THESIS

AUTOMATED METHODS FOR QUANTIFYING THE TORTUOSITY OF

MICROVASCULAR NETWORKS

Submitted by

Melody Dodd

Department of Mathematics

In partial fulfillment of the requirements

For the degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2012

Master's Committee:

    Advisor: Vakhtang Putkaradze

    Stuart Tobet
    Patrick Shipman

ABSTRACT

AUTOMATED METHODS FOR QUANTIFYING THE TORTUOSITY OF

MICROVASCULAR NETWORKS

Networks of microscopic blood vessels can be studied for changes in morphology that correlate with biological abnormalities. Tortuosity, or vessel twistiness, is one of these morphological properties, and it can be surprisingly difficult to quantify. The purpose of this thesis is to present the development, testing, and analysis of new automated methods to measure and quantify the tortuosity of microvascular networks. We will explain necessary automated image processing techniques and background information before presenting our new metrics for measuring network tortuosity. Experiments using the methods will be presented, including a full analysis of the results. We will use the results from these experiments to justify our final conclusions and recommendations regarding the performance of the methods.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

The quantification of various morphological properties of blood vessels and vessel networks has in recent decades become of vast importance in the medical and biological sciences. While tortuosity is of primary interest in this thesis, characteristics such as vascular density, length, and branching have also been studied extensively, as researchers and physicians become increasingly aware that these features can provide insight into various biological and pathological processes. Abnormalities in these features can be indicators of disease, while variation between experimental groups can signify underlying biological differences. Vessel tortuosity, which can be thought of as a measure of "twistiness," in particular has earned much attention in the biological community over the past three decades.

It has long been known, for example, that increased tortuosity and density of retinal vasculature is an indicator for diseases such as hypertension, diabetes, arteriosclerosis, retinopathy of prematurity, plus disease, and other retinopathies [8, 9, 13, 15, 17, 18, 20, 23, 28, 30]. Increased tortuosity of the larger arteries such as the femoral artery or aorta has been shown to correlate with athlerosclerotic changes and hypertension [11, 16, 24], while tortuosity changes in capillaries have been used to diagnose rheumatic diseases [21]. The tortuosity of vascular networks within cancers and other tumors can provide great insight into tumor growth and treatment options [4, 10], and biological changes associated with aging are often manifested in vascular tortuosity [6, 29]. Furthermore, recent research has concentrated on features of microvascular networks in the brain as indicators of neurological diseases [12, 22].

It is these microvascular networks that are indeed the focus of this thesis, and more will be explained regarding their features and significance.

Much of the literature regarding vascular tortuosity is concerned with not only how tortuosity correlates with the aforementioned biological changes, but how best to quantify vascular tortuosity, as there is no universally agreed-upon method. Tortuosity, it turns out, can be surprisingly difficult to measure and quantify. This is often the case with morphological properties of vasculature in general; while various properties are usually quite tangible and often easily seen with the eye, it can be difficult for a human to precisely and objectively quantify such features. Moreover, while the human eye can often detect that two images differ in some way, it is often less clear how to define and measure those differences. These difficulties are increased considerably when one wishes to quantify features of a network of microvessels, as opposed to those of a single large vessel such as the aorta.

As an illustration, consider the images of microvascular networks presented in Figure 1.1. Even untrained observers will usually agree that the two images exhibit significant differences. What exact morphological properties of the vessels contribute to this disparity? Is it mainly due to a difference in vessel density, thickness, or perhaps the number of branch points? Is there any difference in the length of the vessel segments? Most importantly for us, is there a difference in tortuosity? If so, how would we measure that difference?
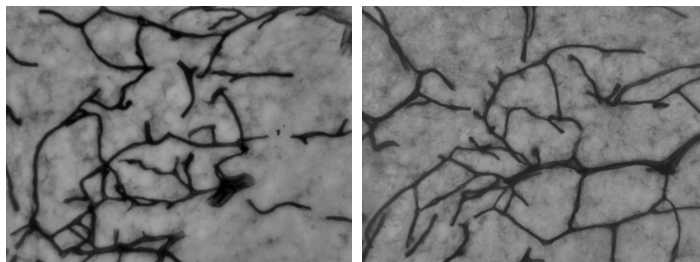

Figure 1.1: Two example images of microvascular networks.

It seems there are several features that contribute simultaneously to the overall effect, and it becomes unclear how best to accurately measure the difference between the two images. The fact is, the human eye takes in many features, including vessel density, thickness, branching, and tortuosity, at the same time. This information is then processed by the brain in lightning speed to reach an almost instantaneous conclusion about the two images. This process is remarkable in many ways, but it is highly subjective and imprecise. The eye is easily distracted by features of the images that have nothing to do with the morphology of the vessels, such as differences in lighting, image contrast, sharpness, background features, and noise. Since the image on the left has darker-colored vessels, this may inadvertently affect our interpretation of their shape. Our eyes are not very good at separating out one individual feature, such as tortuosity, for objective evaluation. Furthermore, while it may be possible to make some qualitative statements about the vascular networks based on visual perception alone, it would be nearly impossible to assign a concrete quantitative value to our conclusions.

This example was meant to illustrate some of the difficulties associated with visually comparing differences between images of vascular networks. How much more complicated do things become when the goal is to quantify a great many images, to contrast differences between two groups in an experiment, for example? To appreciate this challenge, consider the set of twelve images shown in Figure 1.2. Is is possible, using visual perception alone, to make concrete statements about the tortuosity differences among the images? Can the images be ranked or classified in some way, based on their tortuosity? Now imagine performing this task for dozens or hundreds of images. Assuming a person is able to do this to his or her own satisfaction, this classification is highly subjective; it is likely that a different individual may draw markedly different conclusions. Furthermore, it is completely unclear what this classification quantifies. Since tortuosity differences can reveal important information

about biological properties, it is therefore necessary to obtain well-defined metrics to quantify and categorize this property of microvascular networks.
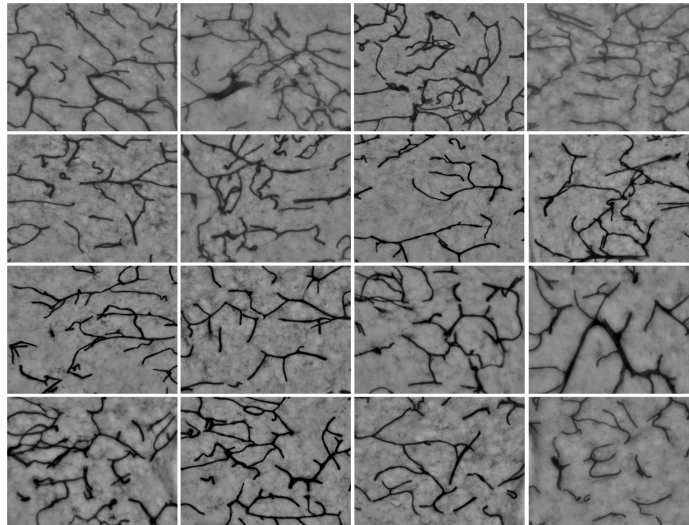

Figure 1.2: Twelve example images of microvascular networks.

Once precise metrics are defined, the next obvious problem is how to best obtain the necessary measurements. If the goal is to measure features of a large number of images, it becomes highly desirable to automate as much of this process as possible using computerized techniques. This not only speeds up the process tremendously, but ensures that human error is largely eliminated from the process of obtaining measurements.

Finally, once metrics have been established, algorithms have been implemented, and data has been collected, processed, and organized, the question remains: does the quantification serve the intended purpose? In other words, does it measure what we want it to measure, and is it robust enough to be useful in a large number of situations? We will provide a detailed assessment of whether the research explained in this thesis was successful, but at this point we feel confident that these goals were largely achieved.

Chapter 2

OVERVIEW OF THE PROBLEM

2.1 Goals and Desired Outcomes

The primary focus of this research project was to quantify the tortuosity of microvascular networks in and around the brain region known as the paraventricular nucleus (PVN) of the hypothalamus in mice. As summarized by Frahm, Schow, and Tobet in [12], the PVN is involved in many critical biological functions. These include hunger, stress responses, energy balance, cardiovascular and neuroendocrine function, and nervous system regulation. Neurons within the PVN contain many important chemical messenger molecules known as neuropeptides, which are known to influence various crucial life-sustaining functions as well as specific behaviors. Additionally, receptors for various critical hormones and neurotransmitters are concentrated within the PVN [12].

Currently, analysis of the PVN has been limited to the clustering of neurons that characterizes this region, rather than the vasculature within the PVN [12]. However, new research has brought to light important and revealing links between the nervous and vascular systems, referred to as the neurovascular unit. Quaegebeur, et al., for example, authored a 2011 paper explaining how "neurovascular crosstalk," or communication between the neurons and vasculature, plays an integral role in the function of both systems. The study suggests that, as blood vessels provide oxygen and nutrients to sustain proper neuronal function, reduction in neuronal access to microvessels may result in abnormalities in the neurons and hence neurological disorders [22].

Frahm, Schow, and Tobet are therefore engaged in an ongoing project to characterize the dense vasculature in and around the PVN, and to discover how changes in vascular morphology correlate with disease. It is hoped that these advances may reveal important insights into neurological diseases such as major depressive disorder, and how prenatal stress can cause hypertension and other problems in offspring. In order to facilitate these goals, large numbers of images of the vasculature in the PVN must be examined for morphological differences [12].

Our main goal was therefore to provide tortuosity metrics that could indicate biological differences between groups of animals used in these studies. For example, it is desirable to establish measurable differences between a group of experimentally treated animals and a control group. Often the samples collected from these groups will be visibly different from each other, but quantifiable validation is needed to confirm these observations. Furthermore, it is necessary for the process of collecting measurements to be automated as much as possible due to the large number of samples collected.

In a study conducted by Frahm, Schow, and Tobet and described in the paper "The Vasculature Within the Paraventricular Nucleus of the Hypothalamus in Mice Varies as a Function of Development, Subnuclear Location and GABA Signaling," (see [12]), total vessel length and branch point counts were used to quantify differences in the PVN and surrounding brain tissue between two groups of mice. One group (designated the "KO" group) was genetically modified so as to "knock out," or inhibit, the development of gamma aminobutyric acid (GABA) receptors in the PVN, specifically $GABA_B$. GABA is the chief inhibitory neurotransmitter in adults, although in early life, GABA plays an important role in the formation of various structures in the brain, including neurons. A normal, healthy animal has a dense concentration of $GABA_B$ receptors within the PVN, and mice lacking these receptors are known to exhibit neuronal changes. Since disruption in GABA signaling during

development can have long-term consequences for both physical and mental health, it is important to gain insight into the mechanisms behind these changes [12].

In the study, the experimental KO group was compared against a control "wild type" (WT) group, and sample images of microvasculature were collected throughout the PVN and surrounding cortex at various stages of development. To quantify differences between the experimental and control groups, branch points were counted manually in the sample images, and total vessel length was obtained using the commercially-available Angiogenesis Tube Formation module in the Metamorph software suite. Total vessel length is a property that can be used as a measure of vascular density. The conclusion was that the KO group had decreased vessel length in the PVN and cortex as compared to the WT group, as well as decreased vessel branching in the mid PVN region. This is important since these changes may indicate links between the PVN neurons and vasculature [12].

Although these quantification methods did provide statistically significant results that indicate a clear change in vessel morphology as a result of inhibited $GABA_B$ signaling, vessel tortuosity was not measured for this study. At the time, there were no methods immediately available to the researchers to measure or quantify this property.

Frahm, Schow, and Tobet are currently engaged in another study of vascular changes in the PVN, which examines the action of glucocorticoids on prenatal development and the resulting long-term implications for the offspring. It is known, for example, that prenatal stress is linked to hypertension in the offspring later in life, although the mechanisms behind this process are not fully understood (see [19]). It has also been found that links exist between neuronal function in the PVN and the regulation of blood pressure (see [26]). Therefore, Frahm, Schow, and Tobet wish to study how the vasculature of the PVN relates to this phenomenon as well as other diseases that correlate with prenatal stress.

In one experiment conducted by the research team, pregnant mice were injected with dexamethasone, a synthetic glucocorticoid steroid, to test whether this treatment would lead to changes in vascular density and tortuosity within the PVN of the offspring. A control group of pregnant mice were injected with an inert vehicle. Exposure to dexamethasone creates a situation in the body that simulates cases of prenatal stress in which glucocorticoid levels are elevated. In both groups, the brains of the offspring were prepared, viewed, and imaged to look for vascular changes within the PVN.

Experiments such as this create changes in the vascular morphology of the PVN. While some of these changes can be measured using total vessel length and branch point counts, the researchers involved wanted to measure tortuosity differences as well. It was their desire, therefore, to develop one or more quantitative metrics to measure vascular tortuosity, in which an index could describe the composite tortuosity of all the vessels within an image. It was further desired to automate the measurement techniques to minimize human labor, and to then apply the metrics to the images obtained from this and other experiments. It was hoped that these methods could measure average tortuosity differences between groups of animals used in these studies, and therefore provide biological insights that would otherwise be hidden.

2.2 Description of the Images

The images used throughout this project were obtained from Frahm, Schow, and Tobet, and were 2-dimensional light-illuminated photomicrographs of regions in the PVN and the surrounding cortex. Please refer to [12] for more information on the exact methods and preparations used to prepare samples and create the images. The images from the GABA study, which were used for testing and validation purposes throughout this project, were taken using a 40× objective, and represented brain

regions of size $300\mu m \times 224\mu m$. These images were originally stored as 1600×1200 pixel RGB images in TIFF format. The images from the dexamethasone study were taken using a $10\times$ objective, and were originally 1600×1200 pixels, but were cropped to a region of interest resulting in an image of size 616×960 pixels. Images were then converted to grayscale and light corrected using Image J prior to further processing, thus reducing problems caused by uneven illumination throughout the images.

Some key points should be noted concerning the limitations of the images used. First, we note that microvascular networks in the brain are fully 3-dimensional structures, while the prepared samples are thin slices of these structures. Furthermore, since each brain slice is thicker than one vessel width, each image represents a 2-dimensional projection of the 3-dimensional structures within each slice. There are also inherent limitations with the depth-of-focus of the microscope, so that some vessels may appear out of focus, or may simply not be visible. Given the methods necessary to obtain lasting, durable samples of the brain regions, the extremely small size of the microvessels, and the fact that the vasculature differs in size throughout the anatomy of the PVN, it was deemed impractical and not cost-effective to prepare and analyze 3-dimensional representations of the PVN.

Therefore, vessels that appear to end in the images may actually continue outside the plane of focus or outside the slice. Other vessels may run perpendicular to the slice, so that we see only their cross-sectional area, and they appear as roughly circular shapes or tiny disconnected segments. Additionally, it can sometimes be difficult for even human eyes to distinguish true branch points from vessel "cross-overs," defined as points where two vessels overlap in the image and appear to intersect each other. These cross-over points are the result of the 2-dimensional projection effect discussed earlier. Examples of these features can be readily seen upon inspection of the images in Figure 1.2. These limitations do introduce unavoidable errors into the analysis,

although for cost-effective analysis this error is deemed acceptable, and we can indeed still determine a great deal about the properties of the vascular networks.

2.3 Comparison to Other Studies

To the best of our knowledge, there is little or no information in the literature on how best to quantify the tortuosity of similar microvascular networks in 2-dimensional images. Several well-established methods exist for quantifying the 2- or 3-dimensional tortuosity and other features of major arteries (see [11, 16, 23, 24], for example). However, characterizing these properties for a single large vessel is quite a different problem than ours, for multiple obvious reasons.
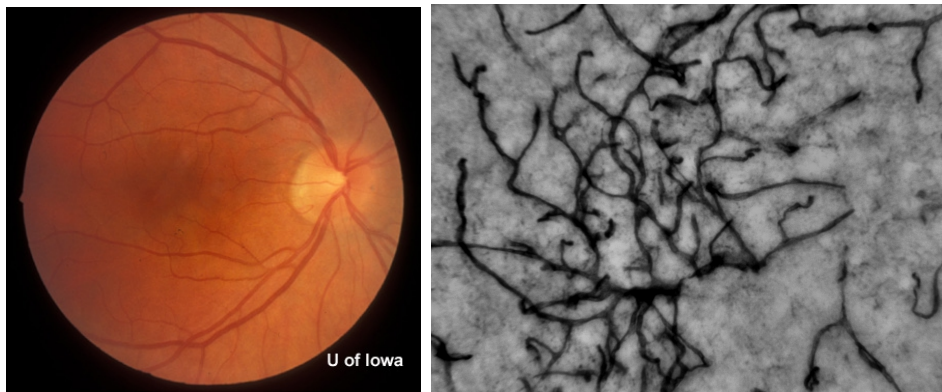

Figure 2.1: Vasculature of the human retina (left) vs. microvessels within the PVN of a mouse (right). Eye fundus image from: J.M. Graff and E.M. Stone; "Unilateral Retinitis Pigmentosa: Visual field changes in a 31 year old woman;" EyeRounds.org

As mentioned in Chapter 1, there is a great wealth of information concerning the quantification of properties (especially tortuosity) of retinal vasculature. However, retinal vessel networks are fundamentally dissimilar to the microvasculature of the PVN and other brain regions, as can be seen in Figure 2.1. The retinal vessels are less dense and have more clearly defined beginning and ending points. They exist mainly in the fairly 2-dimensional "shell" of the eye fundus, so that the image better captures the true structure of the network. Indeed, it is easy to capture an entire

vessel as it travels outward from the optic nerve, as opposed to the disjointed vessel pieces that tend to occur in the PVN images. Additionally, the retinal vessels have a clean, tree-like structure, extending radially from the optic disc, and branching occurs in a somewhat predictable manner. There are far fewer cross-overs, at least of the larger vessels, and vessels have a more consistent and predictable shape. The notion of tortuosity, even, seems much clearer in the retinal vessels, which twist and turn in a gentle, sinusoidal pattern. Compare all these features to the messier, "spaghetti-like" quality of the PVN vasculature, and it is not hard to imagine the difficulties that might occur.

Additionally, many of the techniques outlined in the literature are not completely automated. Many papers describe manual extraction of the vessel midline by tracing, as in [3, 8, 13, 18], or involve manually determining end points of vessel segments in the image, as in [1, 3, 13, 8, 18]. Even in those manuscripts in which completely automated vessel extraction is described, the focus tends to be on determining properties of individual vessels, rather than an entire vessel network.

Another area of intense interest in recent literature has been properties of vascular networks based on 3-dimensional reconstructions obtained through MRA or other scans. This is the case in the work of Elizabeth Bullit et al. [4, 5, 6], and while these methods no doubt provide greatly improved biological insight, for the reasons outlined earlier, 2-dimensional imaging was deemed the best choice for this project.

Chapter 3

VESSEL EXTRACTION AND SEGMENTATION

3.1 Digital Image Representation

In order to take desired measurements from images, a series of digital image processing, vessel-extraction, and segmentation algorithms must be applied. This chapter outlines these steps in detail, and for the benefit of those unfamiliar with computer representations of digital images, we present here an overview of this topic.

A digital image is created when a sensor records a discrete representation of intensity information in response to a continuous signal. Since the images used in this study were taken using a light-illuminated microscope and a digital camera, the signal is visible light, and the sensor is the camera. The sensor discretizes the continuous signal into a rectangular mesh of units known as *pixels* that record intensity, or color, information.

This rectangular mesh is typically stored in computer memory as a matrix of numerical values, where each value corresponds to an intensity level. In the case of a grayscale image, a 2-dimensional matrix of size $m \times n$ contains integer values ranging from 0 (black) to 255 (white), with values between these two extremes representing various shades of gray. An even simpler case is a *binary* image, in which pixels are designated as either "on," given by 1, or "off," given by 0. We typically take the "turned-on" pixels to be the foreground, or area of interest, in the image, while the background of zeros is ignored.

Color images can be represented in various formats, but RGB is common, and RGB is the format in which the images used in this study were originally stored. An

RGB image has 3 color *channels*: red, green, and blue, which additively form other colors in the visible spectrum. Each channel is stored in its own $m \times n$ matrix, so that the data structure containing the entire image is a 3-dimensional matrix of size $m \times n \times 3$. Integer values in each plane of the matrix represent intensity levels for the corresponding color channel. Since we convert images to grayscale prior to further processing, we will not further discuss RGB images. Conversion to grayscale can be accomplished using freely-available scripts or commercial image-processing software.

Throughout this paper we will use the following convention: given a 2-dimensional matrix A of size $m \times n$ representing a digital image, we designate by $A(1, 1)$ the entry at the upper-left corner of the matrix, which corresponds to the pixel at the upper-left corner of the image. Then the pixel in the ith row, j^{th} column of the image, measured from the upper-left corner, corresponds to $A(i, j)$.

Images used in this project had typically been previously light-corrected using Image J at the time they were recieved from the research team. This step may mitigate processing issues caused by uneven illumination, but experimentation with "raw" images that had not been so altered also yielded good results. All subsequent processing was done using algorithms written in Matlab. It should be noted that we did not use the Matlab Image Processing Toolbox, making it necessary to generate code for several basic image processing algorithms. We will therefore briefly outline some of these algorithms within this chapter, as the ones we used may give slightly different results than algorithms in the Image Processing Toolbox.

3.2   Thresholding

Thresholding is an operation performed on the individual pixels of an image to isolate the foreground "pixels of interest" from the background. In our case, we are interested in isolating the darker-colored vessels in an image while discarding the

features of the surrounding lighter-colored brain tissue. In a grayscale representation of a vessel network, the pixel values of the vessels will therefore be closer to 0 than the pixel values corresponding to the background. Thresholding works on the assumption that the foreground pixels in a grayscale image have significantly different pixel values than the surrounding background, and can therefore be isolated by use of a threshold value. For the simplest case, the threshold value T is chosen by the user, and we perform the steps shown in Algorithm 1.

Algorithm 1 Simplest Thresholding Algorithm

1: **for all** i, j **do**
2:      **if** $A(i,j) > T$ **then**
3:          $A(i,j) \leftarrow 0$
4:      **else**
5:          $A(i,j) \leftarrow 1$
6:      **end if**
7: **end for**

This can, of course, be done quite easily in Matlab. However, choosing the threshold value T is not always a trivial matter, and often requires fine-tuning with a human eye. Additionally, when batch-processing dozens or hundreds of images, it is highly unlikely that the same threshold value will work for all images in the batch. Moreover, many images exhibit significant contrast differences among regions within the image, so that a single image may require multiple threshold values across different regions. One partial solution is to have a human manually threshold all images in the batch, using commercially-available image processing software to select regions and appropriate threshold values. This is, of course, very labor-intensive, and we wish to avoid this as much as possible.

Another partial solution is the use of iterative thresholding algorithms with sub-divided processing, which proved useful for many of the images used in this project. In this algorithm, the image matrix is first subdivided into $k \times l$ sub-matrices, and each sub-matrix $B_{k,l}$ is processed individually to find an optimum threshold value.

The values k and l should be chosen to match the nature of the images; for example, if it is known that most images in a batch exhibit a purely vertical contrast gradient, we might choose $k = 3, l = 1$. For more general-purpose applications, we could choose $k = l = 3$. If k or l is too large, the thresholded image will contain artifacts on the edges of the sub-matrices and within sub-matrices where there are few foreground pixels.

For each sub-matrix $B_{k,l}$ we next apply an iterative process to determine an optimum threshold value based on the image *histogram*. This is a plot of the pixel values in an image vs. their frequency in the image, which can provide insight into the distribution of foreground and background pixels. Optimally for thresholding purposes, the histogram is bimodal, so that there is a clear difference between foreground and background. In this case, the theoretically best threshold value occurs at the lowest point between the two spikes in the histogram. Otherwise, when the division between foreground and background is less clear, the iterative method will minimize the error of incorrectly categorizing pixels. While there are more sophisticated iterative methods that will accomplish these goals, the method used for this project is a simpler version given in Algorithm 2.

Algorithm 2 Iterative Thresholding Algorithm

1: $\{p_1, ...p_N\} \leftarrow N$ pixels sampled at corners of matrix $B_{k,l}$
2: $T \leftarrow \text{mean}\{p_1, ...p_N\}$
3: **while** true **do**
4:     $m_1 \leftarrow \text{mean}\{B_{k,l}(i,j) : B_{k,l}(i,j) > T\}$
5:     $m_2 \leftarrow \text{mean}\{B_{k,l}(i,j) : B_{k,l}(i,j) \leq T\}$
6:     $Tnew \leftarrow \frac{1}{2}(m_1 + m_2)$
7:     **if** $T \neq Tnew$ **then**
8:         $T \leftarrow Tnew$
9:     **else**
10:         break
11:     **end if**
12: **end while**

A threshold value T is thus determined for each sub-matrix $B_{k,l}$, thresholding is performed on each $B_{k,l}$ to obtain a binary sub-matrix, and these binary sub-matrices are reassembled to create a binary representation of the original image.

This method proved very useful for many of the images used, such as the sample shown in Figure 3.1. However, thresholding of biological images is notoriously difficult, and some images with poor contrast or uneven lighting still required manual thresholding. Automatic thresholding was always tried first, and a human inspected the output for quality to determine which images needed manual thresholding.
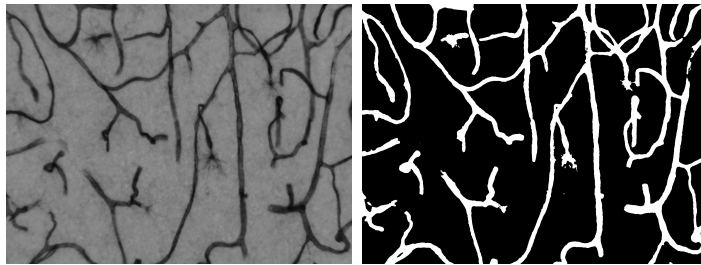

Figure 3.1: Sample image of microvasculature before and after automatic thresholding.

3.3 Morphological Processing

Once an image is thresholded, morphological operations can be performed to further improve the image for processing. A morphological operation on a binary image is a process that alters the shape of features within the image by changing certain pixels from foreground to background or vise versa. The two most basic morphological operations are *dilations* and *erosions*. As noted by Chris Solomon and Toby Breckon in [25], most morphological operations can be reduced to a sequence of dilations and erosions.

The notion of *connectedness* is at the heart of morphological operations. The definition of connectedness that is most relevant to our project is this: we define foreground pixels p_1 and p_2 to be connected if the 3×3 neighborhood of pixels centered at

p_1 contains p_2. This definition of connectedness is often called 8-connectedness, since we look at the "8-neighborhood" of pixels surrounding each foreground pixel. For future reference, we will designate by $\mathcal{N}_8(p_i)$ the 8-neighborhood centered at a pixel p_i. There is also the notion of "4-connectedness," in which we consider foreground pixels p_1 and p_2 to be connected only if p_2 is either immediately above, below, to the left, or to the right of p_2. We will similarly designate the 4-neighborhood centered at p_i by $\mathcal{N}_4(p_i)$ [7, 25].

An *object* in a binary image is a collection of connected foreground pixels. The effect of dilation is to expand the edges of the objects in the binary image. This results in a thickening of the foreground objects, and a closing-off of small "holes" in the foreground. Erosion, on the other hand, "eats away" at the edges of the foreground. This thins the foreground objects, and may delete small objects entirely. This is a desirable effect if we want to reduce background noise in the thresholded image, for example [7, 25].

In both dilation and erosion, a *structuring element* is used to modify pixels in the image. A structuring element is a (typically small) matrix of ones and zeros whose elements can be thought of as binary pixels. Generally, the size and shape of the structuring element should be chosen based on the size and shape of the foreground structures in the image, the definition of connectedness that is used, and the effect that is desired. For our purposes, we will consider a structuring element S to be an $r \times r$ matrix where r is odd. The *center element* c is defined by $c = S(\frac{r+1}{2}, \frac{r+1}{2})$, and the remaining elements of S are called the *neighborhood* of c. In general, the value of c can be either 0 or 1, but for our purposes we will take it to be 1. So, for example,

a 3×3 structuring element that represents 4-connectedness is given by

$$S = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

In both dilation and erosion, we typically begin by padding the original $m \times n$ image matrix A with a border of zeros on the edges to accommodate the structuring element; the width of this border is equal to $\frac{r-1}{2}$. This "pad" is removed after the algorithm completes. We will designate this new padded matrix $\tilde{A}$. For each pixel $\tilde{A}(i,j)$, $i \in [\frac{r-1}{2} + 1, \frac{r-1}{2} + m]$, $j \in [\frac{r-1}{2} + 1, \frac{r-1}{2} + n]$, we define $N_{i,j}$ to be the $r \times r$ matrix representing the neighborhood centered at $\tilde{A}(i,j)$. Let $G = S \circ N_{i,j}$ denote the Hadamard, or component-wise, product of the structuring element S and the neighborhood matrix $N_{i,j}$. Then our algorithms for dilation and erosion are given by Algorithms 3 and 4.

Algorithm 3 Morphological Dilation

1: $B \leftarrow \tilde{A}$ ▷ Copy $\tilde{A}$ into B
2: **for all** $i \in [\frac{r-1}{2} + 1, \frac{r-1}{2} + m]$, $j \in [\frac{r-1}{2} + 1, \frac{r-1}{2} + n]$ **do**
3:      **if** $\tilde{A}(i,j) = 0$ **then**   ▷ Current pixel is background pixel
4:          $G \leftarrow S \circ N_{i,j}$ ▷ Compute Hadamard product G
5:          $s \leftarrow \sum\limits_{\substack{1 < p < r \\ 1 < q < r}} G(p,q)$ ▷ Sum all entries of G
6:          **if** $s > 0$ **then**
7:             $B(i,j) \leftarrow 1$ ▷ Change current pixel to foreground pixel
8:          **end if**
9:      **end if**
10: **end for**
11: $\tilde{A} \leftarrow B$ ▷ Replace $\tilde{A}$ with modified matrix
---

The basic idea behind dilation is therefore to process all background pixels $\tilde{A}(i,j) = 0$, checking the entries of $N_{i,j}$ that correspond to ones in the matrix S. If any of these entries is a 1, we change $\tilde{A}(i,j)$ to a 1. Likewise, in erosion, we con-

---
Algorithm 4 Morphological Erosion

1: $B \leftarrow \tilde{A}$ ▷ Copy $\tilde{A}$ into B
2: $\sigma \leftarrow \sum\limits_{\substack{1<p<r \\ 1<q<r}} S(p,q)$ ▷ Sum all entries of structuring element S
3: **for all** $i \in [\frac{r-1}{2}+1, \frac{r-1}{2}+m]$, $j \in [\frac{r-1}{2}+1, \frac{r-1}{2}+n]$ **do**
4:      **if** $\tilde{A}(i,j) = 1$ **then**   ▷ Current pixel is foreground pixel
5:          $G \leftarrow S \circ N_{i,j}$ ▷ Compute Hadamard product G
6:          $s \leftarrow \sum\limits_{\substack{1<p<r \\ 1<q<r}} G(p,q)$ ▷ Sum all entries of G
7:          **if** $s < \sigma$ **then**
8:              $B(i,j) \leftarrow 0$ ▷ Change current pixel to background pixel
9:          **end if**
10:     **end if**
11: **end for**
12: $\tilde{A} \leftarrow B$ ▷ Replace $\tilde{A}$ with modified matrix
---

sider only the foreground pixels $\tilde{A}(i,j) = 1$. We then look for zeros in the entries of $N_{i,j}$ that correspond to ones in the matrix S and change $\tilde{A}(i,j)$ to a 0 if any are found.

We can combine multiple erosions and dilations together in a sequence to achieve a desired effect. Note that while erosion and dilation seem to be "opposite" operations, applying one after the other, even with the same structuring element, will not restore the original image, since objects that are completely removed by erosion cannot be recovered by dilation, and holes that are completely filled by dilation will not be remade by erosion [25]. Dilation followed by erosion is known as morphological closing, while the inverse operation of erosion followed by dilation is known as opening.

The majority of images used in this project underwent a dilate $\rightarrow$ erode $\rightarrow$ erode sequence to remove small holes and small background objects. Resulting images were quality-checked and the processes were adjusted as needed for different batches of images. An example of the results of this process can be seen in Figure 3.2, in which

a 9×9 structuring element with a roughly circular pattern of ones in the center was applied on each step.
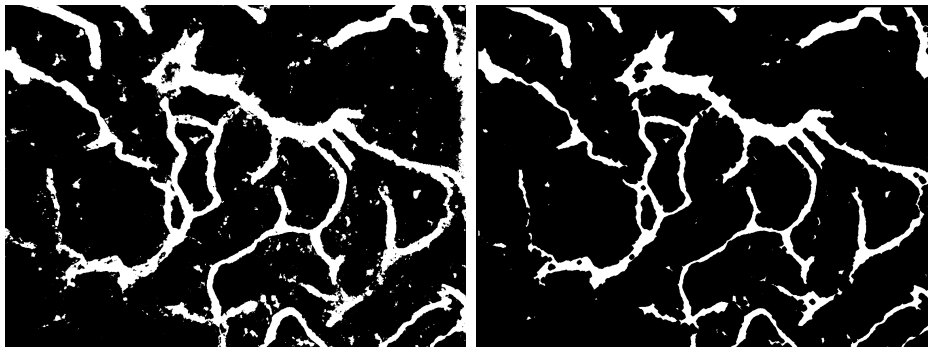

Figure 3.2: A thresholded image with considerable noise before (left) and after a dilate $\rightarrow$ erode $\rightarrow$ erode sequence with a 9×9 structuring element.

3.4 Midline Extraction

In order to measure features such as vessel length and tortuosity, vessel midlines must be isolated. We are, of course, working under the assumption that the length of the vessel midline is a good way to measure the length of the vessel. As previously mentioned, in much of the literature, especially those papers written more than a decade ago, vessel midline extraction was accomplished by manually tracing the vessels. More modern papers describe various computerized techniques to accomplish this procedure, and for the project at hand we used *skeletonization*.

Skeletonization can be thought of as the process of reducing a binary object to its most basic features, resulting in a pixel-wide "skeleton," while maintaining the 8-connectedness of the pixels within the object. There are various algorithms available for the implementation of skeletonization; the one employed in this project was skeletonization via an iterative process of morphological erosions known as *thinning*. The algorithm was based on the Zhang-Suen Parallel Thinning Algorithm (see [31]), with some slight modifications to better suite the nature of the images.

20

In this algorithm, we make multiple passes through the image, on each pass finding pixels on the edges of the foreground objects. There are two sub-iterations on each pass, in which a decision is made whether to delete or preserve each edge pixel based on a set of criteria. Define P_1 to be the current edge pixel, and $P_2, ..., P_9$ to be the pixels in $\mathcal{N}_8(P_1)$, with enumeration beginning with the pixel immediately above P_1 and moving clockwise around P_1.

Let β denote the number of occurrences of $0 \rightarrow 1$ patterns in the sequence $P_2, P_3, ..., P_9, P_2$. On the first sub-iteration, we delete P_1 if all of the following apply:

1. $\beta = 1$

2. $4 \leq \sum_{i=2}^{9} P_i \leq 7$

3. $P_2 P_4 P_6 = 0$

4. $P_4 P_6 P_8 = 0$

After the first sub-iteration is complete, the prescribed changes are made, and the second sub-iteration begins on the modified binary image. The conditions for deletion on the second sub-iteration are almost identical, with changes only made to conditions 3 and 4. P_1 will be deleted on the second sub-iteration if P_1 meets all of the following:

1. $\beta = 1$

2. $4 \leq \sum_{i=2}^{9} P_i \leq 7$

3. $P_2 P_4 P_8 = 0$

4. $P_2 P_6 P_8 = 0$

After both sub-iterations are complete, the process repeats, with iterations terminating when no new deletions are performed. Note that this algorithm differs from the

Zhang-Suen algorithm in the definition of condition 2, as this change was found to give better results in our particular application (see [31]).

The main drawback of most skeletonization algorithms is that the resulting skeletons tend to be highly sensitive to small perturbations in the morphology of the original object. This often results in undesirable "spurs" in the skeleton, which are short skeleton segments that stick out of the main skeleton and can appear in seemingly random places. To mitigate this effect, following skeletonization we implement a "despurring" algorithm. The premise is simple: we locate all end points on the skeleton and follow each spur back to the branch point from which it originates, counting pixels as we go (note that we will discuss end point and branch point identification in the next section). If the spur contains fewer pixels than a certain threshold value, we remove it. This process is then repeated on the modified image, since many spurs may themselves have their own smaller spurs, and two passes are needed to remove them completely. This method is far from perfect, as larger spurs may be left behind, and some smaller true vessel branches may be deleted, but it tends to produce reasonable results for a majority of images. The skeletonization process is illustrated in Figure 3.3.

3.5 Branch and End Point Identification

Identifying branch and end points in the skeletonized image is a fundamental step in the vessel extraction and segmentation process. Not only is this identification necessary for despurring as previously described, but it is also needed to divide the vessel network into segments for the length and tortuosity computations that will be described later. Additionally, since branch points in the skeletonized image typically correspond to actual branch points or cross-over points in the vessel network, branch

22

(a) Original grayscale image.

(b) Thresholded image.

(c) Image after skeletonization; note the unwanted spurs.
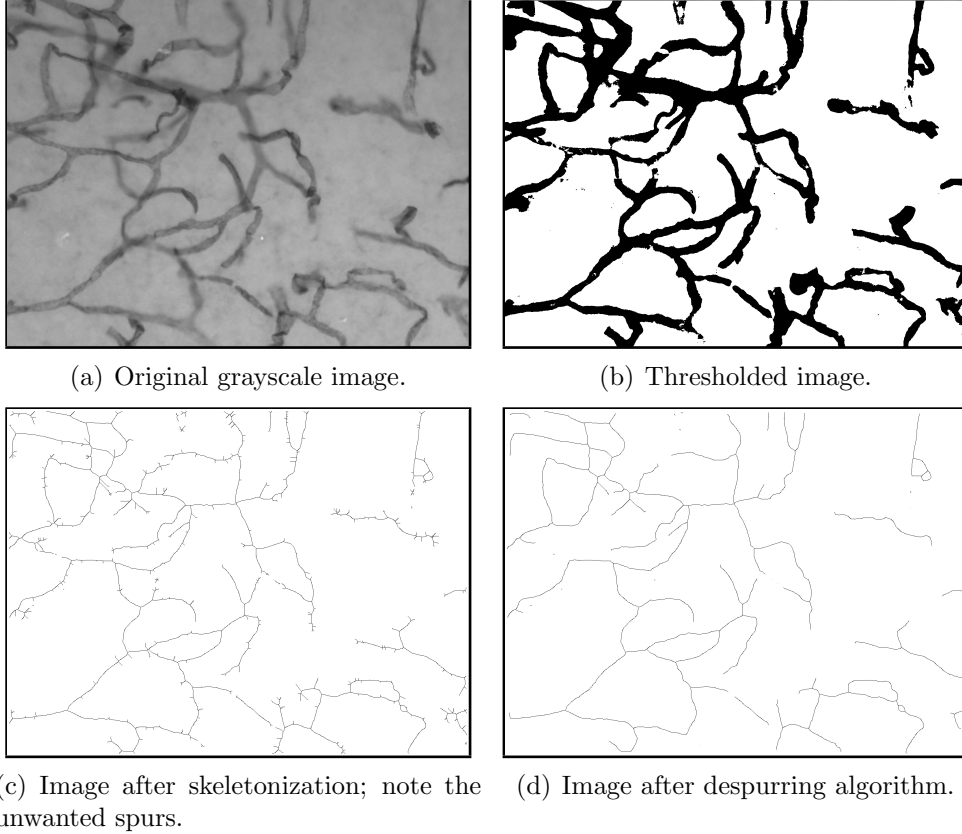
(d) Image after despurring algorithm.

Figure 3.3: The skeletonization process. Binary images have been color-inverted for improved visibility.

point counting can be a useful metric for obtaining information about vessel network morphology.

Locating branch points and end points in a binary skeleton involves iterating through all foreground pixels in the image and examining the surrounding 8-neighborhood $\mathcal{N}_8$ of each pixel. As in the skeletonization example, denote by $P_2, ..., P_9$ the pixels in $\mathcal{N}_8(P_1)$, ordered clockwise with P_2 being the pixel immediately above P_1. Denote again by β the number of $0 \to 1$ patterns in the sequence $P_2, P_3, ..., P_9, P_2$. Then a pixel P_1 is an end point if either of the following apply:

1. $\sum_{i=2}^{9} P_i = 1$

2. $1 < \sum_{i=2}^{9} P_i \leq 3$ and $\beta = 1$.

Note that in the second condition, if $\sum_{i=2}^{9} P_i = 3$, then the point is a spur of length $= 1$. On the other hand, we define P_1 to be a branch point if either of the following are met:

1. $\sum_{i=2}^{9} P_i > 2$ and $\beta \geq 3$

2. $\sum_{i=2}^{9} P_i \geq 5$ and $\beta = 2$ and there exists a $1 \rightarrow 1 \rightarrow 1$ pattern in the sequence $P_2, P_3, ..., P_9, P_2$.

Here, if the second condition is met, then the point is part of a 4×4 block of foreground pixels that are each considered to be a branch point.

Once branch points are identified, we can pare down the count to eliminate adjacent branch points and therefore obtain a more accurate count of true branch and cross-over points in the vascular network. This is accomplished by scanning through the branch points identified by the previous algorithm and deleting any branch points in the surrounding 8-neighborhood of each branch point before moving on to the next pixel.

3.6   Vessel Segmentation

In order to measure features such as tortuosity, we need to next divide the vessel network into individual vessel segments. Once we locate branch and end points, the skeletonized vessels connecting these points are the segments. Define a *nodal point* as either a branch point b_i or end point e_i in the skeletonized image. Then a *vessel segment* is a contiguous arc of 8-connected pixels, with nodal points as beginning and ending points, and containing no nodal points other than these.

Note that if a segment has the same beginning and ending point b_i, we call it a *loop*. For our purposes, we do not consider *isolated* loops which contain no branch points, i.e. the loop must branch off from another segment if it is to be recognized.

It also does not make sense to talk about a loop beginning and ending at one of the end points e_i. When isolated loops do occur in the skeletonized images, they tend to be artifacts of the image processing rather than being representative of the actual vessel network, and they tend to be very small compared to the overall image. These features can thus be ignored without introducing much error. If a large isolated loop was discovered in a skeletonized image, we could manually add a tiny spur to the loop, which would allow the structure to be recognized by the algorithms.

We can isolate the individual segments by performing the steps shown in Algorithm 5, which has here been simplified considerably, but still retains basic features.

Algorithm 5 Vessel Segment Isolation

 1: **for all** (branch points b_i and end points e_j) **do**
 2:     Find starting points s_k of segments originating from b_i or e_j
 3:     **for all** s_k **do**
 4:         **if** (s_k not previously processed) **then**
 5:             Start a new segment, following s_k back until we find another b_l or e_l
 6:             Record segment information in appropriate data structures.
 7:         **end if**
 8:     **end for**
 9: **end for**

In this algorithm, we perform the operation of following segments back to their ending points by inspecting $\mathcal{N}_8(p)$ of the current pixel p for other unprocessed pixels. An unprocessed pixel $q \in \mathcal{N}_8(p)$ is the next pixel in the segment if one of the following is true:

1. $q \in \mathcal{N}_4(p)$

2. The 2 adjacent neighbor pixels $q_-, q_+ \in \mathcal{N}_8(p) \backslash \{p\}$ are both zero.

When we find the next pixel in the segment, we mark the current pixel as processed and update our variables so that the found pixel becomes the current pixel. We designate the nodal points on either end of each segment as the "starting" and "ending" points, based on the order in which they were found and added to the segment.

As the algorithm progresses, we record segment information in data structures. Each segment is given an integer designation $1, 2, 3, ..., M$, where M is the total number of segments in the image, and we store each pixel's position in each segment, measured from the starting point of the segment. We also record each pixel's arc length distance from the starting point. The arc length $l_{i-1,i}$ between any two adjacent pixels $A(x_{i-1}, y_{i-1})$ and $A(x_i, y_i)$ is defined to be their Euclidean distance, given by

$$l_{i-1,i} = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}.$$

The arc length L of an actual vessel segment whose midline is represented by $k + 1$ pixels $A(x_0, y_0), A(x_1, y_1), ..., A(x_k, y_k)$ can therefore be approximated by

$$L \approx \sum_{i=1}^{k} l_{i-1,i} = \sum_{i=1}^{k} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}, \tag{3.1}$$

where L is given in pixels. Since images used in any one study will all represent anatomical areas of the same dimensions, it is safe to leave this length measurement in pixels. If, however, we wanted to compare images representing areas of differing dimensions, it would become necessary to scale the length values by an appropriate factor to obtain length in microns.

By recording not only each segment's total arc length, but also the progressive arc length between the starting point and each subsequent pixel in the segment, we are able to easily divide each segment into smaller sub-segments at a later time if desired. As we will explain later, to compute tortuosity it became necessary to further divide the image into sub-segments based on some maximum length threshold. By storing the above mentioned information at the time of original segmentation, we can refer to the already-constructed data structures to perform the sub-segmentation process.

Once the process of isolating segments and recording arc lengths is completed, we can obtain three possibly useful measures of the microvascular morphology of an image:

1. The total arc length of all segments,

2. The total number of segments between nodal points,

3. The average arc length γ of the segments between nodal points.

Total length is one of the quantities used by Frahm, Schow and Tobet in the work already described in Chapter 2, and this measure did indeed prove useful in their study [12]. Clearly, total length is one way to measure of the density of the vessel network. The number of segments also appears to be an indicator of density; it is directly related to the number of branch points in the image, which is an indicator of density. The average arc length quantity γ will become important later in this thesis.

Chapter 4

TORTUOSITY METRICS

4.1 Background and Motivation

Tortuosity is an intrinsic morphological characteristic of all blood vessels, and vessels within the PVN region of the brain are naturally quite tortuous. While there is no universally agreed-upon definition of tortuosity, it can be thought of as a measure of how sinuous and twisted a vessel appears. This is a property that is easily seen, but proves somewhat difficult to quantify with a reliable metric. As outlined in Chapter 2, the quantification of tortuosity in the microvasculature of the PVN, and the determination of whether this was a useful measure for revealing differences between groups in the studies, was one of the primary goals of this project.

A tortuosity metric for a curve ν is defined to be a real-valued function τ that takes ν as an input. The value $\tau(\nu)$ therefore gives an indicator of how tortuous a curve appears, with the hope that $\tau(\nu)$ corresponds to observers' notions of tortuosity. In the literature, there are a variety of different metrics for measuring tortuosity in both 2 and 3 dimensions; each has its own advantages and disadvantages.

Before discussing the metrics tested in this project, we present in the next section an overview, in chronological order, of some major contributions to the quantification of vascular tortuosity. Since it would be unwieldy to present all major developments in this area of research, we present here developments concerning the 2-dimensional tortuosity metrics that have been most widely used and studied, and the papers that are of greatest relevance to this thesis.

4.2　A Literature Review

4.2.1　The Pioneering Work of Lotmar, Freiburghaus, and Bracher

It seems Lotmar, Freiburghaus, and Bracher were the first to assign quantitative values to vascular tortuosity in their 1979 paper, "Measurement of Vessel Tortuosity on Fundus Photographs" [18]. They manually subdivided retinal vessels into a series of circular arcs with roughly constant curvature, then measured the chord length c_i and arrow height h_i of each arc, where c_i is defined to be the line segment connecting the end points of the arc, and h_i is the length of a line segment perpendicular to this first segment and connecting the first segment to the arc. Tortuosity τ was then computed by approximating what they called the *relative length variation*:

$$\tau = \frac{\Delta c}{c} \approx \frac{8}{3} \sum_i \left(\frac{h_i}{c_i} \right)^2 . \tag{4.1}$$

While the exact computation proposed by Lotmar et al. did not become a popular tortuosity metric, the idea behind their computation is the foundation for other metrics which have been in use for nearly three decades.

4.2.2　The Distance Factor Metric

The metric that is by far in widest use today is known as the *distance factor metric*, DF, which is indeed similar to Equation 4.1. It was first described by David Williams in his 1982 paper, "Quantification of Arteriolar Tortuosity in Two Normotensive Age Groups," in which Williams computed the tortuosity of retinal arteries and used the data to make biological comparisons between two human age groups [29]. The distance factor metric today remains one of the most straight-forward and reliable tortuosity metrics. The idea is simple: let L denote the arc length of a vessel ν, which can be thought of as the length of the vessel midline, and let C denote

29

the chord length measured as the distance between the end points of the vessel (see Figure 4.1). Note that the vessel can be of arbitrary shape. Then the distance factor metric for tortuosity is given by

$$\tau(\nu) = \frac{L}{C} - 1, \tag{4.2}$$

where the -1 is added so a perfectly straight vessel will have $\tau(\nu) = 0$. The DF metric is based on the assumption that a non-tortuous vessel will be close to a straight line, while a tortuous vessel will have an arc length significantly longer than its chord length.
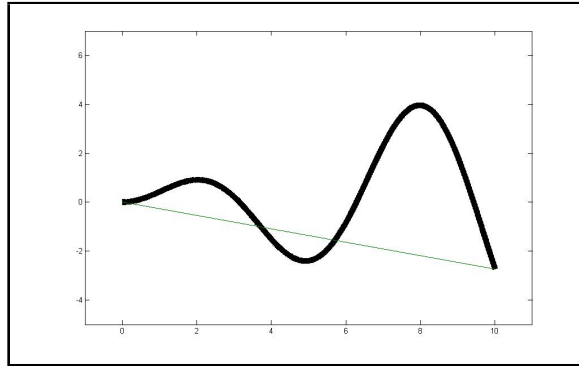

Figure 4.1: Example illustrating distance factor metric. The thick curve represents a vessel midline, and its arc length L is the length of the curve. The length of the thinner line segment connecting the end points gives the chord length C.

4.2.3 Metrics Based on Curvature

Smedby et al. in 1993 were the first to describe metrics based on the mathematical concept of *curvature*, which is the derivative of the tangent direction of a curve with respect to the arc length [24]. In calculus and differential geometry, the unsigned curvature $|\kappa|$ of a function $y = f(x)$ is given by

$$|\kappa| = \frac{|y''|}{(1 + y'^2)^{\frac{3}{2}}}, \tag{4.3}$$

30

and the *total curvature TC* of $y = f(x)$ is then the integral of $|\kappa|$ with respect to arc length s,

$$TC = \int |\kappa| \mathrm{d}s = \int \frac{|y''|}{(1 + y'^2)^{\frac{3}{2}}} \, \mathrm{d}s = \int \frac{|y''|}{1 + y'^2} \, \mathrm{d}x. \tag{4.4}$$

Smedby et al. proposed that the total curvature quantity can be used as a tortuosity metric, with discrete approximations given by symmetric difference quotients and summations. The authors test this metric against the distance factor metric DF, along with several other methods described in the paper, including a scaled TC value, and the number of curvature changes of the vessel midline. The authors conclude that, of the methods tested, the TC and DF methods were the most true to observers' perceptions of tortuosity, and that of these two the DF method was certainly the simpler and more easily explainable method [24].

Indeed, one of the main drawbacks of the TC metric and other similar metrics is their computational difficulty, and the fact that the computations are based on abstract mathematical formulas that may require extensive mathematical training to fully understand. It is, for example, very difficult to explain the TC metric in a simple, intuitive fashion. Since it is often the case that physicians and biologists, who most need to use and understand these metrics, typically do not have extensive mathematical backgrounds, the TC metric loses some of its desirability.

4.2.4   A Defect of the Distance Factor Metric

In 1995, Capowski, Kylstra, and Freedman pointed out a major flaw of the distance factor metric, in that a vessel that curves gently in a large bend will often exhibit a higher DF τ value than a vessel with many smaller bends [8]. This phenomenon can be observed by examining the curves $f(x) = \sqrt{1 - x^2}$ and $g(x) = 0.25 \sin(5x)$ on the interval $[-1, 1]$, plotted in Figure 4.2. Most people would agree that a vessel having a midline represented by $g(x)$ would appear more tortuous than one whose

midline was given by $f(x)$. Since $f(x)$ is a semicircle of radius 1, its arc length on $[-1, 1]$ is just $\pi \approx 3.14$, so that by the DF metric, $\tau(f) = \frac{\pi}{2} - 1 \approx 0.57$. For $g(x)$, we can compute arc length on $[-1, 1]$ by the standard formula from calculus:

$$L = \int_{-1}^{1} \sqrt{1 + g'(x)^2} \, \mathrm{d}x \approx 2.60,$$

so that, contrary to our instincts about observable tortuosity, $\tau(g) \approx 0.30 < \tau(f)$. Capowski et al. do suggest a metric that presents a fix to this problem, but it is very specialized for the particular problem of the diagnosis of plus disease using retinal vascular tortuosity, relying heavily on the particular frequencies with which retinal vessels tend to oscillate [8].
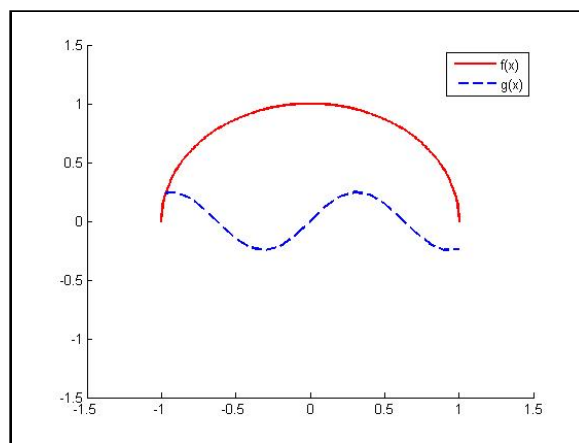

Figure 4.2: Plots of $f(x) = \sqrt{1 - x^2}$ and $g(x) = 0.25 \sin(5x)$ on $[-1, 1]$.

4.2.5 Network Tortuosity

Until the 1998 publication of the paper "Quantification of the Morphological Features of a Full Microvascular Network" by Bidiwala et al., the literature discussed only quantifying properties of individual vessels, rather than vessel networks [3]. The problem of assigning quantitative values to describe the composite features of multiple vessels in a network is rather a different problem than describing features of one vessel.

Measuring the tortuosity, for example, of any one vessel segment in a network is not necessarily representative of the overall appearance of the network. A network may contain both very tortuous and very non-tortuous vessels, but it is the overall effect that needs to be measured [3].

Bidiwala et al. used manual tracing of microvessels in the latissimus dorsi muscle of mice to quantify features of the vessel networks. They used the distance factor method DF to measure tortuosity of individual vessels, and a weighted additivity metric to gauge the tortuosity of vessel networks, given by

$$\tau(Network) = \frac{\sum_i \tau(\nu_i) L_i}{\sum_i L_i}, \tag{4.5}$$

where L_i is the arc length of the ith vessel segment ν_i in the network. The authors report general success of this metric, although they admit that the results were likely adversely affected by errors introduced in the vessel extraction process. It should be noted that the images used in Bidiwala's study were of significantly different quality than our images, with much more background noise (interested readers may refer to [3]).

4.2.6 Theoretical Considerations

In their 1999 paper "Measurement and Classification of Retinal Vascular Tortuosity," Hart et al. test several metrics against each other and define some desirable theoretical properties of a tortuosity metric [14]. The authors suggest that an ideal tortuosity metric should have the following properties, based on human observers' notion of tortuosity:

1. Invariance to translation and rotation, so that tortuosity does not depend on a curve's location or orientation.

2. Multiplicative response to scaling, so that if ν is scaled (i.e. dilated) by a factor α, then we have $\tau(\alpha\nu) = \phi(\alpha)\tau(\nu)$ for some function ϕ, where $\phi(\alpha) = 1$ in the ideal case of invariance to scaling.

3. Compositionality, which is defined to be the property that if a curve ν is composed of two arcs ν_1 and ν_2 with $\tau(\nu_1) \leq \tau(\nu_2)$, then $\tau(\nu_1) \leq \tau(\nu) \leq \tau(\nu_2)$.

While not all tortuosity metrics meet these properties, Hart et al. use these properties as guidelines for defining, characterizing, and evaluating a variety of metrics, including the DF metric and total curvature TC, and several closely related to TC. Both the DF and TC metrics exhibit translational and rotational invariance. The DF metric exhibits invariance to scaling, while the TC metric exhibits multiplicative scaling. Both metrics lack the compositionality property. It should be noted, however, that the use of the weighted additivity technique for quantifying networks of vessels basically forces the compositionality property. Hart et al. do discuss network tortuosity, and they use the weighted additivity formula 4.5 with each of the metrics tested to determine tortuosity of vessel networks [14].

Tests were performed to see if the various metrics matched observers' classification of retinal vessels and vessel networks as being either tortuous or non-tortuous. The end determination was that TC performed poorly overall, and although they did not find significant evidence to make a strong recommendation for any one metric, the metric that performed best was "total squared curvature," which is the same formula as TC (Equation 4.4), but with the integrand squared [14].

Their determination was that the DF metric worked very well for networks when vessel segments were automatically extracted. Their assessment (which matches intuition when one considers the aforementioned drawback associated with the DF metric) was that the DF method works better on shorter vessel segments. Since

automatically extracted vessel segments tended to be shorter than those that are manually extracted, the DF method proved to be a good choice in this scenario [14].

4.3   The Implemented Metrics

For our project, we needed a tortuosity metric that could be applied to the specific problem of detecting morphological differences in the microvasculature within the PVN between groups of mice. In most of the literature, the goals are quite different. For instance, in the retinal vasculature studies, determining a grading system for tortuosity is important, since the severity of retinal disease tends to be an increasing function of vascular tortuosity (see [30], for example). The same is true when arterial tortuosity is used as an indicator for atherosclerosis [24]. Microvasculature within the PVN, however, tends to be naturally sinuous and twisted; tortuosity is not necessarily an indicator of disease. Increases or decreases in tortuosity that come as a result of experimental procedures are of interest to us; our main objective is to see if a particular procedure invokes *any* change in the vascular morphology. If changes are indicated, they will need further study to determine their causes and implications.

The two basic tortuosity metrics used in this study were the DF metric given by 4.2, and a novel metric which we designate the *convex hull* metric for tortuosity, or CH. Due to the computational challenges associated with the metrics based on curvature, we did not implement any of these methods as of the writing of this thesis. We will in this section present a detailed overview of the metrics used, beginning with their definitions for measuring individual vessels and moving on to their extension to vessel networks. The DF metric for individual vessels has already been thoroughly explained, so we will skip the preliminary introduction of this metric, and instead focus our attention on the new CH metric.
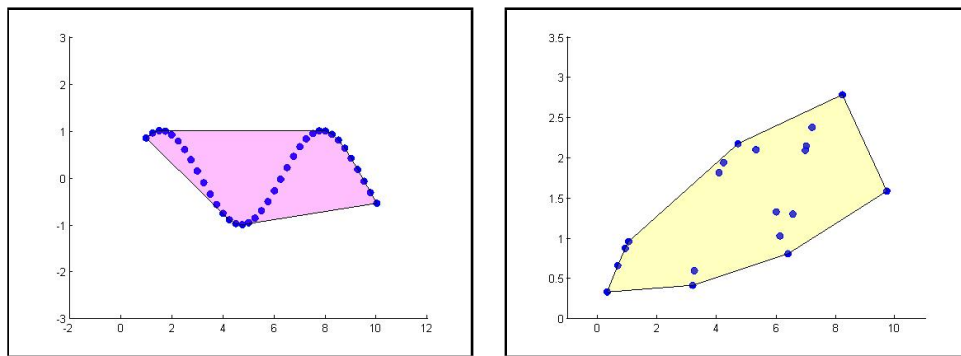
4.3.1 The Convex Hull Metric

In mathematics, a set $U \subseteq \mathbb{R}^n$ is *convex* if, given any two points $x, y \in U$, the line segment connecting x and y is contained entirely in U. The *convex hull V* of a set $\Omega \subseteq \mathbb{R}^n$ is defined by

$$V = \bigcap_\alpha U_\alpha$$

where $U_\alpha \subseteq \mathbb{R}^n$ is any convex set such that $\Omega \subseteq U_\alpha$. Since any intersection of convex sets is itself convex, the convex hull is a convex set [27].

For our problem, we consider only 2-dimensional convex hulls; we therefore may take n to be 2. Intuitively, a 2-dimensional convex hull of a finite set of points $\Omega \subseteq \mathbb{R}^2$ is the smallest convex polygon that entirely contains Ω. This is often explained through the "rubber band analogy:" if each point in Ω is represented by a nail driven halfway into a piece of plywood representing the plane, then the convex hull of the set of points can be seen by stretching a rubber band around all the nails and observing the polygon that is formed. The *convex hull area, H*, is the area contained inside the rubber band. For an illustrative example, see Figure 4.3, where convex hulls for two sets of points have been computed, and the resulting convex hull areas are shaded.

(a) Convex hull area (shaded) of a set of points arranged in a sinusoidal shape.

(b) Convex hull area (shaded) of a set of scattered points.

Figure 4.3: Examples of convex hull areas.

The convex hull (CH) metric for computing the tortuosity τ of an individual vessel or vessel segment ν is defined by

$$\tau(\nu) = \frac{H}{C^2}, \tag{4.6}$$

where H is the convex hull area of the discrete digital representation of the vessel midline, and C is the chord length of the vessel. In other words, we compute the convex hull area of the *pixels* that constitute the vessel, and then divide this area by the square of the distance between the end points of the vessel, thus obtaining a dimensionless quantity. These quantities are illustrated in Figure 4.4. Note that since typically $H << C^2$, the convex hull metric usually gives small values between 0 and 1.
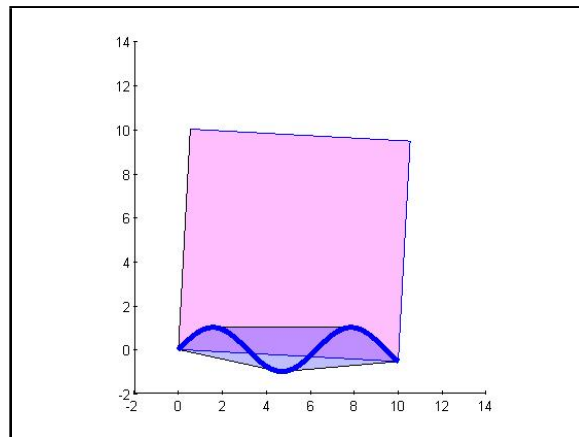

Figure 4.4: Illustration of convex hull metric. The thick curve represents a vessel midline. Its convex hull area is shaded blue, and the large shaded square represents the area of the square of the chord length. In the convex hull metric, the smaller shaded area is divided by the larger area of the square.

What happens when the points representing the vessel segment are collinear, i.e. the vessel segment is perfectly straight? The convex hull of a finite set of collinear points is defined to be a straight line segment containing the points, and its area is

therefore 0. Thus the tortuosity of a perfectly straight vessel will be 0 by the CH metric.

The computation of convex hulls is a standard problem in computational geometry, and there is a wealth of freely-available code for computing convex hulls. For our implementation, we used the standard Matlab *convhull* function. Initial implementation of this project was performed using Matlab version 7.8.0, release R2009a; in this release and in several previous versions, *convhull* itself uses the program Qhull, which is freely-distributed open-source convex hull code [2].

Since Qhull does not compute 2-dimensional convex hulls of collinear points, and attempts to do so will generate an error message, we developed an error-handling technique to address this issue. To construct a 2-dimensional convex hull, Qhull selects 3 initial points to use as an initial simplex (i.e. a triangle, when working in 2 dimensions). If the 3 points lie in one dimension, the algorithm fails. To remedy this problem, it was necessary to invoke the Qhull arguments Qs and QJ when calling the *convhull* function. Qs is useful when the set of points are *not* collinear, but merely close to collinear; this forces Qhull to search all points in the set to form the initial simplex. QJ is used as a last resort, when the points are truly collinear. In this case, Qhull "joggles" the points slightly, based on an error tolerance, to force the set to be 2-dimensional [2].

While this does inevitably introduce error into the computations, since a perfectly straight vessel will now have a computed convex hull area greater than 0, it was determined that this error was acceptable for our purposes. Take, for example, the set of collinear points shown in Figure 4.5. Using the method described above, Qhull computed their convex hull area to be $H \approx 5.01 \times 10^{-9}$. Furthermore, when we divide by the square of the chord length, the tortuosity of the line segment connecting these points will be $\tau \approx 2.51 \times 10^{-11}$, which for our purposes is extremely negligible.

Since more recent versions of Matlab no longer support Qhull-specific parameters in the *convhull* function, a future update of the algorithms for the CH metric will include a different error-handling mechanism, which is yet to be determined.
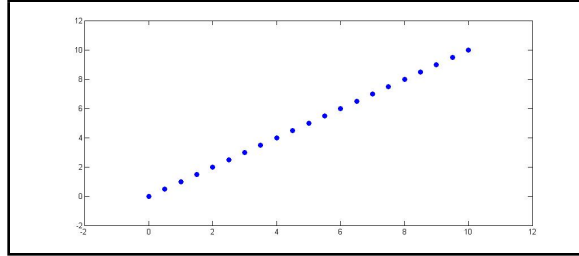

Figure 4.5: A set of collinear points in the plane. Use of the Qhull parameter QJ allows computation of a tortuosity value near 0 for a line segment connecting these points.

The CH metric is admittedly very similar to the DF metric. Both metrics are based on ratios of geometric dimensions associated with a curve. Like the DF metric, the CH metric exhibits rotational and translational invariance and invariance to scaling, but lacks the compositionality property discussed in the previous section. Additionally, the CH metric also tends to give larger values for curves with a large, gentle bend than for curves composed of many smaller bends. The logic behind choosing the CH metric as an alternative to the DF metric was that the convex hull area for a set of pixels tends to be very stable with respect to small perturbations in the location of the pixels. Arc length, on the other hand, can change more dramatically if a few pixels are perturbed, which may increase sensitivity to errors introduced by the image processing and skeletonization processes.

One obvious flaw of the CH metric is that we can draw many different curves of varying visible tortuosity that all have the same convex hull area and the same chord length. The solution to this problem, as we will see, will be the same as the solution to the previously noted problems with the DF metric: proper segmentation of the vessels into small enough pieces. For both the DF and CH metrics, it was hypothesized that

39

we would get more accurate results by dividing the image into segments small enough to obtain roughly constant curvature throughout each segment.

4.3.2 Extension to Vascular Networks

For both the DF and CH metrics, we extend the metric to entire vessel networks using two different weighted additivity methods. The first of these, which we will designate *Method 1*, is the same that was used by both Bidiwala and Hart, in which the individual vessels in a network are weighted according to their arc lengths [3, 14]. The second, *Method 2*, involves using vessel chord lengths as the weights, rather than arc lengths. To the best of our knowledge, Method 2 does not appear in the literature. Admittedly, Method 2 may seem counterintuitive, since vessels with higher tortuosity tend to have smaller chord lengths; therefore vessels with higher tortuosity are often weighted less than vessels with low tortuosity. As we shall see, however, Method 2 stood up quite well in our experiments.

To clarify our definitions of the two methods, consider a vessel network that has been divided into M segments, $\nu_1, ..., \nu_M$. Then the tortuosity τ_1 of the network using Method 1 will be given by

$$\tau_1(Network) = \frac{\sum_{i=1}^{M} \tau(\nu_i) L_i}{\sum_{i=1}^{M} L_i}, \tag{4.7}$$

while for Method 2 the computation will be

$$\tau_2(Network) = \frac{\sum_{i=1}^{M} \tau(\nu_i) C_i}{\sum_{i=1}^{M} C_i}, \tag{4.8}$$

where L_i and C_i are the arc lengths and chord lengths, respectively, of segment ν_i.

We will, throughout the remainder of this thesis, denote by $DF1$ and $DF2$ the distance factor metric applied to networks using Method 1 and Method 2, respectively, and similarly $CH1$ and $CH2$ will denote the corresponding convex hull metrics. The formulas for the 4 network tortuosity metrics are therefore given by:

$$\tau_{DF1} = \frac{\sum_{i=1}^{M} \left(\frac{L_i^2}{C_i} - L_i \right)}{\sum_{i=1}^{M} L_i}, \tag{4.9}$$

$$\tau_{DF2} = \frac{\sum_{i=1}^{M} \left(L_i - C_i \right)}{\sum_{i=1}^{M} C_i}, \tag{4.10}$$

$$\tau_{CH1} = \frac{\sum_{i=1}^{M} \frac{H_i L_i}{C_i^2}}{\sum_{i=1}^{M} L_i}, \tag{4.11}$$

and

$$\tau_{CH2} = \frac{\sum_{i=1}^{M} \frac{H_i}{C_i}}{\sum_{i=1}^{M} C_i}, \tag{4.12}$$

where L_i and C_i are defined as before, and H_i is the convex hull area for the ith segment. Now that the metrics have been thoroughly defined, we will in Chapter 5 move on to a discussion of their performance as tested in a series of experiments.

Chapter 5

PERFORMANCE AND RESULTS

In this chapter we will evaluate each metric's capacity to quantify tortuosity in the images of microvascular networks that are of interest to this thesis. This will be accomplished by analyzing a series of experiments in which we applied the metrics to the images described in Chapter 2. While it should be clear from the previous discussion that the $DF1$ metric has been used for many years to quantify network tortuosity, it has never (as far as we know) been applied to images resembling those used in this project. The $DF2$ metric, on the other hand, is a previously untested variant of the $DF1$ metric, and the CH metrics are completely novel.

When evaluating the usefulness of the metrics, there are two basic considerations. The first is whether the metrics provide a measure that matches our notions of tortuosity based on visual perception, and the second is whether the metrics are successful at indicating significant differences between groups of images. Note that it is possible for a metric to indicate a clear difference between sets of images, and yet not provide a grading scale that matches our visual perceptions of tortuosity.

5.1 Refined Segmentation of Networks

Given that we must divide a vessel network into many small vessel segments in order to apply any of our metrics, the next logical question is how to define a vessel segment. In many papers on the subject, vessels and vessel networks have been manually divided into roughly equal segments or segments with roughly constant curvature. This is, of course, extremely labor-intensive, and was absolutely out of

the question for this project, given the large number of images that needed to be processed. An automated segmentation technique was therefore necessary.

There is no discussion in the literature, as far as we know, regarding the effect of varying segment arc lengths on the computation of tortuosity, either for individual vessels or for networks. Hart et al. did touch on this issue in their discussion of how the performance of the DF metric was improved when shorter segments were used, but this phenomenon was never fully explored [14]. It seems rather surprising to us, given the high sensitivity of the DF metric and similar methods to variations in segment length, that this has not been investigated further. One of the focal points of this project therefore became to analyze the effect of varying segment lengths on the composite network tortuosity.

As discussed in Section 3.6, we initially segmented a skeletonized image of a vascular network by defining a segment to be any length of uninterrupted vessel connecting two nodal points in the skeletonized image. In order to apply the metrics, however, we first had to deal with the presence of closed loops in the images. A closed loop occurs when a segment begins and ends at the same point; the chord length is therefore 0, resulting in division by 0 when applying either the DF or CH metrics. The solution to this problem was to divide each closed loop into two segments of roughly equal length before beginning the tortuosity computation.

Our first preliminary tests with the metrics therefore used the convention that segments were defined as lengths between nodal points, except in the above noted case of closed loops. We also specified a minimum length threshold, L_{min}, which for most of the experimentation was set to $L_{min} = 5$, so that segments with arc length $L < L_{min}$ were ignored for computing tortuosity. Initial experimentation using this segmentation method served to highlight the necessity for more refined segmentation. Images that, on average, contained longer segments inevitably exhibited higher tortuosity values than images containing shorter, rapidly twisting segments. This,

of course, is counterintuitive to the notion of tortuosity. To illustrate this problem, consider the images of vessel networks shown in Figure 5.1. Computing the τ values for these two images using this segmentation method yielded the results shown in Table 5.1, which have been rounded to 4 decimal places.
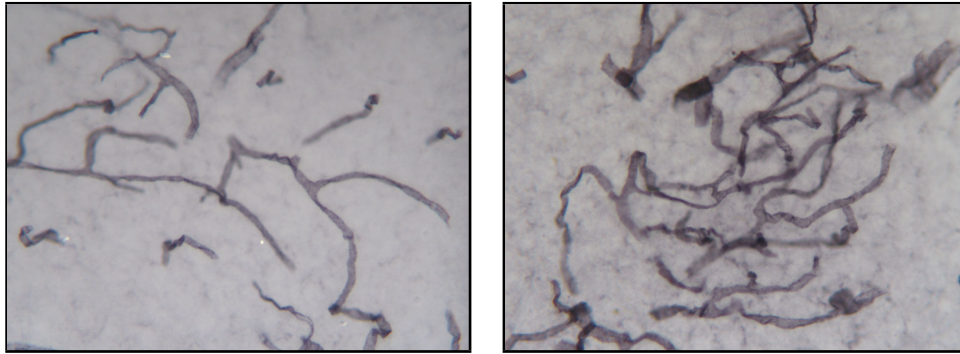
(a) Image A: a network with apparent low tortuosity.

(b) Image B: a network with apparent high tortuosity.

Figure 5.1: Two microvascular networks with visibly disparate tortuosity.

Table 5.1: Tortuosity values τ computed using the 4 basic metrics, with segments defined by curves between nodal points, for the vessel networks shown in Figure 5.1.

	Image A	Image B
$DF1$	0.5195	0.3524
$DF2$	0.4055	0.2863
$CH1$	0.1999	0.1462
$CH2$	0.1571	0.1179

As one can see, the tortuosity values τ for all four metrics are greater for Image A than for the clearly more tortuous Image B. The cause of this is clear to us: the average segment length between nodal points, γ, is much higher in the less tortuous image. Indeed, since increased tortuousity tends to increase the number of cross-over points in an image, it is often the case that more tortuous images contain shorter segments between nodal points. Shorter segments viewed by themselves tend to appear less tortuous, with fewer bends, than longer segments. In fact, not only do they *appear* more linear, but the metrics *measure* them as being more linear. This is true even when the overall network containing the shorter segments exhibits greater tortuosity.

Using the methods outlined in Section 3.6, we can compute the average segment length between nodal points, γ, for each image. For Image A, $\gamma \approx 204.8$ pixels (measured as described in Section 3.6), while for Image B, $\gamma \approx 123.4$ pixels. This significant disparity in segment lengths creates a situation where the network tortuosity is being computed in a fundamentally different manner for each image. This example illustrates the necessity for more refined segmentation techniques.

In Section 3.6 it was discussed that the information for the vessel segmentation is stored in a manner that allows us to easily divide segments into smaller sub-segments based on a maximum length threshold L_{max}. This feature was added to the algorithms in response to the above problem, allowing a user to manually select L_{max} for a set of images. If a segment ν between nodal points has arc length $L > L_{max}$, then we divide the segment into a set of smaller sub-segments $\{\nu_i\}$, each having arc length $L_i \leq L_{max}$. With this convention, we can think of the values shown in Table 5.1 as being computed using $L_{max} = \infty$.

To fully clarify the sub-segmentation process, let $p_0, p_1, ..., p_e$ denote the pixels in the segment, such that p_0 denotes the segment starting point and p_e denotes the ending point. Let $L(p_i)$ denote the computed approximate arc length between p_0 and p_i. We then locate the smallest j such that $L(p_j) > L_{max}$, and pixel p_{j-1} becomes the ending point for the first sub-segment. We repeat this process, on each iteration using the updated p_j as the starting point for the new segment, until we reach p_e. The last sub-segment is therefore typically shorter than the others; this is unavoidable when using this method.

Recomputing the metrics using various values for L_{max} yields the results plotted in Figure 5.2, where the τ values for the two images have been computed as a function of L_{max} for each metric. As one can see, the CH metrics with various $L_{max} < \infty$ now consistently give larger τ values for Figure 5.1(b), which seems to better match what the human eye can see in the images. Results for the DF metrics are less good;

for some L_{max} values, the size of the τ values are still reversed. This example serves to illustrate a fact that was repeated in many experiments: for certain images, the DF metrics are less consistent than the CH metrics in making a distinction between two visually different images. However, when the disparity in apparent tortuosity is extremely large (larger than in this example), the DF metrics seem to do a better job distinguishing tortuosity differences. Additionally, as we shall see in a later section, the DF metrics are generally successful at making distinctions in *average* tortuosity differences between sets containing multiple images, which is more the goal of this project.
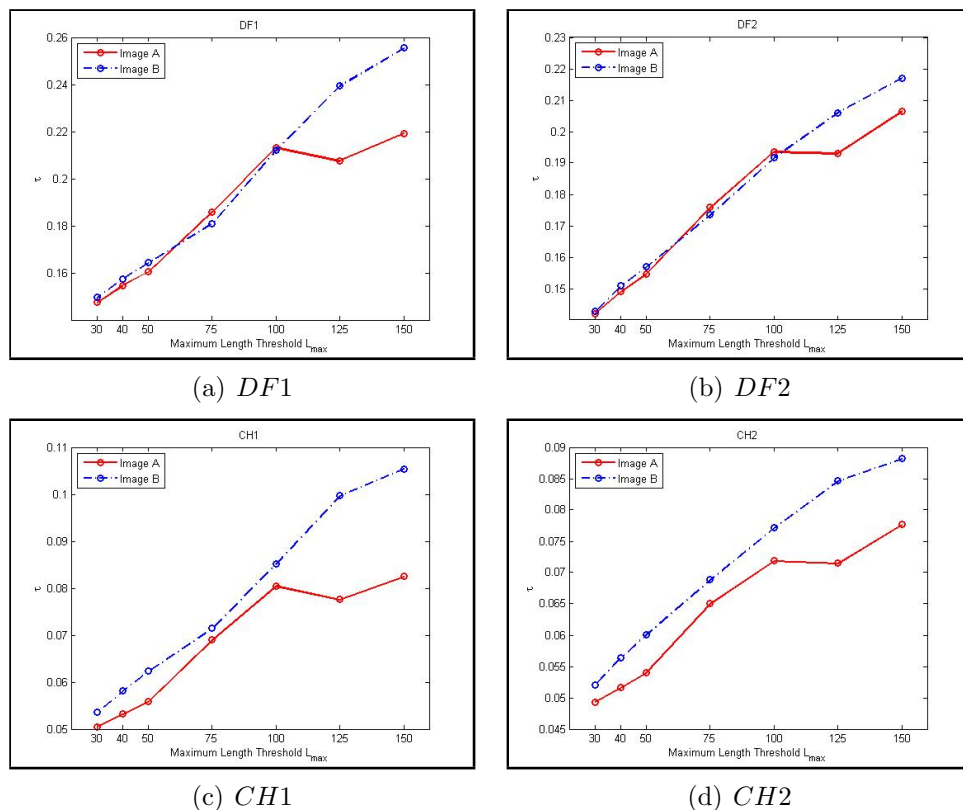

(a) $DF1$
(b) $DF2$
(c) $CH1$
(d) $CH2$

Figure 5.2: Plots of L_{max} vs. τ for each of the 4 metrics, computed for the images shown in Figure 5.1. The solid line represents values for Image A; the dashed line gives data for Image B.

The reader will notice that, for the case of Image B at least, τ is an increasing function of L_{max}. It is easy to infer that this is a result of the "local linearity" property

of smooth curves; i.e. if you zoom in close enough on small portion of a smooth curve, it will look like a straight line. This is essentially the same effect that is presumed to cause the problem with comparing the two images in the first place, since shorter segments inevitably appear more linear. This effect results in smaller τ values for smaller segments, so that the overall network τ value is decreased when we choose a smaller L_{max} value.

As for Image A, the toruosity curves plotted in Figure 5.2 seem to be mostly monotonic, but there is an apparent aberration in all 4 metrics at $L_{max} = 100$. This phenomenon, in which there is a sudden "spike" in tortuosity at one particular L_{max} value, was observed many times in many images, and even occurs in average tortuosity computations across multiple images. At this time we do not fully understand why this occurs, except that it is a response to some geometric property in the images, and it seems to depend somewhat on the value of γ. We also know that, overall, the $DF2$ and $CH2$ metrics seem less sensitive to this phenomenon.

From this example, it should be clear that dividing an image into smaller sub-segments is preferable to using the original segments between nodal points; this certainly seemed to work for the CH metrics, at least. An obvious question, then, is what value of L_{max} will give optimum results for a set of images? This question is difficult to answer, and at this time we do not have a complete solution. Testing the validity of τ values for a particular image set is subjective and highly reliant on observers' intuition of tortuosity. Therefore, finding the optimum maximum length threshold L_{max} for a set of images is a problem that is unlikely to have an analytic, mathematically provable solution. Furthermore, assuming that we can obtain the optimum L_{max} value for a set of images, we cannot expect to apply this same L_{max} value to a different set and obtain good results, since there can be a great deal of variation between image sets. In the next section we will take a second look at the issue of finding a good L_{max} value.

5.2   An Experiment

Our main goal in this project was to detect average tortuosity differences between sets containing multiple images, not necessarily between individual images. When comparing two sets of images, we compute tortuosity values for each image in each set and then take averages over each set. We can then determine whether the difference is statistically significant by using a t-test to obtain a p-value.

To test the validity of the metrics and their ability to serve our purposes, we created a set of 20 images with very high visible tortuosity (designated the "tortuous" group), and a second set of 20 images with very low visible tortuosity (designated the "non-tortuous" group). These images were selected from a sample of 124 images; they were a mix of images from various mice and from various PVN regions as well as the cortex, and all represented the same anatomical dimensions. The images were chosen by a human selector based only on their outstanding visible tortuosity or lack thereof; an attempt was made to ignore features such as vessel density or average length between nodal points. A sample of some of the images from each set is provided in Figure 5.3.
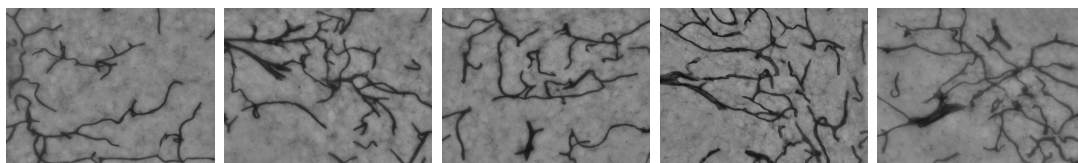
(a) Images from the "tortuous" group.

(b) Images from the "non-tortuous" group.

Figure 5.3: A sampling of images from the tortuous and non-tortuous groups. Images were selected for each group based on their apparent visible tortuosity.

Images were processed according to the methods described in Chapter 3. Tortuosity values τ were computed for all images using the $DF1$, $DF2$, $CH1$, and $CH2$ metrics. Each metric was computed using various maximum length thresholds: $L_{max} = 30$, 40, 50, 75, 100, 125, and 150. A minimum length threshold $L_{min} = 5$ was maintained for each computation. Upon completion of the computations, the images were separated into the two sets, and averages were taken of the τ values in each set. The average τ values obtained for each set using the various metrics are plotted as functions of L_{max} in Figure 5.4.
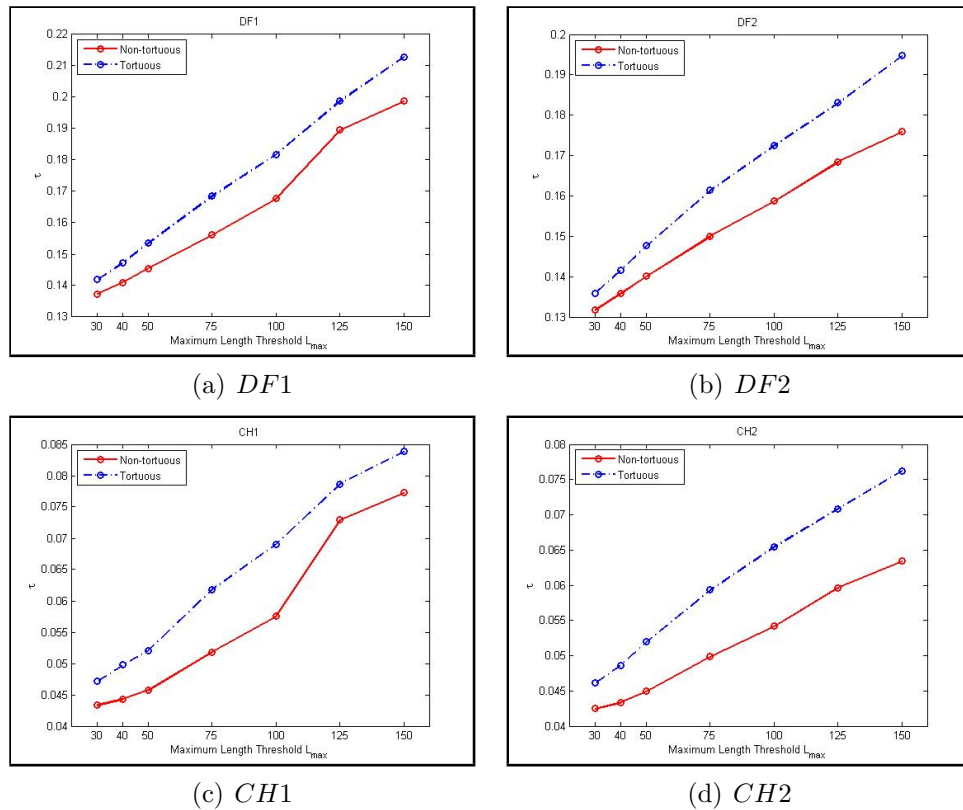

(a) $DF1$ (b) $DF2$

(c) $CH1$ (d) $CH2$

Figure 5.4: Average τ values for the tortuous (dashed line) and non-tortuous (solid line) sets computed using each of the 4 metrics, plotted as a function of L_{max}.

Clearly, all four metrics consistently give "correct" results, in that the average τ values for the visibly tortuous images are higher than those for the visibly non-tortuous images; this is true across all L_{max} values for each metric. To determine

whether the average τ values for each group suggested a statistically significant difference between the groups, we performed a 2-tailed t-test to obtain p-values for the difference in mean for each computation. The t-test results for each metric are shown in Table 5.2, displayed to 4 significant digits.

Table 5.2: P-values indicating the statistical significance of the difference in mean τ values between the tortuous and non-tortuous sets.

L_{max}	$DF1$	$DF2$	$CH1$	$CH2$
30	2.504×10^{-2}	4.010×10^{-2}	4.591×10^{-11}	2.520×10^{-11}
40	2.632×10^{-3}	4.685×10^{-3}	1.145×10^{-13}	6.830×10^{-14}
50	5.754×10^{-4}	7.793×10^{-4}	3.698×10^{-13}	1.594×10^{-13}
75	3.377×10^{-5}	3.640×10^{-5}	1.190×10^{-11}	3.310×10^{-12}
100	1.036×10^{-4}	8.891×10^{-6}	5.379×10^{-10}	1.362×10^{-12}
125	3.534×10^{-1}	1.022×10^{-4}	4.858×10^{-1}	1.111×10^{-8}
150	1.371×10^{-1}	6.387×10^{-5}	3.947×10^{-1}	3.350×10^{-8}

Using the convention that a p-value < 0.05 is an indicator of a statistically significant difference, one can see that when we fix $L_{max} \leq 100$, we obtain statistically significant results using all 4 of the metrics. Thus, all 4 metrics were successful in making a distinction between the two visibly different sets. This supports the hypothesis that the metrics match our intuitive definition of tortuosity.

These results also further support the hypothesis that the results are highly dependent on our choice for L_{max}. When $L_{max} \geq 125$, the $DF1$ and $CH1$ metrics do not indicate statistically significant differences between the sets. At the other exteme, when $L_{max} = 30$, the p-values for the $DF1$ and $DF2$ metrics start to become larger again. It seems that the best choice for L_{max} is somewhere between these two extremes, at least for these image sets.

Further examining Table 5.2, we see that the $CH2$ metric consistently gives the smallest p-values. The CH metrics in general give far smaller p-values than the DF metrics. This may suggest that the CH metrics are better (in this case, at least) at making a clear distinction between the two sets.

5.3 Application to the GABA Study

With the validity of the metrics partially confirmed, we now turn to their application to an actual experimental data set. Our earlier experiment indicated that the metrics performed well when tortuosity differences between two sets of images were very pronounced. The images used for this first experiment were hand-picked to represent extremes in the tortuosity spectrum. It is usually the case, however, that average tortuosity differences between experimental and control groups are not as visually apparent. If the metrics indicate statistically significant differences between groups containing more subtle visual differences, this is a good measure of their robustness.

In the GABA study conducted by Frahm, Schow, and Tobet, it was determined that mice whose $GABA_B$ receptors were eliminated exhibited decreased vascular density and branching within the PVN as compared to the wild type (WT) group [12]. The reader may observe this phenomenon in the sample images provided in Figure 5.5. It seemed to the researchers involved (and perhaps the reader, as well) that the experimental KO group also had decreased tortuosity as compared to the WT group, but at the time of submission of the results for publication, they did not have an effective means to measure this property.

An experiment was therefore performed using bilateral images of the mid-PVN region of 8 mice from the WT group and 9 mice from the KO group, for a total of 16 WT images and 18 KO images (see Chapter 2 for more details regarding the images). Images were processed in the usual way described in Chapter 3. We then computed τ values for each image using all 4 metrics, with $L_{max} = 30, 40, 50, 75, 100, 125,$ and 150. Averages were computed over the images in each group, and these average τ values were plotted against the L_{max} values as before, to obtain the graphs shown in Figure 5.6.
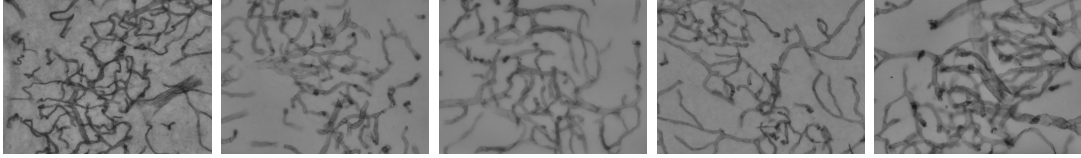
(a) Images from the WT group.
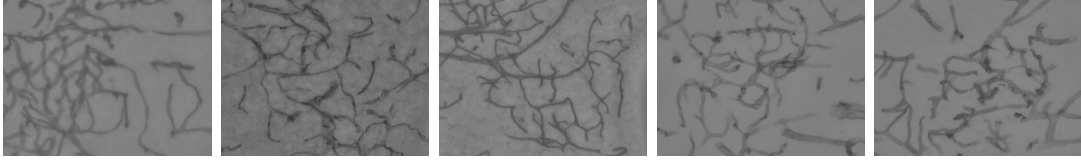
(b) Images from the KO group.

Figure 5.5: A sampling of images from the WT and KO groups.

As one can see, the metrics overall seem to indicate that the WT group indeed exhibits higher tortuosity than the KO group. The reader will also notice that the τ values obtained in this experiment are overall higher than the values obtained in the previous tortuous vs. non-tortuous experiment; indeed, the majority of images used in this experiment exhibited higher visible tortuosity than any of the sample images used in our previous experiment. This is likely due to the age of the animals from which the samples were taken; older animals tend to have denser, more tortuous vessels within the PVN.

There is an obvious anomaly in the $CH1$ metric for the KO group when $L_{max} = 100$, and for the WT group when $L_{max} = 125$. The expected monotonicity of the curve is disturbed here, and the average τ value for the KO group is greater than for the WT group at $L_{max} = 100$. This anomaly is somewhat mirrored in the $DF1$ metric, where we can see a similar, but less pronounced, departure from monotonicity in the KO group when $L_{max} = 100$. This phenomenon seems to be the same one that was noted in our experiment with the two images in Section 5.1.

Define $\bar{\gamma}$ to be the average γ value over all images in a set. Then we note that $\bar{\gamma} \approx 108.7$ pixels for the WT group, and $\bar{\gamma} \approx 119.4$ pixels for the KO group. Choosing $L_{max} = 100$, then, is not likely to be small enough to eliminate problems caused by
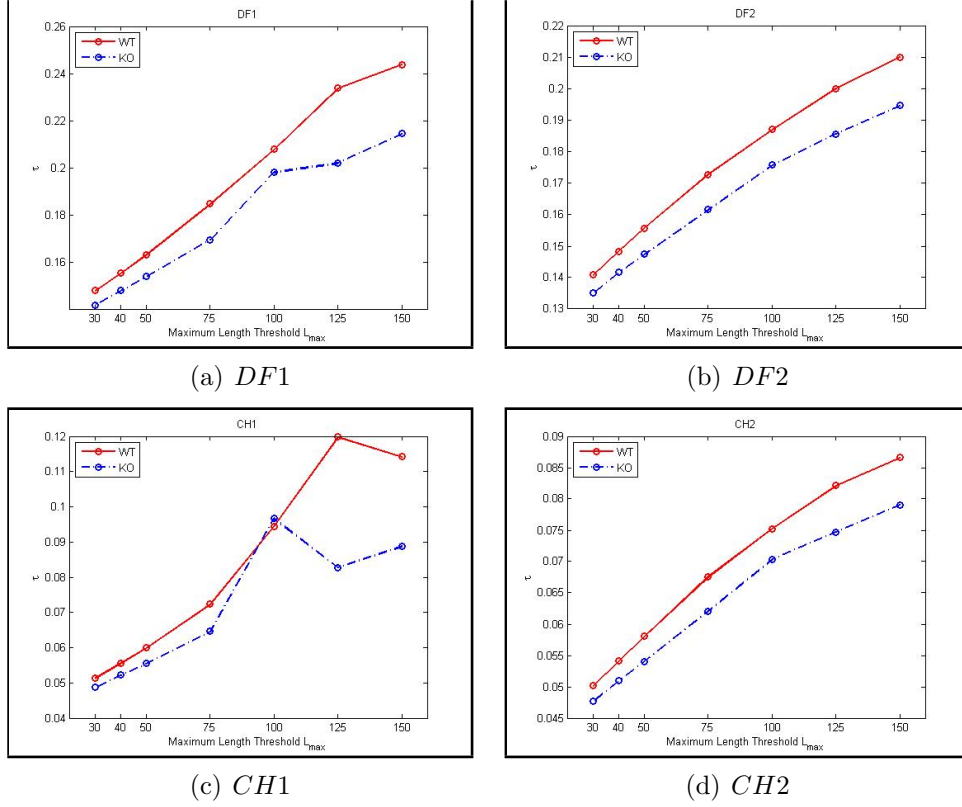
Figure 5.6: Average τ values for the WT (solid line) and KO (dashed line) sets computed using each of the 4 metrics, plotted as a function of L_{max}.

differences in segment lengths between the two sets. This reinforces our hypothesis that L_{max} must be sufficiently small to ensure proper performance of the metrics.

It should be further noted, however, that the $DF2$ and $CH2$ metrics here seem robust to whatever phenomenon is causing the noted aberration. This suggests that the phenomenon is likely related to the arc lengths of the segments, rather than the chord lengths, since these two metrics use chord lengths to weigh individual vessel segments. This fact also serves to promote the desirability of the $DF2$ and $CH2$ metrics over the other two.

If we assume that the τ values computed using $L_{max} < 100$ represent valid tortuosity data for the WT and KO groups, the next question is whether the results indicate a statistically significant difference between the two groups. We again used a

2-tailed t-test to determine whether the difference in mean τ values between the two groups could be considered statistically significant; p-values from this test, rounded to 4 decimal places, are shown in Table 5.3.

Table 5.3: P-values indicating the statistical significance of the difference in mean τ values between the KO and WT groups.

| L_{max}	$DF1$	$DF2$	$CH1$	$CH2$
30	0.0166	0.0238	0.0124	0.0163
40	0.0131	0.0183	0.0106	0.0127
50	0.0044	0.0073	0.0059	0.0076
75	0.0020	0.0055	0.0012	0.0054

As expected, the p-values obtained for these two groups were not as small as those obtained for the previous experiment with the tortuous vs. non-tortuous groups, but all p-values suggest a statistically significant tortuosity difference between the two groups. The smallest p-values this time come from the $CH1$ metric, and the measured difference across all metrics is generally more significant when $L_{max} = 50$ or 75. Fixing $L_{max} < 50$, however, results in increased p-values (and therefore decreased statistically significant difference) for all 4 metrics. This was also observed in the tortuous vs. non-tortuous experiment.

Along with highlighting some important properties of the metrics and the necessity for a good choice of L_{max}, the results seem to confirm that there was indeed a measurable decrease in vascular tortuosity for the KO group as compared to the WT group. This was a difference that could be seen by human eyes, but that was far more subtle than in the case of the tortuous vs. non-tortuous experiment. These results indicate that the metrics are indeed capable of discerning more subtle differences between sets of images.

5.4 Application to the Dexamethasone Study

For our final experiment presented here, we apply the metrics to the images obtained from the dexamethasone study described in Chapter 2. The experiment was conducted using a total of 68 PVN images from the experimental (DEX) group, whose mothers received a dexamethasone injection, and 50 from the control (VEH) group, whose mothers received a vehicle injection (see Chapter 2 for more details).

A sample of the images from the DEX and VEH groups can be seen in Figure 5.7; the reader will notice that these images are of different dimensions and scale than the other images presented in this thesis. While it is clear that overall these images are the most visibly tortuous of any we have presented thus far, most human observers will agree that the DEX images seem more tortuous on average. The images were processed as usual, and τ values were computed as in the previous two experiments. It should also be noted that the average arc lengths between nodal points for these images were *much* smaller than for the previous experiments: $\bar{\gamma} \approx 40.2$ pixels for the DEX group, and $\bar{\gamma} \approx 46.2$ pixels for the VEH group. This difference between the groups, in fact, was found to be quite statistically significant (p-value ≈ 0.0001). Plots of the results of our experiment are shown in Figure 5.8.
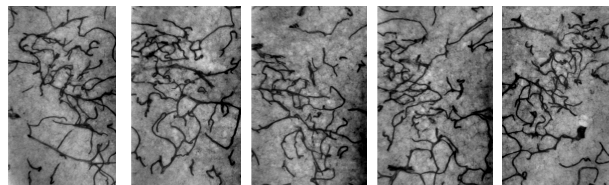
(a) Images from the VEH group.
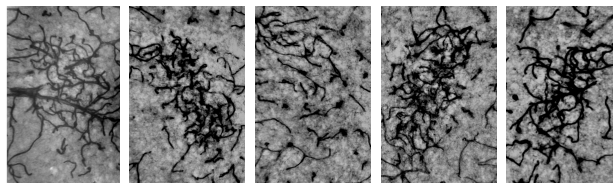
(b) Images from the DEX group.

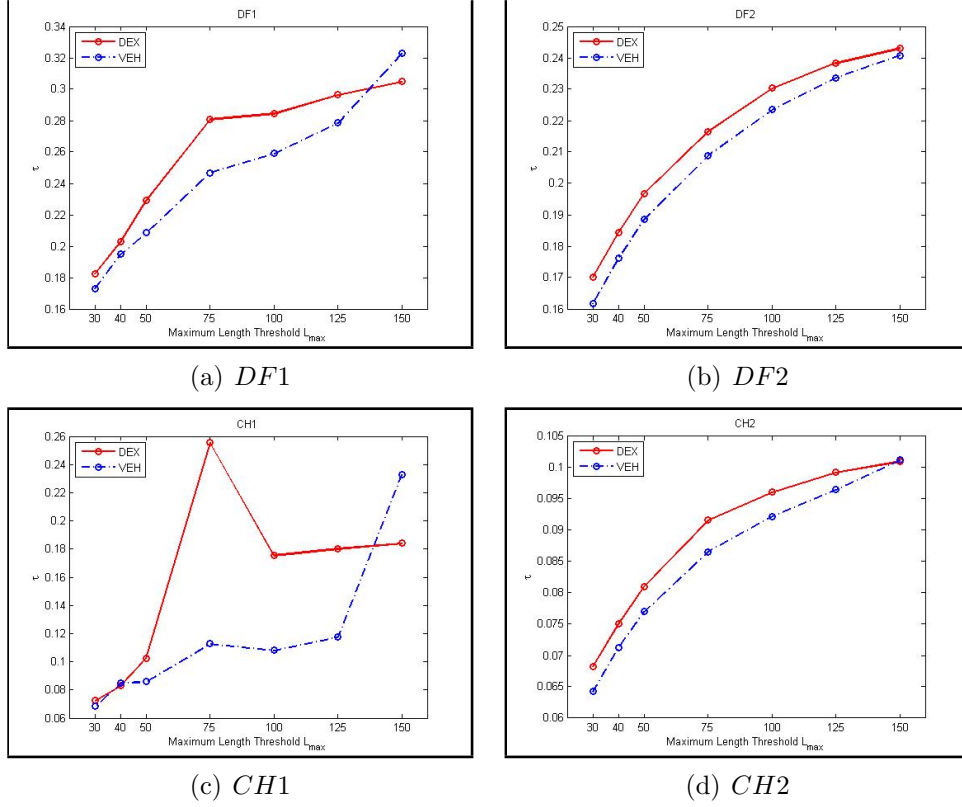Figure 5.7: A sampling of images from the VEH and DEX groups.

Figure 5.8: Average τ values for the DEX (solid line) and VEH (dashed line) sets computed using each of the 4 metrics, plotted as a function of L_{max}.

The reader will immediately notice that the aberration that occurred in the GABA study with the $DF1$ and $CH1$ metrics is repeated in this experiment, but this time at $L_{max} = 75$. A sudden and seemingly discrepant spike is seen in the τ values for the DEX group at this one data point in both the $DF1$ and $CH1$ metrics; this phenomenon is again much more pronounced in the $CH1$ metric. Additionally, the τ values obtained from the $CH1$ metric overall seem inconsistent and unpredictable as compared to the results from the other metrics.

We note too, that the τ values for $L_{max} = 150$ seem to be reversed for each metric as compared to the τ values computed using smaller L_{max}. This is also likely due to the very small lengths between nodal points in these images. Once again, the $DF2$ and $CH2$ metrics seem to give the most consistent, robust results. Overall, all

four metrics gave higher τ values than we have seen in the previous two experiments, which was to be expected.

As in the previous two experiments, p-values were computed (once again using a 2-tailed t-test) for the various τ values to determine the statistical significance of the measured differences between the sets. These p-values are shown in Table 5.4; we have again computed these only for the smaller values of L_{max}. Note the especially bad p-value for the $CH2$ metric when $L_{max} = 40$. In the plots we can see that this corresponds to a point where the two tortuosity curves are reversed in absolute value. Furthermore, we can see that the $DF1$ and $CH1$ metrics in general yield higher p-values, which is consistent with the perturbed shape of the graphs observed in Figure 5.8. Between the $DF2$ and $CH2$ metrics, which measured statistically significant differences across all 4 L_{max} values, the $DF2$ metric this time yields slightly smaller p-values. The best L_{max} values this time seem to be the smaller values; this should not be surprising when one considers the very small γ values in these images.

Table 5.4: P-values indicating the statistical significance of the difference in mean τ values between the VEH and DEX groups.

L_{max}	$DF1$	$DF2$	$CH1$	$CH2$
30	0.0003	1.94×10^{-5}	0.0027	3.17×10^{-5}
40	0.1050	0.0002	0.7830	0.0011
50	0.0017	0.0006	0.0322	0.0018
75	0.1709	0.0111	0.1688	0.0366

5.5 Discussion

Based on the results from the described experiments, we feel we can draw some important conclusions. First, the $DF1$ and $CH1$ metrics seem far too unpredictable to really be of use to us. The $CH1$ metric most noticeably is apparently quite sensitive to geometric properties that seem to be causing the observed perturbations in the metric. While we are still not certain of the exact mechanism behind this

phenomenon, the results from these metrics seem to be untrustworthy. Especially with the $CH1$ metric, performance seems to rely heavily on the maximum length threshold value L_{max}, and small changes in L_{max} may cause wild perturbations in the τ values obtained using this metric. A similar, though less pronounced, effect was observed in the $DF1$ metric.

In Section 4.3.2, we mentioned that it may seem counterintuitive to assign weights to the vessels according to chord lengths, since this will cause more tortuous vessels to be weighted less than vessels that are closer to straight lines. Weighting vessels according to arc length therefore *seems* to be the more logical choice. Previous researchers must have drawn this conclusion as well, as we could find no evidence that network tortuosity has ever before been computed using chord lengths as weights. However, it has been clearly demonstrated that this method, which is used in the $DF2$ and $CH2$ metrics, gives more consistent results than the arc length weighting method, when observed over various L_{max} values.

While the reasons for these observations are not completely understood at this time, we can make a few educated guesses. Since we are comparing the metrics over varying maximum arc length thresholds, the use of a metric that relies heavily on arc lengths within the computation is likely to magnify tortuosity differences among various L_{max} values. We have demonstrated that the reliability of the metrics depends largely on an appropriate choice of L_{max}; by choosing a metric that does not itself rely heavily on arc lengths of individual segments, we reduce this variability. Upon examination of formulas 4.9, 4.10, 4.11, and 4.12, it becomes clear that the $DF2$ metric relies minimally on arc length, while for the $CH2$ metric, arc length does not even enter into the compuation.

Between these two metrics, in fact, our preference at this time is in favor of the $CH2$ metric. We feel that both the $DF2$ and $CH2$ metrics provide accurate tortuosity computations and are quite robust in their ability to discern differences

between sets of visibly different images. However, we have seen that the $CH2$ metric often gives better results when applied to individual images, and the $CH2$ metric did a much better job of making a clear distinction between the image sets in the first experiment.

The issue of how to make an appropriate choice of L_{max} is still not completely solved, but we do have some recommendations, based on all the available information at this time. First, different sets of images will require different L_{max} values. Secondly, it seems that we need to choose $L_{max} \leq \bar{\gamma}$ where $\bar{\gamma}$ is the average of all γ values over the images in a set. If we choose L_{max} too small, however, then the individual segments become close to linear in shape, and the overall morphology of the images are lost. Additionally, it may be wise to take an average over a few different L_{max} values to "smooth out" perturbations between different L_{max} values. Our suggestion, therefore, given 2 or more image sets each containing one or more images, is to obtain tortuosity values for each image using the following procedure:

1. Determine $\bar{\gamma}_j$ for each set j. Let $\bar{\gamma}_* = \min\{\bar{\gamma}_j\}$.

2. Choose a range of L_{max} values, $\{L_{max_1}, ..., L_{max_N}\} \subset [0.5\bar{\gamma}_*, \bar{\gamma}_*]$.

3. Compute the network tortuosity $\tau_{i,k}$ of the kth image, using $CH2$, for each L_{max_i}.

4. Compute the "smoothed" network tortuosity for the kth image: $\tau_k = \frac{1}{N} \sum_{i=1}^{N} \tau_{i,k}$.

We acknowledge that this will result in different tortuosity scales for different experiments. If this is unacceptable, and a uniform tortuosity scale is desired that can be applied across multiple images (assuming the images are of the same level of magnification), then the above procedure can be modified so that it is not dependent on $\bar{\gamma}_*$. This might be accomplished by choosing other upper and lower bounds for the L_{max} values that can apply across all images. Based on the results from the previous

experiments, good results can be obtained using the $CH2$ metric across a fairly wide range of L_{max} values.

Chapter 6

CONCLUDING REMARKS

After considering the results of the experiments, we feel we have largely established the validity of the recommended techniques and have made considerable progress toward the goals outlined in Chapter 2. There are still many questions to be answered concerning some of the results presented here, but in general it seems that we have succeeded in our efforts to quantify tortuosity of microvascular networks and to automate the measurement process.

Further testing of the metrics may be desired, but at this time we feel we have presented considerable evidence to support our conclusions that the $DF2$ and $CH2$ metrics are robust and accurate. The $CH2$ metric, in particular, shows much promise. Since the $CH2$ metric can be applied to simple 2-dimensional images of microvascular networks, we have established there may be little need for expensive and potentially difficult 3-dimensional imaging of these networks, thus providing a solution that is highly cost-effective and practical. Implementation of the metric is also not difficult, especially considering the wide availability of free convex hull code. Additionally, barring some initial image processing to obtain a thresholded image, the process is completely automated, and requires minimal human interaction. It is therefore our sincere hope that the scientific and medical communities may benefit from this research, and that the $CH2$ metric might become a useful tool in the identification of microvascular changes in medical and biological studies.

Bibliography

[1] Hind Azegrouz, Emanuele Trucco, Baljean Dhillon, Thomas MacGillivray, and I.J. MacCormick, *Thickness dependent tortuosity estimation for retinal blood vessels*, Proceedings of IEEE EMBS Annual International Conference, 2003.

[2] C.B. Barber, *Qhull website*, 2012, www.qhull.org.

[3] S. B. Bidiwala, M. S. Mansour, C. K. Stengel, S. A. Klein, S. M. Carroll, S. C. Koenig, A. H. Desoky, G. R. Tobin, C. Maldonado, and J. H. Barker, *Quantification of the morphological features of a full microvascular network*, Medical and Biological Engineering and Computing **36** (1998), 621–626.

[4] Elizabeth Bullitt, Matthew Ewend, James Vredenburgh, Allan Friedman, Weili Lin, Kathy Wilber, Donglin Zeng, Stephen R. Aylward, and David Reardon, *Computerized assessment of vessel morphological changes during treatment of glioblastoma multiforme: Report of a case imaged serially by MRA over four years*, NeuroImage **47** (2008), 143–151.

[5] Elizabeth Bullitt, Guido Gerig, Stephen M. Pizer, Weili Lin, and Stephen R Aylward, *Measuring tortuosity of the intracerebral vasculature from MRA images*, IEEE Transactions on Medical Imaging **22** (2003), 1163–1171.

[6] Elizabeth Bullitt, Donglin Zeng, Benedicte Mortamet, Arpita Ghosh, Stephen R. Aylward, Weili Lin, Bonita L. Marks, and Keith Smith, *The effects of healthy aging on intracerebral blood vessels visualized by magnetic resonance angiography*, Neurobiology of Aging **31** (2010), 290–300.

[7] Wilhelm Burger and Mark James Burge, *Digital image processing: an algorithmic introduction using Java*, Springer, 2008.

[8] J. Capowski, J. Kylstra, and S. Freedman, *A numeric index based on spatial frequency for the tortuosity of retinal vessels and its application to plus disease in retinopathy of prematurity*, Retina **15** (1995), 490–500.

[9] K. V. Chandrinos, M. Pilu, R. B. Fisher, and P. E. Trahanias, *Image processing techniques for the quantification of atherosclerotic changes*, VIII Mediterranean Conference on Medical and Biological Engineering and Computing, 1998.

[10] Joëlle Dennie, Joseph B. Mandeville, Jerrold L. Boxerman, Scott D. Packard, Bruce R. Rosen, and Robert M. Weisskoff, *NMR imaging of changes in vascular morphology due to tumor angiogenesis*, Magnetic Resonance in Medicine **40** (1998), 793–799.

[11] Geoffrey Dougherty and Jozsef Varro, *A quantitative index for the measurement of the tortuosity of blood vessels*, Medical Engineering and Physics **22** (2000), 567–574.

[12] K. A. Frahm, M. J. Schow, and S. A. Tobet, *The vasculature within the paraventricular nucleus of the hypothalamus in mice varies as a function of development, subnuclear location and GABA signaling*, Hormone and Metabolic Research (2012).

[13] Enrico Grisan, Marco Foracchia, and Alfredo Ruggeri, *A novel method for the automatic grading of retinal vessel tortuosity*, IEEE Transactions on Medical Imaging **27** (2008), 310–319.

[14] William E. Hart, Michael Goldbaum, Brad Côté, Paul Kube, and Mark R. Nelson, *Measurement and classification of retinal vascular tortuosity*, International Journal of Medical Informatics **53** (1999), 239–252.

[15] Conor Heneghan, John Flynn, Michael OKeefe, and Mark Cahill, *Characterization of changes in blood vessel width and tortuosity in retinopathy of prematurity using image analysis*, Medical Image Analysis **6** (2002), 407–429.

[16] Michael J. Johnson and Geoff Dougherty, *Robust measures of three-dimensional vascular tortuosity based on the minimum curvature of approximating polynomial spline fits to the vessel mid-line*, Medical Engineering and Physics **29** (2007), 677–690.

[17] J. A. Kylstra, T. Wierzbicki, M. L. Wolbarsht, M. B. Landers, III, and E. Stefansson, *The relationship between retinal vessel tortuosity, diameter, and transmural pressure*, Graephe's Archive for Clinical and Experimental Ophthalmology **224** (1986), 477–480.

[18] W. Lotmar, A. Freiburghaus, and D. Bracher, *Measurement of vessel tortuosity on fundus photographs*, Graephe's Archive for Clinical and Experimental Ophthalmology **211** (1979), 49–57.

[19] S. Maccari, M. Darnaudery, S. Morley-Fletcher, A.R. Zuenac, C. Cinquec, and O. Van Reeth, *Prenatal stress and long-term consequences: implications of glucocorticoid hormones*, Neuroscience and Biobehavioral Reviews **27** (2003), 119–127.

[20] M. Elena Martinez-Perez, Alun D. Hughes, Alice V. Stanton, Simon A. Thom, Neil Chapman, Anil A. Bharath, and Kim H. Parker, *Retinal vascular tree morphology: A semi-automatic quantification*, IEEE Transactions on Biomedical Engineering **49** (2002), 912–917.

[21] Mariusz Paradowski, Halina Kwasnicka, and Krzysztof Borysewicz, *Capillary blood vessel tortuosity measurement using graph analysis*, Knowledge-Based and Intelligent Information and Engineering Systems, Part II (J.D. Velásquez et al, ed.), Springer-Verlag, 2009.

[22] Annelies Quaegebeur, Christian Lange, and Peter Carmeliet, *The neurovascular link in health and disease: Molecular mechanisms and therapeutic implications*, Neuron **71** (2011), 406–424.

[23] Z. Martin Rodriguez, P. Kenny, and L. Gaynor, *Improved characterisation of aortic tortuosity*, Medical Engineering and Physics **33** (2011), 712–719.

[24] Örjan Smedby, Nils Högman, Sven Nilsson, Uno Erikson, Anders G. Olsson, and Göran Walldius, *Two-dimensional tortuosity of the superficial femoral artery in early atherosclerosis*, Journal of Vascular Research **30** (1993), 181–191.

[25] Chris Solomon and Toby Breckon, *Fundamentals of digital image processing: A practical approach with examples in Matlab*, John Wiley and Sons Ltd, 2011.

[26] Srinivas Sriramula, Jeffrey P. Cardinale, Eric Lazartigues, and Joseph Francis, *Prenatal stress and long-term consequences: implications of glucocorticoid hormones*, Cardiovascular Research **92** (2011), 401–408.

[27] James W. Vick, *Homology theory*, Springer-Verlag, 1991.

[28] David Wallace, *Computer-assisted quantification of vascular tortuosity in retinopathy of prematurity*, Trans Am Ophthalmological Society **105** (2007), 594–615.

[29] David Williams, *Quantification of arteriolar tortuosity in two normotensive age groups*, American Journal of Optometry and Physiological Optics **59** (1982), 675–679.

[30] Clare M. Wilson, Kenneth D. Cocker, Merrick J. Moseley, Carl Paterson, Simon T. Clay, William E. Schulenburg, Monte D. Mills, Anna L. Ells, Kim H. Parker, Graham E. Quinn, Alistair R. Fielder, and Jeffrey Ng, *Computerized analysis of retinal vessel width and tortuosity in premature infants*, Investigative Ophthalmology and Visual Science **47** (2008), 3577–3585.

[31] T. Y. Zhang and C. Y. Suen, *A fast parallel algorithm for thinning digital patterns*, Communications of the ACM **27** (1984), 236–239.