

Image Analysis for Describing Meningeal Lymphatic Vessel Morphology

A
Dissertation
Presented to
the faculty of the School of Engineering and Applied Science
University of Virginia

in partial fulfillment
of the requirements for the degree

Doctor of Philosophy

by

Nazia Tabassum

May 2021

APPROVAL SHEET

This
Dissertation
is submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Author: Nazia Tabassum

This Dissertation has been read and approved by the examining committee:

Advisor: Scott T. Acton

Advisor:

Committee Member: Zongli Lin

Committee Member: Daniel Weller

Committee Member: Tom Fletcher

Committee Member: Barry Condron

Committee Member:

Committee Member:

Accepted for the School of Engineering and Applied Science:



Craig H. Benson, School of Engineering and Applied Science

May 2021

Image Analysis for Describing Meningeal Lymphatic Vessel Morphology

A Dissertation
Presented to
The Academic Faculty

By

Nazia Tabassum

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy in the
Charles L. Brown School of Electrical and Computer Engineering

Advisor: Dr. Scott T. Acton
Virginia Image and Video Analysis Lab
University of Virginia

December 8, 2020

Abstract

Neuroscientists have recently discovered the existence of meningeal lymphatic vessels in the brain and have shown their importance in preventing cognitive decline. With age, lymphatic vessels narrow, poorly draining cerebrospinal and interstitial fluids, which leads to plaque accumulation, a hallmark of Alzheimer’s disease. The analysis and detection of these vessels is performed by hand, and thus suffers from quantification variability. Furthermore, the only existing complexity measures currently extracted from images of these vessels are width and area, which are insufficient to capture morphological differences. This dissertation details the first automated segmentation and analysis methods developed for lymphatic vessels. The proposed segmentation approach, called LyMPhi, is a level set segmentation method featuring hierarchical matting to pre-determine foreground and background regions. The resultant approach eliminates the need for user-defined initialization, an advantage over competing methods, and produces smooth segmented contours. The level set force field is modulated by the foreground information computed by matting, while also constraining the segmentation contour to be smooth. Segmentation output from this method has a higher overall Dice coefficient and boundary F1-score compared to that of competing algorithms. The algorithms are tested on real and synthetic data generated by our novel shape deformation based approach. LyMPhi is also more stable under different initial conditions than comparative level set segmentation methods. Analysis can also be extended to studying elastic deformation for lymphangiogenesis as well as vessel narrowing. The deformation model can additionally be used for stretching existing vessel data into realistic synthetic data. Machine learning is explored and tested with application to segmenting the meningeal lymphatic vessels, showing promising results for the future. Manual segmentation, which is used as labels for neural network training, as well as a comparison for measuring segmentation accuracy, is analyzed with respect to statistical measures of variance. The last step in vessel analysis is to create relevant measures, many of which can be extended from already existing measures for other vessel networks. Given this work, neuroscientists have the first image analysis tools for meningeal lymphatic vessels, and subsequently can accurately capture information that will enable them to explore the vessels’ role in neurodegenerative diseases.

Acknowledgements

First of all, thank you to my collaborators, who are former and current members of the Center for Brain Immunology and Glia (BIG) at the University of Virginia, Department of Neuroscience. I would like to particularly thank Dr. Jasmin Herz, Dr. Michael Dong, Dr. Antoine Louveau, and Dr. Jonathan Kipnis. This research would not have been possible without their discoveries and interest in pursuing image analysis for the meningeal lymphatics.

I would like to thank my advisor, Dr. Scott Acton, for his continuous support and encouragement throughout the duration of my time at the University of Virginia. You have helped me grow in ways I did not think was possible. I owe any scholarship I have gained before today entirely to you. Thank you so much for teaching me everything I know about research, and of course, the beautiful subject of image processing.

Thank you to all of the members of both VIVA and VITAL labs, whose knowledge, generosity, and curiosity have made the journey incredible. You all are the best! The best of friends, and the best of colleagues, both current and future. I look forward to working with all of you.

Michael Ferguson, my undergraduate mentee, I owe a special thank you to. The deep learning chapter was our wonderful experiment together. I know your future will be so bright.

Thank you to my committee members for giving me their time so that I can share my work with them. Each of my committee members, both past and present, have influenced this work and my growth as a doctoral student. I would like to especially remember Dr. Malathi Veeraraghavan, who believed in me when no one else did, and gave me the strength and courage to keep fighting, whenever things got tough. She was as tough as they make them, and I hope to honor her memory.

I am forever indebted to my loving family and friends for being with me every step of the way. You know just when to push me, and just how to give me loving warmth when I need it. Abbu, Ammu, Bhaiyya, Zohair, and Allah (SWT) - this dissertation is for you. I hope it makes you proud.

Contents

1	Introduction	11
1.1	The Meningeal Lymphatic Vessels	11
1.2	MLV Image Acquisition	11
1.3	Problem Formulation	12
1.4	Contributions of this Dissertation	13
1.4.1	Level-Set Segmentation of MLV	13
1.4.2	Shape-based Data Augmentation	13
1.4.3	Application of Deep Learning to MLV	13
1.4.4	Analysis of Manual Segmentation	13
1.4.5	Complexity Analysis of the MLV	13
1.5	Dissertation Outline	13
2	Background on Microscopy	14
2.1	Cryo-EM	14
2.1.1	Single-Particle Reconstruction (SPR)	14
2.1.2	<i>Ab Initio</i> Modeling	15
2.1.3	Cryo-EM Structure Determination	16
2.1.4	Fitting an Atomic Model	16
2.2	Light Microscopy	17
2.2.1	Image Analysis with Light Microscopy	17
2.2.2	Technological Advancements for LM	17
2.2.3	Poisson Noise Model - Similarities with SAR	18
2.2.4	Particle Detection	19
3	Background on Meningeal Lymphatics	20
3.1	Immune Cells within the CNS	20
3.2	Disorders of the Central Nervous System	20
3.2.1	Alzheimer's Disease	20
3.2.2	Multiple Sclerosis	23
3.2.3	Parkinson's Disease	23
3.2.4	Autism	24
3.2.5	Amyotrophic Lateral Sclerosis	24
3.2.6	Huntington's Disease	25
3.2.7	Stroke	25
3.3	Lymphatic System	25
3.3.1	Glymphatic System	25
3.3.2	Function of the Meningeal Lymphatic Vasculature (MLV)	26
3.3.3	Complexity of the MLV	28
4	Level-Set Segmentation	31
4.1	Background on Segmentation	31
4.2	Proposed Method: LyMPhi	32
4.2.1	Matting	34

4.2.2	Level Set Formulation	35
4.2.3	Curve Evolution	38
4.3	Experiments Performed and Results	39
4.3.1	Description of the Datasets Used	39
4.3.2	Creation of Manual Annotation	40
4.3.3	Comparative Methods Tested	41
4.3.4	Measuring Segmentation Accuracy	41
4.3.5	Results	43
4.4	Analysis	47
4.4.1	Precision and Recall on the Boundary	47
4.4.2	Per-Image Scores	48
4.4.3	Stability of LyMPhi	48
4.4.4	Isolating the Effect of Foreground Propagation	49
4.5	Discussion	51
5	Data Augmentation	53
5.1	Background on Data Augmentation	53
5.1.1	Generative Adversarial Networks	53
5.1.2	Elastic Deformation	55
5.2	Shape Analysis	55
5.3	Synthetic Vessel Data	57
5.3.1	Vessel Shape Generation	58
5.3.2	Synthetic Image Formation	62
5.4	Segmentation Results	65
5.5	Discussion	66
5.5.1	Conditional Generative Adversarial Networks	66
5.5.2	Other Uses for Morphing	66
6	Deep Learning	68
6.1	Motivation for Deep Learning	68
6.2	Background on Deep Learning	68
6.3	Choice of Network	70
6.4	Training Data	71
6.5	Assessing Segmentation Results	73
6.6	Discussion	76
6.6.1	Working with a Small Training Set	78
6.6.2	Limits of Deep Learning	79
7	Analysis of Manual Segmentation	80
7.1	Analysis of Agreement between Annotators	80
7.2	Performance of LyMPhi and Each Annotator as Measured by STAPLE	81
7.3	Failure to Annotate Boundary by Manual Segmentation	82
7.4	Discussion	82

8	Complexity Analysis	84
8.1	Motivation for Complexity Analysis	84
8.2	Background on Vessel Complexity	85
8.3	Proposed Complexity Measures	86
8.3.1	Skeletonization	87
8.3.2	Vessel Length	87
8.3.3	Vessel Ramification Index	87
8.3.4	Vessel Porosity	89
8.4	Complexity Results	89
8.4.1	Measuring Segmentation Accuracy	91
8.5	Discussion	91
8.5.1	Possible Future Complexity Measures	91
9	Conclusions	93

List of Acronyms

3DZD	Three dimensional Zernike descriptors
Aβ	Amyloid beta
AD	Alzheimer's disease
AI	Artificial intelligence
AIM2	Absent in melanoma 2
ALS	Amyotrophic lateral sclerosis
ASD	Autism spectrum disorder
BACE1	Beta-secretase 1
BF or BF-score	Boundary F1-score
BIG	Center for brain immunology and glia
CCR7	C-C chemokine receptor type 7
CLEM	Correlative light-electron microscopy
CNN	Convolutional neural network
CNP520	Umibecestat
CNS	Central nervous system
Cryo-EM	Cryo-electron microscopy
CSF	Cerebrospinal fluid
CT	Computerized tomography
DAPI	4',6-diamidino-2-phenylindole
DD-SRAD	Distance-driven speckle reducing anisotropic diffusion
EAE	Experimental autoimmune encephalomyelitis
EEG	electroencephalogram
EM	Electron microscopy
fMRI	Functional magnetic resonance imaging
GAN	Generative adversarial network
GFP	Green fluorescent protein

GPU	Graphics processing unit
HD	Huntington's disease
High-NA TIRF SIM	High numerical aperture total internal reflection fluorescence structured illumination microscopy
Hz	Hertz
ISF	Interstitial fluid
JI	Jaccard index
L2S	Legendre level sets
LEC	Lymphatic endothelial cells
LLSM	Lattice light sheet microscopy
LM	Light microscopy
LoG	Laplacian of Gaussian
LyMPhi	Lymphatic matted phi
LYVE-1	Lymphatic vessel endothelial hyaluronan receptor 1
MAP	Maximum a posteriori
MLV	Meningeal lymphatic vasculature
MPP	Marked point process
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
MST	Minimum spanning tree
MVEE	Minimum volume enclosing ellipsoid
NL-means	Non-local means
NN	Neural network
NREM	Non-rapid eye movement sleep
PALM	Photo-activated localization microscopy
PaQ-2-PiQ	Patches to pictures
PCA	Principal component analysis

PD	Parkinson's disease
PFC	Prefrontal cortex
PT	Photothrombosis
RAM	Random-access memory
R-CNN	Region based convolutional neural network
RELION	Regularised Likelihood Optimisation
ResNet	Residual neural network
RI	Ramification index
SAR	Synthetic-aperture radar
SD	Sleep deprivation
SIM	Structured illumination microscopy
SIMPLE	Single-particle image processing Linux engine
SNR	Signal-to-noise ratio
SPR	Single particle reconstruction
SRV or SRVF	Square root velocity function
SSS	Superior sagittal sinus
STAPLE	Simultaneous truth and performance level estimation
STARmap	Spatially-resolved transcript amplicon readout mapping
STED	Simulated emission depletion
STORM	Stochastic optical reconstruction microscopy
SWA	Slow-wave activity
SWS	Slow-wave sleep
tDCS	Transcranial direct current stimulation
TEM	Transmission electron microscope
tMCAO	Transient middle cerebral artery occlusion
TuFF	Tubularity flow field
TV	Total variation

VEGFc	Vascular endothelial growth factor C
VEGFR-3	Vascular endothelial growth factor receptor 3
VFC	Vector field convolution

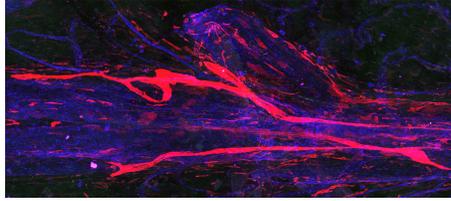


Figure 1: Image of meningeal lymphatic vessels, stained in red. Blood vessels are stained in blue. This image was acquired by Kipnis Lab, University of Virginia, Neuroscience.

1 Introduction

1.1 The Meningeal Lymphatic Vessels

In recent years, study of the lymphatic system present in our central nervous system has come to the research forefront. The lymphatic vessels surrounding the brain, which were unknown just six years ago, have been shown to drain waste from our brains. If this waste is not drained, it accumulates to form plaque, which contributes to neurodegenerative disease. Studies on mice have shown loss of memory and impairment in cognitive ability in cases where the vessels do not drain wastes adequately. [1]

The meningeal lymphatic vasculature (MLV) is present around major sinuses in the meninges, three membranes that line the skull and vertebral canal, enclosing the brain and spinal cord. These lymphatic vessels surround the cerebral cortex. Both the fluid part of the CSF (cerebrospinal fluid) and immune cells present in the CSF, such as T cells and B cells, are drained through the surrounding lymphatic vessels, all the way to the deep cervical lymph nodes.

Recent work shows that in old age, lymphatic drainage is reduced by shrinking of the lymphatic vessels. This can lead to cognitive decline and loss of memory which is common in Alzheimer’s disease. Studies on mice have shown loss of memory and impairment in cognitive ability if the vessels’ draining capacity does not function adequately. [1] A counter-effect is visible in the case of Multiple Sclerosis (MS). It is evident in MS that the lymphatic vessels are used to transport the immune cells responsible for autoimmune attack [2]. The attacks are directed to the myelin sheaths of neurons and can result in paralysis.

An underlying question remains: can we predict onset of these diseases or diagnose patients by studying an individual’s lymphatic vasculature? [3] Fueled by advances in microscopy, image processing will play an important role in the reverse engineering of the central nervous system [4].

As these vessels were only recently discovered, there is no tool available for automatically analyzing these images. By building the first, we can provide more measures for understanding these vessels. An image of the vessels, stained in red, is shown in Figure 1.

1.2 MLV Image Acquisition

The steps described following were applied for meningeal whole-mounts. Incubation with appropriate dilutions of primary antibodies was delivered: anti-LYVE-1–eFluor 660 or anti-LYVE-1–Alexa Fluor 488, anti-CD31, and anti-GFP. After incubating for 10 min DAPI,

the tissue was mounted with Aqua-Mount (Lerner) and glass coverslips. Preparations were stored at 4 °C for no more than one week until images were acquired either using a wide-field microscope (Leica) or a confocal microscope (FV1200 Laser Scanning Confocal Microscope, Olympus). [1]

1.3 Problem Formulation

There are currently no image processing tools for analyzing the meningeal lymphatic vasculature. This dissertation proposes the first.

The first step in analysis is segmentation, or detection of vessel objects in an image. Neuroscientists currently analyze these image data by hand, laboriously marking the boundaries of vessels using mouse clicks to draw a polygon boundary around the edges of a vessel. The time required to process one image, that is, to segment the image into its background and foreground, is around two hours. Manual processing is an incredibly slow and fallible method, and serves as the main hindrance to further research involving this data. Hand quantification tends to overestimate vessel content and cannot account for small holes or loops in vasculature. Furthermore, hand quantification results vary from person to person and even varies on a single image using the same individual, which can be measured. The work hopes to alleviate the time taken to process the meningeal lymphatic vessel images, and therefore speed up research in the areas using these images.

The image data used in analysis is difficult and expensive to produce, as each image of vessels represents one sacrificed mouse. So, it would be useful to create a synthetic image database of meningeal lymphatic vessels for use in further study. Also, a synthetic database has the added benefit of not containing the errors discussed above in manually annotated ground truth. Creating a model for generating realistic vessel images will allow neuroscientists to perhaps use machine learning techniques for analyzing the vessel data; these learning algorithms require more data than is currently available. As neuroscientists learn more about the morphology changes of lymphatic vessels due to disease or age, these changes can also be modeled, replicated, and further studied.

Quantification and complexity analysis of the meningeal lymphatic vessels is currently limited to calculating area of the vessels and average width of vessels in one image. The average width is calculated by sampling the vessel at 100 points and averaging over the widths found at these points. As different users will select different points to sample, the average width computed from user to user for the same image can vary up to 30 percent. This is unreasonable for a metric that should provide a reliable comparison from image to image. Also, averaging the width of vessels that are highly variable in size does not provide enough information on vessel morphology throughout the mouse brain. Quantitative measures already used to characterize other vascular networks should be used to describe the meningeal lymphatics, as well as new metrics developed to capture details unique to these newly discovered vessels.

This dissertation is for the first tool-set developed specifically for the above mentioned unique challenges of analyzing meningeal lymphatic vessels.

1.4 Contributions of this Dissertation

The contributions of this dissertation are listed here.

1.4.1 Level-Set Segmentation of MLV

The first contribution is to develop a level-set based segmentation method for accurately segmenting meningeal lymphatic vessels within a highly cluttered background.

1.4.2 Shape-based Data Augmentation

The second contribution is to build a rich database of realistic synthetic meningeal lymphatic vessel images that can serve as ground truth for image analysis studies [4].

1.4.3 Application of Deep Learning to MLV

Third, smaller images of the MLV are segmented using a deep learning approach, despite the challenge of a small training dataset. Promising segmentation results are shown compared to traditional segmentation methods, such as level-set methods.

1.4.4 Analysis of Manual Segmentation

Fourth, an in-depth analysis of our manual segmentation is performed, to study annotator error and reliability.

1.4.5 Complexity Analysis of the MLV

Finally, the fifth contribution is to develop a set of informatics for characterizing meningeal lymphatic vasculature, aiding in distinction between groups that have differences in lymphatic form and/or function. These informatics can also be used for assessing segmentation accuracy.

1.5 Dissertation Outline

Chapters 2 and 3 are background chapters, on microscopy methods and the meningeal lymphatics, respectively. Chapter 4 covers our level-set segmentation method for the MLV, called LyMPhi. Chapter 5 details a novel shape deformation based data augmentation method for generating MLV data complete with ground truth. Chapter 6 goes over testing of a CNN with the meningeal lymphatic images. Statistical analysis of our hand annotation is pursued in Chapter 7, highlighting important variation and errors in the manual segmentation. The final contribution, complexity analysis of the MLV, is discussed in Chapter 8. Finally, Chapter 9 concludes this dissertation.

2 Background on Microscopy

This section contains a technical background on microscopy, specifically the imaging and image processing procedures followed. The two main types of microscopical imaging are highlighted here, cryo-electron microscopy and light microscopy.

2.1 Cryo-EM

Cryo-electron (cryo-EM) microscopy is essential for the study and functional understanding of non-crystalline macromolecules, such as proteins. These molecules cannot be imaged using X-ray crystallography or other popular methods. Cryo-EM has been successfully used to visualize molecules such as ribosomes, viruses, and gap junction channels, for example. Obtaining structural models of these at various conformational states leads to insight on how these molecules interact with others. The drive for using electron microscopy is to enlighten us on the roles of individual proteins so as to create a complete and measurable model of the workings inside (and outside) a single cell. [4]

Recent advances in imaging technology have given cryo-EM a scientific rebirth, achieving near-atomic resolution of non-crystalline specimens [5]. Because of imaging improvements, image processing and analysis of the resulting images have increased the resolution such that molecular structures can be resolved at the atomic level. Cryo-EM is ripe with stimulating image processing challenges. In this section, we will touch on the most essential information in order to build an accurate structural three-dimensional model from noisy projection images. [4]

2.1.1 Single-Particle Reconstruction (SPR)

Cryo-electron microscopy is often used for the purpose of single-particle reconstruction. The imaging process is described as follows: A plastic tube contains the isolated molecular sample of interest. Inside the tube, nanoparticles are free to move. The sample must be frozen in order for the molecule to be imaged. A miniscule drop ($3 \mu\text{L}$) of the sample is placed onto a 3 mm copper grid. [6]

To move from sample in solution to frozen molecules, the sample is loaded into a machine with a chamber that maintains 100% humidity and a temperature of 4° Celsius. There are blotting papers on either side of the sample, which close on the grid and leave a thin layer of molecules in solution. The flattened sample is then rapidly plunged into liquid ethane, surrounded by a ring of liquid nitrogen. The sample is frozen so quickly that ice crystals cannot form. This amorphous, thin ice is called “vitreous ice”, and is essential to the imaging process, as there is no crystalline order to interfere with the frozen particle structures. [6, 7].

Following is the imaging procedure, to reconstruct a 3D structure from the frozen molecules. The grid containing the frozen particles is loaded into a transmission electron microscope (TEM). Underneath the frozen sample is a detector, such as photographic film. The molecules are fixed in different orientations. When the molecules are shot through with an electron beam, the orientation of the particles leaves a unique “shadow.” These “shadows”, or projections, contain all the 3-dimensional information of the molecule, compressed into a 2-dimensional image. An image collected using the electron microscope is called a micrograph,

and it contains many noisy 2D projection images of a single molecule. [6]

The molecules are trapped in the ice in different orientations, or viewing directions [8]. Projections from the same 3D orientation can be clustered, often using K-means. Projections from the same cluster are rotated so that all projections are in the same in-plane orientation. These “aligned” molecules are then averaged together. This aggregate image (called a “class average”) provides a less noisy view of the molecule in this 3D orientation [9]. We perform this same process for the different 3D orientations of the molecule. Combining 2D templates leads to an estimated 3D structure for each particle. [5,6]

One major challenge of SPR is that there is no way to objectively measure the quality of the resulting particle reconstruction. (Although, more recently, [10] proposes that their *ab initio* model could be used for such quality comparisons.) There are imaging challenges, such as: overexposure to electrons causing radiation damage to specimens, but underexposure resulting in the noisy projections. [5]

After collecting all averaged views, they are combined computationally. [6] If no similar previously solved structures are available, an initial particle model (*ab initio*) is generated based on physical principles, and then, through iterative refinement, the final 3D structure is reconstructed.

2.1.2 *Ab Initio* Modeling

Ab initio modeling estimates the base structure of a molecule. This is the initial step of a procedure called iterative refinement, where the base structure of a molecule is then refined until convergence to reconstruct a single particle (SPR.) If done manually, this initialization can lead to bias and slow convergence. [10] In regards to initial volume estimation, most recent methods use stochastic optimization or geometric constraints imposed by the Central Slice Theorem. [4]

Most *ab initio* modeling algorithms use the idea of common lines and the Fourier slice theorem. The 2D Fourier transform of a projection image is equivalent to a slice through the origin of the 3D Fourier transform of the molecule’s potential function. For asymmetric molecules, any two projections share just one “common line.” To find these common lines, [11] discretize the Fourier space into 360 possible lines per projection, and pick the line that has the highest cross-correlation between each pair of projection images. Then, the authors find the rotations of the projections that best satisfy all of the common lines. The reconstruction is achieved by using the 3D pseudopolar Fourier transform. The approach is robust to noise and is not dependent on the number of projections. One area of concern is the use of class averages for projection images, as class averages do not represent true projections and often lead to misclassification of common lines. [11]

The following two methods are software tools that can be used for solving cryo-EM structures. SIMPLE is a software for *ab initio* reconstruction of heterogeneous single-particles. The projection images are first clustered using a probabilistic PCA based algorithm, which has features of K-means. Then class averages are taken. The main novelty of SIMPLE is in finding the orientation of each projection using a simulated annealing approach. To reconstruct the volume, the authors use a Fourier inversion algorithm with the aligned projections. [9]

PRIME is a probabilistic initial 3D model generator. The creators of PRIME also cre-

ated SIMPLE the previous year, but have improved their techniques to create a faster and more advanced mechanism for creating 3D cryo-EM maps. This time, they use stochastic hill-climbing in local neighborhoods to find the best matches for orientation given a single particle image. These best orientation matches are given weights, so that the final orientation assigned to a particle image is actually a weighted average of all discretized orientations. This is called a sparse orientation weighting approach, because the majority of the weights are zero. The 3D molecule is reconstructed as usual with Fourier inversion. The main benefit to using this algorithm is that there is no alignment, clustering, or averaging necessary - PRIME works well on noisy data. PRIME can be used to create final high resolution cryo-EM maps, although some detail may be missing, so the method is more useful for generating *ab initio* models. [12]

2.1.3 Cryo-EM Structure Determination

The process used for refinement if using *ab initio* modeling (when no known templates exist) is common across many algorithms. Refinement is done using 3D projection matching monitored by the Fourier shell correlation curve. There is no “gold standard” for refinement, particularly because of over-fitting to noise - this recalls a major limitation that there is no complete standard for assessing validity of the reconstruction results. [5]

RELION is a popular implementation of a Bayesian approach for cryo-EM structural determination. RELION follows a Maximum a posteriori (MAP) estimation approach to find an optimized solution to the reconstruction problem. The novelty of RELION is in implementing the well-defined mathematical method into software that is efficient computationally. There are a number of ways in which the authors decrease computational time - such as only searching for orientations similar to the orientations of neighboring projections, instead of over the whole set. Another improvement introduced here is to prevent over-fitting of the reconstruction to noise by performing refinement on two halves of the data and then combining the two converged refinements. The RELION framework is notable for containing many image processing procedures for cryo-EM in one pipeline, without the need for excessive tuning - and doing so quickly, efficiently, and without increasing reconstruction error. There is still a need for structure initialization with RELION, however, which makes the *ab initio* modeling discussed previously of necessity. [13]

2.1.4 Fitting an Atomic Model

We have discussed methods for building and refining a 3D reconstruction of the molecule being imaged. This reconstruction is in reality a density map. What researchers are truly interested in is an atomic level structural model of the molecule. After providing a 3D reconstruction, structural features down to the atomic level can be fit inside the molecule. A model fitting is performed for secondary structure elements (SSEs) such as α -helices and β -sheets. [14,15] Fitting such a model allows for understanding of how the molecule functions inside a cell. [6] Model fitting is performed by introducing some constraints, by using mass spectroscopy, proteomics, and chemical cross-linking. Having this knowledge at the atomic level can aide in analysis of possible deformation paths of molecules. [4].

2.2 Light Microscopy

2.2.1 Image Analysis with Light Microscopy

Both light microscopy and electron microscopy are used to investigate the interior of organisms. Light microscopy (LM) has limitations to overcome, just like EM. The limitations of LM are imposed by the characteristics of photons, such as fluorescence properties, by aberrations in the optics, and the overall resolution of the microscope. Also, the available photon budget of the specimen being imaged also limits the image reconstruction quality. LM is characterizable by the nature of the object being observed, such as the cell or organelle, and also by the association or recombination of those quantities, such as through membrane fusion. [4]

Low signal-to-noise ratio (SNR) and multitudes of objects in the field of view complicate tracking problems, and it is difficult to solve the correspondence problem [16]. Some image analysis problems of interest include tracking in bioimaging [17–19], sub-diffraction limited imaging [20,21], and parametric and geometric active contour based segmentation [22]. Other important topics include shape analysis [23], neuronal tracing [24], image deconvolution [25], and spot detection [26], such as in STARmap [27]. A research group from Stanford has developed a technology called STARmap for identifying the proteins expressed at single-cell resolution in **three-dimensional** intact mouse tissue samples. The different proteins can be seen by staining each nucleotide (or amino acid) a different color, so they can read the tRNA from a multi-channel image. By observing the protein expressed in a single-cell, the authors can determine what type of cell it is - excitatory neurons, inhibitory neurons, microglia, etc. Not only can the authors view overall numbers of cell types in different brain regions, distributions of cell types can be also be characterized, down to single cell nearest neighbors. Unexpected cell groupings and even new cell types have been found. STARmap is less noisy and error-prone than previous methods and can handle large quantities of gene markers. The experiment is performed immediately after some stimulus is given to the mouse, and so can capture important information about neuronal response and brain function. [27]

The MLV specimen is tagged with a fluorescence protein which emits photons when illuminated by some light source [28]. The photons are detected by a sensor to produce an image. Laser scanning confocal microscopes are used for fast imaging of MLV from mice. Co-localization between two proteins, such as VEGFR-3 and LYVE-1, can provide stronger evidence for the presence of lymphatics in an image.

2.2.2 Technological Advancements for LM

Recent technological advancements include multi-photon, STED (simulated emission depletion) microscopy [29], PALM (photo-activated localization microscopy) [30] and STORM (stochastic optical reconstruction microscopy) [31]. Other advancements include coherent beam sources, faster detectors, more sensitive molecular probes, and automated imaging procedures. [4]

Resolution that surpasses the 200 nm diffraction limit has been a running goal of LM. Structured illumination microscopy (SIM) [32,33], produced 3D reconstructions of specimens with resolution close to 100 nm. STED [29], STORM [31,34], and PALM [30,35], which won the 2014 Nobel prize in chemistry, can provide single molecule localization within cells -

which is at the scale of just dozens of nanometers. Lattice light sheet microscopy, or LLSM, also allows for super-resolution [36]. PALM, STORM, and SIM improve resolution with the cost of slowing acquisition rate. A 28x28 micron square image with resolution 60 nm takes about one minute to record using PALM [37]. This mostly restricts application to fixed samples, instead of live imaging. High-NA TIRF SIM (numerical aperture) [38] and LLSM [39] allowed imaging with high resolution of the associations between cortical filaments and proteins and three dimensional intracellular interactions. [4]

Another active field in microscopy imaging is finding ways to combine EM with LM, such as with CLEM microscopy [40]. This brings together the advantages of live fluorescence with the high resolution that EM provides. [4]

2.2.3 Poisson Noise Model - Similarities with SAR

Live-cell fluorescence imaging is limited by time, in order to prevent photo-bleaching or phototoxicity. Image denoising can be used to mitigate these effects, and allow for longer time-series of images to be taken [41–43]. The major advantage of denoising algorithms over frame-averaging is to retrieve denoised time series from acquisitions that have low SNR [4]. This can be observed outside of microscopy, in SAR (synthetic-aperture radar) or satellite imaging as well - using methods such as DD-SRAD [44] provides a denoised time-series from low quality data, without losing the time information.

Poisson and Poisson-Gaussian noise is generally measured when using fluorescence microscopy [4]. The imaging set-up consists of the optical system followed by a photodetector and accompanying filters. The photodetector converts the optical intensity in the form of photons into electrons. However, the signal is generally disrupted by various noise sources during the acquisition process. Poisson noise is particularly degrading when there are low-lighting conditions, shorter exposure times, and inefficient photon detectors - all of which are especially found in LM. Signals are thus known to be damaged by intensity dependent Poisson noise. [4]

A key challenge in Poisson estimation is that the observed counts have different variances. The Poisson noise effect increases as the mean intensity value in the image decreases - and the overall SNR decreases as a result. [4] Again, this is similar to the effects of speckle or multiplicative noise present in satellite imagery - the signal and noise are not independent, and thus not easily partitioned [44].

Variance stabilization techniques are used to combat Poisson noise effects, by converting Poisson noise into Gaussian noise with unit variance [45–48]. Subsequently, commonly used denoising methods for additive white Gaussian noise can be applied. Similarly, with speckle, a logarithmic transform is often used to convert the multiplicative denoising problem into an additive one [44]. Stabilization algorithms provide effective denoising capabilities given the collected photon count is high [47, 48]. For low SNR, there are other denoising methods that avoid "Gaussianization" of the signal [4].

While the NL-means filter [49] is considered efficient and effective for additive white Gaussian noise, in the case of Poisson noise, the method needs to be combined with a Fisz transform [156,157] to provide the same noise reduction while preserving image geometry. The NL-means, or non-local means algorithm, exploits natural image redundant patches to restore the signal. For SAR imagery, simple NL-means also does not remove all high intensity

speckle and also leaves behind texture artifacts [44].

For Poisson noise reduction, wavelet-based methods (including ridgelets and curvelets) [46, 50, 51], such as the Haar wavelet [52, 53], are also used for image denoising. Wavelets, and similar formulations such as shearlet transforms [54], have also shown success in the satellite imaging domain.

Compressive sensing is also at the research forefront to minimize the number of measurements needed during biological imaging [55, 56].

For preservation of live samples during imaging, the illumination intensity is set to lower levels. This causes an increase in noise and resolution loss, to the effect that image components, particularly sub-cellular structures, become blurred. [4] Popular deconvolution methods for fluorescence microscopy minimize an energy functional that includes two terms: a data fidelity term based on noise statistics, point spread function, etc., and a regularization term based on a prior. The solution is generally constrained to be positive. Regularizers to suppress noise while restoring details and structure include Total Variation (TV) methods [57], also used in denoising of SAR imagery [44, 58]. Deconvolution methods also are used for LLSM, for recovering structural details and improving spatial resolution, even when illumination is pushed to extremely low levels in order to limit photo-damage/toxicity [36].

2.2.4 Particle Detection

In LM, the assumption generally is that the background is static, and the object of interest has higher intensity. To extract the particle, thresholding may be used. However, choosing an appropriate threshold is still an unsolved problem. Histogram analysis can be used to estimate a threshold, perhaps by using Otsu's method [59]. If SNR is low, however, simple thresholding cannot capture objects of interests, as some pixels belonging to particles can appear dim. The background is not generally uniform, as well. Cluttered backgrounds including structures of varying size, such as what is observed in images of MLV, generally create poor detection results when global thresholding is applied. Considering spatial coherence between pixel neighbors, such as in Hierarchical Image Matting [60], greatly improves detection and segmentation performance. This helps avoid misclassification of isolated pixels in the image. [4]

Local maxima intensity where the local curvature is high (determined by the Hessian), can also be used to determine objects of interest [61]. The Laplacian of Gaussian, or LoG, also shows a strong ability to detect object location [62]. The LoG is a bandpass filter that only enhances objects of a specific size, thus reducing the presence of lower or higher frequencies (either background clutter or noise.) This is similar to the determination of vesselness proposed by Frangi [63]. The choice of bandwidth for both the LoG filter and for Frangi's method is critical, and difficult to tune for images with highly variable object size. If the object scale is parameterized too small, noise induces over-detection. If the parameter is too large, objects become smoothed or merge together.

This chapter has been a general overview of microscopy methods and some interesting image processing and analysis challenges as they relate to this dissertation topic. The next chapter will provide an in-depth understanding of the biology of the lymphatic system, as well as accompanying diseases of the CNS.

3 Background on Meningeal Lymphatics

3.1 Immune Cells within the CNS

The traffic of T-cells between the meninges and the cerebrospinal fluid is of great interest, because these activated T-cells are involved in many diseases of the central nervous system. The first *in-vivo* to demonstrate communication within the meningeal compartment that drives the property of T-cells to cause disease was [64]. Multiphoton imaging was used to show the meningeal macrophages communication with encephalitogenic T-cells prior to infiltrating the brain parenchyma. Similar molecular mechanics were shown to play a role in the parenchymal infiltration of T-cells and in regulating contact with meningeal macrophages. [64, 65]

The first study that used single-cell approaches in order to analyze the CNS-immune compartment in both normal and pathological conditions was published in 2018 [66]. Using single-cell mass spectrometry, the authors shine light on the diversity and complexity of the immune cells present in the CNS. The authors also demonstrate the presence of resident neutrophils within the meninges. [65, 66]

Meningeal T-lymphocytes produce cytokines (such as interleukin-4 and interferon-gamma) that regulate cognitive and social behavior in mice. There are many different immune cells in the brain. During neuroinflammation (modeled by experimental autoimmune encephalomyelitis in mice), lymphoid cells 1 and mast cells play an important role. Lymphoid cells 1 generate a pro-inflammatory environment and encourage filtration of cells through the brain parenchyma. The immune cells in the meninges activate pathogenic T-cells and regulate their migration into the central nervous system (CNS.) [65]

Glia occupy some 80% of the human brain. Microglia are the tissue resident macrophages of the brain parenchyma and have varied roles in brain development, homeostasis, injury, and disease. [67] There is a growing need to study microglia (the specialized immune cells of the CNS) and test anti-inflammatories as treatment for neurodegeneration caused by inflammation [68].

3.2 Disorders of the Central Nervous System

3.2.1 Alzheimer's Disease

Alzheimer's disease (AD) is *characterized* by the formation of β -amyloid into plaques in the brain. Studies have shown that fifty percent of normal elder people have abnormal amyloid without symptoms of dementia. Dementia occurs about twenty years post development of amyloid abnormality; thus, AD is considered to be progressively neurodegenerative. A person may live for a long time with AD with no major cognitive disruption. [69]

Trials are currently being held to test the efficacy of amyloid-modulating drugs on cognitively normal patients with abnormal amyloid biomarkers. Other treatment methods and targets should be studied, such as **lymphatic vessels**, to prevent abnormal amyloid buildup. [69]

Policy makers will have to consider the opportunity cost of treatment - many left untreated will pass on before any cognitive decline. But once brain damage occurs, treatment

methods become ineffective. With early treatment, it may be possible to live with AD and never develop dementia. [69]

Mesquita *et al.* [1] found that using a compound in aged mice to grow lymphatic vessels and promote drainage helps aged mice learn and retain better. Obstructing the lymphatic vessels led to the accumulation of amyloid plaques similar to those present in human Alzheimer's patients. The scientists are now working on a drug for human use, and are hopeful that by maintaining healthy lymphatic vessels, the effects of AD will be mitigated in the future. [70]

It is well known that AD patients complain about sleep disturbance. The relationship between sleep and $A\beta$ plaques is bidirectional. Mouse and human models have found plaques increase with sleep disruption (measured by $A\beta$ in interstitial fluid.) During sleep is when the glymphatic system clears wastes - it cleans metabolic waste during slow-wave sleep (SWS). Neuronal firing also promotes $A\beta$, and there is less of this firing during SWS. [71]

$A\beta$ increase occurs during all types of interrupted sleep: insomnia, sleep apnea, etc. But this interrupted sleep is always self-reported, and thus not quantitative. $A\beta$ also disrupts the CNS clocks' synchronization, which makes sleep worse. On mouse models, chronically shortened amount of sleep time (sleep loss) led to finding more $A\beta$ in the cerebral cortex. Evidence has also been found that mice may not be able to acclimate to a stressor due to sleep loss. Sleep deprivation is also a stressor in and of itself, which impairs cognitive function on its own. Chronic sleep deprivation (SD) could be a risk factor for AD, even to those who are non-genetically predisposed, or those who develop sporadic AD (late-onset, most common form.) Similar results as discussed above have also been found in fruit flies. [71]

In humans, hippocampal activation captured by fMRI in cognitively impaired patients shows that the SD is augmenting neuronal excitability. Also in humans, short term total SD leads to non-removal of $A\beta$ (which is normally cleared overnight.) The levels remain high through the day. The above studies have found other possible biomarkers for AD - CSF orexin is one example, tau protein another. We are moving "beyond amyloid": AD is multifactorial, and depends on the responses from many causal factors interacting, such as astrocytes, dysfunctional microglia and vasculature (blood and lymphatic.) Simple proof of this: removing the plaques does not improve cognitive function. Both $A\beta$ without AD (as in [69]) and AD without $A\beta$ are possible. Presence of the glymphatic system (perivascular for CSF connected to downstream lymphatics) in the human brain has not yet been proven. [71]

Slow wave activity (SWA) measured by EEG in Hertz (Hz) decreases with age. Memory impairment is related to continuous hippocampal activation and reduced connection between the hippocampus and prefrontal cortex (PFC.) Studies have now shown correlation between $A\beta$ in the medial PFC with disrupted non-REM (NREM) SWA, also with impaired memory transform (short to long-term) and retention overnight. The SWA waveforms also change. In rodents, bidirectionality of SD and $A\beta$ has been shown with shorter, fragmented NREM. [71]

Potential treatments targeting sleep are sleep hygiene (exercise, dim light, etc.), melatonin (didn't improve cognitive performance in drug tests), and transcranial direct current stimulation (tDCS) during sleep. The last is effective for memory consolidation in multiple patient types (schizophrenia, etc.) However, long-term evidence of efficacy is yet to be shown. We need to better understand the different sleep stages and the progression of AD. With this knowledge, non-invasive treatments targeting sleep restoration can be developed. [71]

Thus, a major takeaway for the future of AD research is the move away from the classical

amyloid cascade hypothesis. [71]

In a news article published by Scientific American [72], it was reported that slow wave sleep (non-REM, slow electrical oscillations) helps memory consolidation. It also drives CSF and blood flow through the brain. We know wastes are cleared during sleep. But in the Science article [73], Laura Lewis from Boston University asks *why* during sleep? [72]

The neuroscientists used EEG and fMRI to measure electrical activity, blood oxygenation, and CSF flow, all at the same time in sleeping adults. As brain activity affects blood flow, volume of blood in the brain is reduced, CSF comes in to fill this volume. They measured CSF coming into the fourth ventricle (produces CSF by filtering blood plasma.) The wave flow pushes CSF around ventricles and into the meninges (space between membranes surrounding brain and spinal cord), which carries away toxic wastes. The CNS combines “making decisions” (like a boss in a company) about memories with “cleaning” (like janitors.) In nature, everything is done simultaneously instead of separated and is thus more efficient. Sleep disturbance is common in Alzheimer’s (characterized by a build-up of amyloid- β , toxic proteins not cleared by CSF) as well as depression and schizophrenia, so there are many related disorders of interest. In the future, neuroscientists want to do studies manipulating neural activity in mice (during sleep) and observe the consequences. Modulating brain activity could be a treatment for humans. Studying sleep signatures can also be used as a diagnostic. One major takeaway: sleep is important for this “housekeeping function” of the brain. [72]

Amyloid plaques and neurofibrillary tangles are markers of Alzheimer’s disease. In the past, treatment methods for Alzheimer’s disease have been focused on breaking up plaques. But by the point that plaques have been formed, cognitive decline has already set in and cannot be undone. In reality, the proteins that make up the plaques and tangles are the deep-rooted problem: amyloid- β and tau. Toxic amyloid- β makes good tau protein unfold (like a paperclip) and become toxic. This in turn kills neurons and causes failure at synapses. Tau needs to remain healthy for Alzheimer’s prevention. [74]

Likely this whole process begins many years before the onset of cognitive decline. Humans have so many neurons that our brains can compensate for this toxic conversion until the rate of neuron death increases to an unsustainable rate. This is when Alzheimer’s becomes symptomatic, but the damage is already too great. Dr. George Bloom, a professor with UVA Biology and the UVA Brain Institute, suggests genetic testing to see who is at risk, to then offer treatment up to 15 years prior to symptom onset. His work has most recently been published in the Journal of Alzheimer’s Disease [75]. Memantine (Namenda) is one such potential drug that might prevent tau unfolding into its toxic state. [74]

Of course, better tests need to be developed to determine who is at risk for Alzheimer’s - unnecessary treatment should not be given to those who will not develop the disease. Imaging, such as MRI, of the brain and meningeal lymphatic system could possibly reveal early changes in the brain and risk factors. We are many years from effective prevention of Alzheimer’s, as there are long-term clinical trials that need to be undergone for any drugs that are developed. Studying this disease is expensive, but the consequences of not adequately researching prevention are far greater. Alzheimer’s disease affects 6 million Americans today, and costs \$300 billion dollars for healthcare and treatment overall. The National Institutes on Aging predict that by 2050, these numbers will increase to 16 million people and \$1.1 trillion dollars. Clearly, we need to work towards a solution. [74]

In 2019, drug trials ceased as the cognitive function of Alzheimer’s patients worsened [68]. Trials were performed by Amgen and Novartis Neuroscience Collaboration (launched in August 2015) working with Banner Alzheimer’s Institute. The trials were testing BACE1 inhibitor CNP520 (umibecestat.) The inhibitor blocks the BACE1 enzyme that forms amyloid plaques. The failed trials mark another proof against amyloid theory of Alzheimer’s disease. [68]

The current thinking is that the plaques are triggering inflammatory immune reactions that causes neurodegeneration. It is suggested that **simple drainage of lymphatics** is not the answer; the **shape and branching** also matters in how the vasculature interacts with its environment. Lymphatic vessels carry the immune cells responsible for the possible autoimmune attack mentioned above; treatments can target these vessels’ transport as well. [68]

3.2.2 Multiple Sclerosis

The relationship between lymphatic vessels and disease is complicated. Vessels may help develop conditions for Multiple Sclerosis (MS), which is not yet well understood. MS is an autoimmune disorder where the immune system attacks the myelin sheath around nerves, which disrupts communication throughout the body. A nerve affected by MS has a myelin sheath that has been destroyed and the nerve fiber is exposed. A patient affected with MS may lose the ability to walk.

When CNS immune privilege is undermined, autoimmune attacks occur, such as in multiple sclerosis. In this disease, immune cells specific to the CNS self-antigens cause damage to the CNS. This results in paralysis. [3]

The MLV promote fluid homeostasis, but like other lymphatic networks, also aid in the recycling of immune cells. In multiple sclerosis models, ablating the MLV has a positive effect, because this allows the antigens, including those that attack the myelin sheath of neurons, do not travel as freely. [65] In new research, the Kipnis lab, formerly at UVA, has found that lymphatic vessels are linked to multiple sclerosis, and possibly other neuroinflammatory diseases. There is a signal sent from the brain to the lymph nodes that requests immune cells to reenter the brain. These immune cells travel through the lymphatic vessels and are responsible for autoimmune attack, which can result in paralysis. [76] The positive effect is shown by lowering paralysis in mice [65].

Blockading the MLV decreases pathogenicity and associated pathology in EAE, which is the mouse model of multiple sclerosis [2]. Removing the vessels does decrease the number of immune cells present, but does not completely remove MS from a mouse. We need to further understand how the signal is sent (cellular, molecular, etc.) before researching treatment options. Contrary to in Alzheimer’s disease, during neuroinflammation, there is little change in the size or complexity in the lymphatic vessels. [76]

3.2.3 Parkinson’s Disease

Parkinson’s disease (PD), like Alzheimer’s disease (AD), is another neurodegenerative disease that displays abnormal accumulation of proteins in the brain parenchyma. Inadequate removal of faulty brain cells during neurodevelopment can lead to long-term behavioral

issues [77]. Modulating the MLV function does affect both PD and AD in terms of pathophysiology: the reduced drainage leads to more protein buildup for both diseases, which further incurs neuronal damage and behavioral degeneration. [65]

The cellular cleaner is the AIM2 inflammasome, which has previously been associated with infection immune response, but was first studied in the brain by [78]. The inflammasome plays an important role in making sure the developing brain is properly assembled and correctly functioning. This form of cleaning, by causing cell death, plays a direct role in removing the unwanted cells from the CNS to establish a healthy brain, which has the correct connections and the appropriate number of cells. [77]

Ataxia is a condition where individuals lose control of movement, a common symptom of PD. This pathway could be contributing to neuronal loss that is prevalent with ataxia. The cleanup is necessary, but over-cleaning can have negative consequences - potentially, ataxia. Early-onset neurodegenerative diseases are associated with mutations in DNA damage repair proteins, so this pathway could also be involved. [77, 78]

3.2.4 Autism

Autism spectrum disorder (ASD) refers to a wide range of conditions characterized by challenges with social skills, speech, and nonverbal communication, as well as exhibiting repetitive behavior [79].

An unexpected form of cellular cleanup takes place in developing brains, as discussed in Section 3.2.3. If too little or too much cleaning takes place, permanent changes take place in the wiring of the brain. In experiments performed on laboratory mice, these changes result in anxiety-like behavior. The changes may play a role in neurological conditions in humans, such as autism. [77]

Brain cells that have genomic compromises, or damaged DNA, are normally expelled from cells instead of being incorporated into the CNS. However, when the damage goes unrecognized, cells containing DNA damage live on in the CNS and are evidenced by the accumulation of DNA damage in the brain. Over half the neurons created during development of the brain end up dying, so proper removal is necessary. Again, too much or too little cleanup is hypothesized to underlie autism and even intellectual disability - or any other type of neurodevelopmental disorder. [77, 78]

3.2.5 Amyotrophic Lateral Sclerosis

Amyotrophic lateral sclerosis (ALS) is a progressive nervous system disease that affects nerve cells both in the brain and spinal cord, which causes loss of muscle control. ALS is often referred to as Lou Gehrig's disease. [80]

The discovery of the AIM2 inflammasome pathway's involvement in cell cleanup (discussed in Sections 3.2.3 and 3.2.4) came about somewhat by accident. Researchers [78] were observing the behavior of laboratory mice during an investigation of traumatic brain injury. That unexpected lead gave the scientists a better understanding of the developing brain, which may in the future yield new treatment possibilities for neurological disease. [77, 78]

Such treatment options are likely far in the future, but such a therapy could have widespread benefits. A treatment strategy for the AIM2 inflammasome pathway in the

adult brain would likely prove efficacious for most neurodegenerative diseases that are a result of DNA damage. That includes the major diseases discussed in Sections 3.2.1 and 3.2.3, as well as ALS. [77, 78]

3.2.6 Huntington’s Disease

Huntington’s disease (HD) is a fatal genetic disorder. HD causes progressive breakdown of nerve cells in the CNS. This breakdown causes deterioration of an individual’s physical and mental abilities, usually during their prime working years. HD has no cure. [81]

After ablation of the MLV, the authors of [1] found fold changes in genes, particularly those associated with Huntington’s, Parkinson’s, and Alzheimer’s diseases.

3.2.7 Stroke

Stroke occurs when the blood supply to a part of the brain is interrupted or reduced, preventing the brain tissues from receiving oxygen and nutrients. Brain cells can begin to die in just minutes. A stroke is considered a medical emergency, and prompt treatment is crucial. Early action could reduce brain damage or other complications. There are two overarching types of stroke: ischemic, due to a lack of blood flow, and hemorrhagic, due to bleeding, or hemorrhaging. [82]

In stroke, the method of stroke affects the morphology of the MLV. For example, photothrombosis (PT) induces lymphangiogenesis, while in transient middle cerebral artery occlusion (tMCAO), lymphangiogenesis is not observed. On the other hand, the absence of MLV during tMCAO worsens the stroke conditions, but this is not the case during PT. [65] We still need more understanding on the relationship of vessel structure with different diseases.

3.3 Lymphatic System

3.3.1 Glymphatic System

The glymphatic system, or paravascular system, refers to the peripheral lymphatics, outside of the CNS. This system facilitates the flux of CSF (cerebrospinal fluid) in and out of the brain parenchyma to allow waste removal. [65, 83]

Lymphatic vessels transport pathogens to lymph nodes so they can start creating antibodies and build immunity. Lymphatic vessel endothelial cells making up the vessels express hyaluronan receptor 1 (LYVE-1). The lymphatic valve system made of flaps; opening a valve lets in pathogens and other cells. These are on initial lymphatic vessels where cells enter through “button” pattern junctions between cells. This implicates that the amount of *surface area* of these branched vessels is most important for determining amount of uptake. The smooth muscle of collecting lymphatic vessels (after the initial) uses contractions to propel lymph towards the lymph nodes. The *diameter* should be measured here, at these wider, less branched vessels. [84]

Lymphangiogenesis is the remodeling of peripheral lymphatic networks during inflammation. [65] Lymphangiogenesis is seen in cancer and inflammatory disease. Hypothetically, the spreading network can allow for more fluid and cell flow. Macrophages secrete growth factors

(VEGF) that lead to sprouting and multiplying of vessels. New vessels have wider diameter but no “button” like junctions, so there may be less pathogen and cell entry. Inflammation reduces lymphatic pumping, and the vessels become more permeable, allowing lymph to leak out. (Excess fluid pressure outside the vessel walls due to leakage could cause propulsion of lymph without needing contractions.) However, it has been shown in [85] that lymph flow velocity is reduced when pumping of vessels is suppressed. Can this decrease in velocity be shown in meningeal lymphatic vessels as well? [84]

The first, along with Louveau et al. [86], to molecularly characterize the meningeal lymphatic vasculature and its function in the drainage of CSF-derived constituents through the cervical lymph nodes, was Aspelund et al. in 2015 [87]. The authors used the reporter mice Prox1-GFP and lymphatic deficient mice K14-sVEGFR3 to demonstrate the importance of the MLV in draining brain-injected macromolecules to the peripheral lymphatic system. [65,87]

CT-lymphography provides three-dimensional imaging of the lymphatic system in humans, allowing for visualization of small lymphatic channels, down to 0.7 or 1.2 nanometers in diameter. This is higher resolution than using other imaging techniques, although CT-scans do expose patients to radiation. [88,89] This 3D imaging helps researchers particularly understand the connection between the meningeal lymphatics and the peripheral lymphatics. It provides the ability to map defective vessels and augment drainage of the lymphatics. [90] As this is a different imaging modality entirely from what is used here (confocal microscopy), the noise and artifacts present in these images will be different.

[91] discusses the vascular permeability of lymphatic endothelial cells. An interesting experiment is performed with young and old human skin and the lymphatic vessels imaged there, complete with information about how the two types of skin exhibit different markers or proteins. [91] It is interesting to note that there is a difference in permeability correlated with age even in the peripheral lymphatic endothelial cells. Perhaps the permeability of the endothelial cells also affects the ability of the MLV to effectively drain fluid, the discussion of which takes place in Section 3.3.2.

3.3.2 Function of the Meningeal Lymphatic Vasculature (MLV)

The meninges communicate to the brain periphery via the lymphatic system, because the meningeal compartment lacks a blood brain barrier. The meningeal lymphatic system is a central regulator of CNS homeostasis. [65] It was first described in 1787 by Paolo Mascagni [92] but was excluded from anatomical textbooks, and only recently re-discovered in 2014 [3].

Meningeal lymphatic vessels maintain proteostasis or proper levels of proteins in brain fluids. The brain doesn’t have its own lymphatic vessels to drain waste, so proteins and waste are transported via the interstitial fluid (ISF) out of the brain parenchyma. The ISF travels along blood vessel walls to reach the CSF and the lymphatic vessels. Waste clearance occurs across the blood brain barrier, through a process called transvascular clearance. However, with age, transvascular clearance is inhibited, and more responsibility falls to the lymphatic vessels. If the lymphatic vessels also narrow, then plaques accumulate. [93]

A seminal paper written by Antoine [3] really elucidates the importance of the meningeal lymphatics. The meningeal lymphatics are present around blood vessels in the meninges. The fluid part of the CSF (cerebrospinal fluid) drains through those blood vessels, but

immune cells present in the CSF, such as T cells and B cells, are drained through the surrounding lymphatic vessels, all the way to the deep cervical lymph nodes. This drainage occurs in steady state. [3]

The central nervous system has immune privilege, meaning that antigens, such as self-reactive T cells, can be introduced to the CNS without causing an allergic reaction. Drainage may also contribute to keeping this peace. [3]

The underlying question then remains: do patients with MS or Alzheimer's exhibit changes in their meningeal lymphatics? If so, could we predict onset or of these diseases or diagnose patients by studying an individual's lymphatic vasculature? [3]

The authors of [86] used a new dissection method to study the meninges using multiphoton imaging, which allows them to demonstrate the extent of CSF uptake by the MLV. [65,86]

The first *in vivo* imaging of the MLV in humans was performed by [94]. The authors developed an MRI approach for visualization of the enrichment of gadolinium within the MLV of primates and humans. The presence of these genuine lymphatic vessels was confirmed using immunohistochemistry. [65,94] As the images of humans are taken using MRI, different image processing techniques may be required for analyzing the MRI images compared to the microscopy images used in this thesis.

Meningeal lymphatic networks (MLV) develop postnatally, unlike other lymphatic networks. Development of the MLV depends on the Vascular Endothelial Growth Factor c - Vascular Endothelial Growth Factor Receptor 3 (VEGFc-VEGFR3) pathway, just as peripheral lymphatic networks do (those outside the meninges.) The MLV network remains dependent on this pathway well after the developmental stage. [65] [95] provides the first description of postnatal development of MLV, and shows that the VEGF_c-VEGFR3 pathway is required for maintenance of the MLV. The authors demonstrate that cranial and spinal lymphatic vessels initiate development postnatally, around day eight, and further grow until day twenty-five. This work also presents the first evidence that the MLV, contrary to the peripheral network, requires constant signaling through the VEGFR3 pathway in order to be maintained [65,95].

A salient discovery that strongly advocates for the research in this thesis, was published in 2018 by [1]. This paper provided the first demonstration of the implication of MLV dysfunction in the context of Alzheimer's disease and aging. With aging, the MLV decrease in function. [65]

Using chemical and physical approaches, the authors demonstrate that the meningeal lymphatic vessels are a key regulator of the overall glymphatic system [65]. In their experiments, [1] injected mice with a vessel-damaging drug, and found that the CSF no longer reached the deep cervical lymph nodes located in the mouse neck [93]. Mesquita et. al. [1] impaired lymphatic vessels in mice in three ways: ablating them by injecting visudyne, surgically ligating them, and using mice that are genetically impaired. These impairment methods do not affect the blood vessels in the meninges. Using all three methods, the authors find a decreased amount of CSF draining through the lymph nodes, i.e., reduced molecular diffusion in the brain parenchyma. This impeded fear memory and spatial learning. They also found transcriptional alterations associated with synaptic plasticity in hippocampal neuronal transmission. [1]

The MLV is capable of uptake of CSF, which was found in studies performed by [1], where the MLV was ablated, resulting in less CSF derived macromolecules found in the cervical

lymph nodes. However, this did not cause buildup in CSF pressure, so unlike peripheral lymphatics, the MLV is not responsible for recycling of interstitial fluids - that amount of fluid would be too large for MLV uptake. [65]

Still, the MLV is definitely a regulator of the glymphatic system, which removes wastes from the CSF into the brain parenchyma. During aging, restoring impaired MLV using VEGFc results in improvement in drainage, glymphatic function (waste or protein removal) and cognitive decline, which was observed in [1]. Amyloid buildup is reduced, and the behavioral effects can be demonstrated in mice. [1] These observations are not fully understood. When MLV are impaired, this might change the composition of the CSF, meaning cell concentrations will be shifted, changing the cells' ability to remove wastes. [65]

The experiments are performed on male and female wild type mice, whose fresh frozen brain sections are fixed and stained with antibodies before imaging. The Alzheimer's mouse models do not have plaques in the meninges, but if their lymphatics are ablated, then the plaques grow in regions similarly to plaques in humans whom they imaged. This is significant: human plaques are due to lymphatic impairment, and current Alzheimer's mouse models do not capture the whole picture. [1]

One interesting discovery to explore is that inducing local lymphangiogenesis may promote drainage and waste clearance [93]. Restoring drainage functionality in aged mice rescued intra-parenchymal CSF recirculation and improved age-induced cognitive decline [65]. Aged mice can be treated with vascular endothelial growth factor C to improve lymphatic drainage and peripheral lymphatic sprouting. The drainage could help performance of current AD antibody treatments. [1]

The images presented in this thesis were captured by the authors of [1]. The green-channel images are meningeal whole-mounts, from the top of a mouse head. All of the cranial MLV images are of the superior sagittal sinus (SSS).

In [2], the implication of the MLV in the CNS immune response is described. The authors use genetic and chemical methods to discover the contribution of the MLV in immune cell recirculation and immune response. It is demonstrated that CCR7 expression by meningeal immune cells is required for recirculation through the MLV. [2, 65]

[96] provides evidence of the extent and function of the MLV at the skull base. The authors show that CSF enters at specific locations at the base of the skull, where the dura mater is thinner. The authors also demonstrate the dorsal and basal lymphatic vessels behave differently during aging, which provides insight into the regional specificity of the MLV. [65, 96] The different behavior is observed phenotypically in the MLV structure, which will be discussed in Section 3.3.3.

3.3.3 Complexity of the MLV

[95] present some complexity measures on the MLV. By measures in this section, we simply refer to measures of some biological quantity, not a measure on a set. The authors mention thinning in areas, and calculate the percentage area of lymphatic vessels over image size, for whole mount images. Discontinuity of vessels is noted, measured by the number of endpoints in a vessel, or the number of connected components. This number of breakpoints or discontinuities may be important to understanding the development or degradation of a network. For calculating all of these measures, including area, accurate segmented boundaries

are necessary. [95]

The two main measures used by [1] for measuring changes in lymphatic vessels are percentage area, or coverage, and average diameter of vessels. The diameter is averaged across one hundred vessels and fifty measurements per vessel. It is unclear if the measurements are taken across the width of the main vessel, or across secondary branches. Different researchers measure vessels, which results in a variation of 15 percent in absolute diameter. This measurement variance is unacceptable, since vessel diameter should be a **fixed**, measurable metric. [1]

In [2], the total length and the number of sprouts are measured for MLV complexity. The number of sprouts is measured relative to the distance from the pineal gland, along the transverse sinus. LYVE-1 percentage of coverage is also measured over the image, but this measure must contain a large amount of noise, as LYVE-1 uptake is heavily present in the image background. Lymphatic ablation is measured by either LYVE-1 area over the total area of the sinus, or the length of the lymphatics compared to the length of the sinus. [2]

The spinal region contains lymphatic vessels; a network of collecting lymphatics is observed at each nerve root between the lumbar and cervical regions of the spinal cord. These are the lymphatic vessels we will analyze in our images of the spinal meninges. Analysis of these spinal lymphatics showcases a permeable lymphatic phenotype, which scales positively with complexity. Less complex junctions are found to be impermeable, such as the zipper-like junctions found along the transverse and superior sagittal sinuses [2, 65, 95, 96].

Measures taken by [96] include the number of branches per millimeter of vessel length. Branches are defined as sprouting tips with length over 100 microns. This measure is taken on whole-mount images only. The count is taken manually, and normalized by the total vessel length in a given image. The other measures taken are the number of zipper-type junctions and the number of button-type junctions, the latter of which are more permeable. The cell boundaries need to be visible to determine the junction type. Junction type is related to the location of the MLV, with zipper-like junctions found in the transverse and superior sagittal sinus (SSS), and button-like junctions found in spinal lymphatics. Lymphatic vessel coverage of the SSS is also measured, by taking the amount of LYVE staining around the SSS. Diameter of the vessels is sampled every 20 microns. Lymphatic regression, a negative phenotype associated with age, is measured by comparing the length of the lymphatic vessel to the length of the SSS, with the confluence of sinuses taken as the origin (center of the junction in a whole-mount image.) [96]

Transcriptomic analysis of the MLV shows that the genes regulating development and maintaining of the vasculature are unique from those of the peripheral lymphatic system. This implies that while the MLV and peripheral lymphatic endothelial cells (LEC) share molecular features, atypical pathways are involved in forming and maintaining the MLV. [2] This is extremely important. Therefore, it is reasonable to conclude that the structure could also be different from regular lymphatic networks, since the genes encoding development and maintenance are different. [65]

During the inflammatory state, there is no change in the dorsal or spinal MLV. However, the cribiform plate MLV expands, similar to peripheral lymphangiogenesis [2, 97]. With age, the dorsal MLV degenerate and the basal MLV undergo hyperplasia [1, 96]. Hyperplasia is enlargement, a morphological change. [65]

Based on this review of what is currently known about the meningeal lymphatics, it

is clear that much remains unknown. The current approaches for studying the MLV are not transferable to clinics and human patients as of yet. Thus, researching the mechanics and molecular dynamics of how the MLV change the meningeal and parenchymal brain compartments is of great necessity. [65] It is clear also to see that structurally, there are great changes in the MLV in different conditions, and that these structural changes result in great molecular changes in the central nervous system, which can either worsen or improve disease prognosis depending on the disease. To study the vascular structure, and in the future hopefully model these dynamic causal effects, accurate segmentation of the MLV is of the utmost importance.

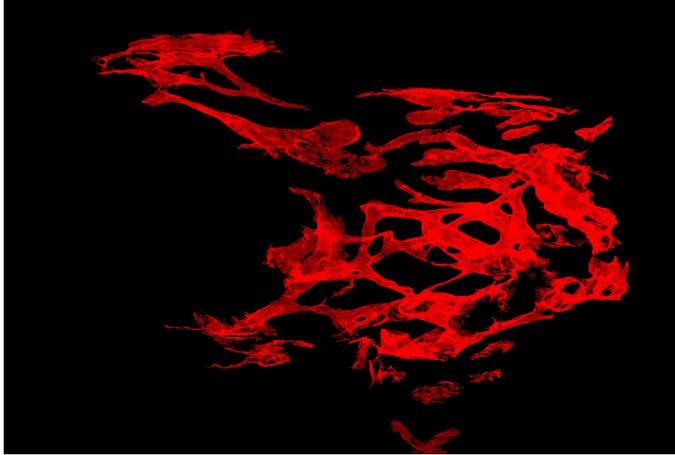


Figure 2: Example of segmentation using LyMPhi on an image of spinal lymphatics after contusion injury. The result shown here is the segmentation binary mask multiplied by the red channel of the original image, to show the full variation of intensity and shadowing in these microscopy images.

4 Level-Set Segmentation

4.1 Background on Segmentation

As our proposed method, called Lymphatic Matted ϕ (LyMPhi), where ϕ represents the signed distance function used to evolve the level set boundary, is fundamentally a level set based method, it is important to describe the intensive history of level set segmentation. The power of this level set segmentation can be seen in Figure 2.

Level set segmentation is a geometric adaptive technique that benefits from the ease of varying the topology of active contours by automatic splitting and merging [98]. This geometric segmentation utilizes a higher-dimensional function ϕ where the zero level set represents the object boundary [99]. The motion of the object contour is performed as the evolution of the zero level set by minimizing the energy functional $\mathcal{E}(\phi)$ [98–100]:

$$\mathcal{E}(\phi) = \mathcal{E}_{external}(\phi) + \mathcal{E}_{internal}(\phi) \quad (1)$$

where $\mathcal{E}_{external}$ represents the energy of the external force that drives the movement of the contour, while $\mathcal{E}_{internal}$ regularize the smoothness of the contour.

Numerous level set segmentation methods have been proposed over recent decades for tubular structures. [101] uses regional kernel for variational level set formulation [102] to maintain the continuity of segmented retinal vessels with variant intensity, where the kernel parameters are manually adjusted for optimal performance. Post-segmentation using a region growing technique is required to recover the missing thin vessels and remove the falsely detected signal. L2S [103] models the inhomogeneity of tubular structures and intra-region illumination variations using Legendre Polynomials. It can preserve the overall structure automatically, but is biased to the initialization of the level set functions. TuFF [104] incorporates the idea of region growing mentioned earlier. Within the curve evolution, an attraction force is proposed to reconnect the disjoint vessel components that are lost using

traditional level set segmentation methods, mainly due to thresholding based initialization. TuFF is a stable work on segmenting tubular structures; however, in the context of segmenting lymphatic vessels, it is challenging to initialize the level set functions to cover all branches which are present throughout the image. Another concern that affects the performance of TuFF on lymphatic vessels is the tradeoff between smoothness of the curve and details of the signals. Because of the variant thickness of vessels, inhomogeneous intensity within the vessels and clutter in the background, smoothing parameters are hard to control, and oversmoothing often results.

The level set methods described above [101–104] require an initial guess, or offset/threshold value for the level set boundary. Most level set algorithms use intensity thresholding, combined with search of specific object scales within the image. Neither of these base methods are substantial especially for meningeal lymphatic segmentation because of the wide variability of intensity and scale within even one single vessel. Reduction of clutter appearing in the segmentation result, using existing techniques, inevitably removes vessel information that cannot always be fully recovered by region growing methods.

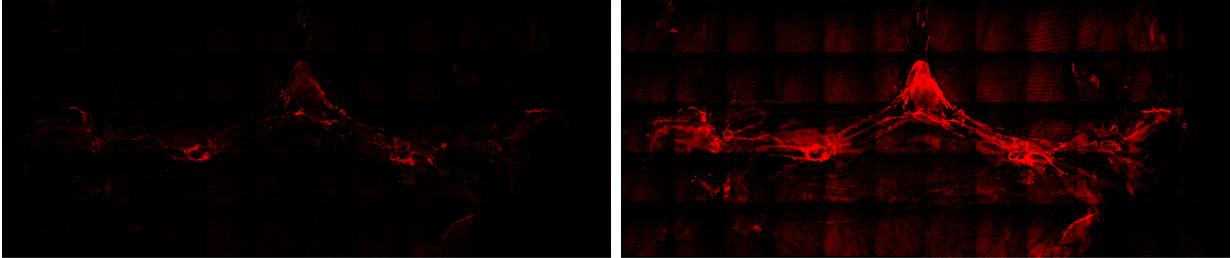
The following are the main roadblocks to successfully segmenting meningeal lymphatic vessel images: the variant thickness of vessels, in-homogeneous intensity within the vessels, noise, and background clutter. The background clutter is perhaps the most challenging aspect; it consists of blood vessels with which the lymphatics vessels are intertwined, as well as other cell types that are also stained. The clutter cannot be removed with thresholding because the vessel staining varies widely in intensity. Existing level set methods do not accommodate these unique challenges of lymphatic vasculature images. In order to remove background clutter, TuFF, for instance, relies on consistent vessel size, which cannot be used in images of lymphatic vasculature. Lymphatic vessel structure varies widely in shape, thickness, and structure. In response to these challenges, we propose a new method which we call Lymphatic Matted ϕ , or LyMPhi to perform image analysis of lymphatic vasculature. LyMPhi utilizes a technique called matting which has been used by others as applied to retinal images [60]. However, to our knowledge, this is the first ever reported work that further develops and uses matting applied to the image analysis of lymphatic vasculature. Furthermore, LyMPhi is the first level-set method powered by matting to reduce clutter in segmentation.

Many of the major segmentation and analysis challenges are shown in Figure 3. Even for a scientist, determining which parts of the image belong to the lymphatic vessel set is difficult. In Figure 3c, there are small endothelial cells at the bottom middle cluttering the image. The boundaries of the vessels are not always well-defined, and are surrounded by noise. Hand annotation, thus, is unreliable as well as time-consuming. Removing these repeating noisy artifacts while accurately segmenting the vessel boundary manually is challenging, as the difference in intensity between the noise and the vessel boundary is small.

4.2 Proposed Method: LyMPhi

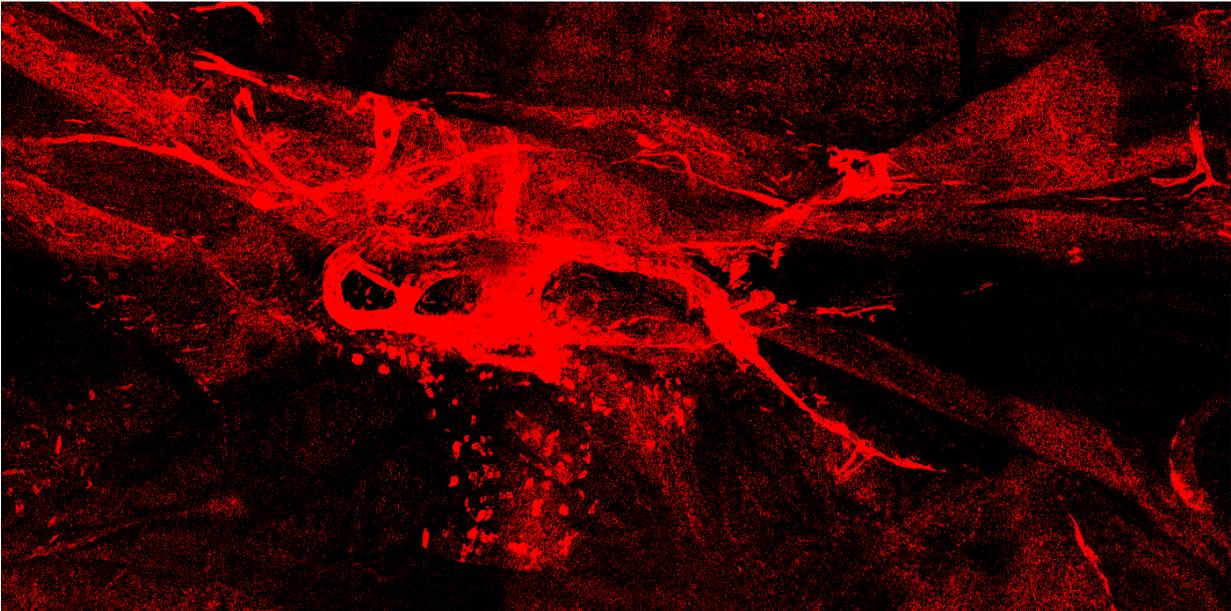
We propose a level set method called LyMPhi. The proposed method is based on the TuFF algorithm, while incorporating hierarchical image matting to remove background clutter.

Level set segmentation provides smoothness of object boundaries, while powerful separation of background and foreground is produced by matting. Since both TuFF and hierar-



(a) Image of meningeal lymphatic vessels.

(b) Contrast enhanced version of the image.



(c) Detail of the contrast enhanced image.

Figure 3: One example image. The original image is shown on the left (a), and the contrast enhanced image is shown on the right (b). The contrast enhancement displays the challenges of noise, artifacts, intensity variation, and background clutter. The final image (c) is a detail image from (b).

chical matting [60] were developed for segmenting vasculature, they are appropriate parent methods to LyMPHi.

4.2.1 Matting

Matting is a technique used to separate foreground pixels from background pixels, which is necessary to remove the background clutter mentioned above. The images of lymphatic vessels share some similarity with retinal fundus images, which capture blood vessels growing on a surface. Since the lymphatic vessels also grow on a surface (blood vessels) which is the main contribution to the clutter, a matting procedure developed for retinal images is adapted for use in LyMPHi.

Fan *et al.* [60] propose the first matting algorithm for segmentation of fundus images. Matting was previously not possible for fundus images because of the amount of time it would take for a user to generate a trimap. The initial trimap generated by the authors is automated, under some size constraints dictated by the known properties of the blood vessels imaged. This initial trimap is then refined using the hierarchical matting procedure, which stratifies unknown pixels into hierarchies based on distance from known pixels and then updates pixel labels in order of hierarchy. Unknown pixels are assigned labels based on their correlation with labeled neighbors. This method has been tested against leading supervised and unsupervised methods for retinal image segmentation, yielding high accuracy. The proposed matting model also has the benefit of being less computationally complex than leading deep learning methods, as well as having higher performance. [60]

This matting algorithm is called hierarchical image matting [60]. It can effectively remove clutter as well as retain small vessel details. The matting procedure is composed of three steps.

Initial Trimap: Image matting begins with an initial trimap, composed of foreground vessel (V), background (B), and unknown (U) pixels. In the original method, the trimap is generated using size criteria on the blood vessels. Since lymphatic vessels vary widely in size and shape, and enough studies have not been performed to set a bound on maximum or minimum vessel width, our method uses an automated approach to initial trimap generation. In LyMPHi, the initialization step is performed using simple Otsu multithresholding for the spinal dataset, to provide two threshold levels: low and high [59]. For the partial and whole mount datasets, the lower threshold is 75, and the upper threshold is 100; out of 255 for a grayscale image (single channel.) The lower threshold is used as the cutoff for B , while the upper threshold is used for demarking V . The pixels with intensity between the two threshold levels fall in the set U .

In [105], Otsu multithresholding is also used as a pre-segmentation step for iris segmentation, which is then optimized using geodesic active contours, as Otsu’s method does not provide the best results if the image histogram is not easily separable. In LyMPHi, Otsu’s method can be used to initialize the hierarchical matting procedure, which will be used to drive level set evolution.

Stratification: The minimum distance from a vessel pixel in V , d_i , is calculated for each pixel in U . The pixels in U are stratified into hierarchies based on this minimum distance.

Unknown pixels with the same distance, d_i are collected into one hierarchy. The hierarchies are ordered based on distance, with the lowest distance hierarchy being assigned label 1. There are m hierarchies, H_1 to H_m . Each hierarchy has the form:

$$H_j = \{u_i^j | i \in 1, 2, \dots, n_i\}. \quad (2)$$

n_i is the number of unknown pixels in each hierarchy that need to be labeled.

Hierarchical Update: The hierarchies are updated in order, from closest distance to V to farthest. In a grid centered on each unknown pixel, correlations are calculated between the unknown pixel and surrounding known pixels. The correlation function is composed of a color cost function and a spatial cost function. The color cost function is defined as

$$\beta_c(u, k) = \|c_u - c_k\| \quad (3)$$

where c_u and c_k are the intensities of the unknown and known pixels, respectively. The intensity is taken in the red channel of the original image. The spatial cost function is

$$\beta_s(u, k) = \frac{\|x_u - x_k\| - x_{min}}{x_{max} - x_{min}} \quad (4)$$

where x_{max} and x_{min} are normalization terms; namely, the maximum and minimum distances between the unknown pixel u and any known pixel k in the surrounding grid. The combined correlation function is

$$\beta(u, k) = \frac{1}{\beta_c(u, k) + \beta_s(u, k)}. \quad (5)$$

The final decision rule for labeling an unknown pixel is described by

$$I_v = \begin{cases} 1 & \text{if } \beta(u, V) > \beta(u, B) \\ 0 & \text{else} \end{cases} \quad (6)$$

I_v representing the mask of vessel pixels. [60]

The matting algorithm provides good separation between background and foreground elements. However, the matting result is not smooth because the method is not iterative and has no smoothness constraints, which is one reason why it is used in combination with level-set segmentation.

4.2.2 Level Set Formulation

The energy functional in 1 can be tuned to the problem at hand. To address the challenge of segmenting lymphatic vessels, we exploit TuFF [104] as the framework of our proposed level set segmentation method. LyMPHi can preserve the vessel-like tubular structure by using a local attraction force (\mathcal{F}_{attr}) to join the discontinuous components:

$$\mathcal{E}(\phi) = \mathcal{E}_{reg}(\phi) + \mathcal{E}_{evolve}(\phi) + \mathcal{E}_{attr}(\phi) \quad (7)$$

Here, inside of the zero level set is denoted as $\phi > 0$, while the outside is $\phi < 0$.

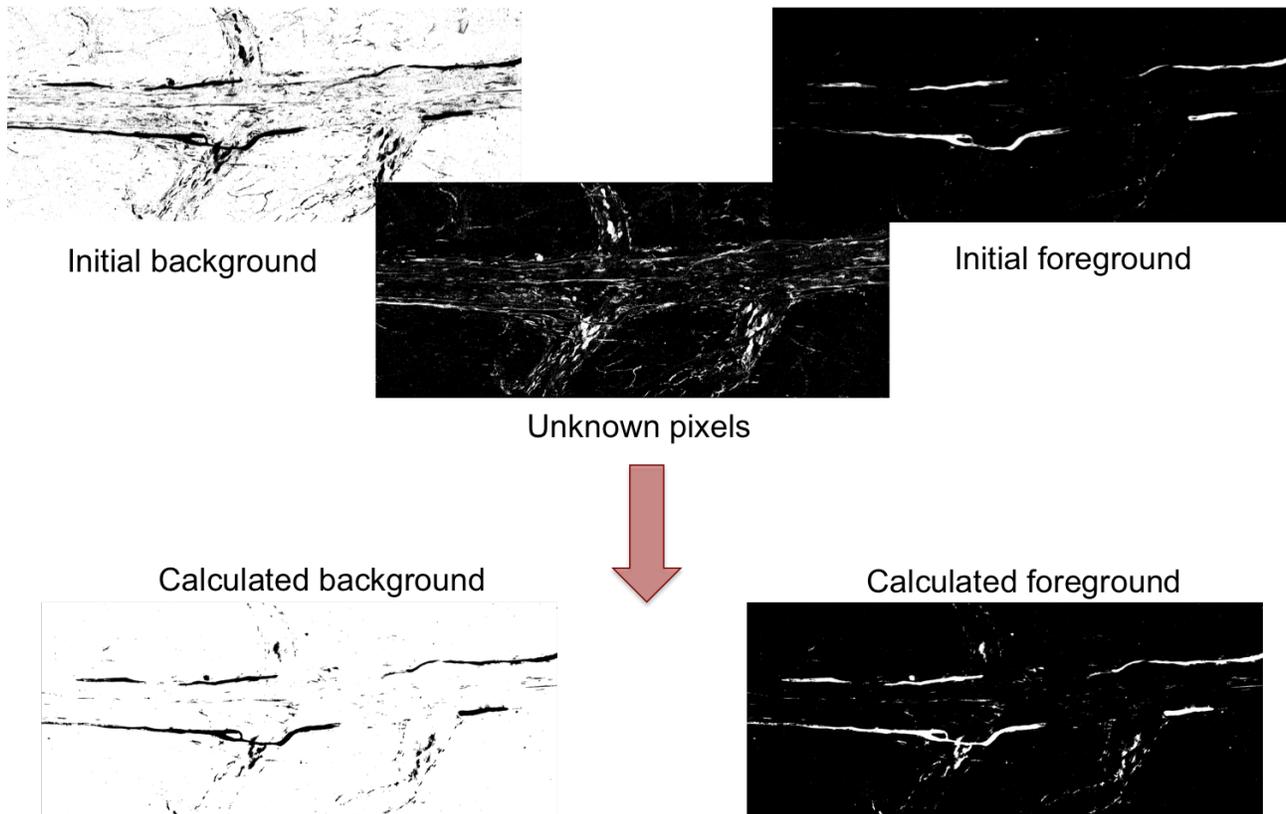


Figure 4: The top shows the initial trimap, consisting of background, unknown pixels, and estimated foreground. Images are binary, white indicating belonging to the set in question, i.e., background, foreground, unknown. After image matting, the calculated background and foreground have converged.

The \mathcal{E}_{reg} term regularizes the length of the zero level set, which controls the smoothness of the contour. \mathcal{E}_{evolve} drives the evolution of the curve where a combination of axial and orthogonal vector field is used to achieve the curve propagation perpendicular to the vessel boundary and along the vessel axis, respectively. The former design fulfills the width of the tubular structure and the latter one traces the growing length along its major axis. \mathcal{E}_{attr} acts for connecting the discontinuous fragments in a way of creating an attraction force field using vector field convolution (VFC) [106]. As a result, two disconnected components will be attracted to connect if the two separate curves lie in the neighboring force field where the forces (\mathcal{F}_{attr}) are opposite and hence mutually attractive.

Minimization of the energy functional is solved using gradient descent, using

$$\frac{\delta\phi}{\delta t} = \mathcal{F}_{reg}(x) + \mathcal{F}_{evolve}(x) + \mathcal{F}_{attr}(x). \quad (8)$$

\mathcal{F}_{reg} and \mathcal{F}_{evolve} are the forces due to their corresponding energy functionals, described above. \mathcal{F}_{reg} , \mathcal{F}_{evolve} , and \mathcal{F}_{attr} are forces, which are defined and applied with respect to the normal direction of the implicit function, $\frac{\partial\phi}{\partial t}$ (as described below.) The \mathcal{F}_{reg} term regularizes the length of the zero level set, which controls the smoothness of the contour. \mathcal{F}_{evolve} drives the evolution of the curve where a combination of axial and orthogonal vector field is used to achieve the curve propagation perpendicular to the vessel boundary and along the vessel axis, respectively. In TuFF, this vector field is created using a edgemap of the vesselness response map, produced by [63], and tuned by the scale parameter determining the thickness of the vessels. As tuning this parameter is difficult for lymphatic vessels, where the vessel scale varies greatly, in LyMPhi, the edgemap is instead computed on the original image.

\mathcal{F}_{attr} acts for connecting the discontinuous fragments in a way of creating an attraction force field using vector field convolution (VFC) [106]. The vector \mathbf{x} is a position in the image. $\mathcal{F}_{reg} = \nu_1 \text{div}[\mathbf{n}(\mathbf{x})]\delta_\epsilon(\phi)$ while

$$\mathcal{F}_{attr}(\mathbf{x}) = \nu_2 \sum_{i=1}^p \sum_{j \neq i}^p \mathcal{F}_{attr}^{(i,j)}(\mathbf{x}), \forall \mathbf{x} \in \Omega. \quad (9)$$

The full expression of \mathcal{F}_{evolve} is not included here due to size constraints. The full forms and derivations of all forces can be found in [104].

For \mathcal{F}_{reg} , ν_1 is a smoothing parameter, $\mathbf{n}(\mathbf{x})$ is the inward normal unit vector to ϕ , and δ_ϵ is the regularized Dirac delta function [107]. For \mathcal{F}_{attr} , ν_2 decides the effect of the attraction force, and p is the number of disjoint connected components that can potentially be attracted to one another. Ω is the image, and d is the dimension of said image. The local attraction force is defined as follows: $\mathcal{F}_{attr}^{(i,j)}(\mathbf{x}) = \kappa_i \langle \Gamma_i(\mathbf{x}), -\mathbf{n}(\mathbf{x}) \rangle \theta_j(\mathbf{x})$. i, j refer to two connected-components. κ_i is the normalized mass of the “child” component. The indicator function $\theta_j(\mathbf{x})$ determines if a child component is within the convex hull of the “parent.” Γ_i is the attraction force field computed via the VFC technique [106].

Level set evolution is an iterative process, where the level set boundary moves closer to the object boundary over many iterations. A novel modification we make to level set curve evolution takes place during the iteration process. We design a force field over the image that is made up of the component forces:

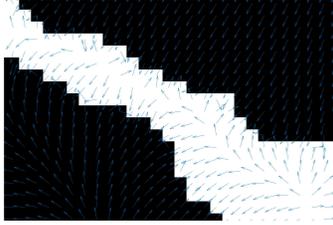


Figure 5: Example force field around vessel. The blue arrows inside the vessel point out towards the vessel boundary, and the arrows outside drive the level set in towards the boundary.

$$F(i, j) = F_{reg}(i, j) + F_{evolve}(i, j) + F_{attr}(i, j). \quad (10)$$

This force field is the velocity of the level set evolution, or $\frac{\partial \phi}{\partial t}$. The velocity at each point (i, j) has a magnitude and a direction. The magnitude is the speed at which we move the zero level set inwards or outwards towards the object boundary (moving the level set elevation up or down), calculated from the component forces. However, we modify the direction of the velocity according to the matting result, as explained below.

4.2.3 Curve Evolution

As the level set evolves, the level set boundary can be pulled towards the foreground pixels and away from the background pixels, to more tightly constrain the final segmentation to the foreground pixels.

At each pixel in the image, there is an overall force determined by the separate forces in 8. This overall force has either positive or negative sign, denoting whether this pixel, at that iteration, is thought to be within the object boundary, or not. LyMPhi modulates this overall force at each pixel location by performing a check with the initial matting result. An example force field is shown in Figure 5, drawn using software from [106].

To realign the level set boundary towards the desired foreground, LyMPhi changes the sign of the signed distance function at each pixel according to whether it is labeled foreground or background by the matting procedure. All calculated foreground pixels now have a positive signed distance to indicate that they lie within the object boundary, whereas all background pixels are given a negative signed distance to show that they lie beyond the boundary. Restraining the level set evolution to calculated vessel foreground enables LyMPhi to more closely adhere to small vessel details, instead of obscuring these details on the quest for overall smoothness, as in other level set methods.

After computing (8), the following is used to drive *foreground propagation*:

$$\frac{\delta \phi}{\delta t}_{i,j} = \begin{cases} \left| \frac{\delta \phi}{\delta t}_{i,j} \right| & \text{if } u_{i,j} \in V \\ - \left| \frac{\delta \phi}{\delta t}_{i,j} \right| & \text{if } u_{i,j} \in B \end{cases} \quad (11)$$

to change the sign of $\frac{\delta \phi}{\delta t}$ at each pixel location (i, j) . u is the pixel at (i, j) and V, B are the foreground and background maps produced by matting. An example is shown in Fig. 6 at

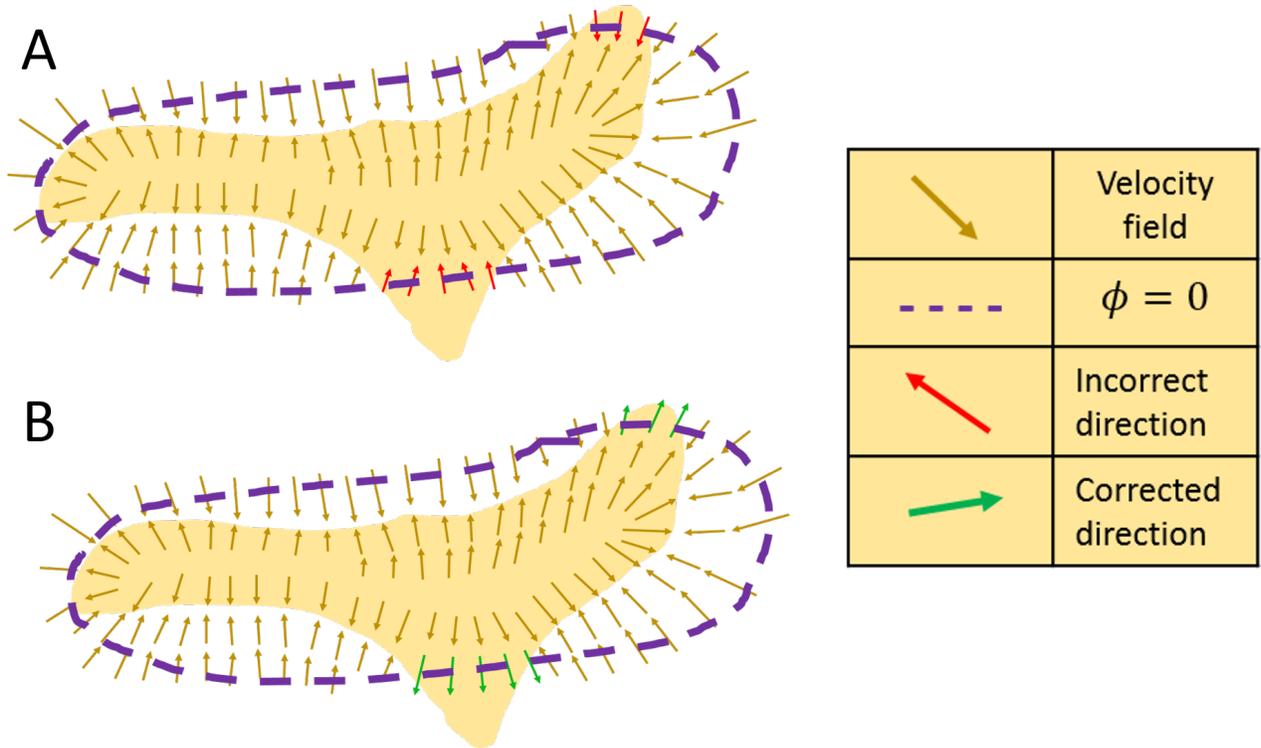


Figure 6: Modifying the velocity field that drives curve evolution. The level set contour is shown with a purple dashed line. The yellow shape represents the foreground calculated by hierarchical matting. In (A), the red arrows along the level set contour are driving curve evolution further towards the vessel interior. In (B), this is corrected using the decision rule 11. The corrected velocity field drives curve evolution at subsequent iterations towards the outer vessel boundary.

iteration t . At time t , the original velocity field, calculated from the TuFF forces, is shown by brown arrows. The yellow shape represents the calculated foreground, V , from hierarchical matting. In (A), the red arrows represent an incorrect direction for curve evolution. These arrows are driving ϕ inside the vessel. These directions are corrected in (B), using the matting label calculated. The now green arrows of the contour evolution speed have been changed in the normal direction.

4.3 Experiments Performed and Results

4.3.1 Description of the Datasets Used

There are three real image datasets used for the experiments performed. All three datasets contain confocal microscopy images taken of lymphatic vessels in mice, stained with LYVE-1. The images were acquired by the Kipnis laboratory at the Center for Brain Immunology and Glia (BIG), then at the University of Virginia School of Medicine, Department of Neuroscience. The current imaging technology produces 2D images, as the lymphatic vessels are laid on a slide and then imaged. In the future, we hope to extend LyMPhi, particularly the

matting algorithm, to 3D images of lymphatics.

There are 39 images in total. Seventeen of these are images of the superior sagittal sinus in the mouse brain, with nine images taken in whole-mount, and eight in partial-mount. An example of the whole-mount sinus is shown in Fig. 8. The images are 16 bits per pixel for the spinal dataset, and 8 bits per pixel for the whole and partial mount datasets. The image size varies within each dataset. The whole-mount images are 2643x948 pixels, with one image having dimensions 2311x948 pixels. The resolution is 0.33 pixels per micron. The partial-mount images vary more widely in dimension, with the minimum width and height being 4631 and 1941 pixels, respectively. The maximum width and height are 4653 and 1993 pixels. Resolution is 1.61 pixels per micron.

The twenty-two remaining images are of lymphatic vessels in the mouse spinal cord post-injury. These images are referred to as the spinal dataset, an example of which is shown in Fig. 7. While these vessels do not originate from the spinal meninges, they are phenotypically related to the meningeal lymphatic vessels. The spinal dataset original images vary the most widely in dimension. The minimum width and height are 1344 and 702 pixels, and the maximum width and height are 4195 and 3114 pixels. The resolution for these images is 0.805 pixels per micron.

It is valuable to have three datasets with different resolutions because different levels of vessel details are captured. The vessel size and structure varies widely for each dataset. The whole-mount images contain the most clutter, as they have the entire structure of the blood vessel network present in the sinus, showing the lymphatic vessels growing on top of and around the blood vessel network (shown in Fig. 8.) The partial-mount images provide finer detail along within the sinus, generally displaying the overall vessel structure at finer detail - which necessarily removes higher level structure, such as branching or looping. In the spinal dataset, lymphatic vessels aggregate in rounder shapes, with a large amount of holes or gaps in the vessels, as shown in Fig. 2. The vessels in the spinal images are overall not as thin or elongated as the meningeal lymphatic vessels shown in the other two datasets.

The driving biological interest in this approach is in segmenting meningeal lymphatics in the sinus, but because of limited data (17 images), testing is also performed on the spinal images. High performance of LyMPHi on the spinal dataset may imply promising application to generalized lymphatic vessel images.

4.3.2 Creation of Manual Annotation

Manual segmentation was performed for comparison with automated segmentation methods. Three operators created separate annotations for each dataset, and the images were merged using majority voting to create one consensus segmentation image for comparison. The vessels are primarily distinguished by brightness or color, and also by shape and relation to surroundings. This manual quantification was performed using the software FIJI (ImageJ), a Java-based image editing software package. The manual annotation is generated by using the “Clear” tool under “Edit,” with “Freehand” selection. The user draws regions around objects in the image that are not considered lymphatic vessel, and these are cleared successively, until only lymphatic vessel content remains in the image.

The important discrepancies were of both interobserver and intraobserver nature; both of which could cause issues with replication. With interobserver discrepancy, there is a poten-

tial for variation in segmentation from different observers, due to the occasional ambiguous definition of a vessel in accordance with its surroundings. This could cause inconsistencies in the resulting cropped image as the criteria (mostly pixel group brightness) for a vessel could be interpreted differently. Even the same person will create slightly different manual segmentation for the same image if created at separate times. The time taken to perform hand segmentation ranged from roughly 30 minutes to an hour; therefore, there is a high potential for fatigue, and subsequently, error. This was the main cause for intraobserver discrepancy; usually a day or two after originally segmenting images, we reprocessed the images, due to the fact that some clutter was originally missed. Annotating in the vicinity of small holes and gaps in the vessel presents difficulty; zooming into these smaller regions means that the intensity variations between vessel and non-vessel are less easily detectable by the human eye, due to noise. Often, these holes or gaps are overlooked or improperly manually segmented. All of the above problems attest to the fact that hand-segmentation is far less than ideally efficient as the state-of-the-art method for lymphatic vessel quantification; a better, more accurate way is needed if the data is to be analyzed in larger quantities. We thus add a synthetic dataset to our testing data, discussed in Section 5, to provide unimpeachable ground truth.

4.3.3 Comparative Methods Tested

Four methods are compared on the datasets described above.

Simple Hierarchical Image Matting: This is a matting based segmentation method developed for blood vessels in fundus images, as described in Section 4.2.1. Initialization is performed using Otsu multithresholding, and the grid size used for hierarchical update is 9x9, as in the original paper [60]. We use our own implementation of the method, based on the equations given in the referenced paper.

L2S: The earliest method tested is L2S, a Legendre polynomial based active contour segmentation method, developed to deal with intensity inhomogeneity along object boundaries [103]. The upper bound on iterations was set to 1,000 (the default.) For this method, initial contours are set by the user using an elliptical boundary surrounding each vessel. The user also must indicate how many separate vessels to segment per initial contour. There are several initial contours for each image, because the vessels are found in different image regions.

Simple TuFF: This method is Tubularity Flow Field, a level set method developed for neuron segmentation, as discussed above in Section 4.1. TuFF is run for 80 iterations per image (the default.) The scales parameter is set to 10 for these images, which defines the width of the Gaussians used for the Frangi vessel enhancement. This choice is used in order to enhance the most prominent vessels in the images. All other TuFF parameters are used in their default settings. Simple morphological post-processing is used to remove small objects from the final segmentation.

4.3.4 Measuring Segmentation Accuracy

Each result was compared to either the manual majority voted annotation or the generated ground truth using the Sørensen-Dice coefficient to measure accuracy [108]. This measures

the pixel overlap between the annotation and the experimental results (or for the synthetic case, ground truth vs. experimental), and combines evaluation metrics such as sensitivity, specificity, and accuracy that are also often used for evaluation segmentation results [109]. The results are shown in Table 1, with the Sørensen-Dice coefficient averaged over each dataset per method. The Dice coefficient has a range from zero to one, with one signifying a perfect match.

The Dice coefficient is complemented by the Boundary F1-score, a measure computed from precision and recall on the segmented boundary [110]. This is denoted by *BF-score* in Table 1. The Boundary F1-score also ranges from zero to one, one meaning a perfect boundary match.

The BF (Boundary F1) contour matching score is based on precision and recall on the boundary only, developed for semantic segmentation in supervised learning, not just foreground/background segmentation. They show that the BF-score is complementary with the Jaccard index, or Dice coefficient. One limitation of Jaccard Index (JI) is that it evaluates the amount of pixels correctly labelled, but not necessarily how accurate the segmentation boundaries are. Therefore, JI alone is not sufficient to compare different segmentation methods. [110]

For some applications, the contour quality greatly contributes to the perceived segmentation quality. Popular contour-based measures for segmentation include the Berkeley contour matching score [111], which computes the F1-measure from precision and recall values with a distance error tolerance θ to decide whether a boundary point has a match or not. The BF-score is modeled off the Berkeley contour matching score, but extended for multi-class. [110]

Given two boundary maps, of the predicted segmentation and the ground truth: for each pixel in the predicted segmentation (for precision), search the ground truth within a distance to see if you can find a matching positive pixel in the manual segmentation. Basically, build a window (surrounding Euclidean distances 19 pixels, depending on image size, away) and search within that window in the manual annotation. Precision is over all positive pixels in the prediction; Recall is iterated over every pixel in the hand segmentation, and over all positive pixels in the ground truth (and the search is in the prediction). [110]

Using Spearman’s correlation coefficient, the authors observe low correlations between BF and other measures (for example JI). The aim is to create a measure to complement the Jaccard index, which is standard for segmentation. [110] As the measures are complementary, this dissertation considers mainly JI and BF.

From a human study conducted by [110]: we can see that the JI is more correlated with the human ranking than the BF. This is consistent with an observation made during the study: accurate contours are less important in the first place than having the right categories. However, for relatively good and similar segmentations, the BF becomes more relevant to rank segmentations. Have also shown that the proposed BF score is complementary with JI as it more carefully takes the contours into account, and both measures should be considered simultaneously. [110]

The BF-scores are also averaged over each dataset. Standard deviations are provided in Table 1.

Table 1: Average Dice Coefficient and BF-score for All Methods

Method	TuFF		L2S		Mat.		LyMPhi	
Metric	<i>Dice</i>	<i>BF-score</i>	<i>Dice</i>	<i>BF-score</i>	<i>Dice</i>	<i>BF-score</i>	<i>Dice</i>	<i>BF-score</i>
<i>Partial</i>	0.31±0.15	0.43±0.12	0.55±0.18	0.58±0.10	0.58±0.09	0.60±0.07	0.71±0.09	0.84±0.07
<i>Whole</i>	0.22±0.10	0.14±0.15	0.35±0.13	0.48±0.17	0.46±0.10	0.74±0.07	0.54±0.09	0.86±0.05
<i>Spinal</i>	0.39±0.23	0.35±0.27	0.59±0.27	0.62±0.28	0.61±0.15	0.79±0.18	0.67±0.10	0.79±0.14
<i>Synthetic</i>	0.06±0.04	0.15±0.08	0.58±0.25	0.79±0.20	0.81±0.13	0.79±0.12	0.89±0.09	0.96±0.05

Table 2: Median Dice Coeff. for All Methods

Method	TuFF	L2S	Mat.	LyMPhi
<i>Partial</i>	0.30	0.59	0.59	0.74
<i>Whole</i>	0.20	0.34	0.48	0.56
<i>Spinal</i>	0.39	0.67	0.65	0.70
<i>Synthetic</i>	0.05	0.56	.91	0.91

4.3.5 Results

Three level-set methods, including LyMPhi, are compared on the datasets described above. Each method is initialized by thresholding a small amount of noise and background clutter, in order to retain the majority of the foreground. For TuFF, the scale parameter is set to 10 for these images, which defines the width of the Gaussian functions used for the Frangi vessel enhancement [63]. This is the best choice for scale enhancement in these images, in order to enhance the most prominent vessels. The other level-set method tested is L2S. All three level-set methods are run for 200 iterations. Hierarchical image matting is tested as described in Section 4.2.1. The initial trimap is found using Otsu multithresholding [59].

The methods described above in Section 4.3.3 are run on each of the four datasets. Including creation of the matting initialization, LyMPhi run-time is approximately 2 hours for a spinal lymphatics image with size 1500x2700 pixels, on a standard Windows x86 desktop with 8GB RAM. This is similar to the computing time for the other level-set methods tested.

For all four datasets, LyMPhi has the maximum average and median Dice coefficient measured from all four methods, as shown in Tables 1 and 2. The median Dice coefficients are also shown as these values are less affected by outliers, such as images with overall weaker staining. TuFF performs well when vessel scales do not vary widely. However, TuFF performs poorly when the scale is highly variable: this method retains too little of the vessel in the segmentation result. As lymphatic vasculature varies so widely in size compared to neuron branches, the scale enhancement used in TuFF tends to discount larger more blob-like vessel regions. If a larger scale parameter is used, thinner vessel structures are omitted from the segmentation result. L2S, when initialized finely, performs better, as on the spinal dataset. However, if initialization is coarser (and thus inaccurate), L2S retains much of the

background clutter. Mean BF-scores are also shown in Table 1. Again, LyMPHi outperforms or has equivalent performance to the other methods tested. LyMPHi’s BF scores are close to 1.

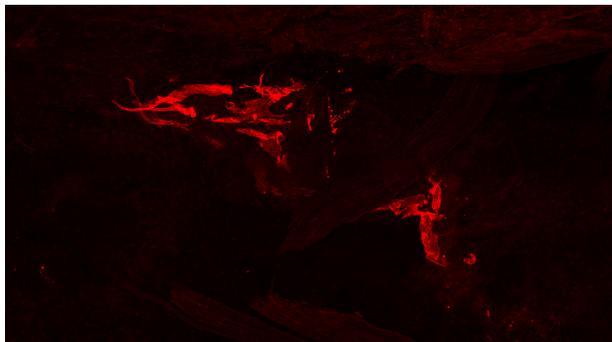
Additionally, the standard deviation of Dice coefficient and BF-score is presented in Table 1. LyMPHi has the lowest standard deviation for both Dice coefficient and BF-score. Paired with the average scores reported in Table 1, this shows that LyMPHi is the most consistently high-performing method across all datasets. For the spinal dataset, where hierarchical image matting also has a high BF-score, LyMPHi by comparison has a slightly lower standard deviation, meaning that the performance on the spinal dataset in terms of BF-score is better by LyMPHi than hierarchical image matting. One reason why hierarchical image matting performs well on the spinal dataset is because, in the spinal images, the vessel mass is larger than in meningeal images. This allows for more information when comparing unknown pixels to the higher intensity foreground - increasing the number of connected known foreground pixels improves the labeling of the unknown pixels. It is important to note, however, that the overall vessel boundary in the segmentation is not smooth as a result of image matting, and appears pixelated, even with these high resolution images. Additionally, although hierarchical image matting partially provides thin branches that are ignored by other methods, this inclusion comes at the expense of including irrelevant cell types.

The vessel size extracted by LyMPHi most closely matches the manual annotation. TuFF results in an undersegmentation, or thinning of vessels, and L2S frequently oversegments, not adhering to the outer vessel boundary. Matting, while more correctly extracting vessel size, includes extraneous cellular matter, which obfuscates the measured amount of lymphatic vasculature. Thus, LyMPHi, due to its incorporation of matting, is the most robust method of the three for detecting thinning of meningeal lymphatic vessels, a precursor to Alzheimer’s disease. Sample segmentation results are shown in Fig. 7. It can be seen that LyMPHi has captured less of the background clutter while preserving the thickness of the vessels. In Fig. 7, it is clear from the displayed results that LyMPHi has retained less of the clutter present at the top and bottom right of the image. LyMPHi has the most consistent results (maximum Dice coefficient) across all datasets.

The results are poorest, for all methods, on the whole-mount dataset (out of the real datasets) due to poor channel separation. In these images, the lymphatic vessels are present in the green channel, but the background clutter of blood vessels and endothelial cells also have strong representation in the green channel, as shown in Fig. 8. This leads to poor segmentation by all methods. However, LyMPHi still outperforms the others in this challenging circumstance.

LyMPHi has the most consistent results (highest Dice coefficient and BF-score) across all datasets. The segmented vessels are smooth and do not contain large gaps due to improper thresholding. The attraction force is also able to join some disjoint vessel components.

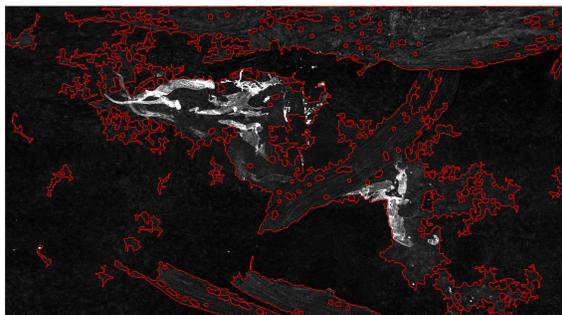
It must be noted that the manual segmentation being compared to is only a best approximation of the vessel boundaries. There is a fair amount of error involved in hand segmentation, as discussed in Section 4.3.2, which are certainly affecting the Dice coefficient calculations. Further analysis on the reliability of manual annotation will be provided in Section 7. It is possible that the LyMPHi segmentation result is closer to the *true* segmentation than our manual annotation. Sample segmentation results are shown in Figure 9. It can be seen that the LyMPHi segmentation result (second from the bottom) has captured



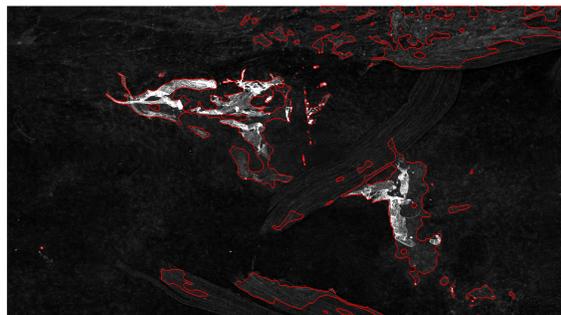
(a) Original Image



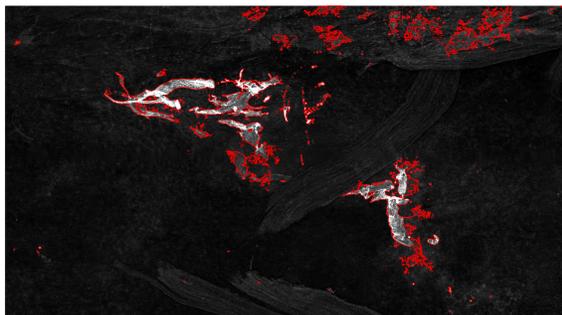
(b) Manual Segmentation



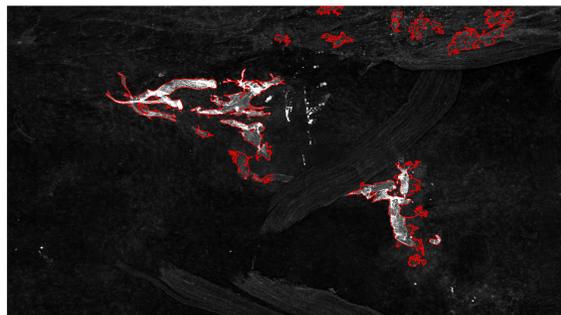
(c) TuFF: Dice = 0.16



(d) L2S: Dice = 0.38

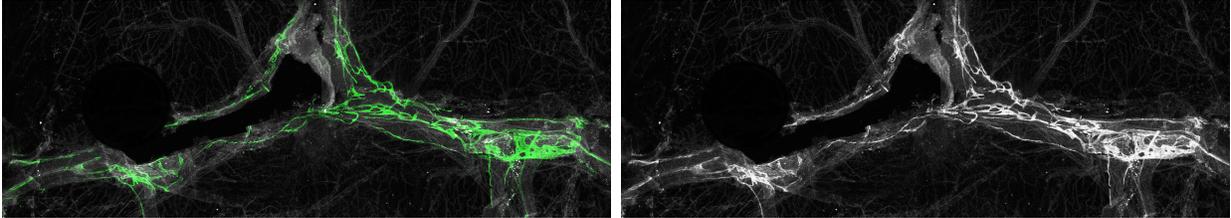


(e) Matting: Dice = 0.48



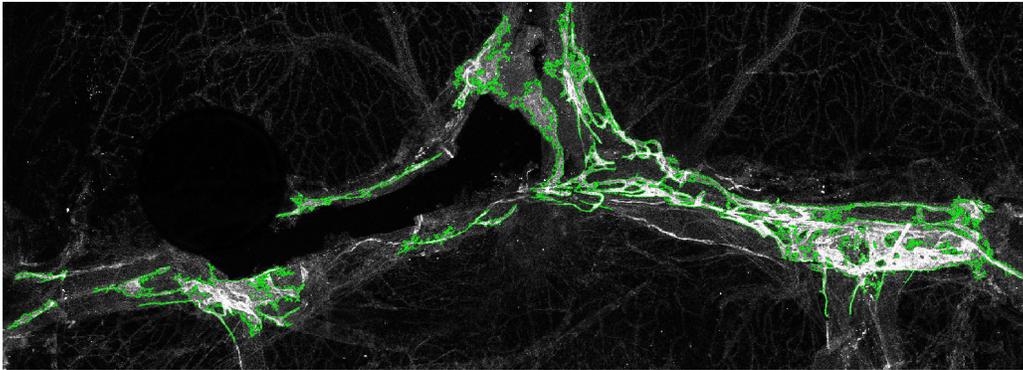
(f) LyMPhi: Dice = 0.73

Figure 7: Segmentation results on one image from the spinal dataset. The results are displayed as gray-scale image data and overlays with the final segmentation contours in red. Captioned is the Dice score



(a) Original Image

(b) Green Channel



(c) LyMPHi Result: Dice = 0.69

Figure 8: Segmentation results on one image from the whole-mount dataset. The results are displayed as gray-scale image data and overlays with the final segmentation contours in green.

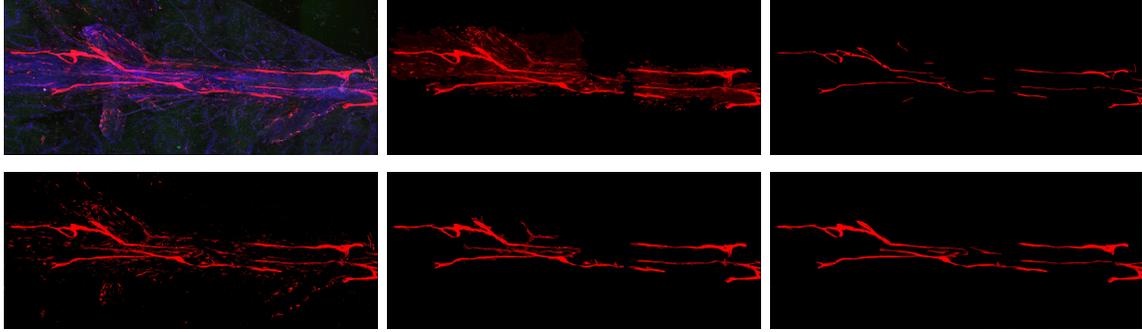


Figure 9: Segmentation results on one image from the Partial-Mount dataset. Results are shown as binary segmentation masked with the original red channel, to show distinction between lighter intensity background and brighter foreground that are included in the result. Top image is the original image, with meningeal lymphatics shown in red. Moving from left to right, top to bottom, the results are: L2S, TuFF, Matting, LyMPhi, and Manual Annotation.

PRECISIONS

Dataset/Method	L2S	TuFF	Matting	LyMPhi
Partial Mount	0.41	0.29	0.43	0.79
Whole Mount	0.40	0.14	0.56	0.75
Spinal	0.55	0.28	0.73	0.81
Synthetic	0.61	0.08	0.85	0.88

Figure 10: LyMPhi provides the highest precision values for all datasets. This precision result means there is a higher proportion of relevant boundary pixels returned by LyMPhi than by the other methods. This performance is explained by the other methods returning too many positive pixels, as evidenced by the recall numbers in Fig. 11.

less of the background clutter while preserving the thickness of the vessels better than the competing methods.

4.4 Analysis

4.4.1 Precision and Recall on the Boundary

When analyzing why the BF-score for LyMPhi is higher compared to the other methods, it is observed that LyMPhi has the highest precision values on the boundary for all datasets. This means there is a higher proportion of relevant boundary pixels returned by LyMPhi, because the other methods return too many positive pixels. These results are shown in Fig. 10 and Fig. 11. Sometimes precision is more important than recall, which could be argued in this case. Precision is important for getting more accurate complexity measurements. Including

RESULTS - RECALLS

Dataset/Method	L2S	TuFF	Matting	LyMPhi
Partial Mount	0.99	0.94	0.99	0.90
Whole Mount	0.63	0.35	0.99	0.85
Spinal	0.80	0.57	0.95	0.70
Synthetic	0.98	0.99	0.99	0.99

Figure 11: Looking at the recall values, the answer is: yes, the other methods do return too many positive pixels. If you mark every pixel as positive, by sheer luck, a higher percentage will match with the ground truth. LyMPhi has lower recall in some cases than the other methods, meaning LyMPhi is a more conservative method.

too much background matter in the lymphatic segmentation could grossly overestimate the volume of lymphatics present.

4.4.2 Per-Image Scores

Measures computed over the whole dataset do not enable distinguishing an algorithm that delivers a medium score on all images from an algorithm that performs very well on some images and very poorly on others. Plotting the histogram of per-image scores enables making such a distinction. Two such histograms have been shown in Fig. 12. The histogram of Dice scores calculated after using LyMPhi is heavier on the upper end of scores, so there are fewer low or medium scores when using LyMPhi as compared to Hierarchical Image Matting. Per-image scores reduce the bias w.r.t. large objects, as missing or incorrectly segmented small objects have a lower impact on the global confusion matrix. Average per-image scores, however, can be used for overall assessment, as shown in Section 4.3.5. [110]

The set of per-image scores to evaluate the percentage of images with a performance higher than a threshold, to compare the percentage of images where one method performs better than another or to analyze the statistical difference of two segmentation algorithms with t-test. When using a paired t-test on Dice coefficients from using LyMPhi and Hierarchical Image Matting, at the $\alpha = 0.05$ significance level, Dice coefficients using LyMPhi are higher than when using Hierarchical Image Matting. Using per-image scores allow comparison to a threshold which can be useful in real applications where the user expects a minimum level of quality.

4.4.3 Stability of LyMPhi

One important facet of LyMPhi is its stability and robustness to various starting points due to initialization. This robustness cannot be offered by other level-set methods tested on the lymphatic data. We have found that initialization plays an enormous part in the final accuracy of the results found by using L2S or TuFF on our data, and this initialization is

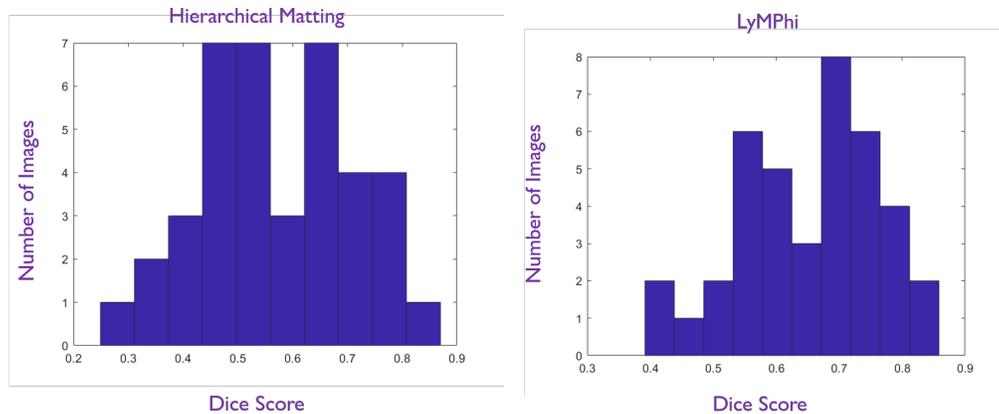


Figure 12: Segmentation results on all microscopy images. Results are shown as histograms of Dice coefficients. Left graph contains the results from Matting, right shows the results from LyMPhi.

difficult to tune from image to image. The staining varies in each vessel sample, leading to variable signal strength. Using too low an initial threshold retains too much clutter, while using too high a threshold removes part of the vessel that is lower in intensity. LyMPhi can segment lymphatic vessels even with a coarser initialization, which may be necessary to retain signal across a dataset.

The matting procedure has already produced a vessel foreground that has removed much of the background clutter, so the first iterations can quickly move the level set boundary close to the calculated vessel foreground. This is done by changing the force field, or velocity, of ϕ . Subsequent iterations smooth the zero level set contour.

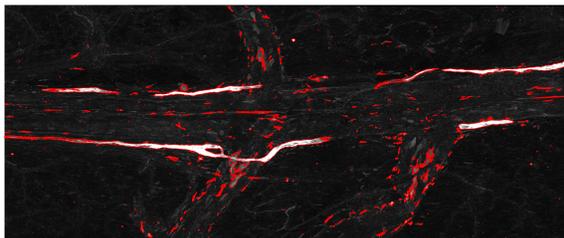
Two variants of initialization are shown in Fig. 13, with their respective segmentation results using TuFF and LyMPhi.

The TuFF result using the finer initialization indeed only captures a few elements of clutter. But the level set boundary tends to overall creep inside the true vessel boundary. This could be due to intensity inhomogeneity and lower signal strength at the vessel boundary, as well as difficulty tuning the vessel scale parameter. If the scale parameter is made larger to prevent such undersegmentation, this will lead to missing thinner vessel segments in the segmentation result. The TuFF result using coarser initialization picks up much more background clutter, as expected. The level set boundary has extended at more points outside the vessel boundary, leading to oversegmentation.

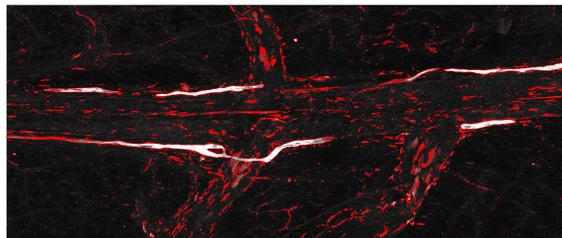
The LyMPhi result using the finer initialization still picks up some background clutter, with a bit more clutter retained when using the coarser initialization. However, the main advantage of LyMPhi is that it captures the true vessel width which TuFF cannot. Width is a measure of significant interest to neuroscientists, as they use it to quantify lymphatic presence.

4.4.4 Isolating the Effect of Foreground Propagation

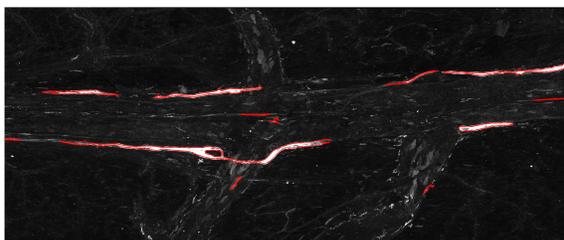
One might wonder whether simply using matting to initialize a level-set algorithm, and then separately running the level-set method without further consideration to the matting



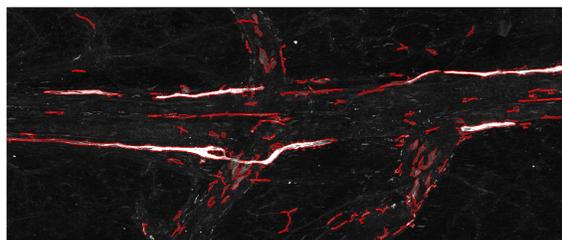
(a) Fine Initialization



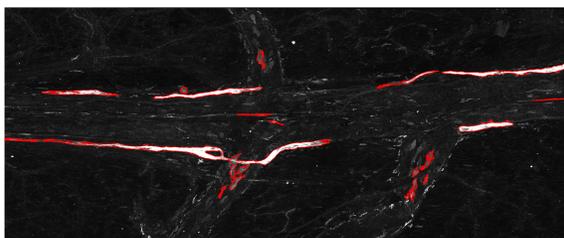
(b) Coarse Initialization



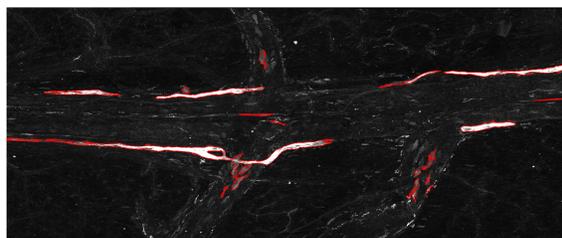
(c) "Fine" TuFF: $D = 0.74$



(d) "Coarse" TuFF: $D = 0.55$



(e) "Fine" LyMPhi: $D = 0.83$



(f) "Coarse" LyMPhi: $D = 0.80$

Figure 13: The first row shows two initial segmentation images of a partial mount image: (a) is fine-grained initialization, and (b) is a rougher thresholding keeping much of the background clutter. The second row shows two TuFF results after 200 iterations, using each initialization. The third row shows two LyMPhi results after 200 iterations. The respective Dice coefficient (denoted with "D") is displayed underneath each result. There is a marked decrease in Dice coefficient of almost 0.2 depending on the initialization used, whereas LyMPhi changes in Dice coefficient by only 0.03 even when the initialization becomes much coarser.



Figure 14: Close-up of vessel detail. From the left: original image, segmentation using matting to initialize TuFF, segmentation using LyMPhi. The loop encircled in yellow is more accurately captured by LyMPhi - the pipelined approach distorts the shape of the vascular loop. Zoom from original image is approximately 30X to show fine detail. Again, intensity variation is retained in the segmentation result to enable visualization of shadowing and detail.

Table 3: Change in Dice Coefficient Due to Foreground Propagation

Dataset	Matting+TuFF	LyMPhi
Partial-Mount	0.64	0.71
Whole-Mount	0.32	0.54
Spinal	0.62	0.67

result will produce similar accuracy as using our intertwined approach, LyMPhi. For this reason, we isolate the effects of foreground propagation, by comparing the segmentation results between using matting plus TuFF as a simple pipeline and LyMPhi. The results show that manipulating the sign on the level-set to pull the contour towards previously marked foreground does prevent over-smoothing of level-set segmentation. LyMPhi is able to capture more details than the pipelined approach, an example of which is shown in Figure 14.

In Table 3, comparisons of the average Dice coefficient with and without foreground propagation are recorded. Adding the foreground propagation to manipulate the sign on the signed distance function does positively impact the accuracy of the segmentation, as shown by the higher Dice coefficient produced by LyMPhi.

4.5 Discussion

The main contributions of Objective 1 can be summarized as follows. First, a matting based level-set segmentation approach, LyMPhi, is proposed to obtain robust and accurate meningeal lymphatic vessel segmentation by removing background clutter and retaining both smoothness and complexity of the vessels. Second, LyMPhi is fully automated, unlike the current hand segmentation used by neuroscientists, or traditional level-set or matting segmentation procedures, which may require difficult to tune user-defined intensity and scale

thresholds. Third, extensive experiments are conducted on three types of lymphatic vessel datasets to validate the performance of LyMPHi using Dice’s coefficient compared to other state-of-the-art segmentation algorithms.

As shown with the Whole-Mount dataset, LyMPHi struggles to segment images with poorly separated channels, or images where the staining intensity is low overall. A correction stage could be placed before the segmentation algorithm, to enhance channel separation and overall intensity, leading to stable segmentation results even when input image quality varies. This could aid in extending LyMPHi to segmentation of other vascular images with high clutter and varying vessel width, such as blood vessel images. Additionally, entropy-based or saliency-based thresholding can be used to determine the initial trimap, in addition to learning more information from neuroscientists about what sizes and shapes the vasculature reasonably maintain.

To improve the segmentation performance further, any number of denoising approaches could be applied prior to using LyMPHi. Anisotropic diffusion [44, 112, 113], or other approaches used for Gaussian and non-Gaussian noise could be used for smoothing. Some of the methods listed in Section 2.2.3, such as NL-means or TV-based denoising specifically for Poisson noise corrupted microscopy images, may work well to mitigate the effects of noise on vessel segmentation.

LyMPHi may not be only limited to use on meningeal lymphatic vessels. Other datasets that include vasculature, such as neuron datasets, or blood vessel datasets, could benefit from the clutter avoiding properties of LyMPHi. Even road networks in satellite images, provided the noise is not too obscuring, could be segmented using LyMPHi.

As has been discussed in Section 4.3.2, the manual annotation used may lose some intricacies in the vasculature due to the laborious process of hand-labeling. In Chapter 5, we will discuss a novel way to create synthetic data that will be used for testing, which intrinsically provides a solid ground truth for comparison.

5 Data Augmentation

5.1 Background on Data Augmentation

One of the central problems of using machine learning in animal brain imaging in regards to analysis is the size of the data set. Most images are hard to come by because each image represents perhaps hours of experimentation and often, an animal being sacrificed. Among data augmentation methods, the following are worth mentioning in this dissertation: (1) Generative Adversarial Networks [114] and (2) u-Net elastic deformation [115].

5.1.1 Generative Adversarial Networks

A first option for data augmentation involves designing a custom generative adversarial network (GAN) to produce images indistinguishable from the original data set. This is a highly popular method in recent years for data augmentation. In this approach, two networks are used to generate and subsequently evaluate image data. The first network is a generative network, typically deconvolutional. The second is a discriminative network, and is conversely convolutional. The generative network creates data from the training data and the discriminative network classifies input data as synthesized or true (from the original data distribution.) Using backpropagation, the generator learns how to create more realistic images to increase the classification error of the discriminator. [116]

Generative Adversarial Networks are often used towards evaluating the robustness of CNNs. The generator creates attacks for realistic looking data with some indistinguishable flaws, to fool the discriminator. [117] There are also *deepfakes*, in images and video, which are again distorted image or video content generated by deep learning algorithms. Research on detecting adversarial attacks and deepfakes is experiencing high current activity. [118]

Cycle-consistent adversarial networks learn the mapping between two images, to apply the correct amount of deformation to one image, in terms of textures and colors, in order to make that image look like something else. For example, the cyclic GAN in [119] learns the mapping between a zebra and a horse, and can “paint” an image of a horse to look like a zebra. [120] uses [119] to add snow to summer scenes and remove snow from winter scenes in remote sensing imagery. This learning of mapping could be useful for modeling degeneration of vasculature, for example.

Of particular interest within cyclic-GANs to biological segmentation problems are the creation of conditional adversarial networks. These networks learn the relationship between a set of labels and a real image, complete with textures and intensity variation. The established relationship is extremely useful, as any data that is augmented needs to include ground truth to be of future use for training or comparison.

One such network, using cascaded refinement networks, learns the how to color objects and features in a city landscape based on labeled data from vehicular cameras. [121] We have tried using this trained network on our MLV data, but found that the network weights are too attuned to the Cityscapes dataset used, so the 256x256 pixel output generated by the GAN highly resembles the images of buildings and vehicles.

A conditional adversarial network named pix2pix was also trained on the Cityscapes dataset and other datasets, and has spawned many similar networks trained on a variety

of datasets and at different resolving power. The original pix2pix network offers very low resolution, smaller generated images. Interestingly, the pix2pix generator network utilizes the u-net architecture [115, 122]. [123]

Conditional GANs are also used for the method called pix2pixhd [122], which is based on pix2pix but offers higher resolution output, with generated image size up to 2048x1024 pixels, thus including finer details. With pix2pixhd, labels for objects can be interactively changed by users in real-time, and one set of labels can generate different images by adding different textures and colors. Object instance segmentation is used to enable changing an object category, such as a tree to a building, and even adding or removing objects. The network also boasts a new adversarial loss function, as well as new multiscale generator and discriminator architectures. [122]

The motivation for conditional GANs can be to create virtual environments, and also training data, as in our case. Existing data can be segmented, the labels can be modified, and new images generated from the changed labels. In the case of pix2pixhd, input to the generator can even be in the form of an edgemap. Pix2pixhd proposes a coarse-to-fine generator and multiscale discriminator architectures to achieve high resolution output. The authors' adversarial loss is proven better than using L_1 loss at learning the image to image pairing. Pix2pixhd was not the first GAN to generate larger images, but due to its advanced objective function, more realistic images are possible. [122]

The generator network from the original pix2pix is split into two sub-networks: G1, a global generator, and G2, a local enhancer network, with larger output. G2 takes input from G1 to integrate global and local information. Both G1 and G2 are trained separately and then fine-tuned together. Three scale discriminators prevent over-fitting and over-loading memory. First, the generated images are downsampled into pyramids, and then each discriminator network handles a different scale. If scales are not incorporated, the generated images have many repeating patterns. [122]

The improved loss function mentioned earlier is a feature matching loss, which extracts features from multiple scales and matches these (learns the mapping) from real to synthetic features. The synthetic features must be statistically realistic at various scales. The feature matching loss is combined with standard GAN loss. An instance boundary map, or edge map, is used to separate objects that are of the same class, such as different types of trees. This prevents bleeding of objects into one another within the generated images; even if all objects are people or cars, they need to remain distinct. This also allows for object based image manipulation instead of only global texture changes, the first GAN to have this functionality. During the training stage, the networks use features from the image and labels to generate the original input image, which enables G1 and G2 learn which features are important for realistic reconstruction. The encoder learns features from the ground truth, which are the real city traffic images. [122]

The features are clustered within each category using k-means clustering (for example, asphalt vs. cobblestone road texture). A cluster center is selected to use these new features in modifying an image. The quality of generated results is assessed using segmentation accuracy of labeled objects, in terms of pixel accuracy and mean intersection over union (IOU). Essentially, if the objects can still be accurately segmented according to their original ground truth labels, the generated image is of high quality. [122]

However, when attempting to test the network, it was simply too large for our available

hardware (Nvidia Titan X and Nvidia Titan XP x2.) Pix2pixhd was developed by researchers from NVIDIA, which means they have access to the largest GPUs. The pix2pixhd network has so many sub-networks, with even more layers and weights, that even testing the existing trained model is not possible given our hardware, let alone trying to train the model on our data. This is the case even when attempting to generate the smallest image size, 256x256. For training, thousands of images are used - which begs the question: are there even enough MLV images to train a conditional GAN, in order to generate more data? [122]

Based on the pix2pix network, [124] is trained on retinal images. However, when using this network on the MLV, even though the retinal images are closer to our vasculature than the Cityscapes dataset, there are still too many differences for the generated data to look like the MLV images. We also do not have enough MLV images for retraining the network.

5.1.2 Elastic Deformation

For this thesis, deformation methods for data generation are more of interest, as having a morphing model for lymphatic vessels can have more uses than just creating additional data (see Section 5.5.2).

In the u-net paper, elastic deformation is used to stretch and bend images of cells into new, biologically plausible shapes - thus creating more data. The authors build a convolutional neural network for segmenting closed loop structures, i.e., cells, in biological microscopy images. The second major contribution the authors make is to use their own elastic deformation model to create new training data from the limited data available. The deformations are created by randomly displacing pixel intensities along a 3x3 window. The interior pixel displacements are then interpolated. Using the augmented data and the trained u-net, segmentation results are vastly improved compared to other competing methods. [115]

Elastic deformation of shapes is a much explored area of interest within image processing. Similar object morphing approaches are used throughout the image processing literature.

5.2 Shape Analysis

The fundamental drawback of using elastic deformation as described in the u-net paper is that the deviations take place in Euclidean space. These deformations are not guaranteed to be meaningful, i.e., have any similarity to real biological shapes. To ensure that the deformations we use to generate new data are biologically sound, we move to deformations along a new space, called the *shape space*.

Shape analysis can be as simple as modeling cell shape with a known shape model, such as an ellipsoidal cell shape model [125], to as complicated as shapes that have yet to be described. There is even learning of shapes being performed using dynamic Graph CNNs [126].

[127] defines 3D shapes of proteins from cryo-EM images. Instead of comparing shapes by their main chain or backbone sequence, the authors compare based on surface shapes, because this is where protein interaction occurs. Proteins with very different backbones can have highly similar surface shapes. 3D Zernike Descriptors (3DZD) are used to represent the shapes, with order $n = 20$ as the order of the descriptor [128]. These are mathematical moment based invariants of 3D functions - they are rotation invariant. This represents

the 3D shape by a weighted combination of 3D basis functions. 3DZD also provides the Euclidean distance between shapes; two shapes are similar if this distance is below 10. The authors compare the 3DZD distance to the Procrustes distance and the TM-score [129] - other distances to measure similarity between shapes. [127]

The resulting shape description is mapped to 3D space using PCA, with 3 eigenvectors. All shapes can be plotted using PCA to view clusters of shapes. For single-chain proteins (and complexes), the shape is most defined by *eccentricity*. The eccentricity is approximated by finding the minimum volume enclosing ellipsoid (MVEE) for each protein. Protein complexes are more spherical than single-chains, as measured by eccentricity. [127]

In the future, this work could allow for taking a protein “census.” Types of proteins can be counted in organisms and compared to other organisms. Simulations can also be designed using this newfound knowledge of specific shapes. [127]

In [115], deformation is used to stretch and bend images of cells into new, biologically plausible shapes - thus creating more data. The authors build a convolutional neural network for segmenting closed loop structures, i.e., cells, in biological microscopy images. The second major contribution the authors make is to use their own elastic deformation model to create new training data from the limited data available. The deformations are created by randomly displacing pixel intensities along a 3x3 window. The interior pixel displacements are then interpolated. Using the augmented data and the trained u-net, segmentation results are vastly improved compared to other competing methods.

The fundamental drawback of using deformation as described in the u-net paper is that the deviations take place in Euclidean space. These deformations are not guaranteed to be meaningful, i.e., have any similarity to real biological shapes. To ensure that the deformations we use to generate new data are biologically sound, we move to deformations along the *shape space*. In [130], the authors use this idea of shape spaces to show that shapes reside on high dimensional manifolds. Manifolds can be built on a variety of objects, such as images of faces. For example, recent research has shown that searching a face manifold of high resolution images that have a low resolution image similar to a low resolution input image, can aid in upsampling the input image [131]. Between different shapes on the shape manifold, a geodesic distance can be taken, which is the path along the curves of the manifold between the two shapes. Sampling points along this geodesic showcases the evolution of one shape to another shape.

An example of this is shown in trajectory using lymphatic vessel data is shown in Figure 15. The sample shapes that lie in between are deformations of the original shapes. [132] We propose to build a large lymphatic dataset by sampling these deformations found between the true vessel data available.

[133] introduces the square root velocity (SRV) transform for shapes to remove translation, scaling, and rotation discrepancies. Their preshape space includes all the rotations and scaled versions of a shape on the shape space. The authors also introduce a path straightening approach to find the geodesics between shapes on the shape space. This method straightens a path between two shapes into a geodesic path. We propose using the square root velocity function (SRVF) to transform shapes of lymphatic vessels into new shapes, as a method for data augmentation.

Prior to computing the SRV, unit length curves are enforced on closed curves (which begin and end at the same location) to remove scaling effects. To preserve orientation, or

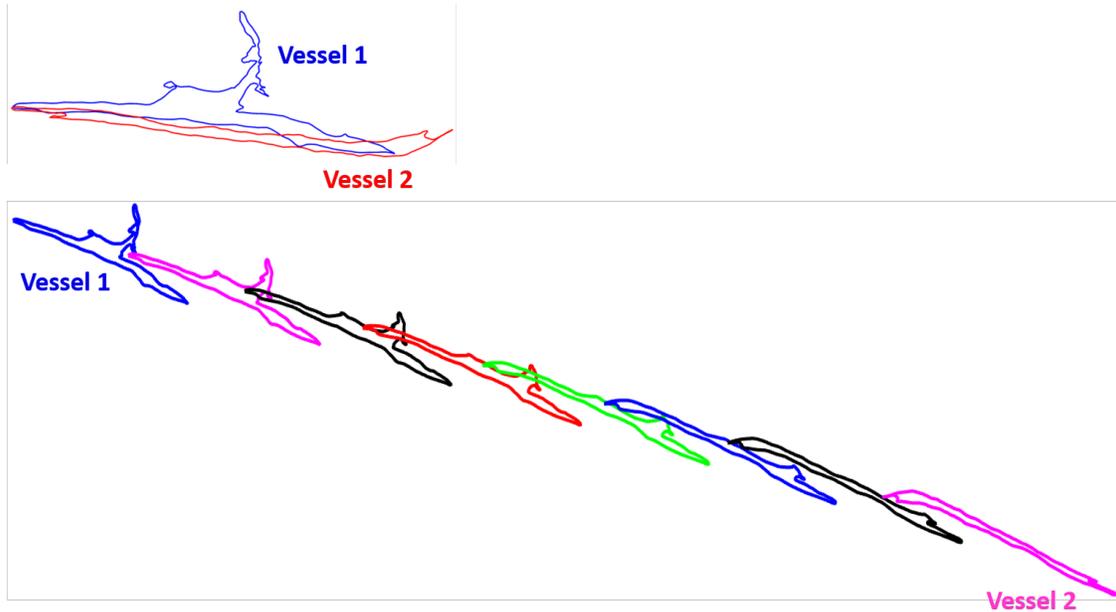


Figure 15: Showing the evolution of one vessel shape (Vessel 1) to another shape (Vessel 2) along a geodesic path, i.e., morphing between one shape to another.

remove rotation effects, reparameterization of shapes is done based on a circle. The sampled points along the curve surface are compared to a circle's points which also start and end at the same place. The SRV is the only shape transform of its type where we can still use Euclidean coordinates to describe the transformed shape. The SRV supports elastic deformation, unlike previous methods, such as [130], which only include bending, and not stretching energy. The SRV transform also enables transfer of deformation from one shape to another shape. The SRV transform will be more extensively discussed in the following section, where it will be used to create synthetic vessel images of the meningeal lymphatics.

5.3 Synthetic Vessel Data

We test segmentation performance on a synthetic dataset of 100 images. We choose to test using synthetic data in addition to real data, because there are errors in the hand annotation, due to not removing the many holes present in the vessels. Thus, only testing with real data means that instead of comparing our results to ground truth, we are comparing the similarity to another flawed measurement. Our method, LyMPhi, may in fact capture more accurately the true shape of lymphatic vasculature. Therefore, we test all methods on a synthetic dataset of our own creation, where ground truth perfectly captures the true vessels so that any comparisons are more meaningful.

There is also a limited amount of data for analysis purposes. Data collection is expensive in three ways: time, money, and animal life - as each image of meningeal lymphatics requires the sacrifice of one mouse. In this section, we propose a method for creating synthetic vessel data so that everyone who wishes to study the meningeal lymphatics can have access to these data.

We build our own deformation model for lymphatic vessels, allowing us to morph the vessels into reasonable formations. All of our deformations can be run by neuroscientists to ensure they are in line with what changes the vessels naturally undergo. Elastic deformation of shapes is a much explored area of interest within image processing. Similar object morphing approaches are used throughout the image processing literature, as discussed earlier.

The following workflow explains how the synthetic images were created. To begin, new vessel shapes must be generated. Elastic deformation is used to stretch existing vessel shapes into new vessels [133].

5.3.1 Vessel Shape Generation

First, from the manual segmentation of partial mount images, isolate each individual vessel segment. The manual segmentation must be used in this step, to prevent bias from any automated segmentation methods. Then, use the square root velocity function (SRVF or SRV, interchangeably) to represent each shape, thus accounting for rotation, shift, and scale variance [133]. These are just the outer contours.

To use the SRVF, first, the shape boundary is sampled in x, y -space. Three hundred sample points are used to achieve a smooth shape boundary. Using fewer sample points will not allow for a smooth boundary, but using more than three hundred sample points will greatly slow the overall shape generation process without much benefit in terms of smoothness. These sample points, denoted β , are locations, which are transformed by

$$q(t) = \frac{\beta'(t)}{\sqrt{\|\beta'(t)\|}} \tag{12}$$

the SRV equation [133, 134]. $\beta'(t)$ is the gradient at each point $\beta(t)$ on the original curve.

Prior to using SRV, unit length curves are enforced on closed curves (which begin and end at the same location) to remove scaling effects. Each location on the shape is a function of the parameter $t \in [0, 1]$, with $t = 0$ and $t = 1$ being the beginning and end of the curve. The process is illustrated in the schematic shown in Fig. 16. After the SRVF has been applied, the new points on the curve, $q(t)$, have been transformed from $p(t)$ into the shape space, where elastic bending, stretching, and shrinking is possible. SRV is a shape transform where we can still use Euclidean coordinates to describe the transformed shape. SRV supports elastic deformation, unlike previous methods, such as [130], which only include bending, and not stretching energy. The same deformation process is repeated to represent any holes, or capillary loops in a vessel.

From any two original vessel shapes, morph along a shape geodesic to find new interpolated vessel shapes. The interpolated vessels are stretched and bent versions of the original vessels, i.e., intermediate deformations [134]. Essentially, sample the geodesic path to get closed curve shapes, representing the vessel exterior. The sampling can be performed finely or sparsely to create any number of new shapes.

Because the manifold containing SRV transformed shapes is locally Euclidean, affine transformations are possible. Shape interpolation is performed by taking convex combinations of q : $\alpha q_1 + (1 - \alpha)q_2$, where q_1 and q_2 are two SRV transformed shapes, and $\alpha \in [0, 1]$ (intermediate algorithmic time-steps.) q_1 and q_2 must have the same dimensions. For a

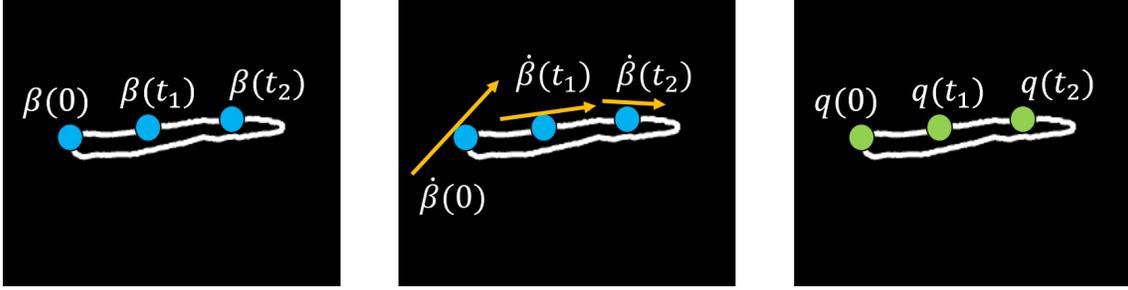


Figure 16: A representation of the SRVF computed at locations on the closed vessel boundary shown in white. The function allows the curve to stretch elastically into another shape. On the left, the original points are labeled $\beta(t)$, starting at $t = 0$ at one end of the shape and ending at $t = 1$ after traversing the whole shape. In the middle, the gradient, or tangent slope, at each point is calculated. The rightmost shape has the labeled points $q(t)$ after the SRV function, 12, has been utilized.

two-dimensional shape, the dimension of each q is $n \times 2$, where n is the number of samples taken to characterize the shape (in our case, 300.)

The SRV transform also enables transport of deformation from one shape to another shape. Using deformation transport [133], the hole shapes can be morphed with the same amount of stretching and bending as the interpolated vessel (as used for the exterior.) Essentially, we use the same bending and stretching energy on a new shape. Given an SRV-transformed shape q_1 and an intermediate deformation to a second known SRV-transformed shape, the shooting velocity v_1 can be calculated. This v_1 signifies which direction and how far to travel along the manifold before reaching the desired deformation. “Transport” means finding the parallel translation of v_1 (call it v_2) for morphing the new shape – in this case, a vessel hole. Fig. 17 shows this intuitively.

Let $[q_1^a]$ and $[q_1^b]$ be the shapes of a lymphatic vessel, at two points along a geodesic path. $[q_1^a]$ is the original vessel shape in SRV-space (initial time τ_a), and $[q_1^b]$ is some arbitrary deformation of the original shape. We can think of this deformation having occurred after a certain amount of time, say after time τ_b . This contour deformation depends on the geometry of the lymphatic vessel. Now, take a different object, a vessel hole from the original vessel shape, which is similar to the lymphatic vessel, but obviously not identical geometrically. Given its initial shape $[q_2^a]$, prior to any deformation, we wish to predict its shape $[q_2^b]$ after the same amount of time, τ_b as for the exterior lymphatic vessel. We thus take the deformation that deformed $[q_1^a]$ to $[q_1^b]$ and then apply this deformation to $[q_2^a]$, using the following:

1. Let $\alpha_1(\tau)$ be a geodesic between the shapes $[q_1^a]$ and $[q_1^b]$ in the SRVF transformed shape space and $v_1 \equiv \dot{\alpha}_1(\tau_a)$ be its initial velocity. (q_1 and q_2 represent two different shapes, and a or b mean a deformation.)
2. Using forward parallel translation, we *transport* v_1 to $[q_2^a]$. Let $\alpha_{1-2}(\tau)$ be a geodesic from $[q_1^a]$ to $[q_2^a]$ in the same shape space. Construct a vector field ω_τ such that $\omega_0 = v_1$ and $\frac{D\omega}{d\tau} = 0$ for all the points along α_{1-2} . Please see [133] for additional details. Fig. 17 shows the relationship between q_1 and q_2 for more clarity.

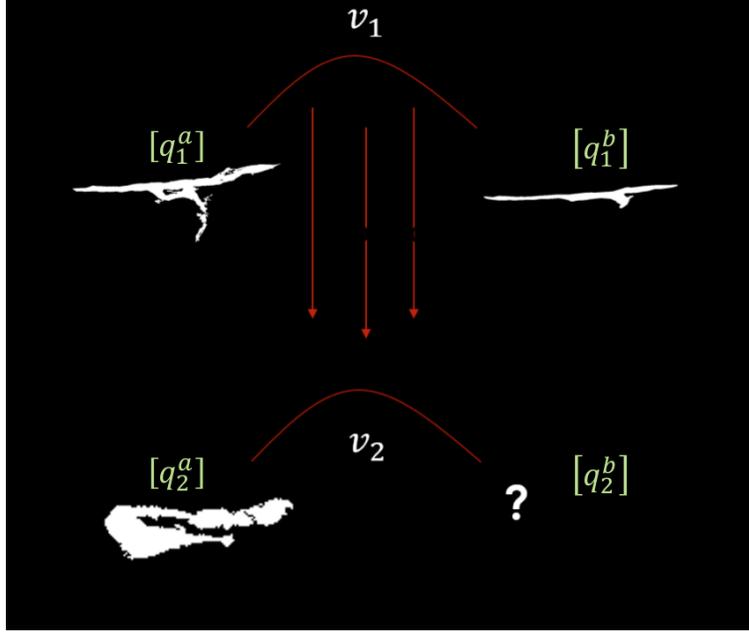


Figure 17: If v_1 is the velocity used to reach the intermediate deformation, v_2 is the velocity to apply the same amount of deformation on the new shape, in this case, a hole. The lower left panel is the large hole within the vessel in the upper right panel. Since shape spaces are overall nonlinear manifolds, the deformations of one shape cannot simply be applied to another. The manifold is only linear when the shape is of the same type.

3. Then, $v_2 \equiv \omega(1)$ is a parallel translation of v_1 .
4. Using v_2 as the initial velocity, form a geodesic starting at $[q_2^a]$, which at time-step τ_b will end in deformation $[q_2^b]$.

The original work [133] on deformation transport is for applications where viewing angles of objects change. Here, we substitute the change in “viewing angle” for a moment in time, either before or after deformation. This is the first use of the theory for modulating the interior of a an object, since a shape necessarily only retains the boundary.

A new, complete vessel shape includes an interpolated shape plus similarly deformed holes placed proportionally (semi-randomly) according to vessel size within the interpolated vessel. A vessel skeleton is generated using the methodology in [135], and the radius of the vessel at each skeleton point is calculated, using size-constrained inscribed spheres [136]. Considering the original size of the hole prior to deformation, the point within the vessel is found where the vessel diameter best matches the original hole size, and the deformed hole is placed at the best diameter match. This ensures the deformed hole fits within the boundary of deformed vessel. Placing the holes based on the original hole location unfortunately does not scale as the vessel may shrink at certain points to where the hole will no longer fit.

New (final) binary vessel images are produced at this stage. An example is shown in Fig. 18. Bioinformatics can be computed on these new shapes. The full diagram illustrating the shape interpolation process is shown in Fig. 19.

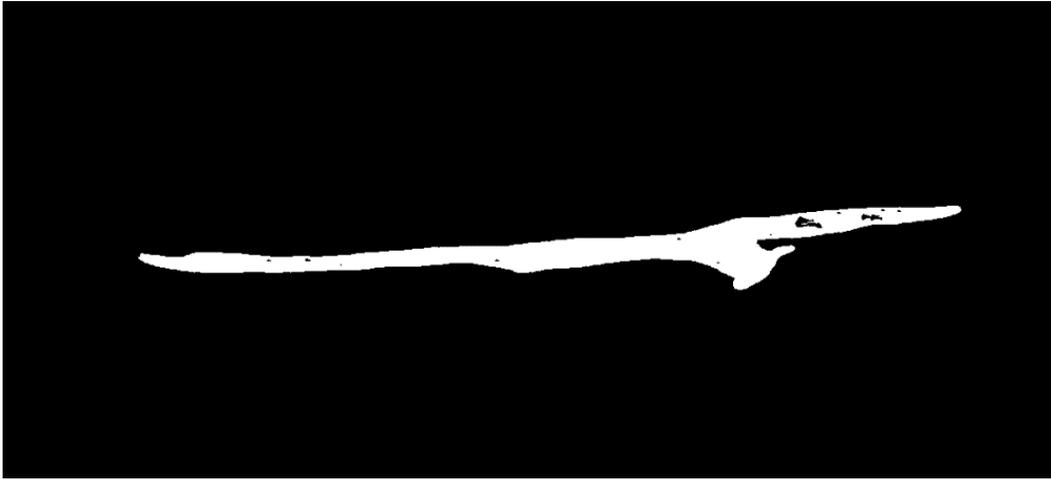


Figure 18: Output from the shape deformation process is an individual binary vessel, with fitted deformed holes.

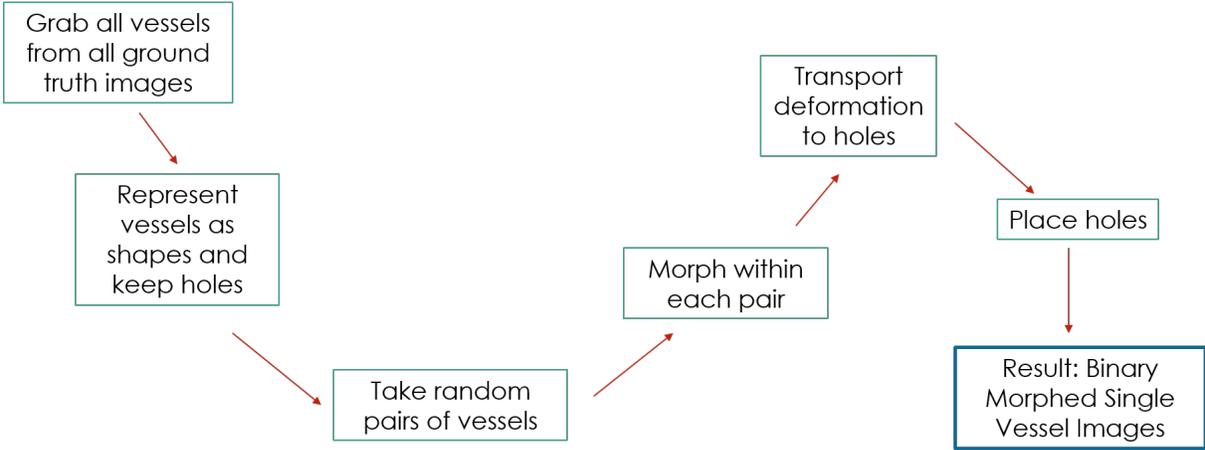


Figure 19: Shape interpolation – from an image to a shape. Streamlined, generalizable workflow from images to new, complete vessel shapes.

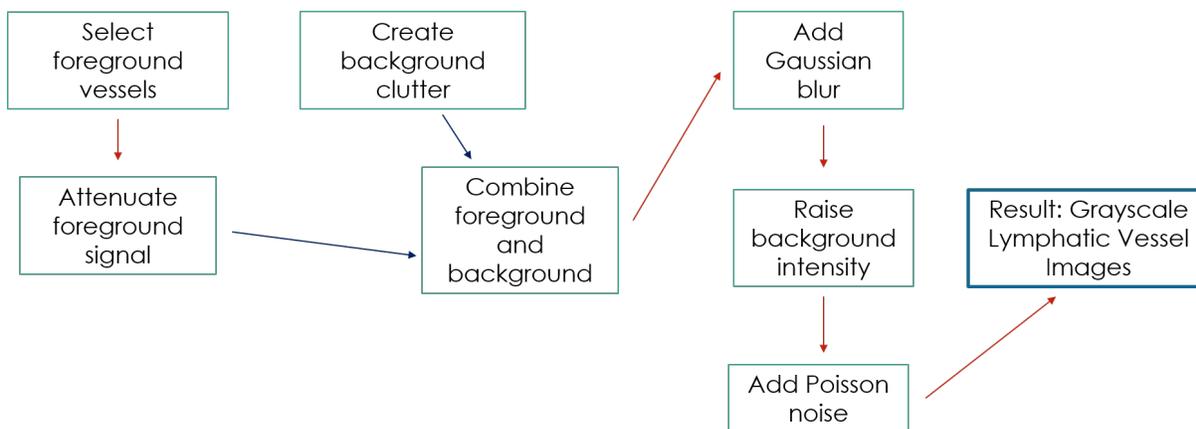


Figure 20: Generating the full synthetic images. Streamlined, generalizable workflow from morphed vessels to realistic, complete vessel images.

In summary, we can loop through combinations of vessels - to create a new vessel, morph between two vessel shapes on the shape space. The two original vessel shapes are picked from all possible vessel pairs. For example, in a dataset of 18 images, where there are 451 original vessels, the total number of vessel pairs is 101,475! To create a full meningeal lymphatic image, combinations of these new vessel shapes are used. As the total number of possible interpolated vessels is incredibly large (there are multiple interpolations possible for each of those 101,465 pairs), and new vessel shapes can be endlessly combined in different selections or groupings, truly, the amount of synthetic data that can be generated through this process is nearly infinite.

There is a possibility for using shape deformation to simulate lymphangiogenesis, thinning of vessels, or lymphatic regression, which is the shortening of the lymphatic vessels along the sinus with age. This could be formulated similarly to neurodegeneration, which has already been shown in [134].

5.3.2 Synthetic Image Formation

A flowchart of the image generation process is shown in Fig. 20. As both the shape generation and image generation processes are generalizable, other areas where synthetic data would be of benefit could use these procedures to augment existing datasets.

Selections of vessel shapes are used to create foreground and background layers of a new synthetic image. Each layer is 4649x1967 pixels, which is also the size of the final synthetic images. This is a similar size to the images found in the partial mount dataset. Select a random number of vessels to include (based on real images.) These vessels are the generated vessels from the shape deformation discussed previously. The images containing these random assortments of vessels are layers for our final image. First, there are some steps needed to produce the foreground layers of the final images.

The foreground layer is used as ground truth, before being convolved with a Gaussian filter to attenuate the signal at vessel edges. This mimics the true images, and is the reason why simple thresholding will not solve the segmentation problem. Set the intensity of the vessels in this layer to the maximum intensity, at 255. The maximum intensity vessel image

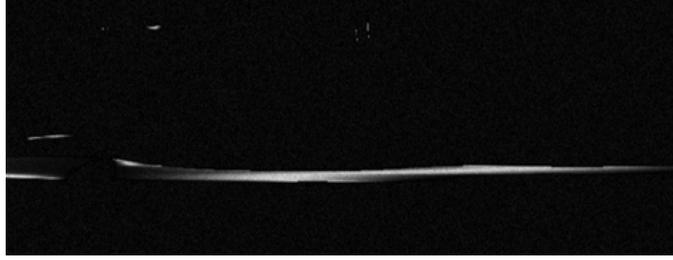


Figure 21: A trimmed Gaussian filter is used over the foreground image layer to attenuate signal intensity. This mimics the true variation in intensity in the microscopy images.

is used as the ground truth.

Then, a trimmed Gaussian filter is applied over each row in the image where the vessel exists to attenuate the intensity of the vessel at boundaries. This mimics the true images, and is the reason why simple thresholding will not solve the segmentation problem. The result is shown in Fig. 21.

For each vessel in the foreground layer, the rows in the layer where the vessel exists must be identified. Then, for each of those identified rows, take the length (the number of pixels in that row that belong to vessel foreground) as n . A 1D Gaussian filter of size $2m + 1$, where $m = n$ if n is even, and $m = n + 1$ if n is odd, is used to attenuate the vessel at each row. Centered at the center pixel in the vessel row, this longer Gaussian filter gives the trimmed Gaussian effect.

Next, on the background layers, which simulate clutter, some additional perturbations need to be added. Create three background layers with the generated vessels but set these intensities at 20 percent of the possible higher foreground intensity (*5 to 1 signal to clutter ratio*), so that there is some overlap between the foreground and background within the lower intensities. The background layers represent the clutter of blood vessels in the original images.

We then combine the foreground and background layers together. The following are adding blur and noise to the combined image. Gaussian blur and Poisson noise [137] are added to complete the synthetic images. Gaussian blur is added with sigma between 0.2 and 0.5 (range is normalized.) The static background intensity is raised up to 19 out of 255, which is similar to the background levels in real data. This is done to make sure that when Poisson noise is applied to the image, the background also becomes noisy - if the background intensity is zero, the Poisson noise in the background will have no effect. Poisson noise is based on the signal intensity in the image.

We use the Poisson noise model on the synthetic data because in confocal microscopy using fluorescence emission to stain biological specimens, Poisson noise is a significant variable. The Poisson noise model is used to reflect the small number and extreme variation of detected photons. [137]

Comparisons in terms of noise are shown in Fig. 22. The noise levels are not exactly the same, but are similar in level and spread given the variations in noise across the many images in the three microscopy datasets. In the histograms, the frequency count for the microscopy region is higher because the region is slightly larger. The intensity in the histograms has been scaled from 0 to 1.

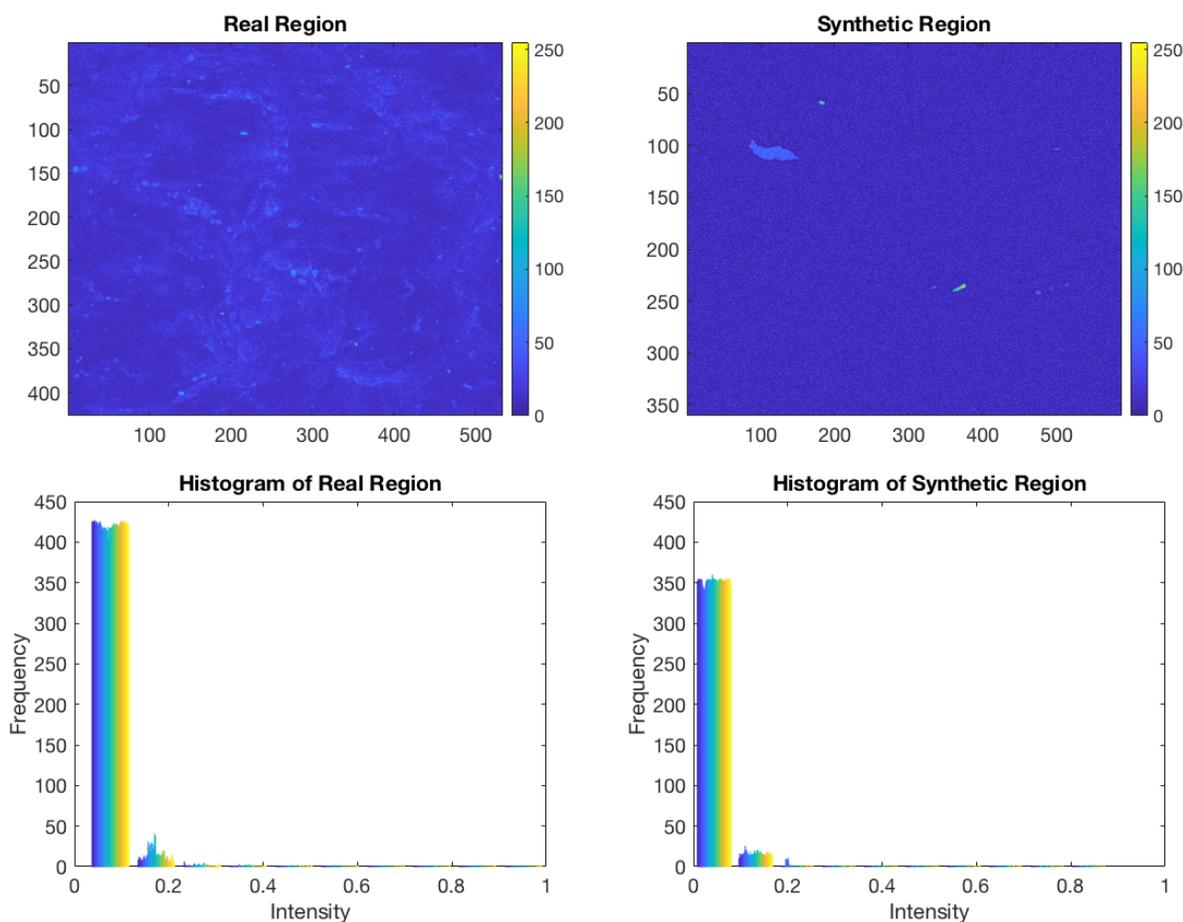
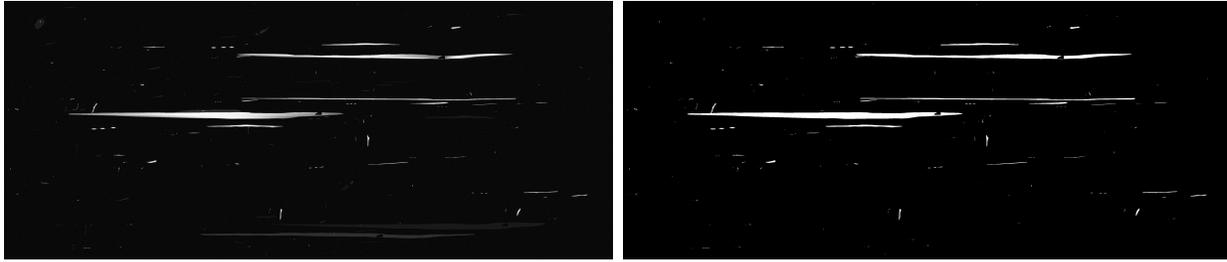
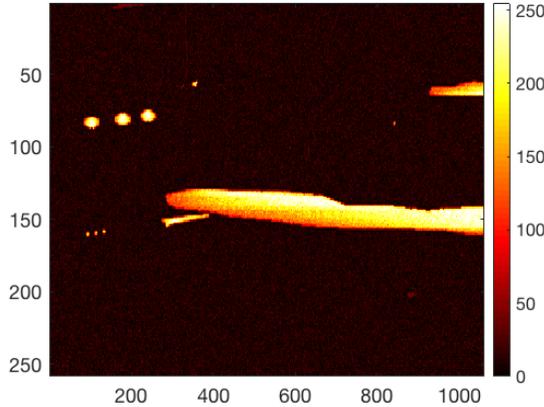


Figure 22: Plotting the small background region in a microscopy image (left) and a synthetic image (right.) There is more faint background clutter in the original image that we do not simulate, but the background intensity is in a similar range. In the synthetic images, the background intensity cannot be lowered too much, otherwise the Poisson noise effects are no longer apparent; which is why the static background in the synthetic image is higher. Histograms of a near constant intensity region in both images are presented on the bottom panel: in a microscopy image (left) and a synthetic image (right.)



(a) Synthetic Image

(b) Ground Truth



(c) Detail of Synthetic Image

Figure 23: One example image from the synthetic dataset. The ground truth is shown on the right, and a close-up of the synthetic image in a red-based colormap in (c), in order to show the added noise, blur, intensity variation, and background clutter.

Fig. 23 showcases one synthetic image. The ground truth is shown as well to display where the clutter resides.

5.4 Segmentation Results

All methods discussed in Section 4.3.3 are tested on the synthetic dataset developed in this Chapter. Average/median Dice coefficients and BF-scores are presented in Tables 4 and 5, complete with standard deviations.

On synthetic data, any noise that is not eradicated in thresholding is magnified by TuFF, where the attraction force (even when used at a minimum) joins noisy elements together.

Table 4: Average Dice Coefficient and BF-score for All Methods

Method	TuFF		L2S		Mat.		LyMPhi	
Metric	<i>Dice</i>	<i>BF-score</i>	<i>Dice</i>	<i>BF-score</i>	<i>Dice</i>	<i>BF-score</i>	<i>Dice</i>	<i>BF-score</i>
<i>Synthetic</i>	0.06±0.04	0.15±0.08	0.58±0.25	0.79±0.20	0.81±0.13	0.79±0.12	0.89±0.09	0.96±0.05

Table 5: Median Dice Coeff. for All Methods

Method	TuFF	L2S	Mat.	LyMPHi
<i>Synthetic</i>	0.05	0.56	0.91	0.91

LyMPHi performs the highest on the synthetic data, as it is the best at separating the varied intensity foreground from the background. The ground truth compared to here is infallible, so there is resiliency in these results.

5.5 Discussion

In this chapter, the first SRVF based data augmentation method has been developed, which should be used for creation of MLV data, as well as any other data where shapes undergo elastic deformation. Our method of synthetic data augmentation is also a fast ground truth generation, which is of necessity in a variety of fields [138].

5.5.1 Conditional Generative Adversarial Networks

Can more realistic data be generated using deep learning methods, as discussed in Section 5.1.1? Images of red blood cells were generated using conditional generative adversarial networks in [139], the workings of which have been discussed previously. This work shows promise for augmentation of MLV images. The red blood cells were first “grown” in the images using a shape based model of red blood cells, much like our vasculature is modeled by known shape. Then noise, intensity variation, etc. is added using the conditional GAN. However, using the current machines available, these networks are unable to be trained or tested - there is not enough memory to store the weights of a network necessary to output a high resolution, full size generated image. There are also simply not yet enough data to train a conditional GAN in order to make more realistic MLV data.

5.5.2 Other Uses for Morphing

Aside from data augmentation, there are two alternative ways to use the morphing paradigm described in Section 5.3. During dissection of the meningeal lymphatic vessels, prior to imaging, the vessels often become disconnected at points. This does not reflect the real structure of the vessels inside the mouse brain. The bending and stretching energies used in elastic deformation can be used to merge disjointed vessels, resulting in a complete vessel network captured for each mouse.

Another way to use morphing is to model the change in vessel structure from young vessels to old vessels, i.e., model the thinning and degeneration of vessels. This can be performed by finding the intermediate stages along the shape manifold between a healthy vessel and an aged vessel. Quantifying the morphing cost between a healthy vessel and other vessels of different ages could provide a metric on how “aged” the vessels are. This work can also be extended to morph healthy vessels to lymphatic vessels in a state of infection.

Neuroscientists can use the above morphing tool to generate more data of lymphatic vessels whenever needed. Perhaps, in the future, models can be built for generating healthy versus aging vessels, and studies can be simulated before testing with live mice.

In the next chapter, a deep learning study on smaller images of the MLV will be performed in compared to traditional segmentation methods.

6 Deep Learning

This chapter showcases a u-net based architecture for the automated segmentation of images of meningeal lymphatic vessels. We are preparing for the time when there is sufficient data available for training. Until that time, LyMPhi is expected to have higher segmentation accuracy. In the future, with enough training data and labels, deep learning methods could perform more efficaciously. We chose the u-net architecture because of its proven success in biomedical segmentation problems, especially for retinal vessel images [140]. U-net also has built-in methods to augment data, such as rotation, cropping, adding noise/blur, etc. These are useful for our dataset, as the experimental data are limited.

6.1 Motivation for Deep Learning

As the meningeal lymphatic vasculature itself is a recent discovery, there is no available software tailored for automatically segmenting these images, other than the one proposed in this thesis, LyMPhi (Chapter 4). Instead, segmentation must be performed by hand, which is a tedious and error-prone process. By building an automatic segmentation tool for these vessels, we can provide more measures for understanding and researching them, in a quick and reliable way. A convolutional neural network, called u-net, is adapted to the vessel segmentation application, with the goal of teaching the network how to segment the vessels. Segmentation using u-net is compared to traditional non-learning based segmentation methods using Dice coefficient.

Examples of traditional segmentation methods for vascular datasets (neurons, blood vessels, etc.) include level-set segmentation methods [103, 104, 107] and image matting [60]. These methods are proven to work well on vascular data, but in the case of retinal images, deep learning has shown new promise. Among the most prominent methods of computer vision and machine learning are convolutional neural networks, or CNNs. This chapter explores the application of CNNs to segmenting MLV in the brain via microscopy images, by applying existing medical-based CNN architecture to improve the time taken to process MLV images. An example result from our work is shown in Fig. 24.

Our research question thus is: can we improve the accuracy of segmenting the meningeal lymphatic vessels by using a CNN as compared to traditional segmentation methods?

6.2 Background on Deep Learning

Deep learning has a wide range of applications. Some networks are used for classification, some for enhancement prior to segmentation [141]. Picture quality can be assessed at the patch level using the network PaQ-2-PiQ [142]. Deep neural networks built for segmentation are essentially classifiers at the pixel level.

Convolutional neural networks are also used for object detection in tracking problems. In [143], the authors find that the main reason one stage detectors struggle with accuracy is class imbalance during training. Class imbalance here refers to having thousands of potential box (detection) candidates and very few useful ones. In a one stage detector, there is no intermediate task which must be performed in order to produce an output. This one stage detector leads to a simpler and faster model architecture. By adding a factor in front of

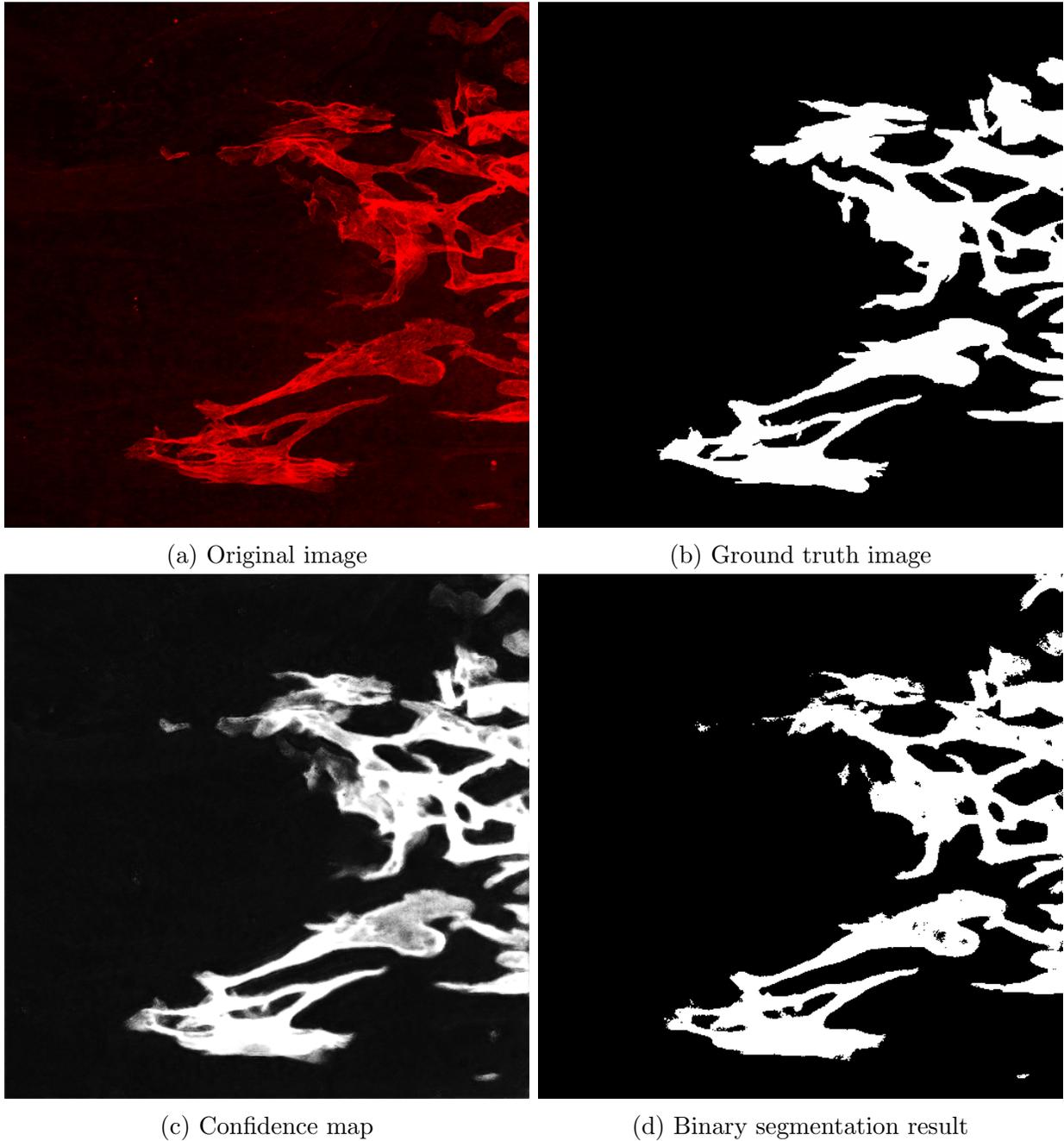


Figure 24: Example result. The original red channel image (cropped to size 512x512 pixels) is shown in (a), the hand created ground truth in (b). U-net generated confidence map is shown in (c), and the binary segmentation result in (d) is produced by thresholding the confidence map at 50%. The Dice coefficient for this image is 0.89.

the standard cross-entropy loss function, the authors of [143] were able to get a classifier that is more accurate than one stage detectors, but is faster than 2 stage detectors. Cross-entropy loss, or log loss, measures the performance of a classification model whose output is a probability value between 0 and 1. The authors call their technique focal loss, which refers to downweighting the loss assigned to well classified examples in a standard cross-entropy loss function. The network they build and train is called RetinaNet, a network featuring dense layers; layers in which each node is connected to a node in the next layer. [143] The RetinaNet has also shown its efficacy in segmenting the optic disk in fundus images [144].

Other work on neural networks seeks to address some common problems, such as the vanishing gradient problem, which often arises when a given network is too deep (has too many layers.) Residual neural networks, or ResNets, are able to address this issue by introducing some skip connections between layers [145].

Region Based Convolutional Neural Networks, or R-CNNs, extract local regions of interest for classification [146]. R-CNNs have shown promise for segmentation problems, particularly where same-class objects repeat throughout the image [142].

Newer work on neural networks (NN) allows for user-input for loss term coefficients, meaning that the overall network need only be trained once. The authors of [147] train a network with multiple loss terms. The NN learns over a distribution of the loss term coefficients, and then the user selects the coefficients to be used as a type of input to the model. For example, the user input can allow for having a network trade-off between image compression and quality. [147]

Many neural networks have been built for segmentation of the retinal vessels in fundus images, not just the optic disk. For example, a three-stage deep learning model was proposed by [109], which segments thinner blood vessels separately from thicker ones.

The authors of [115] build a convolutional neural network for segmenting closed loop structures, i.e., cells, in biological microscopy images. The u-net includes an expander network before the output layer to increase the resolution of the segmentation output. The second major contribution the authors make is to use their own elastic deformation model to create new training data from the limited data available, as discussed in Chapter 5. The deformations are created by randomly displacing pixel intensities along a 3x3 window. The interior pixel displacements are then interpolated. Using the augmented data and the trained u-net, segmentation results are vastly improved compared to other competing methods. [115]

6.3 Choice of Network

Images of MLVs are similar to fundus images in some respects, so using deep learning for segmentation may further enhance the quality of segmentation results. The main challenge with using learning in this area is that the available datasets are not large enough to train networks. For this reason, we choose u-net [115] as the network of choice, because a necessary data augmentation step is built-in to the network pipeline. u-net was designed to handle datasets with few training images. This thesis documents the first time application of deep learning to images of this kind.

6.4 Training Data

One of the central problems of biomedical images in regards to machine learning is the size of the data set; most images are hard to come by because each image represents one (expensive) mouse being sacrificed. The original images of meningeal lymphatic vessels are 2D confocal microscopy image acquired by the Kipnis lab in the University of Virginia department of Neuroscience. There are 39 images in total, which does not constitute a sufficient dataset to train a deep neural network.

Once the data had been processed and augmented, training was ready to begin. The version of u-net we modeled can be found in the following repository [140]. A three-operator majority voted ground truth was created (by Nazia Tabassum and colleagues) and used for training labels.

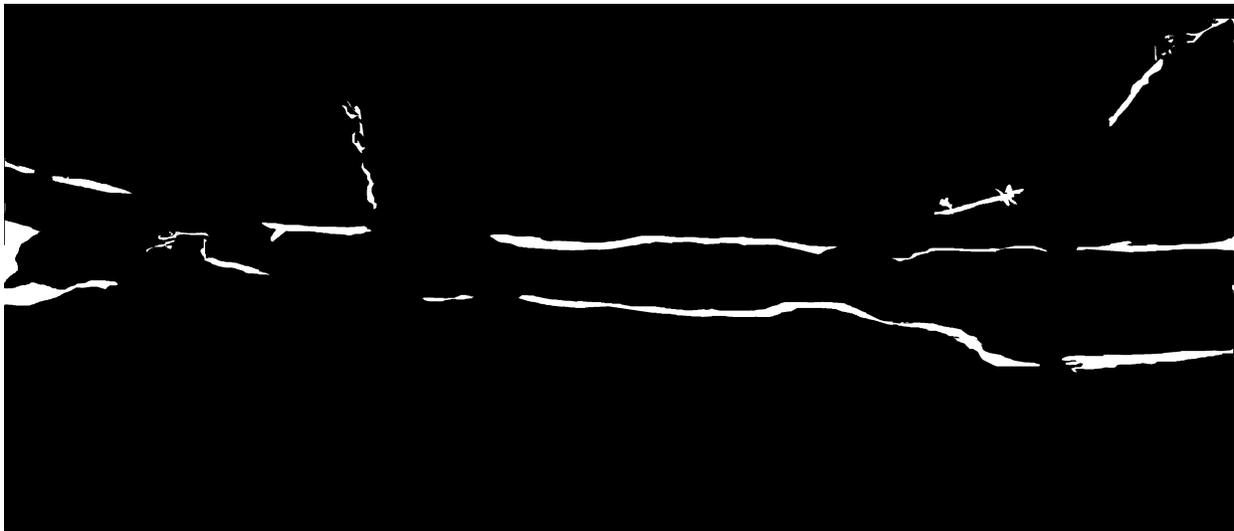
We ran our deformation pipeline to generate 1000 images of size 512x512 to act as training and testing for the ground truth masks and original images, and split those up into 750 training and 249 testing, based off of the original 39 images split into 30 train/9 test. Splitting of the data into train and test datasets occurred prior to data augmentation, making sure not to include any original images or deformed images from training within the test data set. The network then read in and trained on the 750 image pairs, and then was tested on the 249 test images. Data normalization was added after each convolutional layer to further reduce the noise of output images. The network hyperparameters were as follows, as well as the training times and hardware:

1. Training Time: 2 Hours
2. Hardware: Nvidia Titan X and Nvidia Titan XP x2
3. Optimizer: ADAM with $\alpha = 1 * 10^{-5}$ learning rate
4. Loss: Binary Cross Entropy
5. Implemented On: Tensorflow with Keras
6. Number of Epochs: 100
7. Steps/Epoch: 10

The original images varied in size from 4000x2000 to 2000x10000, and the network was trained on images of size 512x512. This training on subsampled images was done to aid computation speed. Results are promising with the above configuration on 512x512 size images, as demonstrated below. However, increasing the image size to 1024x1024, which is closer to the original image size, did not output meaningful segmentation results. This method of data augmentation does not generate enough new features for segmenting images of larger size, possibly leading to vanishing gradients within the CNN [148].



(a) Whole Mount Ground Truth



(b) Partial Mount Ground Truth

Figure 25: To train the CNN, we need labeled data. Using three annotators, we hand segmented the images in ImageJ, by clearing out any background clutter around the vessels. The three annotations are combined using majority voting to reduce noise in the output. Majority voting only counts a pixel as foreground if 2 out of 3 annotators have marked it as such. Shown are two examples of majority voted ground truth labels we produced. These are used for our training labels.

6.5 Assessing Segmentation Results

The trained u-net was tested on 249 test images, augmented from the real dataset. The u-net output is a confidence map in grayscale, higher confidence represented by higher intensity (on a scale of 0 to 1.) 100 percent confidence that a pixel belongs to the foreground set is denoted by intensity 1, and the color white. Thresholding is performed on the outputted confidence map at 0.5, or 50 percent confidence, to produce a binary segmentation. The confidence maps and binary results are shown in Fig. 24 and Fig. 26.

Fig. 26 contains images with segmentation results. The original red channel image of meningeal lymphatics is shown in the top left corner. The majority voted ground truth is shown on the top right. The u-net output is a confidence map, which we threshold at confidence level 0.5 to get a binary segmentation result, with resulting Dice coefficient of 0.89. This is a high segmentation result, as the highest value the Dice coefficient can take is 1, which indicates a perfect match with the ground truth. Dice coefficients above 0.8 show strong segmentation performance.

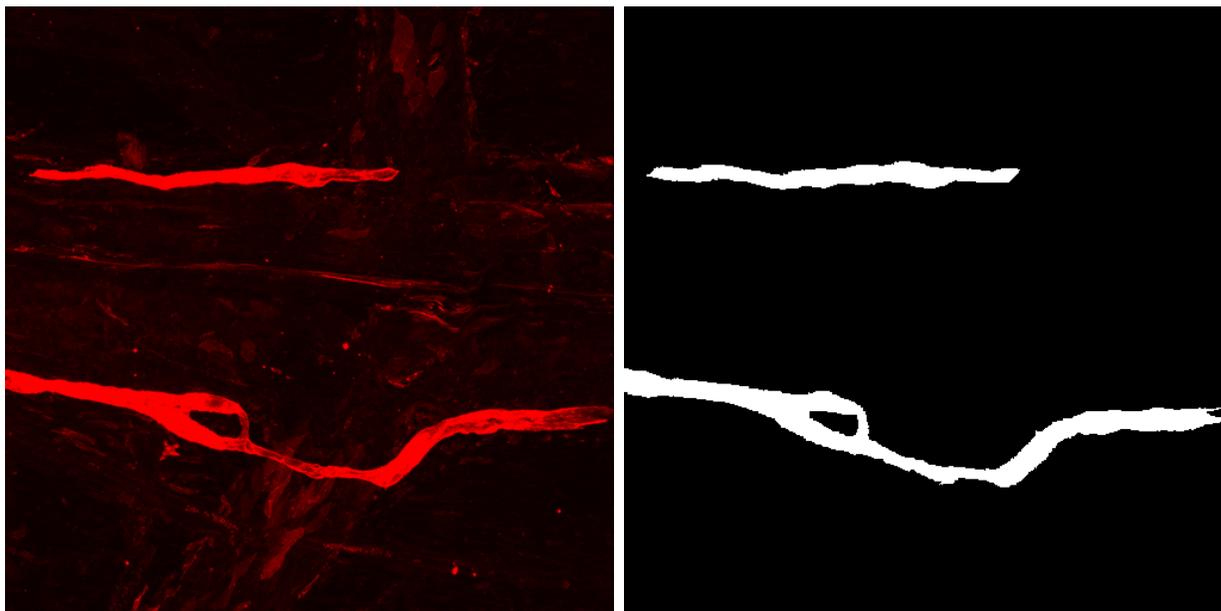
Table 6: Dice and Hausdorff Distance for Test Set

Metric	Dice Coefficient	Hausdorff Distance (microns)
<i>Average</i>	0.72	5.87
<i>Standard Deviation</i>	0.15	1.16
<i>Mode</i>	0.78	4.8
<i>Median</i>	0.76	5.86

This binary segmentation was evaluated compared to the hand-labeled ground truth by using Dice coefficient [108]. A higher Dice coefficient indicates closer alignment with the ground truth. Shown here are the quantitative results for the results on our test dataset using u-net segmentation. Though the average Dice coefficient using u-net is 0.72, in Fig. 27, it is seen that the most frequent Dice coefficient values are between 0.8 and 0.9, which are reasonably high scores in terms of segmentation accuracy. The Hausdorff distance is shown in microns [149], also in Table 6. Lower Hausdorff distance means a better match with the ground truth.

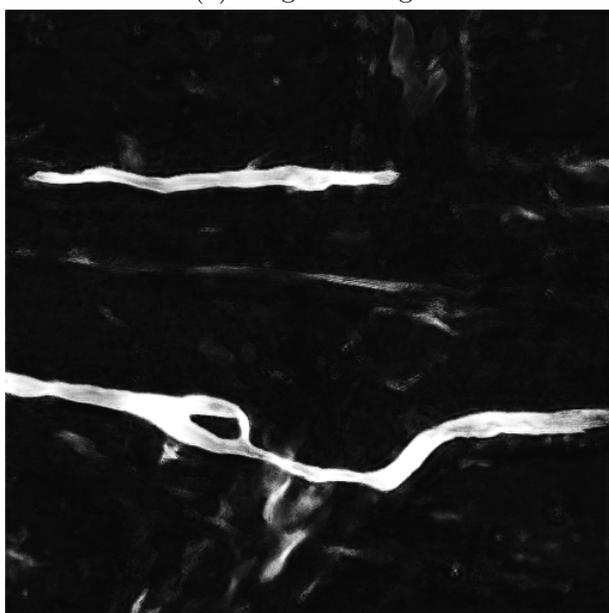
The same 249 images were used to test four separate methods for vessel segmentation: Chan-Vese [107], L2S (Legendre Level Set) [103], Hierarchical Image Matting [60], and TuFF (Tubularity Flow Field) [104]. The Dice coefficient was computed for all the results, reading across the top row (in bold), with statistics reported in the subsequent rows. The average Dice coefficient for the segmentation results is 0.72, as shown in Fig. 28. It is clear to see that the average Dice score for u-net, is much higher than any of the competing algorithms, and that the standard deviation is also comparable or lower. Hausdorff distance for the other methods is shown in Table 7 as well, and it is clear to see that using the other methods produces higher Hausdorff distance than when using u-net for segmentation.

In Fig. 29 we show results on a spinal lymphatic image with other state of the art segmentation methods. Chan-Vese, L2S, and TuFF are all level-set segmentation methods. Hierarchical image matting, in the bottom right corner, is a correlation based segmentation

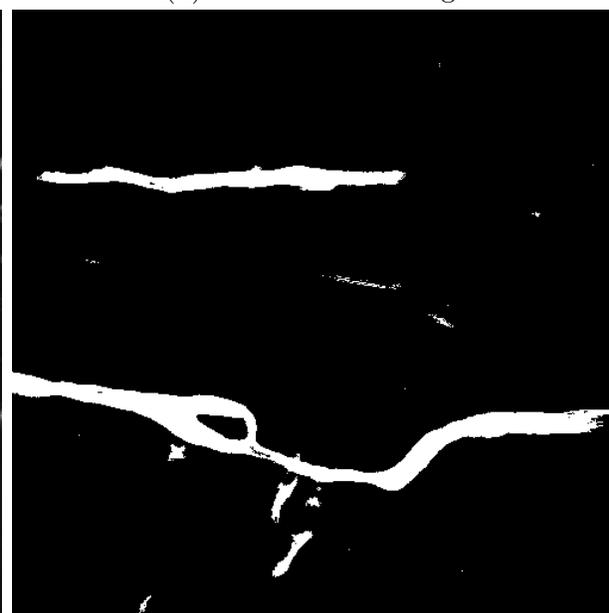


(a) Original image

(b) Ground truth image



(c) Confidence map



(d) Binary segmentation result

Figure 26: Example result on the partial mount dataset. The original red channel image (cropped to size 512x512 pixels) is shown in (a), the hand created ground truth in (b). U-net generated confidence map is shown in (c), and the binary segmentation result in (d) is produced by thresholding the confidence map at 50%. The Dice coefficient for this image is 0.89.

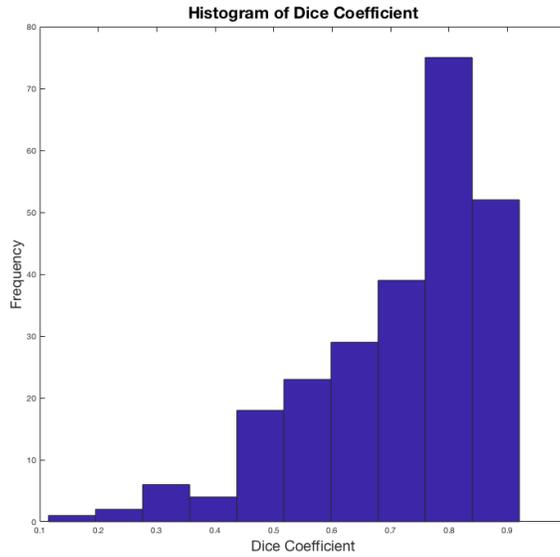


Figure 27: Histogram of Dice coefficients for test data.

Table 7: Dice Coeff. and Hausdorff Dist. (microns) for Competing Methods

Method	Chan-Vese		L2S		Mat.		TuFF	
Metric	<i>Dice</i>	<i>Hau.</i>	<i>Dice</i>	<i>Hau.</i>	<i>Dice</i>	<i>Hau.</i>	<i>Dice</i>	<i>Hau.</i>
<i>Avg.</i>	0.63	7.07	0.62	6.99	0.58	6.98	0.50	8.07
<i>Std. Dev.</i>	0.19	1.86	0.18	1.77	0.15	1.59	0.23	2.16
<i>Mode</i>	0.68	6.09	0.56	5.19	0.53	5.09	0.57	6.70
<i>Median</i>	0.68	6.64	0.64	6.86	0.61	7.02	0.51	7.63

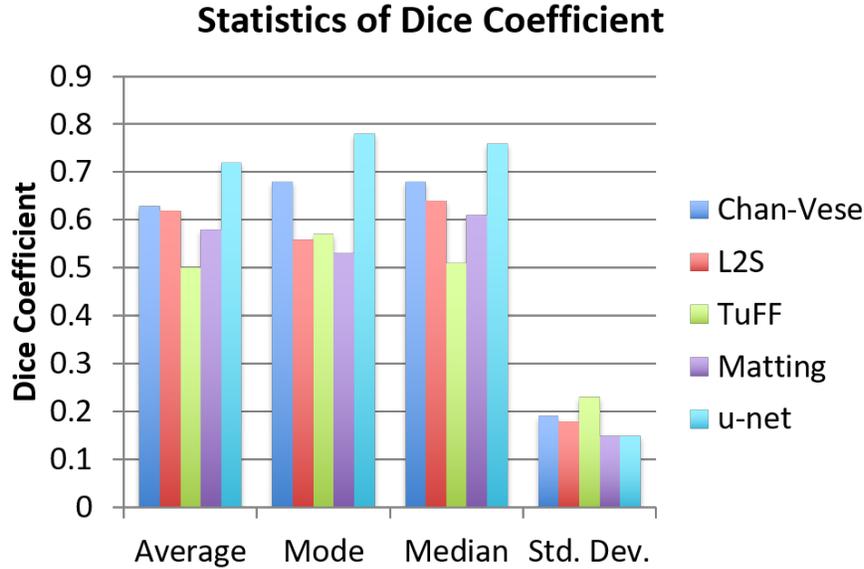


Figure 28: Statistics of Dice coefficient calculated for all segmentation results.

method developed for segmentation of retinal images. Using u-net provides a Dice score of 0.92, a very high Dice score in its own right, and much higher than all the other competing methods. The details in the u-net segmentation result are much clearer, even though the staining in the original image is weak between branches of the vessel, as seen in Fig. 30.

6.6 Discussion

The main outcome revolves around the improved processing times for the data given. Neuroscientists using the tool need not waste valuable resources on hand quantifying the image data. Another glaring issue is that of accuracy and reproducibility; different observers can and do segment the images differently. Therefore, a lack of consistency among the data analysis is prevalent.

Our adaptation of CNNs to the segmentation problem for vessels is also useful for anyone in the future wishing to segment vessel data. Our trained network can be used quickly and efficiently on any new images. Scientists using our tools can spend more time developing a better understanding of the lymphatic system’s relationship with disease. There is hope that the onset of disease could be predicted by studying lymphatic vasculature [3]. If so, drugs can be developed (indeed neuroscientists have already started such studies) to improve vessel function and possibly cure diseases such as Alzheimer’s.

If segmentation of human MLV is desired, 3D CT-lymphography is available [90]. The network will however need to be retrained on the CT images, and there will be the challenge of working with a larger 3D dataset, requiring more images and memory to train the network than used in this example.

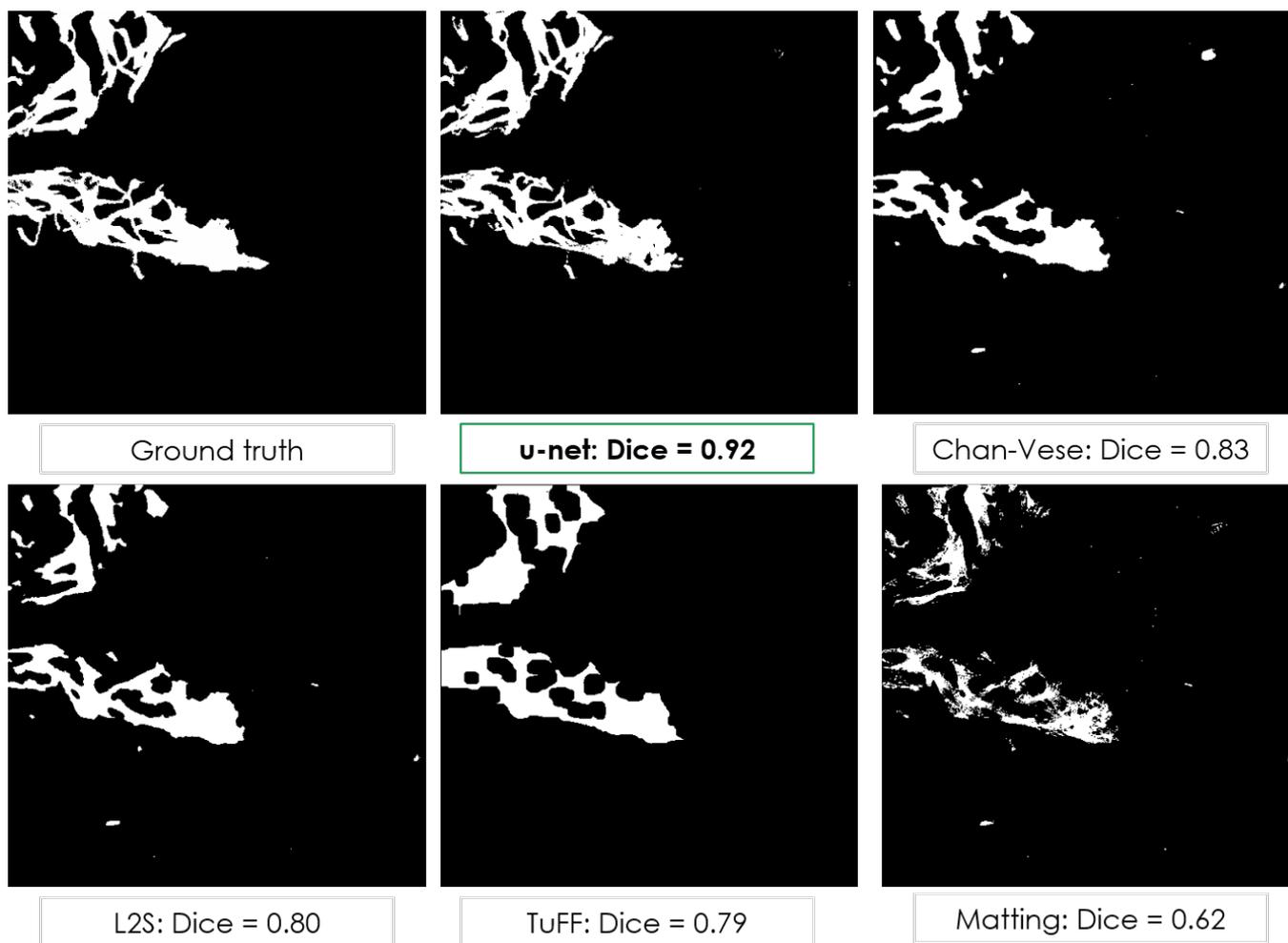


Figure 29: Example result on the spinal dataset. The original ground truth image (cropped to size 512x512 pixels) is shown in the first panel, the segmentation results following. U-net generated binary segmentation result in the top row is produced by thresholding the confidence map at 50%. The Dice coefficient for this image is 0.92.

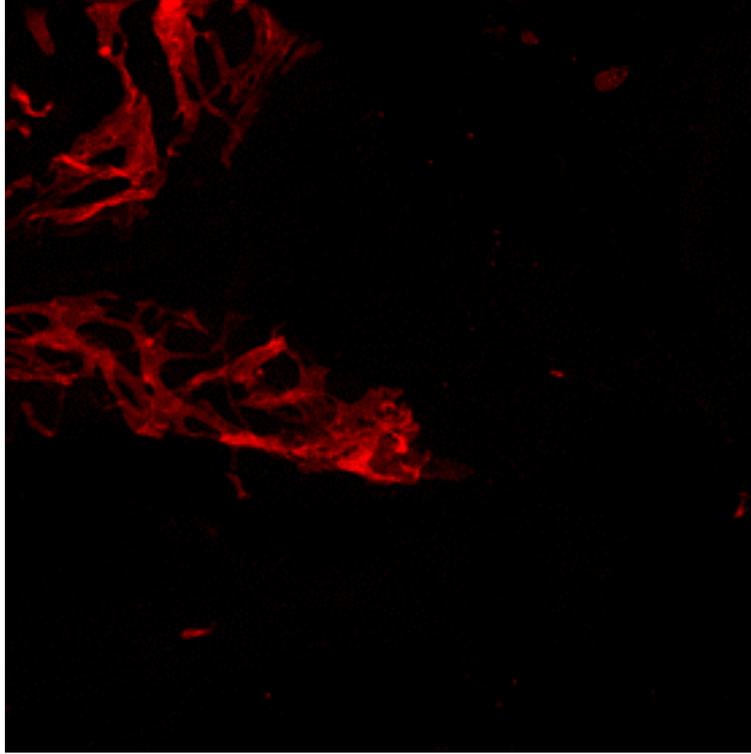


Figure 30: The original red channel image (cropped to size 512x512 pixels) is shown here for the segmentation results in Fig. 29.

6.6.1 Working with a Small Training Set

Working with a small amount of images for training poses challenges when experimenting with deep learning methods. The data augmentation possibilities discussed in Section 5.5.1 could be used with more hardware capabilities. But there are other possibilities for exploration as well. [150]

Patch based training and testing could potentially be implemented to boost the amount of data available for training. Doing so could save computational time if using a uniform size of image. [142] There are other images from different neuroscience labs that have not yet been collected and annotated for training, because the annotation procedure is so time-consuming. For deep learning, however, many weaker (or less accurately labeled) labels are better than fewer labels - so even poorly annotated ground truth is better than not enough examples. This could cause problems however when restitching the output to the original full image size - boundary artifacts may appear. A pretrained network could be used to boost performance (transfer learning), but some thought must be put into choosing the correct dataset. As discussed in Section 5.5.1, fundus images, although vascular in nature, are too fundamentally different to be used as the pretraining dataset. Unfortunately, many microscopy images that look similar to those of the MLV, such as images of microglia, suffer from the same problem - smaller size and lack of labels.

Self-supervised learning is another new area of interest, which helps deep learning methods when there is a dearth of labeled data. Self-supervised learning is good for problems

where transfer learning is not that applicable, such as when there are no similar enough datasets to use for transfer learning - a possibility discussed above. Self-supervised learning uses unlabeled data. First, a network is trained on surrogate tasks, where the surrogate task is something obvious, such as the relative location of two patches in an image, or whether two patches are from the same image. These surrogate tasks are easily labeled, as the labels can be generated automatically. After the network has learned these surrogate labels, the features that are learned during this training process are used for the actual classification task. This approach has been shown to work effectively for classification of spinal MRI images, for example [151]. In this case, again, much thought needs to be put into formulating appropriate surrogate tasks to solve the segmentation problem.

6.6.2 Limits of Deep Learning

There are some broader concerns about relying too heavily on deep learning methods for analysis. Deep learning methods currently do not incorporate time, space, or causality into their algorithms, the foundations for common sense. There are no methods that can understand how an object's *shape impacts its function*. AI does not yet exist under the assumptions that we have, such as that an organism cannot be in two spaces at the same time, for example. Before we can introduce AI into our hospitals, we need to train them to view the world with our common sense assumptions. [152] It would be interesting to be able to train a machine to understand how lymphatic vascular shape impacts its function in transporting immune cells and draining cerebrospinal/interstitial fluid.

In reality, we have just recently created the resources for deep learning, which are massive amounts of data. The methods currently used for neural networks have existed for a number of years already. There is much work to be done to truly make use of the new resources available, in the form of rich, high-dimensional images.

The next chapter will analyse the validity of using manual segmentation as ground truth for comparison, as well as for training labels.

Table 8: Agreement and Fleiss’s Kappa Between Annotators

Dataset	Agreement	Kappa
<i>Partial</i>	0.98	0.62
<i>Whole</i>	0.97	0.60
<i>Spinal</i>	0.96	0.75

7 Analysis of Manual Segmentation

7.1 Analysis of Agreement between Annotators

The results shown prior only compare automated segmentation results between different segmentation algorithms. Since the manual annotation obtained is combined input from three annotators, in this section the agreement and kappa statistics between the annotators is compared, as well as the statistics between the chosen segmentation algorithms and the annotations.

Fleiss’s Kappa is used for comparing the three annotations created by three separate operators for each vessel image. Fleiss’s Kappa measures the reliability of the agreement between the three annotators in this case. [153] Instead of creating a voting matrix and updating it for each pixel in a dataset, counts were kept of voting proportions and probabilities. Percent agreement and Fleiss’s Kappa are shown in Table 8.

The Fleiss’s Kappa values between annotators are between 0.6 and 0.75, indicating moderate agreement between annotators on where a vessel is present. For the spinal dataset, the Kappa value is highest, indicating better agreement than for the other two datasets. However, none of the Kappa values have reached higher levels (.8 or .9 for example), which are more preferred Kappa values - showing that there is still considerable disagreement between annotators, and that creating accurate ground truth for these images is not a trivial problem. [154] Surely, manual analysis has proven inconsistent, which motivates the automated analyses of this thesis.

Agreement and the kappa statistic are calculated for all images in each dataset, between the segmentation algorithm and the majority voted groundtruth. Cohen’s Kappa is used in this case because there are two raters being compared, the automated segmentation algorithm in question, and the consensus segmentation. Cohen’s Kappa is considered to be more reliable of a measure than percent agreement, as the Kappa measure takes into consideration the probability of random agreement. [155] Percent agreement and Cohen’s Kappa are shown in Table 9. Agreement is denoted by *Agr.* and Kappa by *Kap.*

From Table 9, agreement between LyMPhi and annotators largely outperforms inter-annotator agreement in the partial mount dataset. Thus, for this dataset, it can be stated that LyMPhi is superior to manual segmentation. In the other datasets, the Kappa value is lower than the Kappa values shown in Table 8, meaning that LyMPhi has not yet surpassed hand annotation for these datasets. However, the Kappa values for LymPhi do outperform the agreement between annotation and the other segmentation algorithms tested. This further affirms that LyMPhi is better suited for segmentation of meningeal lymphatic vessels

Table 9: Agreement and Cohen’s Kappa Between Algorithms and Annotators

Method	TuFF		L2S		Mat.		LyMPhi	
Statistic	<i>Agr.</i>	<i>Kap.</i>	<i>Agr.</i>	<i>Kap.</i>	<i>Agr.</i>	<i>Kap.</i>	<i>Agr.</i>	<i>Kap.</i>
<i>Partial</i>	0.87	0.23	0.94	0.44	0.97	0.56	0.98	0.70
<i>Whole</i>	0.68	0.14	0.84	0.28	0.93	0.43	0.95	0.52
<i>Spinal</i>	0.65	0.19	0.84	0.39	0.96	0.64	0.96	0.64



(a) Consensus Segmentation via STAPLE (b) Consensus Segmentation via Majority

Figure 31: Consensus segmentation results on a portion of an image from the spinal dataset. The results are displayed as binary. The left image shows a consensus segmentation obtained via the STAPLE algorithm. The right image is the manual annotation we used for comparison, obtained by majority voting.

than other automated methods.

7.2 Performance of LyMPhi and Each Annotator as Measured by STAPLE

The STAPLE algorithm combines ground truth annotations according to expectation maximization and also measures specificity and sensitivity of each ground truth annotator. The algorithm can be used to compare automated segmentation methods to ground truth as well. [156]

An example of where disagreement between annotators can occur is shown in Fig. 31. When combining three annotations using the STAPLE algorithm [156], noisy regions, such as in the top of the image in Fig. 31a can still remain. The images are difficult to annotate due to low-level noise that must be removed. Due to the low image contrast, patches of noise may be missed by an individual annotator. This, and other challenges, leads to annotator disagreement. Using majority voting, as shown in Fig. 31b, can reduce the effects of the disagreement due to noisy background; however, some small amounts of noise do remain, which create divergence when comparing with LyMPhi and other automated algorithms.

Using the STAPLE algorithm, an estimate of sensitivity and specificity of raters can be obtained [156]. In Table 10, the sensitivity and specificity of the three annotators as

Table 10: Sensitivity and Specificity

Dataset	Partial		Whole		Spinal	
Rater	<i>Sensitivity</i>	<i>Specificity</i>	<i>Sensitivity</i>	<i>Specificity</i>	<i>Sensitivity</i>	<i>Specificity</i>
<i>Rater 1</i>	0.77	0.99	0.91	0.98	0.82	0.99
<i>Rater 2</i>	0.77	0.99	0.56	0.99	0.96	0.97
<i>Rater 3</i>	0.76	0.98	0.81	0.98	0.93	0.99
<i>LyMPhi</i>	0.8	0.99	0.74	0.98	0.60	0.99

well as LyMPhi are compared. Sensitivity refers to the number of true positives that are measured, and specificity the number of true negatives [157]. This is in comparison to the STAPLE estimated consensus segmentation, which has been shown to have some flaws - so these sensitivity/specificity measurements are also only estimates.

Specificity is so high across all raters and datasets because the number of negatives, or non-lymphatic pixels, is so large, and it is easier to mark these correctly. High sensitivity, however, is more difficult to achieve. LyMPhi outperforms on sensitivity for the partial mount dataset, which is in accordance with the discussion of Kappa values in Section 7.1. This means that LyMPhi really is better at distinguishing the lymphatic vessels than either of the raters individually, or when they are considered together. Although this is not the case for the remaining two datasets, it is interesting to note that for the whole mount dataset, the sensitivity value for Rater 2 is lower than that of LyMPhi (0.56 compared to 0.74). So, in cases where multiple annotators cannot be reached to perform a consensus segmentation, there is great risk relying on a single annotator for a correct segmentation of the lymphatic vessels. As mentioned before, manual annotation takes a great deal of time; a more accurate segmentation may be found in less time by using LyMPhi, instead of relying on a potentially flawed hand segmentation.

7.3 Failure to Annotate Boundary by Manual Segmentation

Combined errors of sensitivity and specificity at the fine-grained level along the vessel boundary shows that our level-set method, LyMPhi, better captures the true vessel boundary than manual segmentation. This can be shown in Fig. 32, where the red circle in Fig. 32c shows that the boundary found by LyMPhi better matches that in Fig. 32a - the boundary in Fig. 32b has been flattened and does not provide accurate curvature of the vessel edge.

7.4 Discussion

This chapter has shown the risks associated with trusting manual segmentation fully and completely, particularly when relying on a single annotator, as is often the case in segmentation of the MLV. Chapter 8 will discuss the possibilities when a correct segmentation of the lymphatics is in hand: namely, structural complexity analysis.

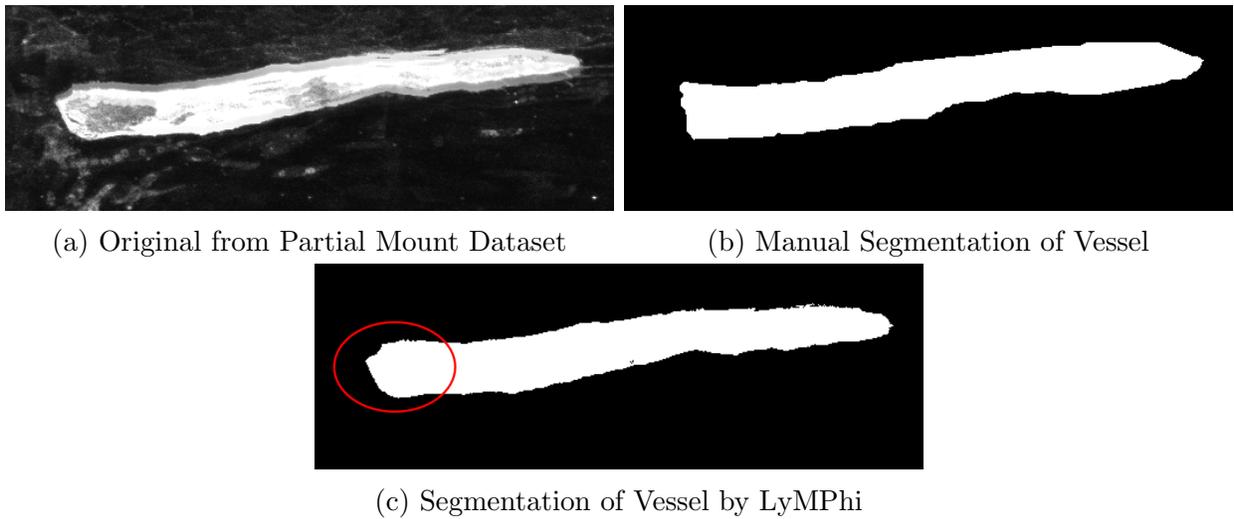


Figure 32: Consensus segmentation results on a portion of an image from the partial mount dataset. The segmentation results are displayed as binary. The top left image shows the original vessel from the cropped image. The top right image is the manual annotation, obtained by majority voting. The bottom image is segmentation performed by LyMPhi (Dice = 0.94).

8 Complexity Analysis

Now that segmentation is in hand, what can we do with the segmented vessels? This chapter discusses complexity measures for the meningeal lymphatic vessels. These complexity measures serve a dual-purpose: comparison between different vessels, and also providing a new way to measure segmentation accuracy. In the future, complexity may be a key bioinformatic used to assess and quantify disease progression and drug efficacy.

8.1 Motivation for Complexity Analysis

Since the meningeal lymphatic vessels were only recently discovered, there is no tool available for automatically analyzing these images. By building the first such tools, we can provide more measures for understanding these vessels.

After segmentation, neuroscience experiments require comparison of lymphatic vessel width, because during aging, lymphatic vessels thin and do not adequately drain wastes from the brain. The comparison of vessels needs to be extended beyond width comparisons, however - there are other interesting features of lymphatic vessels that could characterize their complexity or networking ability. For example, in Figure 33, there is branching structure present in the vessels, stained in green. This structure may have a functional relationship with drainage. We propose to build a set of measures for analyzing and characterizing lymphatic vessels.

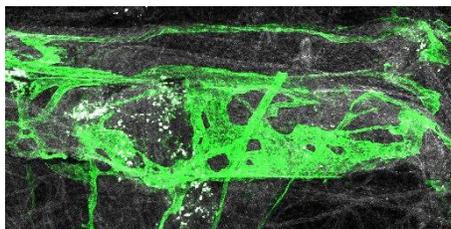


Figure 33: Lymphatic vessels stained in green.

Neuroscientists currently use hand quantification and analysis, which takes up to 7 hours for a single image and is not consistent from person to person. Width measures in the most recently published work have varied up to thirty percent! Width is sampled over the vessels, with each researcher sampling different points, so the samples are not reproducible. These widths are averaged over an entire image, smoothing over considerable vessel variability. Furthermore, quantification by hand consistently over-quantifies vessel quantity because there are fine details that are difficult to capture using mouse clicks. This again is incredibly time consuming, and any measurements would be more accurate if automated. Most current vessel analysis methods are semiautomated, like the approaches used in [158]. A fully automated method to capture lymphatic vessel structure is needed, so that scientists can have a reliable tool for lymphatic analysis. Furthermore, new relationships between vessel growth and function could be understood by introducing new ways of analyzing vessels. The ultimate question is can we predict these diseases or diagnose people with autoimmune disorders just by studying their meningeal lymphatic network?

A wider choice of measures would also yield more information about lymphatic vessels. Also, can complexity measures be tuned specifically for lymphatic vasculature in order to assess and analyze segmentation results?

8.2 Background on Vessel Complexity

Complexity is measured throughout the image processing world on a variety of different types of images. For example, diagram images, which are black and white images consisting of thin-lined pen and pencil diagrams, use something called an image signature to capture the information of the diagram. [159] Vascular skeletons do look similar to thin-lined diagrams, so vessel network signatures could potentially be captured using these methods.

Protein complexes also have measures of complexity. Protein complexes have holes, while single chains generally do not. The holes are places for subunits to bind. Larger complexes have larger holes, as they have more space for them. The authors take the protein volume to convex hull ratio: V_p/V_c . A smaller ratio means more holes. The problem is, this ratio is not just counting holes, which makes this ratio difficult to apply to lymphatic vasculature. A simple chain with skinny branches curving in different directions also has a low ratio. [127]

In neuron analysis, the relationship between the morphology and function of neurons was proposed by Santiago Ramón y Cajal. Morphological analysis of neurons and sub-components, such as dendritic spines, synapses, and mitochondria has shown promise in understanding neurodegenerative diseases. Comparing neuronal structures is also of great importance. A comparison involves computing a similarity score for a type of neuron based on a training dataset. The training set is built by categorizing neurons based on their functionality. As more datasets of traced neurons from different organisms become available, robust comparisons can be made. [4]

Most existing neuron tracers assume a tree-like structure, with no loops. This produces some difficulty when attempting to translate neuron analysis methods to lymphatic analysis, as lymphatic vasculature is ripe with capillary looping. However, techniques used to analyze dendritic spine morphology could be translatable to lymphatic branch types. Spine shape is used to classify the spine, based on maximum curvature, convexity, area, and other attributes. These are used to predict the spine's functional category. [4]

The most recent work published in this area uses hand quantification and analysis for lymphatic vessels [1]. Since these vessels have not been well studied from an image processing perspective, methods for other vascular structures, such as neurons, are of interest. The first step is to create an accurate vessel skeleton. There are many methods for producing accurate vessel skeletons, some based on marked point process, such as [160]. An interesting way to build a neuron graph, and subsequently, a skeleton uses marked point process (MPP) to birth marks along neuron centerline until the minimum energy arrangement is found. The birth and death process proceeds according to simulated annealing. Marks are assessed by their neuriteness (location on centerline) and interaction with other marks (not too crowded, not so spread out that they don't capture enough of the neuronal path.) [160]

The MPP solution is then used to build a graph, with each mark as a node. Gradient vector field speedmap is used to drive fast marching to iteratively find the minimum paths from start node to terminal node. This results in the minimum spanning tree (MST) of the graph, which is the reconstructed neuron skeleton. When initializing, the user must

overestimate the number of marks you need for all the objects (neurons, vessels) in the image. [160]

We propose using a method called hierarchical skeleton abstraction, where an initial skeleton is simplified using minimization over the number of branchpoints [135]. This method provides a smooth skeleton, with fewer false branchpoints, and is more generalizable to lymphatic vessel structure - which is wider than neuronal structure. After obtaining the skeleton, analysis measures similar to those found in common neuron analysis papers, such as [161], can be used, such as calculating the number of branchpoints and total vessel length. More branchpoints signify higher complexity.

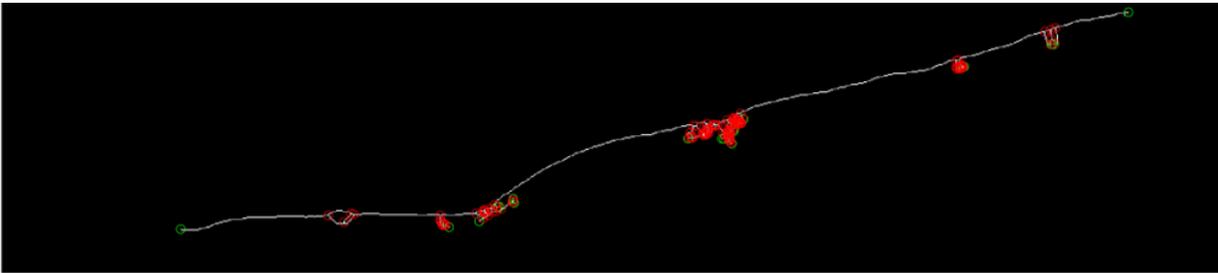


Figure 34: Branchpoints plotted in red and endpoints plotted in green on a vessel skeleton.

Another surveillance measure is the ramification index. The ramification index is calculated for microglia and is defined as the ratio of the perimeter of the cell to the area, normalized by the same ratio for a circle of that area. This index quantifies how ramified, a cell is, or how spread out the branches (processes) of the microglia are. [162] While this index does not directly translate to lymphatic vessels, as the vessels are not a single cell with processes, calculating a similar index and seeing how it scales with lymphatic complexity is still of interest. A similar measure of interest is porosity.

8.3 Proposed Complexity Measures

To further assess the segmentation results, we introduce three measures into the study of lymphatic vasculature: vessel length, vessel ramification index, and vessel porosity. Vessel length is widely used to study neurons and blood vessels, but vessel ramification and vessel porosity (as an image analysis measure) do not exist in the literature.

The following are results shown on a couple of images from the datasets described in 4. The images have been cropped to focus on one connected vessel at a time, as the vessels are often disjoint due to dissection.

Fractal dimension was calculated as a potential measure of how much detail changes in a pattern compared to the measuring scale. It can measure the “space-filling capacity of a pattern.” While the fractal dimension does increase with more complex vessels, the difference in fractal dimension is not very discriminative between simpler and more complex vessels. From our experiments, the range of fractal dimension found is only from .9 to 1.4, approximately. A more discriminative measure is desired.

8.3.1 Skeletonization

The next step to much of vessel analysis after segmentation is skeletonization. The method used in this thesis is called hierarchical skeleton abstraction [135]. This is a generalized skeletonization algorithm for objects designed in 2004. An initial skeleton is created from a simplified silhouette of the object. The initial skeleton is then simplified under two conditions: similarity to the initial guess and having the fewest branching points possible. Both boundary and inner branches are considered. (If an inner branch is removed, the two sub-skeletons are joined.) The skeleton simplification is an optimization problem, where the two constraints are that the two endpoints remain fixed and that eliminated branches do not reappear. A few hundred breadth first sweeps of the skeleton are needed before convergence. [135]

8.3.2 Vessel Length

We can first build a skeleton from the segmentation to start the automated complexity analysis. Skeletons are one-pixel wide backbones of image objects. Displayed in 34 is a lymphatic vessel skeleton with endpoints marked in green, and branchpoints marked in red. One measure of interest for lymphatic vessels is vessel length. This can measure lymphatic regression; regression is where the vessel grows shorter in length over time, with age. The vessel length is calculated as the total number of pixels along the skeleton, as the skeleton is one-pixel wide.

Vessel length was computed for all binary segmentation output by first skeletonizing. This complexity measure is commonly used for neuronal analysis [161].

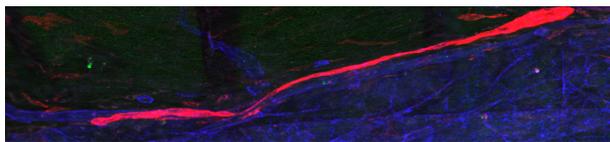
8.3.3 Vessel Ramification Index

Another measure of interest is the ramification index. A ramification index has been calculated for cells such as microglia and is typically defined as the ratio of the perimeter of the cell to the area, normalized by the same ratio for a circle of that area. Again, this index quantifies how ramified a cell is, or how spread out the branches (processes) of the microglia are. [162] This index does not directly translate to lymphatic vessels, as the vessels are not a single cell with processes.

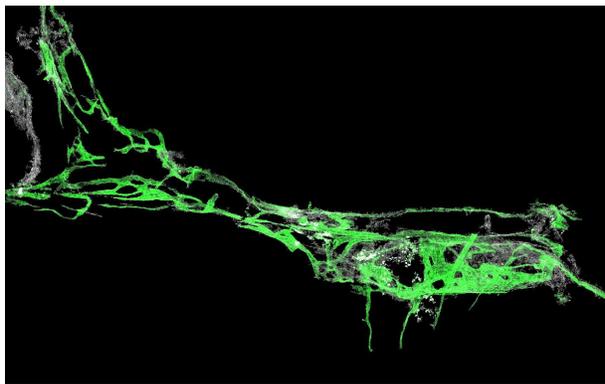
Calculation of the ramification index was performed next on two types of vessels: one simpler vessel, and a “more ramified” vessel. These are both shown in Figure 35. The RI does indeed scale with complexity within vessels. However, there could be more meaningful ways of calculating complexity that are more directly applicable to lymphatic vessels.

The ramification measure was then re-designed to better fit lymphatic vessels. Instead of comparing to a circle of the same area, the comparison is performed with a simpler vessel – a vessel with one smooth contour all along the outside, a filled vessel. The boundary was traced roughly in Fiji to create a smooth vessel approximation, for visual purposes only. A depiction of the process is shown in Figure 36. As it was done by hand, the boundary is much wider than the vessel (boundary in yellow.) In the experiments, the vessel boundary lies directly on the edge of the vessel.

The vessel ramification index is defined as

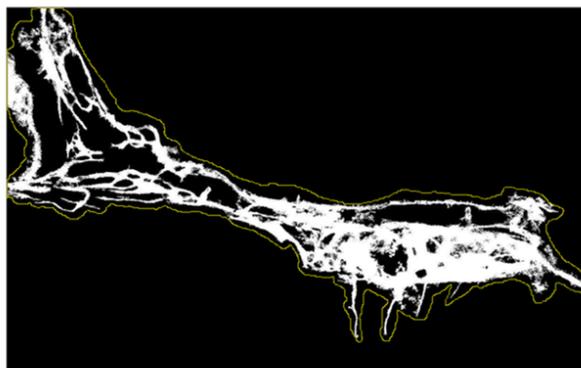


(a) Simple Vessel



(b) Complex Vessel

Figure 35: Two types of vessels.



(a) Tracing Vessel Boundary



(b) Simplified Vessel Result

Figure 36: Simplification of vessel.

$$RI_{vessel} = \frac{Peri_1}{Area_1} \div \frac{Peri_2}{Area_2}, \quad (13)$$

where $Peri_1$ and $Area_1$ refer to the perimeter and area of the original vessel, and $Peri_2$ and $Area_2$ are the perimeter and area of the simplified vessel. This scales with the original ramification index.

8.3.4 Vessel Porosity

Another similar measure we developed is inspired by porosity for materials. Porosity, or void fraction, is a measure of the void or empty spaces in a material. It is the fraction of the volume of voids over the total volume. Porosity is always a number between 0 and 1, unless it is described as a percentage. A material has high porosity if it contains large spaces, such as the figure on the left in Fig. 37.

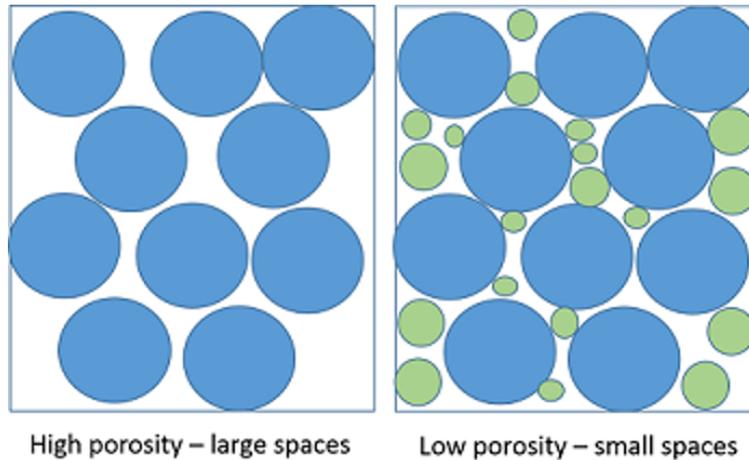


Figure 37: Depiction of porosity.

Similar to ramification index, vessel porosity is calculated as the area of capillary loops over the total surface area of the vessel, akin to how porosity is calculated for materials.

Vessel porosity is calculated as the area of capillary loops, which are the gaps where a vessel branches and reconnects to the larger network, over the total surface area of the vessel. We use area instead of volume because the images are 2D. Pictured in Fig. 38 are some images of the capillary loops as shapes. Vessel porosity is similar to ramification index in that it measures how much space is filled within the larger vessel boundary.

8.4 Complexity Results

The new RI is able to discriminate distinctively between simpler and more complex vessels. The simpler vessel has a ramification index (RI) of 7.50 and the more complex vessel has an RI of 31.28, with a new RI of 2.99 and 21.56, respectively. These results are depicted in Table 11. While the previous RI for microglia also is discriminative between two vessel types, the new one developed makes more sense for this problem.

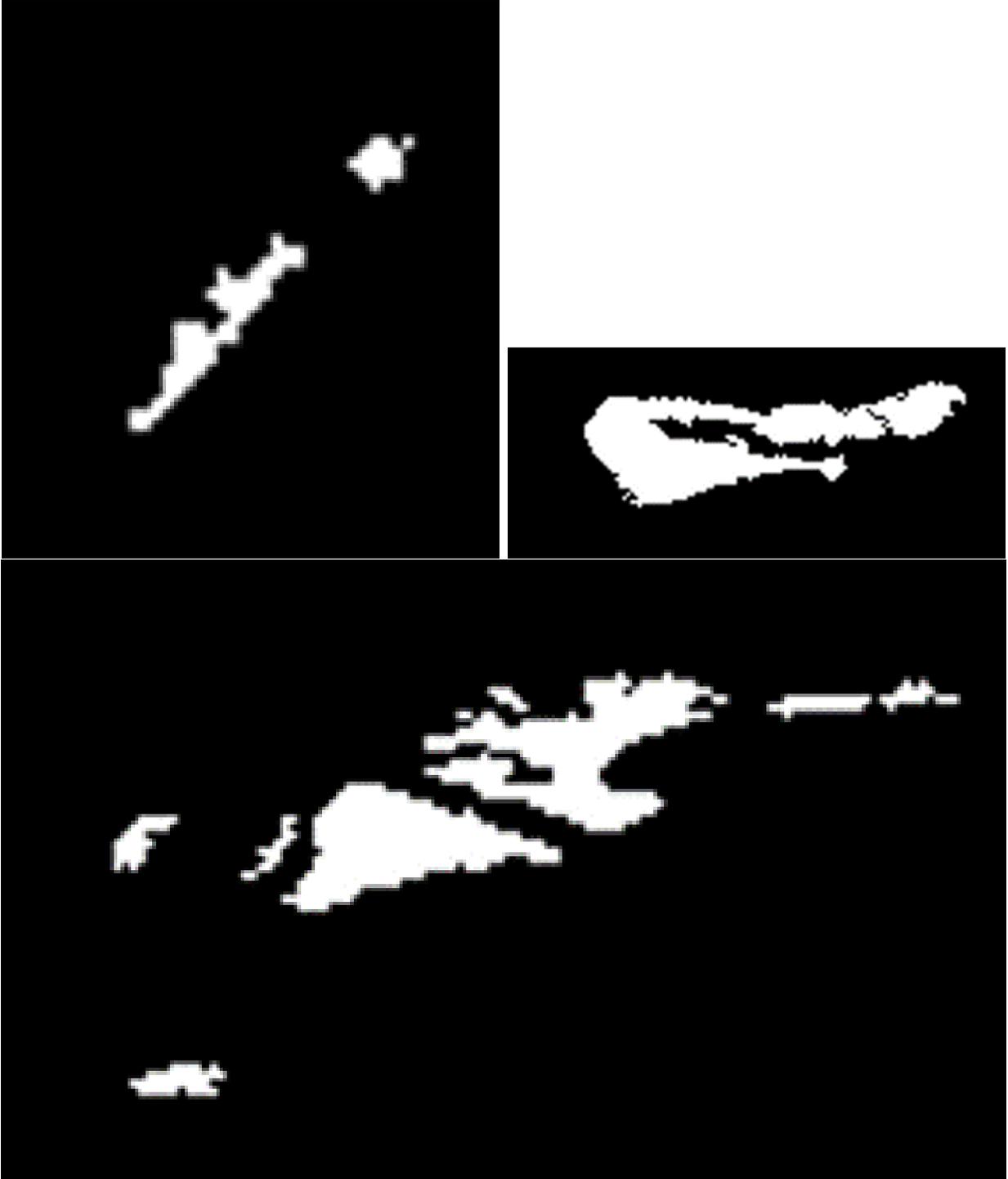


Figure 38: Three examples of capillary loops.

Table 11: Ramification Indices

Vessel Type	RI	New RI
Simple Vessel	7.50	2.99
Complex Vessel	31.28	21.56

8.4.1 Measuring Segmentation Accuracy

All three measures were run on all segmentation results, and compared to the ground truth using mean squared error. The u-net segmentation results have the lowest error for all three measures compared to the ground truth complexity values, as shown in Table 12. Vessel length varies widely because the skeleton produced changes enormously with changes in the segmentation. As porosity is always a value between 0 and 1, the MSE for porosity is lowest across all the methods.

Table 12: Mean Squared Error of Complexity Measures

Measure	Chan-Vese	L2S	TuFF	Matting	u-net
<i>Length</i>	3.77×10^6	4.05×10^6	4.24×10^6	2.41×10^7	1.63×10^6
<i>Ramification</i>	0.508	0.515	0.538	0.478	0.228
<i>Porosity</i>	0.015	0.015	0.015	0.015	0.009

8.5 Discussion

The complexity measures proposed in this paper are automatically calculated, and would serve to benefit the scientists that wish to use the data to draw conclusions. A main limitation of this work is the ability to correlate complexity measures with some real physical meaning, such as more/less drainage of waste. This is difficult to perform without directly working with neuroscientists to design experiments and make measurements.

In the future, the assessment of whether these measures are strong enough to analyze lymphatic complexity would be valuable. The work can be extended to building feature vectors with discriminative power between impaired and healthy MLV for use with simple classifiers, like in other works [134, 163].

8.5.1 Possible Future Complexity Measures

The percentage length of the superior sagittal sinus compared to the whole MLV network is also used as a measure. An extension of this could be to calculate the percentage length to the blood vessel network. [95]

A graph can be built from the skeleton, with branchpoints and endpoints as graph nodes. Within the graph signal processing framework, other measures can be taken on the vessels

that have not yet been applied: tortuosity, graph diameter, density of branches, etc. Tortuosity measures how twisted a path from one node is to the next, compared to the straight line distance between them. Graph diameter and density can provide an idea of network surveillance of the meninges. [134]

The above are primarily adapting known approaches to lymphatic vessels. A possible new way to describe vascular complexity by ascribing physical meaning to our measure in terms of flow rate. Ultimately, whatever measure we use to describe a network of vessels should allow scientists to infer how the structure influences the function of these vessels - in this case, how fluid drains through the vessels. For a single image, analyzing fluid flow is a particularly challenging endeavour. Traditionally, in biomedical data, fluid flow through vessels is estimated from video sequences. As our data is comprised only of individual images, we propose a novel way of describing fluid flow by modeling the vessel network as an electrical circuit. At various sample points, a thickness can be measured along the vessel. This thickness, combined with the viscosity coefficient and compressibility of the passing fluid (using Poiseuille's Law and Navier-Stokes equations), is named as the *resistance* of the vessel at that point [164, 165]. This resistance represents how difficult it is for fluid to flow through a vessel. A wider vessel implies lower resistance. If we imagine a circuit with various resistances along its length, these resistances can be combined to find the equivalent resistance of the vessel network [166]. This overall resistance describing the network would be used to compare the ease with which fluid flows through different vessel configurations. Using the vessel width in this way provides more meaning than averaging, because connection patterns (series, loops, parallel) has a strong impact on the equivalent resistance calculated.

Implementing this flow model will come with challenges, however, as the lymphatic vessels cannot be modeled with a traditional tree-structured graph. A graph structure allowing for self-looping must be built before attempting to describe fluid flow through graph analysis.

Measuring the implications of vessel structure could be more valuable than simple measuring vessel structure alone. This idea could lead to other interesting findings, such as calculating the maximum possible flow velocity at various points. This type of single image fluid flow analysis is novel, and is applicable to various types of vasculature, such as blood vessels. Indeed, through surgical research, Dr. Scott A. Berceci at the University of Florida studied different vein graft models in vivo, and measured blood flow. This research can show which graft models have higher flow vs. lower flow. [167]

We have shown how to segment the meningeal lymphatic vessels, as well as how to calculate complexity measures for the vasculature and compare different segmentation results. The future of lymphatic complexity analysis is ripe with possibility, some avenues of which have been discussed here. The next chapter will conclude this dissertation.

9 Conclusions

The main contributions of this dissertation can be summarized as follows. The first deliverable of this dissertation is a level-set based segmentation method for accurately segmenting meningeal lymphatic vessels from a highly cluttered background. The matting based level-set segmentation approach, LyMPhi, is proposed to obtain robust lymphatic vessel segmentation by removing background clutter and retaining vessel smoothness. Essentially, LyMPhi has enabled the segmentation of a recently discovered anatomy (lymph in the brain), allowing quantification of the delicate vessels that vary in width and intensity. LyMPhi is automated, unlike the current manual segmentation used by neuroscientists, or traditional level-set segmentation procedures, which may require difficult-to-tune intensity and scale thresholds. Extensive experiments are conducted on four types of lymphatic vessel datasets to validate the performance of LyMPhi compared to other state-of-the-art segmentation algorithms.

The second deliverable is a shape morphing procedure which creates a rich database of realistic synthetic meningeal lymphatic vessel images. A novel approach to synthetic data augmentation is proposed, based on shape deformation of the real lymphatic vessels. Segmentation on the synthetic data is performed using LyMPhi and other competing methods, with the added benefit of comparison to completely accurate ground truth. Segmentation output from LyMPhi has a higher overall Dice coefficient compared to that of competing algorithms as well as a higher BF-score, whilst being stable under different initial conditions.

Third, a convolutional neural network was applied to the MLV images, which is a challenging problem considering the lack of microscopic data with labels of the MLV. The u-net architecture was used for segmentation, and promising segmentation results are shown compared to more traditional methods. The fourth major contribution is an in-depth analysis of the manual segmentation used, using kappa statistics, showing the fallacy in holding manual segmentation as the gold standard in lymphatic analysis. In the short range, segmentation can be performed automatically, saving human labor and preventing human error between images. Finally, the fifth deliverable is a set of metrics for characterizing meningeal lymphatic vasculature. The LyMPhi segmentation results will be used to perform complexity analysis of the vessels, using well-known complexity measures for other vasculature, as well as developing new metrics which incorporate features unique to lymphatic complexity, such as capillary looping. [168, 169]

In the long-term, studying the segmented vessels in terms of quantity and complexity, may lead to future discoveries on the role our meningeal lymphatic system plays in diseases of the central nervous system. Vessel quantification can be quickly performed using the automated segmentation result, potentially yielding fast and reliable numbers on lymphatic drug impact during future studies. There is hope that the onset of disease could be predicted by studying lymphatic vasculature [3]. If so, drugs can be developed (indeed neuroscientists have already started) to improve vessel function and if not cure diseases like Alzheimer's, then at least mitigate them. Complexity measures built on the reliable segmentation provided by LyMPhi can lead to fundamental understanding of these vessels: how they grow, how they decline, and ultimately, the intricacies of their function.

References

- [1] S. Da Mesquita, A. Louveau, A. Vaccari, I. Smirnov, R. C. Cornelison, K. M. Kingsmore, C. Contarino, S. Onengut-Gumuscu, E. Farber, D. Raper, *et al.*, “Functional aspects of meningeal lymphatics in ageing and alzheimer’s disease,” *Nature*, vol. 560, no. 7717, p. 185, 2018.
- [2] A. Louveau, J. Herz, M. N. Alme, A. F. Salvador, M. Q. Dong, K. E. Viar, S. G. Herod, J. Knopp, J. C. Setliff, A. L. Lupi, *et al.*, “Cns lymphatic drainage and neuroinflammation are regulated by meningeal lymphatic vasculature,” *Nature neuroscience*, vol. 21, no. 10, p. 1380, 2018.
- [3] A. Louveau, T. H. Harris, and J. Kipnis, “Revisiting the mechanisms of cns immune privilege,” *Trends in immunology*, vol. 36, no. 10, pp. 569–577, 2015.
- [4] C. Kervrann, C. O. S. Sorzano, S. T. Acton, J. C. Olivo-Marin, and M. Unser, “A guided tour of selected image processing and analysis methods for fluorescence and electron microscopy,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, pp. 6–30, 2016.
- [5] Y. Cheng, N. Grigorieff, P. A. Penczek, and T. Walz, “A primer to single-particle cryo-electron microscopy,” *Cell*, vol. 161, pp. 438–449, 2015.
- [6] G. C. Lander, *Structural Architecture of Bacteriophage Described Via Streamlined Cryo-electron Microscopy: A Thesis Presented*. PhD thesis, Scripps Research Institute, La Jolla, California, 2009.
- [7] A. McDowell, J.-J. Chang, R. Freeman, J. Lepault, C. Walter, and J. Dubochet, “Electron microscopy of frozen hydrated sections of vitreous ice and vitrified biological samples,” *Journal of microscopy*, vol. 131, no. 1, pp. 1–9, 1983.
- [8] L. Wang, A. Singer, and Z. Wen, “Orientation determination of cryo-EM images using least unsquared deviations,” *SIAM Journal on Imaging Sciences*, vol. 6, pp. 2450–2483, 2013.
- [9] D. Elmlund and H. Elmlund, “SIMPLE: software for *ab initio* reconstruction of heterogeneous single-particles,” *Journal of Structural Biology*, vol. 180, pp. 420–427, 2012.
- [10] E. Levin, T. Bendory, N. Boumal, J. Kileel, and A. Singer, “3D *ab initio* modeling in cryo-EM by autocorrelation analysis,” in *2018 IEEE 15th International Symposium on Biomedical Imaging*, pp. 1569–1573, April 2018.
- [11] Y. Shkolnisky and A. Singer, “Viewing direction estimation in cryo-EM using synchronization,” *SIAM Journal on Imaging Sciences*, vol. 5, pp. 1088–1110, 2012.
- [12] H. Elmlund, D. Elmlund, and S. Bengio, “PRIME: probabilistic initial 3D model generation for single-particle cryo-electron microscopy,” *Structure*, vol. 21, pp. 1299–1306, 2013.

- [13] S. H. W. Scheres, “RELION: implementation of a Bayesian approach to cryo-EM structure determination,” *Journal of Structural Biology*, vol. 180, pp. 519–530, 2012.
- [14] D. Si, S. Ji, K. A. Nasr, and J. He, “A machine learning approach for the identification of protein secondary structure elements from electron cryo-microscopy density maps,” *Biopolymers*, vol. 97, pp. 698–708, 2012.
- [15] J. Esquivel-Rodríguez and D. Kihara, “Computational methods for constructing protein structure models from 3D electron microscopy maps,” *Journal of Structural Biology*, vol. 184, pp. 93–102, 2013.
- [16] N. Chenouard, I. Smal, F. De Chaumont, M. Maška, I. F. Sbalzarini, Y. Gong, J. Cardinale, C. Carthel, S. Coraluppi, M. Winter, *et al.*, “Objective comparison of particle tracking methods,” *Nature methods*, vol. 11, no. 3, pp. 281–289, 2014.
- [17] E. Meijering, I. Smal, and G. Danuser, “Tracking in molecular bioimaging,” *IEEE signal processing magazine*, vol. 23, no. 3, pp. 46–53, 2006.
- [18] E. Meijering, O. Dzyubachyk, and I. Smal, “Methods for cell and particle tracking,” in *Methods in enzymology*, vol. 504, pp. 183–200, Elsevier, 2012.
- [19] I. Smal and E. Meijering, “Quantitative comparison of multiframe data association techniques for particle tracking in time-lapse fluorescence microscopy,” *Medical image analysis*, vol. 24, no. 1, pp. 163–189, 2015.
- [20] B. Rieger, R. Nieuwenhuizen, and S. Stallinga, “Image processing and analysis for single-molecule localization microscopy: Computation for nanoscale imaging,” *IEEE Signal Processing Magazine*, vol. 32, no. 1, pp. 49–57, 2014.
- [21] R. J. Ober, A. Tahmasbi, S. Ram, Z. Lin, and E. S. Ward, “Quantitative aspects of single-molecule microscopy: Information-theoretic analysis of single-molecule data,” *IEEE signal processing magazine*, vol. 32, no. 1, pp. 58–69, 2014.
- [22] R. Delgado-Gonzalo, V. Uhlmann, D. Schmitter, and M. Unser, “Snakes on a plane: A perfect snap for bioimage analysis,” *IEEE Signal Processing Magazine*, vol. 32, no. 1, pp. 41–48, 2014.
- [23] Z. Pincus and J. Theriot, “Comparison of quantitative methods for cell-shape analysis,” *Journal of microscopy*, vol. 227, no. 2, pp. 140–156, 2007.
- [24] E. Meijering, “Neuron tracing in perspective,” *Cytometry Part A*, vol. 77, no. 7, pp. 693–704, 2010.
- [25] P. Sarder and A. Nehorai, “Deconvolution methods for 3-d fluorescence microscopy images,” *IEEE Signal Processing Magazine*, vol. 23, no. 3, pp. 32–45, 2006.
- [26] I. Smal, M. Loog, W. Niessen, and E. Meijering, “Quantitative comparison of spot detection methods in fluorescence microscopy,” *IEEE transactions on medical imaging*, vol. 29, no. 2, pp. 282–301, 2009.

- [27] X. Wang, W. E. Allen, M. A. Wright, E. L. Sylwestrak, N. Samusik, S. Vesuna, K. Evans, C. Liu, C. Ramakrishnan, J. Liu, *et al.*, “Three-dimensional intact-tissue sequencing of single-cell transcriptional states,” *Science*, vol. 361, no. 6400, p. eaat5691, 2018.
- [28] J. Chen and B. G. Condron, “Branch architecture of the fly larval abdominal serotonergic neurons,” *Developmental Biology*, vol. 320, no. 1, pp. 30 – 38, 2008.
- [29] S. W. Hell and J. Wichmann, “Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy,” *Optics letters*, vol. 19, no. 11, pp. 780–782, 1994.
- [30] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, “Imaging intracellular fluorescent proteins at nanometer resolution,” *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.
- [31] M. J. Rust, M. Bates, and X. Zhuang, “Stochastic optical reconstruction microscopy (storm) provides sub-diffraction-limit image resolution,” *Nature methods*, vol. 3, no. 10, p. 793, 2006.
- [32] M. G. Gustafsson, “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy,” *Journal of microscopy*, vol. 198, no. 2, pp. 82–87, 2000.
- [33] M. G. Gustafsson, L. Shao, P. M. Carlton, C. R. Wang, I. N. Golubovskaya, W. Z. Cande, D. A. Agard, and J. W. Sedat, “Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination,” *Biophysical journal*, vol. 94, no. 12, pp. 4957–4970, 2008.
- [34] M. Heilemann, E. Margeat, R. Kasper, M. Sauer, and P. Tinnefeld, “Carbocyanine dyes as efficient reversible single-molecule optical switch,” *Journal of the American Chemical Society*, vol. 127, no. 11, pp. 3801–3806, 2005.
- [35] S. T. Hess, T. P. Girirajan, and M. D. Mason, “Ultra-high resolution imaging by fluorescence photoactivation localization microscopy,” *Biophysical journal*, vol. 91, no. 11, pp. 4258–4272, 2006.
- [36] J. Wang, T. Batabyal, M. Zhang, J. Zhang, A. Aziz, A. Gahlmann, and S. T. Acton, “Lcuts: Linear clustering of bacteria using recursive graph cuts,” in *2019 IEEE International Conference on Image Processing (ICIP)*, pp. 1575–1579, IEEE, 2019.
- [37] H. Shroff, C. G. Galbraith, J. A. Galbraith, and E. Betzig, “Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics,” *Nature methods*, vol. 5, no. 5, pp. 417–423, 2008.
- [38] D. Li, L. Shao, B.-C. Chen, X. Zhang, M. Zhang, B. Moses, D. E. Milkie, J. R. Beach, J. A. Hammer, M. Pasham, *et al.*, “Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics,” *Science*, vol. 349, no. 6251, 2015.

- [39] B.-C. Chen, W. R. Legant, K. Wang, L. Shao, D. E. Milkie, M. W. Davidson, C. Janetopoulos, X. S. Wu, J. A. Hammer, Z. Liu, *et al.*, “Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution,” *Science*, vol. 346, no. 6208, 2014.
- [40] A. Sartori, R. Gatz, F. Beck, A. Rigort, W. Baumeister, and J. M. Plitzko, “Correlative microscopy: bridging the gap between fluorescence light microscopy and cryo-electron tomography,” *Journal of structural biology*, vol. 160, no. 2, pp. 135–145, 2007.
- [41] J. Fink, N. Carpi, T. Betz, A. Bétard, M. Chebah, A. Azioune, M. Bornens, C. Sykes, L. Fetler, D. Cuvelier, *et al.*, “External forces control mitotic spindle positioning,” *Nature cell biology*, vol. 13, no. 7, pp. 771–778, 2011.
- [42] R. M. Parton, R. S. Hamilton, G. Ball, L. Yang, C. F. Cullen, W. Lu, H. Ohkura, and I. Davis, “A par-1-dependent orientation gradient of dynamic microtubules directs posterior cargo transport in the drosophila oocyte,” *Journal of Cell Biology*, vol. 194, no. 1, pp. 121–135, 2011.
- [43] Y. Luo, Y. Zhan, and J. H. Keen, “Arf6 regulation of gyrating-clathrin,” *Traffic*, vol. 14, no. 1, pp. 97–106, 2013.
- [44] N. Tabassum, A. Vaccari, and S. Acton, “Speckle removal and change preservation by distance-driven anisotropic diffusion of synthetic aperture radar temporal stacks,” *Digital Signal Processing*, vol. 74, pp. 43–55, 2018.
- [45] B. Zhang, M. Fadili, J.-L. Starck, and J.-C. Olivo-Marin, “Multiscale variance-stabilizing transform for mixed-poisson-gaussian processes and its applications in bioimaging,” in *2007 IEEE International Conference on Image Processing*, vol. 6, pp. VI–233, IEEE, 2007.
- [46] B. Zhang, J. M. Fadili, and J.-L. Starck, “Wavelets, ridgelets, and curvelets for poisson noise removal,” *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1093–1108, 2008.
- [47] J. Boulanger, C. Kervrann, P. Bouthemy, P. Elbau, J.-B. Sibarita, and J. Salamero, “Patch-based nonlocal functional for denoising fluorescence microscopy image sequences,” *IEEE transactions on medical imaging*, vol. 29, no. 2, pp. 442–454, 2009.
- [48] M. Makitalo and A. Foi, “Optimal inversion of the generalized anscombe transformation for poisson-gaussian noise,” *IEEE transactions on image processing*, vol. 22, no. 1, pp. 91–103, 2012.
- [49] A. Buades, B. Coll, and J.-M. Morel, “A non-local algorithm for image denoising,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, vol. 2, pp. 60–65, IEEE, 2005.
- [50] D. L. Donoho, “De-noising by soft-thresholding,” *IEEE transactions on information theory*, vol. 41, no. 3, pp. 613–627, 1995.

- [51] R. M. Willett and R. D. Nowak, "Platelets: a multiscale approach for recovering edges and surfaces in photon-limited medical imaging," *IEEE Transactions on Medical Imaging*, vol. 22, no. 3, pp. 332–350, 2003.
- [52] F. Luisier, C. Vonesch, T. Blu, and M. Unser, "Fast haar-wavelet denoising of multidimensional fluorescence microscopy data," in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 310–313, IEEE, 2009.
- [53] F. Luisier, C. Vonesch, T. Blu, and M. Unser, "Fast interscale wavelet denoising of poisson-corrupted images," *Signal processing*, vol. 90, no. 2, pp. 415–427, 2010.
- [54] A. Morteza and M. Amirmazlaghani, "A novel statistical approach for multiplicative speckle removal using t-locations scale and non-sub sampled shearlet transform," *Digital Signal Processing*, vol. 107, p. 102857, 2020.
- [55] M. M. Marim, E. D. Angelini, and J.-C. Olivo-Marin, "A compressed sensing approach for biological microscopic image processing," in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1374–1377, IEEE, 2009.
- [56] V. Studer, J. Bobin, M. Chahid, H. S. Mousavi, E. Candes, and M. Dahan, "Compressive fluorescence microscopy for biological and hyperspectral imaging," *Proceedings of the National Academy of Sciences*, vol. 109, no. 26, pp. E1679–E1687, 2012.
- [57] F. Soulez, "A "learn 2d, apply 3d" method for 3d deconvolution microscopy," in *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pp. 1075–1078, IEEE, 2014.
- [58] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, and T. Pock, "An introduction to total variation for image analysis," *Theoretical foundations and numerical methods for sparse recovery*, vol. 9, no. 263-340, p. 227, 2010.
- [59] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [60] Z. Fan, J. Lu, C. Wei, H. Huang, X. Cai, and X. Chen, "A hierarchical image matting model for blood vessel segmentation in fundus images," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2367–2377, 2019.
- [61] D. Thomann, D. R. Rines, P. K. Sorger, and G. Danuser, "Automatic fluorescent tag detection in 3d with super-resolution: application to the analysis of chromosome movement," *Journal of microscopy*, vol. 208, no. 1, pp. 49–64, 2002.
- [62] D. Sage, F. R. Neumann, F. Hediger, S. M. Gasser, and M. Unser, "Automatic tracking of individual fluorescence particles: application to the study of chromosome dynamics," *IEEE Transactions on Image Processing*, vol. 14, no. 9, pp. 1372–1383, 2005.
- [63] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *International conference on medical image computing and computer-assisted intervention*, pp. 130–137, Springer, 1998.

- [64] C. Schläger, H. Körner, M. Krueger, S. Vidoli, M. Haberl, D. Mielke, E. Brylla, T. Issekutz, C. Cabañas, P. J. Nelson, *et al.*, “Effector t-cell trafficking between the leptomeninges and the cerebrospinal fluid,” *Nature*, vol. 530, no. 7590, pp. 349–353, 2016.
- [65] N. Frederick and A. Louveau, “Meningeal lymphatics, immunity and neuroinflammation,” *Current Opinion in Neurobiology*, vol. 62, pp. 41–47, 2020.
- [66] D. Mrdjen, A. Pavlovic, F. J. Hartmann, B. Schreiner, S. G. Utz, B. P. Leung, I. Lelios, F. L. Heppner, J. Kipnis, D. Merkler, *et al.*, “High-dimensional single-cell mapping of central nervous system immune cells reveals distinct myeloid subsets in health, aging, and disease,” *Immunity*, vol. 48, no. 2, pp. 380–395, 2018.
- [67] T. Ly, J. Thompson, T. Harris, and S. T. Acton, “The coupled tuff-bff algorithm for automatic 3d segmentation of microglia,” in *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 121–125, 2018.
- [68] M. Terry, “Novartis, Amgen and Banner end Alzheimer’s program after analysis shows patients’ conditions worsening,” *BioSpace*, 2019.
- [69] P. Visser and B. Tijms, “Brain amyloid pathology and cognitive function: Alzheimer disease without dementia?,” *JAMA Network*, vol. 317, pp. 2285–2287, 2017.
- [70] J. Barney, “UVA brain discovery could block aging’s terrible toll on the mind,” *UVA-Today*, jul 2018.
- [71] S. Cordone, L. Annarumma, P. M. Rossini, and L. De Gennaro, “Sleep and β -amyloid deposition in alzheimer’s disease: Insights on mechanisms and possible innovative treatments,” *Frontiers in Pharmacology*, vol. 10, p. 695, 2019.
- [72] S. Makin, “Deep sleep gives your brain a deep clean,” *Scientific American*, november 2019.
- [73] N. E. Fultz, G. Bonmassar, K. Setsompop, R. A. Stickgold, B. R. Rosen, J. R. Polimeni, and L. D. Lewis, “Coupled electrophysiological, hemodynamic, and cerebrospinal fluid oscillations in human sleep,” *Science*, vol. 366, no. 6465, pp. 628–631, 2019.
- [74] F. Samarrai, “Faculty spotlight: George Bloom seeks to untangle Alzheimer’s riddle,” *UVAToday Research & Discovery*, nov 2019.
- [75] L. K. Rudenko, H. Wallrabe, A. Periasamy, K. H. Siller, Z. Svindrych, M. E. Seward, M. N. Best, and G. S. Bloom, “Intraneuronal tau misfolding induced by extracellular amyloid- β oligomers,” *Journal of Alzheimer’s Disease*, no. Preprint, pp. 1–14, 2019.
- [76] J. Barney, “UVA identifies brain’s lymphatic vessels as new avenue to treat multiple sclerosis,” *UVAToday*, sep 2018.
- [77] J. Barney, “Brain discovery could have important implications for neurodegenerative diseases,” *UVAToday*, april 2020.

- [78] C. R. Lammert, E. L. Frost, C. E. Bellinger, A. C. Bolte, C. A. McKee, M. E. Hurt, M. J. Paysour, H. E. Ennerfelt, and J. R. Lukens, “Aim2 inflammasome surveillance of dna damage shapes neurodevelopment,” *Nature*, vol. 580, no. 7805, pp. 647–652, 2020.
- [79] A. P. Association *et al.*, *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.
- [80] E. Zucchi, V. Bonetto, G. Sorarù, I. Martinelli, P. Parchi, R. Liguori, and J. Mandrioli, “Neurofilaments in motor neuron disorders: towards promising diagnostic and prognostic biomarkers,” *Molecular neurodegeneration*, vol. 15, no. 1, pp. 1–20, 2020.
- [81] N. S. Caron, G. E. Wright, and M. R. Hayden, “Huntington disease,” *GeneReviews®[Internet]*, 2020.
- [82] G. A. Donnan, M. Fisher, M. Macleod, and S. M. Davis, “Stroke,” *The Lancet*, vol. 371, no. 9624, pp. 1612 – 1623, 2008.
- [83] J. J. Iliff, M. Wang, Y. Liao, B. A. Plogg, W. Peng, G. A. Gundersen, H. Benveniste, G. E. Vates, R. Deane, S. A. Goldman, *et al.*, “A paravascular pathway facilitates csf flow through the brain parenchyma and the clearance of interstitial solutes, including amyloid β ,” *Science translational medicine*, vol. 4, no. 147, pp. 147ra111–147ra111, 2012.
- [84] S. Liao and P. von der Weid, “Lymphatic system: an active pathway for immune protection,” *Seminars in Cell and Developmental Biology*, vol. 38, pp. 83–89, 2015.
- [85] E. M. Bouta, R. W. Wood, E. B. Brown, H. Rahimi, C. T. Ritchlin, and E. M. Schwarz, “*In vivo* quantification of lymph viscosity and pressure in lymphatic vessels and draining lymph nodes of arthritic joints in mice,” *The Journal of physiology*, vol. 592, no. 6, pp. 1213–1223, 2014.
- [86] A. Louveau, I. Smirnov, T. Keyes, J. Eccles, S. Rouhani, J. Peske, N. Derecki, D. Castle, J. Mandell, K. Lee, T. Harris, and J. Kipnis, “Structural and functional features of central nervous system lymphatic vessels,” *Nature*, vol. 523, pp. 337–341, 2015.
- [87] A. Aspelund, S. Antila, S. T. Proulx, T. V. Karlsen, S. Karaman, M. Detmar, H. Wiig, and K. Alitalo, “A dural lymphatic vascular system that drains brain interstitial fluid and macromolecules,” *Journal of Experimental Medicine*, vol. 212, no. 7, pp. 991–999, 2015.
- [88] A. Khaffif, S. Schneebaum, D. M. Fliss, H. Lerman, U. Metser, R. Ben-Yosef, Z. Gil, L. Reider-Trejo, L. Genadi, and E. Even-Sapir, “Lymphoscintigraphy for sentinel node mapping using a hybrid single photon emission ct (spect)/ct system in oral cavity squamous cell carcinoma,” *Head & neck*, vol. 28, no. 10, pp. 874–879, 2006.
- [89] D. Singhal, B. N. Tran, J. P. Angelo, B. T. Lee, and S. J. Lin, “Technological advances in lymphatic surgery: bringing to light the invisible,” *Plastic and Reconstructive Surgery*, vol. 143, no. 1, pp. 283–293, 2019.

- [90] K. S. Hershenhouse, O. Shauly, D. J. Gould, and K. M. Patel, “Meningeal lymphatics: A review and future directions from a clinical perspective,” *Neuroscience Insights*, vol. 14, p. 1179069519889027, 2019.
- [91] Y. Yoshimatsu, S. Kimuro, J. Pauty, K. Takagaki, S. Nomiya, A. Inagawa, K. Maeda, K. A. Podyma-Inoue, K. Kajiya, Y. T. Matsunaga, *et al.*, “Tgf-beta and tnf-alpha cooperatively induce mesenchymal transition of lymphatic endothelial cells via activation of activin signals,” *PLoS one*, vol. 15, no. 5, p. e0232356, 2020.
- [92] P. Mascagni and C. Sanctius, *Vasorum lymphaticorum corporis humani historia et ichnographia*. Ex typographia Pazzini Carli, 1787.
- [93] M. Sweeney and B. Zlokovic, “Lymphatic waste disposal in the brain,” *Nature*, vol. 560, pp. 172–174, 2018.
- [94] M. Absinta, S. Ha, G. Nair, P. Sati, N. Luciano, M. Palisoc, A. Louveau, K. Zaghoul, S. Pittaluga, J. Kipnis, and D. Reich, “Human and nonhuman primate meninges harbor lymphatic vessels that can be visualized noninvasively by MRI,” *eLife*, vol. 6, pp. 1–15, 2017.
- [95] S. Antila, S. Karaman, H. Nurmi, M. Airavaara, M. H. Voutilainen, T. Mathivet, D. Chilov, Z. Li, T. Koppinen, J.-H. Park, *et al.*, “Development and plasticity of meningeal lymphatic vessels,” *Journal of Experimental Medicine*, vol. 214, no. 12, pp. 3645–3667, 2017.
- [96] J. H. Ahn, H. Cho, J.-H. Kim, S. H. Kim, J.-S. Ham, I. Park, S. H. Suh, S. P. Hong, J.-H. Song, Y.-K. Hong, *et al.*, “Meningeal lymphatic vessels at the skull base drain cerebrospinal fluid,” *Nature*, vol. 572, no. 7767, pp. 62–66, 2019.
- [97] M. Hsu, A. Rayasam, J. A. Kijak, Y. H. Choi, J. S. Harding, S. A. Marcus, W. J. Karpus, M. Sandor, and Z. Fabry, “Neuroinflammation-induced lymphangiogenesis near the cribriform plate contributes to drainage of cns-derived antigens and immune cells,” *Nature communications*, vol. 10, no. 1, pp. 1–14, 2019.
- [98] S. Osher and J. A. Sethian, “Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations,” *Journal of computational physics*, vol. 79, no. 1, pp. 12–49, 1988.
- [99] S. T. Acton and N. Ray, “Biomedical image analysis: Segmentation,” *Synthesis Lectures on Image, Video, and Multimedia Processing*, vol. 4, no. 1, pp. 1–108, 2009.
- [100] V. Caselles, R. Kimmel, and G. Sapiro, “Geodesic active contours,” *International journal of computer vision*, vol. 22, no. 1, pp. 61–79, 1997.
- [101] Y. Q. Zhao, X. H. Wang, X. F. Wang, and F. Y. Shih, “Retinal vessels segmentation based on level set and region growing,” *Pattern Recognition*, vol. 47, no. 7, pp. 2437–2446, 2014.

- [102] C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding, “Minimization of region-scalable fitting energy for image segmentation,” *IEEE transactions on image processing*, vol. 17, no. 10, pp. 1940–1949, 2008.
- [103] S. Mukherjee and S. Acton, “Region based segmentation in presence of intensity inhomogeneity using legendre polynomials,” *IEEE Signal Processing Letters*, vol. 22, pp. 298–302, 2015.
- [104] S. Mukherjee, B. Condron, and S. T. Acton, “Tubularity flow field—a technique for automatic neuron segmentation,” *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 374–389, 2015.
- [105] S. Rapaka and P. R. Kumar, “Efficient approach for non-ideal iris segmentation using improved particle swarm optimisation-based multilevel thresholding and geodesic active contours,” *IET Image Processing*, vol. 12, no. 10, pp. 1721–1729, 2018.
- [106] B. Li and S. T. Acton, “Active contour external force using vector field convolution for image segmentation,” *IEEE transactions on image processing*, vol. 16, no. 8, pp. 2096–2106, 2007.
- [107] T. F. Chan and L. A. Vese, “Active contours without edges,” *IEEE Transactions on image processing*, vol. 10, no. 2, pp. 266–277, 2001.
- [108] T. Sørensen, “A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons,” *Biol. Skr.*, vol. 5, pp. 1–34, 1948.
- [109] Z. Yan, X. Yang, and K.-T. Cheng, “A three-stage deep learning model for accurate retinal vessel segmentation,” *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1427–1436, 2018.
- [110] G. Csurka, D. Larlus, F. Perronnin, and F. Meylan, “What is a good evaluation measure for semantic segmentation?,” in *BMVC*, vol. 27, p. 2013, 2013.
- [111] B. Hariharan, P. Arbeláez, L. Bourdev, S. Maji, and J. Malik, “Semantic contours from inverse detectors,” in *2011 International Conference on Computer Vision*, pp. 991–998, 2011.
- [112] P. Perona, T. Shiota, and J. Malik, “Anisotropic diffusion,” in *Geometry-driven diffusion in computer vision*, pp. 73–92, Springer, 1994.
- [113] Y. Yu and S. T. Acton, “Speckle reducing anisotropic diffusion,” *IEEE Transactions on image processing*, vol. 11, no. 11, pp. 1260–1270, 2002.
- [114] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Advances in neural information processing systems*, pp. 2672–2680, 2014.

- [115] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.
- [116] P. Luc, C. Couprie, S. Chintala, and J. Verbeek, “Semantic segmentation using adversarial networks,” *arXiv preprint arXiv:1611.08408*, 2016.
- [117] N. Carlini and D. Wagner, “Towards evaluating the robustness of neural networks,” in *2017 IEEE Symposium on Security and Privacy (SP)*, pp. 39–57, IEEE, 2017.
- [118] D. Güera and E. J. Delp, “Deepfake video detection using recurrent neural networks,” in *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–6, IEEE, 2018.
- [119] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.
- [120] C. X. Ren, A. Ziemann, J. Theiler, and A. M. Durieux, “Cycle-consistent adversarial networks for realistic pervasive change generation in remote sensing imagery,” in *2020 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, pp. 42–45, IEEE, 2020.
- [121] Q. Chen and V. Koltun, “Photographic image synthesis with cascaded refinement networks,” in *Proceedings of the IEEE international conference on computer vision*, pp. 1511–1520, 2017.
- [122] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, “High-resolution image synthesis and semantic manipulation with conditional gans,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 8798–8807, 2018.
- [123] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134, 2017.
- [124] P. Costa, A. Galdran, M. I. Meyer, M. Niemeijer, M. Abràmoff, A. M. Mendonça, and A. Campilho, “End-to-end adversarial retinal image synthesis,” *IEEE transactions on medical imaging*, vol. 37, no. 3, pp. 781–791, 2017.
- [125] F. Yang and T. Jiang, “Cell image segmentation with kernel-based dynamic clustering and an ellipsoidal cell shape model,” *Journal of Biomedical Informatics*, vol. 34, no. 2, pp. 67–73, 2001.
- [126] J. Castorena and D. Oyen, “Learning shapes on image sampled points with dynamic graph cnns,” in *2020 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, pp. 1–4, IEEE, 2020.

- [127] X. Han, A. Sit, C. Christoffer, S. Chen, and D. Kihara, “A global map of the protein shape universe,” *PLoS computational biology*, vol. 15, no. 4, p. e1006969, 2019.
- [128] L. Sael, B. Li, D. La, Y. Fang, K. Ramani, R. Rustamov, and D. Kihara, “Fast protein tertiary structure retrieval based on global surface shape similarity,” *Proteins: Structure, Function, and Bioinformatics*, vol. 72, no. 4, pp. 1259–1273, 2008.
- [129] Y. Zhang and J. Skolnick, “Tm-align: a protein structure alignment algorithm based on the tm-score,” *Nucleic acids research*, vol. 33, no. 7, pp. 2302–2309, 2005.
- [130] E. Klassen, A. Srivastava, M. Mio, and S. H. Joshi, “Analysis of planar shapes using geodesic paths on shape spaces,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 3, pp. 372–383, 2004.
- [131] S. Menon, A. Damian, S. Hu, N. Ravi, and C. Rudin, “Pulse: Self-supervised photo upsampling via latent space exploration of generative models,” in *The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- [132] J. Wang, Z. Fu, N. Sadeghradehyazdi, J. Kipnis, and S. T. Acton, “Nonlinear shape regression for filtering segmentation results from calcium imaging,” in *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 738–742, IEEE, 2018.
- [133] A. Srivastava, E. Klassen, S. H. Joshi, and I. H. Jermyn, “Shape analysis of elastic curves in euclidean spaces,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 7, pp. 1415–1428, 2010.
- [134] T. Batabyal, B. Condron, and S. T. Acton, “Neuropath2path: Classification and elastic morphing between neuronal arbors using path-wise similarity,” *Neuroinformatics*, pp. 1–30, 2020.
- [135] A. Telea, C. Sminchisescu, and S. Dickinson, “Optimal inference for hierarchical skeleton abstraction,” in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, vol. 4, pp. 19–22, IEEE, 2004.
- [136] M. Zhang, J. Zhang, Y. Wang, J. Wang, A. M. Achimovich, S. T. Acton, and A. Gahlmann, “Non-invasive single-cell morphometry in living bacterial biofilms,” *bioRxiv*, 2020.
- [137] K. R. Spring, T. J. Fellers, and M. W. Davidson, “Confocal microscopy - signal-to-noise considerations,” 2020.
- [138] C. Yang, S. Baireddy, Y. Chen, E. Cai, D. Caldwell, V. Méline, A. S. Iyer-Pascuzzi, and E. J. Delp, “Plant stem segmentation using fast ground truth generation,” in *2020 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, pp. 62–65, IEEE, 2020.
- [139] O. Bailo, D. Ham, and Y. Min Shin, “Red blood cell image generation for data augmentation using conditional generative adversarial networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 0–0, 2019.

- [140] zhixuhao, “unet for image segmentation.” GitHub, 2019.
- [141] C. Chen, Q. Chen, J. Xu, and V. Koltun, “Learning to see in the dark,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3291–3300, 2018.
- [142] Z. Ying, H. Niu, P. Gupta, D. Mahajan, D. Ghadiyaram, and A. Bovik, “From patches to pictures (paq-2-piq): Mapping the perceptual space of picture quality,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3575–3585, 2020.
- [143] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal loss for dense object detection,” *IEEE transactions on pattern analysis and machine intelligence*, 2018.
- [144] M. AlGhamdi, “Optic disc segmentation in fundus images with deep learning object detector,” *Journal of Computer Science*, vol. 16, pp. 591–600, May. 2020.
- [145] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [146] R. Girshick, “Fast r-cnn,” in *Proceedings of the IEEE international conference on computer vision*, pp. 1440–1448, 2015.
- [147] A. Dosovitskiy and J. Djolonga, “You only train once: Loss-conditional training of deep networks,” in *ICLR*, May 2020.
- [148] R. Pascanu, T. Mikolov, and Y. Bengio, “On the difficulty of training recurrent neural networks,” in *International conference on machine learning*, pp. 1310–1318, 2013.
- [149] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, “Comparing images using the hausdorff distance,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 15, no. 9, pp. 850–863, 1993.
- [150] N. Tajbakhsh, L. Jeyaseelan, Q. Li, J. N. Chiang, Z. Wu, and X. Ding, “Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation,” *Medical Image Analysis*, p. 101693, 2020.
- [151] A. Jamaludin, T. Kadir, and A. Zisserman, “Self-supervised learning for spinal mris,” in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 294–302, Springer, 2017.
- [152] G. Marcus and E. Davis, “How to build artificial intelligence we can trust,” *The New York Times Opinion*, september 2019.
- [153] J. L. Fleiss and J. Cohen, “The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability,” *Educational and psychological measurement*, vol. 33, no. 3, pp. 613–619, 1973.

- [154] J. R. Landis and G. G. Koch, “The measurement of observer agreement for categorical data,” *biometrics*, pp. 159–174, 1977.
- [155] M. L. McHugh, “Interrater reliability: the kappa statistic,” *Biochemia medica: Biochemia medica*, vol. 22, no. 3, pp. 276–282, 2012.
- [156] S. K. Warfield, K. H. Zou, and W. M. Wells, “Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation,” *IEEE transactions on medical imaging*, vol. 23, no. 7, pp. 903–921, 2004.
- [157] J. Yerushalmy, “Statistical problems in assessing methods of medical diagnosis, with special reference to x-ray techniques,” *Public Health Reports (1896-1970)*, pp. 1432–1449, 1947.
- [158] H. Silvennoinen, S. Ikonen, L. Soinne, M. Railo, and L. Valanne, “Ct angiographic analysis of carotid artery stenosis: Comparison of manual assessment, semiautomatic vessel analysis, and digital subtraction angiography,” *American Journal of Neuroradiology*, vol. 28, pp. 97–103, 2007.
- [159] D. Oyen, “Evaluation of image signatures for diagram image retrieval,” in *The Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, March 2020.
- [160] S. Basu, W. Ooi, and D. Racoceanu, “Neurite tracing with object process,” *IEEE Transactions on Medical Imaging*, vol. 35, pp. 1443–1451, 2016.
- [161] M. Rubinov and O. Sporns, “Complex network measures of brain connectivity: Uses and interpretations,” *NeuroImage*, vol. 52, pp. 1059–1069, 2010.
- [162] C. Madry, V. Kyrargyri, I. Arancibia-Cárcamo, R. Jolivet, S. Kohsaka, R. Bryan, and D. Attwell, “Microglial ramification, surveillance, and Interleukin-1 β release are regulated by the two-pore domain K^+ channel THIK-1,” *Neuron*, vol. 97, pp. 299–312.e6, jan 2018.
- [163] R. Sarkar, S. Mukherjee, and S. T. Acton, “Shape descriptors based on compressed sensing with application to neuron matching,” in *2013 Asilomar Conference on Signals, Systems and Computers*, pp. 970–974, IEEE, 2013.
- [164] G. G. Stokes, “On the theories of the internal friction of fluids in motion, and of the equilibrium and motion of elastic solids,” *Transactions of the Cambridge Philosophical Society*, vol. 8, pp. 287–341, 1845.
- [165] F. M. White and I. Corfield, *Viscous fluid flow*, vol. 3. McGraw-Hill New York, 2006.
- [166] J. Brittain, “Thevenin’s theorem,” *Ieee Spectrum*, vol. 27, no. 3, p. 42, 1990.
- [167] Z. Jiang, L. Wu, B. L. Miller, D. R. Goldman, C. M. Fernandez, Z. S. Abouhamze, C. K. Ozaki, and S. A. Berceci, “A novel vein graft model: adaptation to differential flow environments,” *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 286, no. 1, pp. H240–H245, 2004.

- [168] N. Tabassum, M. Ferguson, J. Herz, and S. T. Acton, “Complexity analysis and u-net based segmentation of meningeal lymphatic vessels,” in *2020 54th Asilomar Conference on Signals, Systems and Computers*, IEEE, 2020.
- [169] N. Tabassum, J. Wang, M. Ferguson, J. Herz, M. Dong, A. Louveau, J. Kipnis, and S. T. Acton, “Image analysis for neuroscience: lymphatics,” *JPhys Phot*, 2020.