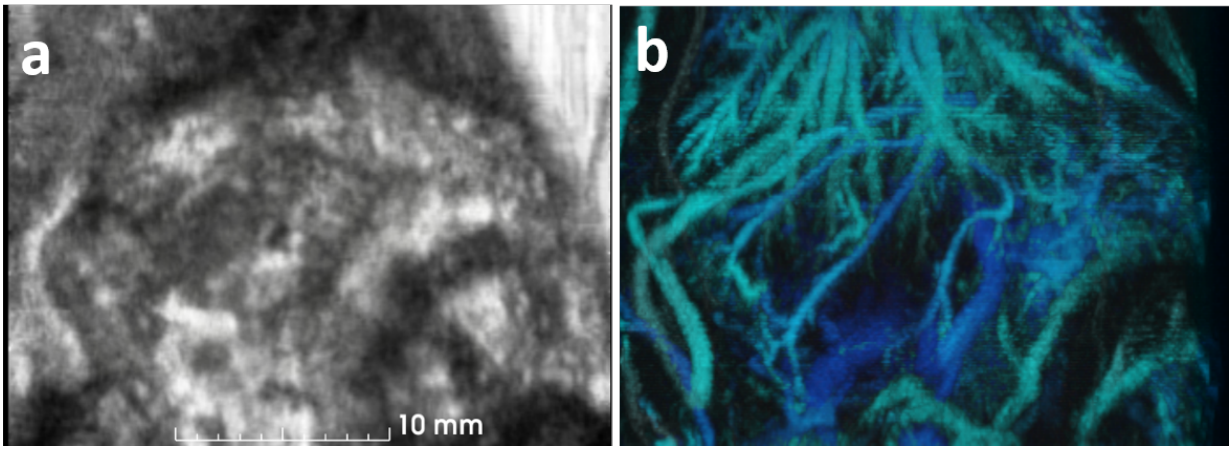

Figure 1.4: Maximum intensity projections from three-dimensional **(a)** B-mode and **(b)** acoustic angiography images of a rodent abdomen. The AA image in **(b)** is color-coded according to the distance between the blood vessel and the transducer (cyan: close, navy: far). Images courtesy of Jordan Joiner.

datasets and complex image processing pipelines.

1.5.1 Acoustic angiography

Acoustic angiography (AA) is a dual-frequency, superharmonic CEUS technique for imaging blood vessels as small as 100 μm in real time (Fig. 1.4). Superharmonic imaging was first developed by Bouakaz and colleagues to improve the SNR of contrast imaging compared to methods such as pulse inversion and amplitude modulation [27]. Superharmonic imaging makes use of wide-bandwidth transducers to detect and isolate the the third and higher harmonics of the transmitted wave (*i.e.*, the superharmonics), which are generated at appreciable levels by microbubble oscillations and far less so by biological tissues. In [27], the authors design and fabricate a dual-frequency linear array transducer where elements with center frequencies of 800 kHz and 2.8 MHz are interleaved with one another. By transmitting with the low-frequency elements and recording echoes using the high-frequency elements, this device is much more sensitive to the higher harmonics of the transmitted pulse than a conventional transducer, which has a typical bandwidth of 70-80%. In their experiments, the authors increased the SNR by 40 dB compared to conventional CEUS imaging.

Later on, this approach was adapted by Gessner *et al.* to develop acoustic angiography, who fabricated a dual-frequency transducer from two confocal annular elements with center frequen-

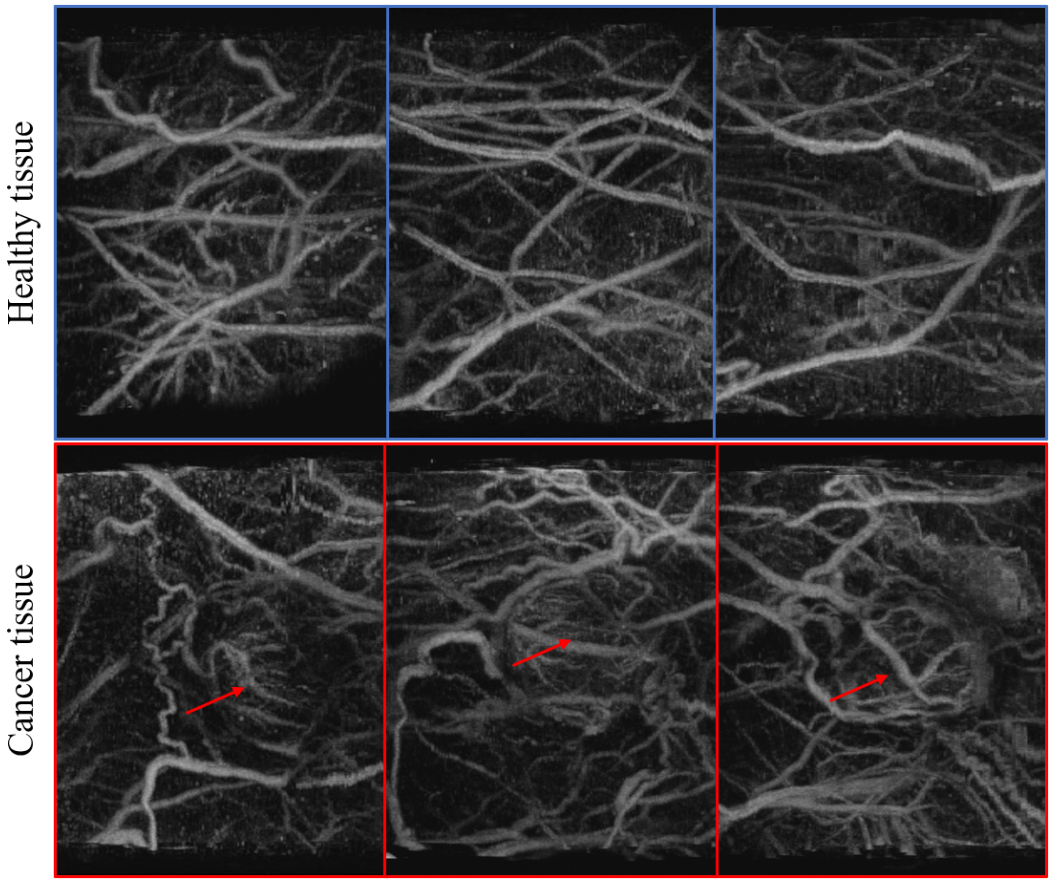

Figure 1.5: The effects of tumors on the local vasculature are easily visible in acoustic angiography images (unpublished data). Note the similarities to the artist's depiction shown in Fig. 1.1. Red arrows indicate the approximate locations of each tumor which are roughly 10 mm in diameter.

cies of 4 and 30 MHz. Assuming a speed of sound $c_0 = 1540$ m·s^{-1}, the wavelength λ of the high-frequency transducer was equal to 51 μm. By scanning this device in the elevational dimension, the authors were able to create three-dimensional maps of the vasculature, including blood vessels as small as 100 μm in diameter. In a follow-up study, the same group demonstrated that AA is sensitive to cancer-induced changes in vascular morphology by quantifying the tortuosity of blood vessels segmented from images of tumors and healthy tissue in rodents [28]. Representative images from healthy and tumor tissues are provided in Fig. 1.5. AA has also been used for superharmonic perfusion imaging [29], as well as monitoring response to treatment by quantifying vascular density after an anti-angiogenic treatment [30].

1.5.2 Ultrasound localization microscopy

Ultrasound localization microscopy (ULM) is a super-resolution method for noninvasively imaging blood vessels smaller than the diffraction limit. It is not a real time approach like AA, though it improves resolution by an order of magnitude. For normal pulse-echo imaging, such as the B-mode and CEUS methods described previously, the resolution of the beamformed image scales with the wavelength λ of the transducer. The result is that two targets nearer than approximately $\lambda/2$ will appear as a single object (see Fig. 1.6 for an example of this phenomenon). For example, consider two small blood vessels that are separated by a distance of $\lambda/4$ – these two channels would appear as a single apparent structure in the beamformed image.

Inspired by the localization microscopy techniques developed by the optical imaging community [31], [32], Couture *et al.* demonstrated that it is possible to recover the resolution lost to diffraction by estimating the precise location of single microbubble contrast agents by fitting parabolas to the echoes in the raw radiofrequency data [33]. Soon after, another group presented an approach for a five-fold improvement in resolution by localizing spatially seaparated microbubbles in CEUS images using the weighted center of mass of each PSF [34]. ULM was demonstrated for the first time *in vivo* in [35], where the authors resolved 19 μm vessels in the ear of a rat. Another group popularized the combination of high-frame rate B-mode imaging and spatiotemporal filtering for ULM shortly thereafter, using this approach to reconstruct the vasculature in a mouse brain [36].

1.6 Quantifying vascular images

There are a variety of techniques for extracting quantitative information from ultrasound vascular images. Perhaps the simplest is measuring the volumetric vascular density (VVD). Typically, VVD is computed within a region of interest by taking the ratio of the number of voxels with intensities greater than a predetermined threshold to the total number of voxels. This approach is useful for measuring tumor-associated angiogenesis because it is sensitive to all blood vessels, including those that are smaller than the diffraction limit. Shelton and colleagues demonstrated that the average VVD values measured from acoustic angiography images of spontaneous

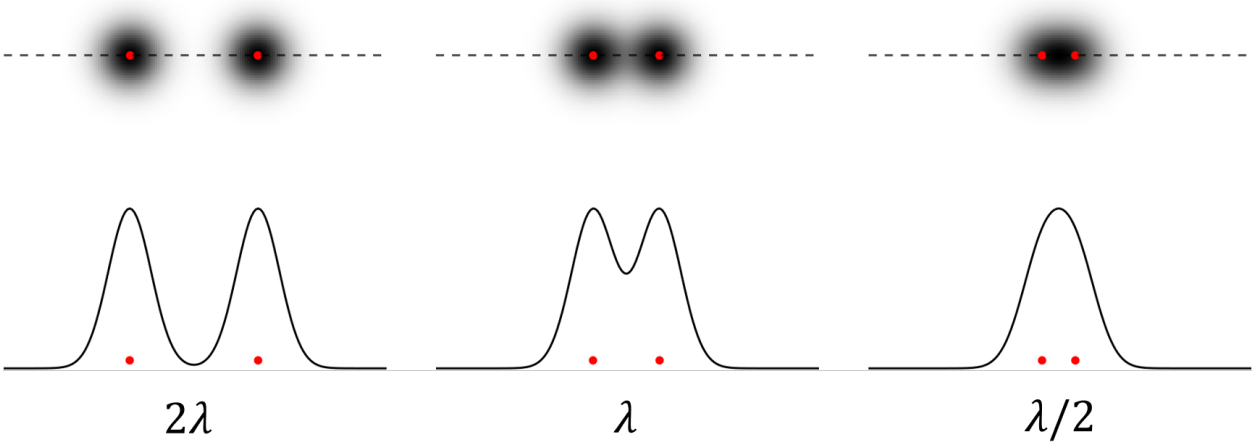

Figure 1.6: Visualization of the Rayleigh criterion for $F\# = 1$. Red dots indicate the locations of point targets, and the locations of the profiles on the bottom row are indicated by the dotted lines. As they move closer together, they become effectively indistinguishable, appearing as a single target.

tumors in a mouse model were more than two times larger than measurements from the control group [37]. Another recent preclinical study demonstrated that tracking VVD indicated response to an anti-angiogenic therapy one week prior to changes in tumor volume [38].

While VVD is typically computed at a global level to produce a single scalar value, it is also possible to analyze these images at a much more granular scale by measuring characteristics of individual blood vessels. Once the vessel centerlines have been segmented from the original image [39], their tortuosity can be measured by computing the distance metric (DM) and sum-of-angles metric (SOAM) [40]. Both the DM and SOAM are measurements of how much a line deviates from a straight path. The distance metric is equal to the total path length of a vessel divided by the distance between its two endpoints. The DM for a perfectly straight vessel is equal to 1. The SOAM is calculated by dividing the sum of all positive angles between consecutive triplets of centerline points by the total path length of the line. This metric is much more sensitive to small-amplitude, high-frequency oscillations than the distance metric.

As mentioned previously, [28] measured statistically significant differences between the tortuosity metrics of vessels segmented from tumor and control AA images in a preclinical study. Afterwards, the same result was achieved in 3-D ULM images, demonstrating that elevated tortuosity is also a characteristic of very small blood vessels [41]. More recently, researchers used

DBSCAN to cluster vessels from each AA image based on their SOAM and DM values. They found that the size of the largest cluster in each image was a strong predictor of cancer, with a final AUC of 0.95 [42]. One limitation to this approach is the requirement for each vessel to be individually segmented from the raw images. A single image can contain hundreds of vessels and require more than an hour to process even with the help of semi-automated software [39]. This issue is one of the major roadblocks to the clinical adoption of these technologies.

1.7 Dissertation scope

At this point in time, acoustic angiography and ultrasound localization microscopy are exciting preclinical technologies with the potential to make an impact in areas such as cancer screening and monitoring. However, there are several challenges still facing the research community which have so far limited the viability of ultrasound microvascular imaging in clinical settings. Therefore, the objective of this dissertation is to develop and validate new methods to improve the performance and utility of AA and ULM imaging. Current approaches for isolating microbubble contrast agents for ultrasound localization microscopy are based on spatiotemporal filtering appraoches, which do not generalize well to data corrupted by physiological motion or very slow flow. Consequently, Chapter 2 examines the use of superharmonic imaging for ULM. In Chapter 3, a novel method for super-resolution ultrasound molecular imaging is presented and validated *in vitro* and *in vivo*, representing the first cases of acoustic molecular imaging beyond the diffraction limit. Chapter 4 explores the use of deep neural networks for improving the precision, recall, and accuracy of contrast agent localization, demonstrating substantial improvements compared to a conventional algorithm in images corrupted by noise. Finally, in Chapter 5, we explore the feasibility of training deep convolutional neural networks to rapidly detect tumors in acoustic angiography images.

REFERENCES

[1] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, "Cancer statistics, 2021," *CA: A cancer journal for clinicians*, vol. 71, no. 1, pp. 7–33, 2021.

[2] J. V. Frangioni, "New technologies for human cancer imaging," *Journal of clinical oncology*, vol. 26, no. 24, p. 4012, 2008.

[3] J. Seely and T. Alhassan, "Screening for breast cancer in 2018—what should we be doing today?" *Current Oncology*, vol. 25, no. s1, pp. 115–124, 2018.

[4] H. U. Ahmed, A. E.-S. Bosaily, L. C. Brown, R. Gabe, R. Kaplan, M. K. Parmar, Y. Collaco-Moraes, K. Ward, R. G. Hindley, A. Freeman, *et al.*, "Diagnostic accuracy of multi-parametric mri and trus biopsy in prostate cancer (promis): A paired validating confirmatory study," *The Lancet*, vol. 389, no. 10071, pp. 815–822, 2017.

[5] R. K. Jain and P. Carmeliet, "Snapshot: Tumor angiogenesis," *Cell*, vol. 149, no. 6, pp. 1408–1408, 2012.

[6] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011.

[7] D. Hanahan and J. Folkman, "Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis," *Cell*, vol. 86, no. 3, pp. 353–364, 1996.

[8] M. L. Robbin, M. E. Lockhart, and R. G. Barr, "Renal imaging with ultrasound contrast: Current status," *Radiologic Clinics*, vol. 41, no. 5, pp. 963–978, 2003.

[9] S. Cheng, T. C. Dy, and S. B. Feinstein, "Contrast echocardiography: Review and future directions," *The American journal of cardiology*, vol. 81, no. 12, 41G–48G, 1998.

[10] S. R. Wilson and P. N. Burns, "Liver mass evaluation with ultrasound: The impact of microbubble contrast agents and pulse inversion imaging," in *Seminars in liver disease*, Copyright© 2001 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New …, vol. 21, 2001, pp. 147–160.

[11] P. Peddu, M. Shah, and P. Sidhu, "Splenic abnormalities: A comparative review of ultrasound, microbubble-enhanced ultrasound and computed tomography," *Clinical radiology*, vol. 59, no. 9, pp. 777–792, 2004.

[12] E. J. Halpern, "Contrast-enhanced ultrasound imaging of prostate cancer," *Reviews in urology*, vol. 8, no. Suppl 1, S29, 2006.

[13] F. Sebag, J. Vaillant-Lombard, J. Berbis, V. Griset, J. Henry, P. Petit, and C. Oliver, "Shear wave elastography: A new ultrasound imaging mode for the differential diagnosis of be-

nign and malignant thyroid nodules," *The Journal of Clinical Endocrinology & Metabolism*, vol. 95, no. 12, pp. 5281–5288, 2010.

[14] C. Demené, T. Deffieux, M. Pernot, B.-F. Osmanski, V. Biran, J.-L. Gennisson, L.-A. Sieu, A. Bergel, S. Franqui, J.-M. Correas, *et al.*, "Spatiotemporal clutter filtering of ultrafast ultrasound data highly increases doppler and fultrasound sensitivity," *IEEE transactions on medical imaging*, vol. 34, no. 11, pp. 2271–2285, 2015.

[15] R. Gessner and P. A. Dayton, "Advances in molecular imaging with ultrasound," *Molecular imaging*, vol. 9, no. 3, pp. 7290–2010, 2010.

[16] J. K. Tsuruta, N. Klauber-DeMore, J. Streeter, J. Samples, C. Patterson, R. J. Mumper, D. Ketelsen, and P. Dayton, "Ultrasound molecular imaging of secreted frizzled related protein-2 expression in murine angiosarcoma," *PLoS One*, vol. 9, no. 1, e86642, 2014.

[17] S. Goss, L. Frizzell, and F. Dunn, "Ultrasonic absorption and attenuation in mammalian tissues," *Ultrasound in medicine & biology*, vol. 5, no. 2, pp. 181–186, 1979.

[18] N. De Jong, A. Bouakaz, and P. Frinking, "Basic acoustic properties of microbubbles," *Echocardiography*, vol. 19, no. 3, pp. 229–240, 2002.

[19] K. Ferrara, R. Pollard, and M. Borden, "Ultrasound microbubble contrast agents: Fundamentals and application to gene and drug delivery," *Annu. Rev. Biomed. Eng.*, vol. 9, pp. 415–447, 2007.

[20] J.-J. Hwang and D. H. Simpson, "Two pulse technique for ultrasonic harmonic imaging," US Patent 5,951,478, 1999.

[21] C. Tremblay-Darveau, R. Williams, L. Milot, M. Bruce, and P. N. Burns, "Visualizing the tumor microvasculature with a nonlinear plane-wave doppler imaging scheme based on amplitude modulation," *IEEE Transactions on Medical Imaging*, vol. 35, pp. 699–709, 2 2016.

[22] S. V. Bachawal, K. C. Jensen, A. M. Lutz, S. S. Gambhir, F. Tranquart, L. Tian, and J. K. Willmann, "Earlier detection of breast cancer with ultrasound molecular imaging in a transgenic mouse model," *Cancer Research*, vol. 73, no. 6, pp. 1689–1698, 2013.

[23] J. Zhou, H. Wang, H. Zhang, A. M. Lutz, L. Tian, D. Hristov, and J. K. Willmann, "Vegfr2-Targeted Three-Dimensional Ultrasound Imaging Can Predict Responses to Antiangiogenic Therapy in Preclinical Models of Colon Cancer," *Cancer Research*, vol. 76, no. 14, pp. 4081–4089, 2016.

[24] M. A. Pysz, I. Guracar, L. Tian, and J. K. Willmann, "Fast microbubble dwell-time based ultrasonic molecular imaging approach for quantification and monitoring of angiogenesis in cancer," *Quantitative imaging in medicine and surgery*, vol. 2, no. 2, p. 68, 2012.

[25] S. Feingold, R. Gessner, I. M. Guracar, and P. A. Dayton, "Quantitative volumetric perfusion mapping of the microvasculature using contrast ultrasound," *Investigative radiology*, vol. 45, no. 10, 2010.

[26] P. Kogan, K. A. Johnson, S. Feingold, N. Garrett, I. Guracar, W. J. Arendshorst, and P. A. Dayton, "Validation of dynamic contrast-enhanced ultrasound in rodent kidneys as an absolute quantitative method for measuring blood perfusion," *Ultrasound in medicine & biology*, vol. 37, no. 6, pp. 900–908, 2011.

[27] A. Bouakaz, S. Frigstad, F. J. Ten Cate, and N. de Jong, "Super harmonic imaging: A new imaging technique for improved contrast detection," *Ultrasound in medicine & biology*, vol. 28, no. 1, pp. 59–68, 2002.

[28] R. C. Gessner, S. R. Aylward, and P. A. Dayton, "Mapping microvasculature with acoustic angiography yields quantifiable differences between healthy and tumor-bearing tissue volumes in a rodent model," *Radiology*, vol. 264, no. 3, pp. 733–740, 2012.

[29] B. D. Lindsey, S. E. Shelton, K. H. Martin, K. A. Ozgun, J. D. Rojas, F. S. Foster, and P. A. Dayton, "High resolution ultrasound superharmonic perfusion imaging: In vivo feasibility and quantification of dynamic contrast-enhanced acoustic angiography," *Annals of biomedical engineering*, vol. 45, no. 4, pp. 939–948, 2017.

[30] J. D. Rojas, V. Papadopoulou, T. J. Czernuszewicz, R. M. Rajamahendiran, A. Chytil, Y.-C. Chiang, D. C. Chong, V. L. Bautch, W. K. Rathmell, S. Aylward, *et al.*, "Ultrasound measurement of vascular density to evaluate response to anti-angiogenic therapy in renal cell carcinoma," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 3, pp. 873–880, 2018.

[31] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.

[32] S. T. Hess, T. P. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophysical Journal*, vol. 91, no. 11, pp. 4258–4272, 2006.

[33] O. Couture, B. Besson, G. Montaldo, M. Fink, and M. Tanter, "Microbubble ultrasound super-localization imaging (musli)," in *2011 IEEE International Ultrasonics Symposium*, IEEE, 2011, pp. 1285–1287.

[34] O. M. Viessmann, R. J. Eckersley, K. Christensen-Jeffries, M. X. Tang, and C. Dunsby, "Acoustic super-resolution with ultrasound and microbubbles," *Physics in Medicine and Biology*, vol. 58, no. 18, pp. 6447–6458, 2013.

[35] K. Christensen-Jeffries, R. J. Browning, M. X. Tang, C. Dunsby, and R. J. Eckersley, "In vivo acoustic super-resolution and super-resolved velocity mapping using microbubbles," *IEEE Transactions on Medical Imaging*, vol. 34, no. 2, pp. 433–440, 2015.

[36] C. Errico, J. Pierre, S. Pezet, Y. Desailly, Z. Lenkei, O. Couture, and M. Tanter, "Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging," *Nature*, vol. 527, no. 7579, pp. 499–502, 2015.

[37] S. E. Shelton, Y. Z. Lee, M. Lee, E. Cherin, F. S. Foster, S. R. Aylward, and P. A. Dayton, "Quantification of microvascular tortuosity during tumor evolution using acoustic angiography," *Ultrasound in medicine & biology*, vol. 41, no. 7, pp. 1896–1904, 2015.

[38] J. D. Rojas, V. Papadopoulou, T. J. Czernuszewicz, R. M. Rajamahendiran, A. Chytil, Y.-C. Chiang, D. C. Chong, V. L. Bautch, W. K. Rathmell, S. Aylward, *et al.*, "Ultrasound measurement of vascular density to evaluate response to anti-angiogenic therapy in renal cell carcinoma," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 3, pp. 873–880, 2018.

[39] S. R. Aylward and E. Bullitt, "Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction," *IEEE transactions on medical imaging*, vol. 21, no. 2, pp. 61–75, 2002.

[40] E. Bullitt, G. Gerig, S. M. Pizer, W. Lin, and S. R. Aylward, "Measuring tortuosity of the intracerebral vasculature from mra images," *IEEE transactions on medical imaging*, vol. 22, no. 9, pp. 1163–1171, 2003.

[41] F. Lin, S. E. Shelton, D. Espindola, J. D. Rojas, G. Pinton, and P. A. Dayton, "3-d ultrasound localization microscopy for identifying microvascular morphology features of tumor angiogenesis at a resolution beyond the diffraction limit of conventional ultrasound," *Theranostics*, vol. 7, no. 1, p. 196, 2017.

[42] S. E. Shelton, J. Stone, F. Gao, D. Zeng, and P. A. Dayton, "Microvascular ultrasonic imaging of angiogenesis identifies tumors in a murine spontaneous breast cancer model," *International Journal of Biomedical Imaging*, vol. 2020, 2020.

CHAPTER 2: SUPERHARMONIC ULTRASOUND LOCALIZATION MICROSCOPY[1]

2.1 Introduction

Recently, super-resolution imaging with ultrasound localization microscopy (ULM) has attracted attention because it resolves blood vessels on the order of a few μm in diameter at centimeters in depth in vivo [1], [2]. A model relating the spatial localization error of microbubble contrast agents (MCAs) to arrival time estimation error predicts that for certain in vivo scenarios, such as human breast imaging, ULM will achieve resolutions on the order of 1 μm [3]. It has long been known that abnormal angiogenesis and vascular morphology are biomarkers for different diseases, including diabetes, inflammatory conditions, and cancer [4], [5]. Recently, imaging abnormal angiogenesis with ultrasound microvascular imaging techniques has been proposed as a method of identifying malignancies [6]–[8]. In this context, ULM has shown diagnostic potential by measuring tortuosity of blood vessel structure in subcutaneous tumors in a rodent model [9]. In addition to providing morphological data, ULM is also able to provide accurate quantification of blood flow velocity, which can be combined with other metrics, such as vessel distances, to create a rich characterization of the imaging target [10].

Many different approaches to ULM are present in the literature, although the method for generating a super-resolved image with ultrasound can be described by three general components [11]. First, microbubble contrast agents are administered intravenously, and a series of frames is acquired (normally hundreds to hundreds of thousands). While some groups have reported success with clinical scanners constrained to lower frame rates (<100 Hz) [2], [10], generally a high frame rate on the order of 1-10 kHz is used to perform accurate velocimetry after target lo-

[1]Portions of this chapter previously appeared as an article in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control. The original citation is as follows: T. M. Kierski et al., "Superharmonic ultrasound for motion-independent localization microscopy: Applications to microvascular imaging from low to high flow rates," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 5, pp. 957–967, May 2020.

calization. Second, the dataset is processed to separate MCA and tissue signals, which overlap in conventional ultrasound imaging. Popular approaches to this step will be discussed subsequently in this section. Finally, MCAs are localized in each frame with sub-wavelength accuracy and the positions are accumulated on a high-resolution grid. The bubble positions are typically tracked between frames to also create high-resolution blood velocity maps [2].

As mentioned previously, a crucial step to the process of generating a ULM image is the separation of MCA signals from background tissue signal. The most popular method of suppressing tissue speckle prior to localization is a filter based on singular value decomposition (SVD). The SVD filter isolates MCAs by taking advantage of the different spatiotemporal coherences of tissue speckle and contrast agents [12], [13]. Though the microbubbles and tissue may be moving with the same velocity magnitude, the fact that the microbubbles are localized in space implies that they have far smaller spatial coherence lengths in the beamformed images. When tissue is relatively static within an ultrafast ensemble, its features tend to be represented in the first singular vectors, where the right singular vectors (also called temporal singular vectors in this context) have most of their energy near 0 Hz [12], [14]. Blood, on the other hand, flows at a range of velocities, and its scatterers decorrelate at varying rates over the course of an acquisition. Crucially, these scatterers decorrelate in spatially localized regions of the image. The energy from these scatterers thus tends to occupy a subspace of higher singular vectors in which the spatiotemporal vectors are higher frequency than those corresponding to tissue. As long as there is sufficient separation between the vector subspaces occupied by tissue and blood flow, a dataset can be filtered to remove tissue.

However, in the slow-flow regime, the tissue and blood singular vector subspaces overlap significantly, especially since the bubble signal can be orders of magnitude smaller than the tissue signal. Due to this low contrast disparity, they can be impossible to tease apart. In [13], it was demonstrated that SVD filtering of an ultrafast ensemble of B-mode images using a commercially-available contrast agent in a flow phantom resulted in contrast-to-tissue ratios (CTRs) of 11 and 25 dB for flow rates of 2 and 20 $mm \cdot s^{-1}$, respectively, when imaging at 3,000

frames per second. Furthermore, [14] has documented the difficulty in determining an appropriate singular vector threshold for the SVD filter in vivo. In their study, the most successful of 13 different threshold estimators was able to achieve a CTR within 10% of the maximum CTR for only 74% of in vivo datasets. The results are even worse for their manual threshold selection, where a relative CTR difference of 10% or less was achieved for only 13% of datasets. Although SVD-based processing has produced many impressive ULM images ([1], [9]), the results of [14] suggest that the performance of such a filter may suffer when applied to the smallest of capillaries where peak blood velocity ranges between 0.2 and 1.7 $mm \cdot s^{-1}$ [15], [16]. Furthermore, SVD may not be appropriate for new applications, such as super-resolution molecular imaging, in which bubbles would exhibit no motion relative to the tissue.

Another approach to contrast enhancement that has been used for ULM is nonlinear imaging [2], [17]. While spatiotemporal processing methods rely on the motion of contrast agents relative to tissue in slow time, nonlinear imaging sequences rely on the fact that microbubble contrast agents generate significantly more harmonic energy than tissue under most circumstances. For example, one study reports that imaging at 1.7 MHz center frequency results in a 2nd harmonic that is 24 dB down compared to the fundamental for tissue and around 9 dB down for microbubbles [18]. To date, methods such as pulse inversion [19], amplitude modulation [20], [21], and more sophisticated combinations of phase and amplitude modulation [22] have achieved CTR on the order of 50 dB with commercially-available contrast agents.

Super harmonic imaging (SHI) is a method of contrast-enhanced ultrasound that reconstructs images using the third and higher order harmonics of the fundamental frequency of the transmit waveform [18]. The advantage of SHI is an improvement in CTR compared to fundamental and second harmonic imaging (40 dB increase reported in [18]), along with an increased resolution from the higher frequencies and reduced sidelobes [23]. These improvements come at the cost of decreased imaging performance at depth because of the rapid attenuation of high-frequency waves in tissue. SHI is used extensively for vascular imaging with an approach called acoustic angiography (AA) [24]. In the previously mentioned study, an AA image is generated by receiv-

ing from the 3rd to approximately the 10th harmonic of the MCA frequency response by using an ultra-wideband dual-frequency transducer (transmit center frequency: 4 MHz, receive center frequency: 30 MHz, both roughly 100% relative bandwidth), producing images of blood vessels with a resolution of approximately 150 μm. It has been shown that AA is able to resolve microvasculature in vivo with high CTR in both rodents and humans [24], [25], although this technique remains fundamentally diffraction limited.

Thus, in this paper, we report a combination of SHI and super-resolution processing, without the SVD filter. In this manner, we eliminate the need to detect microbubbles through their spatiotemporal coherence, and we also overcome the diffraction limitation of SHI. In this paper, super harmonic ULM is performed using a novel dual-frequency array system with transmission at 1.7 MHz and a receive bandwidth centered at 20 MHz [26]. In vitro, we image a 46-μm inner diameter tube and compare the sensitivities of SHI- and SVD-based approaches to ULM with respect to flow velocity. In vivo, we target the rodent kidney and discuss strategies for motion correction in a dual-frequency imaging scheme. For both in vitro and in vivo studies, we compare the resolution of the images to acoustic angiography maximum intensity projections. Finally, we discuss the strengths and weaknesses of SHI for ULM along with areas of future work.

2.2 Materials and Methods

2.2.1 Experimental setup

Contrast agent preparation

Microbubbles were prepared in-house according to [27]. Briefly, a 1 mM lipid solution comprised of 90 mole % 1,2-distearoyl-sn-glycero-3-phosphocholine and 10 mole % 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000] was formulated in phosphate buffered saline containing 15% (v/v) propylene glycol and 5% (v/v) glycerol. Aseptic lipid solution was packaged into 3 mL glass vials, and the air headspace was exchanged with decafluorobutane (C4F10) prior to creating the microbubble emulsion by shaking in a VialMix (Lantheus Medical Imaging, N. Billerica, MA). Concentration and size distribution of the microbubble contrast agent were measured using an Accusizer 780 AD (Entegris, Billerica, MA);

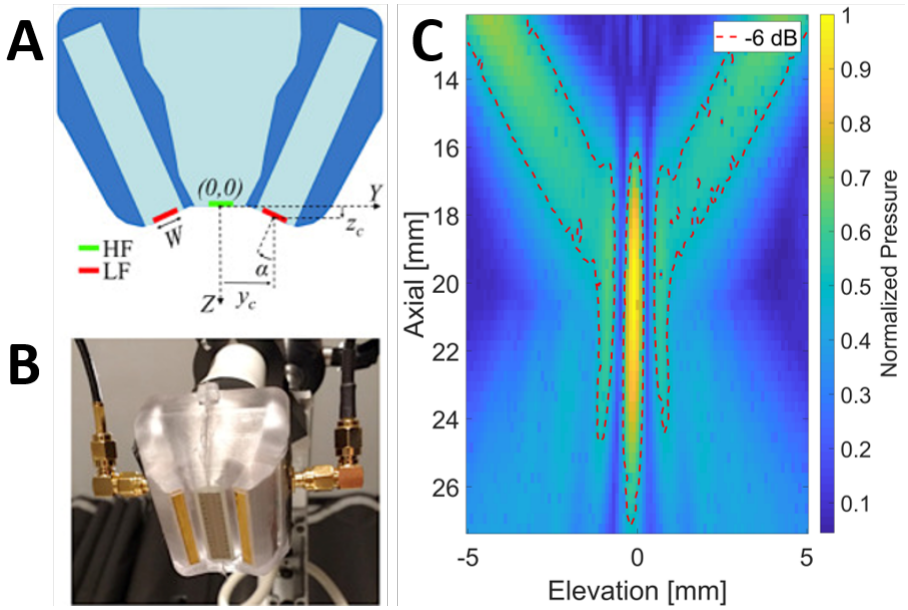

Figure 2.1: Dual-frequency transducer for plane wave imaging. (A) Schematic of the elevation cross-section of the dual-frequency assembly with LF transducers in red and HF array in green. W = 2.9 mm, α = 27°, (y_c, z_c) = (8.45 mm, 0.73 mm). (B) Photograph of the dual-frequency probe used in experiments, illustrating the LF transducers running parallel to the 21-MHz array front face. (C) Hydrophone measurement of LF beam pattern in the elevational-axial plane. The -6 dB contour of the beam is marked with a dashed line. The axial dimension is measured relative to the face of the HF array. Panels (A) and (B) reproduced from Cherin et al with permission.

typical concentration was 3e10 MCA/mL with an average diameter of 0.97 $\pm$ 0.51 μm (mode = 0.6 μm, median = 0.9 μm).

Dual-frequency transducer

A custom dual-frequency (DF) probe described in [26] was used for all imaging in this study (see Fig. 2.1A-B). Briefly, it consists of a commercial 256-element linear array transducer (MS250, VisualSonics, Toronto, Canada) outfitted with two low-frequency (LF) elements (Fig. 2.1). The LF transmit beam has a depth of field of 11 mm with peak pressure at 20 mm in the axial dimension. The high-frequency (HF) array has a center frequency measured at 18 MHz and relative bandwidth of 70%, while the LF elements have a center frequency of 1.7 MHz and relative bandwidth of 78%. This transducer can be operated in DF mode by transmitting with the LF elements and receiving with the HF array and in conventional mode by transmitting and receiving with the HF array. When operating in DF mode, the transmit pulse is a single-cycle, cosine-

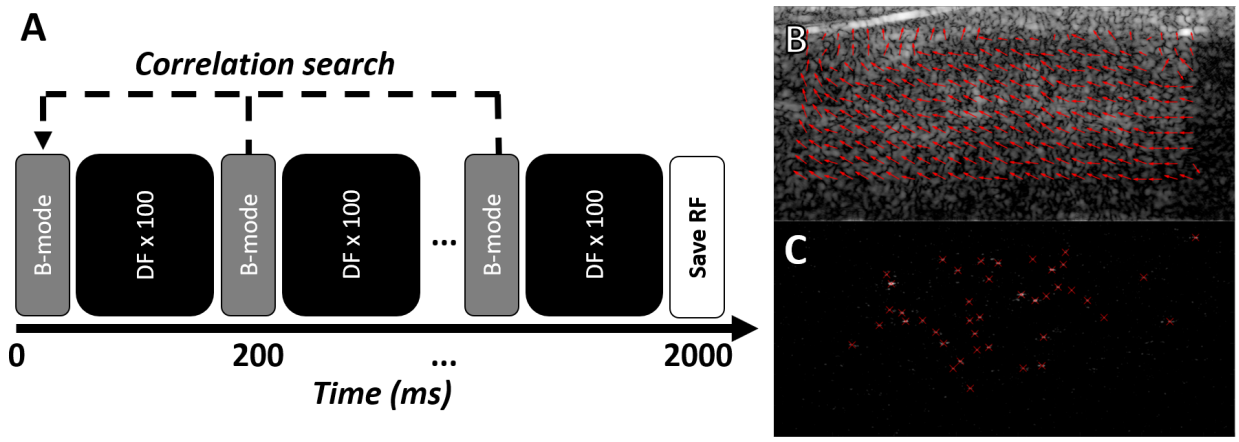

Figure 2.2: An overview of data collection and processing for super harmonic ULM. (A) The imaging sequence used for this study. Chunks of 100 dual-frequency frames collected at a frame rate of 500 Hz are separated by B-mode frames for motion tracking. Radio-frequency data is saved after 1,000 DF frames. (B) Speckle tracking is performed between a manually selected reference frame and each B-mode frame to estimate the non-rigid deformation of the kidney during imaging. (C) Dual-frequency images are processed using a threshold and peak detector to localize microbubbles. These positions are then corrected according to the displacements estimated from speckle tracking or thrown out if the parent B-mode patch is not well-correlated with the reference patch.

windowed sine wave with a center frequency of 1.7 MHz [28]. The LF elements are driven by an arbitrary waveform generator (AWG 2021, Tektronix, Beaverton, OR, USA) connected to a 50 dB radiofrequency power amplifier (240 L, ENI, Rochester, NY, USA). Receiving with the HF array is controlled by a Vantage 256 scanner (high frequency configuration, Verasonics, Kirkland, WA).

2.2.2 ULM imaging scheme

All ULM images in this article were generated using DF mode with a pulse repetition frequency (PRF) of 500 Hz at a mechanical index (MI) of 0.24 for a total of 25,000 frames. Radiofrequency (RF) data were beamformed offline on a 10-μm grid and thresholded to remove background noise (threshold empirically determined). Bubbles were localized using peak detection with an isotropic Gaussian aperture with an RMS width of 100 μm and tracked between frames using a nearest-neighbors approach with a maximum linking distance of 100 μm between frames. For comparison, a super harmonic maximum intensity projection (MIP) was generated from the stack of DF images used to create the ULM image.

2.2.3 Tube imaging in vitro

A resolution phantom was made using two microtubes made of fluorinated ethylene propylene, each with an inner diameter of 46 μm (measured optically with a calibrated microscope). The phantom was submerged in a water bath, and the tubes crossed in an 'X' configuration at a depth of approximately 20 mm. A dilution of microbubbles in saline with a concentration of 1E7 MCA/mL was prepared and infused through both tubes in opposite directions at 10 μL/min using an infusion pump (Harvard Apparatus, Holliston, MA). The tubes were imaged according to the protocol described in section II-B, and average tube profiles were measured within the same ROI for ULM and acoustic angiography images for comparison.

2.2.4 Flow study in vitro

A cellulose tube with an inner diameter of 200 μm was suspended in a water bath at a depth of 20 mm. A dilution of MCA in saline with concentration 1e7 MCA/mL was infused through the tube at volume flow rates ranging between 0.25 and 15.0 μL/min using an infusion pump (Harvard Apparatus, Holliston, MA). These flow rates correspond to mean displacements of 0.27 and 15.90 μm/frame, respectively. Before collecting data for each trial, the tube was flushed with air and water and reinfused with a newly prepared dilution of contrast agent. Infusion was allowed to proceed for a minimum of three minutes before imaging to ensure that the velocity of the contrast agent in the tube had reached steady state. For each trial, one thousand frames were acquired, and three trials for each modality per flow rate were performed. DF frames were collected according to section II-B, while B-mode frames were collected at a MI of 0.11 (center frequency = 15.6 MHz) and a PRF of 500 Hz.

Each batch of B-mode images was SVD filtered as follows: 1.) Arranged the beamformed RF data into the Casorati matrix whose columns are vectorized B-mode frames 2.) Performed a singular value decomposition on this matrix 3.) Zeroed all singular values for the first 15 singular vectors (empirically determined) 4.) Reconstructed the B-mode frames with the new set of singular values. Each set of 1,000 frames was then used to generate a maximum intensity projection, which was normalized and converted to decibels. A reference B-mode frame was used to draw

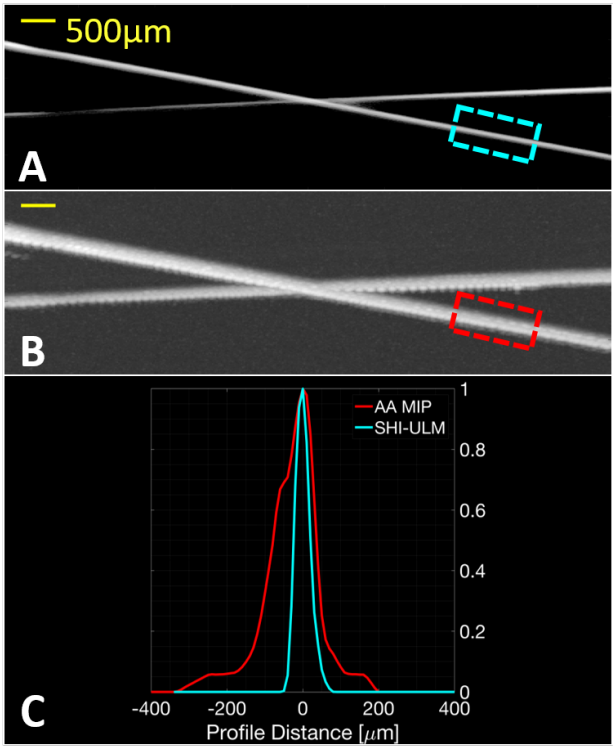

Figure 2.3: A comparison of SHI-ULM and AA using a pair of 46 μm tubes in a water bath. Yellow scale bars are 500 μm. (A) Super-resolution image generated from 25,000 frames. (B) Maximum intensity projection of the SHI frames used to generate the image in panel A. (C) Average profiles within the regions of interest from panels A and B. The full-width half-maximum values of the AA and ULM profiles are 113 and 44 μm, respectively.

a pair of ROIs corresponding to the tube and the background. SNR in this experiment is defined as the difference between the maximum value of the tube ROI and the average magnitude within the background ROI. This definition has been chosen to account for the sparse number of bubbles present in each MIP for the slower flow rates, because averaging within a tube ROI artificially lowers the SNR for each trial by including gaps between bubbles in the average calculation.

2.2.5 Kidney imaging in vivo

In vivo imaging was performed in healthy female Fischer 344 rats (Charles River Laboratories, Durham, NC) according to a protocol approved by the Institutional Animal Care and Use committee at the University of North Carolina at Chapel Hill. A polydisperse population of microbubble contrast agent (mean = 0.97 μm, standard deviation = 0.51 μm) was diluted to 1e9 MCA/mL and administered via a catheter placed in the tail vein at 25 μL/min using a syringe

pump (Harvard Apparatus, Holliston, MA). Infusion was allowed to proceed for 3 minutes prior to any imaging to allow the concentration of contrast in circulation to approach steady state.

DF images were collected and processed according to the parameters in section II-B. To estimate physiological motion, B-mode frames were interleaved between every 100 DF acquisitions (Fig. 2.2A). Two-dimensional speckle tracking was performed on adjacent B-mode frames according to [29] with a square 2-millimeter kernel (approximately 20 HF wavelengths in either dimension), $\pm$ 150-μm search window with 1-pixel step size, and 50-μm steps between adjacent kernels (Fig. 2.2B). The displacement grid for each time step was spatially interpolated to match the 10-μm pixel size of the original image.

To estimate the tissue displacement for a given DF image (Fig. 2.2C), linear interpolation is performed through the slow time dimension between consecutive displacement arrays. Then, for each DF image, detected bubble locations are adjusted based on the estimated deformation of the tissue at that time point. Linear interpolation was utilized for this task because it yielded acceptable results and was computationally inexpensive compared to similar techniques (*e.g.*, Hermite interpolation). Bubble localizations are also weighted in the final image according to the peak correlation coefficient associated with the bubble's parent patch during motion estimation. For example, if the correlation search is able to find a perfect match, the bubble's localization is given a value of 1, whereas a poor match might result in the bubble being weighted at 0.5. Bubbles below a correlation threshold of 0.3 are completely filtered from the analysis. The accuracy of the speckle tracking depends partially on how much the target decorrelates as a result of motion [30]. Therefore, contributions to the final ULM image were weighted by the correlation coefficient from the speckle tracking in order to minimize the effect of inaccurate displacement estimation on image quality. The correlation threshold of 0.3 for completely removing a localization was empirically determined. Between frames, microbubble centroids are linked using a nearest neighbor approach, and these line segments are drawn to create the final image. The diameters of selected vessels in ULM images were determined by taking the average of multiple full-width at half-maximum measurements along the axis of each vessel.

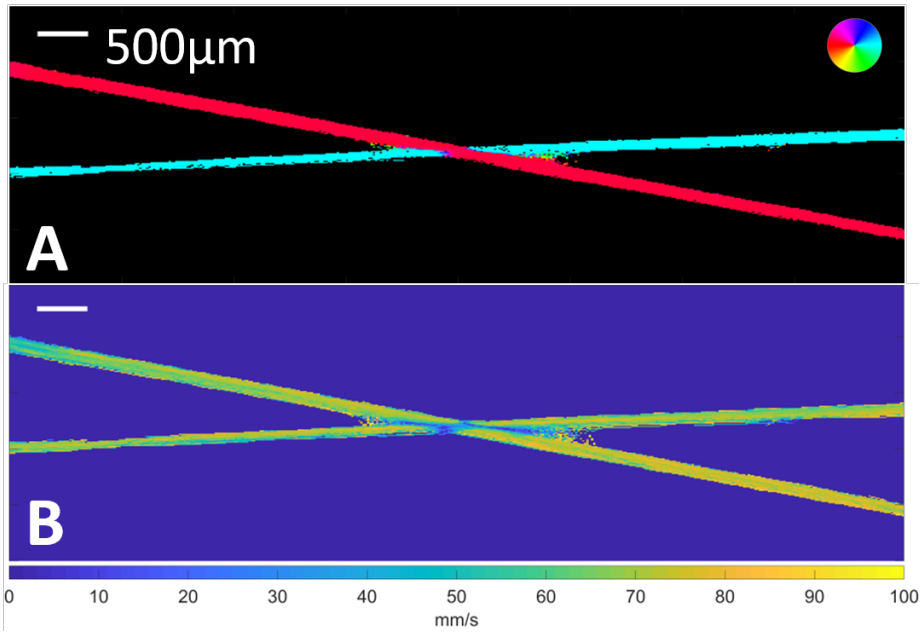

Figure 2.4: Velocity maps of crossed 46 μm tubes in a water bath. White scale bars are 500 μm. (A) Direction map, flow direction indicated by color wheel. (B) Map of the average speed for each pixel.

Three-dimensional imaging was accomplished by using a linear motion stage (XSlide, Velmex, Inc., NY, USA) controlled by a custom LabVIEW program (National Instruments, TX, USA) to mechanically sweep the ultrasound transducer in the elevational dimension. 25,000 DF frames were acquired at each position, and each position was spaced by 500 μm.

2.3 Results

2.3.1 Tube imaging in vivo

A ULM image was generated with 25,000 frames using a 1.7 MHz plane wave transmission and a center frequency on receive of 15.6 MHz (Fig. 2.3A). The average full-width half-maximum (FWHM) values measured within the regions of interest shown in Fig. 2.3 were 44 μm for the ULM image and 113 μm for the super harmonic MIP (Fig. 2.3B). These average profiles are overlaid for comparison in Fig. 2.3C. Maps of velocity direction (Fig. 2.4A) and magnitude (Fig. 2.4B) were also created. From Fig. 2.4A, mean angles of flow for these tubes were measured to be 3.0 degrees and 169.3 degrees, which correspond with the cyan and magenta tubes, respectively. From Fig. 2.4B, the average velocity magnitude within the tubes was measured to be 67.5 mm·s^{-1}. For a 46-μm tube, a volume flow rate of 10 μL/min corresponds to an average

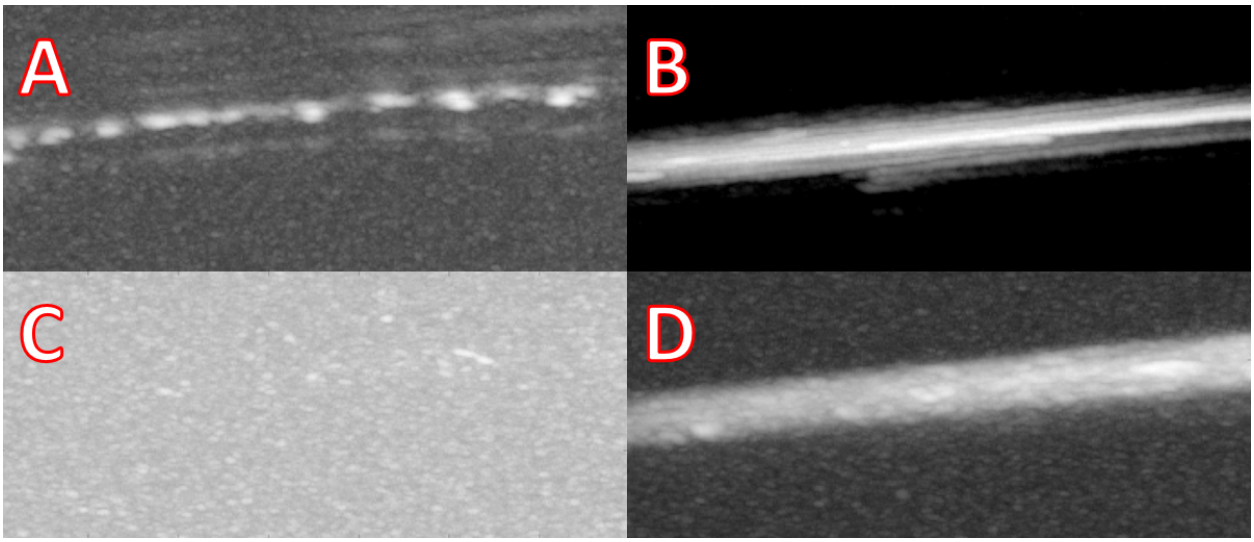

Figure 2.5: Maximum intensity projections for SVD-filtered and super harmonic images of a 200 μmm tube in different flow regimes. All images are displayed on a 25 dB dynamic range for comparison. MIPs of super harmonic images collected at (A) 0.25 μmL/min and (B) 15.0 μmL/min. MIPs of SVD-filtered images collected at (C) 0.25 μmL/min and (D) 15.0 μmL/min.

velocity of 100.3 mm·s^{-1} through a cross-section of the tube. Applying a 2/3 correction factor to account for integrating through elevation [31] predicts the average velocity measured in the ULM imaging plane to be 66.9 mm·s^{-1}, which agrees well with the measurement.

2.3.2 Flow study in vitro

Maximum intensity projections for all of the flow rates and trials were created, and examples of slow- and fast-flow MIPs are provided in Fig. 2.5. The MIPs were generated by envelope detecting the beamformed RF data and taking the maximum through time for each pixel. When infusing a 200-μm tube at 0.25 μL/min and imaging at 500 frames per second, SHI produces an average SNR of 16.5 dB over three trials (Fig. 2.5A). Increasing the volume flow rate to 15 μL/min and holding frame rate constant increases the SNR to 27.4 dB (Fig. 2.5B). SVD filtering produces SNR values of 5.1 dB (Fig. 2.5C) and 18.3 dB (Fig. 2.5D) for the slow- and fast-flow conditions, respectively. Across all flow rates, SHI produces an average improvement in SNR of 10.3 dB compared to SVD filtering (Fig. 2.6).

2.3.3 Kidney imaging in vivo

A super harmonic ULM image of a rodent kidney was generated from 25,000 dual-frequency frames (500 Hz PRF and MI of 0.24) and shows the ability to resolve vessels on the order of 20 μm in diameter (Fig. 2.7). For comparison, a conventional B-mode frame (Fig. 2.7A) and a super harmonic MIP (Fig. 2.7B) are provided. Selected vessels from this ULM dataset have average FWHM values of 20.9, 17.2, and 29.1 μm (Fig. 2.8A-C). The average profiles are provided for comparison in Fig. 2.8D. A ULM image was created from the same dataset without applying motion correction to demonstrate the effects of large magnitude respiratory and cardiac artifacts on image quality (Fig. 2.9A). A qualitative visualization of the performance of the motion correction based on sparsely-interleaved B-mode frames and speckle tracking is provided in Fig. 2.9B. Fig. 2.10 contains the velocity maps corresponding to the motion-corrected ULM frame shown in Fig. 2.7C.

By mechanically scanning the imaging probe in the elevational dimension, three-dimensional ULM datasets were acquired. Fig. 2.11 shows a maximum intensity projection for a rodent kidney dataset (rendered using 3D Slicer 4.10.2, Kitware, Clifton Park, NY). This volume was generated from 17 slices spaced by 500 μm with 25,000 DF frames per slice.

2.4 Discussion

A new approach to ultrasound localization microscopy using super harmonic imaging has been demonstrated both in vitro and in vivo, resolving vessels on the order of 20 μm in diameter in a rodent kidney. SHI offers greater CTR than traditional contrast pulse sequences or SVD filtering while still allowing for motion correction by sparsely interleaving high-frequency B-mode frames into the imaging sequence (1-to-100 ratio). With a dual-frequency arrangement, it is also possible to image slowly moving contrast agents in a cellulose tube in a water bath with much higher SNR than an SVD-based approach. This improvement in SNR may decrease the variance in spatial localizations of slow microbubble contrast agents, which has been modeled as a linear function of the Cramer-Rao lower bound (CRLB) for time delay estimates [3], [30]. The CRLB itself increases strongly as SNR decreases below 10 dB, holding other parameters

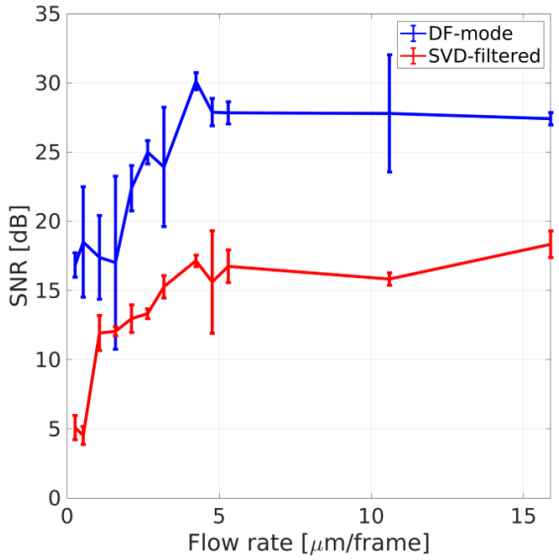

Figure 2.6: SNR vs. flow rate for dual-frequency and SVD-filtered images (blue and red lines, respectively). DF imaging results in an average improvement of 10.3 dB over SVD imaging.

constant. For applications such as molecular imaging, for which the aim is to image stationary bubbles, ULM with SVD filtering may prove challenging even in the absence of physiological motion, assuming that microbubble contrast agents do not decorrelate through slow time.

Interestingly, the results of this flow study revealed a dependence of SNR on flow rate in super harmonic imaging. It is possible that this phenomenon is related to the polydispersity of the contrast agent dilution. The majority of the microbubbles used in this study are around 1 μm in diameter, which have resonance frequencies higher than the 1.7 MHz transmit pulse [31], [32]. For higher flow rates, there is an increased probability that a large bubble with a resonant frequency closer to the LF element center frequency will pass through the field of view during the 1,000-frame acquisition. For slower flow rates, bubbles do not traverse the full length of the tube during a 1,000-frame acquisition (Fig. 2.5A). This means that if a large bubble is not present at the onset of data collection, it is unlikely that one will appear in the tube before all of the frames for that particular trial have been collected. A monodisperse population of bubbles may flatten the SNR vs. flow rate curve for SHI, though this was not investigated.

The study of SNR versus flow rate suffers from some drawbacks, however, such as the rela-

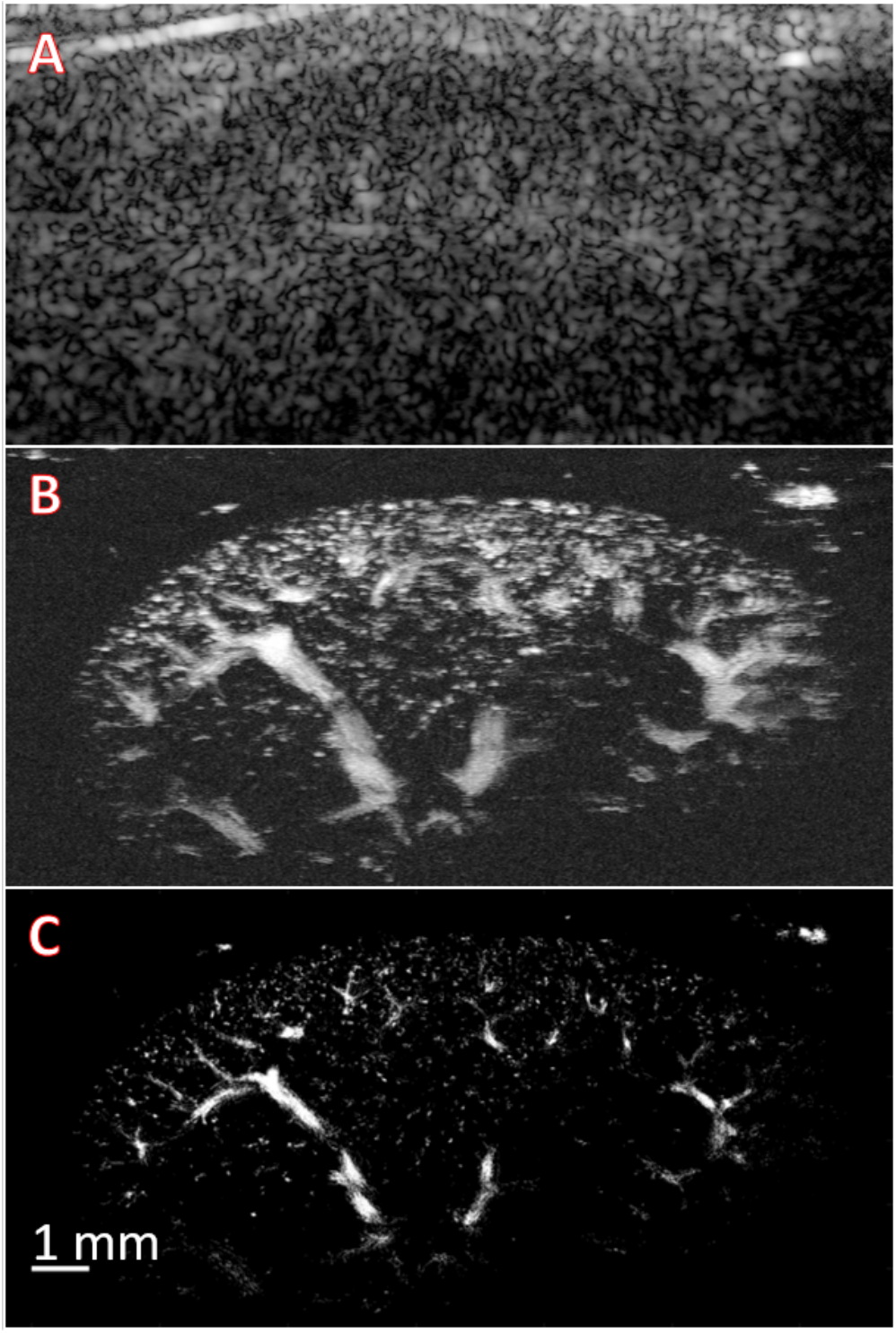

Figure 2.7: Example of super harmonic ultrasound localization microscopy applied to a rodent kidney with motion correction. (A) B-mode scan of the kidney used as a reference for motion correction. (B) Maximum intensity projection of super harmonic images used to generate the ULM image (frames with motion discarded). (C) ULM image generated from 25,000 frames with motion correction applied.

tively low number of trials for each set of parameters (n = 3), which may affect the results shown in Fig. 2.5 and 6. Even with its limitations, the results of this study suggest that dual-frequency imaging outperforms SVD filtering in terms of SNR for all of the flow rates tested between 0.27 and 15.90 μm/frame, and that SHI is better suited for imaging slowly moving contrast agents in a tube when imaging at 500 frames per second. It is important to consider that the performance of the SVD filter depends on both particle speed and frame rate; hence, we report the results as SNR versus μm per frame.

In vitro images of a 46-μm tube resulted in an average FWHM measurement of 44 μm, an error of 4.3%. In vivo, it is quite difficult to assess the accuracy of the ULM imaging without ground truth information regarding the diameter of individual vessels. However, we believe that given the theoretical resolution limit of this system derived in [3] along with the measured error reported earlier, we are justified in assuming the diameters of the selected vessels shown in Fig. 2.8 to be on the order of 20-30 μm, if not smaller. If we assume the resolution error of this system is a fixed 2-μm bias rather than 4.3% of the real value, then the vessels shown in Fig. 2.8 would measure 22.9, 19.2, and 31.1 μm. In any case, these measurements are well below the diffraction-limited resolution of the high frequency array and were collected in a freely breathing rodent without physical constraints.

One limitation of SHI-ULM is the shallow depth of penetration based on the high center frequency of the receiving transducer. This configuration is well-suited for many preclinical scenarios and superficial clinical targets and less so for larger animals and the majority of human organs. However, prior clinical studies have demonstrated super harmonic imaging of microvasculature in the human breast at 25 MHz at less than 2 cm, and we have demonstrated the ability to image microvasculature as deep as 4 cm at 20 MHz in a rodent cancer model [33]. Thus, we hypothesize that SHI-ULM will be relevant for transcutaneous assessment of abnormal angiogenesis or other vascular pathologies in the breast, prostate, thyroid, or other shallow organs and could be used for deeper organs endoscopically. Though the present study was limited to small animal imaging and in vitro experiments, the probe used in this work shows an improvement over previous state

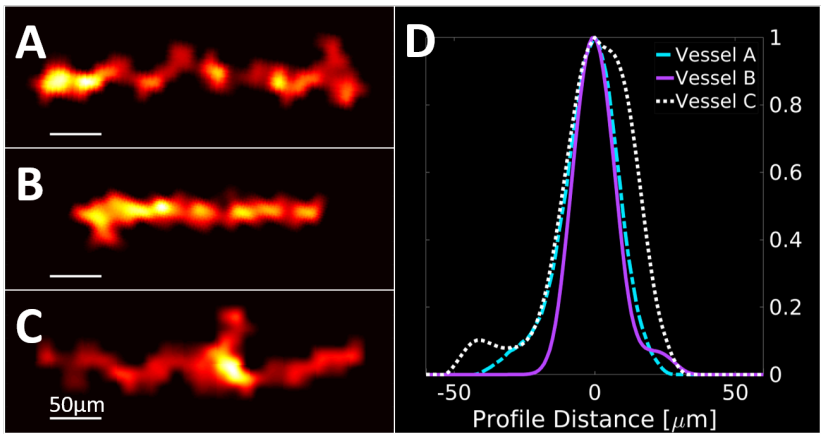

Figure 2.8: Selected vessels from rodent kidney 3-D dataset. (A-C) Example vessels cropped from ULM images. (D) Average profiles of the vessels in panels A-C with full-width at half-maximum values of 20.9, 17.2, and 29.1 μm, respectively.

of the art devices in super harmonic imaging in terms of imaging depth, depth of field, and frame rate. For translation to a clinical population, further study is needed regarding optimal transducer design parameters for an appropriate balance between CTR and imaging depth for dual-frequency ultrasonic imaging.

Another limitation unique to SHI for ULM is the mechanical index (>0.2) necessary to achieve adequate CTR. In these studies, we utilized mechanical indices up to 0.24. While we expect these parameters to be safe based on [34]–[36], this MI is partially destructive to bubbles over repeated pulsing. This might be especially problematic for imaging small capillaries, in which microbubbles may require time scales on the order of minutes to traverse the entire path length of an individual capillary [37]. For this reason, it may provide additional benefit in the future to explore optimization of experimental parameters including frame rate, MI, microbubble formulation and stability, microbubble concentration, infusion rate, and others in an effort to realize the full potential of the SHI approach for ULM.

Another challenge associated with this imaging method is the unique point spread function produced by super harmonic imaging. Under the right circumstances, a single contrast agent will exhibit a point spread function which is multi-modal in the axial dimension due to the strongly nonlinear vibrations of the bubble shell. The presence of such an artifact has a negative impact

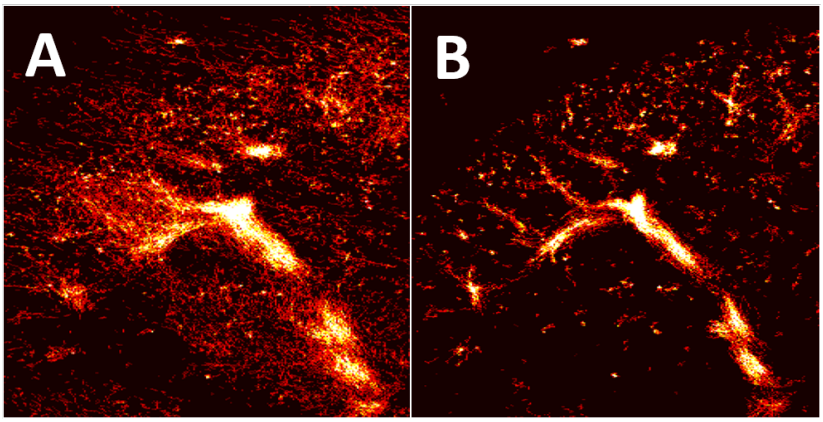

Figure 2.9: A comparison of ULM with and without motion correction based on sparsely interleaved B-mode frames. (A) Rodent kidney vessels are smeared out by respiratory and cardiac motion artifacts. (B) Fine detail of the vessel structure is recovered by a combination of removal of decorrelated frames and using speckle tracking to estimate nonrigid displacements.

on the final image quality if not accounted for because current popular localization methods were not designed with such a phenomenon in mind [38]. In order to control this issue, we have tuned the transmit pressure to attain sufficient CTR for accurate localization while minimizing the multi-modal artifact. This approach, combined with noise thresholding, proved sufficient to mitigate the deleterious effects of the super harmonic artifact. Another approach that can be explored in the future is designing a localization process tailored to the presence of this artifact such that higher mechanical index pulses can be employed to further improve CTR.

It should be noted that the results of the present study are strongly dependent on the characteristics of the contrast agent used during imaging. Recent work has examined the relationship between microbubble parameters and their influence on super harmonic response [39]. One critical parameter is the resonance frequency of the contrast agent, which is largely determined by its diameter [32]. Driving bubbles at or near their resonance frequency leads to strongly nonlinear oscillations of the shell and hence contributes to generating higher harmonics. The results of [39] demonstrate that the in-house bubbles used for the present study are comparable to commercially available contrast agents, such as Definity and Micromarker, in terms of super harmonic backscatter. This finding suggests that the imaging methods described in this work can be replicated in clinical or preclinical settings using commercial bubbles.

One subject that is not studied in this work is the effect of the transducer geometry on ULM image quality. It is certain that the 'X' configuration of the low-frequency elements results in appreciable transmit pressures away from the high frequency array's imaging plane (Fig. 2.1C). While off-target bubbles are sonicated on transmit, hydrophone measurements show the elevational beam-width of the high frequency transducer ranges between 0.5 and 1.0 mm over the main lobe of the low-frequency transmission. It follows that this system is not sensitive to contrast agents that are more than 0.5 mm out of plane. However, we must consider the depth-dependent response of the system imposed by the broadening high frequency beam width. Precisely controlling the contrast concentration in the blood pool ensures that we retain a sparse group of bubbles in each frame even as we receive with a thicker beam at greater depths. Another source of depth-dependence that is not directly accounted for in this study is the variable amplitude of the transmitted pressure in the axial dimension which is given by the degree of overlap between the crossed LF beams. It should be noted, however, that these specific limitations are unique to this sort of transducer design and are not necessarily associated with dual-frequency imaging in general.

As mentioned previously, the current system is suitable for imaging preclinical models, such as rodents, but is not flexible enough for interrogating targets located beyond the mechanically fixed beam pattern. Perhaps future research will focus on the continued development of confocal dual-frequency probes, such as that demonstrated by Van Neer and colleagues in [23], to further improve this imaging method. A fully confocal array design would significantly improve the limited depth of field of a cross-beam transducer (11 mm in this study), allowing for interrogation of larger targets. It is also possible that using dual-frequency transducers with transmit / receive frequencies lower than the 1.7 / 20 MHz used in the present study will allow for deeper super harmonic imaging. While lower frequencies will result in a larger diffraction-limited resolution, we expect to recover resolution with ULM.

Another area that requires further exploration is the parameter space for motion correction based on sparsely interleaved B-mode acquisitions. Fig. 2.9 shows an example of the improve-

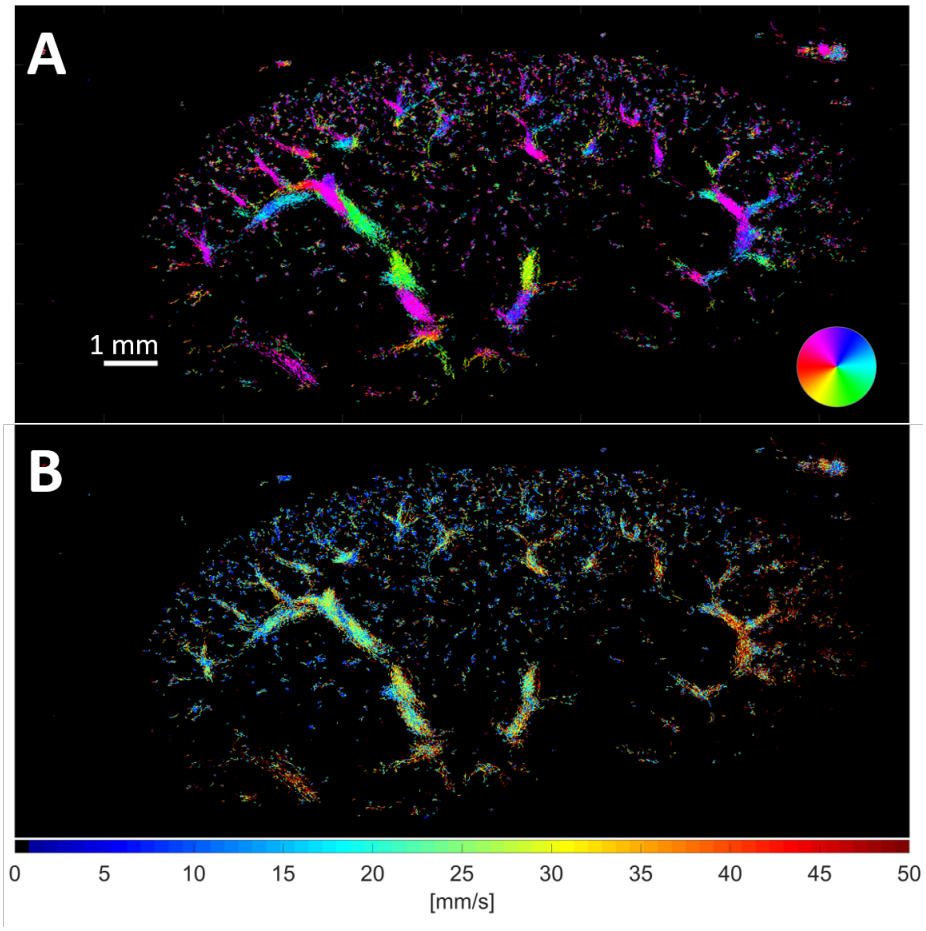

Figure 2.10: Tracking bubbles in vivo allows for the mapping of blood velocity in a rodent kidney. (A) The average direction of microbubbles for the ULM image in Figure 2.7. (B) The magnitude of the velocity for the same dataset.

ment in image quality provided by this algorithm, though we believe that most of the improvement in image quality is derived from simply discarding batches of frames associated with large physiological motions. It is possible that moving to a smaller ratio of SHI to B-mode frames will allow for higher-fidelity speckle tracking based on the smaller decorrelation between B-mode frames of neighboring acquisitions. The in vivo images shown in this article were produced with a 100-to-1 ratio in which B-mode frames were separated temporally by 200 ms. This B-mode frame rate is sufficient for tracking respiratory motion, but must be increased to fully sample the cardiac motion of the rodent model. It is difficult to quantify the performance of this motion correction approach in vivo because we lack ground truth information. Further studies may focus on

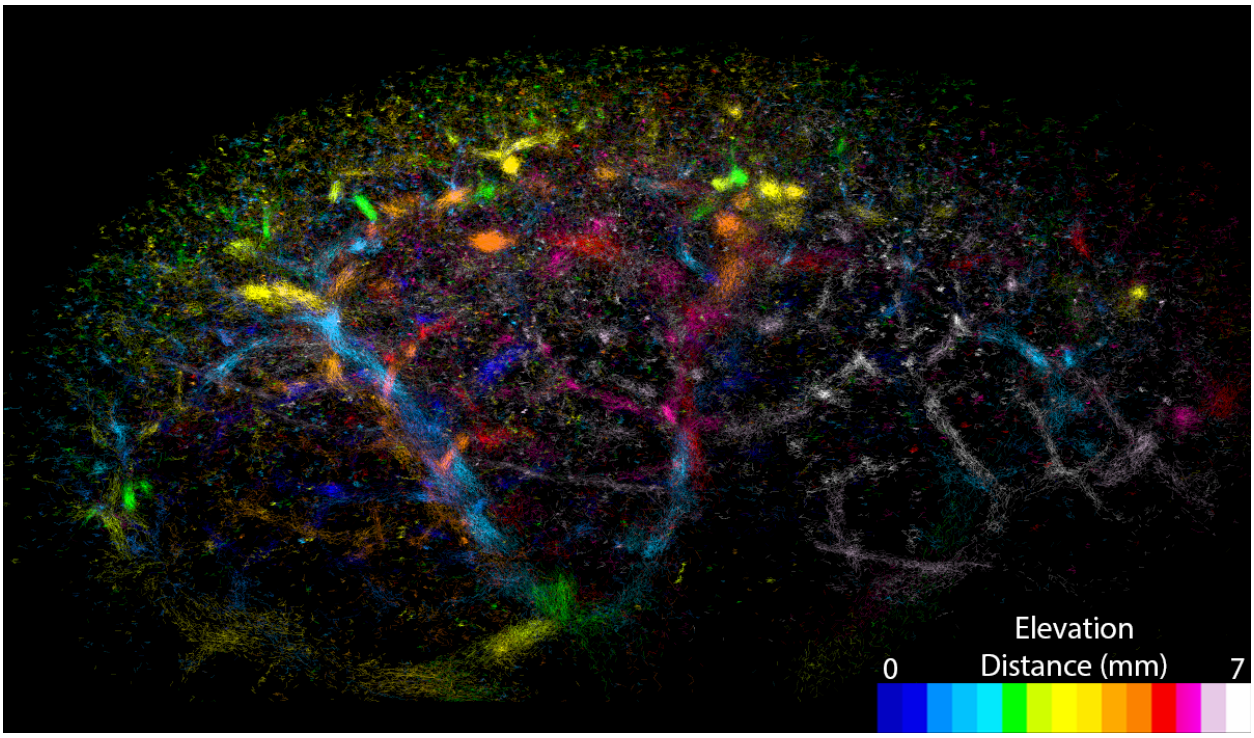

Figure 2.11: An example of 3-D ultrasound localization microscopy with super harmonic imaging by mechanically scanning the transducer in the elevational dimension. This image was generated with 17 slices spaced at 500 μm and contains vessels on the order of 20 μm.

characterizing this approach via simulations and in vitro.

This study also accomplished three-dimensional ULM in a similar fashion to the methods used by Lin and colleagues in [9]. However, because of time constraints during imaging, a relatively large step size of 500 μm was used, which means the elevational resolution was orders of magnitude worse than the axial or lateral resolution. This sort of volume might be useful for evaluating metrics such as vascular density, but will likely fall short for accurately assessing features such as tortuosity. However, this study highlights the potential of utilizing ULM for imaging whole organs in preclinical targets. Improvements in transducer technology might one day lead to fully-sampled matrix arrays capable of ultrafast super harmonic imaging for ULM.

2.5 Conclusion

SHI improves SNR by more than 10 dB in vitro compared to SVD-filtering for average flow rates between 0.3 and 15.9 μm/frame. Since the method does not rely on motion to discriminate contrast from background signal, we expect SHI to work well even when microbubbles are sta-

tionary relative to tissue. Furthermore, SHI operates without the need to tune the singular vector threshold for each dataset, which can be a cumbersome manual process. [14] demonstrated that the most successful automatic threshold estimator for SVD filtering achieves optimal CTR for only roughly 60% of in vivo datasets. SHI, on the other hand, is a robust imaging scheme that requires a simple background noise threshold to produce images suitable for ULM processing. Furthermore, a relatively simple speckle tracking scheme based on [29] applied over sparsely interleaved B-mode frames provides a framework for nonrigid displacement corrections without the need for optimizing a nonrigid transformation estimator such as [40]. Super harmonic imaging therefore offers a straightforward approach to bubble detection for ultrasound localization microscopy, even for challenging imaging scenarios, such as in the presence of slow flow or high frequency physiological motion.

REFERENCES

[1] C. Errico, J. Pierre, S. Pezet, Y. Desailly, Z. Lenkei, O. Couture, and M. Tanter, "Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging," *Nature*, vol. 527, no. 7579, pp. 499–502, 2015.

[2] K. Christensen-Jeffries, R. J. Browning, M. X. Tang, C. Dunsby, and R. J. Eckersley, "In vivo acoustic super-resolution and super-resolved velocity mapping using microbubbles," *IEEE Transactions on Medical Imaging*, vol. 34, no. 2, pp. 433–440, 2015.

[3] Y. Desailly, J. Pierre, O. Couture, and M. Tanter, "Resolution limits of ultrafast ultrasound localization microscopy," *Physics in Medicine and Biology*, vol. 60, pp. 8723–8740, 22 2015.

[4] P. Carmeliet and R. K. Jain, "Angiogenesis in cancer and other diseases," *Nature*, vol. 407, pp. 249–257, 6801 2000.

[5] "Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011.

[6] R. C. Gessner, S. R. Aylward, and P. A. Dayton, "Mapping microvasculature with acoustic angiography yields quantifiable differences between healthy and tumor-bearing tissue volumes in a rodent model," *Radiology*, vol. 264, no. 3, pp. 733–740, 2012.

[7] S. E. Shelton, Y. Z. Lee, M. Lee, E. Cherin, F. S. Foster, S. R. Aylward, and P. A. Dayton, "Quantification of microvascular tortuosity during tumor evolution using acoustic angiography," *Ultrasound in Medicine and Biology*, vol. 41, no. 7, pp. 1896–1904, 2015.

[8] S. R. Rao, S. E. Shelton, and P. A. Dayton, "The "fingerprint" of cancer extends beyond solid tumor boundaries: Assessment with a novel ultrasound imaging approach," *IEEE Transactions on Biomedical Engineering*, vol. 63, pp. 1082–1086, 5 2016.

[9] F. Lin, S. E. Shelton, D. Espindola, J. D. Rojas, G. Pinton, and P. A. Dayton, "3-d ultrasound localization microscopy for identifying microvascular morphology features of tumor angiogenesis at a resolution beyond the diffraction limit of conventional ultrasound," *Theranostics*, vol. 7, pp. 196–204, 1 2017.

[10] T. Opacic, S. Dencks, B. Theek, M. Piepenbrock, D. Ackermann, A. Rix, T. Lammers, E. Stickeler, S. Delorme, G. Schmitz, and F. Kiessling, "Motion model ultrasound localization microscopy for preclinical and clinical multiparametric tumor characterization," *Nature Communications*, vol. 9, pp. 1–13, 1 2018.

[11] O. Couture, V. Hingot, B. Heiles, P. Muleki-Seya, and M. Tanter, "Ultrasound localization microscopy and super-resolution: A state of the art," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 65, pp. 1304–1320, 8 2018.

[12] C. Demené, T. Deffieux, M. Pernot, B. F. Osmanski, V. Biran, J. L. Gennisson, L. A. Sieu, A. Bergel, S. Franqui, J. M. Correas, I. Cohen, O. Baud, and M. Tanter, "Spatiotemporal clutter filtering of ultrafast ultrasound data highly increases doppler and fultrasound sensitivity," *IEEE Transactions on Medical Imaging*, vol. 34, pp. 2271–2285, 11 2015.

[13] Y. Desailly, A. M. Tissier, J. M. Correas, F. Wintzenrieth, M. Tanter, and O. Couture, "Contrast enhanced ultrasound by real-time spatiotemporal filtering of ultrafast images," *Physics in Medicine and Biology*, vol. 62, pp. 31–42, 1 2017.

[14] J. Baranger, B. Arnal, F. Perren, O. Baud, M. Tanter, and C. Demene, "Adaptive spatiotemporal svd clutter filtering for ultrafast doppler imaging using similarity of spatial singular vectors," *IEEE Transactions on Medical Imaging*, vol. 37, pp. 1574–1586, 7 2018.

[15] B. Fagrell, A. Fronek, and M. Intaglietta, "A microscope-television system for studying flow velocity in human skin capillaries," *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 233, pp. 318–321, 2 1977.

[16] C. G. Caro, T. J. Pedley, R. C. Schroter, and W. A. Seed, *The mechanics of the circulation*. Cambridge University Press, 2012.

[17] J. Foiret, H. Zhang, T. Ilovitsh, L. Mahakian, S. Tam, and K. W. Ferrara, "Ultrasound localization microscopy to image and assess microvasculature in a rat kidney," *Scientific Reports*, vol. 7, pp. 1–12, 1 2017.

[18] A. Bouakaz, S. Frigstad, F. J. T. Cate, and N. D. Jong, "Super harmonic imaging: A new imaging technique for improved contrast detection," *Medicine*, vol. 28, pp. 59–68, 1 2002.

[19] J.-J. Hwang and D. H. Simpson, "Two pulse technique for ultrasonic harmonic imaging," US Patent 5,951,478, 1999.

[20] G. A. Brock-Fisher, M. D. Poland, and P. G. Rafter, "Means for increasing sensitivity in non-linear ultrasound imaging systems," US Patent 5,577,505, 1996.

[21] C. Tremblay-Darveau, R. Williams, L. Milot, M. Bruce, and P. N. Burns, "Visualizing the tumor microvasculature with a nonlinear plane-wave doppler imaging scheme based on amplitude modulation," *IEEE Transactions on Medical Imaging*, vol. 35, pp. 699–709, 2 2016.

[22] P. J. Phillips, "Contrast pulse sequences (cps): Imaging nonlinear microbubbles," *Proceedings of the IEEE Ultrasonics Symposium*, vol. 2, pp. 1739–1745, 2001.

[23] P. L. van Neer, G. Matte, M. G. Danilouchkine, N. de Jong, C. Prins, and F. van den Adel, "Super-harmonic imaging: Development of an interleaved phased-array transducer," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 57, pp. 455–468, 2 2010.

[24] R. C. Gessner, C. B. Frederick, F. S. Foster, and P. A. Dayton, "Acoustic angiography: A new imaging modality for assessing microvasculature architecture," *International Journal of Biomedical Imaging*, vol. 2013, 2013.

[25] S. E. Shelton, B. D. Lindsey, P. A. Dayton, and Y. Z. Lee, "First-in-human study of acoustic angiography in the breast and peripheral vasculature," *Ultrasound in Medicine and Biology*, vol. 43, pp. 2939–2946, 12 2017.

[26] "In Vitro Superharmonic Contrast Imaging Using a Hybrid Dual-Frequency Probe," *Ultrasound in Medicine and Biology*, vol. 45, no. 9, pp. 2525–2539, 2019.

[27] J. K. Tsuruta, N. P. Schaub, J. D. Rojas, J. Streeter, N. Klauber-DeMore, and P. Dayton, "Optimizing ultrasound molecular imaging of secreted frizzled related protein 2 expression in angiosarcoma," *PLoS ONE*, vol. 12, pp. 1–21, 3 2017.

[28] B. D. Lindsey, S. E. Shelton, and P. A. Dayton, "Optimization of contrast-to-tissue ratio through pulse windowing in dual-frequency "acoustic angiography" imaging," *Ultrasound in Medicine and Biology*, vol. 41, pp. 1884–1895, 7 2015.

[29] J. Luo and E. Konofagou, "A fast normalized cross-correlation calculation method for motion estimation," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 57, pp. 1347–1357, 6 2010.

[30] W. F. Walker and G. E. Trahey, "A fundamental limit on delay estimation using partially correlated speckle signals," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 42, pp. 301–308, 2 1995.

[31] D. Espindola, R. M. DeRuiter, F. Santibanez, P. A. Dayton, and G. Pinton, "Quantitative sub-resolution blood velocity estimation using ultrasound localization microscopy ex-vivo and in-vivo," *Biomedical Physics & Engineering Express*, vol. 6, no. 3, p. 035 019, 2020.

[32] N. de Jong, M. Emmer, A. van Wamel, and M. Versluis, "Ultrasonic characterization of ultrasound contrast agents," *Medical and Biological Engineering and Computing*, vol. 47, pp. 861–873, 8 2009.

[33] B. D. Lindsey, J. Kim, P. A. Dayton, and X. Jiang, "Dual-frequency piezoelectric endoscopic transducer for imaging vascular invasion in pancreatic cancer," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 64, pp. 1078–1086, 7 2017.

[34] "Optison package insert," GE Healthcare AS, 2016.

[35] "Definity package insert," Lantheus Medical Imaging, 2015.

[36] "Lumason package insert," Bracco Diagnostics, Inc., 2016.

[37] V. Hingot, C. Errico, B. Heiles, L. Rahal, M. Tanter, and O. Couture, "Microvascular flow dictates the compromise between spatial resolution and acquisition time in Ultrasound Localization Microscopy," *Scientific Reports*, vol. 9, no. 1, p. 2456, 2019.

[38] K. C. Jeffries, S. Harput, J. Brown, C. Dunsby, P. Aljabar, M. X. Tang, and R. Eckersley, "Microbubble axial localization errors in ultrasonic super-resolution imaging," *IEEE International Ultrasonics Symposium, IUS*, vol. 64, pp. 1644–1654, 11 2017.

[39] I. G. Newsome, T. M. Kierski, and P. A. Dayton, "Assessment of the superharmonic response of microbubble contrast agents for acoustic angiography as a function of microbubble parameters," *Ultrasound in Medicine and Biology*, vol. 45, pp. 2515–2524, 9 2019.

[40] S. Harput, K. Christensen-Jeffries, J. Brown, Y. Li, K. J. Williams, A. H. Davies, R. J. Eckersley, C. Dunsby, and M. X. Tang, "Two-stage motion correction for super-resolution ultrasound imaging in human lower limb," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 65, pp. 803–814, 5 2018.

CHAPTER 3: ACOUSTIC MOLECULAR IMAGING BEYOND THE DIFFRACTION LIMIT[1]

3.1 Introduction

Given the diversity of neoplastic diseases and their potential to develop into life-threatening conditions, there is a need for safe and effective characterization of tissues across many applications. Assessment of disease biomarkers, such as the vascular "fingerprint" of tumors and their associated microenvironment, is infeasible with many biomedical imaging techniques and typically requires terminal pathology studies. One hallmark of cancer is the deregulation of angiogenic signaling, which results in a chaotic and densely-packed network of blood vessels around the growth [1], [2].

Ultrasound imaging is a good candidate for assessing cancers *in vivo* because it provides excellent spatial and temporal resolutions, does not expose patients to ionizing radiation, and is substantially less expensive and more accessible compared to modalities such as magnetic resonance imaging and positron emission tomography. In addition to imaging anatomical structures, ultrasound may also be used to measure the mechanical properties of tissues [3] as well as the velocity of blood flow [4] in real-time. The clinical utility and versatility of biomedical ultrasound is further increased by microbubble contrast agents (MCAs). MCAs are typically composed of a heavy gas core and a lipid, protein, or polymer shell, with diameters between approximately 1 and 10 μm in diameter. When administered intravenously, these contrast agents serve as blood pool markers. Contrast-enhanced ultrasound imaging (CEUS) has many applications in the context of cancer, such as quantifying perfusion in suspicious lesions [5], [6].

With ultrasound molecular imaging (USMI), it is also possible to visualize biomarker expression *in vivo* by imaging MCAs that have been modified to interact with a specific vascular target.

[1]Portions of this chapter are under review in the IEEE Open Journal of Ultrasonics, Ferroelectrics, and Frequency Control.

Microbubbles can be functionalized for USMI by the addition of one or more ligands to the shell architecture [7]. When targeted bubbles are introduced into circulation, they flow until binding at a site of interest. In comparison, non-targeted microbubbles flow freely throughout the vasculature. Some notable applications of USMI include early detection of cancer [8], classifying breast lesions [9], quantifying inflammation [10], and monitoring responses to treatments [11]–[14].

Differential targeted enhancement (dTE) is a common approach for estimating the distribution of contrast agent binding *in vivo* [7], [15]–[17]. For this method, targeted MCAs are injected and allowed to circulate for a predetermined length of time to allow microbubble targeting and clearance of residual unbound contrast. An image is acquired at this point, after which microbubbles in the field of view are disrupted with a high-amplitude ultrasound transmission. Some time is given for any remaining bubbles to reperfuse the field of view, after which an additional image is collected. The difference between the pre- and post-disruption images provides an estimate of contrast agent binding over a region of interest. While this method is effective for estimating the amount of targeting within a tumor, its resolution is limited by diffraction. Many diagnostically relevant structures exist at spatial scales smaller than the point spread function of a standard clinical imaging system (*e.g.*, capillaries from tumor-associated angiogenesis). Alternative approaches for assessing bound MCAs have been proposed, such as those based on dwell time [14], normalized singular spectrum area [18], and convolutional neural networks [15]. These techniques are promising real-time methods for estimating molecular expression. However, similar to dTE, the resolutions of these methods are largely governed by diffraction.

Ultrasound localization microscopy (ULM) is a super-resolution imaging technique analogous to optical localization methods [19]–[22], and improves resolution by an order of magnitude [23]–[25] enabling noninvasive imaging of capillaries. Briefly, a bolus of MCAs is administered intravenously, and a large sequence of images is collected (normally $N > 1000$). From these images, individual microbubbles are isolated from the background speckle and localized onto a super-resolution grid. These localizations accumulate over the full set of images to produce a map of the underlying vessel structure that is much finer than the point spread function of the

imaging system. To date, a fundamental limitation of ULM has been that most methods for isolating MCAs from tissue before localization rely on the spatiotemporal decorrelation that results from MCAs flowing through the circulatory system while the sequence of images is acquired. However, if a microbubble is bound to an endothelial target, it is in a zero-velocity state relative to the nearby tissue. Hence, spatiotemporal filtering approaches are not feasible for imaging stationary MCAs, such as those in USMI.

In contrast to biological tissues, microbubble contrast agents are resonant structures which oscillate with broadband harmonics when excited by a sound wave near their resonant frequency. Thus, MCAs can be detected spectrally instead of spatiotemporally. Superharmonic imaging (SpHI) is a technique that takes advantage of this phenomenon by recording the higher harmonics (*i.e.*, superharmonics) of the transmitted frequency to create a high-resolution image nearly devoid of tissue speckle. Notably, superharmonic generation is not influenced by the velocity of MCAs, resulting in excellent contrast-to-tissue ratio even for stationary bubbles. Until recently, SpHI had not been applied to ULM because commercially available ultrasound transducers do not have the bandwidth necessary for receiving echoes beyond the third harmonic of the transmitted pulse. Our team recently developed a multi-frequency linear array transducer for plane-wave SpHI [26] and demonstrated its high sensitivity to MCAs independent of their velocity [27]. In this work, we combine this dual-frequency transducer with targeted MCAs, ULM processing, and microbubble tracking, to achieve for the first time super-resolution imaging of molecularly bound contrast agents *in vivo*. We provide a description and validation of this novel method for super-resolution USMI using superharmonic ultrasound localization microscopy to produce co-registered maps of microvessels and angiogenic signaling.

3.2 Materials and Methods

3.2.1 Contrast Agent Preparation

Our in-house non-targeted microbubble contrast agent was formulated from a 1 mM lipid solution that contained 900 μM 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC) and 100 μM 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000]

Contrast agent	Binding target	Mean diameter	Standard deviation	Stock concentration
Control	N/A	1.67 μm	1.10 μm	1.0e10 mL^{-1}
cRGD	$\alpha_v\beta_3$ integrin	1.51 μm	0.92 μm	5.6e9 mL^{-1}
Biotin	Avidin	1.21 μm	0.76 μm	3.3e10 mL^{-1}

Table 3.1: Summary of the ultrasound contrast agents used in the present study.

(DSPE-PEG2000) lipids in 5% (v/v) glycerol and 15% (v/v) propylene glycol in phosphate-buffered saline (PBS). A biotinylated microbubble contrast agent (4.5 mole %) was formulated by replacing 45 μM of the DSPE-PEG2000 in the in-house formulation with 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[biotinyl(polyethylene glycol)-2000] (DSPE-PEG2000-biotin) (Avanti Polar Lipids, Alabaster, AL, USA). A cRGD microbubble contrast agent (0.4 mole %) targeted to bind to $\alpha_v\beta_3$ integrin was formulated from our in-house lipid solution supplemented with 4 μM synthetic azide-activated cyclo-Arg-Gly-Asp (cRGD) peptide (Peptides Int'l, Louisville, Kentucky, USA) conjugated to DSPE-PEG2000-DBCO lipid via click chemistry. The size distributions and concentrations of each contrast agent were measured using an Accusizer FX-Nano (Entegris, Billerica, MA, USA). Table 1 contains a summary of each MCA formulation.

3.2.2 Dual-frequency system description

A dual-frequency array transducer designed for unsteered plane wave transmission was used for all superharmonic imaging [26]. A 1.7 MHz transmit pulse was generated using an arbitrary waveform generator (AWG 2021, Tektronix, Beaverton, OR, USA) and passed through a 50 dB amplifier (240 L, ENI, Rochester, NY, USA) to drive two large low-frequency elements. Superharmonic echoes were detected using an 18 MHz linear array transducer (MS-250, VisualSonics, Toronto, Canada) connected to a high-frequency Vantage 256 research platform (Verasonics, Kirkland, WA, USA). B-scan imaging was performed by both transmitting a 2-cycle plane wave and receiving with the linear array. For both imaging modes, radiofrequency data were collected with 62.5 MHz sampling and a bandpass filter with 15.625 MHz center frequency and 66% bandwidth. SpHI and B-mode imaging were performed at mechanical indices (MI) of 0.24 and 0.11,

respectively (measured in water using HNA-0400, Onda Corporation, Sunnyvale, CA, USA). Volumetric imaging *in vivo* was performed using a motion stage (XSlide, Velmex, Inc., Bloomfield, NY, USA) controlled by a custom LabVIEW program (National Instruments, Austin, TX, USA) to translate the probe in the elevational dimension.

3.2.3 Baseline resolution measurement for the hybrid dual-frequency array

Similar to our previous work [28], we created a very dilute suspension of microbubble contrast agents by adding $1.3 * 10^6$ microbubbles to a 4-liter tank of distilled water. A stir plate (Thermolyne Cimarec, Barnstead International, Dubuque, IO, USA) set to its lowest setting was used to keep the solution well-mixed. This suspension was imaged using the dual-frequency system at 10 frames per second (fps) and MI = 0.24, with the frame rate chosen to guarantee independent arrangements of bubbles across different images. 20 seconds of data were collected for a total of 200 unique images. Delay-and-sum beamforming was performed offline with F# = 2 at depths between 17 mm and 27 mm, and lateral positions between +/- 7.5 mm. Axial and lateral profiles were automatically extracted from microbubbles in the beamformed images using a custom MATLAB routine, and the full-width at half-maximum (FWHM) values were measured.

3.2.4 Data collection for in vitro binding experiment

A cellulose tube (200 μm diameter) was coated with a solution of 5 mg/mL avidin in PBS at room temperature and stored at 4°C for 16 hours. The tube was then submerged in a tank of water in the field of view of an inverted microscope (IX71, Olympus, Tokyo, Japan) using a 60x immersion lens. The microscope was connected to a high-speed camera (FASTCAM SA1.1. Photron, Tokyo, Japan) that digitized images on a 1024x1024 grid at 250 fps. After focusing the microscope on the upper wall of the cellulose tube, B-mode imaging was utilized to align the dual-frequency scanner to the tube.

For all imaging experiments related to tuning the molecular localization algorithm, a contrast agent was diluted to 1e8 MCA/mL in PBS and injected into the tube. The bubbles were allowed to float for three minutes to allow binding to occur, after which a volume flow rate of 10 μL/min was imposed with a syringe pump (Harvard Apparatus, Holliston, MA, USA). A video was cap-

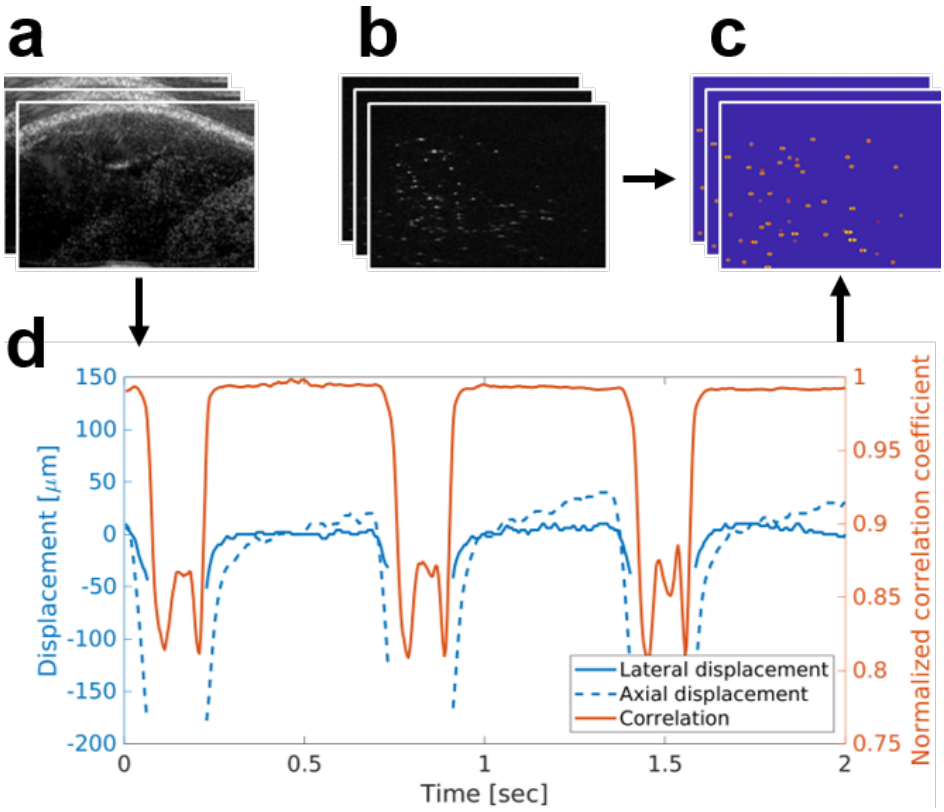

Figure 3.1: A flowchart outlining the steps for molecular ultrasound localization microscopy. (a) B-mode images are interleaved between (b) superharmonic contrast images. (c) Microbubbles are detected in the contrast images and localized. (d) Displacements are measured from the B-modes, and used to correct the coordinates in (c). Motion correction is only performed for *in vivo* imaging. A microbubble is considered bound if it persists locally (after motion correction, if applicable) for 12 or more consecutive time points.

tured at the beginning of the flow to assess contrast-avidin binding. After one minute of flow, 500 superharmonic ultrasound images were collected at 250 fps. Control and targeted imaging were performed using house bubbles and biotin bubbles, respectively.

3.2.5 Animal care and in vivo data collection

In vivo imaging was performed on female Fischer 344 rats (Charles River Laboratories, Durham, NC, USA), and all imaging protocols were approved by the Institutional Animal Care and Use Committee at the University of North Carolina at Chapel Hill. Animals were housed in a cage measuring 140 square inches (individually-ventilated with a static micro-isolator) located within a vivarium with a simulated day-night cycle. Animals received regular daily monitoring and could freely access water and standard rat feed. At the time of imaging, all animals were

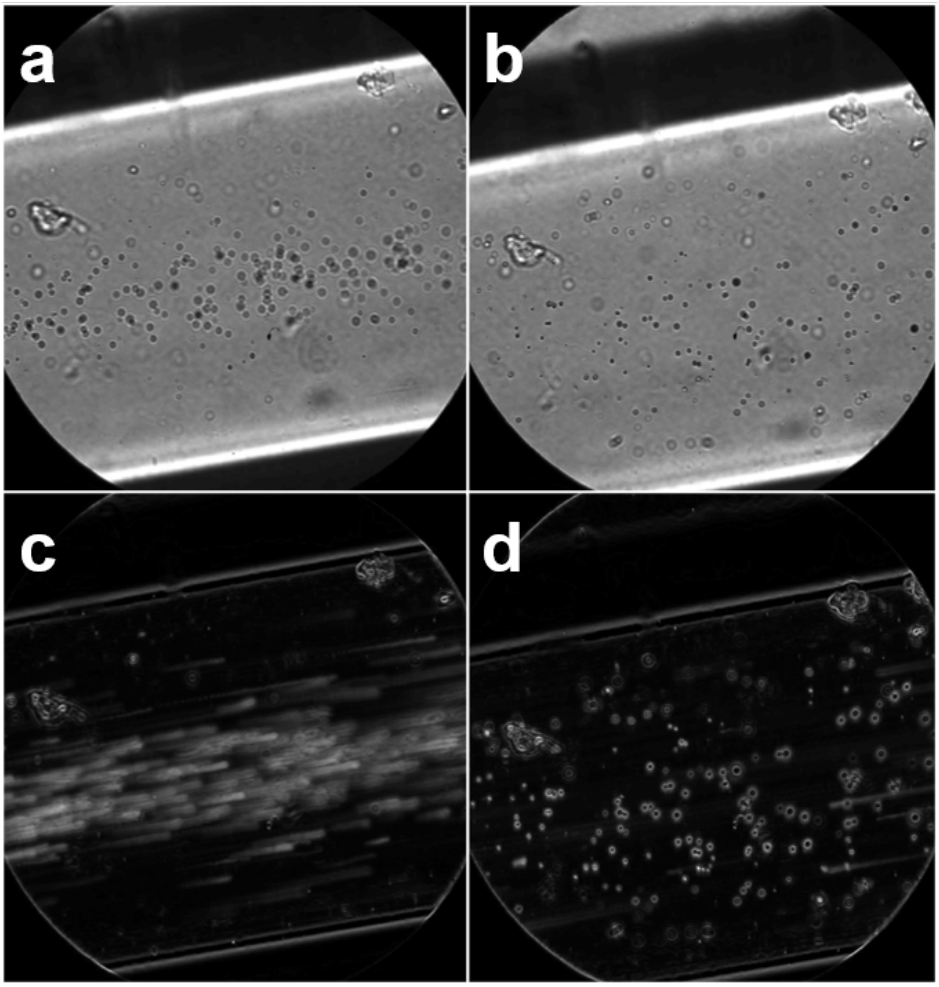

Figure 3.2: Optical comparison of control and biotinylated contrast agents in a microvessel phantom coated with avidin. *Top row*: bubbles near the upper wall of the tube after 3 min of flotation for control (a) and biotin (b) trials. *Bottom row*: Standard deviation images generated from 1 second of optical data captured after introducing flow of saline for control (c) and biotin (d) microbubbles. Streaks in the standard deviation image result from bubble movement.

between 11-12 weeks old and each weighed between 145 and 155 grams.

The fibrosarcoma (FSA) model was prepared as described previously [29], [30]. Small volumes of FSA tissue between 1-2 mm^3 from a donor animal were transplanted into the right flanks of three rats, and tumors were allowed to grow for 10 days before imaging. For imaging experiments, animals were anesthetized with vaporized isoflurane mixed with oxygen, and the target area was shaved. A 24-gauge catheter was placed in the tail vein for administering the microbubble contrast agents.

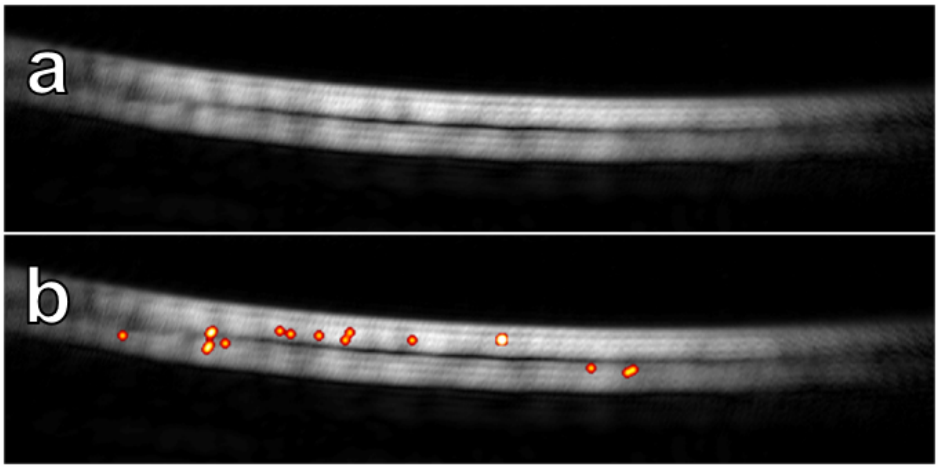

Figure 3.3: Example super-resolution molecular images from control and targeted trials in a microflow phantom. (a) Processing data for the unmodified (control) contrast agent yields no detections. (b) The interaction between the biotinylated microbubbles and the avidin coating produces numerous localizations after processing (**localizations are blurred for visualization only**).

For all animals, a volumetric B-mode scan was acquired before contrast imaging for anatomical reference. Then, a 100 μL bolus of 1e8 cRGD-targeted MCAs was injected via tail-vein catheter, followed by a two-minute wait to allow the bubbles to circulate and bind. Afterward, sets of 500 SpHI frames (250 fps) were acquired at elevational positions spaced by 1 mm. B-mode images were interleaved between SpHI frames for speckle tracking. After the targeted scans, data for a background ULM image was acquired at each elevational position. For each slice, a 100 μL bolus of 1e8 non-targeted MCAs was injected, and 25,000 superharmonic images were immediately acquired at 250 fps with B-modes interleaved.

3.2.6 Localization microscopy processing

For ULM processing, all SpHI images were beamformed offline using delay-and-sum on a 10-μm grid. Each image was thresholded at three times the noise floor (empirically determined), after which bubbles were localized by first convolving with a Gaussian aperture calibrated to the point spread function of the imaging system and detecting local maxima. For *in vivo* datasets, interleaved B-mode images were beamformed on a 10-μm grid, and a 1 mm $\times$ 1 mm region of interest (ROI) beneath the skin was manually selected for each elevational position. Speckle tracking with a normalized cross-correlation (NCC) search was performed on these ROIs to es-

timate displacement during image acquisition. B-modes with estimated displacements less than 100 μm and NCC values greater than or equal to 0.95 were used to correct localization coordinates from their corresponding SpHI frames, and the remaining localizations were discarded.

After motion correction (*in vivo* data only), bubbles were tracked between frames using the Hungarian algorithm (simpletracker, MATLAB). For the molecular targeting data, the max linking distance between frames was set to one pixel, and only tracks with a length of at least 12 frames (48 milliseconds) were considered. The minimum of 12 frames was determined to be sufficient for filtering out moving bubbles while ensuring that bound bubbles were not deflated from repeated sonication.

Tracks with a final coordinate located within one pixel of the first tracked coordinate were considered to be bound bubbles. This threshold of a single pixel was deemed appropriate to prevent filtering out bubbles whose positions might seemingly fluctuate as a result of localization error. The first index of each of these tracks was used as the coordinate of the bound MCA. A summary of the molecular imaging processing is provided in Fig. 3.1. For background ULM images, the max linking distance between frames was set to 100 μm, which corresponds to a maximum velocity of 25 mm·s^{-1}. Only tracks with lengths greater than or equal to 10 frames were considered to reduce noise in the rendered images. The tracks were accumulated to generate the final images. Molecular and background ULM images were co-registered using their respective reference B-mode frames used for motion tracking.

3.2.7 Optical - Ultrasound Calibration

To ensure that the output of the ULM system scales linearly with the true local density of microbubbles, we also performed a calibration by comparing MCA counts in optical and ultrasound images across various concentrations of non-targeted microbubbles. Using the experimental setup described in section II-B, we collected concurrent optical and ultrasound videos of dilutions of non-targeted contrast agents with concentrations between 5e5 and 1e7 MCA/mL flowing through the tube at 5 μL/min. The output of the inverted microscope was digitized with a DSLR camera (X-T2, Fujifilm, Tokyo, Japan) on a 1920x1080 pixel grid at a rate of 30 fps. Ultrasound radiofre-

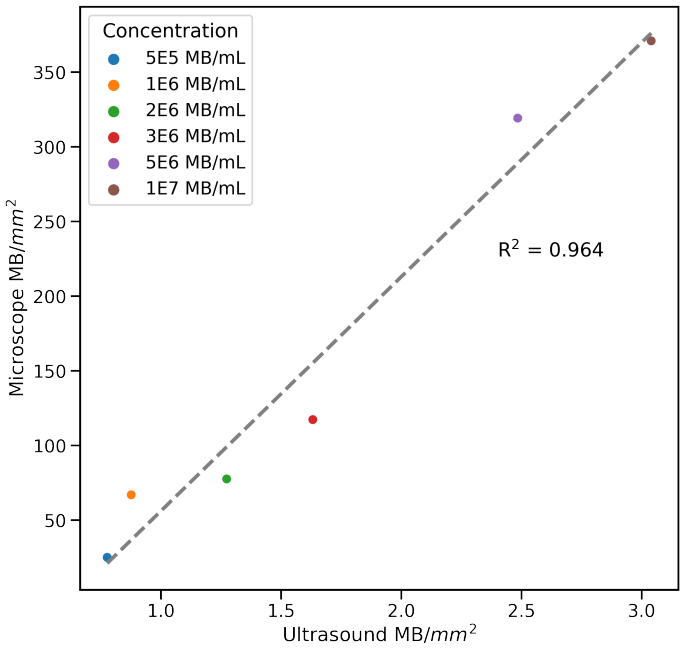

Figure 3.4: Measuring the correlation between optical and dual-frequency ultrasound bubble counts. Counting bubbles in optical and ultrasound videos across a range of different MCA concentrations suggests that the ultrasound bubble count scales linearly with respect to the ground truth bubble density.

quency data were collected at 100 fps and beamformed with a pixel size of 10 μm.

Microbubbles were counted in the ultrasound images using the techniques described in section II-F. The optical images were processed as follows:

1. An 8th order bandpass Butterworth filter with low and high frequency cutoffs of 0.75 and 29 Hz, respectively, was applied to each pixel of the optical data to remove the static background and some high-frequency noise.

2. Each pixel was replaced with its absolute value and then thresholded to 3x the noise floor (empirically measured).

3. The cube root of each pixel was taken to compress the dynamic range of the images.

4. Circles in each frame were detected using a multiscale Hough transform (MATLAB function *imfindcircles*).

Bubble counts were divided by the areas of the tube visible for each imaging modality to convert

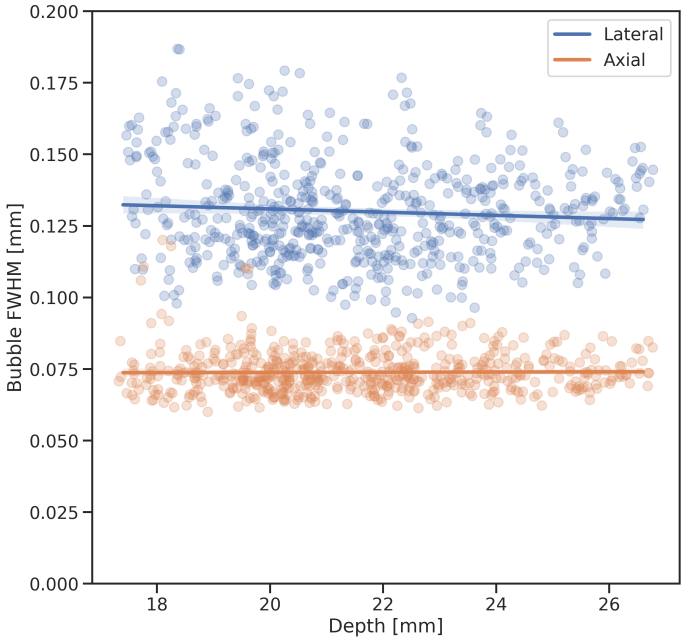

Figure 3.5: Resolution measurements of dual-frequency transducer in SpHI mode. The original resolution of the superharmonic imaging device was empirically determined by measuring the point spread function repeatedly within a region of interest over n = 400 independent images of bubbles floating in a water tank. The mean axial and lateral FWHM values were 73 and 130 μm, respectively.

the raw counts to densities. An ordinary least squares linear model was fit to the data using the *statsmodels* Python package v0.13.

3.2.8 Vessel centerline extraction and analysis

Vessel centerlines were retrieved from each 2-D ULM image using Aylward and Bullitt's algorithm [31]. Vessels were processed in 2-D because of the large elevational step size between each slice. The distance metric (DM) is equal to the total path length of each centerline divided by the straight-line distance between the vessel endpoints. For a centerline composed of n points with coordinates p_x,

$$\mathrm{DM} = \frac{\sum_{x=1}^{n-1} ||p_x - p_{x+1}||}{||p_1 - p_n||}. \tag{3.1}$$

The sum-of-angles metric (SOAM) is equal to the summation of angles between consecutive triplets of points, divided by the total path length of the centerline. For each triplet of points

composing two vectors, v_x and v_{x+1},

$$\text{SOAM} = \frac{\sum_{x=1}^{n-2} \cos^{-1}\left(\frac{v_x}{||v_x||} \cdot \frac{v_{x+1}}{||v_{x+1}||}\right)}{\sum_{x=1}^{n-1} ||p_x - p_{x+1}||}. \tag{3.2}$$

For each vessel in each image, the distance to the nearest molecular localization was computed using the *dsearchn* function (MATLAB, MathWorks, Inc.). Fig. 3.8 shows an example of the pairing algorithm. Outliers were determined using the *isoutlier* function in MATLAB on the full set of DM and SOAM values, and a vessel was removed if it was an outlier for either tortuosity metric.

3.2.9 Measuring the resolution of ULM images with Fourier Ring Correlation

Fourier ring correlation (FRC) is a robust, automatic method for measuring the resolution of images, such as those captured with super-resolution microscopy [32]. FRC measures the resolution of an imaging system by quantifying the agreement in the frequency domain between two images of the same scene with independent noise realizations. This is accomplished by computing the normalized cross correlation between sets of concentric rings (or shells for three-dimensional images) in Fourier space to determine the spatial frequency beyond which true structures and noise are indiscernible. The correlation between the ith frequency bins from the two images is given by

$$\text{FRC}(r_i) = \frac{\sum_{r \in r_i} F_1(r) \cdot F_2(r)^*}{\sqrt{\sum_{r \in r_i} F_1^2(r) \cdot \sum_{r \in r_i} F_2^2(r)}}, \tag{3.3}$$

where r_i is the ith frequency bin, and F_1 and F_2 are the Fourier transforms of the two images. To measure the resolution of our ULM imaging system, we performed the one-image FRC approach described by Koho and colleagues [33], utilizing the open-source repository linked to said publication. The one-image method splits a single image into four independent images, and the average of two FRC measurements is reported. The numeric resolution corresponds to the spatial frequency at which the correlation curve drops below $\frac{1}{7}$ [33]. Our measurement is performed on the background localization microscopy image slice from the center of the tumor (*i.e.*, containing blood vessels) shown in Fig. 3.6h, since the molecular localization maps are sparse, binary

images. We assume that since we are using the same imaging platform, beamforming, and localization algorithms for the background and molecular data, the resolutions of the two approaches should be equal.

3.2.10 Estimating the degree of vessel reconstruction

The degree of vessel reconstruction in ultrasound localization microscopy images was estimated using the methods presented by Dencks et al., who model the accumulation of localizations as a zero-inflated Poisson process [34]. For each of the three tumors imaged during the study, we drew regions of interest (ROIs) around the tumor boundary for each slice of the volume. The degree of reconstruction (DOR) was estimated for each ROI using the equation

$$\mathrm{DOR} = 1 - e^{-\hat{\Lambda}}, \tag{3.4}$$

where $\hat{\Lambda}$ is given by

$$\hat{\Lambda} = W_0\left(\frac{-T_2}{T_1} e^{-\frac{T_2}{T_1}}\right) + \frac{T_2}{T_1}. \tag{3.5}$$

T_1 is equal to the number of non-zero pixels within the ROI, T_2 is equal to the sum of the counts of the non-zero pixels within the ROI, and $W_0(x)$ is the main branch of Lambert's W-function (MATLAB function *lambertw*).

3.3 Results

3.3.1 In vitro studies

Validation of molecular ULM algorithm parameters

We first demonstrated the feasibility of super-resolution USMI using concurrent optical and ultrasonic imaging of control and targeted (biotinylated) MCAs in an avidin-coated microflow phantom. Optical microscopy revealed that targeted microbubbles adhered to the vessel wall and were retained during flow after a few initial loose bubbles dislodged, whereas control bubbles exhibited no retention (Fig. 3.2). We also demonstrated that the molecular localization algorithm described earlier (Section II-F) was not sensitive to free-flowing control contrast agents but detected the biotin-targeted bubbles, which were bound to the walls of the tube (Fig. 3.3).

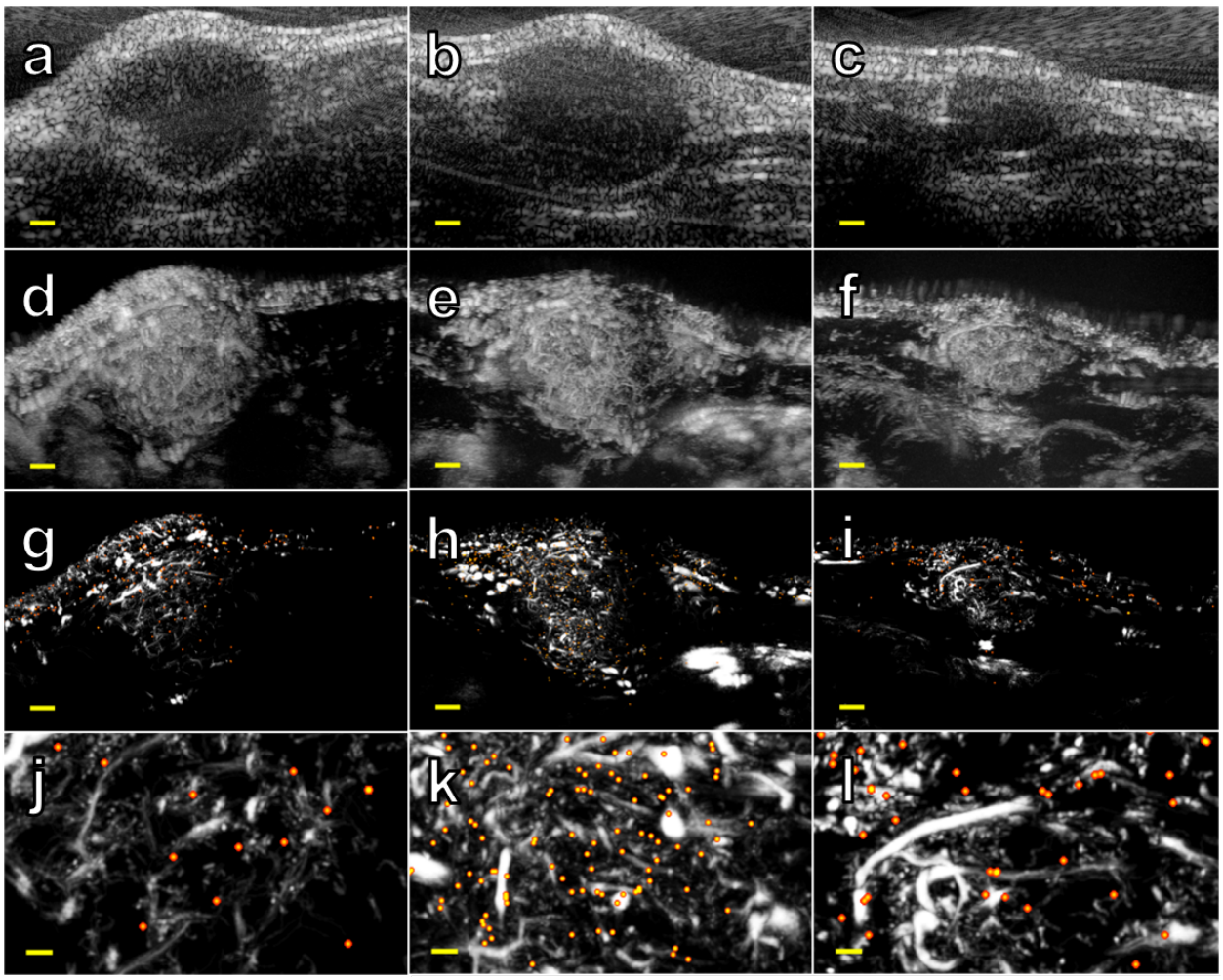

Figure 3.6: Super-resolution ultrasound molecular imaging in three rodent fibrosarcomas. (a-c) B-mode (standard ultrasound) images from the center of each tumor captured with the high-frequency elements of the dual-frequency array (dynamic range = 40 dB). (d-f) Maximum intensity projections generated from dual-frequency superharmonic images acquired across the tumor volumes (dynamic range = 30 dB). (g-i) Maximum intensity projections created with superharmonic ultrasound localization microscopy (gray colormap) with super-resolved molecular signaling overlaid (warm colormap, **localizations are blurred to improve visibility, true size is smaller**). Scale bars are 1 mm. (j-l) 2 mm $\times$ 3.5 mm selections from each ULM image showing microvascular and biomarker detail. Scale bars are 250 μm. Images in the same column are the same tumor.

Optical - Ultrasound Calibration

By counting microbubbles in the ultrasound and optical videos, we measured the relationship between acoustic and optical bubble densities. Fitting an ordinary least squares linear model to the data yields $R^2 = 0.964$, suggesting a strongly linear relationship between the output of the ultrasound imaging and the true local density of microbubbles (Fig. 3.4). We tested higher con-

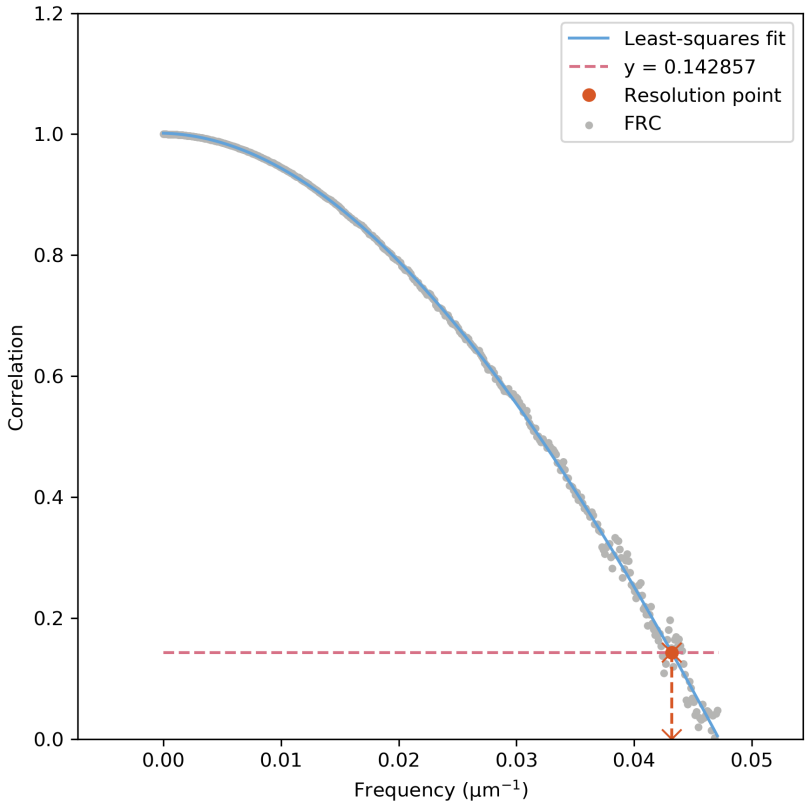

Figure 3.7: Fourier ring correlation plot for ULM image of rat fibrosarcoma vasculature. Single-image FRC plot corresponding to the centermost microvascular image from the center of the tumor in Fig. 3.6b. This plot was created using the open-source library linked to the publication by Koho and colleagues [33].

centrations of microbubbles, but we limit our analysis to the range of concentrations for which the optical counting method is robust (*i.e.*, not too many overlapping bubbles for the Hough transform). It is important to note that because the dual-frequency ultrasound system is only sensitive to a small fraction of the polydisperse bubble population, it is possible to image higher concentrations with ultrasound (up to approximately 1e8 MCA/mL) while maintaining the necessary sparsity for accurate localizations.

Resolution measurements

The superharmonic point spread function was repeatedly measured *in vitro* with the same mechanical index of 0.24 used elsewhere in the study. The mean axial and lateral FWHM values were 73 and 130 μm, respectively (Fig. 3.5). As expected, given the constant F# used during beamforming, there was no change in the average FWHM vs. axial depth.

3.3.2 In vivo imaging

Tumor images and resolution measurement

B-mode images of the rodent fibrosarcoma tumors revealed diameters ranging between 4 and 9 mm (Fig. 3.6a-c). Maximum intensity projections of the superharmonic data confirmed that each tumor was well-vascularized, but the diffraction-limited resolution of the scanner limited the separability of individual microvessels (Fig. 3.6d-f). By overlaying the output of the molecular localization algorithm on the super-resolution image of the same tumor, it was possible to visualize microvessels and biomarker expression at a scale beyond the diffraction limit (Fig. 3.6g-l), with a resolution of 23 μm as measured by Fourier ring correlation (Fig. 3.7). This result is a fivefold improvement in resolution compared to the previously described molecular acoustic angiography [35], and a threefold improvement compared to the axial resolution of the ultrasound system in SpHI mode (Fig. 3.5).

Vessel metrics

Segmentation of vessel centerlines [31] from ULM images allowed for quantification of tortuosity using the sum-of-angles (SOAM) and distance metrics (DM), both of which are elevated by malignant angiogenesis [36]. The combined distribution of DM and SOAM values for the vessels from the three tumors in Fig. 3.6 (n = 698 after removal of outliers) is shown in Fig. 3.9a. From these data, we calculated μ_{SOAM} = 54.9 ± 18.2 and μ_{DM} = 1.2 ± 0.2, both of which are higher than previously reported values for healthy tissues [36]. As expected, a histogram of the distances between segmented blood vessels and their nearest molecular localizations showed that the number of blood vessels decreased as the distance from molecular signaling increased. In other words, the data suggested that the local vessel density was increased in the vicinity of $\alpha_v\beta_3$ integrin. From these data, we computed $\mu_{distance}$ = 529.7 ± 472.8 μm (Fig. 3.9b). We also examined the relationship between vessel tortuosity and distance to the nearest biomarker signal and found that 58.1% and 56.1% of vessels with DM and SOAM values greater than the respective means were located within 500 μm of the nearest targeted MCA, which suggested that tortuosity was increased in the vicinity of $\alpha_v\beta_3$ expression. Finally, the average estimated degrees of vessel

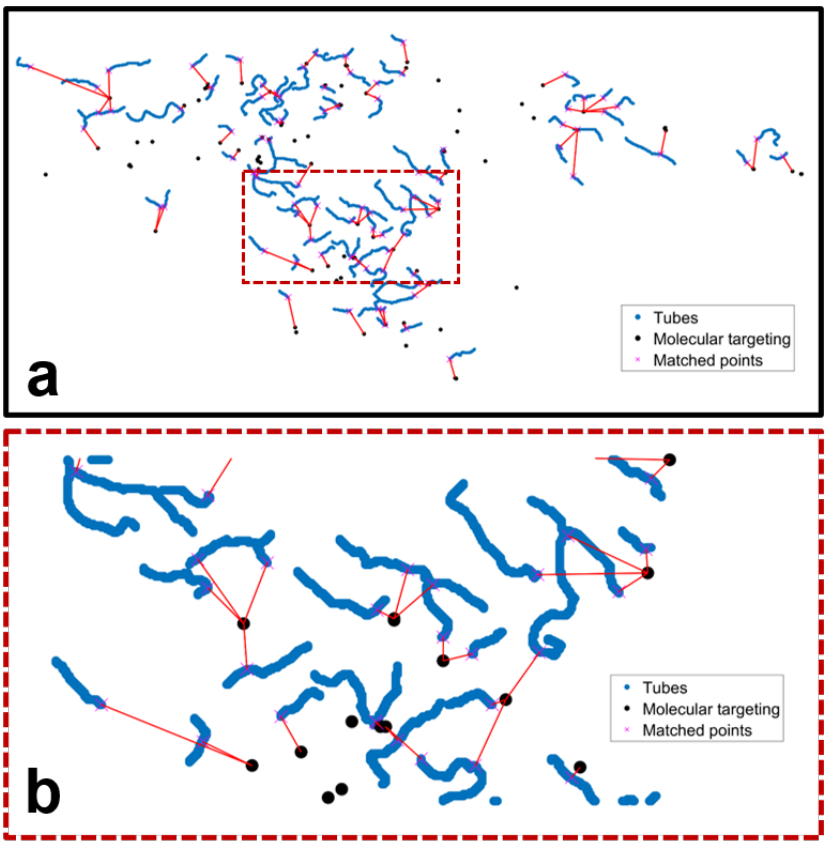

Figure 3.8: Tube-pairing algorithm using in vivo data. Vessel centerlines (shown in blue) are segmented using Aylward and Bullitt's algorithm [31]. Each tube is considered as a set of points, which is compared against the set of molecular localizations (shown in black) to find the distance between that tube and the nearest molecular signal. Shortest path lengths are rendered as red lines. a, Selected slice from center of a tumor. b, Inset from panel A indicated by red dashed line.

reconstruction for each tumor from Fig. 3.6 in order from left to right were 0.69, 0.51, and 0.59 (Fig. 3.10). These values are reasonable given the image acquisition time scale and reconstruction parameters [37].

3.4 Discussion

In this study, we demonstrated a substantial advance in USMI using high-frame-rate SpHI and MCAs targeted to angiogenic biomarkers. Concurrent optical and ultrasonic imaging in a targeted flow phantom demonstrated that this technique is sensitive to stationary bubbles and rejects flowing contrast agents. It is likely that the aggressive tracking thresholds that were used to identify bound vs. moving bubbles excluded some targeted bubbles (*e.g.*, a targeted bubble that deflates in fewer than 12 repeated pulses). However, our thresholds were tuned to minimize false

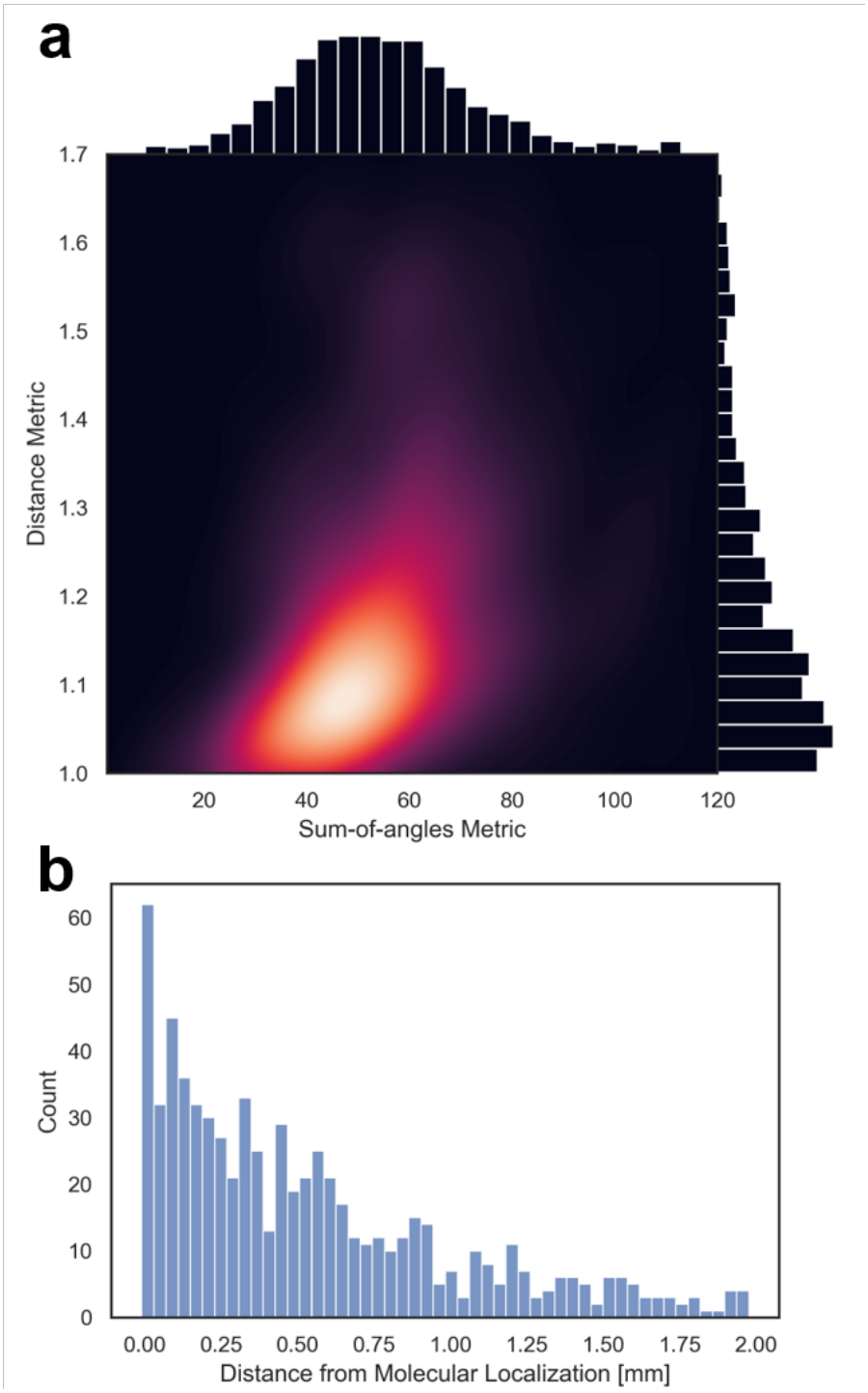

Figure 3.9: Vessel segmentation from tumor images allows for quantification of tortuosity metrics. (a) Kernel density plot of distance metric and sum-of-angles metric for all segmented vessels after removal of outliers ($n = 698$). (b) Histogram of distances between each segmented vessel and the nearest molecular localization.

positives since there are normally some residual circulating bubbles when performing ultrasound molecular imaging *in vivo*.

Furthermore, we demonstrated excellent correlation between detected ultrasound events and the ground truth microbubble density, which suggests that this method can provide superior quantification of adherent MCAs compared to standard USMI, where there are typically numerous microbubbles within a single resolution cell. As with any implementation of ultrasound localization microscopy, it is difficult to guarantee a single bubble per resolution cell at any given point in time since the bubble flow is stochastic. However, based on the linearity of the calibration results, along with our careful control of the concentrations of contrast media for imaging studies, we are confident that the majority of our localizations correspond to single bubbles. As mentioned in the previous section, it is vital to note that the large difference in observed MCA densities between the optical and ultrasound platforms is to be expected. This phenomenon is due to the fact that the ultrasound contrast agent that we manufacture in-house is polydisperse, with diameters ranging between approximately 0.5 and 3 μm. Because the center frequency of the receiving transducer is greater than 10x the transmit frequency, the dual-frequency device is only sensitive to the bubbles which are driven at or quite near their resonant frequency. Bubbles driven off-resonance tend not to produce the wideband echoes necessary for SpHI. In the future, we may explore methods to produce more monodisperse contrast agents so that the imaging device can detect a larger percentage of the injected bubbles.

By combining molecular ULM with a motion correction scheme based on interleaved B-modes, we also imaged subcutaneous tumors in freely breathing rodents to create co-registered maps of microvasculature and biomarker expression. In the present work, we measured a resolution of 23 μm, an almost 5$\times$ improvement in resolution over molecular acoustic angiography [35]. Additionally, we demonstrated that it is now possible to perform image analysis an unprecedented level of granularity, being able to describe the spatial relationships between individual capillaries and points of molecular signalling. While the scope of the present study is limited to the description and validation of the proposed imaging method, we hope to explore in future

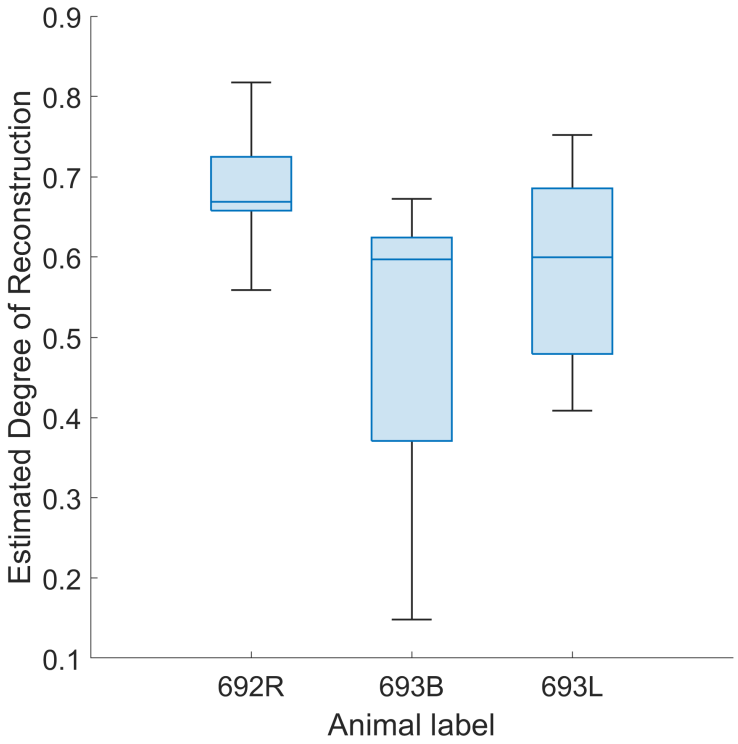

Figure 3.10: Estimated degrees of vessel reconstruction (DOR) for each tumor. Box charts show the distributions of DOR values calculated from individual slices of the tumor volumes, grouped by animal. The order of the charts corresponds to the ordering of the tumors in the first row of Fig. 3.6 (*e.g.*, 692R is the rat in Fig. 3.6a). The mean DOR values in order for each tumor are 0.69, 0.51, and 0.59.

studies how these new data might be used to increase the sensitivity to therapeutic efficacy.

The results of this proof-of-concept study will be further augmented by the advent of ultra-wide bandwidth matrix arrays for volumetric SpHI, which will mitigate the limitations of the current hardware. The transducer used for this study was based on a linear array, hence it was not well-suited for volumetric imaging. To interrogate each tumor, the transducer was mechanically swept in the elevation dimension to acquire a series of ULM images. There is an inherent tradeoff between the length of a ULM acquisition and the integrity of the reconstructed image [34], [37]. Therefore, some small vessels were sacrificed to ensure that the full volumes were imaged within a reasonable timeframe (estimated degrees of vessel reconstruction for the tumors in this study ranged between 0.51 and 0.69). Furthermore, the elevational step size (1 mm) and transmit-receive beamwidth (0.5 mm) were large compared to the in-plane pixel size (10 μm).

This disparity surely affected the measurements of vessel tortuosity, as they were effectively performed on 2-D projections of complex 3-D structures.

Also, the speckle pattern of the high-frequency linear array decorrelates quickly with elevational displacements (*e.g.*, from respiratory motion), which reduced the accuracy of in-plane motion tracking and final image registration. A matrix array designed for SpHI would allow for isotropic voxels and robust elevational tracking for simplified image registration. In addition, volumetric imaging would significantly improve the degree of vessel reconstruction by allowing the entire tumor to be interrogated over the course of the imaging session, thus reducing the number of molecular localizations located outside of reconstructed vessels. We believe that the molecular localizations that are near, but not within, a vessel would be matched with a vessel if it were possible to completely eliminate the errors introduced by physiological motion over the course of the long acquisition time. We also hypothesize that the majority of the molecular localizations associated with larger distances are the result of the non-unity DOR values.

Despite the aforementioned limitations, the results presented herein constitute a compelling argument for the feasibility of super-resolution USMI. The aim of the present study is not to draw any biological conclusions based on the results, but rather to demonstrate the capabilities of superharmonic imaging for imaging stationary contrast agents *in vivo* so that they may be localized and processed. We believe that this new imaging method has the potential to significantly augment the capabilities of ULM and molecular imaging for quantifying features of the tumor microenvironment in preclinical and clinical studies alike. As mentioned earlier, we hope to compare the proposed technique with traditional methods for USMI for applications such as monitoring response to tumor treatment.

3.5 Conclusion

We present a proof-of-concept study for super-resolution ultrasound molecular imaging with superharmonic imaging. *In vitro*, we demonstrated that it is possible to detect and localize stationary ultrasound contrast agents using superharmonic imaging. Our relatively simple algorithm based on bubble tracking can be tuned such that it is sensitive to bound contrast agents while

rejecting freely-flowing bubbles. The method can be extended with a previously described motion correction scheme to produce co-registered maps of the microvasculature and molecular signalling *in vivo* with a resolution of 23 μm. We also propose a variety of approaches to quantification which are currently not possible with diffraction-limited imaging techniques. In future longitudinal cancer treatment studies, we hope to compare the proposed method to traditional USMI.

REFERENCES

[1] "Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011.

[2] R. K. Jain and P. Carmeliet, "SnapShot: Tumor angiogenesis.," *Cell*, vol. 149, no. 6, 1408–1408.e1, 2012.

[3] R. M. Sigrist, J. Liau, A. E. Kaffas, M. C. Chammas, and J. K. Willmann, "Ultrasound elastography: Review of techniques and clinical applications," *Theranostics*, vol. 7, no. 5, pp. 1303–1329, 2017.

[4] A. A. Oglat, M. Z. Matjafri, N. Suardi, M. A. Oqlat, M. A. Abdelrahman, and A. A. Oqlat, "A review of medical doppler ultrasonography of blood flow in general and especially in common carotid artery," *Journal of Medical Ultrasound*, vol. 26, no. 1, pp. 3–13, 2018.

[5] M. X. Tang, H. Mulvana, T. Gauthier, A. K. Lim, D. O. Cosgrove, R. J. Eckersley, and E. Stride, "Quantitative contrast-enhanced ultrasound imaging: A review of sources of variability," *Interface Focus*, vol. 1, no. 4, pp. 520–539, 2011.

[6] D. Y. Huang and P. S. Sidhu, "Focal testicular lesions: Colour Doppler ultrasound, contrastenhanced ultrasound and tissue elastography as adjuvants to the diagnosis," *British Journal of Radiology*, vol. 85, no. SPEC. ISSUE 1, pp. 41–53, 2012.

[7] R. Gessner and P. A. Dayton, "Advances in molecular imaging with ultrasound," *Molecular imaging*, vol. 9, no. 3, pp. 7290–2010, 2010.

[8] S. V. Bachawal, K. C. Jensen, A. M. Lutz, S. S. Gambhir, F. Tranquart, L. Tian, and J. K. Willmann, "Earlier detection of breast cancer with ultrasound molecular imaging in a transgenic mouse model," *Cancer Research*, vol. 73, no. 6, pp. 1689–1698, 2013.

[9] J. K. Willmann, L. Bonomo, A. C. Testa, P. Rinaldi, G. Rindi, K. S. Valluru, G. Petrone, M. Martini, A. M. Lutz, and S. S. Gambhir, "Ultrasound molecular imaging with BR55 in patients with breast & ovarian lesions: First-in-human results," *Journal of Clinical Oncology*, vol. 35, no. 19, pp. 2133–2140, 2017.

[10] N. Deshpande, A. M. Lutz, Y. Ren, K. Foygel, L. Tian, M. Schneider, R. Pai, P. J. Pasricha, and J. K. Willmann, "Quantifi cation and Monitoring of Infl ammation in Murine Infl ammatory Bowel Disease with Targeted Contrast-enhanced US 1 From the Molecular Imaging Program at Stanford, Department of Radiology (N," *Radiology.rsna.org n Radiology*, vol. 262, no. 1, 2012.

[11] J. D. Rojas, F. Lin, Y. C. Chiang, A. Chytil, D. C. Chong, V. L. Bautch, W. Kimryn Rathmell, and P. A. Dayton, "Ultrasound Molecular Imaging of VEGFR-2 in Clear-Cell Renal Cell Carcinoma Tracks Disease Response to Antiangiogenic and Notch-Inhibition Therapy," *Theranostics*, vol. 8, no. 1, pp. 141–155, 2018.

[12] J. Zhou, H. Wang, H. Zhang, A. M. Lutz, L. Tian, D. Hristov, and J. K. Willmann, "Vegfr2-Targeted Three-Dimensional Ultrasound Imaging Can Predict Responses to Antiangiogenic Therapy in Preclinical Models of Colon Cancer," *Cancer Research*, vol. 76, no. 14, pp. 4081–4089, 2016.

[13] F. Zhao, S. Unnikrishnan, E. B. Herbst, A. L. Klibanov, F. W. Mauldin Jr, and J. A. Hossack, "A targeted molecular localization imaging method applied to tumor microvasculature," *Investigative Radiology*, vol. 56, no. 4, pp. 197–206, 2021.

[14] M. A. Pysz, I. Guracar, L. Tian, and J. K. Willmann, "Fast microbubble dwell-time based ultrasonic molecular imaging approach for quantification and monitoring of angiogenesis in cancer," *Quantitative imaging in medicine and surgery*, vol. 2, no. 2, p. 68, 2012.

[15] D. Hyun, L. Abou-Elkacem, R. Bam, L. L. Brickson, C. D. Herickhoff, and J. J. Dahl, "Nondestructive detection of targeted microbubbles using dual-mode data and deep learning for real-time ultrasound molecular imaging," *IEEE Transactions on Medical Imaging*, vol. 39, no. 10, pp. 3079–3088, 2020.

[16] J. R. Lindner, J. Song, F. Xu, A. L. Klibanov, K. Singbartl, K. Ley, and S. Kaul, "Noninvasive ultrasound imaging of inflammation using microbubbles targeted to activated leukocytes," *Circulation*, vol. 102, no. 22, pp. 2745–2750, 2000.

[17] L. Abou-Elkacem, S. V. Bachawal, and J. K. Willmann, "Ultrasound molecular imaging: Moving toward clinical translation," *European journal of radiology*, vol. 84, no. 9, pp. 1685–1693, 2015.

[18] E. B. Herbst, S. Unnikrishnan, A. L. Klibanov, F. W. Mauldin Jr, and J. A. Hossack, "Validation of normalized singular spectrum area as a classifier for molecularly targeted microbubble adherence," *Ultrasound in medicine & biology*, vol. 45, no. 9, pp. 2493–2501, 2019.

[19] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.

[20] S. T. Hess, T. P. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophysical Journal*, vol. 91, no. 11, pp. 4258–4272, 2006.

[21] S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: Stimulated-emission-depletion fluorescence microscopy," *Optics letters*, vol. 19, no. 11, pp. 780–782, 1994.

[22] T. A. Klar, S. Jakobs, M. Dyba, A. Egner, and S. W. Hell, "Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission," *Proceedings of the National Academy of Sciences*, vol. 97, no. 15, pp. 8206–8210, 2000.

[23] C. Errico, J. Pierre, S. Pezet, Y. Desailly, Z. Lenkei, O. Couture, and M. Tanter, "Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging," *Nature*, vol. 527, no. 7579, pp. 499–502, 2015.

[24] O. M. Viessmann, R. J. Eckersley, K. Christensen-Jeffries, M. X. Tang, and C. Dunsby, "Acoustic super-resolution with ultrasound and microbubbles," *Physics in Medicine and Biology*, vol. 58, no. 18, pp. 6447–6458, 2013.

[25] K. Christensen-Jeffries, R. J. Browning, M. X. Tang, C. Dunsby, and R. J. Eckersley, "In vivo acoustic super-resolution and super-resolved velocity mapping using microbubbles," *IEEE Transactions on Medical Imaging*, vol. 34, no. 2, pp. 433–440, 2015.

[26] "In Vitro Superharmonic Contrast Imaging Using a Hybrid Dual-Frequency Probe," *Ultrasound in Medicine and Biology*, vol. 45, no. 9, pp. 2525–2539, 2019.

[27] T. M. Kierski, D. Espindola, I. G. Newsome, E. Cherin, J. Yin, F. S. Foster, C. E. Demore, G. F. Pinton, and P. A. Dayton, "Superharmonic Ultrasound for Motion-Independent Localization Microscopy: Applications to Microvascular Imaging from Low to High Flow Rates," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 5, pp. 957–967, 2020.

[28] I. G. Newsome, T. M. Kierski, G. Pang, J. Yin, J. Yang, E. Cherin, F. S. Foster, C. Carnevale, C. E. M. Démoré, and P. A. Dayton, "Implementation of a novel 288-element dual-frequency array for acoustic angiography: in vitro & in vivo characterization," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2021.

[29] H. Yuan, T. Schroeder, J. E. Bowsher, L. W. Hedlund, T. Wong, and M. W. Dewhirst, "Intertumoral differences in hypoxia selectivity of the pet imaging agent 64cu (ii)-diacetyl-bis (n4-methylthiosemicarbazone)," *Journal of Nuclear Medicine*, vol. 47, no. 6, pp. 989–998, 2006.

[30] J. E. Streeter, R. Gessner, I. Miles, and P. A. Dayton, "Improving sensitivity in ultrasound molecular imaging by tailoring contrast agent size distribution: In vivo studies," *Molecular imaging*, vol. 9, no. 2, pp. 7290–2010, 2010.

[31] S. R. Aylward and E. Bullitt, "Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction," *IEEE transactions on medical imaging*, vol. 21, no. 2, pp. 61–75, 2002.

[32] N. Banterle, K. H. Bui, E. A. Lemke, and M. Beck, "Fourier ring correlation as a resolution criterion for super-resolution microscopy," *Journal of structural biology*, vol. 183, no. 3, pp. 363–367, 2013.

[33] S. Koho, G. Tortarolo, M. Castello, T. Deguchi, A. Diaspro, and G. Vicidomini, "Fourier ring correlation simplifies image restoration in fluorescence microscopy," *Nature communications*, vol. 10, no. 1, pp. 1–9, 2019.

[34] S. Dencks, M. Piepenbrock, and G. Schmitz, "Assessing vessel reconstruction in ultrasound localization microscopy by maximum likelihood estimation of a zero-inflated poisson model," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 8, pp. 1603–1612, 2020.

[35] S. E. Shelton, B. D. Lindsey, J. K. Tsuruta, F. S. Foster, and P. A. Dayton, "Molecular Acoustic Angiography: A New Technique for High-Resolution Superharmonic Ultrasound Molecular Imaging," *International Journal of Biomedical Imaging*, vol. 42, no. 3, pp. 769–781, 2016.

[36] R. C. Gessner, S. R. Aylward, and P. A. Dayton, "Mapping microvasculature with acoustic angiography yields quantifiable differences between healthy and tumor-bearing tissue volumes in a rodent model," *Radiology*, vol. 264, no. 3, pp. 733–740, 2012.

[37] V. Hingot, C. Errico, B. Heiles, L. Rahal, M. Tanter, and O. Couture, "Microvascular flow dictates the compromise between spatial resolution and acquisition time in Ultrasound Localization Microscopy," *Scientific Reports*, vol. 9, no. 1, p. 2456, 2019.

CHAPTER 4: DEEP LEARNING METHODS FOR THE DETECTION AND LOCALIZATION OF MICROBUBBLE CONTRAST AGENTS IN HIGH-CLUTTER ENVIRONMENTS

4.1 Introduction and Background

Biomedical ultrasound is useful in a variety of healthcare settings because it is portable, relatively inexpensive, safe, and provides good spatial and temporal resolutions for soft tissue imaging. With the help of microbubble contrast agents (MCAs), it is also possible to image the blood pool with good sensitivity. While MCAs substantially enhance the echogenicity of blood, they do not change the physics of wave diffraction which govern the spatial resolution limit. That is to say, if a blood vessel is smaller than the point spread function (PSF) of the imaging system, it is not possible to resolve its structure using a normal imaging sequence even with the help of a microbubble contrast agent.

However, in recent years, the ultrasound community has adapted methods from optical microscopy to noninvasively image structures that are smaller than the diffraction limit *in vivo* by repeatedly localizing the positions of individual point targets onto a sub-resolution grid. In optical imaging, these targets are typically fluorescent molecules which are excited such that sparse subsets are stochastically activated within each image [1], [2]. Since each of these molecules is smaller than the wavelength of the imaging system, they appear in the images as PSFs. Assuming that the PSF produced by one molecule is not overlapped with a neighboring PSF, it is possible to precisely estimate the position of the source molecule using various techniques such as radial symmetry, Gaussian fitting, or computing the weighted centroid [3], [4].

The same principle serves as the foundation of ultrasound localization microscopy (ULM). Microbubble contrast agents are much smaller (1-8 μm) than the range of typical clinical wavelengths (0.1-1 mm), a difference of approximately 1-2 orders of magnitude. Hence, MCAs act

much like the fluorescent molecules described earlier, appearing in beamformed images as PSFs. There are some key differences between the optical and ultrasound approaches which should be considered. First, MCAs exhibit a relatively high spatiotemporal coherence except for cases where shell disruption occurs or the flow rate leads to aliasing (*e.g.*, in large vessels). This allows individual MCAs to be tracked as they move between images, improving the reconstruction and yielding estimations of local velocity at a high resolution. Second, the signals from the microbubbles must be separated from those originating from the surrounding tissue so that they can be localized. This can be accomplished during imaging using a nonlinear imaging scheme [5]–[9] or after imaging by filtering the data [10]–[12].

The precision of localization is driven by the signal-to-noise ratio (SNR) [13], [14]. In the context of ultrasound localization microscopy, SNR refers to the difference in amplitudes between the contrast agent signal and the background noise. After the tissue has been suppressed through the means described earlier, the remaining background noise is composed of thermal noise (*i.e.*, electronic noise in the imaging system) as well as residual tissue and clutter echoes which were not completely filtered. As SNR increases, the distribution of localization errors for a stationary contrast agent and independent realizations of noise becomes narrower, and vice versa. The resolution of the reconstructed image is largely determined by the width of this distribution. If the standard deviation of the distribution is much larger than structures in the imaging field, their apparent sizes will be larger. In cases where the SNR is poor, it can be difficult to identify individual contrast agents at all, since they will have a size, shape, and intensity similar to bright points in the background speckle.

There are numerous targets *in vivo* that are difficult to image with ULM because of limitations in SNR. Some examples include transcranial imaging of the brain, as well as imaging organs through the abdominal wall. Complex structures such as the skull and the abdominal wall are known to degrade SNR because of phenomena including attenuation, wave-front aberration, and reverberation clutter [15] (see Figure 4.1). Recently, McCall and colleagues demonstrated that these effects can significantly degrade the resolution of the reconstructed image by a signif-

icant degree, and the results also suggest that the final image might be distorted relative to the ground truth vessel structure [14].

Machine learning models have been successfully trained and validated for a wide variety of image processing tasks, including object detection, segmentation, classification, and upscaling [16]–[20]. Very recently, different groups of researchers have demonstrated the feasibility of detecting and localizing microbubble contrast agents for ULM using neural networks. Van Sloun *et al.* trained an encoder-decoder network to perform pixel-wise predictions of the presence of a microbubble from a single contrast-enhanced ultrasound (CEUS) image [21]. [22] trained a residual convolutional neural network (CNN) to perform the same task, using sub-pixel convolution layers to upscale their images rather than transposed convolutions [20]. Brown and colleagues developed a spatiotemporal framework using a 3-dimensional CNN to classify a time series of image patches that were identified using an *ad hoc* approach [23]. In [24], another spatiotemporal method was presented which predicted a single localization map for a sequence of CEUS images. Finally, Milecki *et al.* trained a 3-D encoder-decoder model to predict the full microbubble trajectories from a stack of consecutive CEUS images, with impressive *in vivo* results [25].

While these studies clearly demonstrate the feasibility of replacing part of the standard ULM imaging pipeline with machine learning models, they all suffer from the same limitation which is that their training data does not include the full description of background noise which was previously described. In many cases, the researchers use simple linear models to synthesize their data (*e.g.*, convolving a PSF with a map of scatterer locations) to produce an image. These models do not include sources of image degradation such as reverberation and aberration, so it is quite difficult to evaluate the performance of the final algorithms on *in vivo* data because there is no access to the ground truth vessel structure. This could prove especially problematic if quantitative metrics related to vessel structure were to be used for diagnostics.

In the present study, we aim to build on the results of the studies described earlier. To do so, we generate a new dataset of realistic microbubble contrast data using the Full-wave simulation tool which allows us to model nonlinear wave propagation in complex heterogeneous media

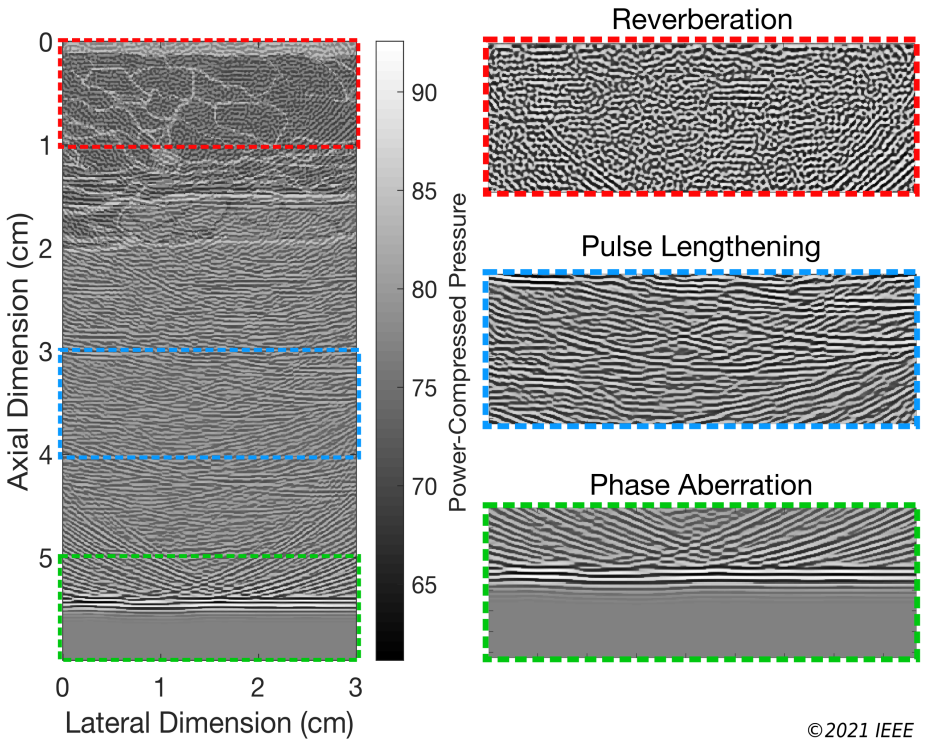

Figure 4.1: Sources of image degradation in ultrasound imaging in heterogeneous media. Reproduced from [14] with permission. In the near-field, sounds can become trapped between layers of different tissues, leading to a complex pressure field at the face of the transducer (*red*). In the far-field, low-amplitude echoes of the transmit escape the laminar structures in the near field and lengthen the effective transmit pulse (*blue*). Finally, variations in the speed of sound affect the uniformity of the transmit beam, leading to a degradation in the point spread function (*green*).

[26]. By coupling the simulation tool with cross-sectional cryosections from the Visible Human Project [27] and micro-CT images of a rodent skull, the resulting dataset includes the image-degrading phenomena described earlier, such as reverberation clutter and attenuation. *In silico*, we test the localization performance of different neural network architectures across a range of SNR values and perform a comparison against a previously published locaization algorithm. Finally, we demonstrate the generalizability of these algorithms to real scanner data using an *in vitro* microvessel flow phantom.

4.2 Methods

4.2.1 Simulation parameters

For this study, we characterized the L11-5 and the P4-1 transducers (Verasonics, Kirkland, WA, USA). Using a needle hydrophone (Onda Corporation, Sunnyvale, CA, USA) in a tank of

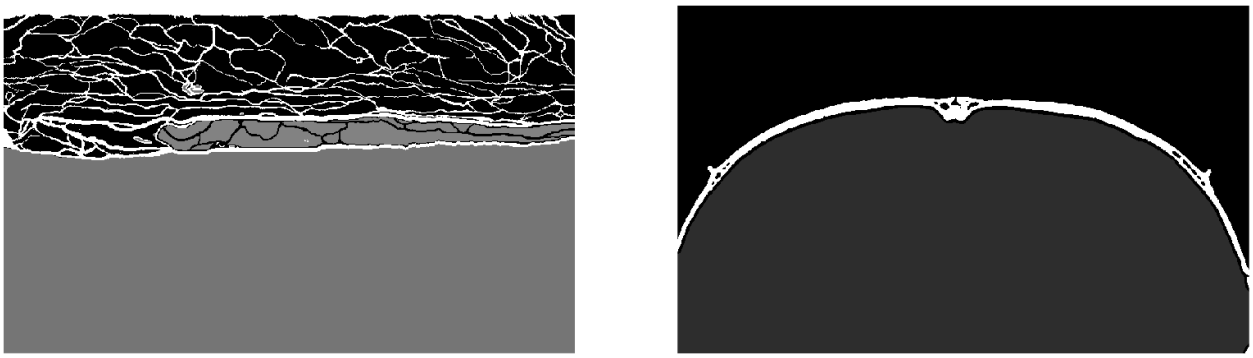

Figure 4.2: Representative abdominal and skull speed of sound maps for Fullwave simulations. Not shown are the corresponding maps of the other material properties listed in Table 4.1

degassed water, we recorded 2-cycle transmissions for both transducers. These waveforms were used as inputs to the simulations described in later sections. The center frequency was set to 2.5 MHz for the P4-1, and 7.8 MHz for the L11-5. No transmit apodization was applied during this calibration.

The simulation grid size was set to $\lambda/12$, where λ is equal to the wavelength assuming $c =$ 1540 m·s^{-1}. The pixel size was 51.3 μm for the P4-1 transducer and 16.5 μm for the L11-5. The tissue maps described in later sections were interpolated to match the pixel size of the given simulation, and these images were either padded or truncated in the axial dimension to achieve a simulation depth of 17 mm for the L11-5 and 50 mm for the P4-1. The images were cropped in the lateral dimension so that the width of the simulation was equal to 1.2$\times$ the width of the transducer aperture.

The time sampling was determined by setting the Courant number to 0.5 (this value was chosen to ensure numerical stability). The Courant number is given by

$$C = \frac{c_0 \cdot \Delta t}{\Delta x}, \tag{4.1}$$

where c_0 is the speed of sound in the medium, Δt is the sampling period, and Δx is the pixel size of the simulation. The resulting sampling frequency was 186.7 MHz for the 7.8 MHz L11-5 transducer and 60.0 MHz for the 2.5 MHz P4-1 probe. The length of each simulation in time was given by $t_{max} = 2.5 \cdot z_{max}/c_0$, where z_{max} was the maximum depth in the simulation field.

Tissue	B/A	c_0 (m·s^{-1})	ρ (kg·m^{-3})	α (dB·cm^{-1}·MHz^{-1})
Fat	9.6	1479	0.937	0.40
Connective	8.0	1613	1.120	0.68
Muscle	8.0	1550	1.070	0.15
Blood	5.0	1520	1000	0.01
Organ	7.6	1570	1.064	0.50

Table 4.1: Acoustic properties of tissues for Fullwave simulations

Channel data was recorded by saving the values of the pressure field at the transducer surface through time. Since the grid size of the simulation was finer than the pitch of the transducer elements, the recordings from groups of adjacent pixels were averaged such that the final channel count was equal to the number of elements for the transducer being modelled (128 channels for the L11-5 simulations, 96 for the P4-1). Channel data were beamformed onto the same $\lambda/12$ grid using delay-and-sum. For a transducer element located at the lateral position x_i, the total time of flight of the unsteered plane wave and its reflection from a pixel located at (x,z) is given by

$$\tau(x_i,x,z) = \frac{z+\sqrt{z^2+(x-x_i)^2}}{c_0}. \tag{4.2}$$

The size of the receive aperture was set according to $F\# = 2$, and no apodization was applied.

4.2.2 Generating reverberation clutter

To simulate realistic reverberation artifacts, we created Fullwave environments for abdominal wall imaging with the P4-1 as well as transcranial imaging with the L11-5. For the abdominal simulations, cross-sections of optical data from the Human Visible Project were manually segmented and labelled as either fat, connective tissue, muscle, blood, or organ tissue (Fig. 4.2). The values of speed of sound, density, nonlinearity, and attenuation for each of these is given in Table 4.1. For organ tissue, the acoustic properties of liver were used. A set of 2-D slices was generated by scanning through the volumetric abdomen data. The distance between slices was chosen to ensure that the speckle patterns from adjacent simulations were completely decorrelated.

To isolate the echoes due to reverberation, all of the tissue beneath the laminar structure of

the abdominal wall was assigned as homogeneous organ tissue (*i.e.*, no sub-resolution scatterers). By configuring the simulation in this manner, any sound wave that escaped the abdominal wall into the tissue below had no way of returning to the transducer because there were not any scatterers in its path. Any echoes recorded by the transducer could thus be attributed entirely to multiple reflections in the imaging of the transducer. The skull simulation environment was created by segmenting micro-CT images of a rat skull using a empirically determined intensity threshold (Fig. 4.2). The acoustic properties of the skull were set according to the results of [28]. Similar to the abdominal simulations, everything beneath the skull bone structure was assigned as homogeneous organ tissue in order to isolate the effects of reverberation.

4.2.3 Microbubble contratst agent simulations

For each simulation, a set of random microbubble coordinates was generated with a densities between 20 – 50 MCA·cm^{-2}. This was achieved by creating an array $M_i \in [0,1]^{Z\times X}$ of uniformly distributed noise, where X is the number of simulation grid points in the lateral dimension and Z is the number of axial grid points. The microbubble position array $\hat{M}_i \in \{0,1\}^{Z\times X}$ for the i^{th} sample was given by

$$\hat{M}_{i,j} = \begin{cases} 1 & \text{if} M_{i,j} \leq T \\ 0 & \text{if} M_{i,j} > T \end{cases}, \tag{4.3}$$

where T is the fraction of pixels that should contain a bubble in order to achieve the desired density of contrast agents.

Apart from the locations of the contrast agents, the material maps for the microbubble simulations consisted of homogeneous organ tissue. Contrast agents were added to the simulation as speed of sound scatterers by element-wise subtracting $0.05 \cdot c_{organ} \cdot (\hat{M}_i * G_1)$ from the $Z \times X$ speed of sound map, where G_1 is a normalized Gaussian kernel with a standard deviation equal to 1.

A random aberrator with a spatial frequency of 5 mm^{-1} and root mean square width between 0 and 100 ns was applied to each sample of microbubble channel data during beamforming to simulate the effects of varying degrees of aberration caused by speed of sound heterogeneity. In this way, it was possible to preserve the original microbubble datasets and modify the strength of

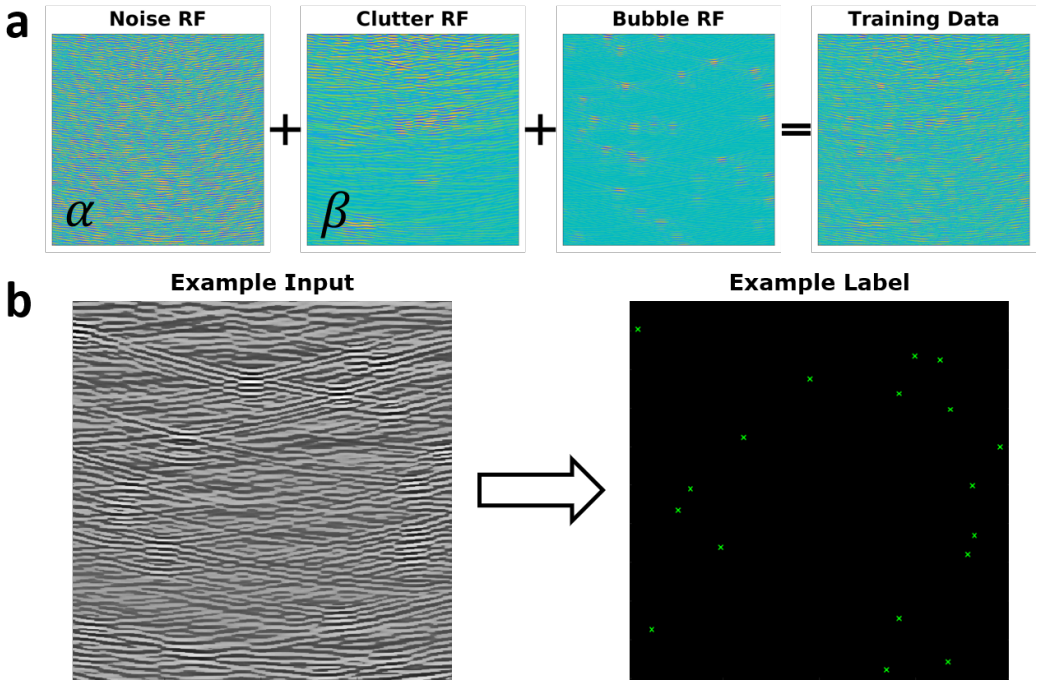

Figure 4.3: Synthetic data for training localization models. **(a)** One sample of training data is obtained by the summation of beamformed noise, clutter, and bubble data. The noise and clutter images are scaled by random factors α and β for each sample. The final image is normalized. **(b)** The label for each sample is a sparse binary matrix where a 1 indicates the position of a microbubble contrast agent.

the aberration effect as desired.

4.2.4 Synthetic channel noise

Channel noise was measured for each transdcuer by recording radiofrequency data from each device for 100 μs without transmitting. Data were recorded using a Vantage 256 ultrasound research platform (Verasonics, Kirkland, WA, USA). To generate synthetic channel noise, a random generator was parameterized according to the distributions of the measured data. This generator was then sampled to populate an array with the same shape as the channel data generated by the simulations described in previous sections. In this way, the properties of the synthetic noise matched those of the real data which is processed by the onboard bandpass filter prior to beamforming. An independent realization of noise was generated for each sample of training data.

4.2.5 Creating the synthetic training data

Images of reverberation clutter, microbubbles and channel noise, denoted $I_{R,i}$ and $I_{M,i}$, $I_{N,i}$ respectively, were beamformed separately and then normalized by the maximum of the absolute

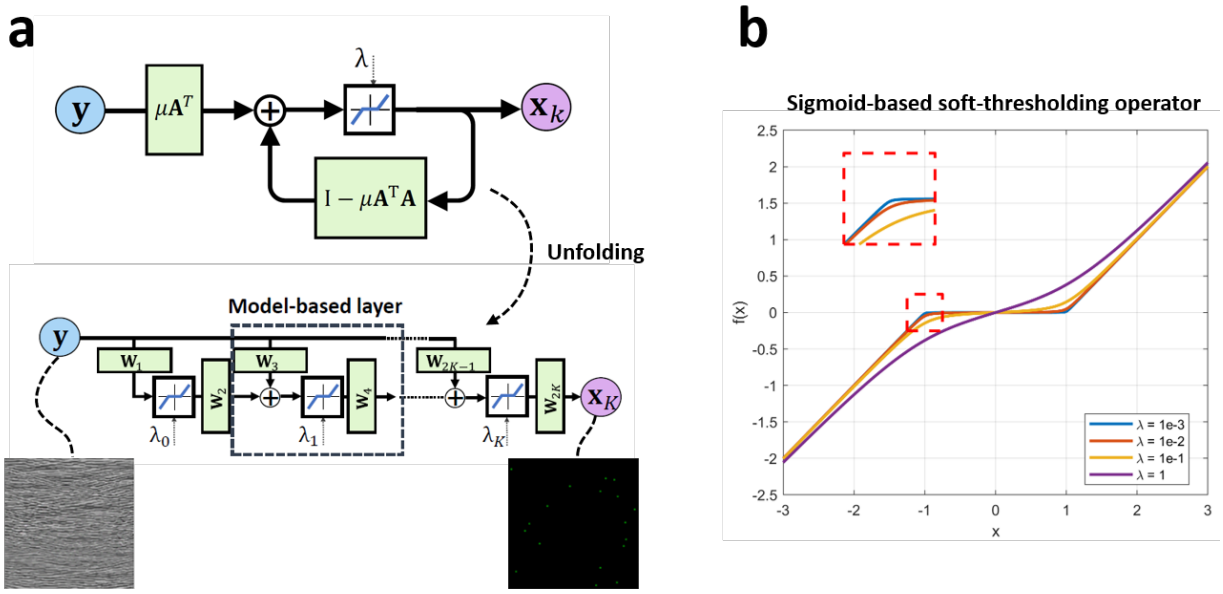

Figure 4.4: Schematic of the deep unfolded architecture used for ultrasound localization microscopy. **(a)** The original ISTA scheme is unfolded into a deep model by replacing the iterative approach with a series of learned convolutional layers. **(b)** Each activation function layer has a learned parameter λ which controls the smoothness of the soft-threshold operator.

value of the beamformed radiofrequency data. A sample of training data S_i is then given by $S_i = \alpha I_{N,i} + \beta I_{R,i} + I_{M,i}$, where $\alpha, \beta \in [0.125, 1]$ are randomly selected for each sample (Fig. 4.3). This produces images with a range of possible SNR values between approximately -6 and +12 dB. Each S_i is normalized, and its corresponding label for training the model is the binary array M_i from Section 4.2.3 which contains the ground truth positions of microbubble contrast agents.

4.2.6 Model implementation and training

In the present study, we utilized the deep unfolded architecture presented in [29] (Fig. 4.4a). In short, this model was designed for sparse recovery problems such as ULM and is based on an unfolded version of the iterative shrinkage-thresholding algorithm (ISTA) for the following inverse problem defined in [29]:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} ||\mathbf{y} - \mathbf{A}\mathbf{x}||_2^2 + \lambda ||\mathbf{x}||_1 \tag{4.4}$$

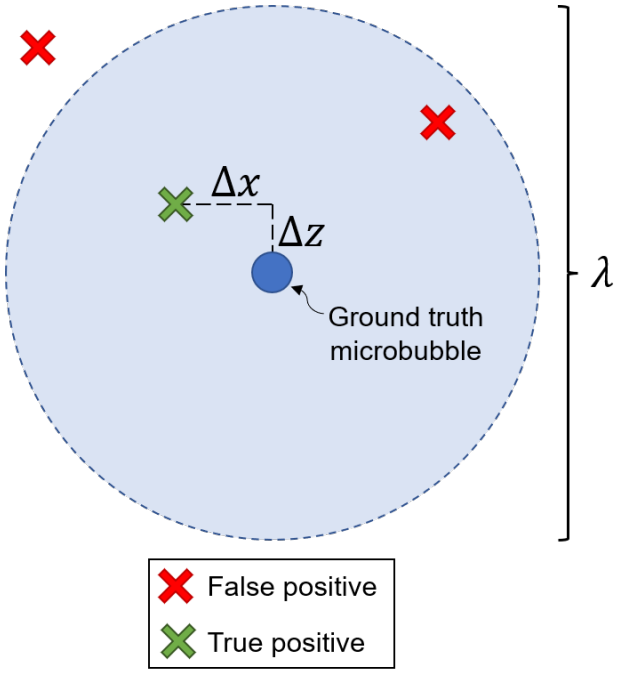

Figure 4.5: Matching algorithm for localization data. Consider a single ground truth microbubble with no other bubbles present nearby. Localizations from the image data are represented with crosses. The green cross is the closest localization within the search radius, and so it is paired with the ground truth microbubble. It has lateral and axial localization errors of Δx and Δz, respectively. The red cross within the search radius is considered a false positive because it is not the closest localization to the ground truth bubble (in this example, we assume there are no other bubbles nearby which it could be paired with). Finally, the red cross outside of the search radius is considered a false positive because it is not within $\lambda/2$ of any ground truth position.

where **y** is the measured CEUS image, **x** is a vector describing the locations of microbubble contrast agents, and **A** describes the PSF of the imaging system. Rather than using an iterative approach such as ISTA to estimate **x**, each iteration of the algorithm can be unfolded into separate layers of a deep neural network. Each model-based layer is composed of two convolutional layers along with a sigmoid-based soft-threshold operator [30] with its own learnable smoothness parameter (Fig. 4.4b). In this way, the parameters of the recovery algorithm are learned rather than manually tuned. We used an architecture with 40 folds and 7×7 kernels (single filter per layer), approximately 2,000 total parameters in size.

We trained two models – one on the abdominal dataset and one on the transcranial dataset. The quantitative analysis described in a later section focused on the results of the abdominal model alone since the total number of transcranial images was only 439. The abdominal images

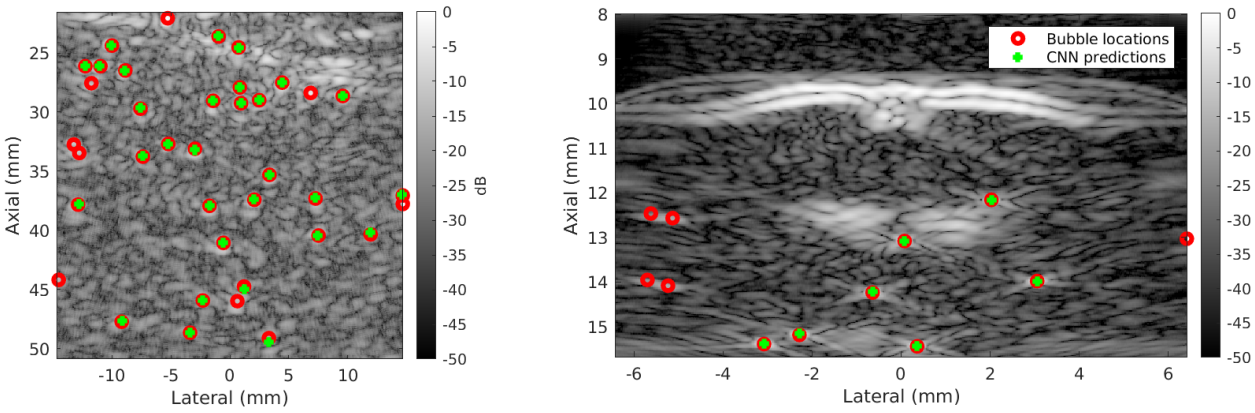

Figure 4.6: Predictions on test data from the deep unfolded network. A model trained on 2.5 MHz data performs inference on images with abdominal clutter (*left*), and a model trained on 7.8 MHz data localizes contrast agents underneath the skull (*right*). In both cases, the model does not localize the bright clutter artifacts.

were split into sets of 1,200, 200, and 100 images for training, validation, and testing, respectively. As a proof of concept, the second model was trained on 373 transcranial images with 65 samples held out for qualitative testing. The models were implemented using TensorFlow [31] and trained for 100 epochs by minimizing the pixel-wise focal loss [16] using the Adam optimizer [32] with a learning rate of $1e{-}3$ and a batch size of 16. Similar to [21], we convolved the ground truth label with a narrow Gaussian kernel ($\sigma = 1$) to stabilize training and improve convergence time. No data augmentation was applied in this study.

4.2.7 Model evaluation

After training the models, they were evaluated on test data. The final predictions were thresholded at 0.5 to remove any low-confidence scores. The local maxima in the thresholded prediction maps were considered to be localizations. For comparison, a standard localization technique based on a peak detection algorithm was also used on the test data [6].

The performances of the conventional localization algorithm and the machine learning model trained on abdominal data were evaluated by measuring the numbers of false positives and false negatives along with the axial and lateral localization errors for true positives. A prediction was considered to be a true positive if it was within a radius of $\lambda/2$ mm from a ground truth microbubble. If multiple predictions were present within the radius for a given ground truth bub-

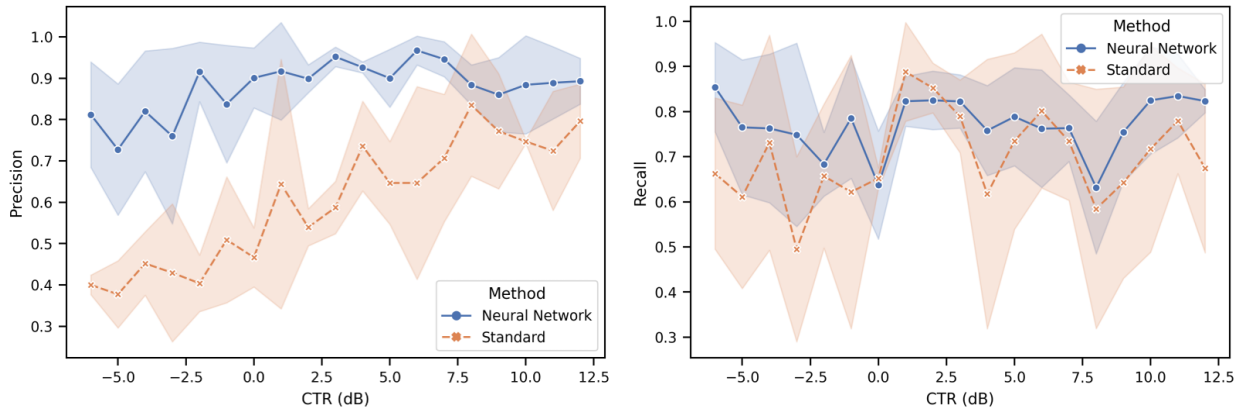

Figure 4.7: Precision (*left*) and recall (*right*) vs. CTR for both localization methods on abdomen test dataset. Error bars show the standard deviation of the data.

ble, then the closest predicted bubble was considered as its pair (Fig. 4.5). Once a true positive pairing was identified, the corresponding prediction and ground truth pixels were removed from the working set, and this process was repeated until no true positives remained. Any remaining predicted localizations were considered to be false positives, since they did not have a ground truth pairing. Conversely, the counts of remaining bubbles from each frame of the ground truth data were summed to measure the number of false negatives. These measurements were then used to calculate precision and recall for each frame of test data according to:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{4.5}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{4.6}$$

where TP, FP, and FN are the counts of true positives, false positives, and false negatives, resepctively.

We also evaluated the performance of the abdominal model on image data of an *in vitro* flow phantom. The phantom was composed of a 50 μm diameter FEP tube embedded in a graphite-gelatin mixture with an attenuation coefficient of approximately 0.5 dB·cm^{-1}·MHz^{-1}. A batch of 25,000 unsteered plane-wave images was collected at 500 frames/s with the P4-1 and the Vantage 256 scanner. These images were beamformed offline using the same parameters used for the simulations described in earlier sections. To isolate the bubbles from the speckle from the gelatin

phantom, we subtracted subsequent images of beamformed radiofrequency data. These differential images were then processed using the deep unfolded model and the standard localization approach to generate a ULM image.

4.2.8 Statistical analysis

All statistical tests were performed on the two sets of true positives from the conventional and neural network localization approaches (n = 2,230 and 2,518, respectively, from the 100 frames of abdomen test data). As mentioned previously, the width of the distribution of localization errors is a driving factor behind the resolution of ULM and similar methods [13], [14]. Hence, we compared the variances of the lateral and axial localization errors with the two-sample F-test (*vartest2* in MATLAB). Another factor that is important for image fidelity is the the bias of the error distributions. We used the two-sample t-test assuming unequal variances to test for significant differences between the means of the lateral and axial localization errors (*ttest2* in MATLAB). Finally, we compared the distributions of the magnitudes of the errors using the Mann-Whitney U-test (*ranksum* in MATLAB).

4.3 Results

After training for 100 epochs, the deep unfolded networks were able to detect and localize individual microbubble contrast agents in the challenging test data from the abdominal and transcranial simulations (Fig. 4.6). An analysis of precision and recall for different SNR values of abdominal clutter demonstrated that the neural network outperformed the standard localization technique, especially in terms of improved precision at SNR values less than 10 dB (Fig. 4.7).

There was no trend between SNR and the magnitudes of the axial or lateral localization error. However, this result was likely due to our definition of a true positive, with bubbles further than $\lambda/2$ not being considered. A comparison of the mean localization error magnitudes for the entire test set resulted in significant differences for both the lateral and axial dimensions ($p < 0.0001$ for both dimensions, Fig. 4.8). The mean axial error magnitudes were 107.2 and 46.1 μm for the standard and neural network methods, respectively, and the average lateral magnitudes were 57.0 and 24.0 μm.

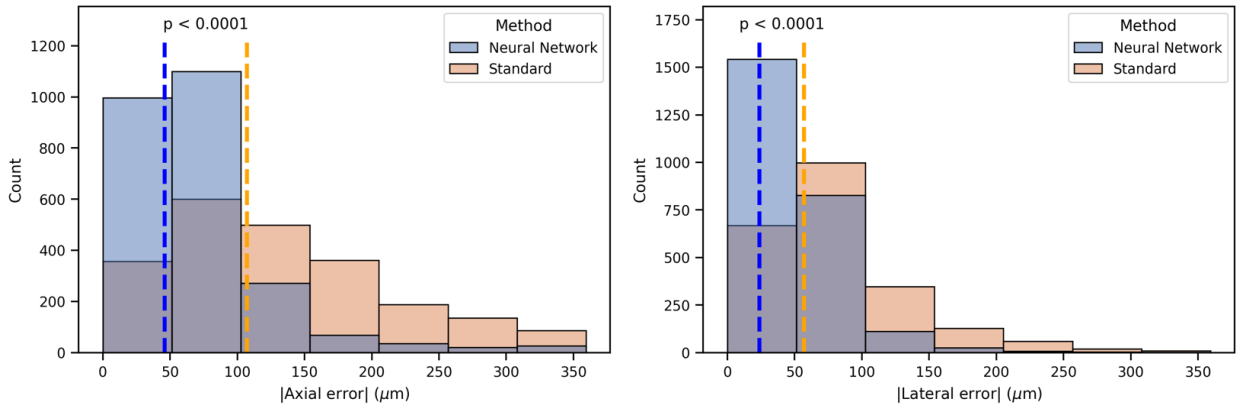

Figure 4.8: Distributions of the magnitudes of errors in the axial (*left*) and lateral (*right*) dimensions for both methods evaluated on P4-1 abdominal data.

Measurements of the standard deviations of the distributions of localization errors were 72.0 and 134.5 μm in the axial dimension for the machine learning model and standard method, respectively ($p < 0.0001$ for comparison of variances, Fig. 4.9). In the lateral dimension, the standard deviation of the unfolded network was 42.3 μm, compared to 73.9 μm for the conventional localization technique ($p < 0.0001$). The average values of the distributions shown in Fig. 4.9, which represent the measured biases of either localization method on these test data, were 2.6 and 19.4 μm in the axial dimension ($p < 0.0001$) for the machine learning and standard methods, respectively, and 2.6 and 29.9 μm in the lateral dimension ($p < 0.0001$).

Finally, we apply the model trained on simulations of P4-1 imaging to data acquired in a microflow phantom *in vitro* (Fig. 4.10). In this simple demonstration without additional clutter, the deep unfolded network yielded an image similar to that produced by the standard localization algorithm.

4.4 Discussion

In this study, we presented a new technique for synthesizing data composed of realistic reverberation clutter, channel noise, and point target echoes. We then used these data to train a lightweight, model-based network to detect and localize individual microbubble contrast agents across a range of clinically relevant signal-to-noise ratios. With a variety of qualitative and quantitative analyses, we demonstrated that the trained model outperformed the standard localization

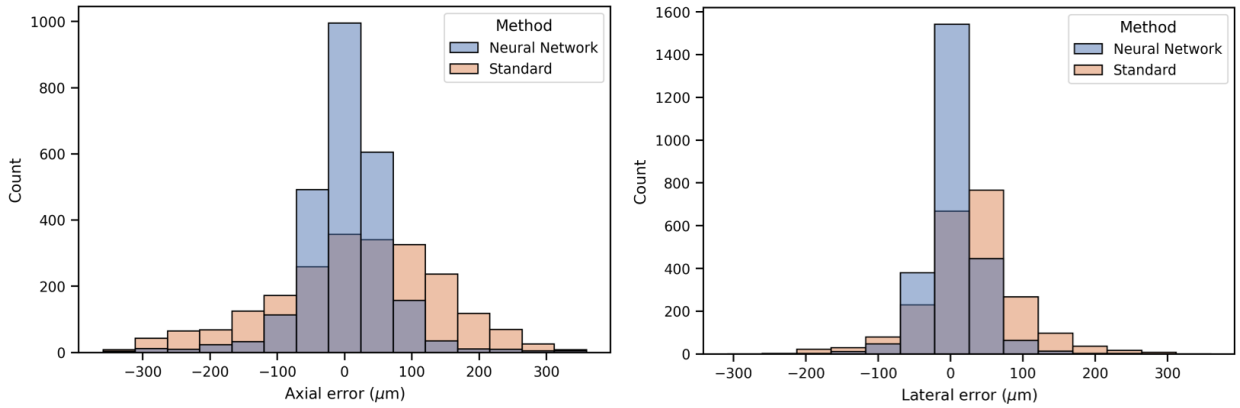

Figure 4.9: Distributions of the axial (*left*) and lateral (*right*) localization errors from the P4-1 test dataset.

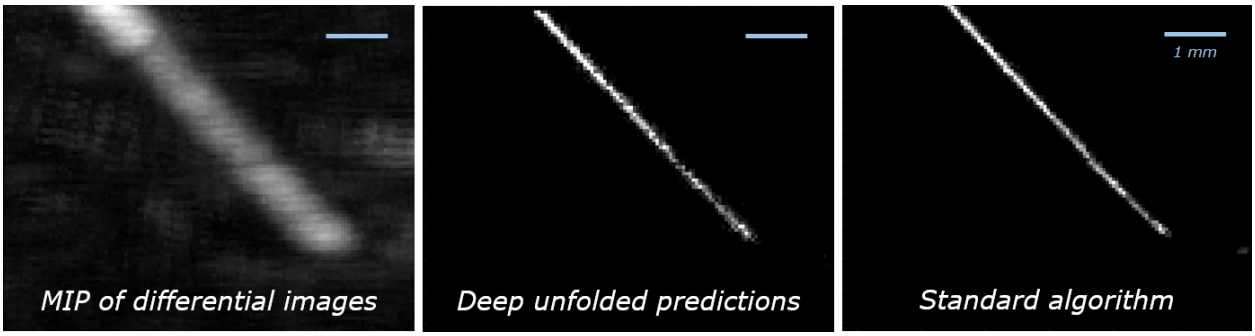

Figure 4.10: ULM images of an *in vitro* microflow phantom using the deep unfolded network and standard localization algorithm. (*Left*) A maximum intensity projection of the sequence of microbubble images demonstrates the original resolution of the imaging system. ULM images from the neural network (*Center*) and standard approach (*Right*) both yield substantial improvements in resolution. The blue scale bars are 1 mm in length.

approach used previously in [6] in terms of precision (the ratio of the number of true positives to the number total predictions), especially for SNR values less than 10 dB. While there was not a large difference between the two methods in terms of recall, the neural network was much more consistent than the standard approach, as is evidenced by the error bars in Fig. 4.7. Together, these results suggest that the unfolded model can improve image quality in low-SNR scenarios, since it is far less prone to producing false positives.

Using the P4-1 test dataset, we also quantified the increased performance of the deep learning approach with respect to localization error for correctly identified contrast agents. We measured significant improvements in the average magnitude of localization error in the axial (61.1 μm) and lateral (33.0 μm) dimensions for the machine learning model (Fig. 4.8). In addition, we

noted an improvement in the widths of the localization error distributions, which are closely related to the resolution of the ULM imaging system. The neural network improved the standard deviation of the residuals by 31.6 and 62.5 μm in the lateral and axial dimensions, respectively.

It is also important to consider the limitations of this study. First, the decoupling of the microbubble and clutter simulations was not ideal from the perspective of aberration modelling, since the echoes from the bubbles did not pass through the abdominal wall or skull (a simplified model of aberration was applied to the microbubble channel data). We found it was necessary to split the two simulations because speed of sound heterogeneity caused distortions in the beamformed images relative to the ground truth microbubble map, which led to difficulties in training the models. It is possible that the small, unfolded architecture simply did not have the capacity necessary to account for this phenomenon. We plan to revisit this specific problem in future work with different models and new strategies for creating the ground truth labels.

Another limitation of this work is the narrow range of imaging conditions that were simulated and tested. It is likely that new models would need to be trained to perform inference on image data from different transducers or from the same transducers using different imaging schemes (*e.g.*, focused rather than plane-wave transmission). Therefore, the our approach is not optimal, since the computational cost of training a new model for each transducer configuration would be considerable. In future work, we hope to adapt our simulation and training strategies to produce models that are more generalizable.

One noteable drawback of our proposed method is the requirement to isolate the contrast agent signal from the background speckle before performing inference with the machine learning model. This is also a limitation of the previously mentioned studies [21]–[25], [29]. Though we demonstrated that the model predictions are more robust to noise than a conventional localiztion technique, many of the widely used filtering approaches for removing tissue speckle are sensitive to factors such as physiological motion. In future work, we plan to test spatiotemporal processing of seqeuences of B-mode images to extract microbubble positions, similar to [25].

4.5 Conclusion

In this study we train deep learning models to detect and localize microbubble contrast agents for ultrasound localization microscopy using realistic simulations of point scatterers and reverberation clutter. *In silico*, we demonstrate a substantial reduction in localization error across a wide range of signal-to-noise ratios compared to a standard localization technique. This methodology has the potential to improve the resolution and fidelity of super-resolution microvascular images in challenging scenarios that are common to human clinical imaging.

REFERENCES

[1] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.

[2] S. T. Hess, T. P. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophysical Journal*, vol. 91, no. 11, pp. 4258–4272, 2006.

[3] R. Parthasarathy, "Rapid, accurate particle tracking by calculation of radial symmetry centers," *Nature methods*, vol. 9, no. 7, pp. 724–726, 2012.

[4] K. Christensen-Jeffries, S. Harput, J. Brown, P. N. Wells, P. Aljabar, C. Dunsby, M.-X. Tang, and R. J. Eckersley, "Microbubble axial localization errors in ultrasound super-resolution imaging," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 64, no. 11, pp. 1644–1654, 2017.

[5] K. Christensen-Jeffries, R. J. Browning, M. X. Tang, C. Dunsby, and R. J. Eckersley, "In vivo acoustic super-resolution and super-resolved velocity mapping using microbubbles," *IEEE Transactions on Medical Imaging*, vol. 34, no. 2, pp. 433–440, 2015.

[6] T. M. Kierski, D. Espindola, I. G. Newsome, E. Cherin, J. Yin, F. S. Foster, C. E. Demore, G. F. Pinton, and P. A. Dayton, "Superharmonic ultrasound for motion-independent localization microscopy: Applications to microvascular imaging from low to high flow rates," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 67, no. 5, pp. 957–967, 2020.

[7] P. J. Phillips, "Contrast pulse sequences (cps): Imaging nonlinear microbubbles," *Proceedings of the IEEE Ultrasonics Symposium*, vol. 2, pp. 1739–1745, 2001.

[8] J.-J. Hwang and D. H. Simpson, "Two pulse technique for ultrasonic harmonic imaging," US Patent 5,951,478, 1999.

[9] G. A. Brock-Fisher, M. D. Poland, and P. G. Rafter, "Means for increasing sensitivity in non-linear ultrasound imaging systems," US Patent 5,577,505, 1996.

[10] C. Demené, T. Deffieux, M. Pernot, B.-F. Osmanski, V. Biran, J.-L. Gennisson, L.-A. Sieu, A. Bergel, S. Franqui, J.-M. Correas, *et al.*, "Spatiotemporal clutter filtering of ultrafast ultrasound data highly increases doppler and fultrasound sensitivity," *IEEE transactions on medical imaging*, vol. 34, no. 11, pp. 2271–2285, 2015.

[11] C. Errico, J. Pierre, S. Pezet, Y. Desailly, Z. Lenkei, O. Couture, and M. Tanter, "Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging," *Nature*, vol. 527, no. 7579, pp. 499–502, 2015.

[12] T. Opacic, S. Dencks, B. Theek, M. Piepenbrock, D. Ackermann, A. Rix, T. Lammers, E. Stickeler, S. Delorme, G. Schmitz, *et al.*, "Motion model ultrasound localization microscopy for preclinical and clinical multiparametric tumor characterization," *Nature communications*, vol. 9, no. 1, pp. 1–13, 2018.

[13] Y. Desailly, J. Pierre, O. Couture, and M. Tanter, "Resolution limits of ultrafast ultrasound localization microscopy," *Physics in medicine & biology*, vol. 60, no. 22, p. 8723, 2015.

[14] J. R. McCall, P. A. Dayton, and G. F. Pinton, "Characterization of the ultrasound localization microscopy resolution limit in the presence of image degradation," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 1, pp. 124–134, 2021.

[15] G. F. Pinton, G. E. Trahey, and J. J. Dahl, "Sources of image degradation in fundamental and harmonic ultrasound imaging using nonlinear, full-wave simulations," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 58, no. 4, pp. 754–765, 2011.

[16] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.

[17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[18] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241.

[19] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International Conference on Machine Learning*, PMLR, 2019, pp. 6105–6114.

[20] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1874–1883.

[21] R. J. van Sloun, O. Solomon, M. Bruce, Z. Z. Khaing, H. Wijkstra, Y. C. Eldar, and M. Mischi, "Super-resolution ultrasound localization microscopy through deep learning," *IEEE transactions on medical imaging*, vol. 40, no. 3, pp. 829–839, 2020.

[22] X. Liu, T. Zhou, M. Lu, Y. Yang, Q. He, and J. Luo, "Deep learning for ultrasound localization microscopy," *IEEE transactions on medical imaging*, vol. 39, no. 10, pp. 3064–3078, 2020.

[23] K. G. Brown, D. Ghosh, and K. Hoyt, "Deep learning of spatiotemporal filtering for fast super-resolution ultrasound imaging," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 67, no. 9, pp. 1820–1829, 2020.

[24] U.-W. Lok, C. Huang, P. Gong, S. Tang, L. Yang, W. Zhang, Y. Kim, P. Korfiatis, D. J. Blezek, F. Lucien, *et al.*, "Fast super-resolution ultrasound microvessel imaging using spatiotemporal data with deep fully convolutional neural network," *Physics in Medicine & Biology*, vol. 66, no. 7, p. 075 005, 2021.

[25] L. Milecki, J. Porée, H. Belgharbi, C. Bourquin, R. Damseh, P. Delafontaine-Martel, F. Lesage, M. Gasse, and J. Provost, "A deep learning framework for spatiotemporal ultrasound localization microscopy," *IEEE Transactions on Medical Imaging*, vol. 40, no. 5, pp. 1428–1437, 2021.

[26] G. F. Pinton, J. Dahl, S. Rosenzweig, and G. E. Trahey, "A heterogeneous nonlinear attenuating full-wave model of ultrasound," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 56, no. 3, pp. 474–488, 2009.

[27] M. J. Ackerman, "The visible human project," *Proceedings of the IEEE*, vol. 86, no. 3, pp. 504–511, 1998.

[28] S. Pichardo, V. W. Sin, and K. Hynynen, "Multi-frequency characterization of the speed of sound and attenuation coefficient for longitudinal transmission of freshly excised human skulls," *Physics in Medicine & Biology*, vol. 56, no. 1, p. 219, 2010.

[29] R. J. Van Sloun, R. Cohen, and Y. C. Eldar, "Deep learning in ultrasound imaging," *Proceedings of the IEEE*, vol. 108, no. 1, pp. 11–29, 2019.

[30] X.-P. Zhang, "Thresholding neural network for adaptive noise reduction," *IEEE transactions on neural networks*, vol. 12, no. 3, pp. 567–584, 2001.

[31] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Y. Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng, *TensorFlow: Large-scale machine learning on heterogeneous systems*, Software available from tensorflow.org, 2015.

[32] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *ArXiv preprint arXiv:1412.6980*, 2014.

CHAPTER 5: DETECTING CANCER IN ULTRASOUND MICROVASCULAR IMAGES USING DEEP CONVOLUTIONAL NEURAL NETWORKS[1]

5.1 Introduction

Cancer is the 2nd leading cause of death in the United States, killing approximately 600,000 individuals annually [1]. According to the same review, the death rate declined 40% and 51% for breast and prostate cancer patients, respectively, between 1989 and 2016, and the authors suggest that these positive trends are due in part to improved screening techniques. However, it is still important to note the shortcomings of current screening technologies. One recent study found that more than half of the women who undergo a biopsy following a mammogram do not have malignant tumors [2]. A similar pattern was noted for men with prostatic lesions [3]. These needless biopsies impose a psychological burden on patients and logistical and financial burdens upon the healthcare system. Hence, there exists a strong motivation to improve the diagnostic power of biomedical imaging technologies for cancer screening.

Many cancers are characterized by unchecked and disorderly angiogenesis [4]. As tumors grow, regions of hypoxic tissue can develop, leading to an overexpression of angiogenic molecules such as vascular endothelial growth factor receptor-2 (VEGFR-2) [5]. The recruitment of new blood vessels to feed tumor growth is so rapid and pronounced that this phenomenon is described as one of the hallmarks of cancer [6]. Based on this knowledge, researchers have used angiographic imaging to gauge malignancy in organs such as the brain [7]. Magnetic resonance (MR) imaging is a common and effective modality for this task, though it suffers from relatively low spatiotemporal resolution and is costly compared to other forms of imaging.

Biomedical ultrasound is a widely-accessible alternative for imaging soft tissues with excellent spatial and temporal resolutions. Contrast-enhanced ultrasound imaging (CEUS) utilizes in-

[1]Portions of this chapter are under review in the IEEE Transactions on Medical Imaging.

jectable microbubbles (MBs) to enhance the echogenicity of the blood pool. MBs are gas spheres with a lipid or protein shell typically between 1 and 8 μm in diameter, allowing them to circulate much like erythrocytes. One clinical application of CEUS is quantitative perfusion imaging. For example, CEUS has been used to quantify the perfusion of neoplastic lesions in the liver [8] and kidney [9] by measuring speckle brightness over time after contrast injections. While useful for estimating vessel density in a region of interest, the resolution of this method is insufficient for resolving structures such as individual arterioles.

Acoustic angiography (AA) is a form of CEUS that has improved contrast to tissue ratio (CTR) and higher resolution (Figure 5.1. Conventional CEUS techniques, such as pulse inversion and amplitude modulation, harness frequencies around the 2^{nd} and 3^{rd} harmonics of the transmitted frequency [10]. AA uses ultra-wideband transducers to record "superharmonics," or the 4^{th} and higher harmonics of the fundamental frequency [11], [12]. The superharmonic echoes result from the strongly nonlinear oscillations of microbubbles driven near their resonant frequencies [13]. Normally, clinical transducers are not sensitive to these signals because of bandwidth limitations. To overcome this challenge, dual-frequency devices have been constructed [14], [15]. Typically, a low-frequency element (*e.g.*, 2-4 MHz) is used for transmitting a pulse into the body, and a high-frequency element (*e.g.*, 25-30 MHz) is used to record the resulting echoes. Under normal imaging conditions, superharmonics are produced at negligible levels by tissues, while microbubbles produce significant broadband content. The result is high-resolution imaging with excellent CTR, allowing vessels as small as 100 μm to be resolved in real time [14].

Acoustic angiography has been used to quantify the differences in vascular density and tortuosity between tumor-bearing and healthy tissues [16], [17]. More recently, a classifier based on a clustering algorithm applied to vessel-level statistics achieved an AUC score of 0.95, outperforming human readers on the test data [18]. To perform these quantitative analyses on AA data, a reader must manually segment individual blood vessels from each image volume. To date, this has been accomplished with custom software based on a ridge-traversal technique for extracting tubular objects [19]. After segmentation, the vessel centerlines are further processed to extract

various metrics of tortuosity. Some drawbacks to this pipeline are that the segmentation is time-consuming (1-2 hours per volume) and subject to inter-operator variability even with a software aid, meaning it is impractical to introduce a similar method into a clinical workflow.

In recent years, advancements in the field of computer vision have made machines competitive with their human counterparts in a variety of complex image-based tasks, such as classification, segmentation, object detection, and captioning. Deep convolutional neural networks, or CNNs, surpassed human performance on the difficult 1,000-class ImageNet challenge nearly 7 years ago [20], and the state-of-the-art has improved steadily since. CNNs are a class of machine learning model with a general design that is analogous to the hierarchical structure of the visual cortex. While many variants of CNNs have been proposed for different tasks, the fundamental unit of these models remains the convolutional layer. Each layer is comprised of a number of filters which are learned from the data and whose outputs are passed through a nonlinear activation function. Typically, layers near the network input learn filters that detect basic structures (*e.g.*, blobs and edges), while deeper layers respond to more abstract features. Scaling the resolution (channel size), width (number of channels per layer), and/or depth (number of layers) of a network can modify its expressive power [21]. A modern CNN might include tens or even hundreds of convolutional layers along with additional operators for normalizing and downsampling the stream of data as it flows through the network. One of the benefits of processing an image with a CNN is that once a model has been trained, inference on new data can typically be performed in a fraction of a second.

Deep learning has been applied to ultrasound tumor imaging. For example, CNNs have been used to improve detection and classification of cancer in B-mode and conventional CEUS images [22]–[24]. Inspired by these and similar examples of deep learning applied to medical image processing, our aim in the present study is to explore the feasibility of training convolutional neural networks to detect cancer in acoustic angiography images. First, we collect and curate a dataset of nearly two hundred acoustic angiography images, which we make available for other researchers to develop their own classification algorithms. Then, we train both two- and three-

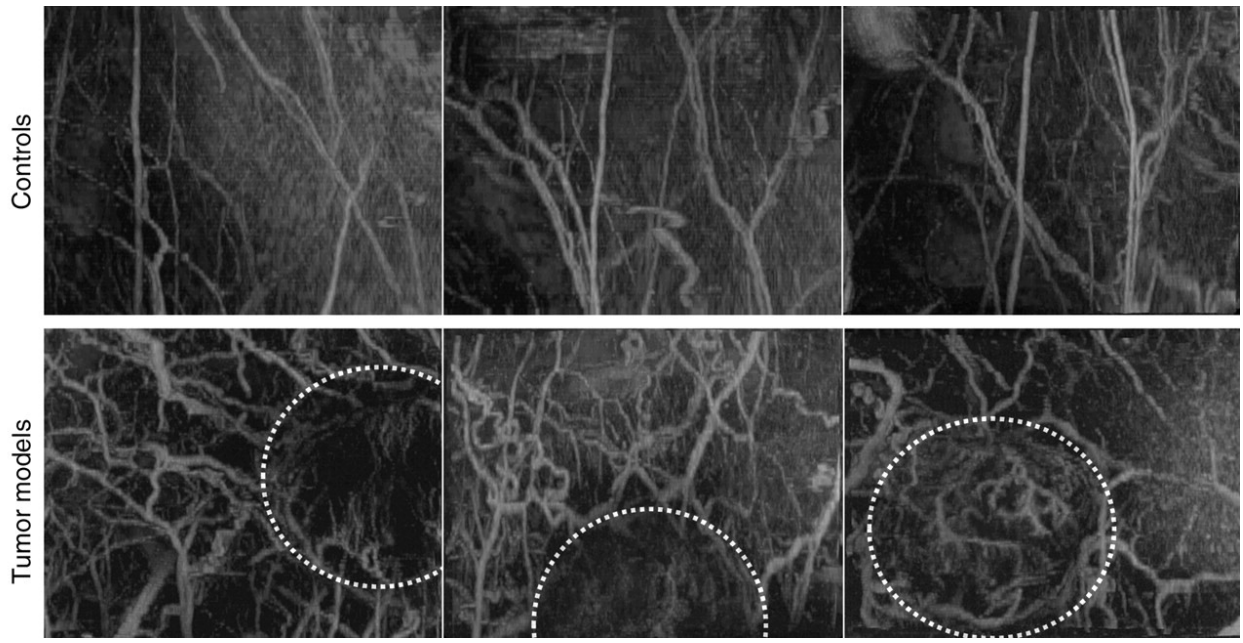

Figure 5.1: Reproduced from Gessner *et al*, 2012 with permission. Maximum intensity projections from representative acoustic angiography images of the same rodent model used in the present study. For healthy animals in the top row, the blood vessels have an orderly, branch-like structure. On the other hand, the vasculature feeding the fibrosarcoma xenografts in the bottom row is tortuous and disorderly. The approximate positions of tumors are indicated by the white dotted lines.

dimensional CNNs to classify ultrasound microvascular images, and compare the performance of these models with the results reported in previous publications. Finally, we explore utilizing saliency maps to gain a qualitative understanding of the inner workings of best-performing machine learning models.

5.2 Materials and Methods

5.2.1 Contrast agent preparation

The in-house non-targeted microbubble contrast agent was formulated from a 1 mM lipid solution comprised of 900 μM 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC) and 100 μM 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000] (DSPE-PEG2000) lipids in 5% (v/v) glycerol and 15% (v/v) propylene glycol in phosphate-buffered saline (PBS). The solution was gas-exchanged with decafluorobutane 4 times and then mechanically agitated for 45 seconds to form a polydisperse population of microbubbles (Vialmix, Lantheus Holdings, Billerica, MA, USA). The contrast agent was characterized using

an Accusizer FX-Nano (Entegris, Billerica, MA, USA). The mean diameter was 0.96 $\pm$ 0.55 μm and the stock concentration was 3.6e10 mL^{-1}.

5.2.2 Animal Care and Tumor Model

All experiments were performed on female Fischer 344 rats (Charles River Laboratories, Durham, NC, USA) and imaging protocols were approved by the Institutional animal Care and Use Committee at the University of North Carolina at Chapel Hill. Animals were housed in groups of 4 and could access water and standard rat feed *ad libitum*. The cages measured 140 in^2 with individual ventilation, and were located in a vivarium with a 12-hour day-night cycle.

The fibrosarcoma (FSA) model used in this study has been passaged on a regular schedule *in vivo* from tissue received from Mark Dewhirst's laboratory [25], [26]. Approximately 1-2 mm^3 of FSA tissue from a donor animal was transplanted into the right flank of each rodent and monitored for 10 days prior to imaging. Animals were humanely euthanized when tumors exceeded 20 mm in diameter.

5.2.3 In vivo imaging

Animals were anesthetized with a mixture of 2.5% isoflurane and oxygen administered at 1 L/min. To prevent hypothermia, the rats were warmed with an infrared lamp and transferred to a heating pad during imaging. Animals were imaged with a Vevo 770 scanner (FUJIFILM VisualSonics, Inc., Toronto, ON, Canada) connected to a RMV 707 transducer modified for dual-frequency imaging [16]. B-mode images were acquired by transmitting and receiving at 30 MHz. For acoustic angiography imaging, the microbubble contrast agent was diluted to 1e10 mL^{-1} and injected at 30 μL/min with a syringe pump (Harvard Apparatus, Holliston, MA, USA) into a 24-gauge catheter placed in the tail vein. AA images were created by transmitting a single cosine-windowed pulse with a center frequency equal to 4 MHz [27] and recording bubble echoes with the 30 MHz transducer. For both B-mode and AA imaging, 3-D images were captured by mechanically sweeping the transducer in the elevation dimension in 102-μm increments for 20 mm using a linear motion stage. At each position of the acoustic angiography scan, 4 frames were acquired at 5 frames/sec and averaged to suppress signals from sub-resolution blood vessels. The

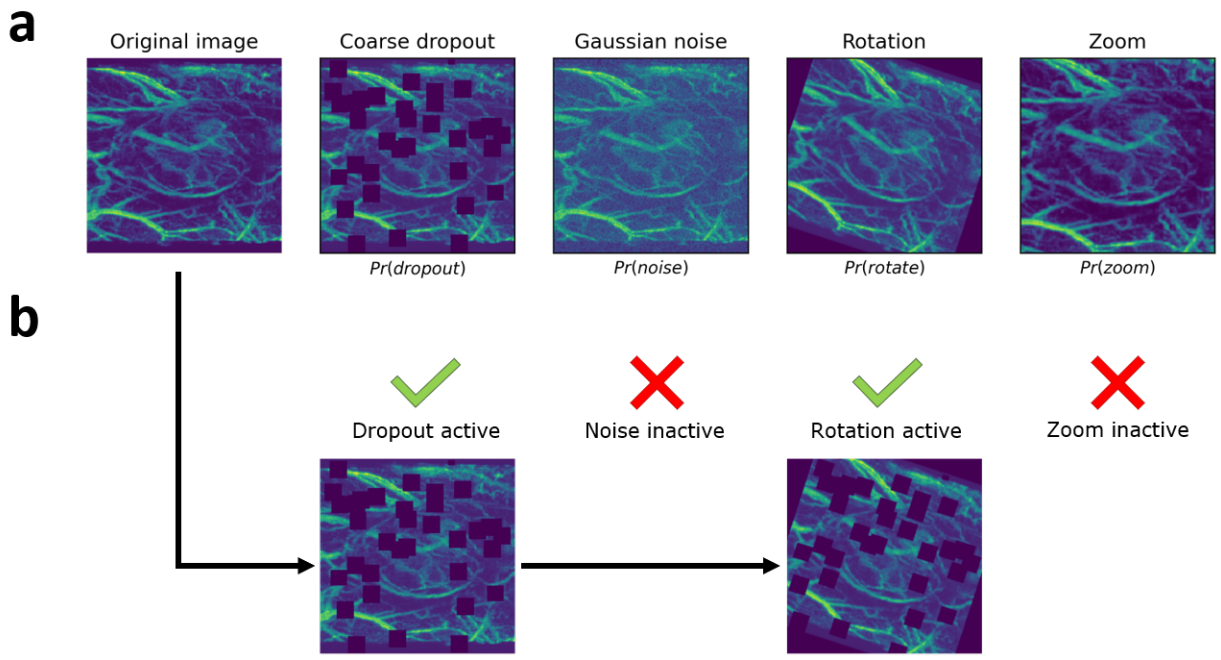

Figure 5.2: An overview of the data augmentation pipeline. (a) Each operator is applied to the same image of a tumor to illustrate its effect. (b) An example of one possible training data augmentation, showing active dropout and rotation but without the addition of noise or zooming in this case. The final image is the result of applying each of the active operators to the input image in order from left to right. The probabilities of each operator being active are hyperparameters of the model being trained.

volume dimensions were $25.6 \times 25.6 \times 20$ mm^3, or $512 \times 512 \times 196$ px^3.

The tumors were imaged at multiple time points spaced by a minimum of 4 days to ensure the dataset was comprised of a relatively uniform distribution of target sizes. The 4-day minimum allowed both significant alterations to vessel structures between adjacent time points and efficient use of the animal model (*i.e.*, producing multiple images per tumor implant). Control images were acquired from the contralateral flank of implanted animals and both flanks of healthy animals. Control images were acquired only once per animal, since the vessel structures were very slow to change in the absence of a tumor. Images of 179 samples were collected in total, with 89 tumors and 90 controls.

5.2.4 Image preprocessing and augmentation

All data preprocessing and augmentation was implemented using PyTorch [28] and MONAI [29]. Raw images were resized to $224 \times 224 \times 224$ px^3 using bilinear interpolation, and an addi-

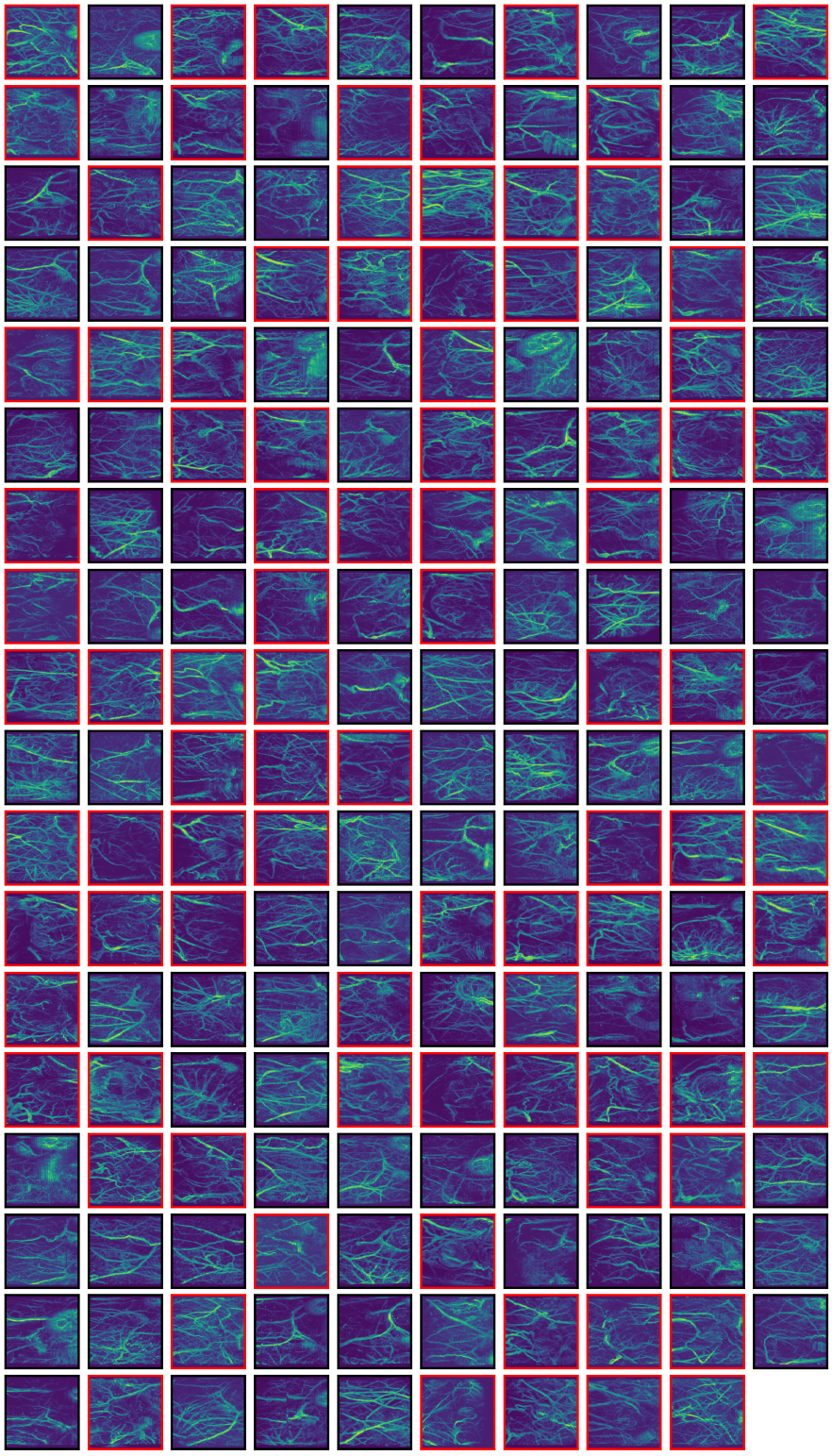

Figure 5.3: Maximum intensity projections of the entire dataset ($n = 179$). Red borders indicate that the image is of a tumor, black borders indicate controls.

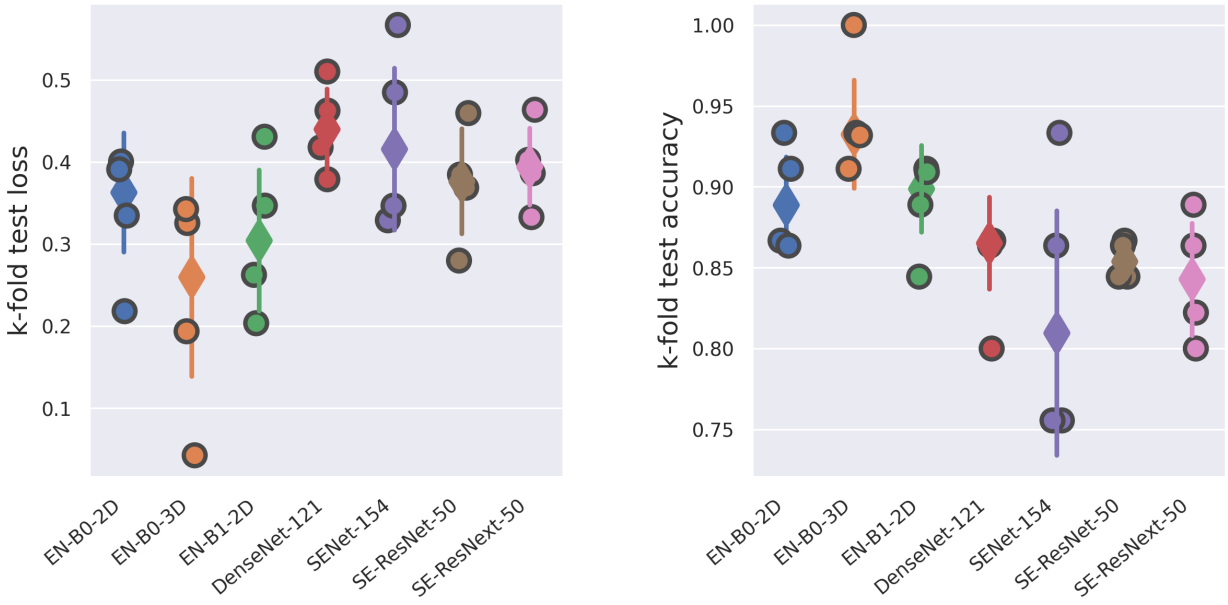

Figure 5.4: The 3-D implementation of EfficientNet-B0 achieved both the lowest test loss and highest test accuracy. The two point plots show the k-fold test loss (*left*) and accuracy (*right*) values for the highest-performing models for each architecture, determined by selecting the model with the lowest mean k-fold test loss across the 200-iteration hyperparameter sweep. Circles denote the per-fold loss and accuracy results, diamonds the median of each group consisting of the four folds, and lines the inter-quartile ranges for each group.

tional set of 2-D maximum intensity projections (MIPs) was generated for training the 2-D CNNs by taking the maximum along the axial dimension for each lateral-elevation position. The full dataset of MIPs is provided in Figure 5.3. Data augmentation was applied to each mini-batch during training to reduce overfitting (Figure 5.2). The augmentation pipeline was composed of operators for coarse dropout, additive Gaussian noise, image rotation, and zooming, in that order. The probabilities of each augmentation operator being active for some mini-batch were given by the current set of hyperparameters from the random search. For example, if $\Pr(\text{Dropout}) = 0.5$ for the current model, the dropout operator had a 50% chance of being active for the current batch of training images. If an operator was inactive, it did not modify the inputs in any way. When an operator was active, it modified images according to the parameters randomly selected from the uniform distributions described in Table 5.3. For example, the rotation operator might apply rotations of $+\frac{\pi}{12}$ and $-\frac{\pi}{5}$ to subsequent batches. After data augmentation, all images were scaled to an intensity range of 0 to 1.

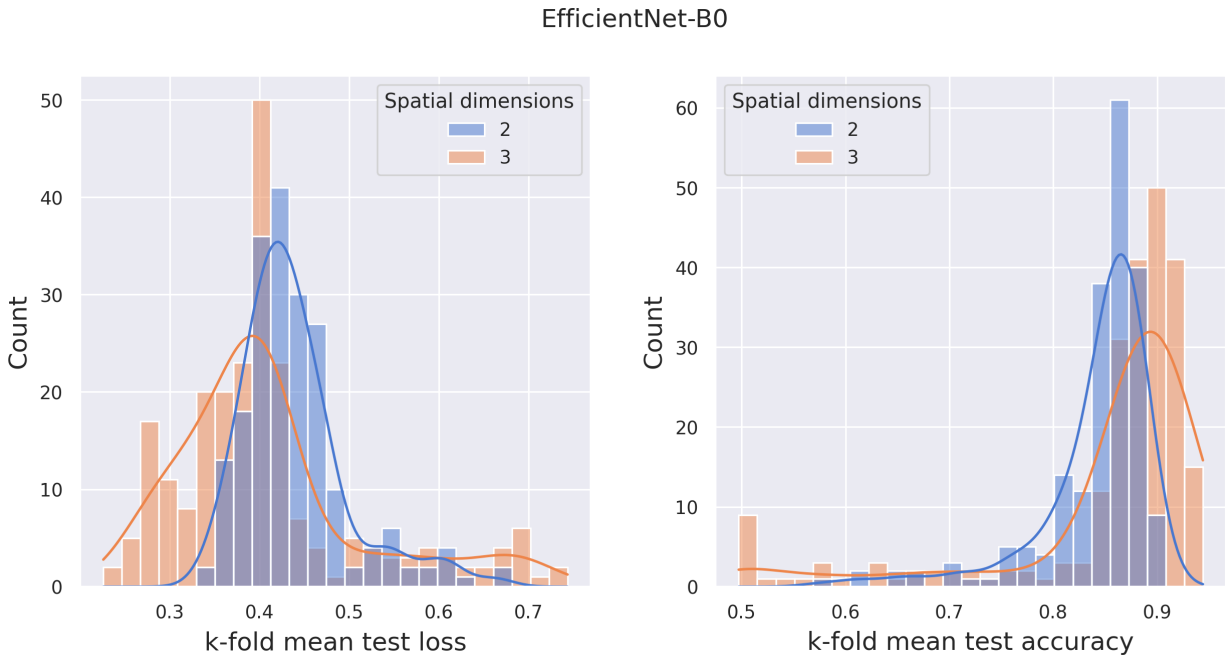

Figure 5.5: A comparison of the 2-D and 3-D implementations of EfficientNet-B0. Distributions of the mean k-fold test loss (*left*) and accuracy (*right*) values for 2-D (blue) and 3-D (orange) EfficientNet-B0 architectures measured over the course of their respective hyperparameter searches. The solid lines plot the kernel density estimates for each distribution. These data suggest that within the hyperparameter space tested in this study, the 3-D models are better suited for classifying AA images.

5.2.5 Model training and evaluation

The model architectures used in this study are listed in Table 5.1. Each model was implemented using PyTorch and MONAI and always initialized using the same random seed to ensure reproducibility. All models were trained on Tesla V100-SXM2 GPUs (NVIDIA, Santa Clara, CA, USA). Training was distributed across multiple GPUs when necessary.

For each architecture, the hyperparameter space defined in Table 5.2 was randomly searched [30] for 200 iterations using Weights and Biases for experiment tracking [31]. Since the dataset was relatively small ($n = 179$), we employed k-fold cross-validation with $k = 4$ and stratified groups to reduce the impact of sampling on the final test scores. For each set of hyperparameters, k new models were identically initialized and trained independently, each using a different group of data for testing and the remaining $k - 1$ groups for training. Each model was trained for 80 epochs by minimizing the cross-entropy loss using the AdamW optimizer, with $\beta_1 = 0.9$, β_2

Table 5.1: Model performance summary

Model	# Parameters	Accuracy	MCC	Sensitivity	Specificity	F_1 score
DenseNet-121 [34]	6.9 M	0.849 ± 0.028	0.703 ± 0.060	0.854 ± 0.058	0.844 ± 0.067	0.849 ± 0.025
SENet-154 [35]	113.0 M	0.827 ± 0.076	0.657 ± 0.150	0.844 ± 0.071	0.811 ± 0.095	0.830 ± 0.073
SE-ResNet-50	26.0 M	0.855 ± 0.010	0.711 ± 0.022	0.820 ± 0.003	0.889 ± 0.020	0.849 ± 0.012
SE-ResNext-50	25.5 M	0.844 ± 0.035	0.692 ± 0.066	0.831 ± 0.088	0.854 ± 0.061	0.839 ± 0.044
EfficientNet-B0 [21]	4.0 M	0.894 ± 0.030	0.788 ± 0.060	0.899 ± 0.038	0.889 ± 0.022	0.893 ± 0.031
EfficientNet-B1	6.5 M	0.888 ± 0.027	0.787 ± 0.054	0.900 ± 0.077	0.878 ± 0.085	0.889 ± 0.026
EfficientNet-B0-3D	4.6 M	$\mathbf{0.944 \pm 0.033}$	$\mathbf{0.892 \pm 0.064}$	$\mathbf{0.955 \pm 0.045}$	$\mathbf{0.932 \pm 0.068}$	$\mathbf{0.945 \pm 0.032}$

"M": million, "MCC": Matthew's correlation coefficient

= 0.99 [32]. The best model for each architecture was determined by measuring the mean test loss across the 4 folds for each of the 200 iterations of the random search. The number of folds was chosen to balance between providing a sufficiently large training set and the computational burden of training $200 \times k$ models for each architecture.

Some models were qualitatively evaluated after training by using GradCAM [33] to generate saliency maps. Each model was quantitatively evaluated by computing the following metrics on the test data predictions for each fold:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{5.1}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{5.2}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \tag{5.3}$$

$$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}, \tag{5.4}$$

$$F_1 = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}, \tag{5.5}$$

where TP is the number of true positive predictions, FP is the number of false positives, TN is the number of true negatives, FN is the number of false negatives, and MCC is the Matthew's correlation coefficient (also known as the ϕ coefficient).

Table 5.2: Hyperparameter search space

Parameter	Minimum	Maximum	Distribution
Batch size (2-D)	4	32	Discrete uniform
Batch size (3-D)	1	10	Discrete uniform
Pr(Dropout)	0.0	1.0	Uniform
Pr(Noise)	0.0	1.0	Uniform
Pr(Zoom)	0.0	1.0	Uniform
Pr(Rotate)	0.0	1.0	Uniform
Rotation range	$\pm\frac{\pi}{12}$	$\pm\frac{\pi}{4}$	Uniform
Learning rate (2-D)	5.0e−5	5.0e−3	Log-Uniform
Learning rate (3-D)	1.0e−5	1.0e−3	Log-Uniform
Weight decay	1.0e−5	1.0e−3	Log-Uniform

5.3 Results

Table 5.1 contains a summary of the performance metrics for each of the model architectures tested in the k-fold cross-validation study. As expected, the 3-D implementation of EfficientNet-B0 outperformed all of the 2-D networks, achieving both the lowest average loss and the highest average accuracy of 94.4% (Figure 5.4). The average sensitivity and specificity of the 3-D CNN were 0.955 and 0.932, respectively, compared with sensitivity and specificity of 0.871 and 0.941 in the prior work where manually segmented images were classified using a clustering approach [18]. The F_1 score of the best 3-D model was 0.945, and its Matthew's correlation coefficient (MCC) was 0.892.

To better understand if the 3-D CNNs were generally better-suited for this classification problem, we also compared them against the 2-D EfficientNet-B0 models across the entire hyperparameter search space (Figure 5.5). By comparing the distributions of mean test loss and mean test accuracy for the sets of 200 EfficientNet-B0 models, it seems that the 3-D architecture performed better on average, except for the occasions when it did not learn its task at all. To gain a better understanding of the circumstances that led to poor generalizability, we examined the effects of varying the various hyperparameters on model performance. Both the 2-D and 3-D EfficientNet-B0 architectures were most sensitive to changes in batch size and learning rate,

with sharp decreases in accuracy as learning rate became too large and/or batch size became too small (Figure 5.6). The results suggest that increasing the maximum batch size for the 3-D models might yield further improvements, since the performance was generally correlated with batch size, especially for higher learning rates. However, we were unable to test this theory in the present study because of GPU memory constraints.

A comparison of the 2-D CNN architectures suggested that the SENets were possibly over-parameterized for the task at hand, given that they produced the lowest scores and were the largest by far in terms of number of parameters (113 M for SENet-154). On the other hand, the lightweight EfficientNet-B0 model achieved a 89.4% accuracy with a sensitivity of 0.899 and specificity of 0.889. Saliency maps generated from one of the 2-D EfficientNet-B0 models suggest that the networks are learning image features associated with elevated levels of vascular tortuosity and density rather than noise or other artifacts (Figure 5.7).

5.4 Discussion

In this study, we demonstrated the feasibility of training both 2-D and 3-D convolutional neural networks to detect tumors in ultrasound images of the microvasculature in a rodent model. Of the architectures tested, the 3-D EfficientNet-B0 produced the best results by far, achieving a k-fold average sensitivity of 0.955 and specificity of 0.932 on test data. This result supported our hypothesis that the 3-D models are better-suited for this classification task, since a great deal of information is lost when the images are represented as maximum intensity projections. However, the results of the 2-D CNNs were also quite promising, considering that the raw volumetric images were projected into a single plane for processing (see Figure 5.3 for examples). This is especially relevant when considering the deployment of trained models to on-site devices. Processing 2-D images rather than full 3-D volumes requires significantly less memory at the expense of some accuracy, which could be an appropriate trade off in circumstances where computational resources are limited. A small, 2-dimensional representation of images could also be ideal for telemedicine applications where data is transmitted off-site to be processed.

One of the limitations of this study was the size of the dataset which was small by traditional

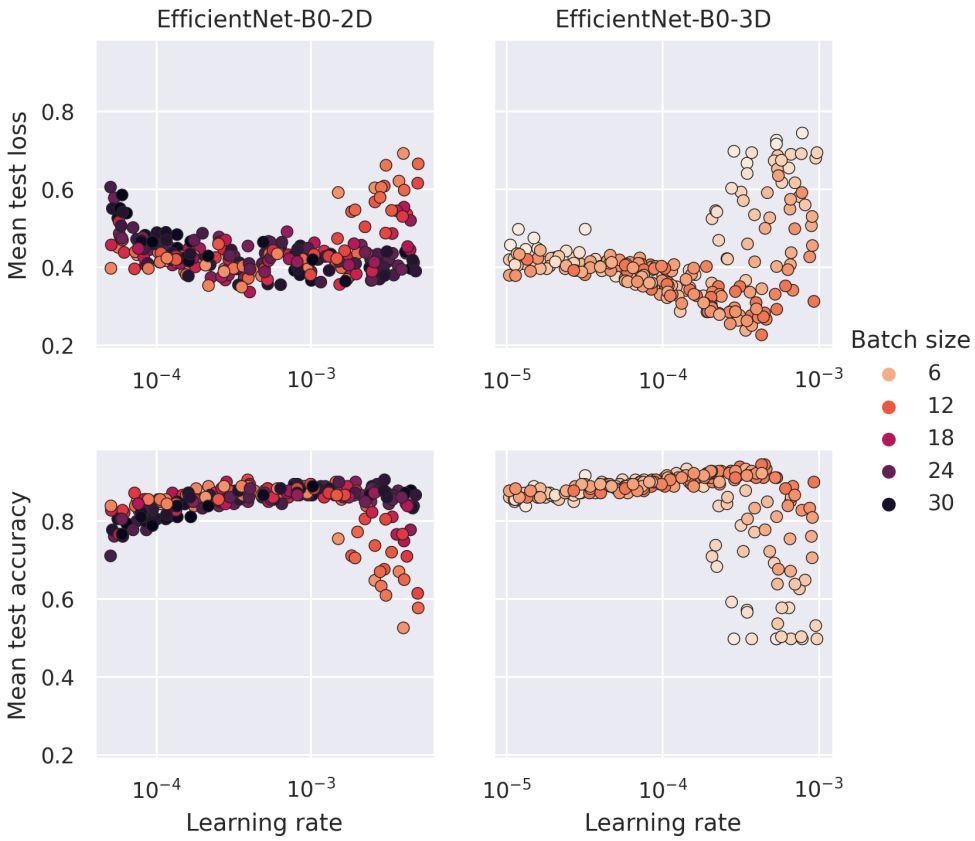

Figure 5.6: Examining the effects of learning rate and batch size on EfficientNet-B0 performance. The left and right columns show results for 2-D and 3-D models, respectively. The top and bottom rows correspond to mean test loss and mean test accuracy, respectively. The batch size during training for each model is given by the color of each data point. The color scale on the right is abbreviated for clarity, the full ranges of possible values of batch size are provided in Table 5.2.

deep learning standards. Another limitation is that all of our images were collected from the same anatomical region in a single species of rodent. Therefore, it is likely that these models would not generalize well outside of the narrow scope of this work. For these reasons, the results presented herein should be considered as a proof-of-concept. In future studies, we would like to explore the potential impact of this methodology by collecting images of different tumor models in a variety of organs, such as the spontaneous model of human breast cancer in C3(1)/Tag mice. Ultimately, our goal is to improve the limit of detection for small solid tumors by achieving similar or better sensitivity and specificity values for tumors less than or equal to 3 mm in diameter.

We also plan to conduct a longitudinal study to explore the ability of CNNs to predict the response of various tumor models to treatment. We hypothesize that training CNNs on combined

Table 5.3: Data augmentation parameters

Transform	Parameter	Value
2-D Coarse dropout	Min. # holes	20
	Max # holes	35
	Min hole size	5 px
	Max hole size	20 px
	Fill value	0.0
3-D Coarse dropout	Min. # holes	25
	Max # holes	50
	Min hole size	5 px
	Max hole size	20 px
	Fill value	0.0
Gaussian noise	Mean	0.1
	Min std. dev.	0.0
	Max std. dev.	0.1
Zoom	Min zoom factor	0.9
	Max zoom factor	1.1

acoustic angiography and molecular images will produce models that can accurately classify samples as responders or non-responders at an earlier timepoint than would be possible by monitoring the changes in tumor volume alone.

In spite of the limitations mentioned earlier, we believe that the results of this study are encouraging for future clinical applications of acoustic angiography for cancer detection and possibly monitoring response to treatment. Bearing in mind that we utilized a different tumor model for the present study, we achieved higher accuracy than in our previous work while completely eliminating the need for time-consuming image segmentation [18]. Each of the CNNs in this study performs inference on new images in less than a second on modern GPUs, so a similar methodology could be used in a clinical setting for near real-time processing of images as they are acquired. Also, we predict that saliency maps and similar techniques for visualizing network attention will be useful tools in a clinical workflow for highlighting regions of interest for further investigation.

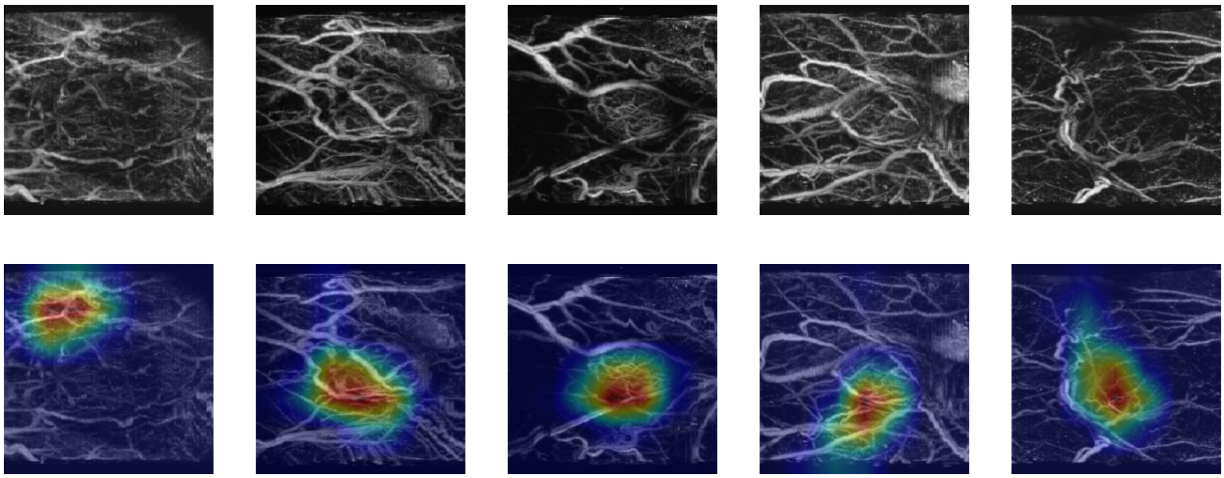

Figure 5.7: Saliency maps created with GradCAM to qualitatively assess the performance of one of the 2-D EfficientNet-B0 models on test data. These images were correctly identified as tumors. In these examples, warm colors highlight regions of increased network attention. These images suggest that the network has learned tumor-associated features.

5.5 Conclusion

In this study, we demonstrate the feasibility of using deep convolutional neural networks to detect tumors in volumetric ultrasound images of small blood vessels. The results of the *k*-fold cross-validation study suggest that lightweight 3-dimensional CNNs such as the EfficientNet-B0-3D have the potential to quickly and reliably process angiographic images, and we expect the performance of these models to improve with more data. Moving forward, our aims are to expand this methodology to a spontaneous tumor model to improve detection, and also to train models to predict the response of cancers to treatment.

REFERENCES

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *CA: A cancer journal for clinicians*, vol. 69, no. 1, pp. 7–34, 2019.

[2] J. Seely and T. Alhassan, "Screening for breast cancer in 2018—what should we be doing today?" *Current Oncology*, vol. 25, no. s1, pp. 115–124, 2018.

[3] H. U. Ahmed, A. E.-S. Bosaily, L. C. Brown, R. Gabe, R. Kaplan, M. K. Parmar, Y. Collaco-Moraes, K. Ward, R. G. Hindley, A. Freeman, *et al.*, "Diagnostic accuracy of multi-parametric mri and trus biopsy in prostate cancer (promis): A paired validating confirmatory study," *The Lancet*, vol. 389, no. 10071, pp. 815–822, 2017.

[4] P. Carmeliet and R. K. Jain, "Angiogenesis in cancer and other diseases," *Nature*, vol. 407, pp. 249–257, 6801 2000.

[5] B. Keith and M. C. Simon, "17 - tumor angiogenesis," in *The Molecular Basis of Cancer (Fourth Edition)*, J. Mendelsohn, J. W. Gray, P. M. Howley, M. A. Israel, and C. B. Thompson, Eds., Fourth Edition, Philadelphia: W.B. Saunders, 2015, 257–268.e2.

[6] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011.

[7] E. Bullitt, D. Zeng, G. Gerig, S. Aylward, S. Joshi, J. K. Smith, W. Lin, and M. G. Ewend, "Vessel tortuosity and brain tumor malignancy: A blinded study1," *Academic radiology*, vol. 12, no. 10, pp. 1232–1240, 2005.

[8] S. R. Wilson and P. N. Burns, "Liver mass evaluation with ultrasound: The impact of microbubble contrast agents and pulse inversion imaging," in *Seminars in liver disease*, Copyright© 2001 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New …, vol. 21, 2001, pp. 147–160.

[9] M. L. Robbin, M. E. Lockhart, and R. G. Barr, "Renal imaging with ultrasound contrast: Current status," *Radiologic Clinics*, vol. 41, no. 5, pp. 963–978, 2003.

[10] R. J. Eckersley, C. T. Chin, and P. N. Burns, "Optimising phase and amplitude modulation schemes for imaging microbubble contrast agents at low acoustic power," *Ultrasound in medicine & biology*, vol. 31, no. 2, pp. 213–219, 2005.

[11] A. Bouakaz and N. De Jong, "Native tissue imaging at superharmonic frequencies," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 50, no. 5, pp. 496–506, 2003.

[12] P. L. Van Neer, M. G. Danilouchkine, M. D. Verweij, L. Demi, M. M. Voormolen, A. F. Van Der Steen, and N. De Jong, "Comparison of fundamental, second harmonic, and super-

harmonic imaging: A simulation study," *The Journal of the Acoustical Society of America*, vol. 130, no. 5, pp. 3148–3157, 2011.

[13] N. De Jong, A. Bouakaz, and P. Frinking, "Basic acoustic properties of microbubbles," *Echocardiography*, vol. 19, no. 3, pp. 229–240, 2002.

[14] R. C. Gessner, C. B. Frederick, F. S. Foster, and P. A. Dayton, "Acoustic angiography: A new imaging modality for assessing microvasculature architecture," *International journal of biomedical imaging*, vol. 2013, 2013.

[15] A. Bouakaz, S. Frigstad, F. J. Ten Cate, and N. de Jong, "Super harmonic imaging: A new imaging technique for improved contrast detection," *Ultrasound in medicine & biology*, vol. 28, no. 1, pp. 59–68, 2002.

[16] S. E. Shelton, Y. Z. Lee, M. Lee, E. Cherin, F. S. Foster, S. R. Aylward, and P. A. Dayton, "Quantification of microvascular tortuosity during tumor evolution using acoustic angiography," *Ultrasound in medicine & biology*, vol. 41, no. 7, pp. 1896–1904, 2015.

[17] R. C. Gessner, S. R. Aylward, and P. A. Dayton, "Mapping microvasculature with acoustic angiography yields quantifiable differences between healthy and tumor-bearing tissue volumes in a rodent model," *Radiology*, vol. 264, no. 3, pp. 733–740, 2012.

[18] S. E. Shelton, J. Stone, F. Gao, D. Zeng, and P. A. Dayton, "Microvascular ultrasonic imaging of angiogenesis identifies tumors in a murine spontaneous breast cancer model," *International Journal of Biomedical Imaging*, vol. 2020, 2020.

[19] S. R. Aylward and E. Bullitt, "Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction," *IEEE transactions on medical imaging*, vol. 21, no. 2, pp. 61–75, 2002.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026–1034.

[21] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International Conference on Machine Learning*, PMLR, 2019, pp. 6105–6114.

[22] S. Han, H.-K. Kang, J.-Y. Jeong, M.-H. Park, W. Kim, W.-C. Bang, and Y.-K. Seong, "A deep learning framework for supporting the classification of breast lesions in ultrasound images," *Physics in Medicine & Biology*, vol. 62, no. 19, p. 7714, 2017.

[23] H. Li, J. Weng, Y. Shi, W. Gu, Y. Mao, Y. Wang, W. Liu, and J. Zhang, "An improved deep learning approach for detection of thyroid papillary cancer in ultrasound images," *Scientific reports*, vol. 8, no. 1, pp. 1–12, 2018.

[24] Y. Feng, F. Yang, X. Zhou, Y. Guo, F. Tang, F. Ren, J. Guo, and S. Ji, "A deep learning approach for targeted contrast-enhanced ultrasound based prostate cancer detection," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 16, no. 6, pp. 1794–1801, 2018.

[25] H. Yuan, T. Schroeder, J. E. Bowsher, L. W. Hedlund, T. Wong, and M. W. Dewhirst, "Intertumoral differences in hypoxia selectivity of the pet imaging agent 64cu (ii)-diacetyl-bis (n4-methylthiosemicarbazone)," *Journal of Nuclear Medicine*, vol. 47, no. 6, pp. 989–998, 2006.

[26] J. E. Streeter, R. Gessner, I. Miles, and P. A. Dayton, "Improving sensitivity in ultrasound molecular imaging by tailoring contrast agent size distribution: In vivo studies," *Molecular imaging*, vol. 9, no. 2, pp. 7290–2010, 2010.

[27] B. D. Lindsey, J. D. Rojas, K. H. Martin, S. E. Shelton, and P. A. Dayton, "Optimization of contrast-to-tissue ratio and role of bubble destruction in dual-frequency contrast-specific "acoustic angiography" imaging," in *2014 IEEE International Ultrasonics Symposium*, IEEE, 2014, pp. 1774–1777.

[28] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds., Curran Associates, Inc., 2019, pp. 8024–8035.

[29] M. Consortium, *Monai: Medical open network for ai*, version 0.6.0, If you use this software, please cite it using these metadata., Mar. 2020.

[30] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization.," *Journal of machine learning research*, vol. 13, no. 2, 2012.

[31] L. Biewald, *Experiment tracking with weights and biases*, Software available from wandb.com, 2020.

[32] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *ArXiv preprint arXiv:1711.05101*, 2017.

[33] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.

[34] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[35] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132–7141.

CHAPTER 6: CONCLUDING REMARKS

In this dissertation, we have presented a variety of improvements to processes for both creating and interpreting ultrasound microvascular images. We have demonstrated that superharmonic imaging is well-suited for detecting microbubble contrast agents for ultrasound localization microscopy. In a comparison against the conventional spatiotemporal filter, we have found that superharmonic imaging produces a better signal-to-noise ratio across the range of physiologically relevant flow rates. Furthermore, we have shown that it is possible to combine high-frame rate speckle tracking for motion compensation with dual-frequency imaging, a strategy which has not been applied previously. The result is high-fidelity localization microscopy imaging in the presence of both slow flow as well as physiological motion.

We have also demonstrated the first ever case of super-resolution ultrasound molecular imaging, using a novel imaging scheme and contrast agent tracking algorithm to precisely locate targeted microbubbles. *In vitro*, we have demonstrated a strongly linear relationship between the output of the molecular imaging scheme and the ground truth microbubble density established using high-frame rate optical microscopy. *In vivo*, we have shown the feasibility of registering pseudo-volumetric super-resolution images of blood vessels and molecular signaling from a fibrosarcoma model. Also, we have shown for the first time that it is possible to measure the relationship between individual blood vessels and molecular markers, paving the way for new approaches to quantitative ultrasound imaging.

In addition, we have extended previous work in the area of deep learning for ultrasound localization microscopy. We have shown that deep neural networks can be trained to recover sparse distributions of microbubble contrast agents from images corrupted by reverberation clutter and thermal noise. The machine learning approach improves the width of the localization error distribution as well as precision and recall compared to a peak detection algorithm previously used in

the literature. These results suggest that AI has an important role to play in the clinical translation of ULM, especially for challenging patients who produce low-SNR images.

Finally, we have shared a new framework for rapidly processing ultrasound microvascular images using deep convolutional networks. In a large k-fold cross-validation study, we have shown that lightweight, 3-D CNNs can differentiate between images of healthy and cancer tissue with excellent sensitivity and accuracy. Furthermore, this approach circumvents the requirement for manual blood vessel segmentation, saving well over an hour per image processed. By highlighting important regions in the input images using saliency maps, We have also demonstrated a potential avenue to provide feedback to the reader. We envision that a similar approach will be very helpful as the scanning volume increases, as in the case of whole-breast imaging for cancer screening.

As mentioned previously, the focus of this dissertation is the development and validation of new technologies for collecting and interpreting images of the vasculature. As we look toward the future, the next step is to answer important biological questions using these techniques. Two goals which are highly relevant to cancer research are improving the limit of detection for very small tumors and predicting the response to treatment at earlier timepoints. The work in Aims 1 and 2 demonstrates that superharmonic imaging improves SNR for slow flow (*i.e.*, small blood vessels) and is capable of super-resolution molecular imaging. We hypothesize that the combination of high-resolution molecular and vascular imaging can be leveraged to outperform earlier imaging technologies with respect to detecting small tumors and classifying cancers post-treatment.

We also believe that additional performance will be gained via the machine learning approaches presented herein. The localization pipeline proposed in Aim 3 was shown to augment both the fidelity and resolution of reconstructed images by reducing false positive localizations. Reducing noise in this manner may improve sensitivity to subtle image features associated with tumorigenesis or future remission after some treatment regimen. Finally, we hypothesize that the CNN classifiers presented in Aim 4 are ideal for detecting very small tumors and will outperform

classifiers based on vessel statistics alone. A similar framework might also be applied to predict whether a tumor will continue to grow or respond to a treatment based on an image collected post-treatment.

Overall, this dissertation contains four avenues to improve ultrasound microvascular imaging by (1) combining superharmonic imaging and localization microscopy, (2) improving the resolution of ultrasound molecular imaging by a factor of five, (3) mitigating the effects of noise on image quality, and (4) overhauling the processing pipeline for classifying images. As high-resolution ultrasound imaging continues to develop, we expect that its impact in both preclinical research and clinical settings will continue to grow.