SHG Imaging and Machine Learning Methods for Feature Extraction and Analysis Tool Development

By

Melissa I. Champer

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Biomedical Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2025

Date of final oral examination: 05/20/2025

The dissertation is approved by the following members of the Final Oral Committee:

Paul Campagnola, Professor, Biomedical Engineering
Beth Meyerand, Professor, Medical Physics and Biomedical Engineering
Kevin Eliceiri, Professor, Biomedical Engineering
Vikas Singh, Professor, Biostatistics and Medical Informatics
Filiz Yesilkoy, Professor, Biomedical Engineering

Abstract

Alterations to the extracellular matrix (ECM) are associated with an array of different diseases including cancers, fibroses, and connective tissue disorders. Within the ECM, collagen is the most abundant protein and is responsible for its structure and tensile support. Thus, understanding collagen in disease may allow for improvements in diagnosis and shed light on pathogenic mechanisms. Second Harmonic Generation (SHG) microscopy is uniquely well suited for imaging fibrillar collagen with optical sectioning. In this thesis, we develop a large library of SHG images from a variety of diseases and animal models. Using analysis tools from CurveAlign and FIJI software, we investigate different morphological features of the fibers such as length, straightness, alignment, and density. These metrics can in turn be used to classify samples using a support vector machine. This large library of samples is then used as training data for a generative adversarial network, StyleGAN2-ADA. We achieve an excellent Fréchet Inception Distance score, indicating that the network is recapitulating patterns and structures from the SHG images. We then apply Semantic Factorization (SeFa) to examine eigenvectors, along which images can be edited within the latent space of the generator. By creating analysis tools in the SeFa GUI, we are able to identify two semantics with linear behavior and strong correlation with fiber length and density. Furthermore, we integrate image projection methods into the SeFa interface. The encoder can project SHG images into the generator's latent space to then apply semantic edits. With further refinement, these tools can be used to create SHG images of collagen fibers with variable morphology based on our analysis on different disease types. Such image creation may allow for improvement of deep learning tools for collagen image analysis, which could yield major improvements in early disease diagnosis.

Acknowledgements

To my brother Jackson, who taught me from my first day of consciousness and no doubt will to my last as well.

To my brother Sam, who lent me hardware galore. Unbelievably, he decided to start coding which used to be just my thing. Sadly, he has far surpassed my own abilities and now lends a helping hand to coding projects far and wide.

To my sister Miriam, who answers nearly every one of my questions. Now we can say we've both studied ovarian cancer. You'll still always be the better doctor.

To my dear friend Nicole, whom without I feel I surely would have been unable to mentally bring myself to finish graduate school.

To my lab members, thank you for working on so many different things with me. A very special thanks to Darian James, Sam Alkmin, and Emily Shelton, who all vastly helped me understand optics, physics, and light itself. Thank you for helping me overcome my unexpected test taking hurdles. Another thank you to Eric Rentchler for being my first mentor in lab.

To my adviser Paul, and my committee members Beth, Kevin, Vikas, and Filiz, thank you for dealing with my litany of surgeries with grace. Health issues are endlessly frustrating, and I surely would have fallen through the cracks without your understanding.

To my doctors, especially my PCP and PT, thank you for reassembling my limbs and making me as close to a functional human being as possible given the constraints.

To my parents, Deb and Rob, you gave me my firmest supporters in three older siblings. I'll happily try to pay you back by carrying heavy things and your corgis as well as taking you to wild places.

To my friends and family near and far, thank you for being you.

Finally, to all those who keep me sane: Dingo, Ruti, Pickles, Dane, Horatio, Dora, and Dyno. And to those that have passed: Simon, Harry, Betsy, Peanut, Buttercup, Nibbles, Billo, Mopsey, Kerry, and my Grandparents.

Table of Contents

Abstract i
Acknowledgements ii
Table of Contents iii
List of Abbreviations v
Chapter 1: Introduction 1
1.1 Summary Statement 1
1.2 Collagen in the Extracellular Matrix 1
1.3 Second Harmonic Generation Microscopy 2
1.3.1 Microscopy Setup 2
1.3.2 Biomimetic Models with SHG 3
1.4 Machine Learning and Applications 4
1.4.1 Applications 8
1.4.2 Conclusions 21
Chapter 1 References 23
Chapter 2: SHG Imaging and Analysis 28
2.1 Summary Statement 28
2.2 General Methods 28
2.2.1 Image Analysis 29
2.3 Collagen remodeling in *Spag17* knockout mice 30
2.3.1 Introduction 30
2.3.2 Methods 32
2.3.3 Results 34
2.3.4 Discussion 42
2.3.5 Conclusions 45
2.4 Collagen in Lichen Sclerosis 45
2.4.1 Introduction 46
2.4.2 Methods 48
2.4.3 Results 50
2.4.4 Discussion 55

2.4.5 Conclusion 57
2.5 SHG Microscope Depth with Colon Samples 58
2.5.1 Results 58
2.6 Tumor SHG analysis with DDR2 knockout 64
2.6.1 Results 65
2.7 Kidney Murine Models 69
Chapter 2 References 78
Chapter 3: Artificial SHG Image Feature Tuning with StyleGAN2-ADA and SeFa Semantic Analysis 83
3.1 Summary Statement 83
3.2 Introduction 84
3.3 Methods 87
3.4 Results 95
3.5 Discussion 107
3.6 Conclusions 111
3.7 Supplemental Material 112
Chapter 3 References 119
Chapter 4: Conclusions and Future Approaches 124
4.1 Future Directions 124
Chapter 4 References 126

List of Abbreviations

SHG	Second Harmonic Generation
GAN	Generative Adversarial Network
ECM	Extracellular Matrix
DDR2	Discoidin Domain Receptor 2
SPAG17	Sperm Associated Antigen 17
SSC	Systemic Sclerosis
LS	Lichen Sclerosus
SVM	Support Vector Machine
ADA	Adaptive Discriminator Augmentation
SeFa	Semantic Factorization
FID	Fréchet Inception Distance
KID	Kernel Inception Distance
MSE	Mean Squared Error
LPIPS	Learned Perceptual Image Patch Similarity

Chapter 1: Introduction

1.1 Summary Statement

This chapter provides an overview of the extracellular matrix and the collagen fibers therein. Second harmonic generation microscopy and its applications to imaging collagen is discussed as well as using these images to fabricate scaffolds. Finally, key machine learning techniques and applications are discussed to show the current landscape of available software and algorithms. Microscopy databases provide a variety of challenges to achieving viable results in deep learning. These approaches provide a starting point for integrating collagen imaging into computational analysis pipelines.

1.2 Collagen in the Extracellular Matrix

Collagen is a protein found across tissues with functions pertaining to structure and cell-matrix interactions. It consists of repeating polypetide chains Glycine-X-Y, where X and Y are typically proline and hydroxyproline. Here, we specifically focus on collagens that assemble into fibers with a triple-helix configuration. These types include I, II, III, V, XI, XXIV, and XXVI. The fibers are the mechanical load-bearing and tensile element for bony, cartilaginous, fibrous, and tubular structures of tissues[1]. Of these, type I is the predominant fibrillar collagen present in vertebrate animals. They are typically organized into arrays that align with mechanical tensile stress. For example, Col I is uniaxial in tendons and ligaments and organized across the plane in skin, cornea, and sclera[2,3]. Here, we will specifically focus on the role of Col I in the extracellular matrix (ECM) and the fiber morphology found in disease states.

The ECM functions to provide mechanical and biochemical support to tissues and cell activity. Within the ECM, collagen synthesis is performed by fibroblasts. Lysyl oxidases crosslink collagen for assembly into fibrils within the extracellular space. The assembly and downstream regulation is associated with integrins, fibronectin, actin, and discoidin domain receptors[2,4] The transition from normal ECM to pathological is often associated with the early stages of disease. In tumor masses, the ECM can comprise up to 60% of the tissue, usually dominated by the accumulation of collagen due to myofibroblast infiltration[5]. The underlying cell-matrix interactions have shown significant changes in local tumors and suggest mechanisms for metastasis[6].

1.3 Second Harmonic Generation Microscopy

SHG is a nonlinear second order coherent process where two lower energy photons are up-converted, emitting a photon at exactly twice the frequency of the incident excitation source. SHG contrast requires non-centrosymmetric assemblies, which is ideal for imaging well-ordered structures such as fibrillar collagen with triple helices[7]. Other structural proteins such as non-fibrillar or symmetric fibrillar collagen (e.g., Col IV), laminin, fibronectin and elastin are transparent by this modality as this criterion is not met[8].

1.3.1 Microscopy Setup

We followed SHG microscopy protocols described in Chen *et al*.[9] and Lien *et al*.[10]. Briefly, imaging was performed with laser scanning galvos (Cambridge Technologies, Bedford, MA, USA), coupled to an upright microscope (BX61; Olympus, Tokyo, Japan). The excitation source was a mode-locked Titanium Sapphire laser (Mira; Coherent, Santa Clara, CA, USA), providing 890 nm excitation. Laser scanning and data acquisition were conducted using custom LabVIEW

code and an FPGA interface board (National Instruments, Austin, Texas). The collected images were 512 × 512 pixels, representing a field of view of 180 × 180 μm. Power at the specimen was controlled by an electro-optic modulator (Conoptics, Danbury, CT, USA), with average power varying according to sample preparation, thickness, and type. Image acquisition time was 3 s per frame with three-frame Kalman averaging.

A 40 × 0.8 NA water immersion lens (LUMPlanFL/IF; Olympus, Tokyo, Japan) was used to focus the laser into the sample, and a 0.9 NA condenser collected the forward-propagating SHG signal. The lateral and axial resolution of the system was approximately 0.7 and 2.5 μm, respectively, sufficient for resolving collagen fibers. The forward-directed and backward-directed SHG emission was collected using identical photon-counting detectors (7421 GaAsP; Hamamatsu, Hamamatsu City, Japan), with the backward detector in the infinity space of the microscope. For each channel, the SHG wavelength (445 nm) was isolated with a dichroic mirror and 10 nm wide bandpass filter (Semrock, Rochester, New York). The excitation wavelength was confirmed using a fiber-optic spectrometer (Ocean Optics, Dunedin, FL, USA). Circular polarization was used for imaging, as this state excites all fiber orientations equally. This polarization of the excitation laser was determined at the focus by imaging dye labelled vesicles[9].

1.3.2 Biomimetic Models with SHG

Effective models for the ECM must include structural features of collagen, which can have significant effects on cell behavior. Properly modeling structure remains a significant challenge, especially given the relatively low availability of imaging data for a variety of diseases. Genetically engineered mouse models have shown the effects of specific mutations but remain inadequate to fully represent human disease at the cellular level[11]. ECM contains complex

morphology with intricate networks of different structures that are typically spatially heterogeneous[12].

SHG images can be fabricated into collagen scaffolds using multiphoton excited photochemistry. This technique produces submicron features on 3D topography into which cells can be seeded and tracked using phase-contrast microscopy[13]. Previous tissue models have used sinusoidal functions to mimic the collagen morphology commonly found in normal and diseased tissue. These models found that topography drove significant changes in cell shape, migration, F-actin alignment, and focal adhesion expression[14]. Since modifying the frequency and amplitude of the functions affects the cellular response, it follows that other structural features could have similar effects.

1.4 Machine Learning and Applications

As optics technologies and imaging techniques continue to grow and change, they provide previously unimaginable volumes of data. Analysis of this data has a vast amount of relevance to any number of biological inquiries such as quantification of cellular processes, examination of structural changes, and tracking of cancer cells. High-content and long-term monitoring of processes can result in a volume of data so great as to be cumbersome. Further, individual analysis may result in intrinsic biases or limit the possible recognition of essential features amongst the results[15]. Machine learning has emerged as a versatile tool to tackle these issues and to efficiently extract findings from high volume biological data.

Computational learning follows a set of statistical methods to classify data in either a supervised or unsupervised manner. Supervised learning begins with a human designer (generally an expert) specifically labeling examples of a given category, and then defining common features in the

data such as range of sizes, shapes, fluorescence intensity, or molecular markers[16]. Unsupervised learning, on the other hand, describes the process by which a fully automated AI can tread through a dataset, identifying classes and features using a series of complex and often convoluted networks of layers that analyze and transform the initial input[17]. Both approaches generally work by first delineating objects in an image via segmentation, detecting objects of interest within the image, spatially analyzing the pattern of segmented and/or detected objects, and in some cases, tracking the object if it is in motion across a defined space. Algorithms may tackle one or more of these tasks, and eventually, this process can reveal quantitative information on shapes, positions, spatial distributions, and statistical probability for a set of potential classifications[18].

Support vector machines (SVM) are one of the most commonly used supervised machine learning algorithms. Their straightforward implementation, logical statistical inferences, and well-studied pitfalls provide an excellent platform for user-friendly machine learning. The first step in training a SVM is to divide a large data into two subsets: a training set and a test set. An *m*-dimensional vector set defines the number of features that characterize each element. The training set is fully categorized by an expert so that the SVM can build an *m* sized hyperplane that creates an ideal margin between nearest-neighbor vector elements. The optimal width is usually determined by minimizing the mean cost of the classification with a cost function: a function that maps one or more variables onto a real number that can be minimized to reduce difference between an estimation and true value. By building these margins, new data can be categorized based on its apparent fit to this feature map (Figure 1.1). Although generally a binary (0 or 1) classifier, SVMs can be modified in a one-vs.-all method to fit data into multiple classes[16]. SVM is often limited to classification that is straightforward with obvious features as its accuracy in multi-class categorization decreases more markedly than the accuracy of other

algorithms[19]. However, SVM is a powerful tool that can be swiftly applied to reduce labor and recognize patterns in high output imaging.

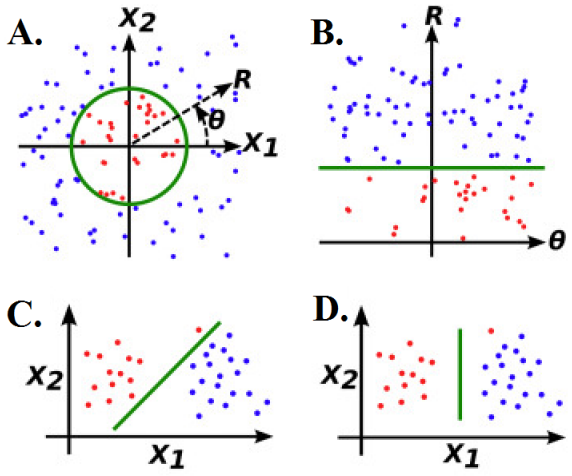


Figure 1.1: Examples of SVM classifying boundaries and margins used to identify red and blue binary classifications. (**A**) Example of SVM's capability for non-linear boundaries. (**B-D**) Boundaries for varying datasets fit to unspecified features. *Modified from Wang and Fernandez-Gonzalez*[16].

Deep learning refers to the use of a multi-layered neural network and is a technique that can be used for either semi-supervised or fully unsupervised learning. The network begins with an input layer, simply the original data in an easily manipulatable form, which is then transformed via a number of hidden layers, each of which has a width defined by the number of neurons (elementary units) in each layer, until a final transformation populates an output layer with predicted probabilities for each category that the network is designed to recognize or create[20]. Weights are parameters corresponding to neurons, which are defined through the minimization of a loss function after training on data. The loss function differs from a cost function through its high-dimensional and non-convex complexity. The simplest methodology employs backward propagation to efficiently train the algorithm with gradient-based descent (Figure 1.2)[21]. One of the most commonly used algorithms in complex biological problems modifies this relatively

simple but computationally expensive method by using multidimensional arrays as the input data. These convolutional neural networks work by convolving each layer with a set of kernels and added biases to generate a new feature map. A non-linear transform to create pooling layers on this map will reduce the total number of parameters in the network[17]. This technique allows learning of large numbers of abstract features while remaining computationally reasonable[22]. Receiver operating characteristic curves are usually used as a measure for classifiers' performance. They are based on a plot of sensitivity versus specificity for diagnostic tests. These measures allow the adjustment of the rate of false positives and false negatives by moving decision thresholds up or down the curve. Changing the operation point on the curve is also associated with the overall accuracy of the test[23].

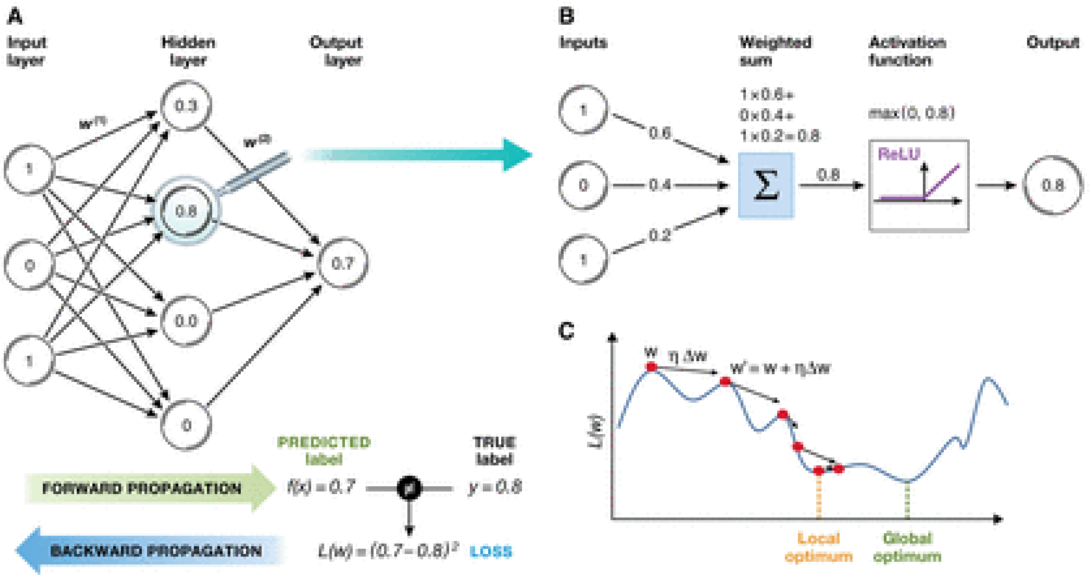


Figure 1.2: A sample neural network generation with depth of one. (**A**) The network is drawn out with all neurons interconnected between non-output layers. The algorithm works first by feed-forward calculation followed by backpropagation to recalculate each weight. (**B**) A sample weighted sum leading to the final algorithm's output. (**C**) Hypothetical example of a feature's valley-rich optimization path demonstrating potential difficulties with modification of learning rate to avoid stalling in local minima. *Modified from Angermueller et al.*[21].

1.4.1 Applications

Phenotyping

High content screening represents one of the most obvious applications of machine learning. Classification of cells through phenotyping requires a combined automation of microscopy with high-throughput biotechnology. Presently, optics technologies have the potential to capture 10^5 images per day. With such large volumes of data, automation can be necessary for feature and data extraction for each assay[24]. Due to their speed, convolutional neural networks are the current favored choice in cellular analyses[25]. In cases where only a subset of cells can be labeled with known classes, multiple instance learning can be used to pool new classifications into the hidden layers of the original network. This allows the algorithm to be down classed from semi-supervised to weakly-supervised[24]. Phenotyping has been used for live cell classification, drug behavior tracking, and organelle observations[26]. Deep learning networks have been shown to surpass human accuracy in classification of proteins, cells, and drug mechanisms[27], even aside from the obvious advantages in terms of speed.

Live imaging presents a challenge in analysis but opens the possibility of introducing temporally-resolved phenotyping. As an example, SAPHIRE (Stochastic Annotation of PHenotypic Individual-cell REsponses) was developed to characterize single-cell shape dynamics across a large variety of assays. This software segments and tracks cells based on concurrent actin and nuclear dynamics imaged with fluorescent reporters LifeAct-eGFP and histone H2B-mCherry. Shape descriptors are then fit into a tensor with shape features, individual cells, and time points as its dimensions. This temporal shape space can be fit to classification parameters and used to identify state-space dynamics. Finally, the data is clustered and fully

classified in the output. SAPHIRE is uniquely programmed to detect diversification of morphology following drug treatment or other manipulated dynamics[28]. Cellular heterogeneity can be detected with downstream visualization of features as well as statistical measures of Euclidean, Mahalanobis or Manhattan distance and similarity such as Pearson's rank correlation[29].

Microfluidics and Moving Morphology

The temporal dimension introduces the ability to image both alteration of cellular morphology and the potential for changing spatial location. Flow cytometry is a high-throughput analysis wherein cells are siphoned through a fluid stream. Generally, data consists of a recording of fluorescent events that occur when fluorochromes in the fluid stream are subject to excitement via laser[30]. The challenge, in terms of the image classification task in this situation, is to design a multiple-object-tracking algorithm that can continue to accept input as the cells change and move across the image[31]. One strategy to overcome this challenge is through time stretch quantitative phase imaging. This is a form of mapping that recognizes spatial information as a time-stretched dispersive Fourier transform. This technique can capture up to 36 million images per second in flow rates of 10 meters per second. Combined with a neural network that is trained over a receiver operating characteristic curve on algal cells based on their lipid contents, mean accuracies at or above 85% for a variety of features have been achieved (Figure 1.3)[32].

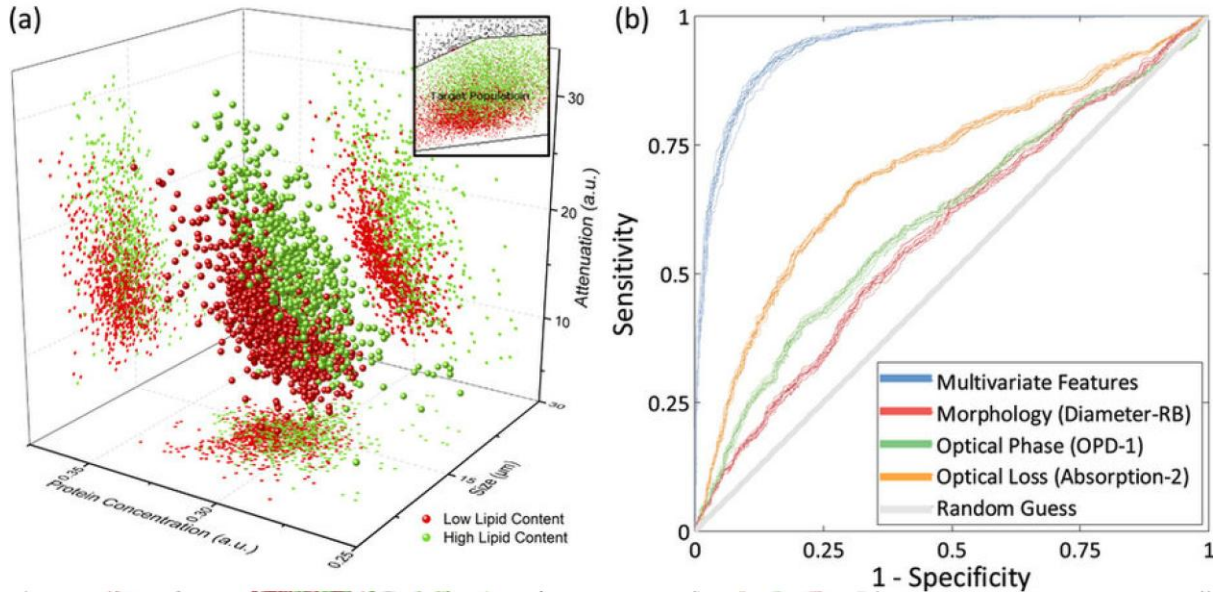


Figure 1.3: Classification of algal cells by lipid content based on a variety of features. (**a**) A 3D plot of cells based on size, protein concentration and attenuation. (**b**) Receiver operating characteristics curves based on all features (blue) versus specific biophysical features. *Modified from Chen et al.*[32].

Optical neural imaging has advanced to provide high content data of deforming and mobile neurons. Although it is possible to individually segment and identify each neuron of interest, a label-free automated algorithm can provide more feasible analysis. One approach that has been used in research on *Caenorhabditis elegans* is non-rigid point-set registration to create vectors of features that relate nearest-neighbor neurons that can be clustered. These feature vectors can work time-independently to track the individual neuron cells A thin-plate spline interpolation is used to correct aberrations and other errors in imaging and segmenting. Over an 8 minute recording, 156 neurons were successfully tracked and classified[33]. Convolution neural networks can also tackle this challenge via 3D imaging and voxel classification. Training with open source labeled images from BigNeuron allowed an algorithm to map neurons by examining and assigning a binary classification to each voxel individually[34].

Algorithms that can sort and track individual motile cells can prove useful in a variety of situations. Interferometric phase microscopy is currently used as a screening method for in vitro fertilization to examine sperm head morphology. A method to automate this analysis can help rid bias and introduce quantitative measures to improve the process. Using 1400 sperm cells, an algorithm was built to track and record 3D image stacks of sperm heads as well as generate features to describe their morphology. These features were then fed into an SVM classifier to sort morphology deemed good and bad. The results had approximately 90% precision and achieved an area 88.59% under the receiver operating characteristic curve[35]. Tracking changes in cellular motion can also lead to diagnoses of pathology associated with abnormal morphology. Microfluidics can be used in conjunction with a convolutional neural network to train on leukemia cell lines. Morphological classification can achieve an accuracy above 90%, and with all features, the network can operate above 95% accuracy[36].

Cancer Diagnostics

Early detection of cancer has long been established as greatly enhancing the chances for successful treatment. Optics technologies continue to reveal new markers for different types of cancer, allowing for the possibility of utilization of an AI to diagnose patients using clinical imaging modalities. This could significantly reduce diagnostic errors, which have been found to contribute to 10% of patient deaths and 6-17% of adverse events in hospitals[37].

Uncontrolled growth of pigment cells is the most serious form of skin cancer. There has been a noted increase in the prevalence of this type of melanoma, especially in young Australians. Skin lesion classification can be tackled with a bag-of-features model that learns high-level representation and maps it to a deep neural network. This algorithm creates independent local

descriptors and maps them to a histogram vector. This histogram contains up to 1000 classes of visual descriptor words that are evaluated for frequency across the image. The neural network's output is then filtered into a new SVM for final classification. Over a test set of 244 images, this model performed with 95% accuracy[38]. Convolutional neural networks can also be pre-training on labeled data sets. A network was trained on 6120 colored images with two convolving layers with a five by five kernel. A total of 70 features were weighed over four layers of convolution and pooling. This methodology achieved an accuracy of 81% while remaining computationally reasonable with a relatively smaller network width[39]. Cancer detection for basal cell carcinoma also has the potential to benefit from deep learning algorithms. Using histopathological images, an unsupervised feature learner was trained via autoencoders. Autoencoders learn features by modelling the output to be as similar as possible to input while minimizing an overall cost function. These features can then be fed into a convoluting and pooling neural network to map features onto the image arrays. Finally, a softmax classifier uses logistic regression (again minimizing an overall cost function) for binary classifications. With 400 features and a pooling layer dimension of 71, the algorithm achieved accuracies between 79% and 92% based on the type of training data utilized[40].

Other types of cancer require optical techniques that are less simplistic than surface imaging. Lung cancer provides a particularly challenging diagnostic problem as the tissue is deep, yet imaging must be swift and accurate, and cannot utilize harmful contrast agents or biodegradation. Coherent anti-Stokes Raman scattering has been used as a diagnostic tool for lung cancer as an emerging label-free imaging technique. Two laser beams fired at the target sample are used to match the frequency difference of the pump field and the Stokes field to the vibrational frequency of Raman molecules in the sample. Large volumes of real-time submicron

resolution data from intrinsic molecular vibrations can be used to provide information on biological and chemical properties of samples[41]. An adaptation of the GoogleNet Inception v3 convolutional neural network was created, using 22 stacked layers[42]. The output set was modified from the 1000 ImageNet classes to four possible lung diagnoses[43]. This modification of generalized object detection was retrained on 512 labeled images and had an overall accuracy of 89.2%. Further specialization to separate adenocarninoma and squamous cell carcinoma was able to achieve a 97% accuracy[44].

Cryo-electron Microscopy

Structural biology has recently seen an abundance of research due largely to the advancement of single-particle electron cryo-microscopy (cryo-EM). The 2017 Nobel Prize in Chemistry was awarded to the developers of cryo-EM technology. It has the capabilities to push biochemistry into a new era with high-resolution images of 3D protein structures. Particle selection from micrographs is required to generate these high-quality images. Particle picking is a time-intensive and laborious task with easily introduced human bias and inconsistency. As it relies on the selection of specific particles, this is an excellent task for machine learning. There are several potential strategies to address this data analysis bottleneck[45].

DeepPicker is a software package that attempts to address the issue with a deep learning framework. The convolutional neural network operates in five steps: scoring, cleaning, filtering, sorting, and iteration. The micrographs are scanned to achieve a standardized box size corresponding to a value slightly larger than the particles. This allows the network to score the full map; that is, it distributes likelihood scores for the full spectrum of the micrograph. The cleaning step discards noisy data from false particles by comparing nearest neighbors to the

previously determined boxing threshold. Filtering further refines the particle candidates and assigns coordinates to the predicted centers. The particles are then sorted by their ultimate prediction scores. Finally, the algorithm iterates the classifier itself to refine the ultimate neural network. DeepPicker can use a stochastic gradient descent to train on existing image sets to operate in a fully automated manner, or users can input ~400 particles in a semi-supervised learning environment. DeepPicker successfully identifies particles that experts may have picked with an 81% recall score. It also avoids a possible issue with identifying noise as a particle. In the sorting step it keeps these ice-generated noise artifacts at a low prediction level, and thus is a "low-quality" particle that will not be selected. Implementation of the final software was based on Torch7, an open source deep learning code library[46].

Cryo-EM has attracted other attempts to apply deep convolutional neural networks. DeepEM differs by specifically setting eight layers for the network while still recursively training the algorithm. The training set for DeepEM was quadrupled by feeding the model the same dataset four times after rotating the dataset at 90-degree intervals. Due to irregularities in cryo-EM such as protein impurities, ice artifacts, overlapping samples, and carbon-rich areas, multiple micrographs are needed for training in order to account for the possibilities of negative particle images. This methodology ultimately resulted in a very similar recall score of 80%. The final program was implemented in a hardware based graphics card accelerated version of Matlab to increase computational efficiency[47].

Beyond particle picking, cryo-EM can benefit from structure determination to assess heterogeneity. A statistical manifold learning framework known as generative topographic mapping can be used to classify datasets. This is a variation of neural networks that derives a low-dimensional sub-manifold from high-dimensional data induced by smooth mapping. A sub-

manifold represents a subset of the represented topological space that locally resembles Euclidean space. The alteration to these schemata for cryo-EM is accomplished with a contrast transfer function which efficiently corrects microscope aberration effect in Fourier space. The model was tested with a sample group of 3961 classified particles, wherein an 88.7% accuracy was achieved. As this software also attempted to accomplish its network in the relatively short time of 143 minutes, this accuracy could be increased with added iterations and computational time. The algorithm was implemented in ROME, an open source software that builds on the functionality of algebraic modeling in Matlab[48,49].

Machine Learning Software

As machine learning has become a significant tool in image analytics, numerous versions of highly-used algorithms have appeared in research. The developed algorithms found success when finely tuned to specific tasks set by the researchers. However, this leaves little room for development on new experiments, even those closely related, as the code is often highly specialized, especially in academic environments[15]. The need for intelligent generalized solutions in software is clear as more and more tasks are relegated to computational processing[18]. To that end, there are several software packages geared towards biological image analysis as well as generalized machine learning software aimed to create user-friendly tools.

CellProfiler Analyst focuses on high content cellular images by quantifying phenotypes in a supervised learning dataset. Its classifier works in tandem with visualization options for field-of-view and cell classification[50]. The software handles multi-class modelling with algorithms from scikit-learn such as Random Forest, AdaBoost, and standardized SVM learning. Random Forest perturbs and combines randomized decision trees, introducing randomness into the classifier

construction to reduce initial specification bias. However, because the best split of the tree is picked from a random subset of features the overall forest statistical bias increases, but the averaging causes variance to decrease. Adaboost is a meta-classifier in that it creates additional copies of an initial classifier but alters the weights progressively to focus on difficult learning cases that may have been initially classified incorrectly[51]. CellProfiler has seen a wide variety of use in research. A recent example tackled the interaction of lipid-laden macrophages and vascular smooth muscle cells using nuclei morphology as the primary classification feature. Typical approaches to characterizing these interactions are wrought with difficulties in fluorescence and membrane permeability. The results yielded a 61.3% accuracy compared to individual analysis with ImageJ/Fiji, showing the potential problems with applying a generalized learning algorithm to a specific scientific inquiry[52].

Advanced Cell Classifier (ACC) is another graphically focused package for large-scale phenotype sorting in cell experiments. Increased volume of high content screening have caused the feasibility of visual interpretation and verification to become highly inefficient. ACC focuses on automation of analysis with minimal user interaction. To that end, ACC methodology relies on semi-supervised learning wherein classes of interest can be defined by the user, but final feature-based statistics are a combination of user and machine input. ACC can utilize the algorithms from CellProfiler Analyst as well as the more computation-reliant multilayer perceptron and logistic boost classifiers[53]. Multilayer Perceptron classifiers are basic feedforward artificial neural networks that learn based on backpropagation and may bridge the gap into deep learning with more than one hidden layer. Logistic boost is a direct successor of Adaboost, in that it uses convex optimization, or logistic regression, to improve each iteration of the meta-classifier[54].

Full automation of cell research, let alone generalized biological models, is not within reach with current technology. However, new frameworks for deep learning are providing opportunities to grow research in this direction. TensorFlow is an open-source software specifically geared towards user-friendly multi-layer convolutional neural networks. Complex learning algorithms are typically exclusively successful with experts in the field so software such as TensorFlow aims to broaden the audience of potential users. A potential problem with these methods is the need for high-capability hardware as well as the requirement for parallelization in order to achieve feasible run times[55]. Complex computational biology questions can certainly benefit from the deep learning approach, and more studies will utilize these tools as automation becomes easier to use and open source software continues to develop. Beyond microscopy, image object detection has grown into fast, real-time, robust frameworks. YOLO9000 is a generalized classifier that works at 67 FPS with more than 9000 classes based on the ImageNet labelled data set[43]. This advancement is possible due to multi-scale training, convolutional boxing, and dimension clusters[56]. The potential for similar training algorithms in well-studied biological tissue could lead to advances in real-time microscopy analyses on live cultures.

Other Applications

Machine learning has the potential to impact many large-volume image analyses. However, it can also be used in a variety of more unique inquiries. Its sorting, feature-recognizing, and classification potential can be applied to problems that remain unconfirmed or simply too nonlinear to approach reasonably in other analytics. Second harmonic generation (SHG) has been established as an excellent tool for imaging collagen fibers in tissue samples. Collagen plays a key role in structural support and is the most abundant protein in mammals[57]. Considering this role, it is important to estimate collagenous elastic properties. The usual method

to quantify mechanical properties is through a stress-strain curve, achieved from direct tissue testing[58]. However, estimating the tissue properties directly from microscopy images could help medical diagnoses and identify problems in morphology or other connective tissue issues. A deep learning approach was used on a series of SHG images from glutaraldehyde-treated bovine pericardium tissue. A convolutional neural network was created with a total of 18 layers and implemented in Matconvnet which is a customizable opensource Matlab toolbox[59]. The fully trained network achieved an accuracy of 84% for the final nonlinear anisotropic stress-strain curve (Figure 1.4)[60].

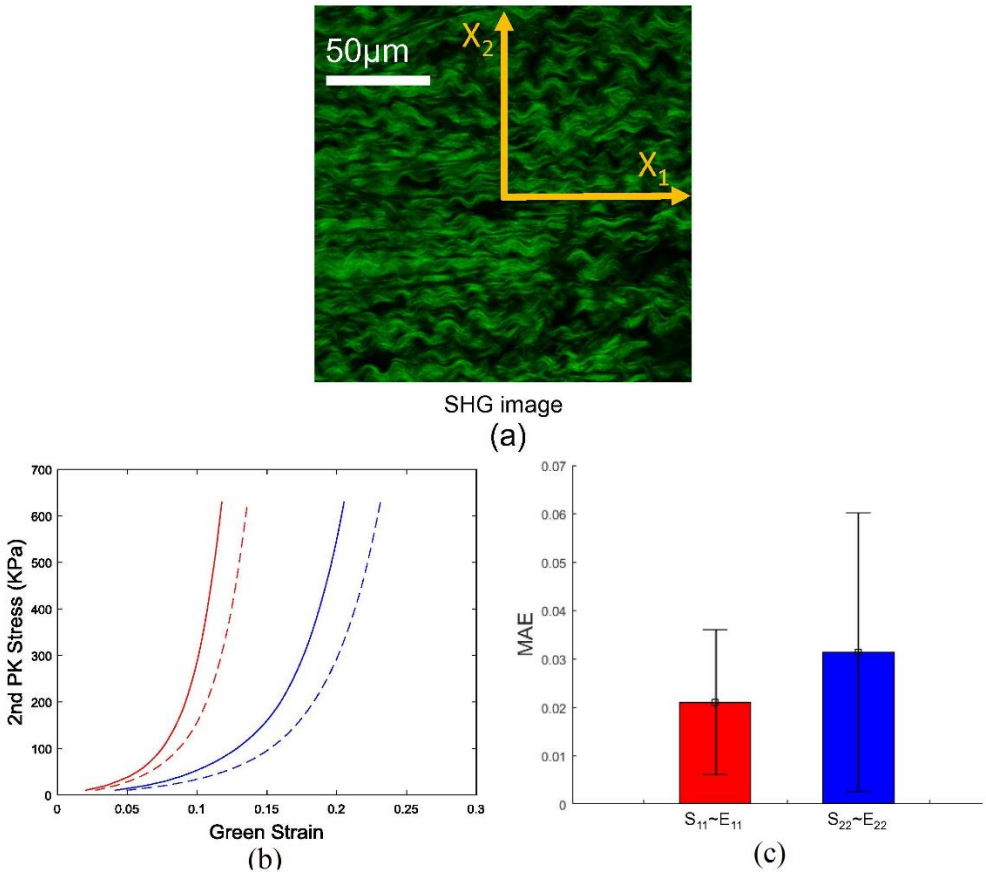


Figure 1.4: Representative SHG image compared to mechanically measured and analytical stress-strain curve. (a) SHG generated image of a collagen fiber network in the pericardium. (b) Stress-strain curves from mechanical testing (bold) and deep learning estimation (dashed) for the *xx* (11) direction (red) and *yy* (22) direction (blue). (c) The mean absolute error distribution. *Modified from Liang et al.*[60].

Deep learning can also be applied to cell-level interactions that may dictate whole animal behavior. Parasite mediated diseases such as toxoplasmosis are found worldwide in humans and other animals. *Toxoplasma gondii* has well studied structure and development, but the overall life cycle can be studied further with other animal models[61]. Microbial parasites have adapted to manipulate behavior in a variety of animal hosts, usually via gene expression through the new host cells. These phenotypes are not currently well studied on a micrometer level[62]. The fungal parasite *Ophiocordyceps unilateralis sensu lato* has been shown to interfere with its host, the carpenter ant. After 16-25 days of growth the parasite influences mandible muscles resulting in atrophy and also seems to cause the ant to bite a plant at a specific elevation and location that is optimal for fungal growth and dispersal. Ants at different stages of fungal growth were imaged with a scanning-electron microscope to observe the 3D structure and distribution of cells. Ants were also infected with a non-behavior manipulating fungus as a positive control for final classification. A deep convolutional network was used to segment and label between fungus and muscle cells, adapted from previous work in biomedical image analysis. The modified U-Net is specifically geared towards low training samples by supplementing with successive layers for upsampling operators. Convolution layers are then able to operate at higher precision and resolution[63]. The final classification model was able to achieve accuracy of 93% compared to experts' labeling. This network was able to identify key communicating fungal networks created in the ant body as well as direct filamentous mycelia invasion of muscle fibers[64].

The Campagnola lab has employed a variety of machine learning techniques to classify alterations in normal collagen structure. We previously used a texton approach to classify a series of tissues related to ovarian cancer. Texture features were identified using image responses to

sets of filter banks, and a representative model was then generated for the tissue types based on the distribution of the respective features in each class[65]. Ultimately, this method was able to achieve a classification accuracy >90% using one vs. the rest classification. Analogously, we used principal component analysis of 2D wavelet transforms of idiopathic pulmonary fibrosis and classified normal vs. fibrotic disease with >95% accuracy. While successful, these classifiers are a "black box," and provide no insight into the identification of the most important and differentiating aspects of each class of tissue.

In recent years, generative adversarial networks (GAN) have been used in a variety of applications for microscopy. However, they remain limited by a lack of training data. Generating new synthetic images that are highly similar to real images could address this issue. Even where data is available, it is not always usable: in some modalities, resolution can be weak, or images can be noisy. GANs can be used to enhance such images by adjusting sharpness, noise, and clarity. By training GANs on labeled data, they can be used to automatically segment datasets by identifying characteristic features such as tumors[66–68]. In addition to these tasks, microscopy modalities can use GANs for virtual staining, generating realistic H&E, Masson's Trichrome, and Jones' images from unlabeled data. Similarly, digital fluorescence labeling can be performed on brightfield or phase contrast images (Figure 1.5)[69,70]. Several approaches have achieved significant results including StyleGAN, CycleGAN, SPADE GAN, and HingeGAN[71]. Typically, these generators are evaluated based on the Fréchet inception distance (FID), which measures the feature vector distance between synthetic and real images. Style-based GANs have achieved particularly good FIDs, indicating successful generation of quality image data reflective of that available in microscopy[72]. Further details on GANs and their architecture is located in Chapter 3.

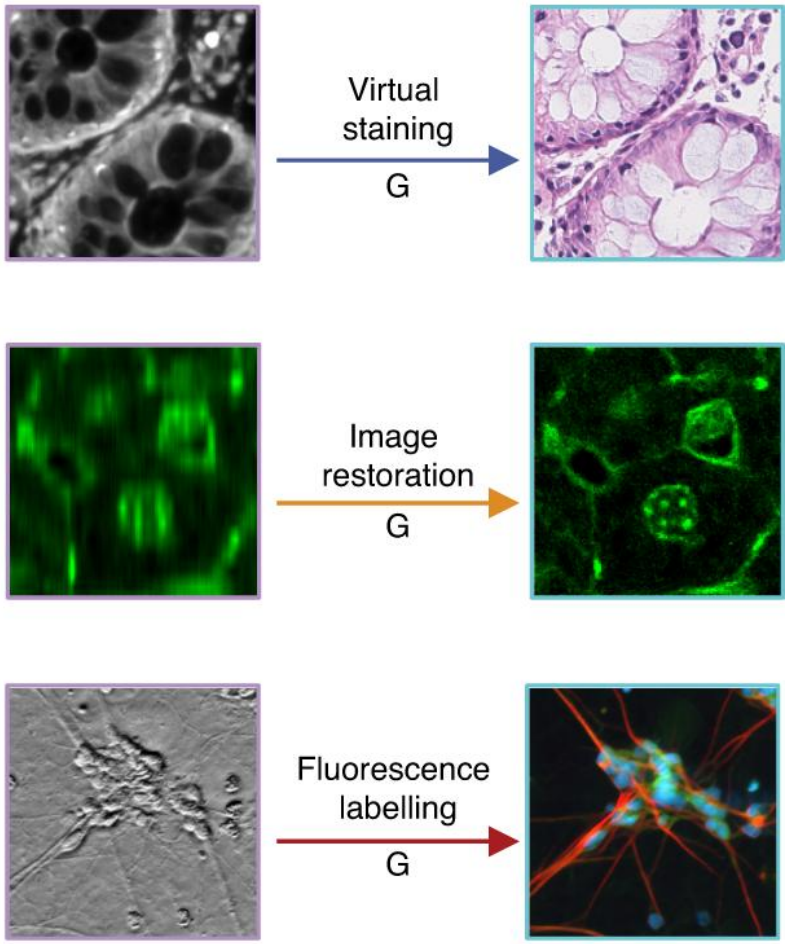


Figure 1.5: Examples of uses for GANs in microscopy. H&E stains can be mimicked, image resolution can be restored, and fluorescence labelling can be applied to greyscale images. Adapted from Li et al., *Light: Science & Applications* 2021.

1.4.2 Conclusions

Machine learning is not necessarily wholly unbiased as specific thresholds, weights, and features may be modified by the user. Therefore it is not sufficient for statistical analysis of problems and data by itself until it has undergone rigorous testing[19]. Additionally, academic work in deep learning has typically yielded test accuracies of 60-93% using optimized algorithms. Although this may be sufficient for proof of concept, these low accuracies cannot be used in clinical diagnostics or even in less critical phenotyping assays. Google's software for object classification generally aims for 90+% accuracy before optimization, showing the potential for

stronger image classification in future work[42]. For clinical diagnostics, accuracy is less important than adjusting the false negative rate by moving the operating point on the receiver operating characteristic curve. Although balancing sensitivity and specificity may improve numbers and seemingly improve results, there must be ethical consideration for patients and their potential pathologies. Ensuring the best outcome for patients means reducing false negative results. Although this may skew the curve to a high degree of false positives, expert pathologists generally have excellent records for low false positive rates and can work in complement to the diagnostic tools developed from machine learning. Recently, cancer diagnostic algorithm competitions have been bringing industry and academia together towards achieving higher accuracy with lower false positives helping to address some of these potential pitfalls[73]. Finally, many current tools available require advanced hardware and long computation times. This can be improved via simpler methods of parallelization and more efficient approaches to neural networks.

Machine learning and deep learning methods are a powerful complement to medical imaging and microscopy. These tools can aid in a large variety of laborious analyses as well as mine important features from data that may not be visible to human observers. Generalized and user-friendly software is creating platforms to apply these tools across many disciplines and laboratories. Generalization may prove to have lower accuracies, but as an initial discovery tool, it is still impressively useful. As biomedical engineering and computer science continue to grow together, there will be many exciting opportunities to apply automation and discovery with these and future learning tools.

Chapter 1 References

1. Holmes, D. F., Lu, Y., Starborg, T. & Kadler, K. E. Collagen Fibril Assembly and Function. *Curr. Top. Dev. Biol.* **130**, 107–142 (2018).
2. Siadat, S. M. & Ruberti, J. W. Mechanochemistry of collagen. *Acta Biomater.* **163**, 50–62 (2023).
3. Bella, J. & Hulmes, D. J. S. Fibrillar Collagens. *Subcell. Biochem.* **82**, 457–490 (2017).
4. De Martino, D. & Bravo-Cordero, J. J. Collagens in Cancer: Structural Regulators and Guardians of Cancer Progression. *Cancer Res.* **83**, 1386–1392 (2023).
5. Hu, M., Ling, Z. & Ren, X. Extracellular matrix dynamics: tracking in biological systems and their implications. *J. Biol. Eng. 2022 161* **16**, 1–13 (2022).
6. Karamanos, N. K. *et al.* Extracellular matrix-based cancer targeting. *Trends Mol. Med.* **27**, 1000–1013 (2021).
7. Campagnola, P. J. & Loew, L. M. Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms. *Nat. Biotechnol.* **21**, 1356–1360 (2003).
8. Mostaço-Guidolin, L., Rosin, N. L. & Hackett, T.-L. Imaging Collagen in Scar Tissue: Developments in Second Harmonic Generation Microscopy for Biomedical Applications. *Int. J. Mol. Sci.* **18**, 1772 (2017).
9. Chen, X., Nadiarynkh, O., Plotnikov, S. & Campagnola, P. J. Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure. *Nat. Protoc.* **7**, 654–669 (2012).
10. Lien, C.-H., Tilbury, K., Chen, S.-J. & Campagnola, P. J. Precise, motion-free polarization control in Second Harmonic Generation microscopy using a liquid crystal modulator in the infinity space. *Biomed. Opt. Express* **4**, 1991–2002 (2013).
11. Russo, A. *et al.* PTEN loss in the fallopian tube induces hyperplasia and ovarian tumor formation. *Oncogene 2018 3715* **37**, 1976–1990 (2018).
12. Kular, J. K., Basu, S. & Sharma, R. I. The extracellular matrix: Structure, composition, age-related differences, tools for analysis and applications for tissue engineering. *J. Tissue Eng.* **5**, (2014).
13. Atry, F. *et al.* Parallel multiphoton excited fabrication of tissue engineering scaffolds using a diffractive optical element. *Opt. Express* **28**, (2020).
14. Alkmin, S. *et al.* Migration dynamics of ovarian epithelial cells on micro-fabricated image-based models of normal and malignant stroma. *Acta Biomater.* **100**, (2019).
15. Eliceiri, K. W. *et al.* Biological imaging software tools. *Nat. Methods* **9**, 697–710 (2012).
16. Wang, M. F. Z. & Fernandez-Gonzalez, R. (Machine-)Learning to analyze in vivo microscopy: Support vector machines. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1865**, 1719–1727 (2017).
17. Litjens, G. *et al.* A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**, 60–88 (2017).

18. Sbalzarini, I. F. Seeing Is Believing: Quantifying Is Convincing: Computational Image Analysis in Biology. in *Advances in anatomy, embryology, and cell biology* vol. 219 1–39 (2016).
19. Kan, A. Machine learning applications in cell image analysis. *Immunol. Cell Biol.* **95**, 525–530 (2017).
20. Kraus, O. Z. & Frey, B. J. Computer vision for high content screening. *Crit. Rev. Biochem. Mol. Biol.* **51**, 102–109 (2016).
21. Angermueller, C., Pärnamaa, T., Parts, L. & Stegle, O. Deep learning for computational biology. *Mol. Syst. Biol.* **12**, 878 (2016).
22. Suzuki, K. Overview of deep learning in medical imaging. *Radiol. Phys. Technol.* **10**, 257–273 (2017).
23. Dendumrongsup, T. *et al.* Multi-Reader Multi-Case Studies Using the Area under the Receiver Operator Characteristic Curve as a Measure of Diagnostic Accuracy: Systematic Review with a Focus on Quality of Data Reporting. *PLoS ONE* **9**, e116018 (2014).
24. Kraus, O. Z., Ba, J. L. & Frey, B. J. Classifying and segmenting microscopy images with deep multiple instance learning. *Bioinformatics* **32**, i52–i59 (2016).
25. Shariff, A., Kangas, J., Coelho, L. P., Quinn, S. & Murphy, R. F. Automated Image Analysis for High-Content Screening and Analysis. *J. Biomol. Screen.* **15**, 726–734 (2010).
26. Nketia, T. A., Sailem, H., Rohde, G., Machiraju, R. & Rittscher, J. Analysis of live cell images: Methods, tools and opportunities. *Methods* **115**, 65–79 (2017).
27. Grys, B. T. *et al.* Machine learning and computer vision approaches for phenotypic profiling. *J. Cell Biol.* **216**, 65–71 (2017).
28. Gordonov, S. *et al.* Time series modeling of live-cell shape dynamics for image-based phenotypic profiling. *Integr. Biol. Quant. Biosci. Nano Macro* **8**, 73–90 (2016).
29. Caicedo, J. C. *et al.* Data-analysis strategies for image-based cell profiling. *Nat. Methods* **14**, 849–863 (2017).
30. Urish, K. L., Deasy, B. M. & Huard, J. Automated classification and visualization of fluorescent live cell microscopy images. *J. Microsc.* **249**, 206–214 (2013).
31. Heo, Y. J., Lee, D., Kang, J., Lee, K. & Chung, W. K. Real-time Image Processing for Microscopy-based Label-free Imaging Flow Cytometry in a Microfluidic Chip. *Sci. Rep.* **7**, 11651 (2017).
32. Chen, C. L. *et al.* Deep Learning in Label-free Cell Classification. *Sci. Rep.* **6**, 21471 (2016).
33. Nguyen, J. P., Linder, A. N., Plummer, G. S., Shaevitz, J. W. & Leifer, A. M. Automatically tracking neurons in a moving and deforming brain. *PLOS Comput. Biol.* **13**, e1005517 (2017).
34. Li, R., Zeng, T., Peng, H. & Ji, S. Deep Learning Segmentation of Optical Microscopy Images Improves 3-D Neuron Reconstruction. *IEEE Trans. Med. Imaging* **36**, 1533–1541 (2017).

35. Mirsky, S. K., Barnea, I., Levi, M., Greenspan, H. & Shaked, N. T. Automated analysis of individual sperm cells using stain-free interferometric phase microscopy and machine learning. *Cytometry A* **91**, 893–900 (2017).
36. Gopakumar, G., Hari Babu, K., Mishra, D., Gorthi, S. S. & Sai Subrahmanyam, G. R. K. Cytopathological image analysis using deep-learning networks in microfluidic microscopy. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **34**, 111–121 (2017).
37. National Academies of Sciences, Engineering, and M. *Improving Diagnosis in Health Care*. (National Academies Press, Washington, D.C., 2015). doi:10.17226/21794.
38. Sabbaghi, S., Aldeen, M. & Garnavi, R. A deep bag-of-features model for the classification of melanomas in dermoscopy images. in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* vol. 2016 1369–1372 (IEEE, 2016).
39. Nasr-Esfahani, E. *et al.* Melanoma detection by analysis of clinical images using convolutional neural network. in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* vol. 2016 1373–1376 (IEEE, 2016).
40. Cruz-Roa, A. A., Arevalo Ovalle, J. E., Madabhushi, A. & González Osorio, F. A. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. *Med. Image Comput. Comput.-Assist. Interv. MICCAI Int. Conf. Med. Image Comput. Comput.-Assist. Interv.* **16**, 403–10 (2013).
41. Gao, L. *et al.* Differential diagnosis of lung carcinoma with three-dimensional quantitative molecular vibrational imaging. *J. Biomed. Opt.* **17**, 066017 (2012).
42. Szegedy, C. *et al.* Going deeper with convolutions. in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* vols 07-12-June 1–9 (2015).
43. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1* 1097–1105 (2012).
44. Weng, S., Xu, X., Li, J. & Wong, S. T. C. Combining deep learning and coherent anti-Stokes Raman scattering imaging for automated differential diagnosis of lung cancer. *J. Biomed. Opt.* **22**, 1 (2017).
45. Voss, N. R., Yoshioka, C. K., Radermacher, M., Potter, C. S. & Carragher, B. DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. *J. Struct. Biol.* **166**, 205–13 (2009).
46. Wang, F. *et al.* DeepPicker: A deep learning approach for fully automated particle picking in cryo-EM. *J. Struct. Biol.* **195**, 325–336 (2016).
47. Zhu, Y., Ouyang, Q. & Mao, Y. A deep convolutional neural network approach to single-particle recognition in cryo-electron microscopy. *BMC Bioinformatics* **18**, 348 (2017).
48. Wu, J. *et al.* Massively parallel unsupervised single-particle cryo-EM data clustering via statistical manifold learning. *PLOS ONE* **12**, e0182130 (2017).

49. Goh, J. & Sim, M. Robust Optimization Made Easy with ROME. *Oper. Res.* **59**, 973–985 (2011).
50. Dao, D. *et al.* CellProfiler Analyst: interactive data exploration, analysis and classification of large biological image sets. *Bioinformatics* **32**, 3210–3212 (2016).
51. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
52. Roeper, M., Braun-Dullaeus, R. C. & Weinert, S. Semiautomatic High-Content Analysis of Complex Images from Cocultures of Vascular Smooth Muscle Cells and Macrophages: A CellProfiler Showcase. *SLAS Discov. Adv. Life Sci. RD* **22**, 837–847 (2017).
53. Piccinini, F. *et al.* Advanced Cell Classifier: User-Friendly Machine-Learning-Based Software for Discovering Phenotypes in High-Content Imaging Data. *Cell Syst.* **4**, 651-655.e5 (2017).
54. Friedman, J., Friedman, J., Hastie, T. & Tibshirani, R. Additive Logistic Regression: a Statistical View of Boosting. *Ann. Stat.* **28**, 2000 (1998).
55. Rampasek, L. & Goldenberg, A. TensorFlow: Biology's Gateway to Deep Learning? *Cell Syst.* **2**, 12–14 (2016).
56. Carroll, J. YOLO9000: Better, faster, stronger. *Eur. Pharm. Contract.* 30 (2014) doi:10.5860/lrts.53n4.261.
57. Mostaço-Guidolin, L., Rosin, N. & Hackett, T.-L. Imaging Collagen in Scar Tissue: Developments in Second Harmonic Generation Microscopy for Biomedical Applications. *Int. J. Mol. Sci.* **18**, 1772 (2017).
58. Sherman, V. R., Yang, W. & Meyers, M. A. The materials science of collagen. *J. Mech. Behav. Biomed. Mater.* **52**, 22–50 (2015).
59. Vedaldi, A. & Lenc, K. MatConvNet - Convolutional Neural Networks for MATLAB. *Proc. 23rd ACM Int. Conf. Multimed.* 689–692 (2015) doi:10.1145/2733373.2807412.
60. Liang, L., Liu, M. & Sun, W. A deep learning approach to estimate chemically-treated collagenous tissue nonlinear anisotropic stress-strain responses from microscopy images. *Acta Biomater.* **63**, 227–235 (2017).
61. Dubey, J. P., Lindsay, D. S. & Speer, C. A. Structures of Toxoplasma gondii tachyzoites, bradyzoites, and sporozoites and biology and development of tissue cysts. *Clin. Microbiol. Rev.* **11**, 267–99 (1998).
62. de Bekker, C. *et al.* Species-specific ant brain manipulation by a specialized fungal parasite. *BMC Evol. Biol.* **14**, 166 (2014).
63. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Miccai* 234–241 (2015) doi:10.1007/978-3-319-24574-4_28.
64. Fredericksen, M. A. *et al.* Three-dimensional visualization and a deep-learning model reveal complex fungal parasite networks in behaviorally manipulated ants. *Proc. Natl. Acad. Sci.* **114**, 12590–12595 (2017).
65. Wen, B. *et al.* 3D texture analysis for classification of second harmonic generation images of human ovarian cancer. *Sci. Rep.* **6**, (2016).

66. Goodfellow, I. *et al.* Generative adversarial networks. *Commun. ACM* **63**, 139–144 (2020).
67. Singh, N. K. & Raza, K. Medical Image Generation using Generative Adversarial Networks. *arXiv* (2020).
68. Sreejith Kumar, A. J. *et al.* Evaluation of Generative Adversarial Networks for High-Resolution Synthetic Image Generation of Circumpapillary Optical Coherence Tomography Images for Glaucoma. *JAMA Ophthalmol.* **140**, 974–981 (2022).
69. Li, X. *et al.* Unsupervised content-preserving transformation for optical microscopy. *Light Sci. Appl. 2021 101* **10**, 1–11 (2021).
70. Tian, L. *et al.* Deep learning in biomedical optics. *Lasers Surg. Med.* **53**, 748 (2021).
71. Chen, Y. *et al.* Generative Adversarial Networks in Medical Image augmentation: A review. *Comput. Biol. Med.* **144**, 105382 (2022).
72. Reich, C., Prangemeier, T., Wildner, C. & Koeppl, H. Multi-StyleGAN: Towards Image-Based Simulation of Time-Lapse Live-Cell Microscopy. in vol. 12908 476–486 (2021).
73. Krittanawong, C. The rise of artificial intelligence and the uncertain future for physicians. *European Journal of Internal Medicine* (2017).
74. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).

Chapter 2: SHG Imaging and Analysis

Elements of this chapter have been published as: Mora EMM, Champer MI, Huang W, Campagnola PJ, Grimes MD. Collagen is More Abundant and Structurally Altered in Lichen Sclerosus. Urology. 2023 Mar;173:192-197. doi: 10.1016/j.urology.2022.11.036. Epub 2022 Dec 9. PMID: 36509210; PMCID: PMC10038846.

2.1 Summary Statement

SHG imaging microscopy has been utilized to quantify changes in collagen organization in a wide array of diseases. SHG is label-free and has great specificity/sensitivity for imaging collagen fibers at high resolution (~0.5 microns). In this chapter, we will examine SHG imaging of a novel murine model for scleroderma, lichen sclerosis in penile and vulvar tissue, normal and cancerous colon samples, discoidin domain receptor 2 knockout mouse tumors, and five different murine modeled kidney disorders (chronic rejection, acute rejection, unilateral ureteral obstruction, tacrolimus, cyclosporin, and a control group). These samples were analyzed with CT-Fire, CurveAlign, and FIJI macros to extract fiber feature metrics. In some sample groups a support vector machine was used to find the features most associated with classification accuracy.

2.2 General Methods

Second Harmonic Generation microscopy was described in Chapter 1 Section 1.3, and will be briefly summarized here as well. Three locations were chosen for imaging in each sample and one image was obtained for each location. Full details of our SHG microscopes are detailed by Chen *et al.*[1]and Lien *et al.*[2] Briefly, imaging was performed with laser scanning galvos, coupled to an upright microscope (BX61; Olympus, Tokyo, Japan). The excitation source was a mode-locked Titanium Sapphire laser (Mira; Coherent, Santa Clara, California), providing 890 nm

excitation. The collected images were 512x512 pixels with a field of view of 180 x 180 μm. Power at the specimen was controlled by an electro-optic modulator. Image acquisition time was 3 s with three-frame Kalman averaging. A 40×0.8 NA water immersion lens (LUMPlanFL/IF; Olympus, Tokyo, Japan) was used to focus the excitation into the sample, and a 0.9-NA condenser collected the forward propagating signal. The lateral and axial resolution of the system was approximately 0.7 and 2.5 μm respectively, where this is sufficient for resolving collagen fibers. SHG emissions were collected using identical photon-counting detectors (7421 GaAsP; Hamamatsu, Hamamatsu City, Japan).

2.2.1 Image Analysis

FIJI, which is an open-source ImageJ platform, was used to calculate gray level co-occurrence matrix (GLCM) textures, along with the use of a custom macro that aided in the automation of the analysis. Using the Texture Analyzer plugin (Julio E. Cabrera, version v0.4 2006/07/07) the custom macro calculated the following texture parameters: Angular Second Moment (ASM), Inverse Difference Moment (IDM), Contrast, Entropy, and Correlation. With the use of Origin 2016 for curve fitting (OriginLab, Northampton, MA, USA) and FIJI, Two-dimensional fast Fourier transform (2D-FFT) analysis was performed, in which two custom macros utilized the Radial Profile Extended (Philippe Carl, version 2017/04/18) and the Oval Profile Plot (Bill O'Connell, version 2012/03/01) plugins. To characterize the individual fiber morphologies, curvelet transform and fiber extraction was performed with the use of CT-FIRE and CurveAlign[3,4].

2.3 Collagen remodeling in *Spag17* knockout mice

Systemic Sclerosis (SSc) can have serious health consequences, especially for patients with high morbidity risk factors. Identifying the underlying genetic links to pathogenesis is crucial for increasing treatment methods and patient detection. Moreover, while some genetic and environmental factors have been implicated, clear correlations to disease have not been established. To better understand the genetic link to collagen remodeling, we used the Sperm Associated Antigen 17 (*Spag17*) knockout mice as a model for the fibrotic effects that occur in SSC, where Second Harmonic Generation (SHG) imaging was used to examine both the dermis of wild-type and *Spag17* deficient mice. Using CT-FIRE and 2D-FFT analysis, the length, density, width, and alignment of the collagen fibers were found to be significantly different between the control and the *Spag17* KO mice in a manner consistent with a pro-fibrotic phenotype. Further, the features characterized by SHG imaging can accurately differentiate these dermal tissues using a linear support vector machine. This study supports the idea that *Spag17* potentially contributes to fibrosis in SSC and that collagen remodeling can be used as a potential diagnostic tool.

2.3.1 Introduction

Scleroderma is a rare autoimmune disease with an approximate prevalence of 17.6 per 100,000 in the global population. Similar to other inflammatory connective tissue disorders, the exact cause of the disease is still unknown.[5] Local scleroderma is characterized by thickening and collagen deposition (fibrosis) within patches of skin. Systemic scleroderma (SSc), in contrast, can have more severe consequences due to the widespread fibrosis that may appear in multiple organs[6,7]. SSc can have a long-term survival rate as low as 53% for patients at increased risk of

morbidity, e.g. having widespread organ involvement and a younger diagnosis[8,9]. Even among patients with lower morbidity risk factors, symptoms can cause complications that result in significantly affected quality of life[10]. Identifying the underlying genetic links to pathogenesis is crucial for increasing treatment methods and patient detection. However, current research on scleroderma's etiology is limited due to the rarity of the disease. Moreover, while some genetic and environmental factors have been implicated, clear correlations to disease have not been established[11].

More recently, some molecular markers associated with SSc have been investigated by several groups. Shin *et al*. found that SSc fibroblasts showed a significant increase in mRNA and protein expression of TGFβ2[12]. This family of growth factors is well-studied and has been known to increase collagen deposition and induce fibrosis in a variety of diseases[13–15]. Sapao *et al*. used RNA sequencing to identify *SPAG17* as the most significantly downregulated mRNA in SSc human skin biopsies. The SPAG17 protein was seen within fibroblasts from healthy patients but reduced within SSc biopsies. Using a *Spag17* KO mouse model, they further showed that the loss of *Spag17* expression can then lead to development of activated myofibroblasts and increased collagen production[16]. Specifically, the aberrant collagen was visualized using Second Harmonic Generation (SHG) microscopy. In this report, we provide more detailed and quantitative analyses of the changes in the collagen structure and organization in the *Spag17* KO mouse model.

SHG imaging microscopy[1] has been utilized to quantify changes in collagen organization in a wide array of diseases including cancers, fibroses, and connective tissue disorders[17,18]. SHG is label-free and has great specificity/sensitivity for imaging collagen fibers at high resolution (~0.5 microns). We have previously used SHG to probe structural differences in collagen in human osteogenesis imperfecta (OI)[19] as well as in the oim mouse model[19]. We and others have utilized

machine learning to successfully classify the fiber morphology in several cancers and fibroses[20–26].

In this work, we employ a multivariate support vector machine (SVM) approach to classify the collagen in wild-type (WT) and *Spag17* knockout (KO) mouse. Our specific approach uses SHG microscopy to image skin biopsies in the mouse WT and KO. Skin tissue provides an ideal target as the systemic pathology is present in the dermis and SHG imaging easily penetrates the extent of the tissue. Importantly, we showed that collagen alterations in human OI and the oim mouse skin[19,27] were identifiable even with limited clinical presentation. By analyzing the collagen fiber morphology using machine learning, this study quantifies the fibrotic effect expected from the loss of *Spag17* expression mouse model skin samples. While this study was performed on *ex vivo* mouse biopsies, the overall approach can be translated to imaging human skin *in vivo*, providing a new diagnostic tool while further advancing our understanding of SSc.

2.3.2 Methods

The B6N(Cg)-Spag17$^{tm1b(KOMP)Wts1/J}$ mouse line was used for this research[16]. Originally, seven samples were received; within those seven samples, three were wild-type (control) and four were knockout *Spag17* deficient. The skin was taken from the middle back area of the mice using a punch biopsy. The ages of the mice varied from 2 months - 2 years, and all but one of the mice were male (the female was of the wild-type group). Samples were between 0.3 - 0.9 mm thick. All samples were fixed in 10% formalin. Later, ten additional samples were received; within those ten samples, one was wild-type (control) and nine were knockout. The ages of the second batch varied similarly, and five of the mice were female (all *Spag17* deficient). Samples had similar thickness (around 0.3-0.9 mm) and were fixed in 10% formalin (Table 2.1).

Table 2.1: Summary table of each batch of samples with sex, genotype, and number of mice from each group.

Batch	**Sex**	**Genotype**	**Number of Mice**
Batch 1	Male	WT	2
		KO	4
	Female	WT	1
		KO	0
Batch 2	Male	WT	1
		KO	4
	Female	WT	0
		KO	5

Sample Preparation

A vibratome was used to section the first batch of samples to remove fat and ensure only the dermal layer would be imaged. The hair follicles were used as a reference point, as they begin slightly within the dermal layer. Samples were placed in a sanitized bin, and Nair was applied to the samples to remove hair for five minutes. Afterward, they were rinsed with deionized water. Slicing and Nair application were not performed for the second batch of samples because they already had Nair applied prior to the punch biopsy. All samples were mounted with PBS (phosphate-buffered saline). Nail polish was used to seal the coverslip onto the slide over the sample.

2.3.3 Results

SHG Imaging

Only forward propagating signals from SHG Imaging were used to analyze fiber morphology in this research. In imaging, the average power at the sample was between 100-110mW. Qualitative observations can be made by observing the images of the skin obtained from wild-type and *Spag17* knockout mice, where Figure 2.1 displays sample images from wild-type (left) and *Spag17* knockout mice (right). Compared to the skin of wild-type mice, skin from *Spag17* KOs demonstrates greater fiber density and alignment, a common feature observed in other fibroses.

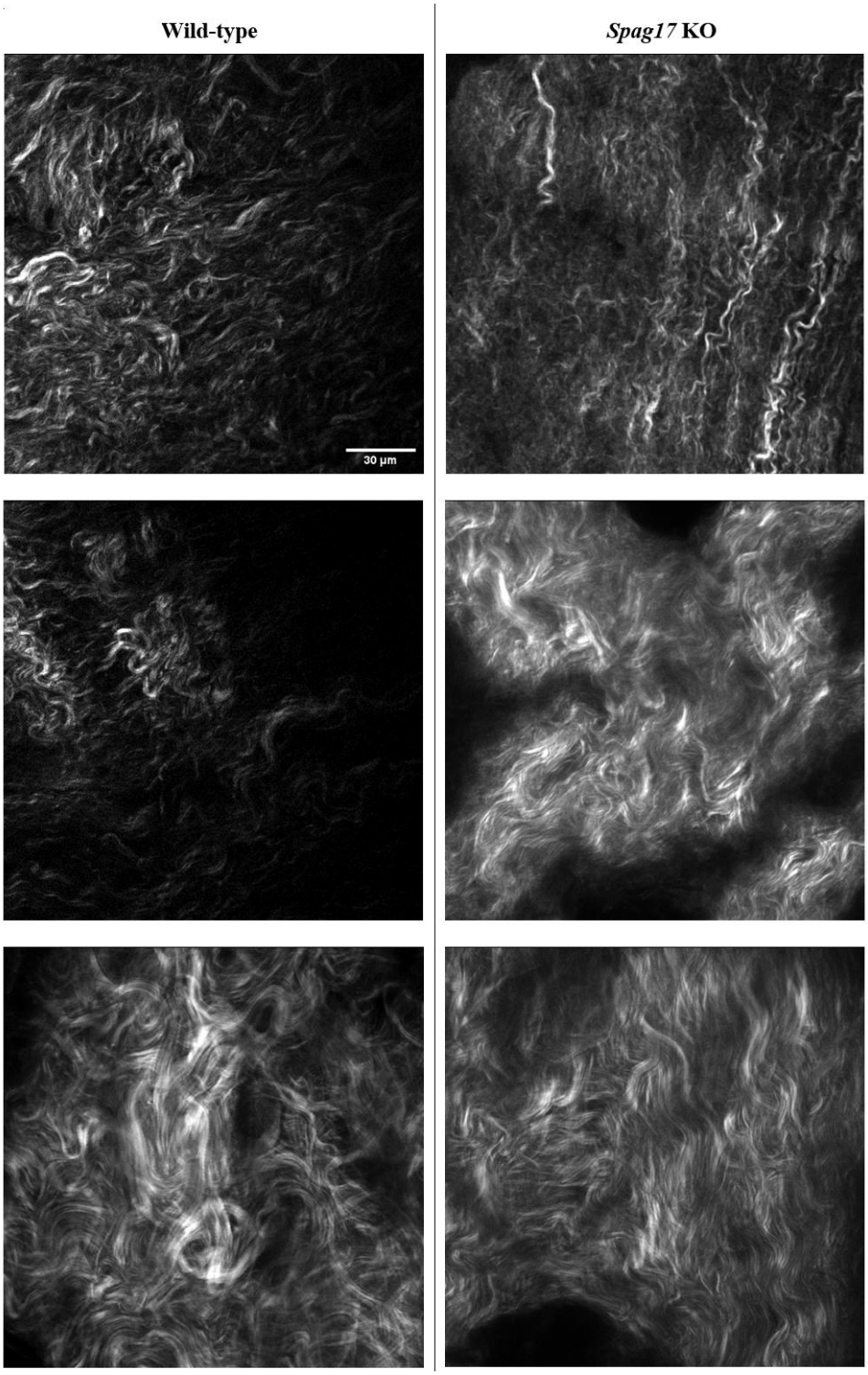


Figure 2.1: Second Harmonic Generation (SHG) forward collected images from wild-type (left) and Spag17 knockout (right) mouse dermis.

Gray Level Co-Occurrence Matrix (GLCM) Analysis

The gray level co-occurrence matrix (GLCM) is a second-order statistical approach to comparing nearest neighbor pixel grayscale values. This analysis and its features are detailed in Rentchler et al.[26], however, the features analyzed in this research are summarized below:

(i) Angular Second Moment (ASM): This is used to measure the number of repeated pairs. If the occurrence of repeated pixel pairs is high, the energy (square root of ASM) will be high.

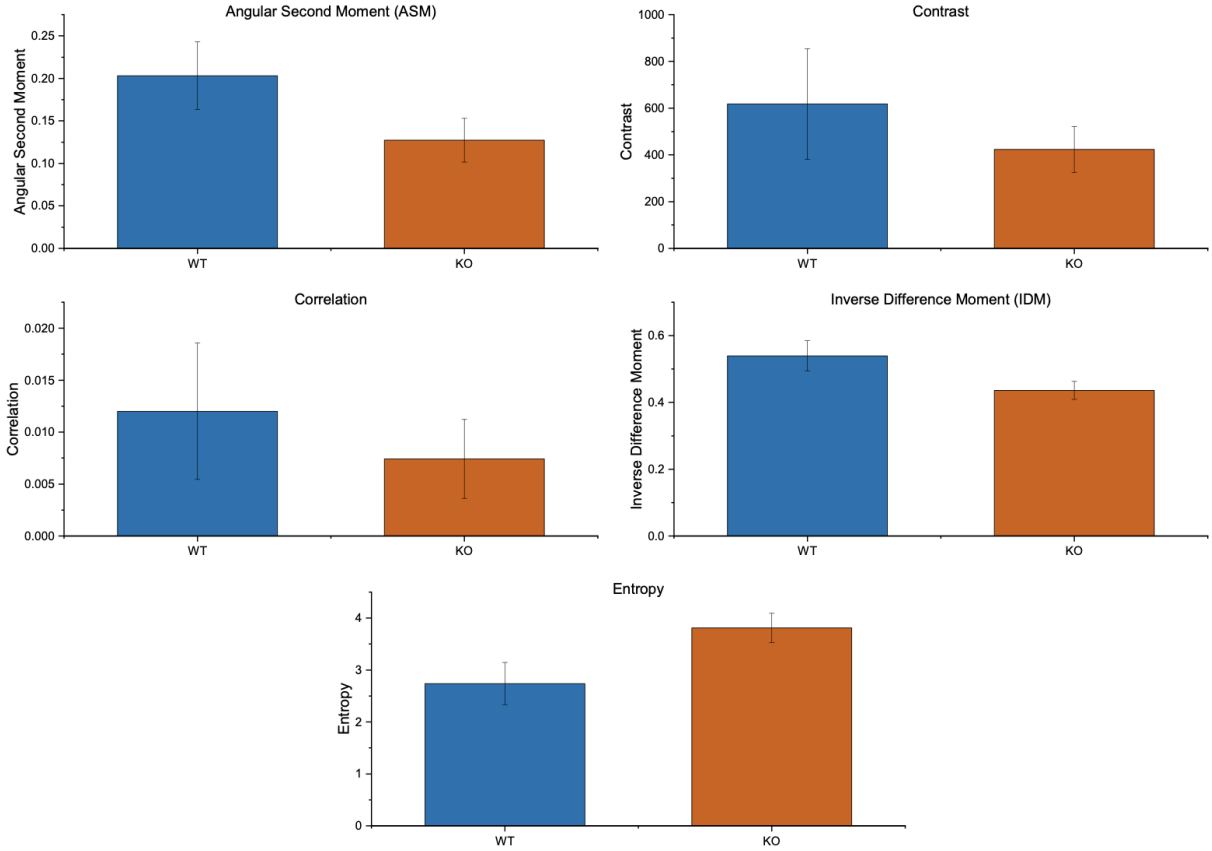
(ii) Inverse Difference Moment (IDM): This relates to the smoothness across the image. If the gray levels of the pixel pairs are similar, IDM will be high.

(iii) Contrast: This is a measurement of the local contrast of an image. If the gray levels of each pixel pair are similar, contrast will be low.

(iv) Entropy: This is a measurement of the randomness of gray level distribution. With greater randomly distributed gray levels throughout the image, the entropy will be high.

(v) Correlation: This measures the linear dependency of gray levels on those of neighboring pixels.

Results from GLCM analysis based on the forward signal from SHG imaging are displayed in Figure 2.2. Based on unpaired t-tests, none of the features yielded statistically significant results. However, IDM and entropy display potential trends with *p*-values of 0.060 and 0.054, respectively (Table 2.2). As displayed below, entropy was higher in knockout samples, and IDM was lower on average.



Figures 2.2*:* Plots show the mean value for each group in all five (SHG forward signal only) of the textures calculated with GLCM analysis.

Table 2.2: Grey Level Co-Occurrence Matrix analysis

GLCM Results	**Angular Second Moment**	**Contrast**	**Correlation**	**Inverse Difference Moment**	**Entropy**
Control WT	0.2032±0.0396	617.7±237	0.01201±0.00655	0.5396±0.0450	2.739±0.405
***Spag17* KO**	0.1273±0.0258	423.3±98.3	0.007418±0.00380	0.4362±0.0271	3.809±0.282
p-value	0.14	0.38	0.55	0.060	0.055

Two-Dimensional Fast Fourier Transform Analysis

Using 2D Fast Fourier Transform (2D-FFT) analysis, the features present in the tissue regions of the forward SHG channels were compared. FFT converts the spatial domain to frequency space, leading to higher spatial frequencies for smaller features. By analyzing the log scale power spectrum of images produced by the FFT function in ImageJ, the overall organization of features is revealed, as more alignment creates larger asymmetry. In contrast, images with more random features will produce a more symmetrical power spectrum[26].

Results from using 2D Fast Fourier Transform analysis to determine alignment are displayed in Figure 2.3. Alignment was assessed with an unpaired t-test and was significantly different between the wild-type and *Spag17* knockout tissues with a *p*-value of 0.031. Overall, wild-type tissues had more randomly oriented fibers, which resulted in less mean difference in the power spectrum compared to the *Spag17* knockout group (Table 2.3).

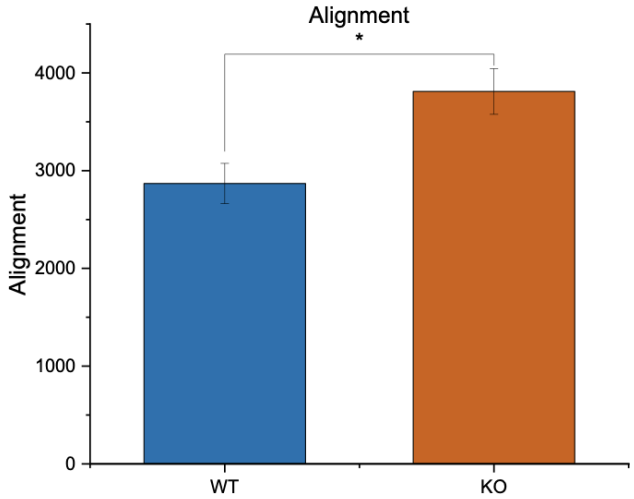


Figure 2.3: Plot shows the mean value for alignment calculated in the 2D-FFT analysis. *$p <$ 0.05.

CT-FIRE Analysis

Curvelet transform-fiber extraction (CT-FIRE) analysis was used to determine length, width, straightness, and density of the collagen fibers in the forward SHG images as displayed in Figure 2.4. Density is defined as the mean distance between a fiber and its nearest neighbors where a lower density value corresponds to closer packing. The curvelet transform (CT) to determine fiber metrics was performed following the fiber extraction (FIRE) algorithm.

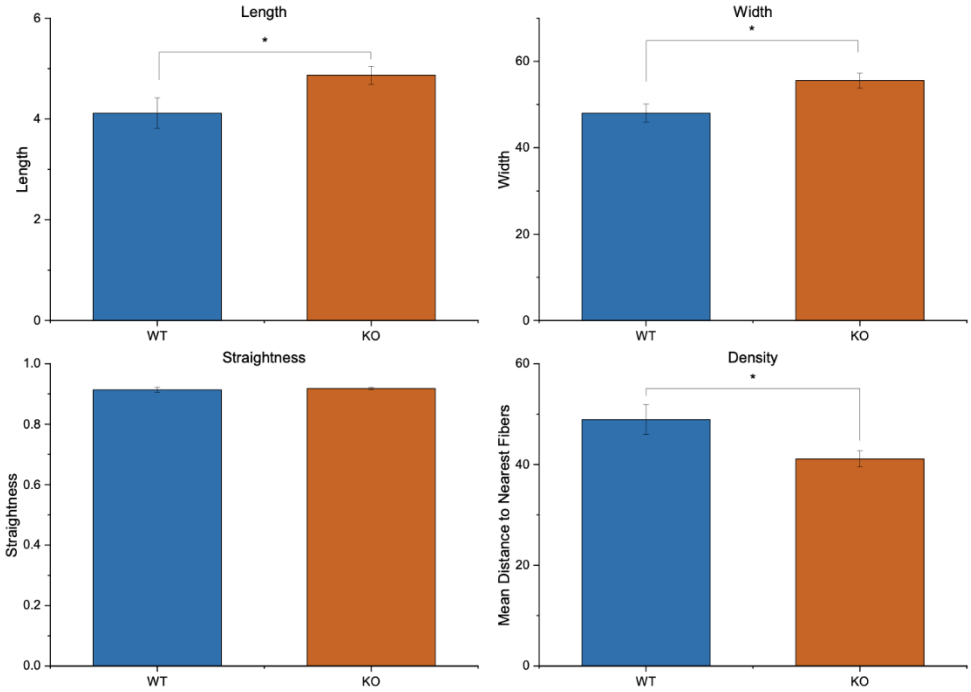


Figure 2.4: Plots showing the mean value for each feature in both groups (SHG forward signal only). Features were obtained with CT-FIRE and CurveAlign analysis. *$p < 0.05$.

The FIRE algorithm was used to de-noise images, enhance fiber edge features, and extract fibers. CT-Fire results were assessed with unpaired t-tests, and show significant differences in length, width, and fiber density between the wild-type and *Spag17* knockout tissues with *p*-values of 0.041, 0.024, and 0.019, respectively. Straightness was not significantly different between the

samples. Overall, *Spag17* knockout collagen fibers are generally longer, wider, and more densely packed than wild-type skin in the analyzed mouse models (Table 2.3).

Table 2.3: CT-Fire and Fourier analysis

	Length	**Width**	**Straightness**	**Density**	**Alignment**
WT	4.12±0.09	48.0±2.1	0.914 ±0.008	48.9 ±3.0	2868±206
***Spag17* KO**	4.87±0.08	55.6±1.7	0.918 ±0.003	41.1 ±1.6	3810±231
p-value	0.041*	0.024*	0.57	0.019*	0.031*

Classification by Support Vector Machine

We used a support vector machine (SVM) to rank the relative feature importance of the CT-FIRE and GLCM metrics. Further details on classification by SVM and its applications have previously been discussed[28]. In this implementation, we used the scikit-learn package in Python to create a c-support vector with a linear kernel. Feature values were normalized and split into training and test sets, consisting of 80 and 20% of the data, respectively. Weight coefficient values were squared to standardize the analysis for positive and negative values. Our model achieved an accuracy of 89%, with alignment, straightness, and density generally being the highest ranked features for classification (Figure 2.5).

Figure 2.5: Coefficient weights assigned to each attribute in an SVM classification model. Negative values are of equal importance to classification as positive.

2.3.4 Discussion

Animal models for systemic sclerosis have been used to advance the understanding of SSc development and pathology, as well as providing a target for drug development. The most widely used induced models have used bleomycin and hypochlorous acid[29]. However, this method generally focuses on the tissue fibrosis but deprioritizes the immunological and vascular pathology of SSc[30]. Moreover, the fibrotic response of bleomycin is not sustainable and resolves quickly. Genetic models for SSc can fill this gap by showing potential pathological cascades that mimic symptoms of the disease. For example, silencing the transcription factor Friend leukemia virus integration 1 (Fli1) resulted in mice with downregulated autoimmunity as well as dermal and esophageal fibrosis[31]. As more models are assessed, a more complete picture of SSc etiology can form. The *Spag17* knockout represents a SSc-like model and may be a good step forward in the study of this disease[16].

Here, we have shown that SHG imaging can identify and quantify the fibrotic response in *Spag17* knockout mice. We found that *Spag17* deficient skin biopsies contained collagen fibers

that were more aligned, wider, and more tightly packed, matching the results of other studies that examined scar tissue[32,33]. Overall, this further suggests a connection between downregulation of *Spag17* and increased risk of fibrosis. Although the GLCM analysis did not yield statistically significant results, it contributed to the SVM analysis, which achieved suitable classification accuracy for potential clinical use.

The results in this study align with similar studies using SHG microscopy to examine fibrotic and desmoplastic responses in other diseases[18]. While the geometric aspects of fibers can vary, e.g. width, length, and straightness, the common thread is an increase in alignment. For example, we have reported this in ovarian cancer[25,26] and its precursors as well as in idiopathic pulmonary fibrosis[24,34]. Other researchers have found this effect in breast cancer and pancreatic cancer, where alignment is negatively correlated with survival[35,36].

Spag17 deficient mice were used as a model due to the gene's significance in fibrotic diseases. Mutations to the *Spag17* gene in mice have been found to lead to skeletal malformations, likely due to its function in ciliogenesis[37]. Cellular receptors localized in the primary cilium carry a significant role in the activation of morphogen signaling pathways that have been associated with organ fibrosis. Primary cilium has been proposed as a mediator of fibrosis pathogenesis in systemic sclerosis[38]. Further, previous research found *Spag17* deficient mice have immotile nasal and trachea cilia, suggesting the gene’s essential role in motile cilia function[39,40]. There appears to be a direct relationship with the *Spag17* gene and fibrosis because mutations in primary cilia have been known to lead to fibrosis in various areas of the body[37]. With future research on *Spag17*, further conclusions may be drawn on the extent it contributes to connective tissue

disorders.

Because this was the first mouse model of its kind, there were limitations due to the low number and differing preparation of the samples. The first batch of seven samples underwent Nair to remove excess fur. However, these tissues were already formalin-fixed, so complete removal of hair was not achieved. Keratin is highly autofluorescent and invisible by SHG imaging[41]. However, keratin autofluorescence can block the SHG signal due to an inner filter effect or shadowing effect. These effects somewhat reduced the optical resolution. The second batch of mice had Nair applied prior to the punch biopsy and sample preparation. This successfully increased fiber resolution. Nonetheless, we were still able to extract fiber features from both preparations and merged the results.

Animal models for scleroderma typically see sex differences in disease response, with female mice exhibiting greater susceptibility to fibrosis compared to males[42]. In previous models such as the bleomycin-induced fibrosis model, female mice developed more severe skin and lung fibrosis[43]. This is reasonably similar to the response in human patients wherein SSc is typically much more common in women, although diffuse cutaneous SSc and worse cardiovascular complications are more likely in men[44–46]. Estrogen has been studied as a possible profibrotic agent in SSc, but further reviews will be needed to study sex differences in patients[47,48]. The immune response could also affect sex differences. Models have shown higher levels of pro-inflammatory cytokines in female mice, which may in turn contribute to the pathogenesis of SSc[49,50]. Given the limited number of specimens available for this new *Spag17* line, the sample size was insufficient to detect notable differences between male and female mice. Future studies

can address potential sex differences in disease responses across species.

Despite limited sample size, we have shown the utility of using SHG to analyze collagen fiber morphology in the *Spag17* mouse model of SSC, where a multivariate approach resulted in good classification accuracy between the KO and WT mice. This opens the door to future applicability in both animal and human studies. Significantly, the collagen fiber alterations are apparent in skin. While this analysis was done with *ex vivo* biopsies, they could also be performed *in vivo*. Here, optical clearing[51] could be used to improve image contrast and allow imaging through the entire human dermis. This could also be combined with other contrast methods such as Third Harmonic Generation (THG), which is sensitive to regions with a change in refractive index, as well as two-photon excited autofluorescence imaging.

2.3.5 Conclusions

Here we used SHG microscopy and multivariate imaging analysis to distinguish collagen fiber morphology in WT and the *Spag17* knockout samples representing Systemic Sclerosis. Significant differences were found in length, width, alignment, and fiber density, and an SVM afforded good differentiation. The overall features in the fibrotic tissues were analogous to those seen in other fibroses and cancers displaying desmoplastic responses. The ability to perform the analysis on skin affords the ability to pursue this approach *in vivo* and provide a new and much needed diagnostic tool.

2.4 Collagen in Lichen Sclerosis

We used picrosirus red to stain and image collagen in human urethral, vulvar, and foreskin specimens with and without LS. Using ImageJ, we quantified and compared 1) collagen proportionate area (CPA) and collagen fiber count, 2) collagen fiber length and width, and 3)

collagen structure using the texture analysis technique gray level co-localization matrix (GLCM) with respect to LS status and tissue type. We analyzed 23 LS specimens (vulva n=9, urethra n=7, foreskin n=7) and 29 non-LS specimens (vulva n=9, urethra n=7, foreskin n=13). Fiber count and CPA were significantly higher in all LS specimens compared to non-LS specimens (CPA: mean 0.971 vs 0.948, $p < 0.007$; fiber count: mean$\pm$SE = 2906$\pm$127 vs 2509$\pm$78.03 fibers; $p = 0.003$). Collagen fiber width and length were similar with respect to LS status. GLCM analysis showed decreased inverse difference moment and increased entropy in LS tissues ($p<0.001$). We found that LS tissues had greater collagen content compared to non-LS tissues which is consistent with other fibrotic diseases. Quantitative assessment of collagen organization, using GLCM, revealed less homogeneity and more disorganization of collagen in LS compared to non-LS tissues. Taken together, our findings suggest that alterations in physical tissue properties seen in LS may be due to both increased collagen abundance and altered structure.

2.4.1 Introduction

Lichen sclerosis (LS) is a debilitating chronic inflammatory and fibrotic disease of the anogenital skin which causes significant urinary and sexual dysfunction. LS is more common in women than men (6:1), and the prevalence is estimated to be between 1:300 and 1:1000[52]. In affected men, LS is associated with severe urethral fibrosis and associated stricture disease in up to 30% of cases[53]. LS associated urethral strictures are refractory to surgical reconstruction (urethroplasty) failing up to 70% of the time[53,54]. The pathogenesis of LS and the mechanism by which it causes urethral fibrosis is unknown and represents a critical knowledge gap that impedes development of alternative treatment strategies for this debilitating disease.

Fibrotic diseases are defined by excessive deposition of fibrillar collagen. Dysregulation of collagen production and metabolism result in collagen accumulation (fibrosis) that disrupt organ structure and function[27,55,56]. For example, excessive collagen type I synthesis drives myocardial fibrosis that results in heart failure[19] and renal collagen deposition sustains pathological responses of the glomerulus and thereby implicated in progression of chronic kidney disease[57]. More recently, the significance of alterations in the collagen and ECM structure of fibrotic tissues has been recognized as a critical determinant of their mechanical properties and thereby a central driver of disease pathogenesis[58–60]. Changes in collagen structure usually result from post-translational crosslinking of collagen fibrils by an enzyme family called LOX/LOXL which culminates in increased tissue stiffness[61]. Alterations in collagen organization, and increased crosslinking in particular, are often more powerful determinants of tissue stiffness than total collagen content as has been demonstrated in both Duchenne Muscular Dystrophy and Idiopathic Pulmonary Fibrosis among other diseases[34,62]. Further, targeting the mechanisms driving altered collagen structure represents an untapped therapeutic target in fibrotic diseases.

Alterations in collagen content and structure have not been well studied in LS and associated urethral stricture disease and those that have been done have yielded mixed results[54,62]. Therefore, we compared collagen content and structure in three human tissue types (vulva, foreskin, and male urethra) both affected and unaffected by LS. We addressed the hypothesis that genital skin and male urethra affected by LS has greater total collagen content and wider collagen fibers (reflective of increased collagen crosslinking) compared to tissues unaffected by LS, recapitulating the pattern seen in other fibrotic diseases.

2.4.2 Methods

Patient identification and sample collection

Following Institutional Review Board (IRB) approval, we used procedural billing data to identify and include adult patients (age >18yrs) who underwent vulvar biopsy, circumcision, urethroplasty, or urethral biopsy at the University of Wisconsin-Madison Hospital from 2015-2020 with available pathologic tissue. Due to the rarity of LS we cross-referenced these results with pathology records to identify specimens with mention of either "lichen sclerosus" or "balanitis xerotica obliterans" (a historic synonym for LS) in the pathology report. Pathology reports were reviewed and specimens with a diagnosis of malignancy or carcinoma in situ were excluded. Ten specimens per tissue type (urethra, vulva, foreskin) without mention of LS or BXO were randomly selected as non-LS samples. All pathologic tissues were then re-reviewed with a Genitourinary Pathologist (WH) who was blinded to the original pathology report to determine the diagnosis of LS based on the presence of epidermal atrophy, loss of dermal rete pegs, hyperkeratosis, dermal hyalinization, and presence of inflammatory infiltrate (Table 2.4). Based on this review, patients were classified as either "LS" or "non-LS."

Table 2.4: Characteristics of included specimens with and without lichen sclerosus.

	No-LS (n=23)	**LS (n=29)**
Gender, no. (%)		
Male	14 (61)	20 (69)
Female	9 (39)	9 (31)
Tissue type, no. (%)		
Vulva	9 (40)	9 (31)
Urethra	7 (30)	7 (24)
Foreskin	7 (30)	13 (45)

LS = lichen sclerosus

Collagen staining

We stained collagen with picrosirius red (PSR) which has a high affinity and specificity for fibrillar collagen and is commonly used to quantify collagen content in multiple tissue types[63]. Five-micron formalin-fixed, paraffin embedded tissue sections were deparaffinized at room temperature in two washes of xylene (3 min) and then rehydrated through a graded series of ethanol (100%, 75%, 50%; all 5 min) to distilled water. Sections were incubated at room temperature for 1 hour in a solution of 0.5g of Direct Red 80 (Sigma-Aldrich, St. Louis, MO) in 500mL of saturated picric acid. The slides were washed twice with acidified water (0.5% acetic acid) for 5 minutes each and then dehydrated in ascending concentrations of ethanol (50%, 75%, 100%; 4 quick dips) and cleared in xylene. Sections were covered with Richard-Allan toluene-based mounting medium (4112APG; Thermo Fisher Scientific) before coverslipping.

Image acquisition, processing, and analysis

Picrosirius red (PSR) stained tissues were imaged using fluorescent microscopy as previously described[64]. Briefly, PSR stained tissues were imaged using an Eclipse E600 microscope (Nikon) with a 20x dry objective lens (Plan Fluor NA = 0.50; Nikon) for collagen content analysis and a 40x dry objective lens (Plan Fluor NA = 0.75; Nikon) for collagen structure analysis. PSR fluorescence was detected using a Y-2E/C TX Red filter cube and tissue autofluorescence was detected using a B-2 E/C FITC filter cube (Chroma Technology Corp; Bellows Fall, VT). The PSR fluorescence and tissue autofluorescence images were captured using a CoolSnap DYNO camera (Photometrics) and NIS-Elements D v4.6 imaging software (Nikon) at 1920 x 1460 pixel resolution. Five randomly selected stromal regions of interest (ROIs) immediately adjacent to the epithelial surface were captured using a 20x objective for content analysis and three ROIs were

captured using a 40x objective for fiber characteristics. Exposure time was kept constant for all images, and brightness and contrast adjustments that optimize signal-to noise-ratio were kept constant using Adobe Photoshop software (Adobe Systems; San Jose, CA).

The PSR fluorescence and tissue autofluorescence images were superimposed using the ImageJ-based analysis package Fiji[65]. Images were processed as red and green images with PSR fluorescence occupying the red channel and tissue autofluorescence occupying the green channel. PSR fluorescence was isolated using Fiji and the PSR Isolation plugin. For collagen content analysis, the resulting monochrome image was processed with Fiji and the Black Background Quantification plugin was used to quantify the total number of black and color pixels. Total collagen content was defined as the proportion of color pixels compared to the total number of pixels for each image. For collagen structure analysis, the PSR isolated monochrome images were analyzed using CT-FIRE software (LOCI; Madison, WI). This software detects and quantifies collagen fiber-like metrics in images, including fiber length, diameter, and density. Values for collagen content, fiber length, fiber count, and fiber thickness obtained for each ROI were used to calculate a mean value for each subject. Mean values for each parameter were compared between LS and no-LS tissues using the Student's T-test and the Mann-Whitney U test where appropriate.

2.4.3 Results

Collagen content

We analyzed 52 total specimens which included 23 non-LS specimens (9 vulva, 7 urethra, 7 foreskin) and 29 LS specimens (9 vulva, 7 urethra, 13 foreskin). Average patient age was 57.5 (range 19–90) years. We hypothesized that total collagen content is greater in LS compared to

non-LS tissues. To test this, we used two quantitative metrics to compare collagen content. The first was collagen proportionate area (CPA) which is the proportion of each region of interest stained with PSR. This method has been validated for use in multiple other fibrotic tissue types (lung, liver, kidney)[66,67]. The second metric was collagen fiber count defined as the number of individual collagen fibers in a given slide area.

Overall, CPA was significantly greater in LS compared to non-LS tissues (mean±SD 0.971±0.03 vs 0.948±0.02, $p < 0.007$). When each tissue type was compared individually, collagen content in urethra and foreskin LS tissues was significantly greater than in non-LS tissues (urethra – 0.979±0.01 vs 0.953±0.02, $p=0.02$; foreskin – 0.974±0.03 vs 0.951±0.02, $p = 0.027$). Vulvar LS tissues also had increased CPA compared to non-LS tissues (0.961±0.06 vs 0.941±0.02) but this difference was not statistically significant ($p = 0.385$) (Figure #). Similarly, there were more collagen fibers in LS tissues than in non-LS tissues (mean±SE = 2906±127 vs 2509±78 fibers; $p = 0.003$). When collagen fiber counts were compared across individual tissue types with respect to LS status we found that LS tissues consistently had more collagen fibers but this difference was statistically significant only for foreskin tissues (2722±154.8 vs 2161±42.24 fibers; $p=0.003$) (Figure 2.6).

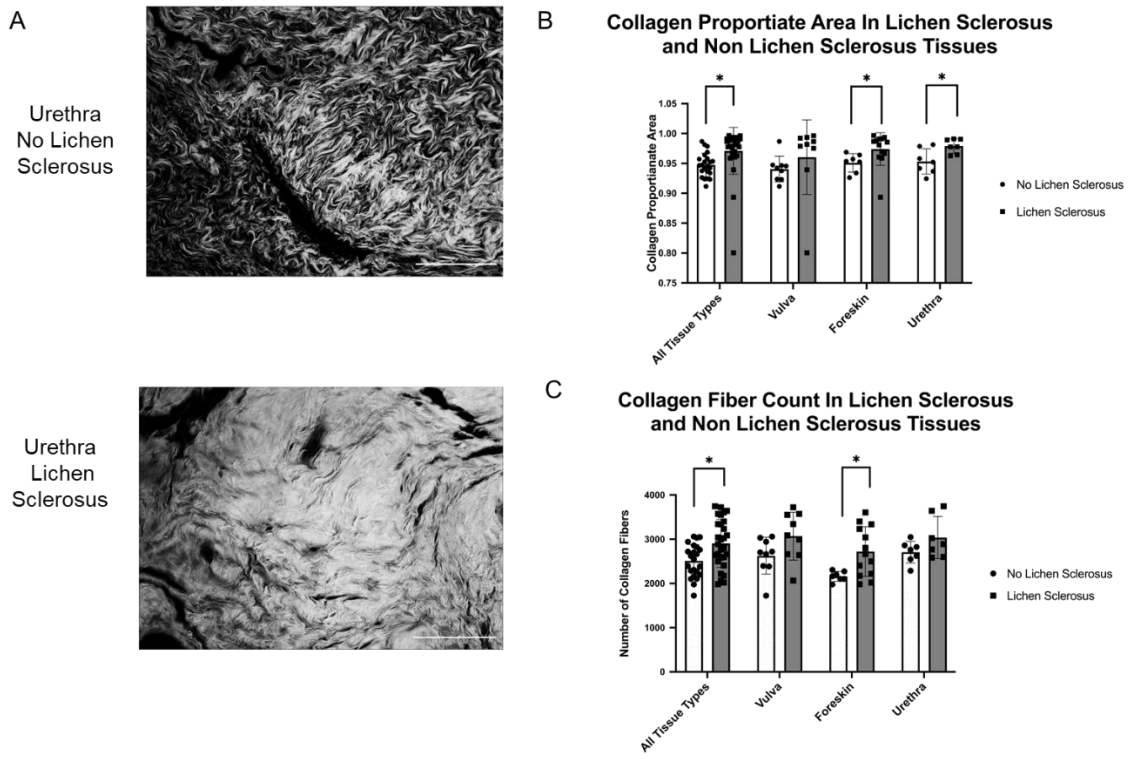


Figure 2.6: Comparison of collagen content in human vulva, urethra, and foreskin with and without LS. (A) Representative fluorescent microscopy images (20x) of PSR stained urethral collagen demonstrating increased collagen content in LS compared to non-LS tissue. (B) Comparison of mean collagen proportionate area in tissues with and without LS. Taken together, LS tissues had higher mean collagen proportionate area compared to non-LS tissues (0.971 vs 0.948, respectively; $p = 0.007$). When segregated by tissue type urethra and foreskin LS tissues had significantly higher collagen proportionate areas (urethra – 0.979 vs 0.953, $p = 0.02$; foreskin – 0.974 vs 0.951, $p = 0.027$) while vulvar tissues were no different with respect to LS status (0.961 vs 0.941, $p = 0.385$). (C) Comparison of collagen fiber count in tissues with and without LS. Taken together, LS tissues had higher mean collagen fiber counts compared to non LS tissues (2906 vs 2509 fibers, $p = 0.003$). When segregated by tissue type foreskin LS tissues had higher mean number of collagen fibers (2722 vs 2161 fibers, $p = 0.004$) while vulva and urethra were similar with respect to LS status (vulva – 3071 vs 2625 fibers, p = 0.07; urethra – 3036 vs 2707, $p = 0.139$). Brackets indicate pairwise comparisons with "*" indicating $p < 0.05$. LS = lichen sclerosus, PSR = picrosirius red. Scale bar = 100 μm.

Collagen structure is altered in lichen sclerosus

We hypothesized collagen fibers are wider and longer in LS compared to non-LS tissues as this is reflective of increased fiber cross linking[58]. We found that collagen fiber width and length are similar between LS and non-LS tissues causing us to reject our hypothesis (Figure 2.7).

However, visual inspection of the PSR stained images showed very high density and overlap of collagen fibers and subjective differences in collagen organization between LS and non-LS tissues. Because of this we were concerned that curvelet analysis may lack sensitivity to differences in individual fiber characteristics due to poor curvelet fitting to individual collagen fibers. Therefore, we used gray level co-occurrence matrix analysis (GLCM) to assess differences in the structural organization of collagen fibers.

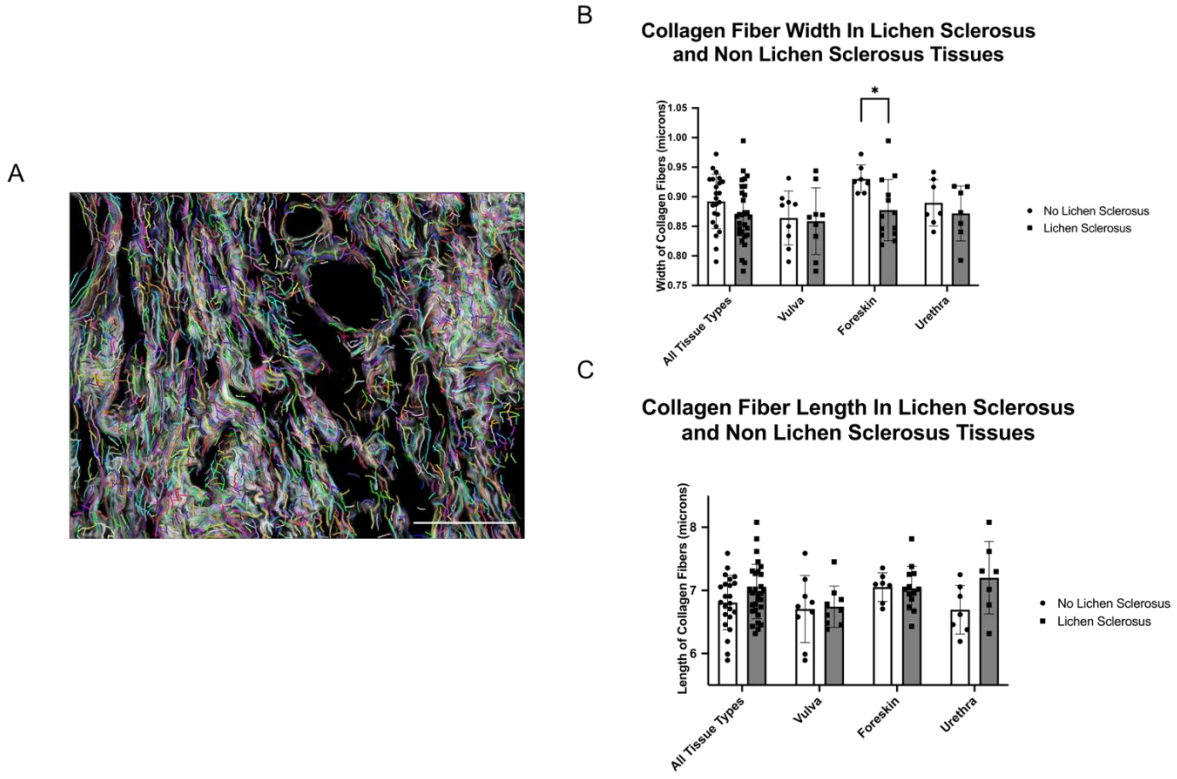


Figure 2.7: Comparison of collagen fiber width and length in human vulva, urethra, and foreskin tissue with and without LS. (A) Representative 40x image demonstrating an area of subepithelial collagen deposition with superimposed curvelets from the CT-FIRE software package. Each colored curvelet is fit to an individual collagen fiber and analysis of these curvelets yield the data presented in the remainder of the figure. (B) Mean collagen fiber width was similar across all tissue types with respect to LS status ($p = 0.111$). However, when individual tissue types were examined foreskin tissues with LS were found to have significantly increased collagen fiber width compared to non-LS tissues (0.877 vs 0.93 microns, $p = 0.006$). (C) Comparison of collagen fiber lengths across vulvar, foreskin, and urethral tissues did not display significant differences with regard to LS status. Brackets indicate pairwise comparisons with "*" indicating

p <0.05. LS = lichen sclerosus. Scale bar = 50 μm.

We used GLCM to assess two metrics. The first is entropy which measures image orderliness by examining the probability of pixel pair occurrences. The value of entropy increases when pair occurrences are less common. The second is inverse difference moment (IDM) which quantifies the homogeneity of an image by weighing the diagonal of the matrix the highest, thus more pixel similarity results in higher IDM. Using GLCM analysis, we found that LS tissue has more entropy and less IDM than non LS tissue. These results suggest decreased homogeneity of collagen structure and more disorganized structure in LS tissues compared to non-LS tissues (Figure 2.8).

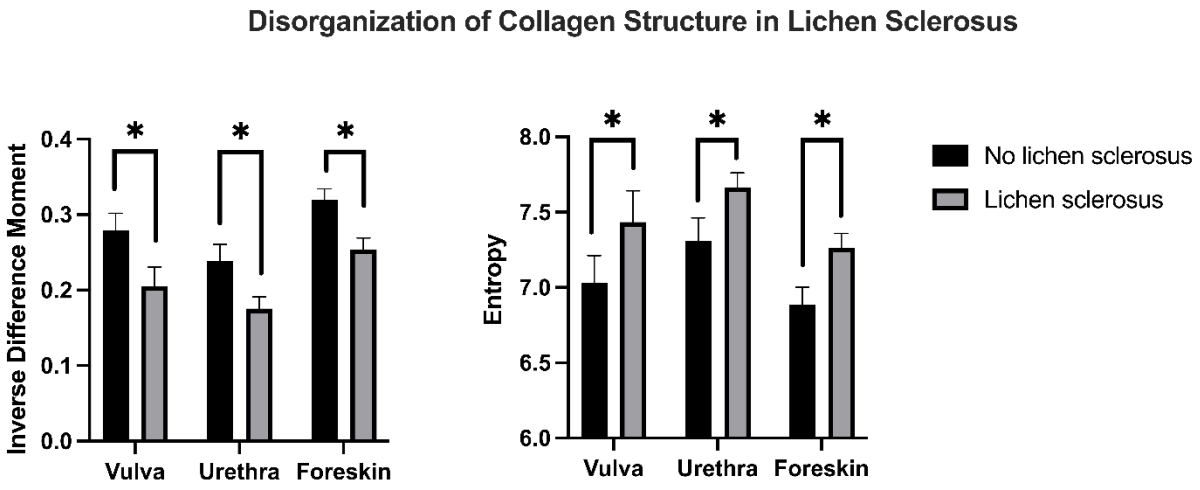


Figure 2.8: Comparison of inverse difference moment (IDM) and entropy in human vulva, urethra, and foreskin tissues with and without LS. There is a consistent decrease in IDM and increase in entropy in LS compared to non-LS tissues. This reflects decreased homogeneity and disorganization of collagen structure and may imply alternative pathways for collagen production and remodeling in LS. * indicates p<0.001 in pairwise comparisons.

2.4.4 Discussion

Collagen content and structure are poorly studied in LS and often with conflicting results. We quantitatively compared collagen content and structure in vulvar, foreskin, and urethral tissues affected and unaffected by LS. We found that LS tissues had greater collagen content compared to non-LS tissues which is consistent with other fibrotic diseases. Quantitative assessment of collagen organization, using GLCM, revealed less homogeneity and more disorganization of collagen in LS compared to non-LS tissues. Taken together, our findings suggest that alterations in physical tissue properties seen in LS may be due to both increased collagen abundance and altered structure.

There are few prior studies characterizing LS collagen content and structure. Electron microscopy studies reveal a greater density of dermal collagen in vulvar LS compared to non-diseased control tissues[68]. An immunohistochemical study revealed an abnormal distribution of collagen I and III in vulvar LS, although individual collagen fibers were not distinguished in this study, perhaps due to the markedly increased fiber density[68]. Godoy and colleagues showed in a case-control format similar to the present study that LS tissues had decreased elastic fibers, a "non-homogeneous" distribution of collagen, and increased expression of collagens I, III, and V with increased density of collagen V on 3D reconstruction[69]. An additional immunohistochemical study examining expression of basement membrane proteins noted more abundant collagen IV and VII (both of which constitute the structural scaffold of the ECM) in LS tissues[69]. In summary, prior studies of LS describe qualitative increases in collagen content and alterations in ultrastructure and our results are concordant with this literature. However, our study is the first to quantify increases in collagen content and alterations in collagen

organization. Further, prior studies have been limited to vulvar and foreskin tissues, and we demonstrate a similar and consistent pattern across multiple tissue types including urethra.

It is controversial if the increased collagen content seen in LS results from increased collagen production, impaired degradation, or both. A study by Oikarinen and colleagues demonstrated an increase in collagen production in LS and colocalization of fibroblasts and type I procollagen cDNA on in situ hybridization[70]. However, the authors did not identify active fibroblasts characterized by TGFb expression in regions of histological inflammation. The authors suggest that alternative inflammatory pathways may account for fibrosis in LS. Other studies have concluded that collagen biosynthesis is not increased in LS. One study of foreskin specimens affected by LS actually reported that new collagen synthesis was decreased in LS compared to unaffected controls on hydoxyproline assay[71]. These prior studies have examined a small number of patients with different tissue types. Therefore, it is possible that variation across multiple tissues effected by LS accounts for differences and/or that selection bias in how patients were recruited for study confounded these results. While our study does not provide data regarding the mechanisms underlying increased collagen abundance, we plan to pursue this as a future direction.

We demonstrate altered collagen structure in LS across multiple tissue types. In other fibrotic disorders, such as idiopathic pulmonary fibrosis, altered collagen structure most often results from post-translational modification. The most common of these is cross linking of collagen fibrils which is positively associated with increased tissue stiffness[72]. Further, there is mounting evidence that increased matrix stiffness can promote self-sustaining progressive fibrosis such as that seen in LS tissues[73,74]. Based on our data demonstrating a consistent increase in collagen content and structural alterations we postulate that LS related fibrosis results from both increased

collagen production and altered post-translational modification of collagen fibrils. To further investigate this, we plan to assess LOX/LOXL, the enzyme responsible for collagen cross linking, in LS tissues and perform ex-vivo mechanical studies with atomic force microscopy to quantify differences in tissue stiffness and correlate these with changes in collagen cross linking.

Our study has several limitations which deserve mention. First as a retrospective study of archival pathology specimens it is subject to inherent selection bias. However, due to the relative rarity of LS in male patients obtaining sufficient tissue for analysis in a prospective manner represents a significant pragmatic barrier. Second, our study used non-LS tissues as a comparison group rather than true normal tissue - the collection of which is limited by clinical equipoise. While this may obscure differences accounted for by variation across tissue types, we predict that comparing diseased tissues affected and unaffected by LS is clinically relevant and that use of true normal tissue as a comparison group would likely amplify the magnitude of observed differences. Third, our microscopy-based approach does not allow for chemical quantification of collagen content nor does it permit analysis of differences in collagen transcription or post translational modification. Prior studies have validated that collagen proportionate area and curvelet analysis reflect chemically measured collagen content[75]. As a future direction we aim to characterize differences in collagen gene expression as well as LOX/LOXL expression which is responsible for collagen fibril cross linking in order to further parse the relative contribution of increased collagen production and cross linking in LS pathogenesis.

2.4.5 Conclusion

Our study demonstrates, for the first time, a quantitative increase in collagen content and altered collagen structure in LS across multiple human tissue types including urethra. These data are

novel in LS, but consistent with findings in other fibrotic disorders. As a future direction we plan to further probe the molecular mechanisms underlying both increased collagen synthesis and post-translational modifications of collagen fibrils leading to altered structure. By coupling these mechanistic studies with ex-vivo study of the mechanical properties of tissue, specifically urethra, we hope to identify pathways which are fruitful therapeutic targets in treatment of LS related fibrosis.

2.5 SHG Microscope Depth with Colon Samples

We received a total of six colon samples, three from normal tissue, and three from patients with confirmed cancerous malignancy. Each sample was embedded in 5% agarose gel and thickness was estimated with calipers. The samples had a wide range of thickness from approximately 1.85mm to 6.2mm. Thinner samples required less power, from 99mW, and thicker tissues could be imaged at up to 501mW without burning. We imaged three different widely spaced areas of each sample. One malignant sample was not particularly wide and contained little to no collagen; only one suitable area was found. Imaging was collected from just before and just after a discernable SHG signal could be acquired. Due to the length of imaging time, the 40x0.8 NA water immersion lens was fed additional droplets at regular 1-hour intervals.

2.5.1 Results

Signal intensity and axial depth were the primary attributes we considered in the analysis. SHG stacks were processed with gray-level thresholding and pixel-based signal averaging. Each sample was considerably different from one another in terms of anatomical appearance, signal intensity, and collagen organization. SHG signal intensity comparison between our sample groups was inconclusive. Our analysis included the background (0-value pixels) in the averaging.

Each sample has different levels of coverage (area of image with collagen signal) in the different areas imaged (Figures 2.9 and 2.10). Therefore, the current results will be skewed by the contribution from the ratio of fiber coverage to background and is not describing the raw contribution from absolute intensity of only the measured collagen. Average power for SHG image collection was 269 mW, and 174 mW for fluorescence images. Maximum intensity across the samples was found at 74 microns for normal tissue, 110 microns for cancerous tissues, and 90 microns across all samples (Figure 2.11 and 2.12).

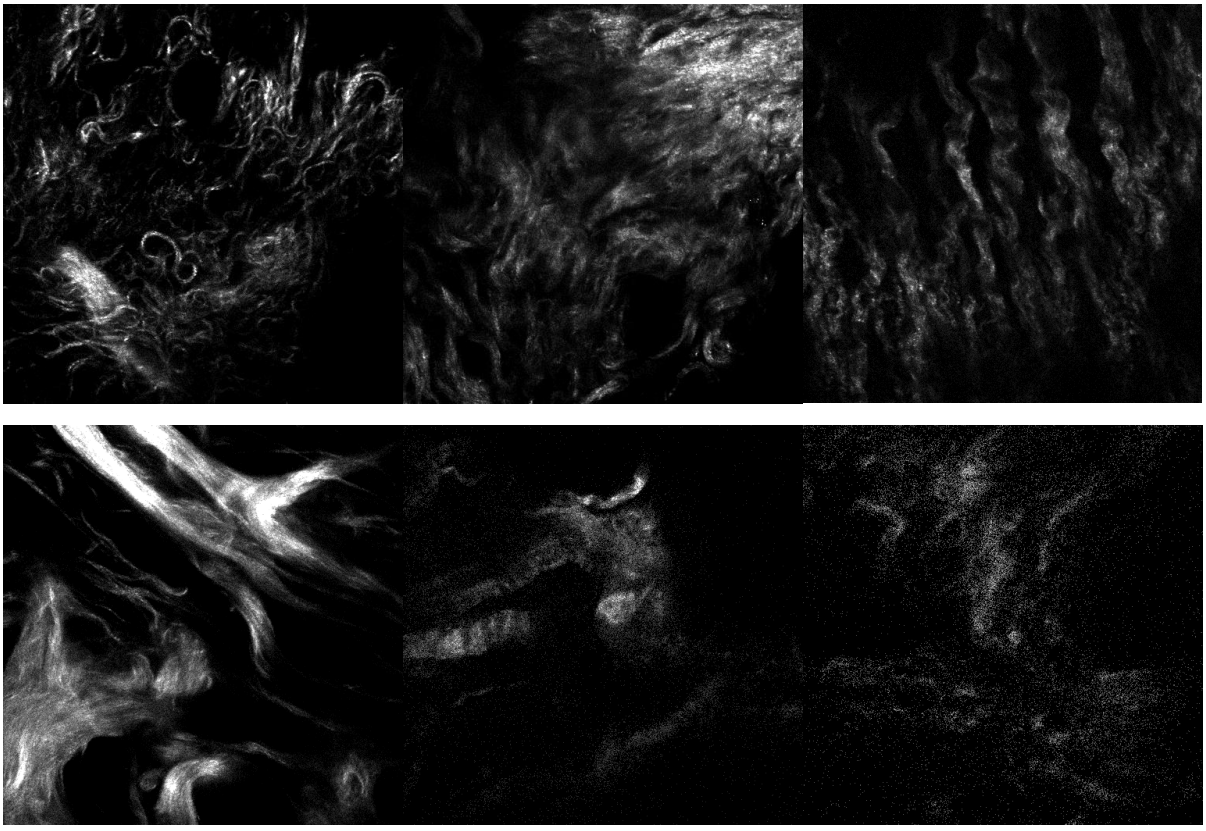


Figure 2.9: Representative SHG images from each sample. *Top:* Normal unlabeled at 97 microns. Normal 721126 at 105 microns. Normal 19203015 at 54 microns. *Bottom:* Unlabeled tumor at 62 microns. Tumor 1923014 at 59 microns. Tumor at 721125 at 84 microns (very low fiber coverage and signal sample).

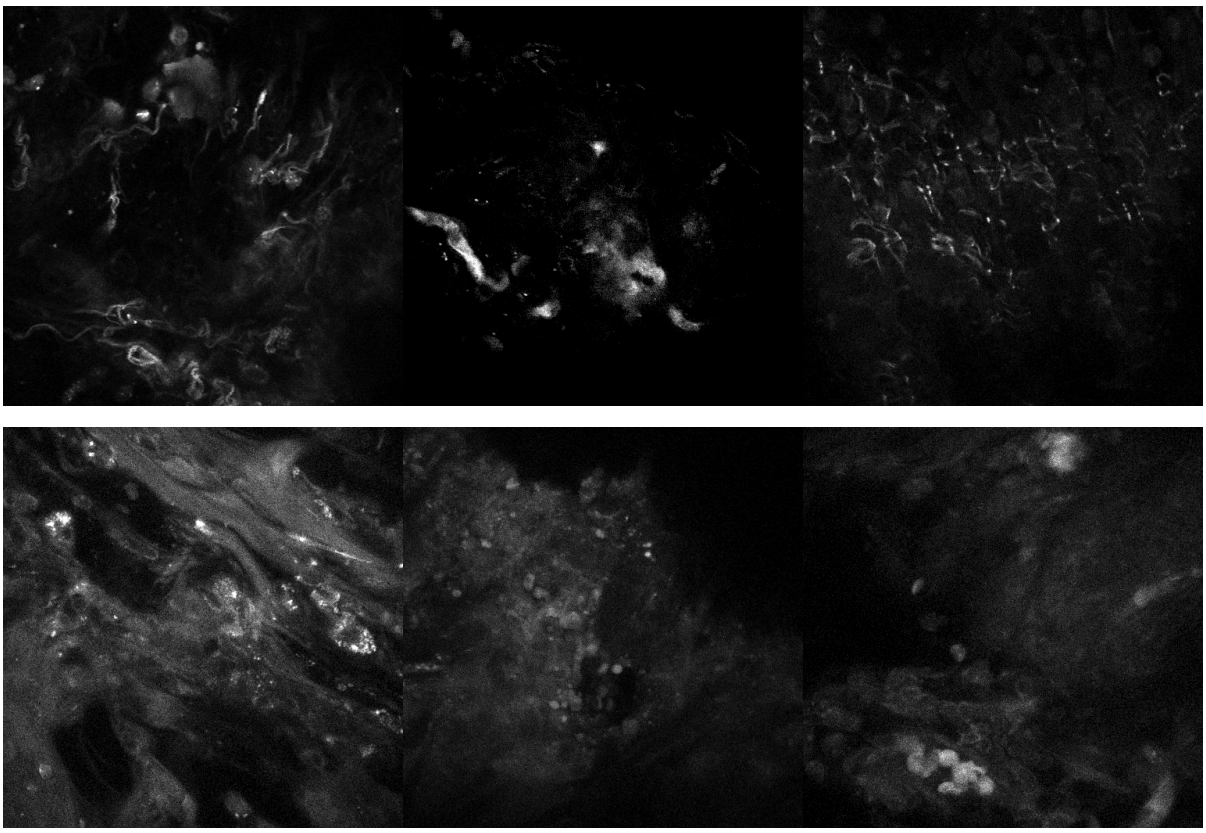


Figure 2.10: Representative fluorescence images from each sample. *Top:* Normal unlabeled at 97 microns. Normal 721126 at 105 microns. Normal 19203015 at 54 microns. *Bottom:* Unlabeled tumor at 62 microns. Tumor 1923014 at 59 microns. Tumor at 721125 at 84 microns (very low coverage sample).

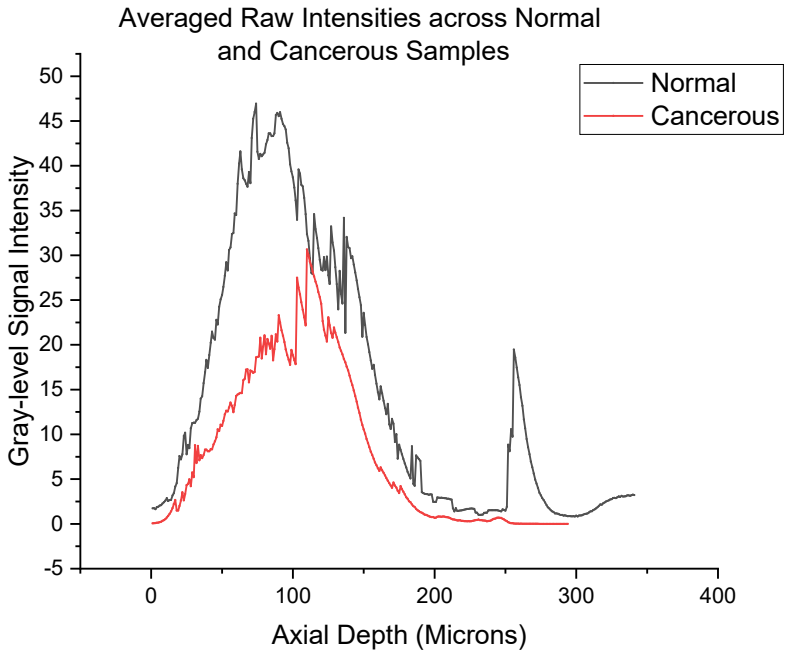


Figure 2.11: Averaged raw data for normal and cancerous samples showing drop off at approximately 200 microns. Comparison is influenced by the image coverage more than overall intensity of collagen.

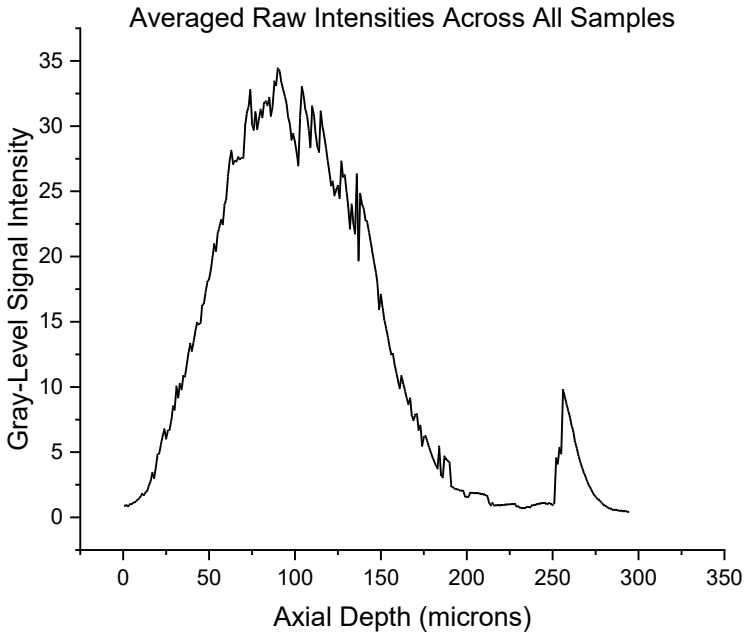
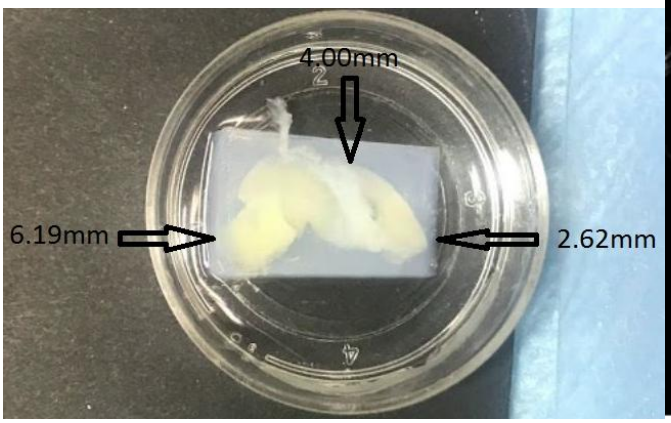


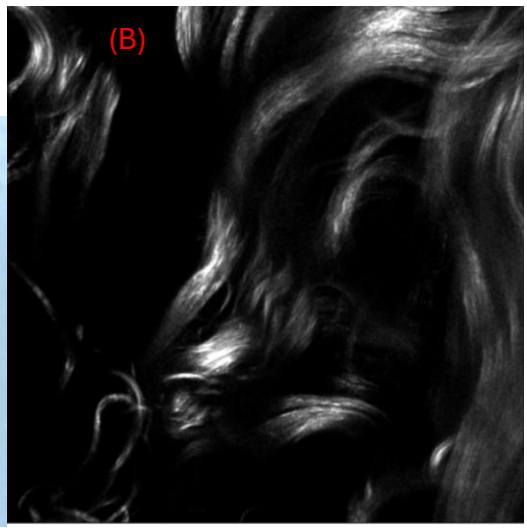
Figure 2.12: Averaged raw intensity data across all image stacks of all groups. Peak intensity and coverage were found at 90 microns.

Differences between our sample groups can largely be attributed to major differences in sample thickness and variability in location of biopsy. Areas with what appeared to be glands/ducts or vessels had little to no collagen in both types of samples. One of the cancerous samples was almost entirely composed of this type of tissue. The highest axial depth was found in the 6mm portion of the thickest normal sample (Figure 2.13). The diseased samples (Figure 2.14) tended to be thinner (approximately ~1.85-3mm), which seemed to affect the achievable maximum depth and signal.

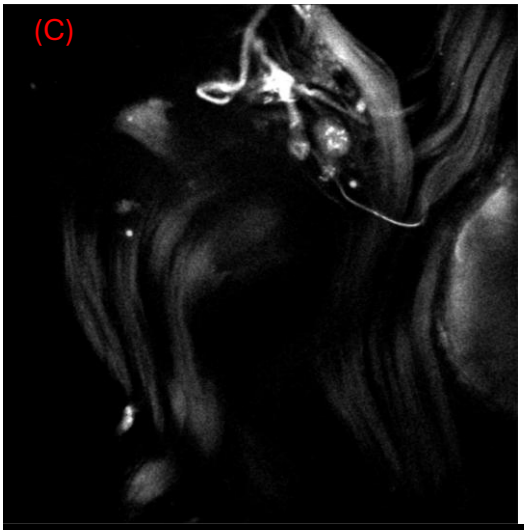
(A)


4.00mm
6.19mm
2.62mm

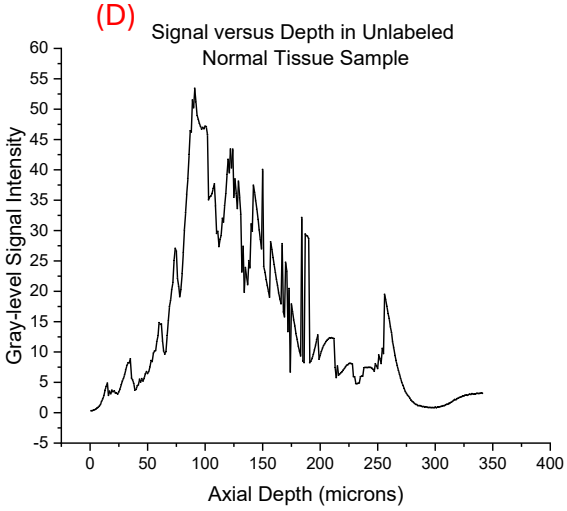
(B)



(C)



(D)


Signal versus Depth in Unlabeled
Normal Tissue Sample
Gray-level Signal Intensity
60
55
50
45
40
35
30
25
20
15
10
5
0
-5
0
50
100
150
200
250
300
350
400
Axial Depth (microns)

(E)

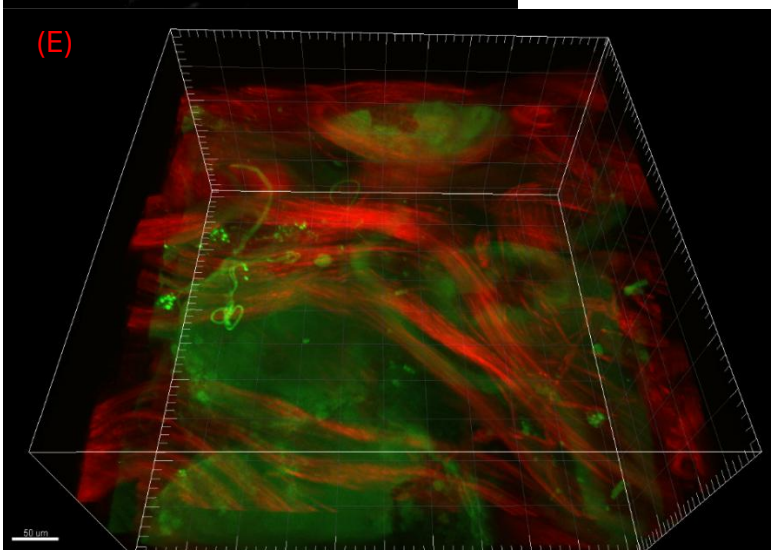

50 um

Figure 2.13: Analysis of the thickest normal unnumbered tissue. (A) Lab image of the top-down (laser entry) view of the embedded tissue. Arrows indicate areas where thickness was estimated. (B) Representative SHG imaged with 465 mW power from an axial depth of 103 microns. (C) Representative autofluorescence, imaged with 347 mW power, showing high intensity for elastin and cellular material. (D) Data analysis indicating significant intensity cut off at approximately 250 microns. (E) 3D top view of SHG (red) and fluorescence (green) modeled with the Imaris software suite.

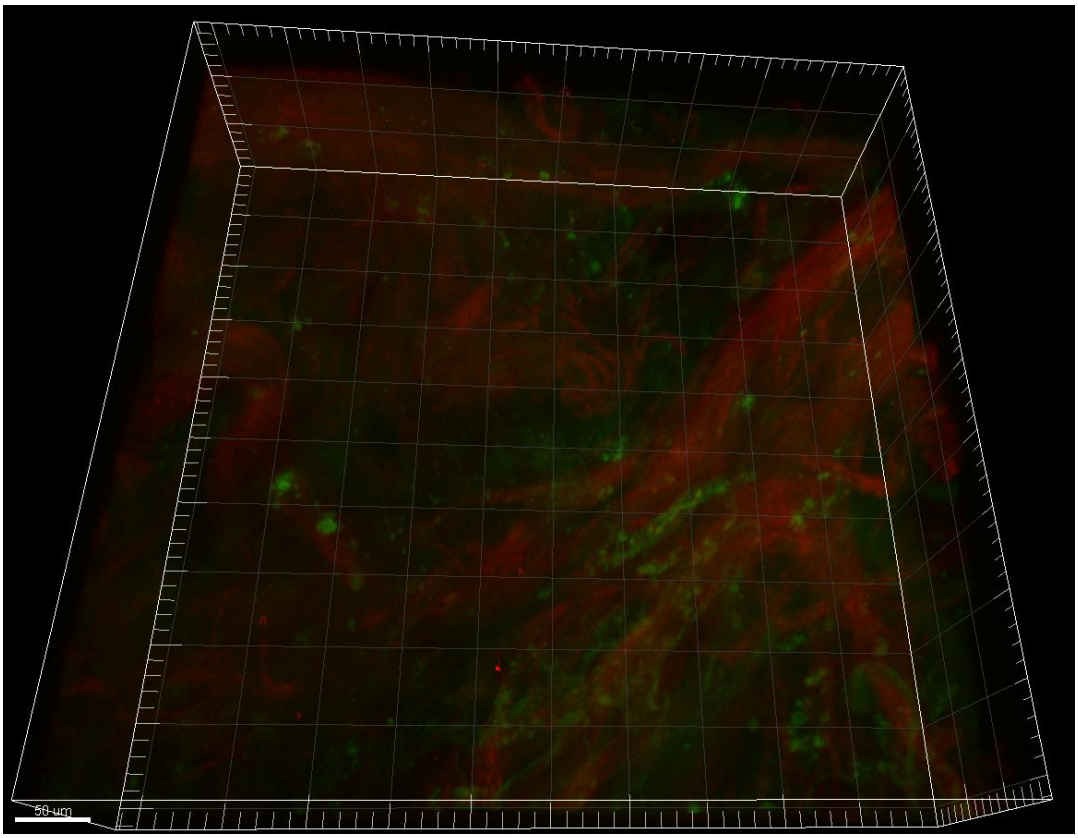


Figure 2.14: Representative 3D top view of SHG (red) and fluorescence (green) for a cancerous sample. Rendering modeled with the Imaris software suite.

2.6 Tumor SHG analysis with DDR2 knockout

Ovarian cancer has the highest mortality of all gynecologic malignancies[76]. Consequently, there is a pressing need to identify molecular mechanisms underlying tumor metastasis in this disease. Elevated expression of receptor tyrosine kinase Discoidin Domain Receptor-2 (DDR2) has been

associated with poorer patient survival. DDR2-mediated signaling networks are known to play a critical role in tumor progression[77]. DDR2 binds specifically to fibrillar collagen[78]. High expression of DDR2 has been linked to increased metastasis in various cancers[77,79–83]. Notably. fibrillar collagen is more abundant in the tumor stroma of ovarian cancer patients than in the ovaries of healthy individuals[84].

To further examine collagen morphology's relationship with DDR2 expression we received samples from Dr. Katherine Fuh. Tumor tissue from C57BL/6J *Ddr2* WT and KO mice were fixed in 10% formalin and sliced with a vibratome. All samples were mounted with PBS (phosphate-buffered saline). Nail polish was used to seal the coverslip onto the slide over the sample.

2.6.1 Results

Overall, 45 wildtype tumors and 51 DDR2 knockout samples were imaged and analyzed. Using CT-Fire curvelet fiber, we have found length, straightness, and density to be statistically significantly different (Table 2.5). For this study, density is defined as the mean distance between collagen fibers, such that larger values indicate a less dense sample. DDR2 KO models showed shorter, less straight, and less dense fiber morphology (Figure 2.15 and 2.16). The GLCM analysis yielded a single statistically significant metric, indicating that DDR2 KO samples had greater contrast than the control model (Table 2.6).

Table 2.5, CT-Fire and Fourier analysis

	Width	**Length**	**Straightness**	**Density**	**Alignment**
Control WT	4.18±0.09	54.0±1.3	0.91 ±0.003	48.6 ±1.5	4445±270
DDR2 KO	4.06±0.08	49.2±0.99	0.89 ±0.003	53.9 ±1.8	4162±291
p-value	0.29	0.004*	0.0005*	0.031*	0.48

Table 2.6, Grey Level Co-Occurrence Matrix analysis

GLCM Results	**Angular Second Moment**	**Contrast**	**Correlation**	**Inverse Difference Moment**	**Entropy**
Control WT	0.512±0.056	462±80.7	0.0162±0.0069	0.722±0.037	1.90±0.27
DDR2 KO	0.520±0.043	838±145	0.0237±0.0088	0.757±0.025	1.42±0.15
p-value	0.91	0.032*	0.52	0.52	0.11

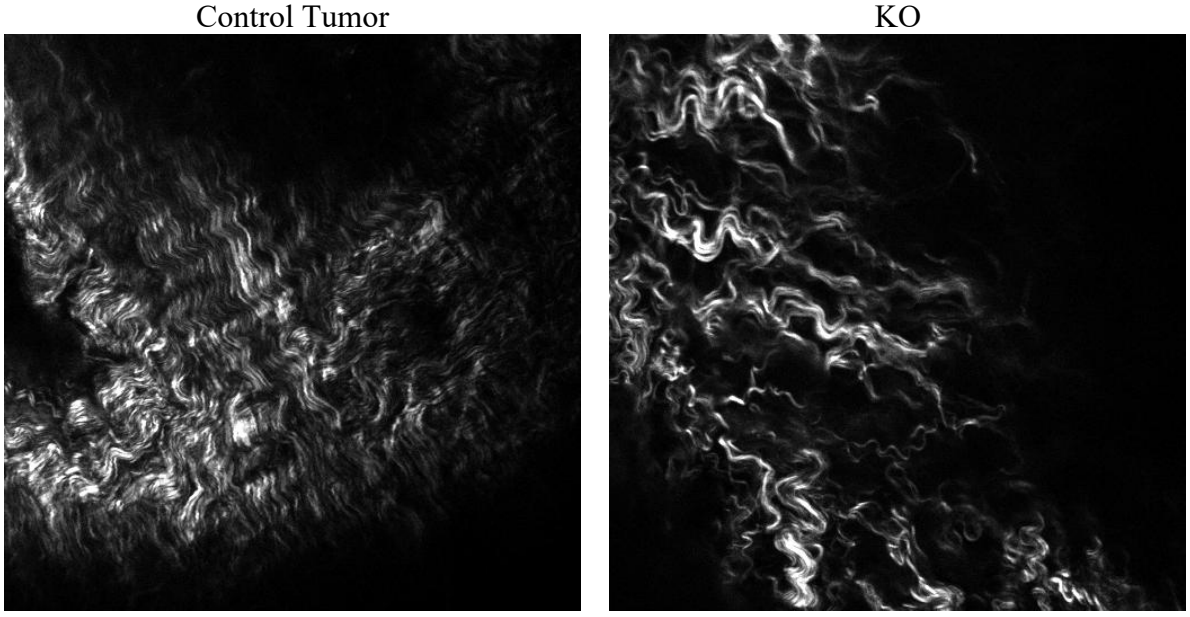
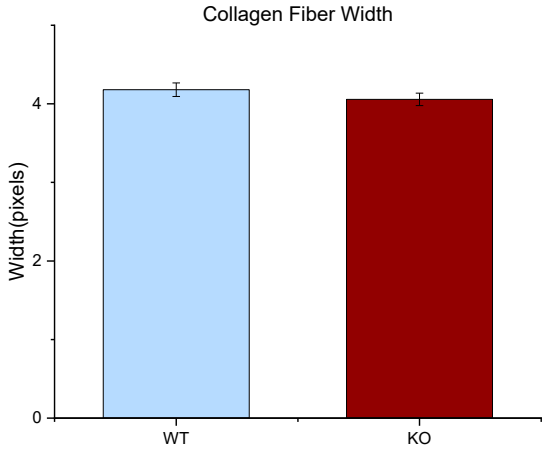
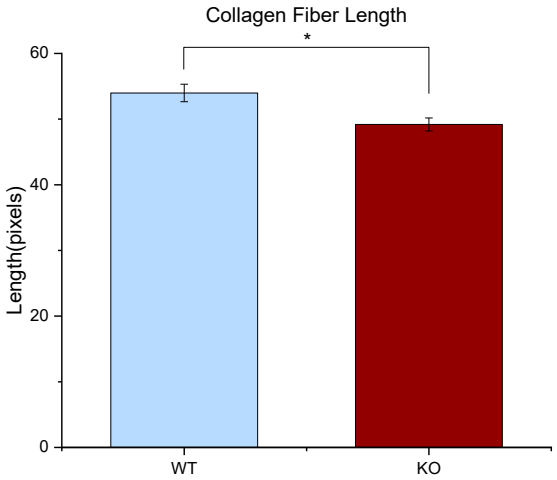
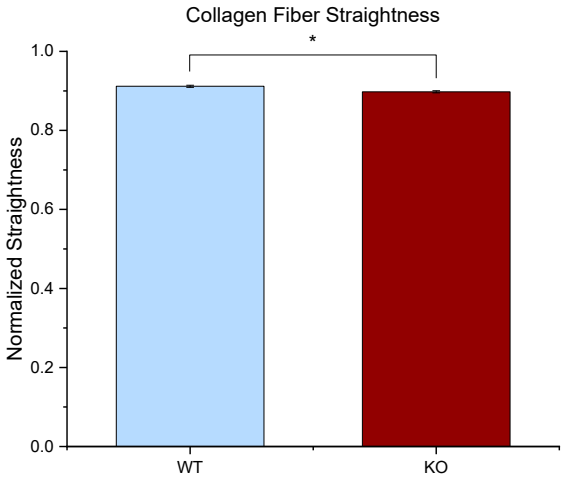
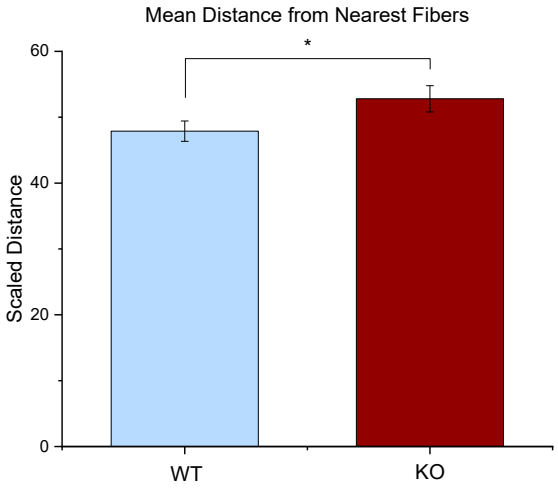


Figure 2.15, Tumor environment compared with DDR2 knockout. A) Control tumor B) DDR2 Knockout tumor.


Collagen Fiber Width
Width(pixels)
0
2
4
WT
KO


Collagen Fiber Length
*
Length(pixels)
0
20
40
60
WT
KO


Collagen Fiber Straightness
*
Normalized Straightness
0.0
0.2
0.4
0.6
0.8
1.0
WT
KO


Mean Distance from Nearest Fibers
*
Scaled Distance
0
20
40
60
WT
KO

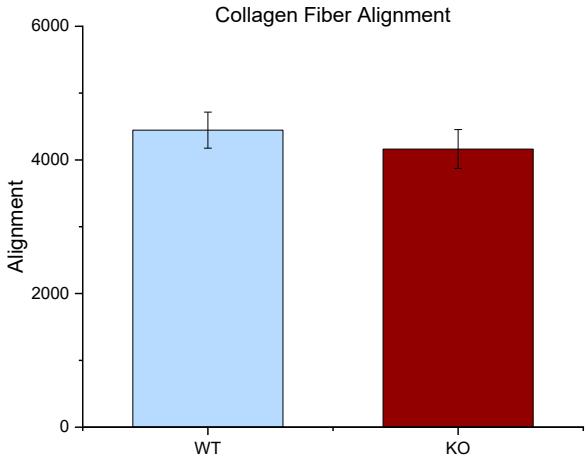

Collagen Fiber Alignment
Alignment
0
2000
4000
6000
WT
KO

Figure 2.16, CT-Fire analysis results. A) Control collagen width had a mean pixel width of 4.179 and KO had 4.057. B) The length of collagen fibers was a mean of 53.98 pixels in the control group and 49.19 in the knockout group. C) The wildtype collagen had a mean straightness of 0.91 and KO had 0.89. D) Control collagen had a mean distance between nearest fibers of 48.6 and KO had 53.9. E) The mean alignment of the wildtype collagen was 4445 while the knockout group was 4162.

A support vector machine (SVM) was used to rank the relative features for linear classification between the two sample groups. The model achieved an accuracy of 81% with entropy, angular second moment, and length ranked highest (Figure 2.17).

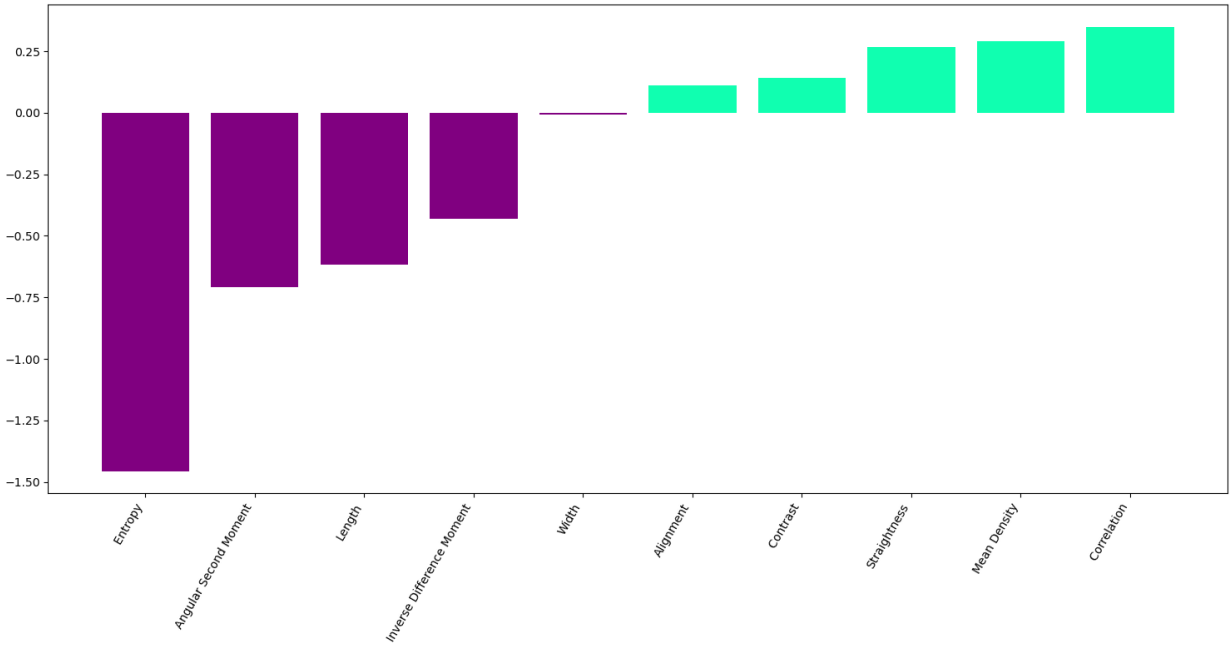


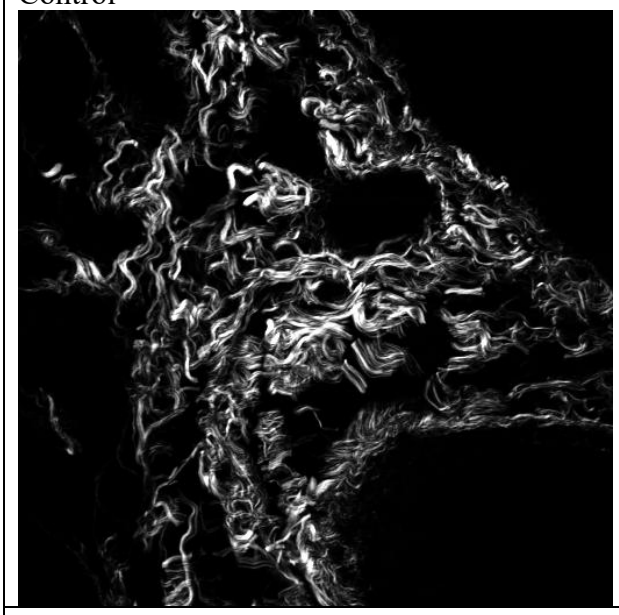
Figure 2.17, SVM classification results**.** Each feature is assigned a coefficient for classification. Negative values are of equal importance to positive. The larger weights indicate higher usefulness for class comparison.

2.7 Kidney Murine Models

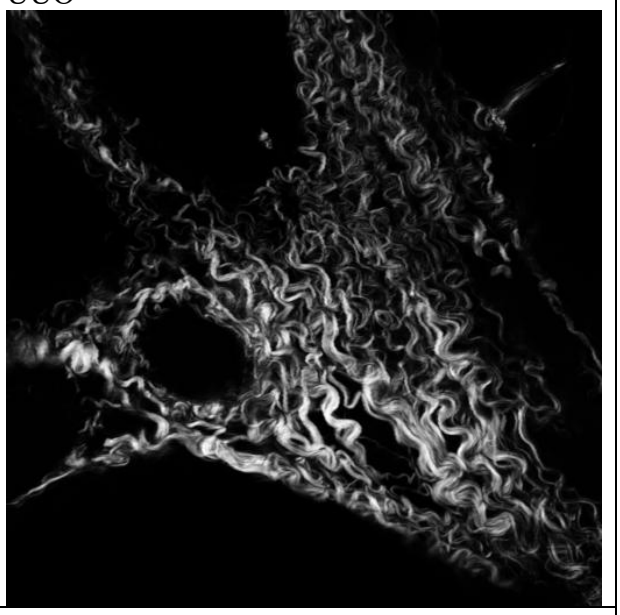
The kidney project features a large array of modeled diseases: chronic rejection, acute rejection, UUO (unilateral ureteral obstruction), TAC (tacrolimus), and CsA (cyclosporin). These have been provided by Dr. Arjang Djamali in H&E slides at three different thicknesses (5, 10, and 15

ums) (Figure 2.18). We used a structural similarity index to verify compatibility of images from the different thicknesses[85]. Data was then pooled together for the CT-Fire analysis. Ultimately 364 acute rejection, 411 chronic rejection, 365 CSA, 361 Control, 175 UUO, and 247 TAC images were used for testing. For alignment analysis with FFT, forwards and backwards signal images were assessed separately. Despite separating the data, we expect similar trends to appear.

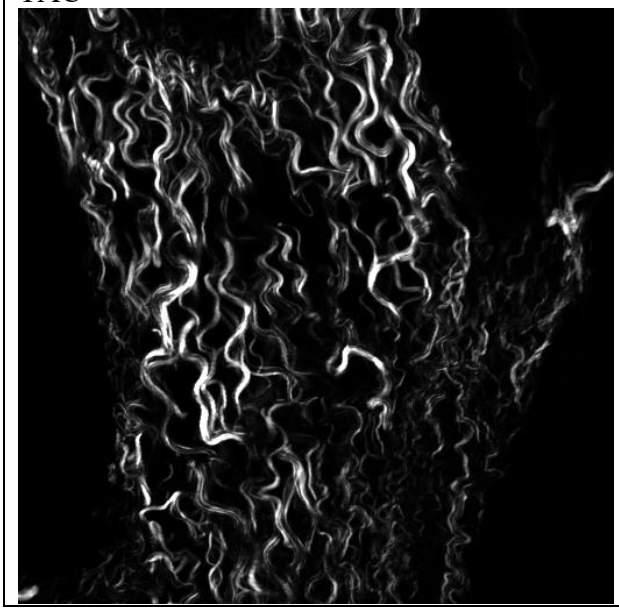
Control



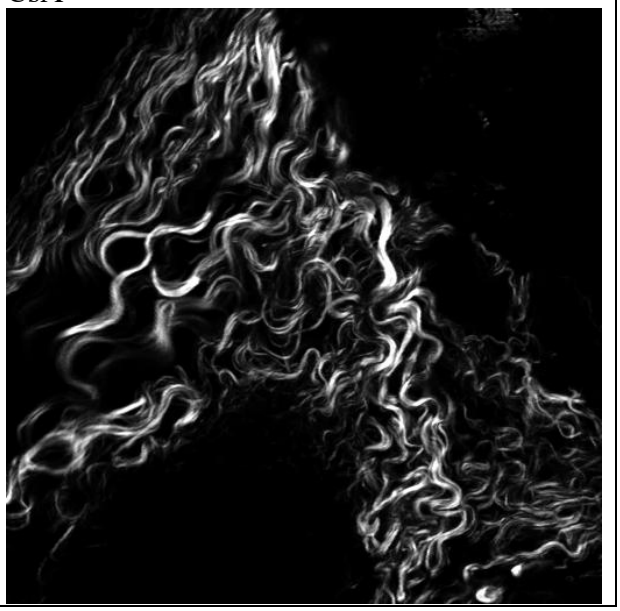
UUO



TAC



CsA



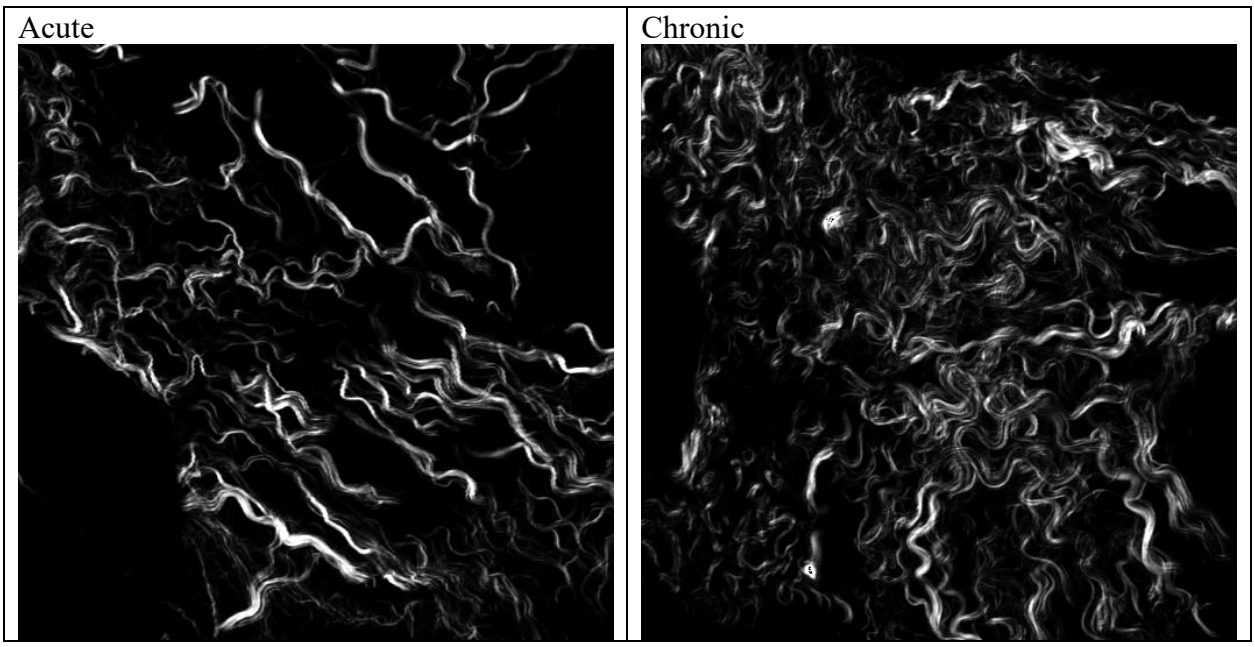
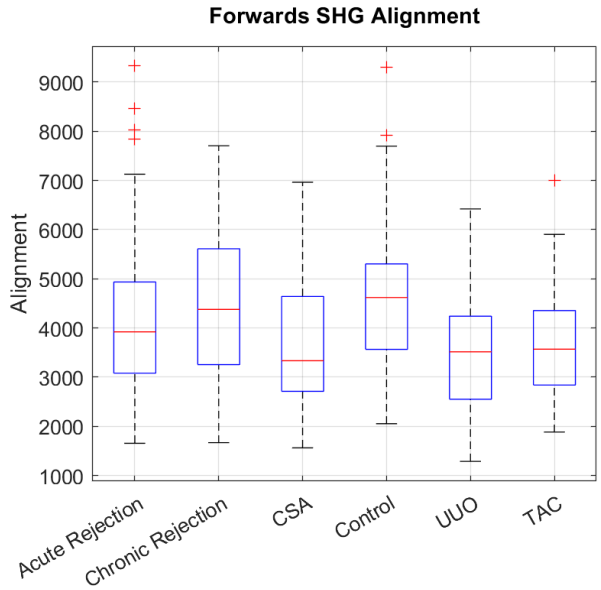
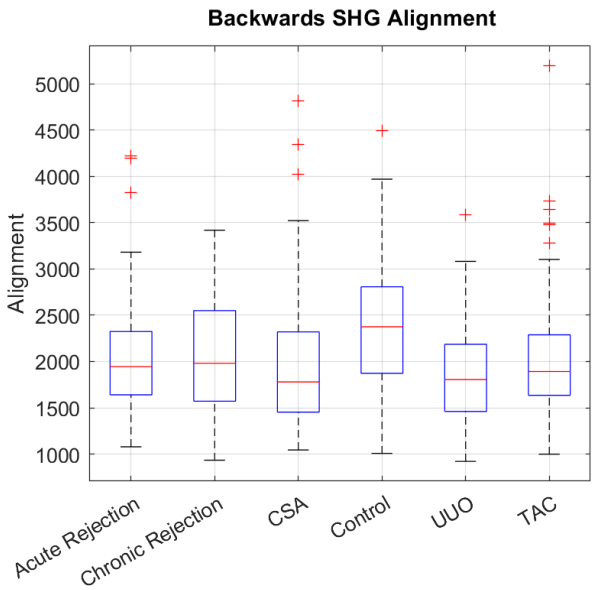
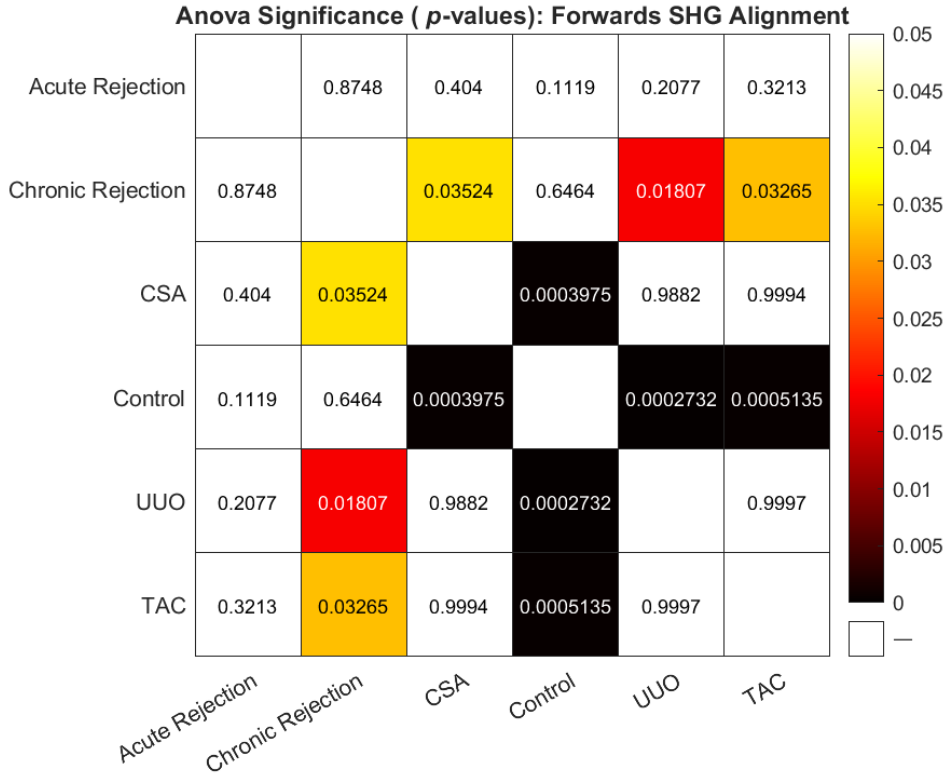


Figure 2.18, SHG Images of Kidney Murine Models. Representative SHG images of all analyzed groups from kidney murine models.

The ANOVA of alignment showed three groups differed from the control: CSA, UUO, and TAC. The backwards images similarly showed significance for both CSA and UUO from the control group (Figure 2.19). The results indicate a potential flaw in using 2D FFT for analyzing collagen fiber alignment. The forwards results tend to show more collagen fibers in these thinner samples and produce higher average alignments than the backwards images with fewer fibers.


Forwards SHG Alignment
Alignment
9000
8000
7000
6000
5000
4000
3000
2000
1000
Acute Rejection
Chronic Rejection
CSA
Control
UUO
TAC


Backwards SHG Alignment
Alignment
5000
4500
4000
3500
3000
2500
2000
1500
1000
Acute Rejection
Chronic Rejection
CSA
Control
UUO
TAC


Anova Significance (p-values): Forwards SHG Alignment
Acute Rejection 0.8748 0.404 0.1119 0.2077 0.3213
Chronic Rejection 0.8748 0.03524 0.6464 0.01807 0.03265
CSA 0.404 0.03524 0.0003975 0.9882 0.9994
Control 0.1119 0.6464 0.0003975 0.0002732 0.0005135
UUO 0.2077 0.01807 0.9882 0.0002732 0.9997
TAC 0.3213 0.03265 0.9994 0.0005135 0.9997
Acute Rejection
Chronic Rejection
CSA
Control
UUO
TAC
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0

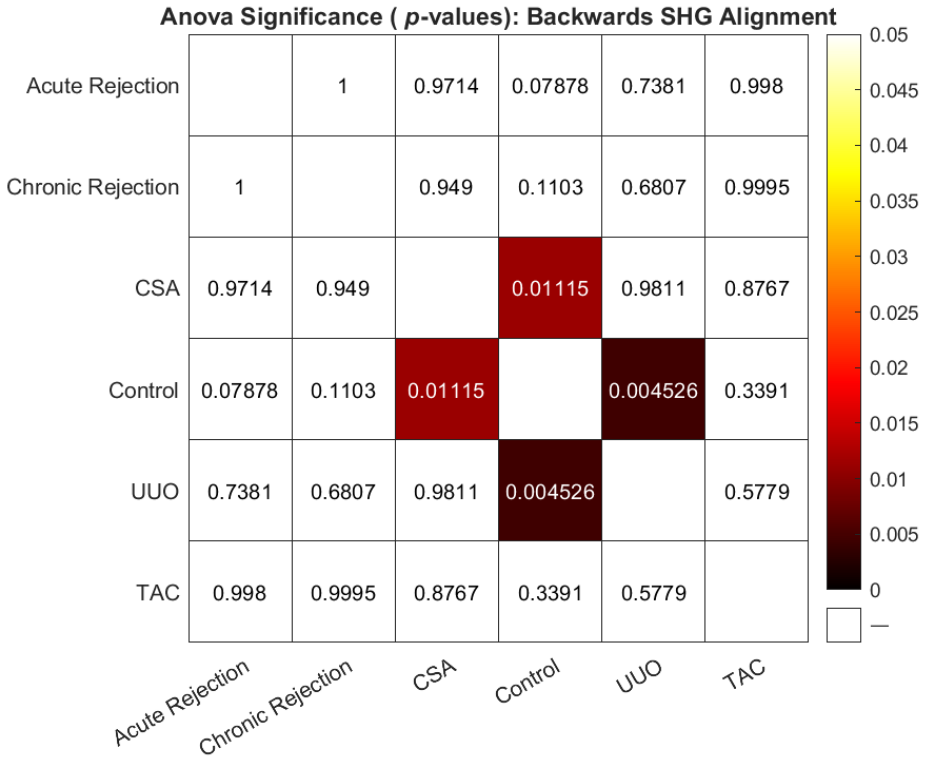
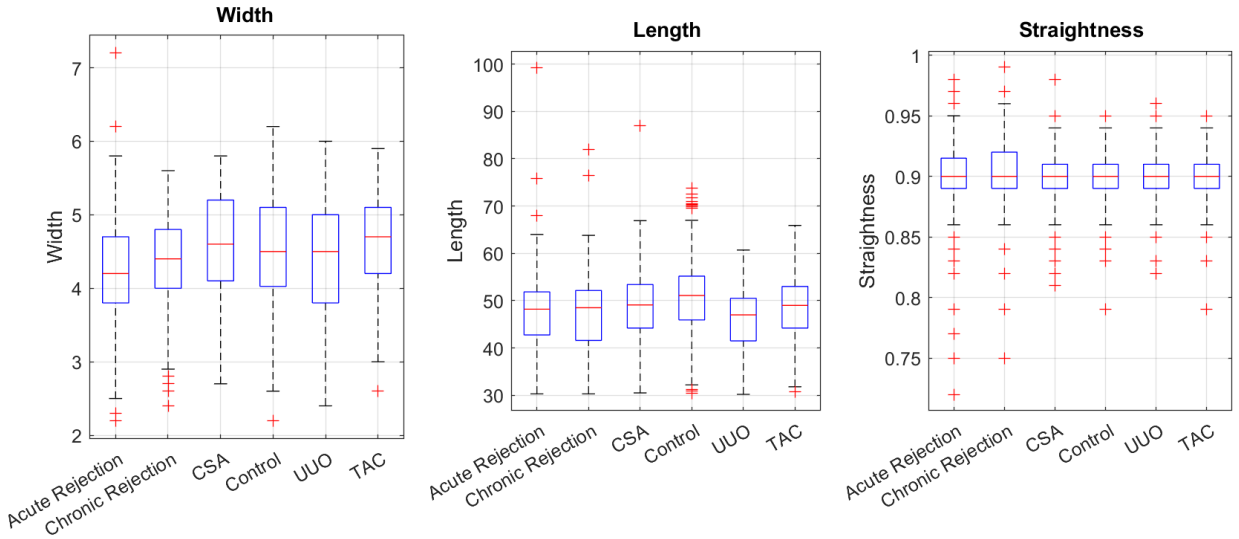
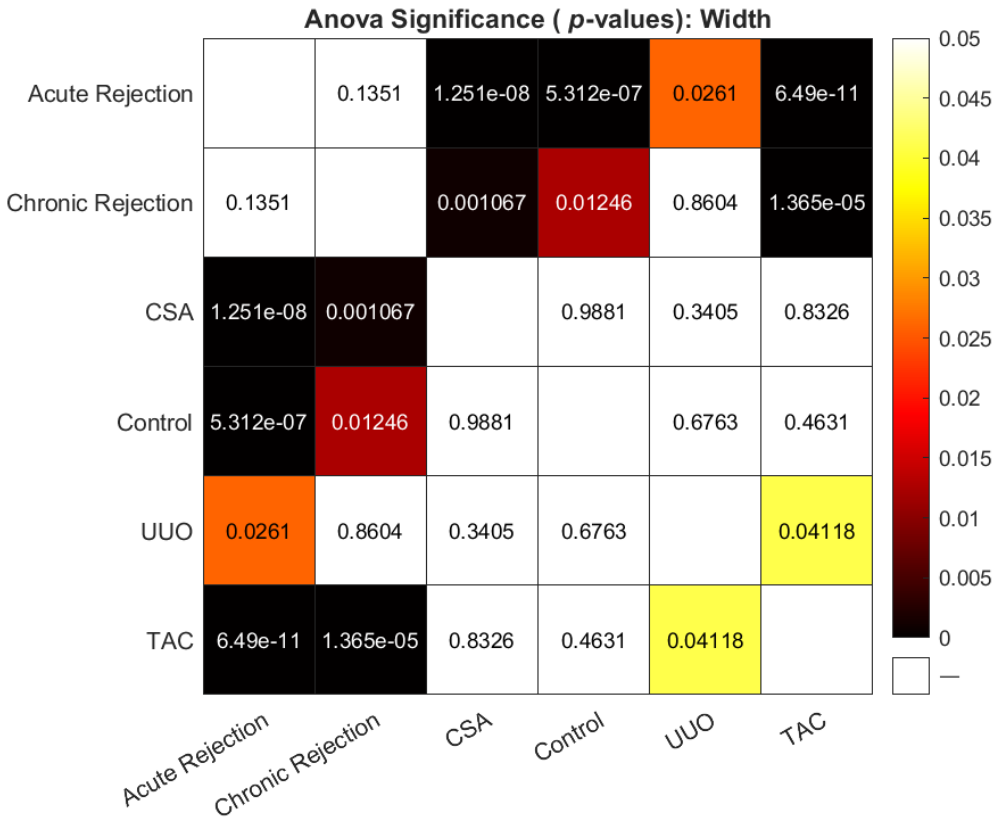
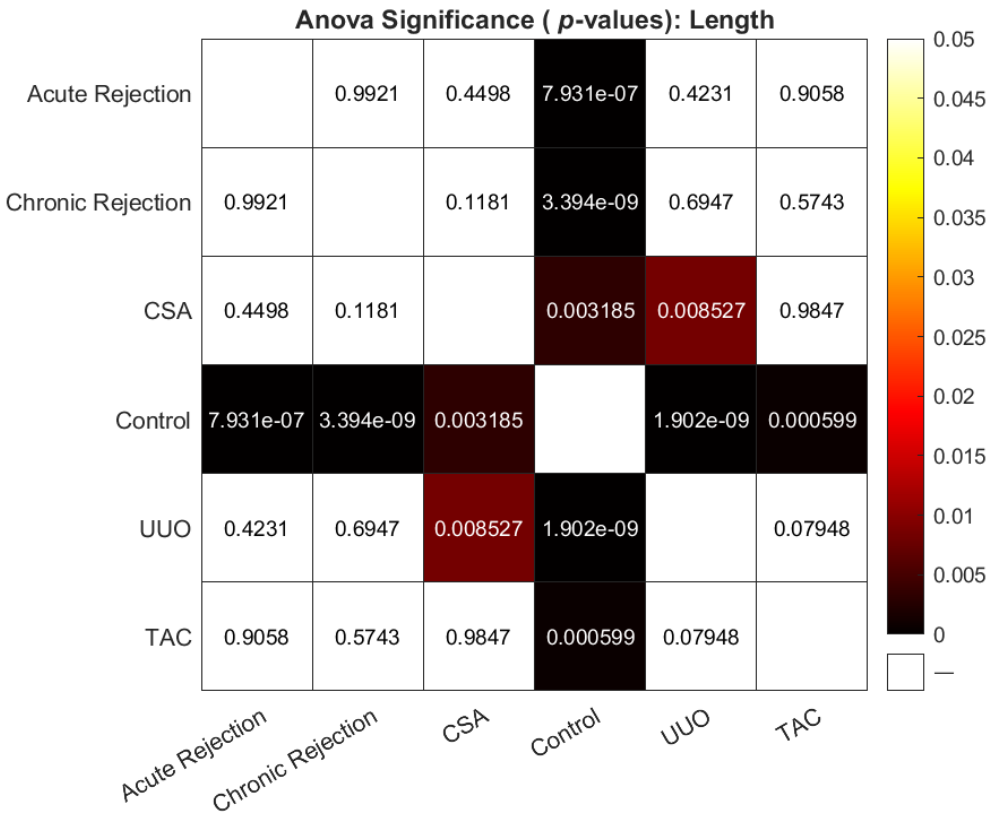


Figure 2.19, Murine Kidney Models Analyzed with 2D FFT. ANOVA results for the alignment feature calculated with a 2D FFT transform. Heatmap of ANOVA *p*-values for forwards and backwards images.

CT-Fire results examined average length, width, and straightness of the collagen fibers. An ANOVA analysis was run on each feature along with creating corresponding *p*-value heatmaps (Figure 2.20). Width had two groups, acute rejection and chronic rejection, that differed from the control. Length was significantly different for all groups when compared to the control. Finally, straightness did not differ from the control group, but there were still a couple of significant results between groups.


Width
Width
Length
Length
Straightness
Straightness
Acute Rejection
Chronic Rejection
CSA
Control
UUO
TAC


Anova Significance (p-values): Width
Acute Rejection
Chronic Rejection
CSA
Control
UUO
TAC
0.1351
1.251e-08
5.312e-07
0.0261
6.49e-11
0.1351
0.001067
0.01246
0.8604
1.365e-05
1.251e-08
0.001067
0.9881
0.3405
0.8326
5.312e-07
0.01246
0.9881
0.6763
0.4631
0.0261
0.8604
0.3405
0.6763
0.04118
6.49e-11
1.365e-05
0.8326
0.4631
0.04118
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0


Anova Significance (p-values): Length
Acute Rejection
Chronic Rejection
CSA
Control
UUO
TAC
0.9921
0.4498
7.931e-07
0.4231
0.9058
0.9921
0.1181
3.394e-09
0.6947
0.5743
0.4498
0.1181
0.003185
0.008527
0.9847
7.931e-07
3.394e-09
0.003185
1.902e-09
0.000599
0.4231
0.6947
0.008527
1.902e-09
0.07948
0.9058
0.5743
0.9847
0.000599
0.07948
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0

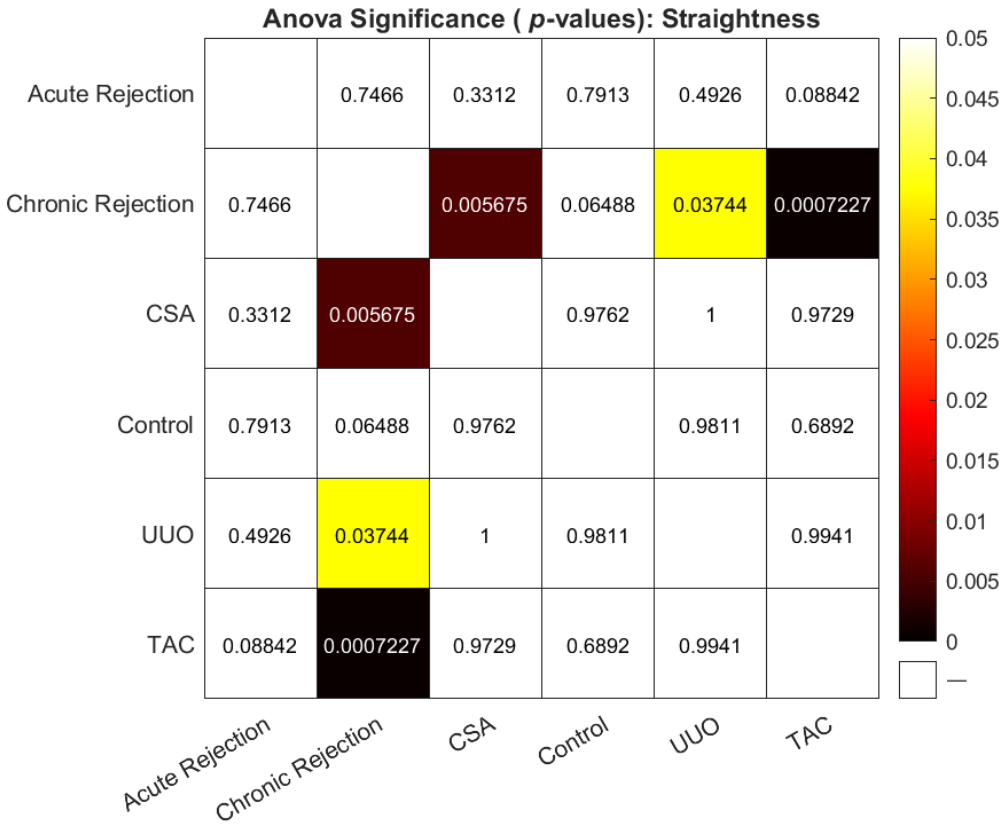


Figure 2.20, Murine Kidney Models Analyzed with CT-Fire. ANOVA results for width, length, and straightness calculated with CT-Fire. Heatmaps of ANOVA *p*-values for all features.

Chapter 2 References

1. Chen, X., Nadiarynkh, O., Plotnikov, S. & Campagnola, P. J. Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure. *Nat. Protoc.* **7**, 654–669 (2012).
2. Lien, C.-H., Tilbury, K., Chen, S.-J. & Campagnola, P. J. Precise, motion-free polarization control in Second Harmonic Generation microscopy using a liquid crystal modulator in the infinity space. *Biomed. Opt. Express* **4**, 1991–2002 (2013).
3. Liu, Y. & Eliceiri, K. W. Quantifying Fibrillar Collagen Organization with Curvelet Transform-Based Tools. *J. Vis. Exp. JoVE* (2020) doi:10.3791/61931.
4. Y, L., A, K., GS, M., CR, D. & KW, E. Methods for Quantifying Fibrillar Collagen Alignment. *Methods Mol. Biol. Clifton NJ* **1627**, 429–451 (2017).
5. Kowalska-Kępczyńska, A. Systemic Scleroderma—Definition, Clinical Picture and Laboratory Diagnostics. *J. Clin. Med.* **11**, 2299 (2022).
6. Varga, J. & Abraham, D. Systemic sclerosis: a prototypic multisystem fibrotic disorder. *J. Clin. Invest.* **117**, 557–567 (2007).
7. Volkmann, E. R., Andréasson, K. & Smith, V. Systemic sclerosis. *Lancet Lond. Engl.* **401**, 304 (2023).
8. Al-Dhaher, F. F., Pope, J. E. & Ouimet, J. M. Determinants of morbidity and mortality of systemic sclerosis in Canada. *Semin. Arthritis Rheum.* **39**, 269–277 (2010).
9. Elhai, M. *et al.* Mapping and predicting mortality from systemic sclerosis. *Ann. Rheum. Dis.* **76**, 1897–1905 (2017).
10. Nassar, M. *et al.* Gastrointestinal involvement in systemic sclerosis: An updated review. *Medicine (Baltimore)* **101**, E31780 (2022).
11. Careta, M. F. & Romiti, R. Localized scleroderma: clinical spectrum and therapeutic update. *An. Bras. Dermatol.* **90**, 62–73 (2015).
12. Shin, J. Y. *et al.* Epigenetic activation and memory at a TGFB2 enhancer in systemic sclerosis. *Sci. Transl. Med.* **11**, (2019).
13. Frangogiannis, N. G. Transforming growth factor-β in tissue fibrosis. *J. Exp. Med.* **217**, (2020).
14. Lafyatis, R. Transforming growth factor β—at the centre of systemic sclerosis. *Nat. Rev. Rheumatol. 2014 1012* **10**, 706–719 (2014).
15. Pohlers, D. *et al.* TGF-beta and fibrosis in different organs - molecular pathway imprints. *Biochim. Biophys. Acta* **1792**, 746–756 (2009).
16. Sapao, P. *et al.* Reduced SPAG17 Expression in Systemic Sclerosis Triggers Myofibroblast Transition and Drives Fibrosis. *J. Invest. Dermatol.* **143**, (2023).
17. Campagnola, P. Second harmonic generation imaging microscopy: applications to diseases diagnostics. *Anal. Chem.* **83**, 3224–3231 (2011).
18. James, D. S. & Campagnola, P. J. Recent Advancements in Optical Harmonic Generation Microscopy: Applications and Perspectives. *BME Front.* **2021**, 3973857 (2021).

19. Nadiarnykh, O. *et al.* Second harmonic generation imaging microscopy studies of osteogenesis imperfecta. *J. Biomed. Opt.* **12**, 051805 (2007).
20. Keikhosravi, A., Bredfeldt, J. S., Sagar, A. K. & Eliceiri, K. W. Second-harmonic generation imaging of cancer. *Methods Cell Biol.* **123**, 531–546 (2014).
21. Miyaaki, H. *et al.* Liver fibrosis analysis using digital pathology. *Med. Mol. Morphol.* **57**, 161–166 (2024).
22. Pichon, J. *et al.* Label-free 3D characterization of cardiac fibrosis in muscular dystrophy using SHG imaging of cleared tissue. *Biol. Cell* **114**, 91–103 (2022).
23. Ricard-Blum, S., Baffet, G. & Théret, N. Molecular and tissue alterations of collagens in fibrosis. *Matrix Biol. J. Int. Soc. Matrix Biol.* **68–69**, 122–149 (2018).
24. Tilbury, K. *et al.* Second harmonic generation microscopy analysis of extracellular matrix changes in human idiopathic pulmonary fibrosis. *J. Biomed. Opt.* **19**, 086014 (2014).
25. Wen, B. *et al.* 3D texture analysis for classification of second harmonic generation images of human ovarian cancer. *Sci. Rep.* **6**, (2016).
26. Rentchler, E. C., Gant, K. L., Drapkin, R., Patankar, M. & J Campagnola, P. Imaging Collagen Alterations in STICs and High Grade Ovarian Cancers in the Fallopian Tubes by Second Harmonic Generation Microscopy. *Cancers* **11**, 1805 (2019).
27. Lacomb, R., Nadiarnykh, O. & Campagnola, P. J. Quantitative second harmonic generation imaging of the diseased state osteogenesis imperfecta: experiment and simulation. *Biophys. J.* **94**, 4504–4514 (2008).
28. Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L. & Lopez, A. A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing* **408**, 189–215 (2020).
29. Bi, X., Mills, T. & Wu, M. Animal models in systemic sclerosis: an update. *Curr. Opin. Rheumatol.* **35**, 364–370 (2023).
30. Asano, Y. Recent advances in animal models of systemic sclerosis. *J. Dermatol.* **43**, 19–28 (2016).
31. Takahashi, T. *et al.* Epithelial Fli1 deficiency drives systemic autoimmunity and fibrosis: Possible roles in scleroderma. *J. Exp. Med.* **214**, 1129–1151 (2017).
32. Mostaço-Guidolin, L., Rosin, N. L. & Hackett, T.-L. Imaging Collagen in Scar Tissue: Developments in Second Harmonic Generation Microscopy for Biomedical Applications. *Int. J. Mol. Sci.* **18**, 1772 (2017).
33. Rosin, N. L. *et al.* Collagen structural alterations contribute to stiffening of tissue after split-thickness skin grafting. *Wound Repair Regen.* **24**, 263–274 (2016).
34. James, D. S. *et al.* Probing ECM remodeling in idiopathic pulmonary fibrosis via second harmonic generation microscopy analysis of macro/supramolecular collagen structure. *J. Biomed. Opt.* **25**, 014505 (2020).
35. Conklin, M. W. *et al.* Aligned collagen is a prognostic signature for survival in human breast carcinoma. *Am. J. Pathol.* **178**, 1221–1232 (2011).

36. Drifka, C. R. *et al.* Highly aligned stromal collagen is a negative prognostic factor following pancreatic ductal adenocarcinoma resection. *Oncotarget* **7**, 76197–76213 (2016).
37. Teves, M. E., Strauss, J. F., Sapao, P., Shi, B. & Varga, J. The Primary Cilium: Emerging Role as a Key Player in Fibrosis. *Curr. Rheumatol. Rep.* **21**, (2019).
38. Teves, M. E. *et al.* Spag17 deficiency results in skeletal malformations and bone abnormalities. *PloS One* **10**, e0125936–e0125936 (2015).
39. Roberson, E. D. O. O. *et al.* Alterations of the Primary Cilia Gene SPAG17 and SOX9 Locus Noncoding RNAs Identified by RNA-Sequencing Analysis in Patients With Systemic Sclerosis. *Arthritis Rheumatol. Hoboken NJ* **75**, 108–119 (2023).
40. Teves, M. E. *et al.* Sperm-associated antigen-17 gene is essential for motile cilia function and neonatal survival. *Am. J. Respir. Cell Mol. Biol.* **48**, 765–772 (2013).
41. Campagnola, P. J. *et al.* Three-dimensional high-resolution second-harmonic generation imaging of endogenous structural proteins in biological tissues. *Biophys. J.* **82**, 493–508 (2002).
42. Christner, P. J. & Jimenez, S. A. Animal models of systemic sclerosis: insights into systemic sclerosis pathogenesis and potential therapeutic approaches. *Curr. Opin. Rheumatol.* **16**, 746–752 (2004).
43. Yamamoto, T. & Nishioka, K. Cellular and molecular mechanisms of bleomycin-induced murine scleroderma: current update and future perspective. *Exp. Dermatol.* **14**, 81–95 (2005).
44. Calderon, L. M. & Pope, J. E. Scleroderma epidemiology update. *Curr. Opin. Rheumatol.* **33**, 122–127 (2021).
45. Gegenava, T. *et al.* Sex-specific difference in cardiac function in patients with systemic sclerosis: association with cardiovascular outcomes. *RMD Open* **9**, (2023).
46. Khanna, D. *et al.* Etiology, Risk Factors, and Biomarkers in Systemic Sclerosis with Interstitial Lung Disease. *Am. J. Respir. Crit. Care Med.* **201**, 650–660 (2020).
47. Ciaffi, J., van Leeuwen, N. M., Schoones, J. W., Huizinga, T. W. J. & de Vries-Bouwstra, J. K. Sex hormones and sex hormone-targeting therapies in systemic sclerosis: A systematic literature review. *Semin. Arthritis Rheum.* **50**, 140–148 (2020).
48. Hughes, M. *et al.* Gender-related differences in systemic sclerosis. *Autoimmun. Rev.* **19**, (2020).
49. Chizzolini, C., Brembilla, N. C., Montanari, E. & Truchetet, M. E. Fibrosis and immune dysregulation in systemic sclerosis. *Autoimmun. Rev.* **10**, 276–281 (2011).
50. Whitfield, M. L. *et al.* Systemic and cell type-specific gene expression patterns in scleroderma skin. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 12319–12324 (2003).
51. LaComb, R., Nadiarnykh, O., Carey, S. & Campagnola, P. J. Quantitative second harmonic generation imaging and modeling of the optical clearing mechanism in striated muscle and tendon. *J. Biomed. Opt.* **13**, 021109 (2008).
52. Granieri, M. A., Peterson, A. C. & Madden-Fuentes, R. J. Effect of Lichen Sclerosis on Success of Urethroplasty. *Urol. Clin. North Am.* **44**, 77–86 (2017).

53. Fergus, K. B. *et al.* Pathophysiology, Clinical Manifestations, and Treatment of Lichen Sclerosus: A Systematic Review. *Urology* **135**, 11–19 (2020).
54. Levy, A. *et al.* Insights into the Pathophysiology of Urethral Stricture Disease due to Lichen Sclerosus: Comparison of Pathological Markers in Lichen Sclerosus Induced Strictures vs Nonlichen Sclerosus Induced Strictures. *J. Urol.* **201**, 1158–1163 (2019).
55. Querejeta, R. *et al.* Increased collagen type I synthesis in patients with heart failure of hypertensive origin: relation to myocardial fibrosis. *Circulation* **110**, 1263–1268 (2004).
56. Duffield, J. S. Cellular and molecular mechanisms in kidney fibrosis. *J. Clin. Invest.* **124**, 2299–2306 (2014).
57. Liu, S. B. *et al.* Lysyl oxidase activity contributes to collagen stabilization during liver fibrosis progression and limits spontaneous fibrosis reversal in mice. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **30**, 1599–1609 (2016).
58. Wollensak, G., Wilsch, M., Spoerl, E. & Seiler, T. Collagen fiber diameter in the rabbit cornea after collagen crosslinking by riboflavin/UVA. *Cornea* **23**, 503–507 (2004).
59. Da-Silva, E. A., Sampaio, F. J. B., Dornas, M. C., Damiao, R. & Cardoso, L. E. M. Extracellular matrix changes in urethral stricture disease. *J. Urol.* **168**, 805–807 (2002).
60. Sahani, R., Wallace, C. H., Jones, B. K. & Blemker, S. S. Diaphragm muscle fibrosis involves changes in collagen organization with mechanical implications in Duchenne muscular dystrophy. *J. Appl. Physiol. Bethesda Md 1985* **132**, 653–672 (2022).
61. Cox, T. R. *et al.* LOX-mediated collagen crosslinking is responsible for fibrosis-enhanced metastasis. *Cancer Res.* **73**, 1721–1732 (2013).
62. Erickson, B. A. *et al.* Urethral lichen sclerosus under the microscope: a survey of academic pathologists. *Can. J. Urol.* **25**, 9328–9333 (2018).
63. Huang, Y. *et al.* Image analysis of liver collagen using sirius red is more accurate and correlates better with serum fibrosis markers than trichrome. *Liver Int. Off. J. Int. Assoc. Study Liver* **33**, 1249–1256 (2013).
64. Wegner, K. A., Keikhosravi, A., Eliceiri, K. W. & Vezina, C. M. Fluorescence of Picrosirius Red Multiplexed With Immunohistochemistry for the Quantitative Assessment of Collagen in Tissue Sections. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* **65**, 479–490 (2017).
65. Bauman, T. M. *et al.* Characterization of fibrillar collagens and extracellular matrix of glandular benign prostatic hyperplasia nodules. *PloS One* **9**, e109102 (2014).
66. Calvaruso, V. *et al.* Computer-assisted image analysis of liver collagen: relationship to Ishak scoring and hepatic venous pressure gradient. *Hepatol. Baltim. Md* **49**, 1236–1244 (2009).
67. Courtoy, G. E. *et al.* Digital Image Analysis of Picrosirius Red Staining: A Robust Method for Multi-Organ Fibrosis Quantification and Characterization. *Biomolecules* **10**, 1585 (2020).
68. Almeida, H. L. de, Bicca, E. de B. C., Breunig, J. de A., Rocha, N. M. & Silva, R. M. e. Scanning electron microscopy of lichen sclerosus. *An. Bras. Dermatol.* **88**, 247–249 (2013).
69. Godoy, C. A. P. *et al.* Unusual remodeling of the hyalinization band in vulval lichen sclerosus by type V collagen and ECM 1 protein. *Clin. Sao Paulo Braz.* **70**, 356–362 (2015).

70. Marren, P., Dean, D., Charnock, M. & Wojnarowska, F. The basement membrane zone in lichen sclerosus: an immunohistochemical study. *Br. J. Dermatol.* **136**, 508–514 (1997).
71. Panizzon, R., Vuorio, T. & Bruckner-Tuderman, L. Collagen biosynthesis and type I and type III procollagen mRNA in lichen sclerosus et atrophicus. *Arch. Dermatol. Res.* **282**, 480–483 (1990).
72. Jones, M. G. *et al.* Nanoscale dysregulation of collagen structure-function disrupts mechano-homeostasis and mediates pulmonary fibrosis. *eLife* **7**, e36354 (2018).
73. Liu, F. *et al.* Mechanosignaling through YAP and TAZ drives fibroblast activation and fibrosis. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **308**, L344-357 (2015).
74. Parker, M. W. *et al.* Fibrotic extracellular matrix activates a profibrotic positive feedback loop. *J. Clin. Invest.* **124**, 1622–1635 (2014).
75. Tsochatzis, E. *et al.* Collagen proportionate area is superior to other histological methods for sub-classifying cirrhosis and determining prognosis. *J. Hepatol.* **60**, 948–954 (2014).
76. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA. Cancer J. Clin.* **72**, 7–33 (2022).
77. Akinjiyan, F. A., Dave, R. M., Alpert, E., Longmore, G. D. & Fuh, K. C. DDR2 Expression in Cancer-Associated Fibroblasts Promotes Ovarian Cancer Tumor Invasion and Metastasis through Periostin-ITGB1. *Cancers* **14**, 3482 (2022).
78. Vogel, W., Gish, G. D., Alves, F. & Pawson, T. The discoidin domain receptor tyrosine kinases are activated by collagen. *Mol. Cell* **1**, 13–23 (1997).
79. Xu, J. *et al.* Overexpression of DDR2 contributes to cell invasion and migration in head and neck squamous cell carcinoma. *Cancer Biol. Ther.* **15**, 612–622 (2014).
80. Chua, H.-H. *et al.* Upregulation of discoidin domain receptor 2 in nasopharyngeal carcinoma. *Head Neck* **30**, 427–436 (2008).
81. Corsa, C. A. S. *et al.* The Action of Discoidin Domain Receptor 2 in Basal Tumor Cells and Stromal Cancer-Associated Fibroblasts Is Critical for Breast Cancer Metastasis. *Cell Rep.* **15**, 2510–2523 (2016).
82. Hammerman, P. S. *et al.* Mutations in the DDR2 kinase gene identify a novel therapeutic target in squamous cell lung cancer. *Cancer Discov.* **1**, 78–89 (2011).
83. Ren, T. *et al.* Discoidin domain receptor 2 (DDR2) promotes breast cancer cell metastasis and the mechanism implicates epithelial-mesenchymal transition programme under hypoxia. *J. Pathol.* **234**, 526–537 (2014).
84. Nadiarnykh, O., LaComb, R. B., Brewer, M. A. & Campagnola, P. J. Alterations of the extracellular matrix in ovarian cancer studied by Second Harmonic Generation imaging microscopy. *BMC Cancer* **10**, 94 (2010).
85. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).

Chapter 3: Artificial SHG Image Feature Tuning with StyleGAN2-ADA and SeFa Semantic Analysis

3.1 Summary Statement

Collagen fibers within the extracellular matrix undergo morphological changes in the presence of many diseases including cancers and connective tissue disorders. Second Harmonic Generation (SHG) microscopy is ideal for visualizing this morphology by its natural optical sectioning of fibrillar collagen. Given the success of deep learning based methods analyzing microscopy datasets, in this study, we trained a generative adversarial network (GAN) model (StyleGAN2-ADA) on a dataset of 1,319 SHG images. When comparing evaluation metrics between the baseline dataset and the generator, we find that the model is well suited to generating novel images that appear perceptually similar to those from our small set of training data. Using Semantic Factorization (SeFa) on the generated images to examine the 10 highest scoring eigenvectors, we find two semantics with strong correlation to fiber morphological features that were quantified via analysis using CT-Fire, CurveAlign, and FIJI). We integrated analysis tools into the SeFa GUI to support an extensible framework for future analyses and image editing. Additionally, we tested two image projection methods for editing SHG images in the GAN's latent space using the identified semantics. The encoder-based approach significantly outperformed the Learned Perceptual Image Patch Similarity (LPIPS) method. With further refinement, these tools can be used to create SHG images of collagen fibers with variable morphology based on disease types.

3.2 Introduction

Fibrillar collagen serves a key role in structural organization, especially within the extracellular matrix (ECM). The mechanical and structural properties of collagen architecture are altered in many diseases including cancers, fibroses, and connective tissue disorders[1-3]. Due to its sensitivity/specificity, Second Harmonic Generation (SHG) microscopy has emerged as a powerful tool to visualize collagen morphology in a wide range of these and other diseased states. Analysis of this morphology can potentially lead to earlier diagnoses and better understanding of disease advancement.

We have investigated the application of machine learning techniques to classify alterations from normal collagen structure. We previously used "textons" (or texture features) to classify a series of tissues related to ovarian cancer. Texture features were identified using image responses to sets of filter banks, and a representative model was generated for specific tissue types based on the distribution of the respective features in each class[4]. Ultimately, this method was able to achieve a classification accuracy >90% using one versus the rest classification. We also used principal component analysis of 2D wavelet transforms of idiopathic pulmonary fibrosis and classified normal versus fibrotic disease with >95% accuracy. Unfortunately, identification methods of the most important and differentiating aspects of each class of tissue in other systems often does not produce this level of accuracy. Our overarching goal is to use SHG images as blueprints to fabricate tissues, and engineering image-based scaffolds to study the cell biology of disease progression[5]. To model and study such behavior, it is essential to go beyond classification – we seek to develop a high throughput pipeline to optimize biomimetic model design, fabrication, and testing. To this end, we describe how trained generative models, can help in

exploring and editing the morphological features of synthetically generated collagen fiber images.

In recent years, GANs have been used in an increasingly broad range of applications, including specific use cases in microscopy. However, training GANs with small sample sizes remains challenging. Even in settings where sufficient data in terms of sample sizes is available, it is not always usable: in some modalities, resolution can be weak, or images can be noisy. But GANs, once successfully trained, can be used to enhance images (e.g., sharpness, noise, and clarity). By training GANs on labeled data, they can be used to automatically segment datasets by identifying characteristic features such as tumors[6,7]. In addition to these tasks, microscopy modalities can use GANs for virtual staining, generating realistic H&E, Masson's Trichrome, and Jones' images from unlabeled data. Similarly, digital fluorescence labeling can be performed on brightfield or phase contrast images[8,9]. Several approaches have achieved strong results including models based on StyleGAN, CycleGAN, SPADE GAN, and HingeGAN[10] architectures. Typically, these generators are evaluated based on the Fréchet inception distance (FID), which measures the feature vector distance between synthetic and real images. Style-based GANs have achieved particularly good FIDs, indicating successful generation of quality image data reflective of that available in microscopy[11]. Importantly, Style-GANs have been used to extract image features and assess their editability. Latent feature analysis via StyleGAN has been applied to human faces where the layers correspond to examples such as head shape, hair style, and color scheme[12]. To our knowledge, latent space feature analysis and editing within the generator for microscopy datasets has remained somewhat limited so far.

Our model is based on StyleGAN2-ADA architecture, in large part due to its broad use in the computer vision community and empirical performance relative to other backbones in our initial testing. Its latent space is also accessible, which makes it well suited for our use case. Briefly, StyleGAN allows the generator network to alter specific features of the image through the use of an intermediate latent space compared to the usual Gaussian space. The latent space (for deep generative models, but also for deep models in general) is typically high-dimensional. It allows us to interpret representations or embeddings of the data (e.g., images) learned by the model[13]. StyleGAN uses so-called style blocks empirically. By encouraging individual style blocks to behave "independently", StyleGAN can achieve a more linear and disentangled latent space[14]. Also, layers in the model network often correspond to different image styles, which consequently allows easier image editability[15]. When StyleGAN is trained on SHG images, we expect the styles to correspond to different features of collagen fibers such as density, length, width, curvature, and alignment[16]. A recent version of this architecture family, StyleGAN2-ADA, is well-suited to microscopy images because it uses adaptive discriminator augmentation to train with limited data, dynamically adjusting input into the generator based on the degree of overfitting, greatly improving training convergence for limited datasets such as those of human and animal tissues[17].

By training StyleGAN on a database of SHG images of both normal and diseased tissue, we generated latent codes that model features for collagen networks. Analyzing the generator's affine layers using an eigenvalue decomposition with SeFa[18], we found interpretable editing directions. Real images can then be encoded and edited along latent directions that correspond to both known and unknown features that reflect disease states. These editable latent directions are often referred to as semantics. In principle, this approach could create more accurate biomimetic

models and yield models for disease progression. It could also alleviate unique difficulties we face in small microscopy-based training sets[19]. When developed further, assessing structural changes between normal and high-grade tissues could lead to better understanding of tumor progression, diagnostic possibilities, or even possible therapeutic targets. Here, we use the external analysis tools for collagen fiber features to quantify the behavior of the GAN-based learned semantics.

3.3 Methods

Sample Library

SHG images (curated from both published and unpublished work) were previously collected from several collagen-containing tissues using our established imaging protocols[20]. An open-source ImageJ platform for image analysis (FIJI) was used to process all images. A custom macro was built to transform raw tiff images to 8-bit greyscale 512x512 PNG images. Brightness and contrast were tuned to reduce overall background speckle. A set of images from an older version of the microscope software had artifacts on the brightest fibers. Custom Matlab2023a code was used to process these artifacts using nearest-neighbor pixel by pixel correction to reduce the possible effect of this error in training.

Generative Adversarial Network

The StyleGAN2 algorithm is described in detail in Karras *et al*. [15,21]. Briefly, a GAN consists of a generator and discriminator network that work in conjunction to create a distribution of synthetic images that will ideally be statistically indistinguishable from the distribution of the real training set. It is the discriminator's role to check whether the two distributions (synthetic versus real) are indistinguishable In StyleGAN, the discriminator is a progressively grown network[22]. The generator is broken into two parts: the mapping network *f*, which inputs a latent vector $\mathbf{z} \in Z$ onto

the auxiliary latent space ***W***, and the synthesis network *g*, which maps vector **w** ε ***W*** to the output image. The distribution of the latent codes in ***W*** is learned by mapping network *f*, which allows for a more disentangled latent space than in Z space. The synthesis network consists of a series of style blocks that each consist of modulation, demodulation (normalization), and convolution. Biases and noise broadcast operation B are added outside the active area of the style so that they affect only normalized data (Figure 3.1).

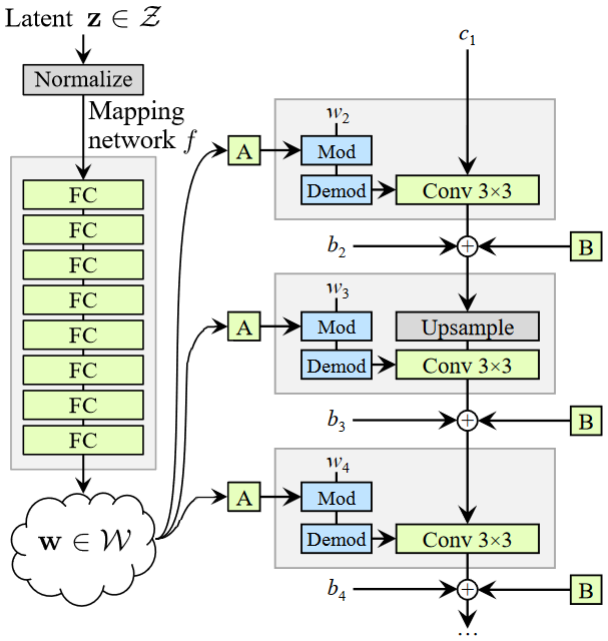


Figure 3.1. An adapted diagram of StyleGAN2 from Karras *et al.*[20, 21]. The latent code **z** from the input space **Z** is mapped to a network (*f*) of fully connected (FC) layers. It is then mapped to an intermediate latent space *W*. In the generator network, "A" is a learned affine transformation, and "B" is a learned per-channel scaling factor to add noise. The style weights are within each *w,* the bias factors are *b*, and the constant input is *c*. Each style block consists of modulation, demodulation (normalization), and convolution.

Training in StyleGAN2 is further improved with adaptive discriminator augmentation (ADA). ADA was developed for small training sets that are augmented by performing image alterations

to increase training data in order to improve results. This approach dynamically tunes image augmentation with two overfitting heuristics. Changing the probability of applying an image filter (e.g. rotation, distortion, cutouts, etc.) based on estimating what portion of the training set will get positive discriminator outputs significantly reduces the chance of overfitting in the generator model[17].

Our modified model, StyleGAN2-ADA-SHG, was implemented in Pytorch with a few alterations from the original code. For example, all inputs, layers, and outputs were changed to greyscale in order to accept SHG images. This changes the input and output size to 512 x 512 x 1, which has the downstream effect of reducing the network dimensionality. Training was conducted using 1319 SHG images, an image batch size of 16, a learning rate of 0.002, and the Adam optimizer. All machine learning training was performed on a GeForce RTX 3080 GPU with 10GB of vRAM.

Latent Space Editing with SEFA

We combined a customized version of semantic factorization (SeFa) with the StyleGAN2-ADA model along with additional GUI tools for analytics and visualization. To identify interpretable directions in the latent space of a GAN model, SeFa computes the eigenvectors of the affine transformation **A** within each style modulation that maps latent codes to modulation parameters. Finding the directions that cause large variation after the projection of **A**, yields information regarding image variation within the GAN model. The eigenvectors therein are then referred to as semantic directions that traverse the latent space of the model. This applies to StyleGAN in the transformation from latent code to style code and can interpret all or any subset of layers. The layer subsets were chosen based on StyleGAN's style blocks where early levels correspond to coarse spatial resolution and control high-level features. Middle layers correspond to smaller

scale features. Lastly, the layers towards the end correspond to higher resolution control such as pixel texture and microstructures. Subsets within the GUI are layers 0-2, 3-4, and 5-7 to represent the early, middle, and end levels respectively.

The factorization yielded a set of ranked eigen-directions, each corresponding to a potential semantic variation. For subsequent analyses, we selected the top 10 components by eigenvalue magnitude.

CT-FIRE and Image Analysis

To facilitate quantitative analysis of results from the model described above, we needed to establish a reference for comparison. We used the curvelet transform-fiber extraction (CT-FIRE) analysis[23] to determine the length, width, and straightness of the collagen fibers in the generated SHG images. This analysis first uses the fiber extraction (FIRE) algorithm[24] to de-noise images, enhance fiber edge features, and extract fibers, followed by the curvelet transform (CT) to determine fiber metrics. FIJI was also utilized in the two-dimension fast Fourier transform (2D-FFT) analysis, where two custom macros utilized the Radial Profile Extended (Philippe Carl, version 2017/04/18) and the Oval Profile Plot (Bill O'Connell, version 2012/03/01) plugins to model collagen fiber alignment[25].

Image Projection Strategies

Two strategies were implemented for projecting real and GAN images into the latent space of the generator: learned perceptual image patch similarity (LPIPS-based optimization and a custom-trained encoder. For the optimization-based method, we implemented a projector that minimizes a combination of pixel-wise mean squared error (MSE) LPIPS loss between the input image and a generated image that is synthesized from the latent code $\mathbf{w}$[26]. The LPIPS loss was computed

using a neural network, VGG, backbone[27]. An initial latent vector was selected from 10 random **z** samples by choosing the one whose mapped **w** yielded the lowest MSE. The optimization ran for 1000 steps with a learning rate of 0.01. A small amount of noise regularization was applied to suppress artifacts (see *projector.py*, available on GitHub for full details).

The encoder was adapted from the architecture used in a Stylegan2 pytorch encoder and modified for single-channel input and output (Figure 3.2) [28]. To simplify version integration, optimization was performed with adaptive moment estimation (Adam) and MSE loss[29]. Training was performed using only images synthesized from the generator. The encoder was trained with GAN images from the model for 300 epochs with a batch size of 4 and a learning rate of 5e-5.

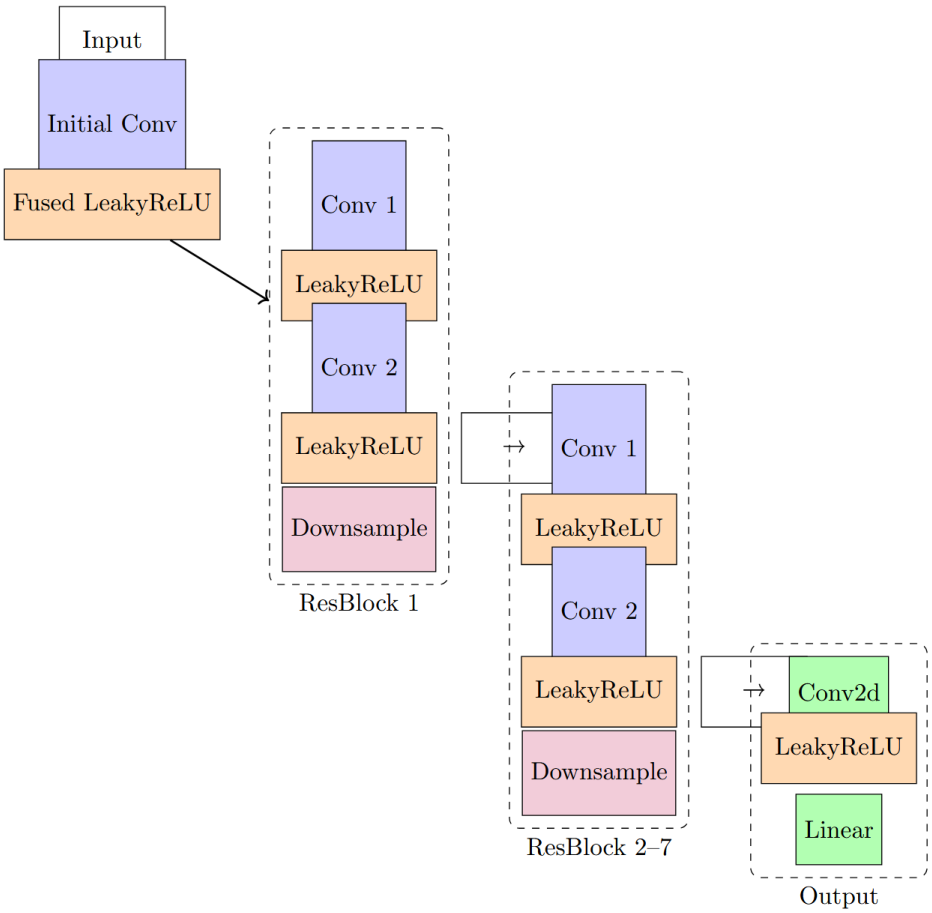


Figure 3.2. Encoder architecture adapted from Stylegan2 encoder[28]. The encoder consists of an initial convolution layer with a fused LeakyReLU activation. Stacks of ResBlock are added based on the dimensionality of the images. In this case, there are 7 total blocks consisting of an initial convolution without downsampling, then one with downsampling. Both use LeakyReLU activations. The final output is fed into an EqualConv2d layer that flattens the projection from spatial to latent space. The output is then reshaped to transfer latent codes to the GAN. Architecture was generated with code available on GitHub: https://github.com/HarisIqbal88/PlotNeuralNet.

Interface Integration and Traversal Evaluation

To inspect the effect of latent semantics on the generated samples, we integrated the encoder and projector, adding fine control to SeFa latent directions into a Streamlit graphical user interface (GUI). Users can select among the top 10 directions, adjust weights, and immediately visualize

the resulting image modifications. An initial grid was generated with 10 samples per semantic from -3 to +3 weights.

An additional tool was implemented to create customizable traversals of selected semantics. For this analysis, traversals were generated from weights −10.0 to +10.0 along a selected latent direction in increments of 0.1. Images were synthesized using latent vectors offset from the original by scalar multiples of the direction vector. These traversals were used to qualitatively evaluate the controllability and entanglement of semantic edits. Code for the interface is available at: https://github.com/Campagnola-BME-Lab/sefa_collagen

Further analysis on selected semantics used 50 random generated images traversed in increments of one. These samples were then analyzed with CT-Fire to measure all relevant features. To assess the correlation between semantic weight and fiber feature, while controlling for other feature variations, partial distance correlation (PDC) was implemented. This approach allows for conditional dependence which effectively captures the relationship between two variables after removing other effects[14]. The PDC implementation was adapted from the code available at: github.com/zhenxingjian/Partial_Distance_Correlation

Network Evaluation Measures

StyleGAN2-ADA and the encoder were both evaluated via Fréchet Inception Distance (FID) and Kernel Inception Distance (KID). FID compares the mean and covariance of real and generated image features by assuming their distribution to be a multivariate Gaussian. A lower score indicates generated images are more similar to the real ones in their distribution[30]. KID uses maximum mean discrepancy along with a polynomial kernel to compare real and generated images. As the score approaches zero the generated images are expected to match the real images. Both FID and KID use InceptionV3 as their feature extractor. This network is trained on

ImageNet, a dataset of 1.2 million RGB images with labels[31]. Due to potential statistical bias in FID's reliance on Gaussian estimation, KID can be used alongside FID as a complementary non-parametric metric for evaluating generative image quality[32]. The torch fidelity package was used to score the SHG image library as a baseline, the GAN generator, and the encoder with 10 stacks of 100 random images. To pass greyscale images to metric evaluation, the channel is repeated for each expected RGB layer. Additionally, StyleGAN2-ADA has built in FID metrics for in-training evaluation.

To further assess the baseline image database, generator, and encoder, structural similarity indexes (SSIM) were implemented for additional scoring. This index is computed locally across smaller windows of the image and then averaged, it evaluates luminance, contrast, and structure[33]. Multiscale-SSIM is an extension that computes across multiple spatial resolutions, where it progressively downsamples the image to mimic the hierarchical nature of vision[34]. MicroSSIM was developed in response to SSIM components behaving unexpectedly in images with low signal to noise conditions. Here it attempts to account for varying noise profiles and contrast levels common in microscopy datasets by incorporating background suppression and adaptive weighting strategies[35].

Fast Fourier transforms (FFTs) were used on all image databases to assess the exponential decay of the image with amplitude, offset, and the time constant. The FFT power spectrum captures the spatial frequencies contained in the image, where the fit provides information on the features in the tissues. A longer time constant will correspond to higher frequencies which in turn corresponds to smaller features in the collagen fibers[16]. Analysis was performed using Matlab2023a and further details can be found in the supplementary section.

3.4 Results

Metrics for overall training goals

To assess our GAN model, we analyzed the learned features and checked associations with currently used collagen analysis tools for features of interest including collagen fiber length, width, curvature, density, and alignment. The training set includes widely varying examples of these features. Other differences in the samples include coverage (the overall signal of fibers across the 512 x 512 image), number of fibers, and brightness. To establish a baseline for comparison, random image sets were scored with both the Fréchet Inception Distance (FID) and the Kernel Inception Distance (KID). The training data had a 74.1 ± 1.44 FID score and a 7.41E-05 ± 6.54E-04 KID score. This relatively high FID score is most likely due to the high heterogeneity inherent to most biological tissues[36,37]. FID is not typically used on the training data itself but is included here to examine the difference after machine learning training.

StyleGAN2-ADA Training and Evaluation

After our model was trained, we first visually inspected generated images alongside real SHG images to assess quality. Specifically, we checked for repetitive artifacts, background noise, and fiber separation and fidelity (Figure 3.3). Additionally, the high variability between our SHG training data means that ideally, the generated images should also be similarly variable – encompassing a variety of collagen fiber densities, sizes, and textures. Our initial visual assessment was promising – it suggested meaningful similarity between SHG and generated images (Figure 3.4). The FID-50k score from training was 25.7. Using torch fidelity, the FID score was 92.2 ± 1.53. Generally, the FID score is not expected to be below the baseline score of the training set. The KID score was 0.00583 ± 0.00090. Comparing this KID value to larger models (e.g. ImageNet and FFHQ) indicates this is an excellent score[38,39]. Small datasets with

biological heterogeneity have shown this discrepancy of higher FID scores with relatively good KID scores[40]. Stability of training was assessed by inspecting the internal FID-50K curve (Figure 3.5), which indicated that 1,319 images were likely sufficient for high generation accuracy. Full evaluation metrics, including the fast Fourier transform data, can be found in the supplemental material.

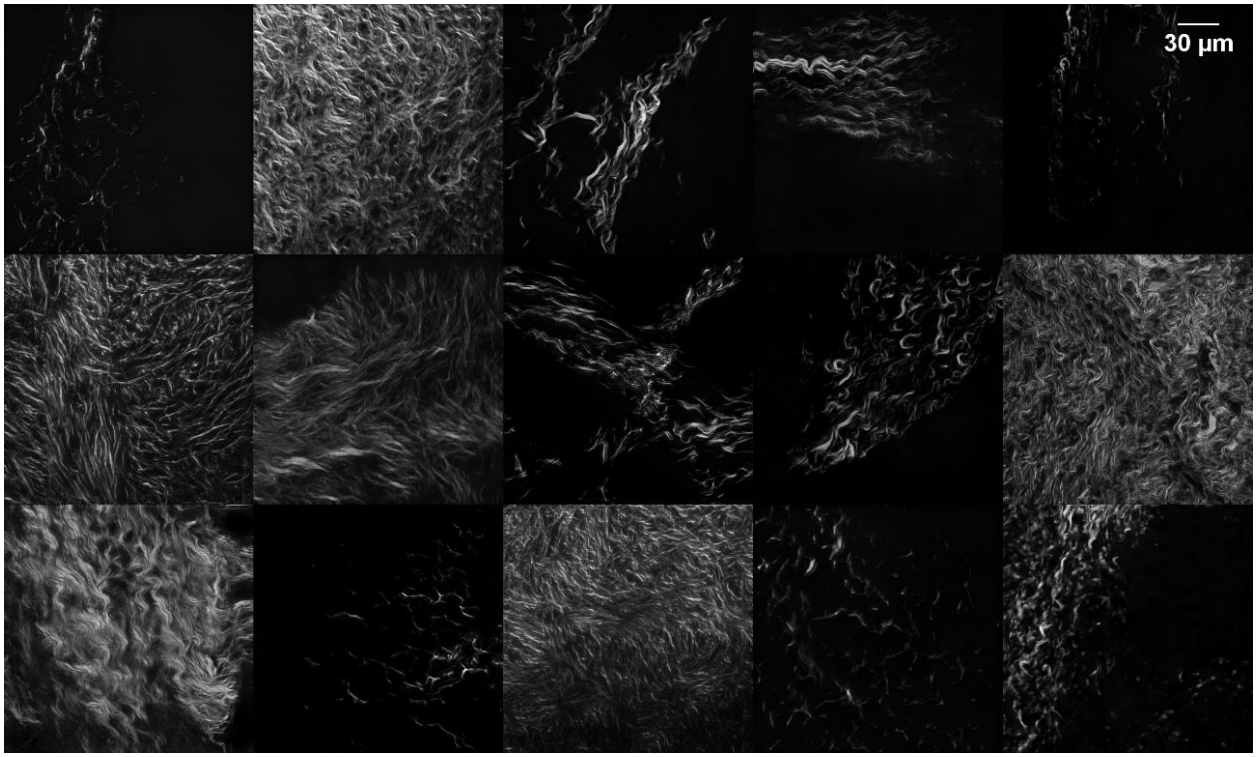


Figure 3.3 Grid of random samples from our StyleGAN2-ADA-SHG model. The generator must capture features found with collagen imaging like varying density, curvature, width, length, and alignment.

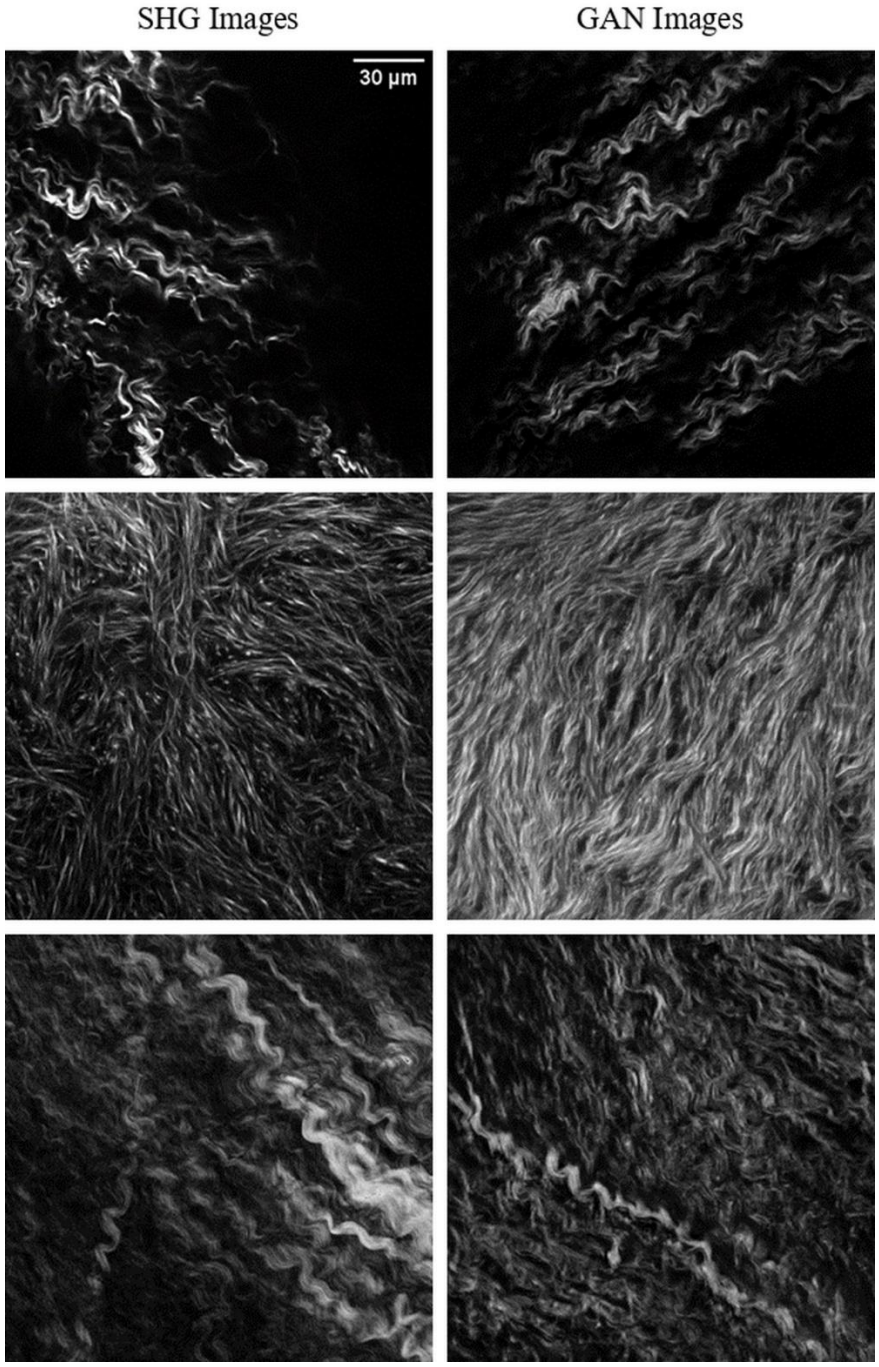


Figure 3.4. Real SHG images (left) and randomly generated StyleGAN2-ADA-SHG images with similar texture (right). Generated images were inspected for similarity to a variety of imaged tissues.

Figure 3.5. Training curve for Fréchet Inception Distance 50K as scored by StyleGAN2-ADA's internal metrics.

Initial SeFa Analysis

SeFa computes the eigenvectors within StyleGAN that transform latent code to style code. These eigenvectors are then referred to as semantic directions that can traverse the latent space of the model. SeFa semantics for the generated images were analyzed with measures derived from CT-Fire, CurveAlign, and FFT to look for semantics with significant changes to at least one feature. Two of the ten semantics showed correlation with one feature metric. All other features, including width, curvature, and mean density, did not show any significant changes. This analysis was performed with all layers (all style blocks) of the GAN (Table 3.1) on 10 random samples per semantic. Effect size was calculated to examine the magnitude of difference between semantics weighted at +3 and -3. Semantic 001, correlated with length, was measured as having an effect size of 1.1, indicating a very large influence that is visually detectable. Semantic 006 is

included for reference as this analysis indicated no significant correlation with any feature metrics.

Table 3.1. Differences between selected semantics at weights of +3 and -3.

Semantic:	Semantic 001	Semantic 006
Mean Difference	Δ Length 2.08 ± 0.96	Δ Density 0.6 ± 1.6
Two-Tailed *P*-Value	0.048*	0.71
Cohen's D Effect Score	1.10	0.0289

Semantic Traversals on Identified Semantics

To further characterize the potential association of SeFa features with length and alignment, we traversed across semantic weights on all layers and all subsets of layers. Generally, in StyleGAN2-ADA, the style blocks of early levels correspond to coarse spatial resolution and control high-level features. Middle layers correspond to smaller scale features, and end layers correspond to higher resolution control such as pixel texture and microstructures.

In this network, Semantic 001 was most correlated with collagen fiber length. Results indicated that layer subset 3-4 best captured this feature with an R^2 value of 0.77 for a linear trendline across weight values (Figure 3.6). However, assessing the partial distance correlation scores (Table 3.2) indicates that layers 5-7 may have higher association with fiber length when controlling for the other variables. A higher score indicates a stronger relationship where one would mean a perfect deterministic dependence. These results indicate that collagen fiber length is captured by this network as a small-scale feature. Linear behavior is expected based on the factorization within SeFa. Certain morphological features are intrinsically coupled, and manipulating one inevitably affects the other. In this case, shrinking fiber length until fibers no

longer register also reduces collagen density (Figure 3.7). These features co-vary due to shared spatial or structural constraints.

Table 3.2. Partial distance correlation (PDC) of length assessed versus semantic weight, while controlling for fiber width, curvature, and density.

Semantic 001 - Length	PDC Score
Layers 0-2	0.203 ± 0.035
Layers 3-4	0.386 ± 0.031
Layers 5-7	0.586 ± 0.031

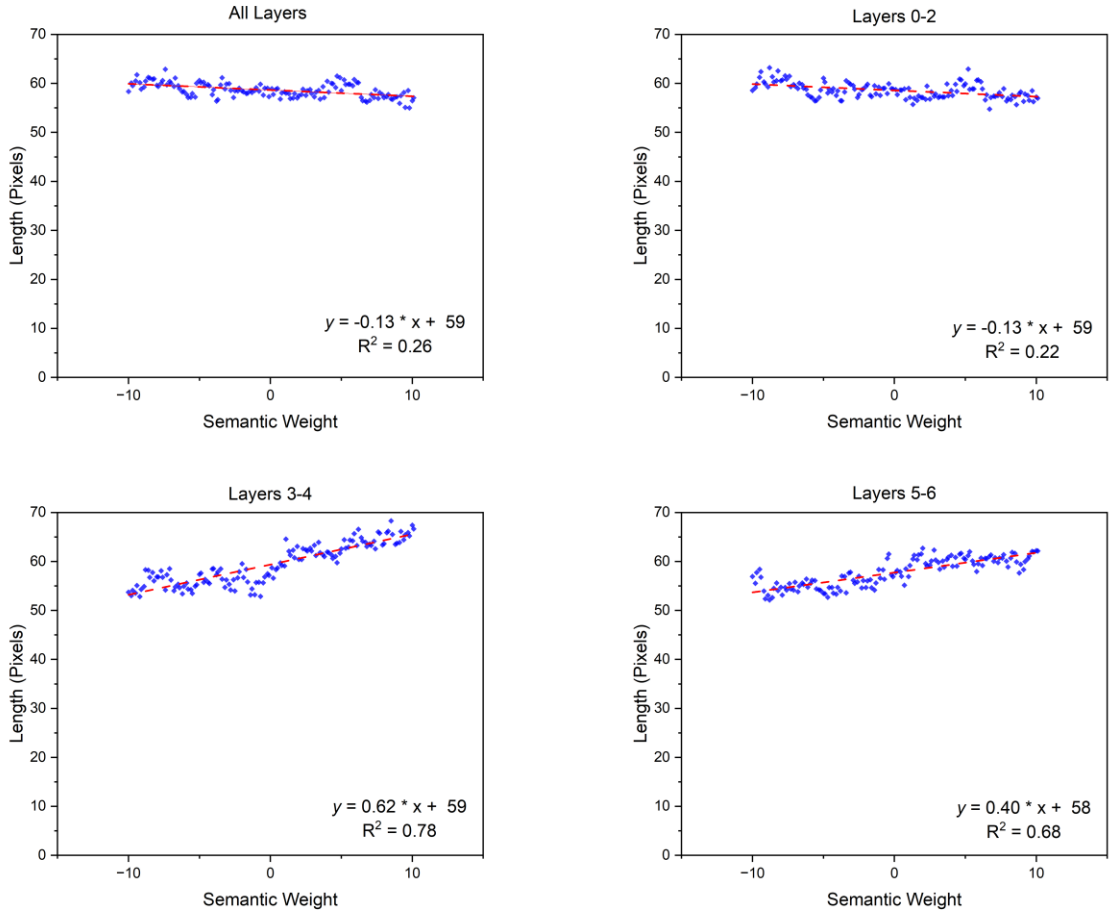


Figure 3.6. Results of CT-Fire analysis on collagen fiber length for Semantic 001 traversal applied to all layers, layers 0-2, layers 3-4, and layers 5-7. Linear trendlines and R^2 values are reported for each subset.

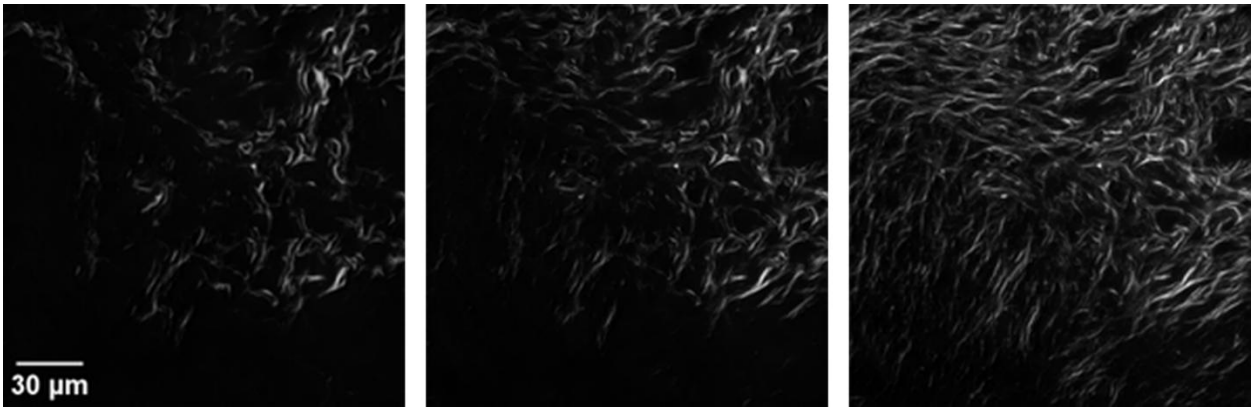


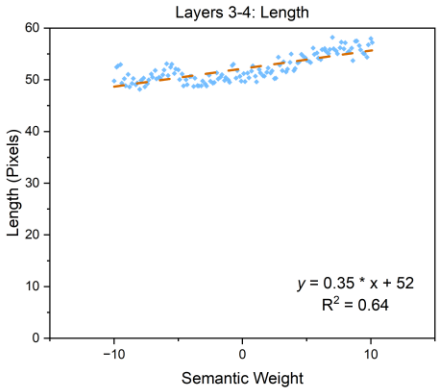
Figure 3.7. Semantic 001 traversal results at layers 3-4 with weights -6, 0, and +6 from left to right. As length is increased and decreased, the overall structure of the image appears to stay the same. Individual fibers can also be seen shrinking and elongating.

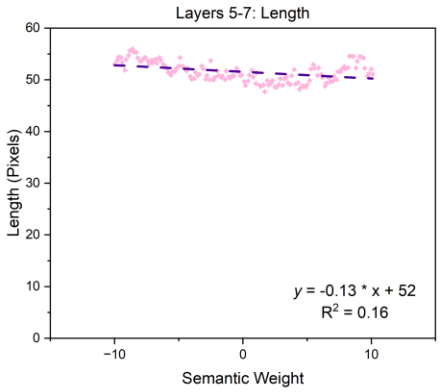
Due to the trend observed in Semantic 001, we suspected that the initial SeFa analysis with all layers was insufficient for observing potential feature-semantic correlations. Therefore, we used our visualization module to explore all 10 semantics on all subsets of layers to observe possible trends. Visual inspection revealed that Semantic 006 showed a promising correlation with fiber density (defined, in this case, as the mean distance between nearest fibers). Layers 0-2 showed few differences. Layers 3-4 and 5-7 were then analyzed to pick the most representative and least entangled feature representation. Both layer subsets yielded similar slope magnitude (though one is inverted) and R^2 values for density (based on the mean distance to the nearest fibers). However, the response of the output density in Layers 3-4 appears to have entangled behavior with both fiber width and length. These measures yielded high R^2 values, confounding our expectation that a semantic would not be correlated with multiple features (Figure 3.8). Thus, Layers 5-7 are the best representation of fiber density for Semantic 006 (Table 3.3). The inverse slopes of these two-layer subsets could hypothetically be the cause of this semantic not producing significant results in the initial SeFa analysis of all layers combined (Table 3.1). If the behavior is of equal and opposite magnitude, the overall effect of the semantic may have gone

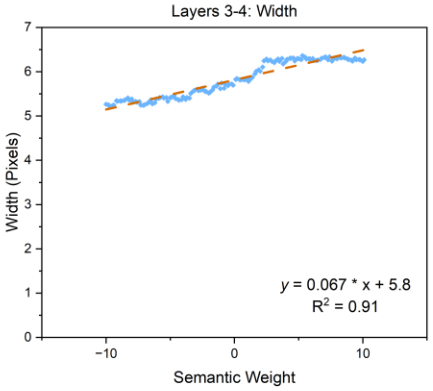
undetected. Visual inspection of sample behavior appears to agree with this quantitative analysis, as Layers 3-4 appear to add more random texture than Layers 5-7. The images from Layer 5-7 appear to better maintain the original structure and arrangement (Figure 3.9).

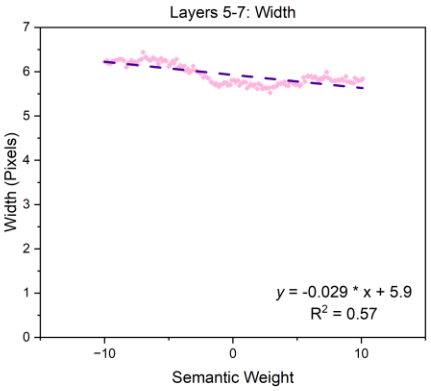
Table 3.3. Partial distance correlation (PDC) of density assessed versus semantic weight, while controlling for fiber width, curvature, and length.

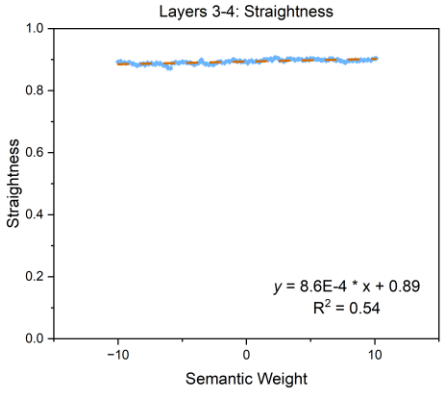
Semantic 006 - Density	Partial Distance Correlation
Layers 0-2	0.303 ± 0.0430
Layers 3-4	0.384 ± 0.0443
Layers 5-7	0.657 ± 0.0405

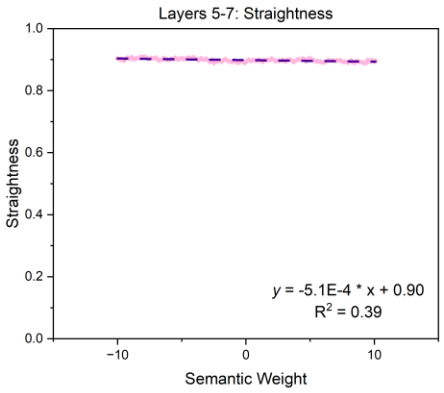

Layers 3-4: Length
Length (Pixels)
Semantic Weight
y = 0.35 * x + 52
R2 = 0.64

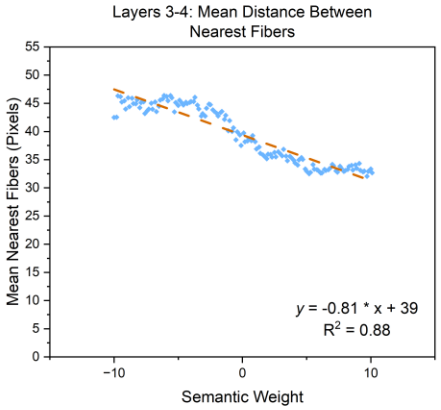

Layers 5-7: Length
Length (Pixels)
Semantic Weight
y = -0.13 * x + 52
R2 = 0.16


Layers 3-4: Width
Width (Pixels)
Semantic Weight
y = 0.067 * x + 5.8
R2 = 0.91


Layers 5-7: Width
Width (Pixels)
Semantic Weight
y = -0.029 * x + 5.9
R2 = 0.57


Layers 3-4: Straightness
Straightness
Semantic Weight
y = 8.6E-4 * x + 0.89
R2 = 0.54


Layers 5-7: Straightness
Straightness
Semantic Weight
y = -5.1E-4 * x + 0.90
R2 = 0.39


Layers 3-4: Mean Distance Between Nearest Fibers
Mean Nearest Fibers (Pixels)
Semantic Weight
y = -0.81 * x + 39
R2 = 0.88

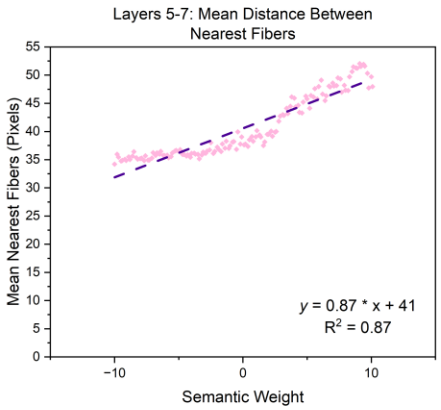

Layers 5-7: Mean Distance Between Nearest Fibers
Mean Nearest Fibers (Pixels)
Semantic Weight
y = 0.87 * x + 41
R2 = 0.87

Figure 3.8. Results of CT-Fire analysis on collagen fiber length, width, straightness, and mean distance to nearest fibers (MDNF) to represent density. Linear trendlines and R^2 values are reported for each feature for both layer subsets.

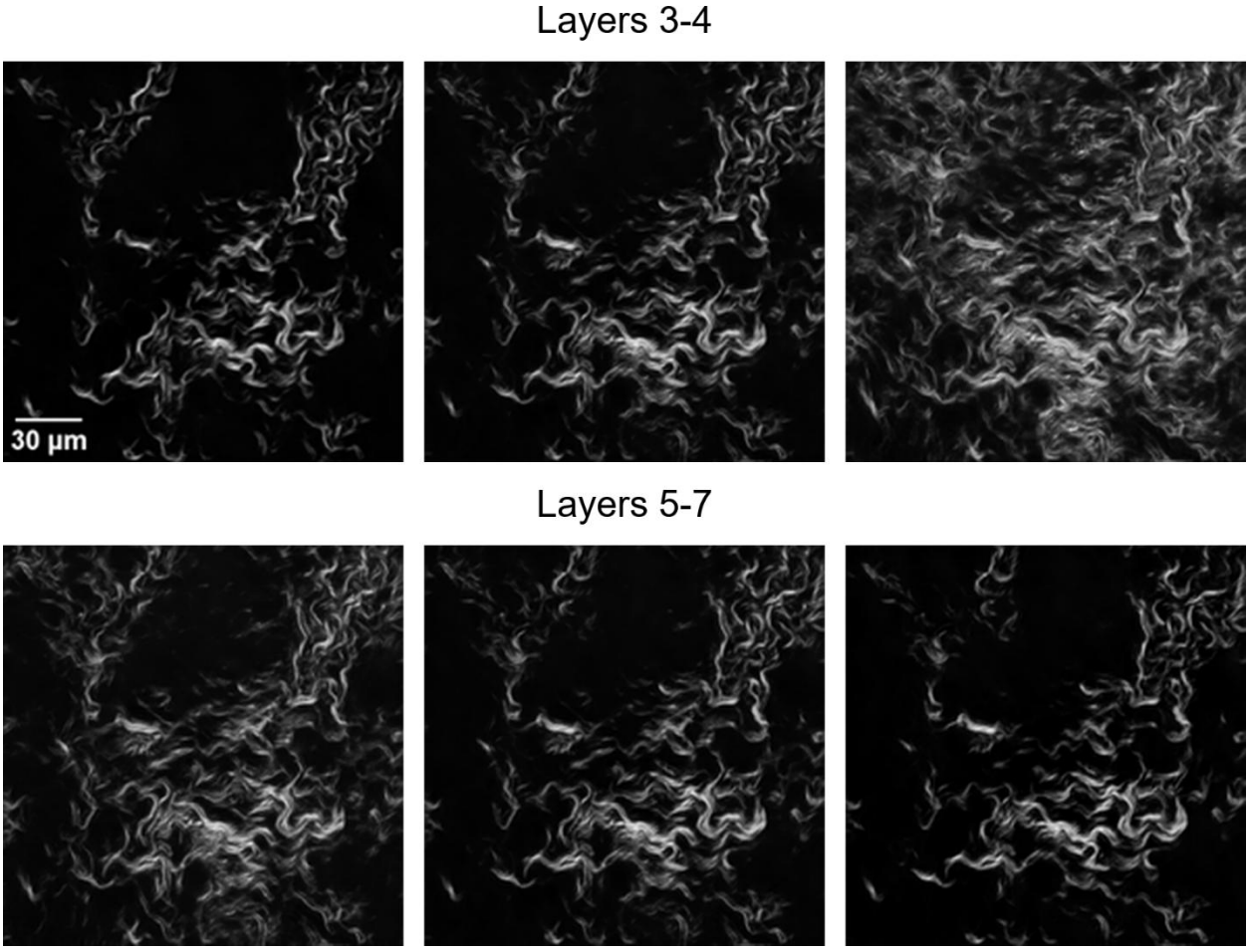


Figure 3.9. Semantic 006 traversal results at weights -6, 0, and +6 from left to right. Images from layers 3-4 are within the top row and appear to add more overall texture than Layers 5-7 (bottom row).

Encoder Training and Projection

Our encoder was trained on GAN images and visually inspected for projection quality with both generated images and real SHG images (Figure 3.10). The encoder achieved an FID score of 68.1 ± 1.2 and a KID score of 0.00190 ± 0.00060. Stability of training was assessed by inspecting the KID curve. Noise in this measurement is due to hardware memory limitations on

batch size, where smaller batches typically lead to greater variance in training (Figure 3.11). Visual inspection shows that projection does not completely capture the imported image, especially with real SHG images. This is expected, as encoders are typically liable to the loss of fine-grained detail[41]. In the projection, curvature is clearly altered, and not all fibers are modeled. Additionally, directionality, and therefore, overall alignment, is not fully preserved.

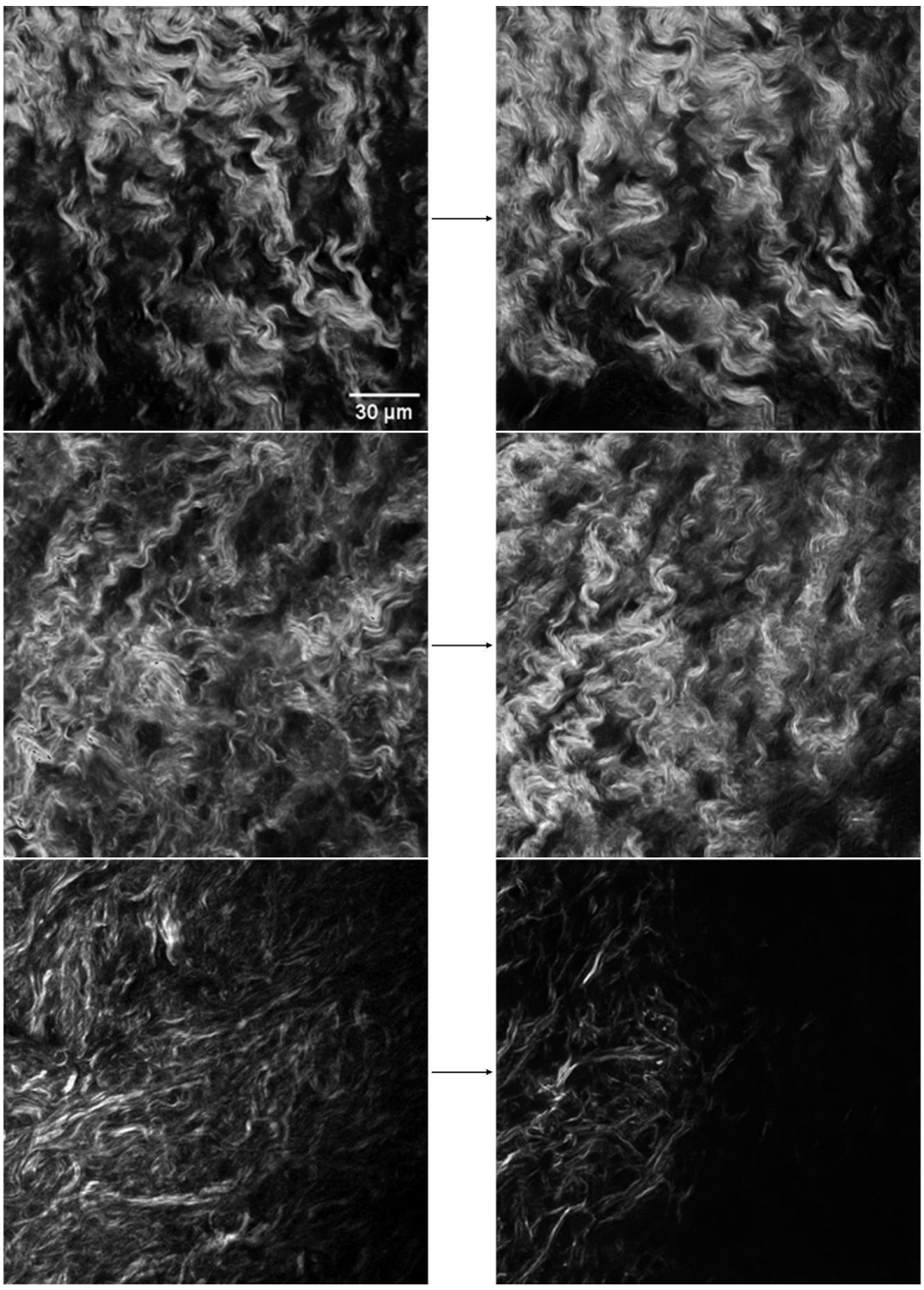


Figure 10. Images were mapped to the generator's latent space using the encoder. To test the encoder's fidelity both GAN images (top left) and real SHG images (middle and bottom left) were encoded.

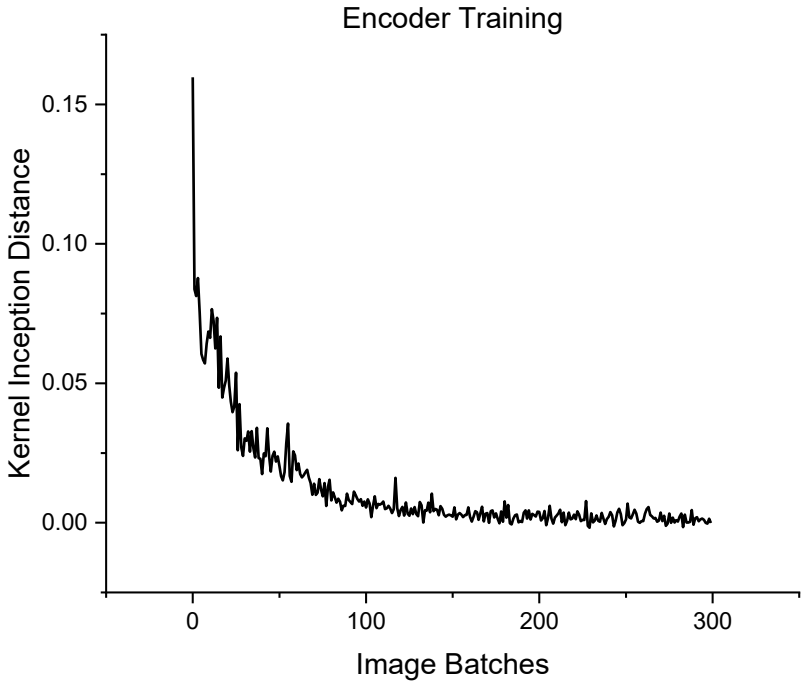


Figure 11. Encoder training was validated with Kernel Inception Distance score.

Results for the learned perceptual image patch similarity-based projection can be found in supplemental section.

3.5 Discussion

We have previously applied machine learning techniques to classify alterations in normal and diseased collagen fiber architectures in the ECM[16,42-44]. This work is relevant since collagen fiber structures are altered in a wide range of pathologies, including cancers and fibroses. The transition from normal ECM to pathological is often associated with the early stages of cancer. Changes in the ECM structure are apparent in disease progress as in advanced tumor masses, the ECM can comprise up to 60% of the tissue, usually dominated by the accumulation of newly synthesized collagen[45]. While successful, our custom designed classifiers are still “black-box”. It

is important to capture the specific features that are associated with diseases for more directed use. The ability to edit in latent space to determine the relative contributions of fiber characteristics, here semantics, adds further richness to the approach.

There are a few approaches to latent space analysis that may improve the number of meaningful semantics available. Alterations to images within latent space may produce image variations that are not visually detectable or captured within the current set of analysis tools. SeFa in particular typically corresponds to linearly separable variations that can affect the image independently. This separation ability is limited when there are higher-order interactions between latent components, such as cases in which a feature is a mixture of multiple latent factors. It also may miss more complex structures in the latent space, such as nonlinear trajectories or features that emerge only after certain activation thresholds are crossed[46,47]. Alternative strategies could address these limitations. For example, GANSpace applies a principal component analysis (PCA) to intermediate feature activations, which hypothetically handle latent combinations better[48]. StyleFlow uses normalizing flows, which are explicitly nonlinear for modeling complex semantic trajectories[49] and could be another useful strategy for finding features that are not currently within the analysis pipeline.

The semantic traversal tool was built to capture the behavior of the eigenvectors given randomly generated images. In general, we expect to correlate semantics to collagen fiber features in the CT-Fire and FIJI analysis including length, width, curvature, density, and alignment. As a neural network trains on a dataset, it is expected to capture hierarchical features of images including edges, shapes, and textures as a natural result of successive convolutional and downsampling layers. However, features are also expected to be connected to the layers of Style blocks within the network. In this study, when semantics with changes were detected in the initial analysis,

they were later found to have more linear behavior in specific subsets of layers. Specifically, we found two semantics correlated with collagen fiber length and density respectively.

The baseline data, generator, and encoder all had relatively high FID scores and relatively low SSIM scores. There are a few potential confounding factors that have influenced these results. Features found on pre-trained FID and KID models are from labeled RGB databases and may have comparatively limited value for realistic microstructure present in microscopy with grey level contrast[50]. FID scores can also artificially change based on how many images are used for assessment[51]. SSIM scores can be lower when there are significant differences in signal to noise ratio across the dataset[52]. In this case, the use of a large variety of different diseases and tissues resulted in a high degree of heterogeneity in the sample library. Full results for all calculated metrics can be found in the supplementary material. Even though FID and SSIM scores indicate room for improvement, examining the difference between baseline and generator scores indicate a relatively close fit. KID scores appeared to offer more stable and meaningful comparisons across methods, potentially due to its less biased approach and suitability for smaller sample sizes[32].

There are several benefits in using a library of SHG images from multiple tissue and disease types. This allows the algorithm to learn from a relatively full scope of collagen morphologies. Increasing the number of images in training also significantly increases the quality of the resulting generator network. Often, in microscopy imaging, the largest issue for achieving high quality machine learning is the small number of images available for training. By combining images of different pathologies, we have significantly decreased the probability of low-data biases in training. In general, converting neural networks to work with greyscale images reduces the dimensionality and therefore the overall complexity of a given network. It also slightly

reduces the number of learnable parameters and computational load compared to a network with RGB input. Hypothetically this also simplifies feature extraction across the model. However, a major trade-off is the inability to effectively use common pre-trained datasets such as ImageNet for transfer learning. Additionally, FID and KID scores are based on a model trained on RGB image features. By triplicating the greyscale input, we can utilize these metrics, but there may be unpredictability in evaluating scores and a less stable correlation to visual inspection.

Explainable machine learning can lead to further study of biological features in disease states. An ideal model is transparent, interpretable, and explainable[53]. Given the rapid pace of innovation in machine learning, it is important to consider alternative generative models to StyleGAN2-ADA that can offer similar capabilities. StyleGAN3 was created in part to address phase artifacts that can occur due to the discriminator's progressive architecture. However, there are substantial hurdles that must be overcome to convert this method to use a greyscale-only network[54].

Variational autoencoders (VAE) aim to model the underlying distribution of data through a compressed latent space. This method uses a Gaussian estimation to approximate the posterior distribution of the data. Although this creates a relatively smooth, continuous, and interpretable latent space, the Gaussian assumption can result in blurry image generation[55,56]. There are methods that address these limitations, but they do not offer the same degree of ease in editing the latent codes. For example, the VQ-VAE model provides high-quality image generation while using discrete hierarchical latent codes rather than continuous ones[57]. NVAE uses convolutions within the generator to improve results and separates the latent space codes with deep hierarchical architecture[58]. However, both approaches require a separate module to render images in response to latent codes being accessed and edited.

Latent Diffusion Models (LDMs) extend upon the VAE framework by building a hierarchy of denoising autoencoders that operate in the lower-dimensional latent space rather than the standard pixel space. They have been particularly adept at modeling complicated multi-modal distributions [59]. The approaches to editing latent space features in latent diffusion models have generally focused on text-to-image generation. However, recent approaches have focused on linear and versatile image-editing that more closely resembles the relatively straightforward process of editing the latent space within GANs[60,61]. Despite advances in data efficiency, diffusion models still generally require at least 5,000 images to achieve excellent results[62]. In contrast, StyleGAN2-ADA achieved an excellent representation with only 1,319 images in our model.

3.6 Conclusions

In this study, we implemented a high-performing GAN for creating SHG images of collagen fibers, alongside GUI tools which allow for fine control of image manipulation that aid in the analysis of the learned semantics within latent space. Using external analysis tools, CT-Fire and FIJI, we directly modeled the GAN response within latent space. By testing each subset of layers using SeFa-based semantic sliders, we were able to directly correlate features to their respective network style block within StyleGAN2-ADA-SHG. This quantitative approach can greatly aid the usage of GANs and potentially allow for new ways to analyze morphological changes in collagen fibers in disease states. Ultimately, although only two meaningful semantics were discovered herein, this work indicates that this is a viable strategy for manipulating microscopy images. With further refinement, these tools can be used to create biomimetic models with variable morphology based on disease types.

3.7 Supplemental Material

Code Availability

Code and pre-trained models used in this study are available at: https://github.com/Campagnola-BME-Lab/sefa_collagen

This codebase uses files that were modified from the following repositories:

https://github.com/genforce/sefa

https://github.com/NVlabs/stylegan2-ada-pytorch

https://github.com/bryandlee/stylegan2-encoder-pytorch

https://github.com/zhenxingjian/Partial_Distance_Correlation

The encoder figure was generated with this repository:

https://github.com/HarisIqbal88/PlotNeuralNet

All code is additionally available by request.

Expanded Results

In the initial SeFa analysis, semantic 008 appeared to have a significant difference in alignment. Semantic 008 was measured as having an effect size of 1.08, indicating a large influence on the output (Table S.1).

Table S.1. All relevant results from initial semantic assessment from all layers.

Semantic:	Semantic 001	Semantic 008	Semantic 006
Mean Difference	Δ Length 2.08 ± 0.96	Δ Alignment 1737 ± 811	Δ Density 0.6 ± 1.6
Two-Tailed *P*-Value	0.048*	0.050	0.71

Cohen's D Effect Score	1.10	1.08	0.0289

Collagen fiber alignment was most correlated with Semantic 008. Results seemingly indicated that layer subset 5-7 best captured this feature with an R^2 value of 0.84 for a linear trendline across weight values (Figure S.1). However, visual inspection revealed that this semantic is modifying overall collagen coverage within the image (Figure S.2). It is coverage, and not alignment, as the fibers do not have altering patterns (e.g. becoming more parallel) as the semantic weight is changed. Inspecting density, the slope and R^2 values are both significantly lower than in Semantic 006. Similarly, the length metric has lower values for slope and R^2 values than Semantic 001. We do not currently have an analysis tool for fiber coverage as this varies based on tissue samples and location. However, as the coverage varies significantly within the training data, it follows that this would be an identified feature within SeFa.

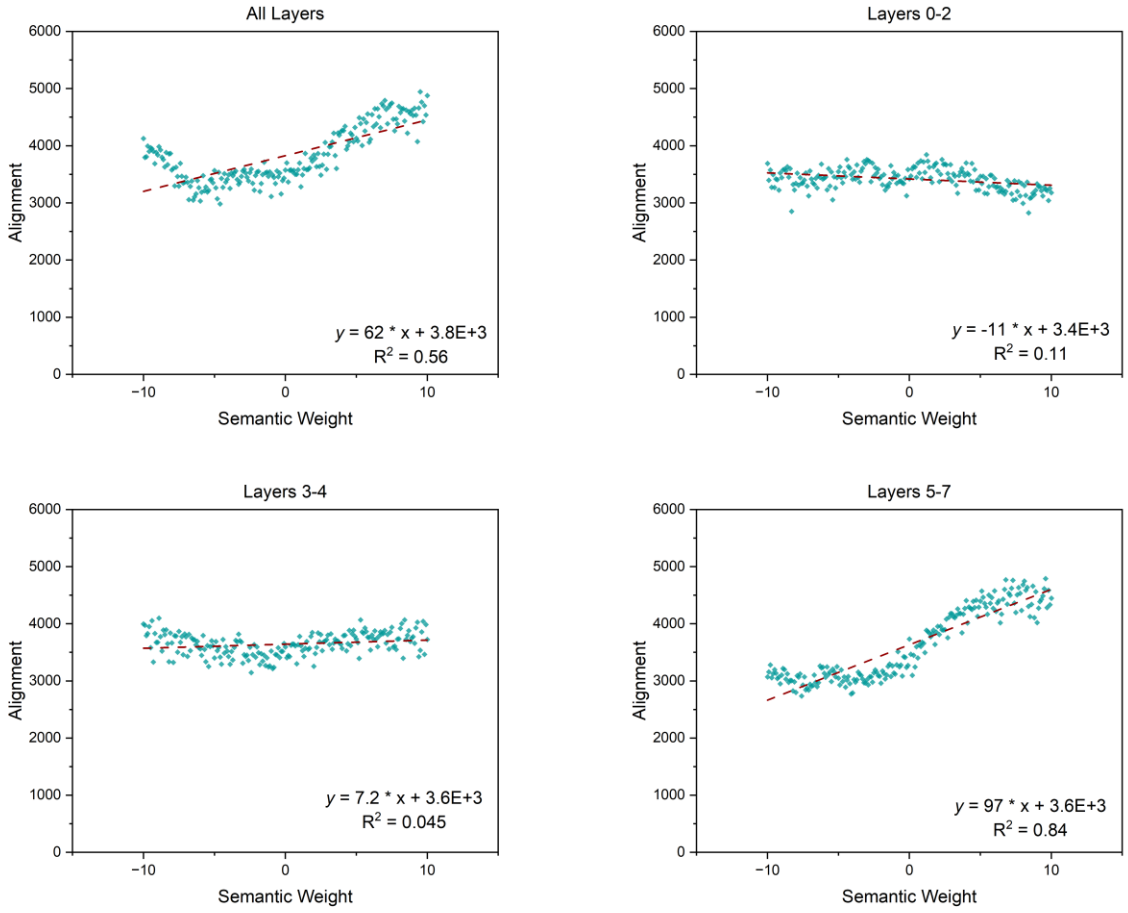


Figure S.1. Results of FFT analysis on collagen fiber alignment for Semantic 008 traversal applied to all layers, layers 0-2, layers 3-4, and layers 5-7. Linear trendlines and R^2 values are reported for each subset.

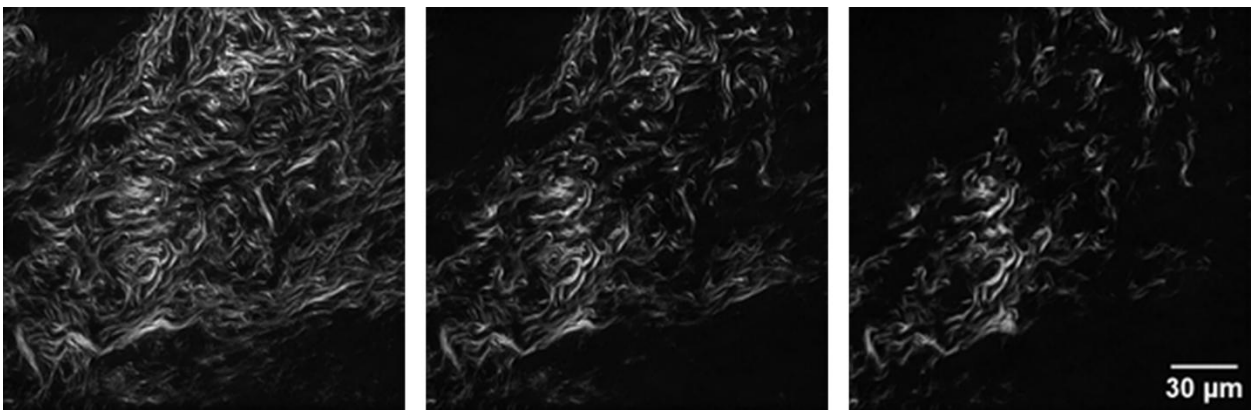


Figure S.2. Semantic 008 traversal results at layers 5-7 with weights -6, 0, and +6 from left to right. As weight is decreased fiber coverage increases within the 512 x 512 images.

Full Evaluation Metrics

Metrics were calculated with ten sets of 100 vs. 100 images. Standard SSIM was compared with MS-SSIM and MicroSSIM. FID and KID scores were calculated using torch-fidelity. StyleGAN uses internal FID scoring as part of its training. The encoder used KID scores as part of its training. Given the hypothetical limit of the generator's FID score should be the baseline score, it follows that the difference, 18.1, shows the relative level of potential future improvement. Similarly, the difference between MicroSSIM scores between the baseline and the generator, 0.015, appears to be a better marker for fidelity of the synthetic images. The encoder was evaluated against generated images only, which is most likely the reason for relatively better scoring.

Table S.2. Evaluation metrics for the baseline database of SHG images, the StyleGAN generator, and the adapted encoder.

	SSIM	MS-SSIM	MicroSSIM
Baseline	0.103 ± 0.0042	0.158 ± 0.0042	0.222 ± 0.0085
Generator	0.112 ± 0.0038	0.178 ± 0.0043	0.237 ± 0.013
Encoder	0.296 ± 0.0078	0.296 ± 0.0070	0.327 ± 0.0064

	FID	KID
Baseline	74.1 ± 1.44	7.41E-05 ± 6.54E-04
Generator	92.2 ± 1.53	0.00583 ± 0.000895
Encoder	68.1 ± 1.21	0.00190 ± 0.000600

We computed frequency-domain metrics from grayscale microscopy images by applying a 2D fast Fourier transform (FFT) to each image. Prior to transformation, all images were normalized to a [0,1] intensity range to eliminate variation due to global brightness.

The magnitude of the FFT output was radially averaged to produce a 1D frequency profile. To characterize the rate of frequency decay the data was fit to an exponential decay model of the form:

$$y = a * e^{-\frac{r}{\tau}} + c$$

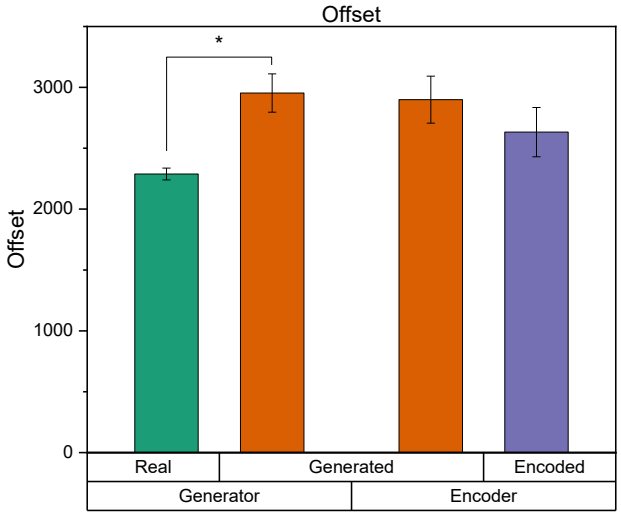
where **r** is the radial frequency, **a** is a scaling coefficient, **c** is the offset term correlated to noise level in the spectrum, and **τ** (tau) is the time constant. This parameter reflects how rapidly high-frequency content decays, with lower values indicating more rapid falloff (i.e., smoother or blurrier textures), and higher values indicating high-frequency structure.

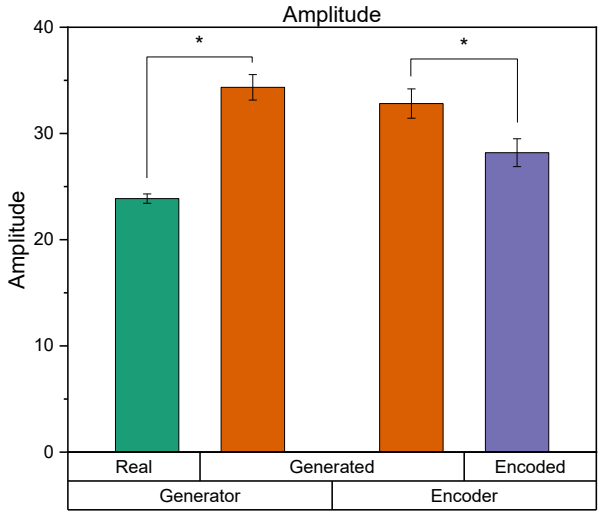
The significantly lower τ value for generated images compared to real images indicates that the generator fails to fully reproduce the high-frequency content present in natural images (Table S.3). The generator likely has a tendency toward smoother or blurrier textures. However, the non-zero τ value also implies that the generator retains some degree of high-frequency structure, indicating that it does capture some of the fiber detail. The encoder has a significantly higher τ value, which could be due to a variety of reasons (Figure S.3). This could reflect mild filtering from the process of inverse mapping, that may result in image sharpening. Artifacts and shifts in the latent space may cause the generator to interpret higher frequency structure.

Table S.3. Results of FFT data radially averaged and fit to an exponential decay model.

	Amplitude	Time Constant (τ)	Offset

Real SHG Images	23.9 ± 0.44	2.79 ± 0.17	2289 ± 48
Generated	34.3 ± 1.2	1.08 ± 0.076	2953 ± 157
Generated for Encoding	32.8 ± 1.4	1.03 ± 0.058	2899 ± 193
Encoder	28.2 ± 1.3	1.22 ± 0.061	2633 ± 202





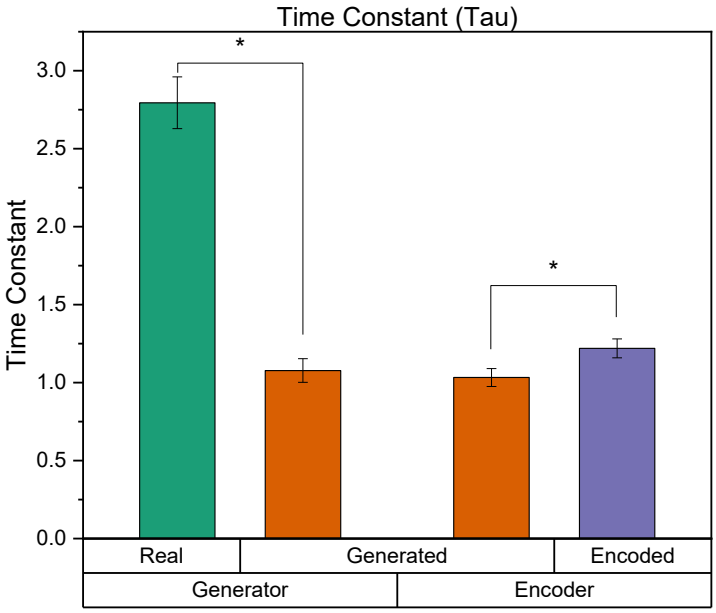


Figure S.3. Radially averaged FFT data fit to an exponential decay model. Encoder images were paired to specific generated images. The generator was compared with an unpaired t-test.

LPIPS Projector Results

For the optimization-based method, we implemented a projector that minimizes a weighted combination of pixel-wise mean squared error (MSE) and learned perceptual image patch similarity (LPIPS) loss between the input image and the synthesized output. The LPIPS loss was computed using a VGG-based neural network. This method did not produce adequate images, as the structure of collagen fibers is not visible (Figure S.4). Some directionality of the fibers is captured, but ultimately more robust methods (as in the implemented encoder) are required to map collagen images into the generator's latent space.

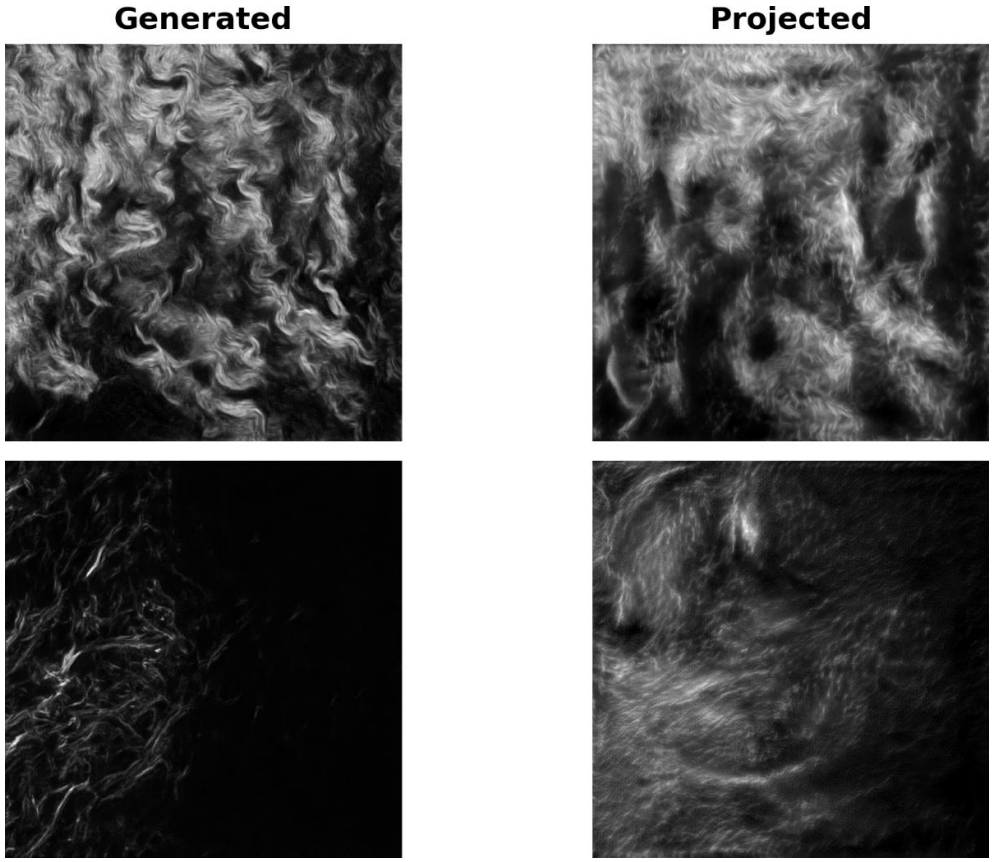


Figure S.4. LPIPs-based projection for encoding generated images. Although some directionality is captured, the structure of collagen is not visually identifiable.

Chapter 3 References

1. Provenzano, P. P., Eliceiri, K. W., Campbell, J. M., Inman, D. R., White, J. G. & Keely, P. J. Collagen reorganization at the tumor-stromal interface facilitates local invasion. *BMC Medicine* **4** (2006). https://doi.org:10.1186/1741-7015-4-38
2. Campagnola, P. Second harmonic generation imaging microscopy: applications to diseases diagnostics. *Analytical Chemistry* **83**, 3224-3231 (2011). https://doi.org:10.1021/ac1032325
3. James, D. S. & Campagnola, P. J. Recent Advancements in Optical Harmonic Generation Microscopy: Applications and Perspectives. *BME Front* **2021**, 3973857 (2021). https://doi.org:10.34133/2021/3973857
4. Wen, B. *et al.* 3D texture analysis for classification of second harmonic generation images of human ovarian cancer. *Scientific Reports* **6** (2016). https://doi.org:10.1038/srep35734
5. Alkmin, S. *et al.* Role of Collagen Fiber Morphology on Ovarian Cancer Cell Migration Using Image-Based Models of the Extracellular Matrix. *Cancers* **12**, 1390-1390 (2020). https://doi.org:10.3390/cancers12061390
6. Goodfellow, I. *et al.* Generative adversarial networks. *Communications of the ACM* **63**, 139-144 (2020). https://doi.org:10.1145/3422622
7. Sreejith Kumar, A. J. *et al.* Evaluation of Generative Adversarial Networks for High-Resolution Synthetic Image Generation of Circumpapillary Optical Coherence Tomography Images for Glaucoma. *JAMA Ophthalmology* **140**, 974-981 (2022). https://doi.org:10.1001/JAMAOPHTHALMOL.2022.3375
8. Li, X. *et al.* Unsupervised content-preserving transformation for optical microscopy. *Light: Science & Applications 2021 10:1* **10**, 1-11 (2021). https://doi.org:10.1038/s41377-021-00484-y
9. Tian, L. *et al.* Deep learning in biomedical optics. *Lasers in surgery and medicine* **53**, 748-748 (2021). https://doi.org:10.1002/LSM.23414
10. Chen, Y. *et al.* Generative Adversarial Networks in Medical Image augmentation: A review. *Computers in Biology and Medicine* **144**, 105382-105382 (2022). https://doi.org:10.1016/J.COMPBIOMED.2022.105382
11. Reich, C., Prangemeier, T., Wildner, C., Koeppl, H. (2021). Multi-StyleGAN: Towards Image-Based Simulation of Time-Lapse Live-Cell Microscopy. In: de Bruijne, M., *et al.* Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. MICCAI 2021. Lecture Notes in Computer Science(), vol 12908. Springer, Cham. https://doi.org/10.1007/978-3-030-87237-3_46
12. Melnik, A. *et al.* Face Generation and Editing With StyleGAN: A Survey. *IEEE Transactions on Pattern Analysis \& Machine Intelligence* **46**, 3557-3576 (2024). https://doi.org:10.1109/TPAMI.2024.3350004
13. Shon, K., Sung, K. R., Kwak, J., Shin, J. W. & Lee, J. Y. Development of a β-Variational Autoencoder for Disentangled Latent Space Representation of Anterior Segment Optical

Coherence Tomography Images. *Transl Vis Sci Technol* **11**, 11 (2022). https://doi.org:10.1167/tvst.11.2.11
14. Zhen, X. a. M., Zihang and Chakraborty, Rudrasis and Singh, Vikas. On the Versatile Uses of Partial Distance Correlation in Deep Learning. *Proceedings of the European conference on computer vision (ECCV)* (2022).
15. Karras, T., Laine, S. & Aila, T. A Style-Based Generator Architecture for Generative Adversarial Networks. *IEEE transactions on pattern analysis and machine intelligence* **43**, 4217-4228 (2021). https://doi.org:10.1109/TPAMI.2020.2970919
16. Rentchler, E. C., Gant, K. L., Drapkin, R., Patankar, M. & Campagnola, P. J. Imaging Collagen Alterations in STICs and High Grade Ovarian Cancers in the Fallopian Tubes by Second Harmonic Generation Microscopy. *Cancers* **11** (2019). https://doi.org:ARTN 180510.3390/cancers11111805
17. Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J. & Aila, T. Training Generative Adversarial Networks with Limited Data. *Advances in Neural Information Processing Systems* **2020-December** (2020).
18. Shen, Y. & Zhou, B. in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* 1532-1540.
19. Khan, A., Lee, C.-H., Huang, P. Y. & Clark, B. K. Leveraging generative adversarial networks to create realistic scanning transmission electron microscopy images. *npj Comput Mater* **9**, 1-9 (2023). https://doi.org:10.1038/s41524-023-01042-3
20. Chen, X., Nadiarynkh, O., Plotnikov, S. & Campagnola, P. J. Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure. *Nat Protoc* **7**, 654-669 (2012). https://doi.org:10.1038/nprot.2012.009
21. Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J. & Aila, T. Analyzing and Improving the Image Quality of StyleGAN. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 8107-8116 (2019).
22. Karras, T., Aila, T., Laine, S. & Lehtinen, J. Progressive Growing of GANs for Improved Quality, Stability, and Variation. *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings* (2017).
23. Bredfeldt, J. S. *et al.* Computational segmentation of collagen fibers from second-harmonic generation images of breast cancer. *Journal of biomedical optics* **19** (2014). https://doi.org:10.1117/1.JBO.19.1.016007
24. Stein, A. M., Vader, D. A., Jawerth, L. M., Weitz, D. A. & Sander, L. M. An algorithm for extracting the network geometry of three-dimensional collagen gels. *Journal of Microscopy* **232**, 463-475 (2008). https://doi.org:10.1111/j.1365-2818.2008.02141.x
25. Rentchler, E. C., Gant, K. L., Drapkin, R., Patankar, M. & J Campagnola, P. Imaging Collagen Alterations in STICs and High Grade Ovarian Cancers in the Fallopian Tubes by Second Harmonic Generation Microscopy. *Cancers* **11**, 1805 (2019). https://doi.org:10.3390/cancers11111805
26. Zhang, R., Isola, P., Efros, A. A., Shechtman, E. & Wang, O. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. Preprint at https://doi.org/10.48550/arXiv.1801.03924 (2018).
27. Simonyan, K. & Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. (2014). https://doi.org:10.48550/arxiv.1409.1556

28. Zhu, J., Shen, Y., Zhao, D. & Zhou, B. In-Domain GAN Inversion for Real Image Editing. Preprint at https://doi.org/10.48550/arXiv.2004.00049 (2020).
29. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. Preprint at https://doi.org/10.48550/arXiv.1412.6980 (2017).
30. Bischoff, S. *et al.* A Practical Guide to Sample-based Statistical Distances for Evaluating Generative Models in Science. Preprint at https://doi.org/10.48550/arXiv.2403.12636 (2024).
31. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 2818-2826.
32. Bińkowski, M., Sutherland, D. J., Arbel, M. & Gretton, A. Demystifying MMD GANs. *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings* (2018).
33. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing* **13**, 600-612 (2004). https://doi.org:10.1109/TIP.2003.819861
34. 34 Wang, Z., Simoncelli, E. P. & Bovik, A. C. Vol. 2 1398-1402 Vol.1392 (2003).
35. Ashesh, A., Joran Deschamps, and Florian Jug. MicroSSIM: Improved Structural Similarity for Comparing Microscopy Data. *ECCV 2024 Workshops* (2024). https://doi.org:https://doi.org/10.48550/arXiv.2408.08747
36. Tian, X., Anantrasirichai, N., Nicholson, L. & Achim, A. The quest for early detection of retinal disease: 3D CycleGAN-based translation of optical coherence tomography into confocal microscopy. *Biological Imaging* **4** (2024). https://doi.org:10.1017/s2633903x24000163
37. Skandarani, Y., Jodoin, P.-M. & Lalande, A. GANs for Medical Image Synthesis: An Empirical Study. *J Imaging* **9**, 69 (2023). https://doi.org:10.3390/jimaging9030069
38. Lee, J., Kim, J.-H. & Lee, J.-S. Demystifying Randomly Initialized Networks for Evaluating Generative Models. arXiv:2208.09218 (2022).
39. Yang, M., Yang, C., Zhang, Y., Bai, Q., Shen, Y. & Dai, B. Revisiting the Evaluation of Image Synthesis with GANs. arXiv:2304.01999 (2023).
40. Jiang, L., Huang, S., Luo, C., Zhang, J., Chen, W. & Liu, Z. An improved multi-scale gradient generative adversarial network for enhancing classification of colorectal cancer histological images. *Front Oncol* **13**, 1240645 (2023). https://doi.org:10.3389/fonc.2023.1240645
41. Tov, O., Alaluf, Y., Nitzan, Y., Patashnik, O. & Cohen-Or, D. Designing an Encoder for StyleGAN Image Manipulation. arXiv:2102.02766 (2021).
42. Tilbury, K., Hocker, J., Wen, B. L., Sandbo, N., Singh, V. & Campagnola, P. J. Second harmonic generation microscopy analysis of extracellular matrix changes in human idiopathic pulmonary fibrosis. *J Biomed Opt* **19**, 086014 (2014). https://doi.org:10.1117/1.JBO.19.8.0860141899297 [pii]
43. Wen, B. L. *et al.* Texture analysis applied to second harmonic generation image data for ovarian cancer classification. *J Biomed Opt* **19**, 096007 (2014). https://doi.org:10.1117/1.JBO.19.9.0960071906043 [pii]
44. Wen, B. *et al.* 3D texture analysis for classification of second harmonic generation images of human ovarian cancer. *Sci Rep-Uk* **6** (2016). https://doi.org:ARTN 3573410.1038/srep35734

45. Hu, M., Ling, Z. & Ren, X. Extracellular matrix dynamics: tracking in biological systems and their implications. *Journal of Biological Engineering 2022 16:1* **16**, 1-13 (2022). https://doi.org:10.1186/S13036-022-00292-X
46. Oldfield, J., Georgopoulos, M., Panagakis, Y., Nicolaou, M. A. & Patras, I. Tensor Component Analysis for Interpreting the Latent Space of GANs. Preprint at https://doi.org/10.48550/arXiv.2111.11736 (2021).
47. Zhang, Z. & Schomaker, L. Optimizing and interpreting the latent space of the conditional text-to-image GANs. *Neural Computing and Applications* **36**, 2549-2572 (2024). https://doi.org:10.1007/s00521-023-09185-6
48. Härkönen, E., Hertzmann, A., Lehtinen, J. & Paris, S. GANSpace: Discovering Interpretable GAN Controls. *Advances in Neural Information Processing Systems* **2020-December** (2020).
49. Abdal, R., Zhu, P., Mitra, N. & Wonka, P. StyleFlow: Attribute-conditioned Exploration of StyleGAN-Generated Images using Conditional Continuous Normalizing Flows. *ACM Transactions on Graphics* **40**, 1-21 (2021). https://doi.org:10.1145/3447648
50. Shah, D., Suresh, A., Admasu, A., Upadhyay, D. & Deb, K. A Survey on Evaluation Metrics for Synthetic Material Micro-Structure Images from Generative Models. (2022). https://doi.org:10.48550/arXiv.2211.09727
51. Chong, M. J. & Forsyth, D. Effectively Unbiased FID and Inception Score and where to find them. Preprint at https://doi.org/10.48550/arXiv.1911.07023 (2020).
52. Nilsson, J. & Akenine-Möller, T. Understanding SSIM. Preprint at https://doi.org/10.48550/arXiv.2006.13846 (2020).
53. Roscher, R., Bohn, B., Duarte, M. F. & Garcke, J. Explainable Machine Learning for Scientific Insights and Discoveries. *IEEE Access* **8**, 42200-42216 (2020). https://doi.org:10.1109/ACCESS.2020.2976199
54. Karras, T. *et al.* Alias-Free Generative Adversarial Networks. Preprint at https://doi.org/10.48550/arXiv.2106.12423 (2021).
55. Marino, J. Predictive Coding, Variational Autoencoders, and Biological Connections. *Neural Comput* **34**, 1-44 (2021). https://doi.org:10.1162/neco_a_01458
56. Connor, C. W. Understanding New Machine Learning Architectures: Practical Generative Artificial Intelligence for Anesthesiologists. *Anesthesiology* **140**, 599-609 (2024). https://doi.org:10.1097/ALN.0000000000004841
57. Razavi, A., Oord, A. van den & Vinyals, O. Generating Diverse High-Fidelity Images with VQ-VAE-2. Preprint at https://doi.org/10.48550/arXiv.1906.00446 (2019).
58. Vahdat, A. & Kautz, J. NVAE: A Deep Hierarchical Variational Autoencoder. Preprint at https://doi.org/10.48550/arXiv.2007.03898 (2021).
59. Rombach, R., Blattmann, A., Lorenz, D., Esser, P. & Ommer, B. High-Resolution Image Synthesis with Latent Diffusion Models. Preprint at https://doi.org/10.48550/arXiv.2112.10752 (2022).
60. Kwon, M., Jeong, J. & Uh, Y. Diffusion Models already have a Semantic Latent Space. Preprint at https://doi.org/10.48550/arXiv.2210.10960 (2023).
61. Pathania, K. Enhancing Conditional Image Generation with Explainable Latent Space Manipulation. Preprint at https://doi.org/10.48550/arXiv.2408.16232 (2024).
62. Wang, Z. *et al.* Patch Diffusion: Faster and More Data-Efficient Training of Diffusion Models. Preprint at https://doi.org/10.48550/arXiv.2304.12526 (2023).

Chapter 4: Conclusions and Future Approaches

Alterations in the extracellular matrix (ECM), particularly in collagen structure, are linked to diseases such as cancer, fibrosis, and connective tissue disorders. This thesis presents a large library of Second Harmonic Generation (SHG) microscopy images of fibrillar collagen across disease models. Morphological features such as fiber length, straightness, alignment, and density were quantified using CurveAlign and FIJI, enabling sample classification via support vector machines. These images also served as training data for a StyleGAN2-ADA model, which achieved a strong Kernel Inception Distance score. Semantic Factorization (SeFa) was used to identify latent directions in the generator associated with interpretable fiber features. Two semantics showed linear behavior and strong correlation with fiber length and density. Finally, projection methods were incorporated to enable image editing via the GAN's latent space. These tools lay the groundwork for generating synthetic SHG images with disease-relevant collagen morphologies.

4.1 Future Directions

Transfer learning is a powerful tool that could significantly improve both StyleGAN2-ADA, and more importantly, the current image encoder. Regardless of biological scale or modality, both microscopy and medical imaging are often single-channel (greyscale) images. For a generative model like StyleGAN2-ADA, as long as the pixel distribution is consistent and the resolution is standardized, the network operates agnostically to whether it's generating collagen fibers or lung nodules. This opens the possibility to train on significantly larger databases such as the NIH chest dataset of 112,120 X-ray images[1] or the Lung Image Database Consortium of 244,527 CT images[2]. By initializing StyleGAN2-ADA on these larger databases we can begin training at

network weights closer to our final task. Most importantly, by training the encoder on this more robust version of the GAN, further fine-tuning should result in significantly better reproductions of SHG images.

To extract more meaningful semantics, methods beyond SeFa must be implemented. The GANSpace approach with principal component analysis would be a promising addition[3]. A new implementation would have to be refactored from the current documented code. Unfortunately, GANSpace uses CUDA version 10.1 which means it is incompatible with StyleGAN2-ADA which is run on CUDA version 11.1. Thus, to fully integrate these two programs, a unified environment would need to be created. This would likely require careful reconciliation of PyTorch versions and plugin builds to maintain stable GPU execution across both frameworks. Ultimately, both programs suffer from version deprecation, especially as newer GPUs are released that require higher CUDA versions. Although arduous, it may be wise to refactor both models to upgrade to the newest version. This would also allow a newer version of the GUI interface Streamlit to be implemented. Streamlit's current version has significantly better development tools such as logical button behavior that requires no additional manual hard-coding. Its continued development has allowed successful integrated user control in deep learning models[4,5].

Further improvements to the work herein could draw from collaboration with projects focused on alternative approaches to modeling collagen fibers through deep learning. The Eliceiri lab created a multi-model approach for fiber centerline tracking. First, a variational autoencoder generates synthetic centerlines with controllable topology. Then, a conditional GAN turns these centerlines into realistic collagen fiber images, creating paired training data. Finally, a centerline extraction network is trained using both the real and synthetic image–centerline pairs[6]. This

approach to linking features (topology of fiber centerlines) could be useful for specific attributes like alignment that may be otherwise entangled in our GAN approach. Stein et al. introduce an algorithm for reconstructing the network geometry of collagen gels from 3D fluorescent images. Using synthetic databases, they show accurate recovery of structural parameters such as fiber length, cross-link density, and shear modulus[7]. This approach, when combined with our proposed GAN pipeline, could serve to validate future biomimetic models. These models could then incorporate tunable matrix stiffness which affects cell migration and motility[8].

Chapter 4 References

1. Kufel, J. *et al.* Multi-Label Classification of Chest X-ray Abnormalities Using Transfer Learning Techniques. *J Pers Med* **13**, 1426 (2023).
2. Armato, S. G. *et al.* The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans. *Med Phys* **38**, 915–931 (2011).
3. Härkönen, E., Hertzmann, A., Lehtinen, J. & Paris, S. GANSpace: Discovering Interpretable GAN Controls. *Advances in Neural Information Processing Systems* **2020-December**, (2020).
4. Buga, R. *et al.* Streamlit Application and Deep Learning Model for Brain Metastasis Monitoring After Gamma Knife Treatment. *Biomedicines* **13**, 423 (2025).
5. Sufian, M. A. *et al.* Enhancing Clinical Validation for Early Cardiovascular Disease Prediction through Simulation, AI, and Web Technology. *Diagnostics (Basel)* **14**, 1308 (2024).
6. Park, H. *et al.* Collagen fiber centerline tracking in fibrotic tissue via deep neural networks with variational autoencoder-based synthetic training data generation. *Medical Image Analysis* **90**, 102961 (2023).
7. Stein, A. M., Vader, D. A., Jawerth, L. M., Weitz, D. A. & Sander, L. M. An algorithm for extracting the network geometry of three-dimensional collagen gels. *J Microsc* **232**, 463–475 (2008).
8. Alkmin, S., Patankar, M. S. & Campagnola, P. J. Assessing the roles of collagen fiber morphology and matrix stiffness on ovarian cancer cell migration dynamics using multiphoton fabricated orthogonal image-based models. *Acta Biomater* **153**, 342–354 (2022).