Fall 2017

Automated image-based estimation of severity and cause of optic disc edema

Jason Agne
University of Iowa

AUTOMATED IMAGE-BASED ESTIMATION OF SEVERITY AND CAUSE OF OPTIC DISC EDEMA

by

Jason Agne

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy degree
in Electrical and Computer Engineering
in the Graduate College of
The University of Iowa

December 2017

Thesis Supervisor: Associate Professor Mona K. Garvin

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Jason Agne

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Electrical and Computer Engineering at the December 2017 graduation.

Thesis Committee: _____
Mona K. Garvin, Thesis Supervisor

_____
Mathews Jacob

_____
Hans J. Johnson

_____
Randy H. Kardon

_____
Ed Ratner

ACKNOWLEDGEMENTS

When you have been given the opportunity to spend four years continually improving upon and repurposing prior work, your work becomes prior work and it is difficult to truly deem your role as finished – perhaps more so when you are passing along the torch to someone else. Yet that is what is about to happen here. The work was exciting and challenging, with enough that remains to be done to take decades, but it would appear that my time spent contributing to it is at an end.

I first came to the University of Iowa reasonably sure that I wanted to contribute to the field of quantum or optic computation. This ambition was based not on the obvious practical applications, but on the simple fact that the topic seemed mysterious and challenging with relatively few people working on it. However, one of the first lessons I learned when looking for a research advisor to work with was that pretty much all research topics are going to be mysterious and challenging – that is why they need to be researched.

On the other hand, one of my favorite things to do is write code. It is so natural and satisfying that I assumed pretty much everyone likes to code, and so any field that requires coding is likely to be more competitive than I wanted to deal with, but it turns out I was wrong about that. When I first met with Professor Garvin to discuss what opportunities she might have available, I was almost expecting a queue. When I explained some version of these assumptions to her, I was expecting confirmation of my suspicions. That is when I learned I was completely wrong about how most of the world felt about coding, and Prof. Garvin became my research advisor.

Advisor is such a misnomer, though. To say that Prof. Garvin gave me 'advice,' is simply an understatement. She taught and mentored me not just through presentations, lectures, lecture notes, and meetings, but by example as well. It was quite a shift from my previous job, when I worked under a general manager who had only a sense of what I do, with minimal understanding of how. It took a while to adjust to

the idea that I could not only explain my methods to Prof. Garvin, but my reasoning as well. And not only would she understand it, but give feedback, offer alternatives, and often suggest a handful of other sources which describe various ways it has been done before. It is accurate to say that without her 'advice,' this thesis would be blank.

Albert Einstein once said "If you can't explain it to a six year old, you don't understand it yourself." And when it comes to the subject of neuro-ophthalmology, at the onset of my doctoral work, I can honestly say my understanding of the field was on par with that of a six year old. Fortunately, Dr. Kardon has precisely the sort of understanding that Einstein was talking about. While I was comfortable with what I was doing, I was a little uncomfortable with not having an understanding of why it needed to be done. However, after patient explanation by Dr. Kardon and being invited to sit in on clinical meetings, the topic has been made as clear as one could reasonably hope for.

Speaking of 6 year olds, I'd like to mention my children, Jared (six years old) and Neil (three years old) who have kept my home life interesting, and my wife, Leah, who held the household together during my late nights at the office and during my trips to various conferences. I'd like to thank Cathy Kern for her positive outlook and overall administrative helpfulness. Prof. Jacob, who was the professor of the first class I took upon my return to academics after five years working an industrial job. Ed Ratner, who has given me invaluable career advice, in addition to teaching me the fundamentals of graphing algorithms. Prof. Johnson, who introduced me to the world of ITK at a pace appropriate for someone who had not programmed anything in C for eight years and helped put my doctoral work into perspective by emphasizing the importance of ensuring doctoral work has practical applications, thus motivating me to learn about the clinical applications. My friends in the lab, Victor, Ray, Bhavna, Chen, Will, Qiao, Saleh, Sam, Caitlin, Abhay, John, Shafkat,

and Yashila for making the environment friendly and helping when needed. My cat, Ayame, for being adorable. And finally, I'd like to thank the National Institutes of Health and the Department of Veterans Affairs for their financial support through R01 EY023279 and the VA Merit Award I01 RX001786.

ABSTRACT

Optic disc edema can arise from a variety of possible causes, some benign and some life threatening. For timely appropriate medical intervention, or to reduce patient anxiety in the event none is needed, it is critical that the cause and severity of a swollen optic disc be determined as soon as possible.

In this doctoral work, several algorithms are pieced together to determine the cause of optic disc edema. The process of determining the cause of swelling involves the extraction of several features, many of which are relatively new to the field of ophthalmology. Included among these are the shape of Bruch's membrane as found semi-automatically from SD-OCT images and the presence and orientation of folds in the retina, which are also most visible in SD-OCT images, and some selected features from 2D fundus images.

One specific cause of optic disc edema, called papilledema, is due to raised intracranial pressure. This, too, has a variety of possible causes, and often urgency (or severity) is rated by the Frisén scale, which is a 0-5 ordinal rating of severity (with 0 being normal). In the event papilledema is found to be the cause of swelling, this doctoral work also seeks to implement a more robust measurement of severity than the Frisén scale. Specifically, the total retinal volume (TRV) of the optic disc has been computed from SD-OCT images in other work. It is believed that the TRV serves as a more reliable means of assessing papilledema severity, as it is a continuous, repeatable measurement that is not subject to observer interpretation. As part of this doctoral work, the TRV is estimated from fundus images, which are faster and cheaper to obtain than SD-OCT images.

Thus, the aims of this thesis consist of finding and quantifying folds in the retina, using folds (and other features) to distinguish between the causes of a swollen optic disc, and, in the event an optic disc is swollen due to papilledema, to assess the severity of the swelling by estimating the TRV from fundus images. While the ultimate goal of

this work would be to entirely diagnose a patient with optic disc edema from fundus images, that is beyond the scope of a single thesis. Thus the efforts here are to build towards that.

PUBLIC ABSTRACT

The optic disc is the region in the back of the eye that connects the eye to the optic nerve, which connects to the brain. It is the only nerve that is visible without undergoing an invasive procedure, and thus can be of immense value for divining problems that may be taking place in the brain or even the rest of the nervous system. Specifically, if the optic disc is swollen, it could indicate any number of problems, some of which could demand urgent medical intervention. Alternatively, a swollen optic disc could mean nothing at all, as some people have optic discs that naturally appear that way.

In clinics, the severity of optic disc edema, or more specifically papilledema, is assessed on a 0-5 scale (with 0 being normal), called the Frisén scale. This is problematic for two reasons. For one, the Frisén scale is subjective and ordinal, meaning that clinicians can disagree and not know how close they are to agreement. Moreover, the second reason is that optic disc edema is not a widely understood condition throughout all of medical practice - indeed the Frisén scale is meant to apply to only one specific cause of optic disc edema (papilledema), but is often applied to anything that looks swollen because only experts in the field can tell the difference. A major component of this doctoral work is to put effort towards creating a classifier that tells that difference automatically. The classifier uses metrics from various filter responses and operations (called features) in a large array of decision trees (called random forest classification) to determine the most likely cause of optic disc edema. Furthermore, in the event the optic disc is deemed to be swollen due to papilledema, another application of random forest that uses averages instead of classification (called random forest regression) is used to also estimate the severity of the swelling on a continuous scale.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1
INTRODUCTION

Evaluation of the optic disc is one of the few ways a medical practitioner can noninvasively obtain some sense of a patient's condition beyond the limit of what is observed. The optic disc is the area where the optic nerve joins with the eye. Because it exists in the eye, it is visible by looking through the pupil. The significance of this is easy to overlook at first glance - a major nerve, one directly connected to the brain, can be viewed by anyone by merely looking. There is no need for invasive sample collections or laboratories to analyze them.

The healthy optic nerve head has a somewhat standard appearance, which ophthalmologists know and understand quite well. However, deviations from this standard appearance could be indicative of any number of conditions, including glaucoma, papilledema, optic neuropathy, optic neuritis, hemorrhaging, and even some viruses.

While there are a variety of ways the optic disc can deviate from a healthy appearance, this thesis focuses on cases in which the optic disc is swollen. Optic disc edema can be due to several conditions, some of which can be life-threatening while others are benign. For this reason, it is critical that an accurate diagnosis is made promptly to alleviate undue anxiety for the patient, or so that appropriate treatment can be given as soon as possible. For example, pseudopapilledema is a condition that causes the optic disc to appear swollen, but is understood to be benign. Whereas papilledema is very similar in appearance to pseudopapilledema, but is the result of raised intracranial pressure, which could be indicative of any number of issues in the nervous system, or more specifically the brain.

Papilledema is commonly assessed by either direct inspection of the back of the eye or through the equivalent evaluation of a captured fundus image, which is a picture of the back of the eye. The swelling observed in fundus images is quantified using the Frisén scale [3] or modified Frisén scale (MFS) [4], which is a 0-5 ordinal rating

of severity with 0 indicating a normal healthy eye. The criteria for the Frisén scale is judged subjectively and requires manual inspection by a clinical expert, making the evaluations susceptible to inconsistencies. Two fundus images of the same eye can differ in quality, which may affect a grade, and two clinical experts grading the same image may interpret the significance of various features differently. For example, under both the traditional Frisén scale and the MFS, a key qualitative difference between a Frisén grade of 1 or a grade of 2 is whether the obscuration of the ONH border is confined to the nasal region of the optic nerve head, or whether the obscuration is circumferential. However, there is the possibility of ambiguity if the obscuration is only nearly circumferential or that a part of the border is only beginning to obscure, but is still distinct. Moreover, there are other equally important criteria for assessing the severity of papilledema, such as vessel swelling, that may not entirely fit the expected progression of the Frisén scale, thus adding uncertainty to a diagnosis. Adding to that, the grading is ordinal, as opposed to continuous, which implies one cannot determine how close a case of optic disc edema is to the next severity level based on a Frisén grading.

Previously, in an attempt to address the subjectivity of Frisén grading, Echegaray et al. [5] proposed an approach to semi-automatically predicting Frisén-scale grades from color fundus image features. The work by Echegary et al. involves attempts to quantify the qualitative criteria of the Frisén-scale stages, by assessing the obscuration of the ONH boundary for stages one and two while also considering vessel obscuration and other features to distinguish between the more severe stages. There were a combined total of 24 vessel, ONH, and peripapillary features used in a random forest classification, but the most significant features were the ones most closely based on Frisén grading criteria, as one might expect. While this alleviates the effects of qualitative human interpretation, it still does not address the problem of the ordinal nature of the Frisén grading scheme.

With the introduction of optical coherence tomography [6], and more recently spectral-domain optical coherence tomography (SD-OCT), the swelling of the ONH has instead been quantified via the total retinal volume (TRV) [7]. This quantification is done through an automated volumetric segmentation of the physical space between the internal limiting membrane (ILM) and retinal pigment epithelium (RPE) within a 6×6mm area, roughly centered on the ONH. A strong correlation has been found between the TRV and expert-defined Frisén scale grades [7]. Given its continuous nature and better consistency across repeated examinations, it is assumed that the SD-OCT-based ONH volume serves as a better measure of the severity of optic nerve edema [4] and is treated, in this doctoral work, as the gold standard for assessing papilledema severity. However, computing this volume requires clinicians to capture the ONH using SD-OCT imagery, which necessitates the use of expensive equipment not commonly available in either an emergency room or telemedical setting. The ability to accurately assess the degree of optic nerve edema in emergency and telemedical settings is highly desirable as many causes of optic nerve edema can indicate serious underlying conditions that need to be diagnosed expeditiously.

A large part of this doctoral work involves using SD-OCT-based 3D volumetric measures as a standard reference for measuring the severity of optic nerve edema. Specifically, a regression-based approach for predicting the 3D SD-OCT-based TRV of the ONH is developed from features extracted from relatively inexpensive 2D color fundus photographs, thereby obtaining a continuous assessment of the severity of optic nerve edema from fundus photographs alone. This is done by taking some of the more significant features found by Echegaray and building on them to obtain a feature set suitable for predicting continuous scale TRV.

While the severity of optic disc edema is important to discern, it is important to first determine the cause of swelling. Papilledema (Section 2.1), for example, can be indicative of severe underlying conditions such as a brain tumor or hemorrhage,

or could effectively mean nothing in the case of pseudo-papilledema (Section 2.2). Anterior ischemic optic neuropathy (AION, Section 2.3) may not be fatal, but can result in vision loss. There are many ailments that can result in optic disc edema, and each has its own cause. An expedient correct diagnosis is essential for starting early treatment.

The theme of this doctoral work is to use both 3D SD-OCT and 2D fundus images to determine the cause of a swollen optic nerve head and to assess the severity of a swollen optic disc by predicting TRV from fundus images alone, thus obtaining the severity and the risk to the patient. **My work can be summarized by the following aims:**

- **Aim 1: Detect and quantify folds around the optic nerve head.** It has been found [8] that several types of folds can exist in cases of optic disc edema, notably peripapillary wrinkles, retinal folds, and choroidal folds. These findings are still somewhat new, so while a concise method for quantifying folds in some standardly accepted way has not yet been developed, it serves as an essential step towards determining a cause and severity for optic disc edema to detect and categorize the resulting folds. These folds are detected in the retina and quantified for use as features in Aim 2.

- **Aim 2: Use fundus and SD-OCT image features to distinguish between various causes of optic nerve edema.** An approach for extraction of a number of features from both fundus and SD-OCT images is developed for classifying a case of ONH edema based on its cause. In addition to the quantification of folds, it has been found that Bruch's membrane bends towards the vitreous in cases of papilledema, and away from the vitreous in cases of pseudo-papilledema [9]. Previous work has been done to model the shape for high definition raster scans, which are a specific imaging mode that is not always used clinically, but this approach is applied to the more standardly available

volumetric images. The causes distinguished depends on the data available, which is discussed in Section 2.5.

- **Aim 3: Develop an approach to assess optic nerve head edema severity by predicting the total retinal volume.** Features from prior work are used and new features are developed from color fundus photographs alone to train a random forest regression analysis for predicting the total retinal volume as computed from 3D SD-OCT volumetric images acquired on the same day.

1.1 Thesis Organization

The remainder of this thesis is organized as follows.

- **Chaper 2** introduces the clinical background, including the various causes of edema to be distinguished and how optic disc swelling is currently assessed.
- **Chaper 3** discusses the technical details including image modalities and established processing techniques used in this thesis.
- **Chaper 4** provides a detailed account of the detection of folds.
- **Chaper 5** continues the work of the previous chapter by discussing how folds are quantified, and what additional features are used in the distinction between causes of optic disc edema.
- **Chaper 6** looks specifically at papilledema cases to predict the severity of optic disc swelling on a continuous scale.
- **Chaper 7**, concludes and summarizes previous chapters to discuss possible directions for future work.

CHAPTER 2
CLINICAL BACKGROUND

2.1 Papilledema

The nervous system can be modeled as a closed container filled with organic material and fluid. This is not biologically accurate, as there is always an exchange of fluid with the rest of the body, but for this thesis, this simple model serves as a suitable illustration. As a closed system, if pressure is applied to any part of the nervous system (particularly the brain), that increase in pressure may be detectable at any other part of the nervous system. This is what is referred to when one mentions a rise in intracranial pressure (ICP). When the optic disc is swollen due to raised ICP, it is called papilledema. There are a number of causes for elevated ICP, such as a head injury or a brain tumor, but raised ICP can also be idiopathic. Symptoms of raised ICP include headaches and nausea, but these can also be indicative of many possible ailments that have nothing to do with raised ICP. Blindness is not a typical symptom during the early stages of papilledema. However, other visual symptoms such as grey vision or double vision can occur. These symptoms could be caused by factors other than raised ICP, so they are not symptoms one can rely on, alone, for diagnosing someone with papilledema.

Among the most sensitive methods of detecting and monitoring an elevation in ICP is a lumbar puncture [10]. This is an invasive procedure, involving the insertion of a catheter into the spine, which carries some significant risks. Fortunately the ONH (Fig. 2.1), which is directly connected to the nervous system, can be observed noninvasively with no notable risk to the patient, and some conclusions about the status of the ICP of a patient can be drawn.

The current standard for assessing the swelling of an ONH is known as the Frisén scale, a 0-5 ordinal rating of severity (Fig. 2.2). However, assigning a grade can be an imprecise practice. The grading is mostly based on the fact that over time a

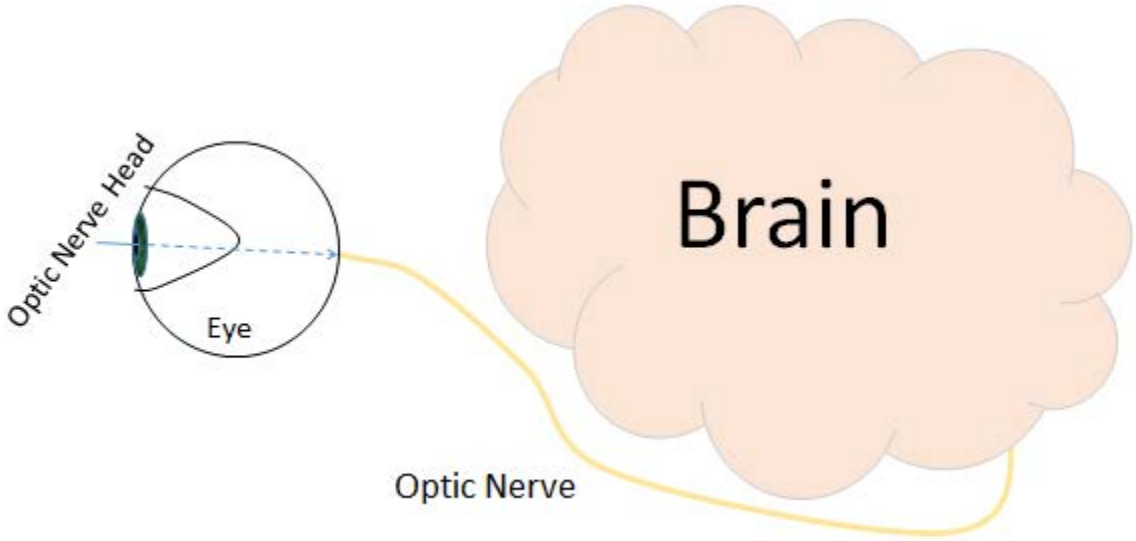


Figure 2.1: Diagram showing the basic anatomy of the eye as it applied to this thesis. The optic nerve head is the region in which the optic nerve meets the eye and can be observed through the pupil.

buildup of axoplasmic fluid obscures the ONH border and/or the veins and arteries in the region. The Frisén criteria qualifies this obscuration by considering the common order in which the obscuration occurs. At Frisén stage 1, there is only an obscuration of the ONH border in the nasal region. At stage 2, the entire ONH border becomes circumferentially obscured. At stage 3, at least one major blood vessel becomes obscured as it leaves the optic disc. At stage 4, the blood vessels on the disc become obscured. Finally, stage 5 is assessed by a total obscuration of all vessels on and leaving the optic disc [4]. While seemingly straightforward, in practice there is no guarantee that events will discretely occur in precisely this order. For example, if the clinician is viewing an image of poor quality, they may be led to believe that there is more swelling than there actually is, or they might overcompensate for the photographic blur of the image and assign a lesser grade. Another example is if the optic disc border is not completely circumferentially obscured before a blood vessel leaving the ONH region starts to show signs of obscuration. In this case, a grading

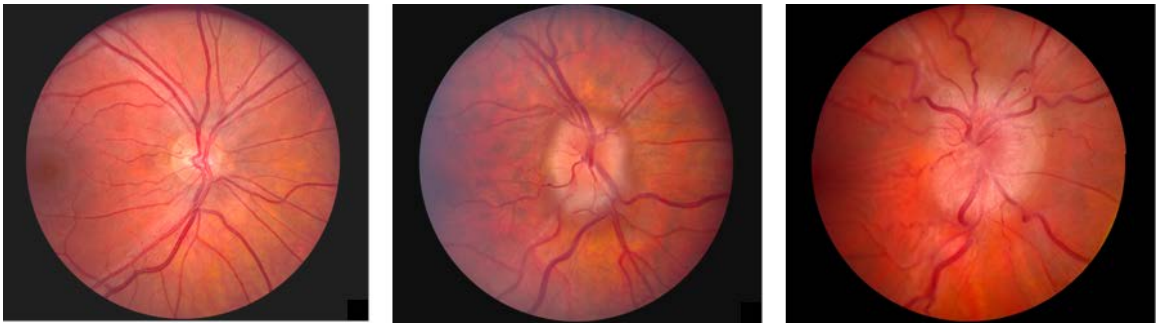
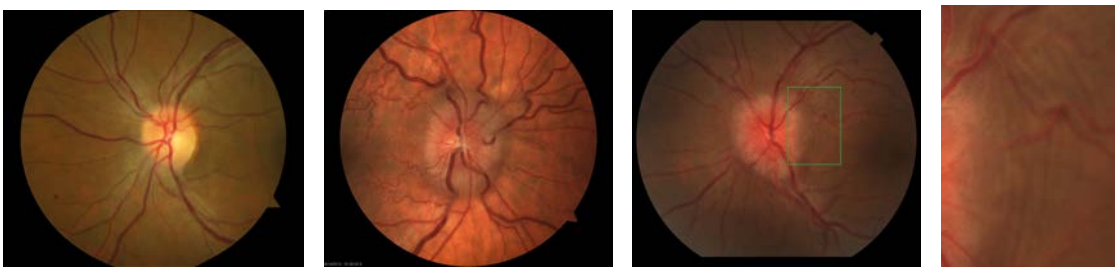


Figure 2.2: Fundus images of the ONH with Frisén grades of 0, 2, and 4, from left to right.

of 1, 2, or 3 could be justified, and for reasons such as these, there tends to be inter- and intra-observer variation in the assignment of Frisén grades [11].

Some other common indicators of optic disc swelling, in general, which are not considered part of the Frisén grading scheme include a twisting and/or swelling of the veins [12] and folds in the retina or choroid [8] (Fig. 2.3). The twist in the veins is a result of the ONH swelling, and does not contribute to the cause [13]. The swollen optic nerve creates a pinching of the veins as they leave the retina, causing a higher pressure in the veins inside the eye. As such, it would seem likely that some quantitative assessment of the twist and swelling of the veins, compared to the arteries, could be a means of assessing the severity of the swelling using only the information provided in the image.

Additionally, folds around the optic disc [8] come about due to tensile pressure generated by the ONH swelling. The presence of folds is an effect rather than a cause of swelling, but some quantitative measurement of fold characteristics could indicate severity or possibly cause. There are three main types of folds, and they are peripapillary wrinkles, retinal folds, and choroidal folds, and each of them is more easily visualized in SD-OCT images than in fundus images [8]. In the case of choroidal folds, this is particularly true as they occur on and between the retinal



(a) (b) (c) (d)

Figure 2.3: A mild case of swelling (a), absent of significantly twisted vessels or retinal folds. An example of an image with significantly twisted and swollen veins (b). An example of retinal folds (c). A blow-up of the green box to more clearly see the folds (d).

pigment epithelium (RPE) and Bruch's membrane (BM) layers, which are layers that are not normally visible in fundus images.

2.2 Pseudopapilledema

Pseudopapilledema can be similar to papilledema in appearance. However, it is often a benign condition that requires only routine review [14]. Not a lot is known about the cause of pseudopapilledema, apart from the fact that it is likely a congenital condition and often includes the presence of calcified deposits called drusen [15]. One thing that is certain is that pseudopapilledema does not come about due to raised intracranial pressure, and thus is not the result of a hemorrhage, tumor, concussion, or any of the possible causes of true papilledema. As such, it is vital to efficiently make this distinction between papilledema and pseudopapilledema early on to avoid unnecessary and costly treatments and, of course, alleviate anxiety in the patient.

One indication of real papilledema is the shape of Bruch's membrane (BM). The rise in ICP tends to cause BM to bend towards the vitreous [9], which can be detected through SD-OCT scans. This, along with the presence of drusen, can aid in differentiating between papilledema and pseudopapilledema.

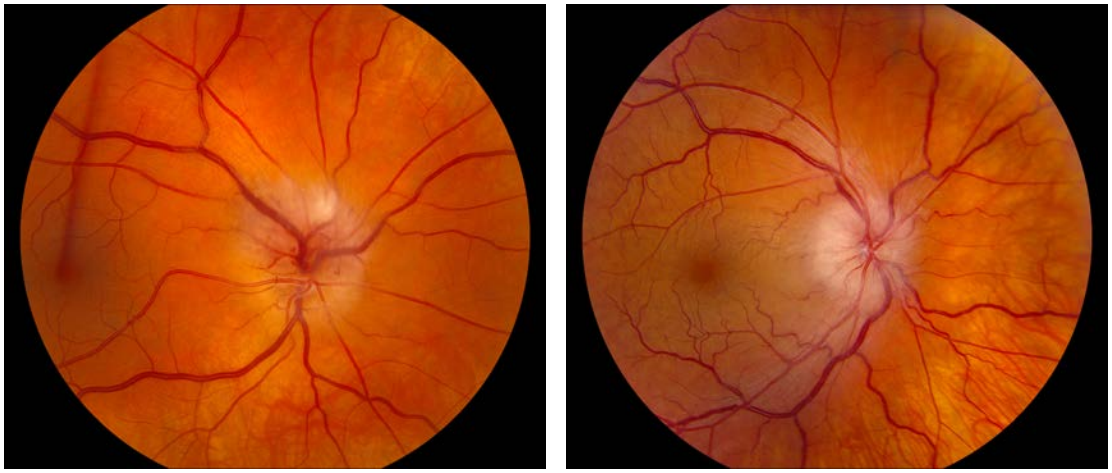
2.3 Anterior Ischemic Optic Neuropathy

Like any part of the human body, a reduction in blood supply leads to problems in the eye. In the case of ischemic optic neuropathy [16], this can lead to blindness or severely impaired vision. Ischemia of the optic nerve is typically categorized as either anterior (AION) [17] or posterior (PION) [18], which corresponds to the region of the optic nerve, relative to the optic nerve head, where the cause of the ischemia is located. While the problem tends to mostly affect people over the age of 50, it is possible for any age to acquire it. AION is the most common form of ischemic optic neuropathy, but both AION and PION can be further divided into either having arteric or non-arteric causes. PION, specifically, has a third possible cause, which is due to surgical complications, however, of the two, this doctoral work deals entirely with AION.

AION has two further classifications, which is determined by whether the cause is arteric. Arteric AION is due to some form of vascular inflammation, typically giant cell arteritis, which requires prompt treatment to avoid vision loss. Non-arteric AION (NAION), which is the most common form of AION, is due to inadequate perfusion due to a rise in intraocular pressure or a drop in blood pressure, which is why its symptoms are most commonly first noted on waking in the morning. Some work has been done on ways to distinguish AION from papilledema by evaluating the RNFL thickness [19]. Additionally, much like pseudopapilledema, in cases of AION it is assumed that the BM would not bend towards the vitreous. And finally, AION tends to result in a chalky white swelling of the optic disc (Fig. 2.4). This gives a few options for singling out a classification for the cause in a case of optic disc edema.

2.4 Other Causes of Optic Nerve Edema

Apart from papilledema, pseudopapilledema, and AION, other causes of optic nerve edema can include infiltrative or inflammatory disorders, such as venous sinus thrombosis, sarcoid optic neuropathy, infiltrative optic neuropathy, subarachnoid



(a) (b)

Figure 2.4: An example of AION (a) and papilledema (b). The ONH total retinal volume (section 3.2) of these two images were 18.14 and 18.04 mm^3, respectively.

hemorrhaging, Grave's compressive optic neuropathy, Leber's hereditary optic neuropathy, neuroretinitis, and optic nerve sheath meningioma. While this doctoral work intends to lay the groundwork for classifying as many of these as possible, no known dataset includes a sufficient quantity of images with swelling due to each of these causes.

2.5 Datasets

The University of Iowa Hospitals and Clinics (UIHC) is involved in an ongoing data collection effort to obtain fundus/SD-OCT image pairs taken on the same day from patients with optic disc edema. As of October 30, 2017, there has been usable SD-OCT/fundus image pairs acquired from a total of 88 patients, 63 with papilledema, 14 with NAION, and 11 with optic disc edema for other reasons. Additionally, the Idiopathic Intracranial Hypertension Treatment Trial (IIHTT) [20] is a multi-centered clinical trial used to ascertain the effects of acetazolamide in cases of papilledema. While the IIHTT dataset does not serve as a way to determine the cause of optic disc edema, as the dataset consists entirely of papilledema, a few of the algorithms in this

doctoral work were developed on it.

CHAPTER 3
TECHNICAL BACKGROUND

3.1 Assigning Frisén Grades to Fundus Images

The work of Echegaray et al. [5] was to determine a feature set for the semi-automatic assignment of Frisén grades to fundus images. The most important feature in their work was the vessel discontinuity index (VDI), which was a simple count of discontinuities in the vessel segmentation. The segmentation was acquired by enhancing the vessels using a method [21] that considers a local second order Taylor expansion to generate a Hessian matrix. An eigenvalue analysis of the Hessian was used to detect the presence and direction of any local tubular structures (i.e. vessels). Once the vessels were enhanced, a segmentation was achieved via a local entropy thresholding method [22].

The second most important feature found was the obscuration around the disc margin. This was found by dividing the ONH into sectors and assessing the 'blurriness' of each sector. Related to the obscuration around the disc margin were the disc margin ratios. These were found by finding the points with the greatest radial intensity gradient from the ONH center, every 2 degrees. These points were fit to an ellipse, and the shortest distance between each point and the edge of the ellipse was found and divided by the distance between the opposite point and its shortest distance to the edge of the ellipse.

All of the most important features found by Echegaray et al. were variants of these features. However, the main drawback of this approach is the fact that it predicts Frisén grade, which is ordinal and subjective (more on this in section 3.2).

3.2 SD-OCT Layer Segmentation

As far as segmentation methods go, dynamic programming is probably the easiest to understand and implement. This method is fast for 2D images, but has three key drawbacks, in that it is time consuming to enforce circularity constraints, cannot be

used to segment multiple boundaries simultaneously, and cannot be feasibly extended to surface segmentations in 3D images. Somewhat recently [23], a new segmentation method has been presented which represents an image as a graph, and uses a min-cost closed set algorithm to generate a segmentation. This has a number of advantages, including ease of enforcing circularity constraints, the simultaneous segmentation of multiple layers, and, of course, the fact that it can be extended to any number of dimensions. As such, it is a natural approach to layer segmentation (Fig. 3.1) in SD-OCT images of the optic nerve.

In fact, this method has already been developed [24], by segmenting a cost image generated through a combination of edge-based and region-based sub-images. The edge-based sub-image uses a combination of a Gaussian and Sobel filters to create a cost image, which is translated into a graph by assigning the cost of each pixel to the cost of a corresponding node. A surface feasibility constraint, which assures connectivity, is applied as well. This can also be performed on images that have been down-sampled in the z-direction, with the results used to localize each higher-resolution search for better processing speed [7].

As the scaling of SD-OCT images in physical space is commonly recorded when the image is taken, a segmentation of the top and bottom layers of an ONH conveniently allows for an absolute measurement of the volume of the ONH. Since the swelling of the ONH is, in a sense, defined by the additional space the ONH takes up, which is another way of saying 'volume,' it stands to reason that the total retinal volume (TRV Fig. 3.2) as computed through this segmentation would serve as a far better measure of optic nerve edema than the Frisén scale [7]. Not only is TRV continuous, but it is also reproducible. While it may be true that there is some natural variation in the TRV due to variation in attributes such as patient size, one of the main assumptions in this doctoral work is that the TRV is the best measure by which optic nerve edema severity can be assessed.

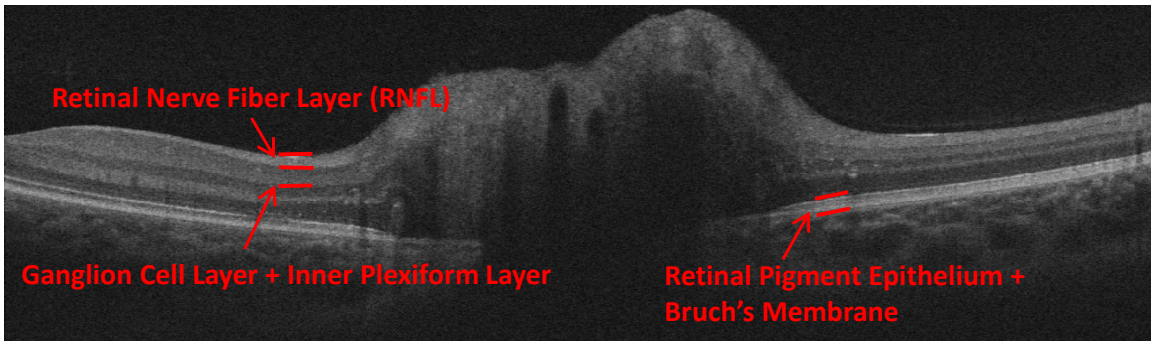
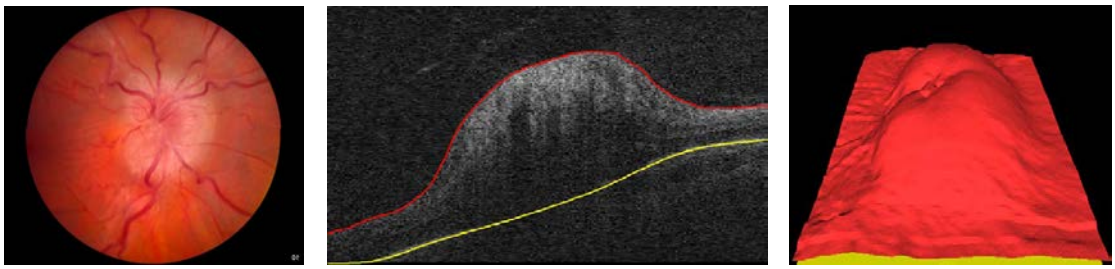


Figure 3.1: Indication of the ONH layers of interest.



(a) (b) (c)

Figure 3.2: Example of how TRV is computed. The original fundus image (a). A slice from the SD-OCT segmentation (b). A 3D rendering of the full segmentation (c). The TRV is computed as the physical space between the red and yellow surfaces - in this example, which corresponds to a Frisén grade of 4, the TRV is 22.61 mm^3.

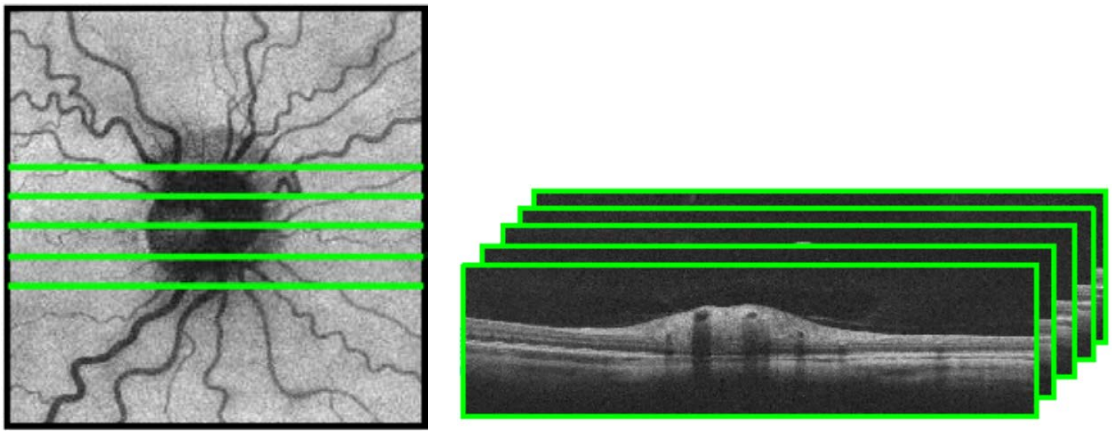


Figure 3.3: A depiction of where a 5-line raster scan comes from.

3.3 Quantification of BM Shape

As mentioned in sections 2.2 and 2.4, a rise in intracranial pressure can manifest at the eye and that increase in pressure tends to bend the BMO towards the vitreous. Some work has been done to quantify this bending [2]. However, the method used requires a somewhat uncommon use of SD-OCT images, called 5-line raster scans [9]. The use of high density 5-line raster scans offers the advantages of higher image quality and simplicity in identifying the center-most scan (Fig. 3.3), where the shape changes of BM are most pronounced.

The shape was computed by eigenvalue decomposition of 20 points marked on the Bruch's membrane boundary (Fig. 3.4). There were three eigenvectors associated with the decomposition, but two are not considered relevant to a diagnosis, and their only purpose is to limit the influence their variation has on the third shape parameter (Fig. 3.5). These two eigenvectors assessed the width (size) of the BMO and the lateral evenness of one side relative to the other. By quantifying these two attributes, it leaves the bending of BM to be quantified without their influence, which constitutes the third[1] eigenvector coefficient. This coefficient describes the bending

[1]Of the three shape parameters, this is commonly listed as the 2^{nd} coefficient, due to the fact they are ordered by weight, and the bending tends to be the second-most descriptive component of

of BM towards or away from the vitreous, which, as described in sections 2.2 and 2.3, may be a significant indication as to whether or not a swollen optic disc is due to raised intracranial pressure.

The points were marked in a semi-automated fashion. That is, only the points corresponding to the BMO were marked. Then, a 2D version of the segmentation algorithm discussed in section 3.2 was used to determine the location of Bruch's membrane, using the marked points as a 'very low cost' location to force the segmentation through them. The remaining 18 points were then automatically generated by placing them on the layer segmentation equidistant from each other, to cover a total of 2.5mm in each direction. Then, using Procrustes analysis and an eigendecomposition on the entire dataset, the shape of each 20-point set was approximated by its most relevant eigenvectors. As it happens, the eigenvector associated with the second-most amount of variance represents the bending of BM towards the vitreous in prior work. In any event, once a shape model has been computed, and the relevant shape parameters determined, any SD-OCT image can be modeled for its shape.

3.4 Vessel Segmentation

Automated vessel segmentation is a topic that has had a rather large amount of work put into it, with a variety of approaches. The method used in this thesis involves a ridge detection algorithm [25]. In an image, a ridge can be thought of in much the same way one might think of a ridge geographically - that is, a thin but long peak among an otherwise homogeneous area. The applications of finding ridges should be somewhat clear in regards to detecting vessels, as that is precisely what vessels tend to be. This approach first applies a Gaussian derivative, then uses the Hessian to determine the direction and strength of an assumed ridge at each pixel. A gradient is stepped once in the direction of minimum Hessian response and the sign of the difference in the gradient before-and-after the step is taken, and multiplied by the

the shape.

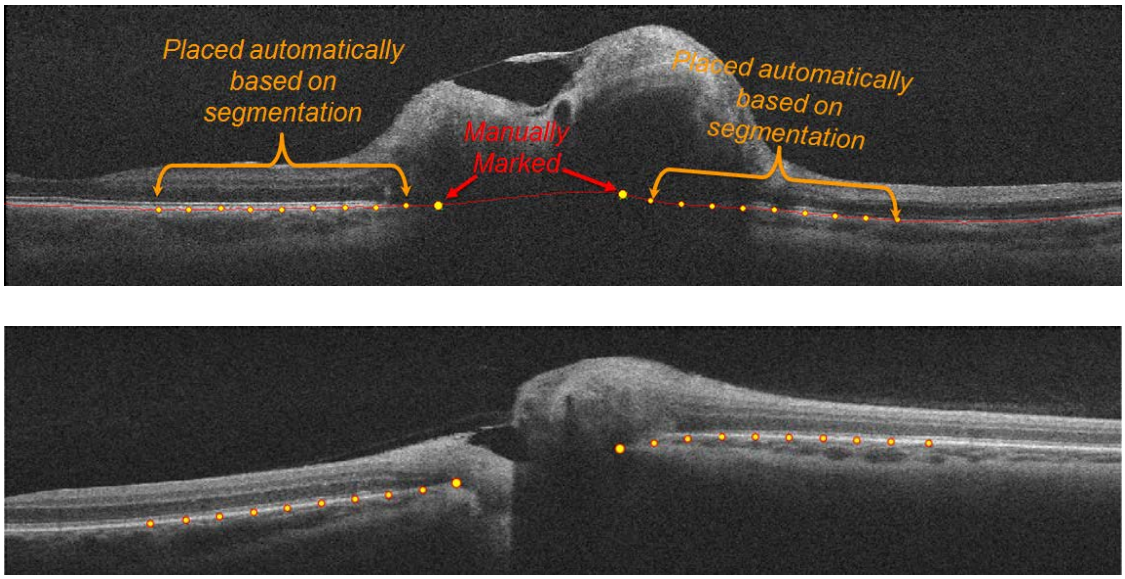


Figure 3.4: Examples of how the shape of BM is modeled in 2D images. An eigenvalue decomposition would be performed on the points to reconstruct their shape using the eigenvectors derived from a Procrustes analysis and PCA across 116 similarly marked images.

	$c_{\mathbf{s}_{Lj}} = -3$	$c_{\mathbf{s}_{Lj}} = -2$	$c_{\mathbf{s}_{Lj}} = -1$	$c_{\mathbf{s}_{Lj}} = 0$	$c_{\mathbf{s}_{Lj}} = 1$	$c_{\mathbf{s}_{Lj}} = 2$	$c_{\mathbf{s}_{Lj}} = 3$
$\mathbf{s}_{L1} = \bar{\mathbf{s}} + c_{\mathbf{s}_{L1}} \cdot \sqrt{\lambda_{L1}}\mathbf{e}_{L1}$							
$\mathbf{s}_{L2} = \bar{\mathbf{s}} + c_{\mathbf{s}_{L2}} \cdot \sqrt{\lambda_{L2}}\mathbf{e}_{L2}$							
$\mathbf{s}_{L3} = \bar{\mathbf{s}} + c_{\mathbf{s}_{L3}} \cdot \sqrt{\lambda_{L3}}\mathbf{e}_{L3}$							

Figure 3.5: Shape parameters and their ranges. The first and third describe the BMO width and lateral evenness, respectively. The second coefficient describes the bending towards or away from the vitreous [2], which is the parameter of most interest in this doctoral work.

sign of the eigenvalue of the maximum Hessian response. Essentially, this looks at the most likely direction of a ridge at each pixel, and checks for ascent in one direction normal to ridge direction, and ascent in the opposite direction. If the ascent/ascent or descent/descent combination is found, that pixel is marked as a 'true' for being a part of a ridge, and the sign of the Hessian eigenvalue is then applied to determine whether it is a local minima or local maxima - effectively resulting in a trinary image of all ridges.

Of course, not every ridge is going to be a vessel, and not every pixel of every vessel is going to be a ridge. The next step involves partitioning every ridge pixel into groups. The rules of the group are designed to check for proximity and direction. There are three metrics that are used to determine if each pixel is to be assigned to a group, or if it should form its own. In addition to proximity and direction, there is also a check for a parallel component. Two parallel ridges may be two vessels, which would share a direction component, by definition, but could also be within the limits of proximity. To test for this, the product is taken of a unit vector in the same direction of the minimum Hessian eigenvector with a unit vector in the direction of the two pixels to be potentially grouped. If the resulting vector magnitude is sufficiently high, it can be said that the two pixels are not parallel and, given the other two metrics being satisfactory, belong to the same group. From there, all non-ridge pixels are further assigned to a group based solely on proximity.

A kNN classifier is applied to each group. Features are built based on a transformation that samples a line through each point in a group perpendicular to the ridge direction. The features include the height (pixel intensity at ridge), width (distance between strongest gradient locations on sampled line through ridge points), height-to-width ratio, edge strength, edge strength-to-width ratio, average edge height, height minus edge height, and the ratio of height to edge height. Additional features include the straight-line distance of each set, length, curvature, local average, standard devia-

tion, green-to-red frame local average ratio, and ridge strength (average eigenvalue).

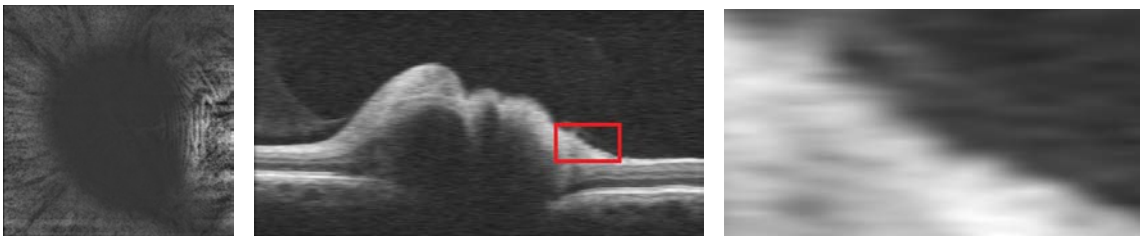
CHAPTER 4
DETECTION AND QUANTIFICATION OF FOLDS IN SD-OCT IMAGES (AIM 1)

4.1 Introduction

The retinal area around the optic disc sometimes responds to the biomechanical physical stress of swelling by exhibiting visible wrinkles (or folds) of the tissue around the optic disc [26]. The presence and variety of retinal and choroidal folds [27] (Fig. 4.1) in cases of papilledema give rise to the possibility of different causes of optic disc swelling resulting in different types and severities of folds. Moreover, if the swelling is congenital (i.e. pseudopapilledema), it is much more likely that there will be no folds at all. In this chapter, the detection of folds is explained. The quantification of folds, however, is reserved for practical use, and is discussed in greater detail in Section 5.2.3 where it is used for distinguishing between causes of optic disc edema.

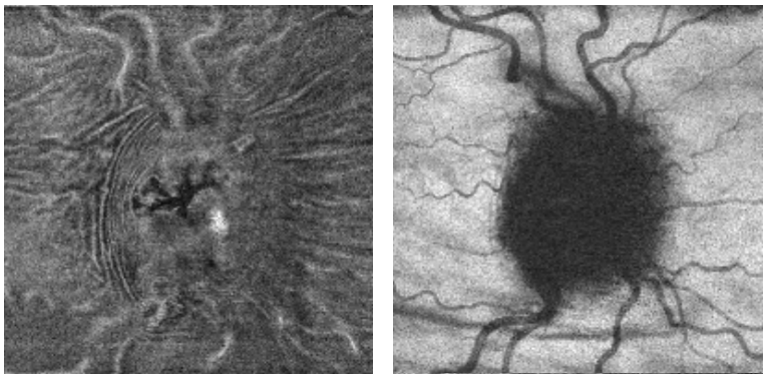
4.2 Fold Enhancement

The layers in the SD-OCT images of the optic disc are first segmented [28] using a graph-based method developed for optic disc swelling. The layer segmentation algorithm has a smoothing feature that can complicate its direct use to visualize folds, as the smoothed surface cuts through the local average of minor fluctuations

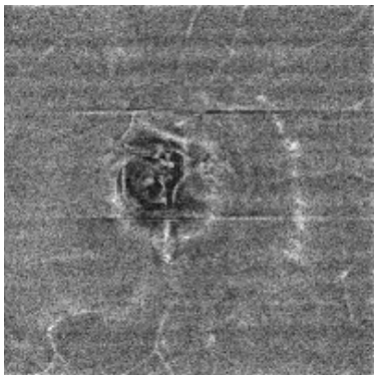


(a) (b) (c)

Figure 4.1: Cross-sectional example of folds. The modified *en face* image (a) shows folds in the topology. The cross sectional image (b) through the central slice, and a magnification of the cross sectional image centered on the folds (c).



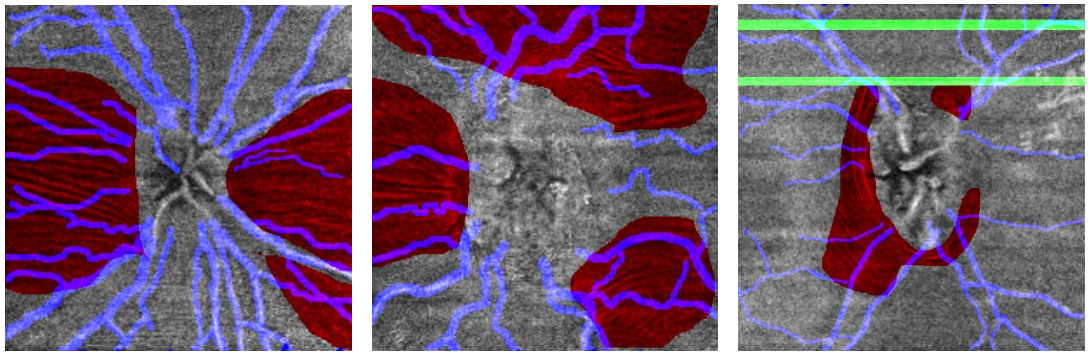
(a) (b)



(c)

Figure 4.2: (a) Fold-enhanced image generated by averaging within 7 pixels of the ILM boundary. (b) Vessel image by averaging across the RPE layer. (c) A fold-enhanced image without visible folds.

along the ILM. As such, a fold-enhanced 2D image is generated by taking the average pixel value within 7 pixels of the ILM (a total span of about 29.3 microns). The result is a 2D image in which folds appear to be thin, tube-like structures. One obstacle is that there are other thin, tube-like structures we expect to see in these images – the blood vessels and artifacts. To robustly account for blood vessels, an *en face* image of the RPE complex layer is also generated (where vessels are most pronounced). Features from both the ILM and RPE (Fig. 4.2) images are used for classification.



(a) (b) (c)

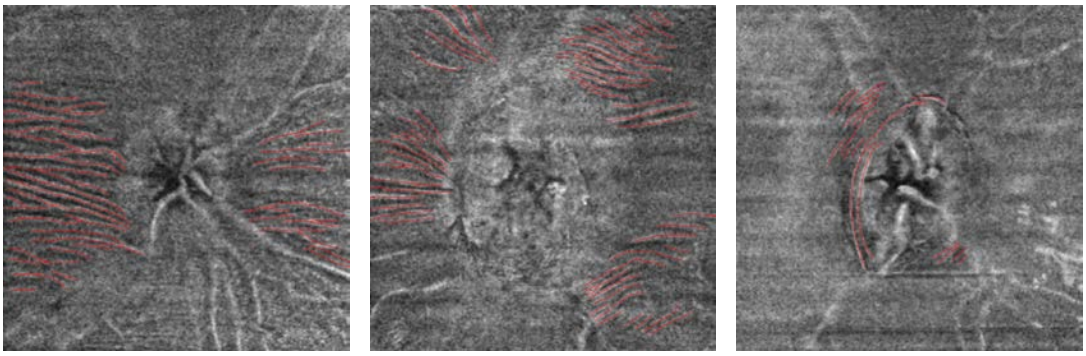
Figure 4.3: Some truths marked for the first revision of this fold detection work as presented at MICCAI in 2017 [1]. Red areas indicate areas with folds. Blue indicates vessels. Green indicates image artifacts.

4.3 Reference Standard

SD-OCT images acquired by the University of Iowa Hospitals and Clinics were converted to fold-enhanced images and marked. This occurred in two phases. The first, initial, phase involved a regional marking (Fig. 4.3) but carried with it the issue of precision being a major limiting factor to the results. It is difficult to precisely define where a folded area stops being a folded area, and marking a region makes it critical that the boundary is as accurate as possible. Since then, the folds were marked again such that the peak of each fold was marked (Fig. 4.4). However, due to limitations in resolution, the fold-enhanced images were upsampled by a factor of four prior to marking.

4.4 Feature Extraction

The thin tube-like appearance of folds in the fold-enhanced image suggests a detection method similar to one would use for a vessel segmentation, but also accounting for parallel repetition. The method employed here involves feature extraction and pixel-based classification using random forest regression. The full feature lists can be seen in Table 4.1 for the initial iteration of this algorithm, and Table 4.2 for the



(a) (b) (c)

Figure 4.4: The truths marked for final fold detection work. While the vessels and artifacts were also marked, a separate classifier was not run on them, and fold pixels aligned with either a vessel or artifact were simply omitted from the truth.

final version. Gabor responses [29] are computed at 15° intervals, with wavelengths ranging from 4 to 25 pixels. These are applied to the fold-enhanced image and the *en face* image of the RPE, each upsampled by a factor of 4 since it is not expected that the physical wavelength of the folds will always be a whole integer value. In physical space, this means the Gabor wavelengths range between 30 and 187.5 microns. The maximum Gabor response across all wavelengths and orientation angles, M_G, for each pixel location in the fold-enhanced image is extracted and summed to the maximum Gabor response from the inverse of the fold-enhanced image, M'_G. The orientation and coherence [30, 31] images are computed for M_G, M'_G, and their sum $M'_G + M_G = I_G$ (Fig. 4.5, 4.6).

A line of 29 sample points are extracted from each pixel location in I_G spanning 435 microns (roughly half-pixel intervals), perpendicular to the orientation. The newest version of this algorithm involves truth images that were upsampled by a factor of 4, and the natural way to work with that is to also upsample the test images themselves by 4. To keep this line of 29 sample points consistent with its purpose, in the updated algorithm it needs to be increased to 113 sample points to

span the same 435 microns. In regions with folds present, this sample would have a sinusoidal element to it. This is quantified this by taking the Fourier transform of each sample, and the Fourier transform of the derivative of each sample (to remove linear components). Disregarding the 0^{th} Fourier coefficient, the highest Fourier response should correspond to the frequency of the sinusoidal element, if one exists. For the first iteration of this work, the highest Fourier element was used to create the magnitude, phase, and frequency of a test signal, and the test signal was removed from the sample. The standard deviation of the leftover signal is then taken as a feature to determine how dominant the sinusoidal element was. However, for the updated version, the highest Fourier element was used to create a fit to a sinusoid, and the RMS error of the fit was used as a feature. Additionally, the Fourier magnitudes are used as features, directly.

In addition to detecting folds, in the initial version of this algorithm, vessels and artifacts were also detected to mask out the fold response. The Gabor response to the vessel image (RPE complex *en face* image) seemed to be sufficient for vessel detection. As the artifacts to be detected tend to occur across the image, rather than in isolated areas, the horizontal average of the vertical derivative is used as a feature across the entire image - i.e. for this feature, every pixel on the same horizontal (similar to Fig. 4.6(f)) has the same value for a given image. Additionally, as the artifacts often appear to be the result of a horizontal shift when the image is captured, the displaced location of the minimum average vertical derivative when each horizontal slice is shifted across its full range is used as a feature. Finally, since the direction of artifacts is close to zero, the cosine of the direction divided by the standard deviation across the horizontal was used as a feature.

Feature count	Description
1	RPE complex image

2	Radial distance from center determined automatically [7]
3-24	Coherence of Gabor responses to upsampled fold-enhanced images (each pixel-wavelength used as a separate feature)
25	Coherence of average Gabor response across all wavelengths
26	Gaussian filter applied to 25
27-48	Coherence of Gabor responses to upsampled inverse of fold-enhanced image
49	Coherence of average Gabor response to inverse fold-enhanced image
50	Gaussian filter applied to 49
51	Local maximum in 7×7 window of fold-enhanced image
52	Local minimum in 7×7 window of fold-enhanced image
53	Range in a 7×7 window (i.e. features 52 - 51)
54	Normalized version of feature 53 (i.e. (52-51)/(52+51))
55	M_G
56	M'_G
57-78	Gabor responses of RPE complex image
79-100	Gabor responses of inverse of RPE complex image
101	Average of 57-78
102	Average of 79-100
103	Orientation of I_G (sum of average Gabor responses)
104	Coherence of I_G
105	Standard deviation of cross section after top frequency removed
106	Standard deviation of derivative of cross section after top frequency removed

107-120	Sorted magnitude of Fourier response of cross section
121-133	Sorted magnitude of Fourier response of derivative of cross section
134	Maximum Fourier response of cross section
135	Difference between max and second rank Fourier responses of cross section
136	Difference between max and third rank Fourier responses of cross section
137 - 139	Features 134-136 repeated for derivative of cross section
140	Vertical displacement measure of fold-enhanced image
141	Local max of 140
142	Vertical displacement measure of RPE complex image
143	Local max of 142
144	Vertical derivative of fold-enhanced image
145	Vertical derivative of RPE complex image
146	Product of features 144 and 145
147	Locally normalized I_G
148	Coherence of thresholded and skeletonized I_G
149	11×11 median filter applied to I_G

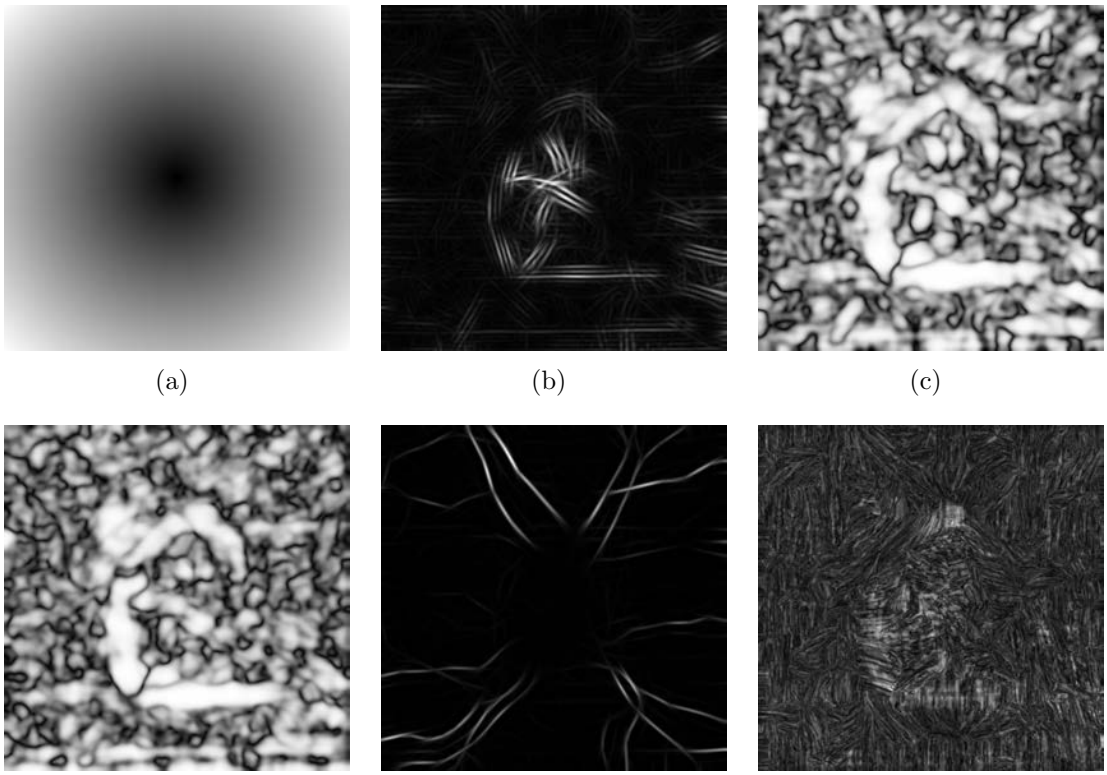
Table 4.1: Complete list of features used for initial fold classification as reported at MICCAI 2017 [1].

4.5 Classification

The vessels, folds, and artifacts were marked for each image to generate truths (Figs. 4.3, 4.4). For the initial work, these were classified using a common feature

Feature count	Description
1	RPE complex image
2	Radial distance from center
3	M_G
4	M'_G
5	I_G
6-51	Gabor response (evens) and coherence (odds) of M_G at increasing frequencies
52-95	Gabor response (evens) and coherence (odds) of M'_G at increasing frequencies
96	Max Gabor response across all wavelengths and orientations of RPE complex image
97	Max Gabor response across all wavelengths and orientations of inverse RPE complex image
98	Sum of (96) and (97)
99	Magnitude of sinusoidal fit using Fourier transforms
100	RMS Error of sinusoidal fit
101	Peak count for each signal
102 - 104	(99 - 101) repeated for derivative of signal
105	Horizontal disjoint measurement for fold-enhanced image
106	Horizontal disjoint measurement for vessel image
107	Cosine divided by standard deviation of direction

Table 4.2: Complete list of features used for final fold classification.



(a) (b) (c)

(d) (e) (f)

Figure 4.5: Some selected features (a) Radius (b) I_G (c) Coherence of M_G (d) Coherence of M'_G (e) Gabor response to vessel image (f) Magnitude of fit to sinusoid normal to direction.

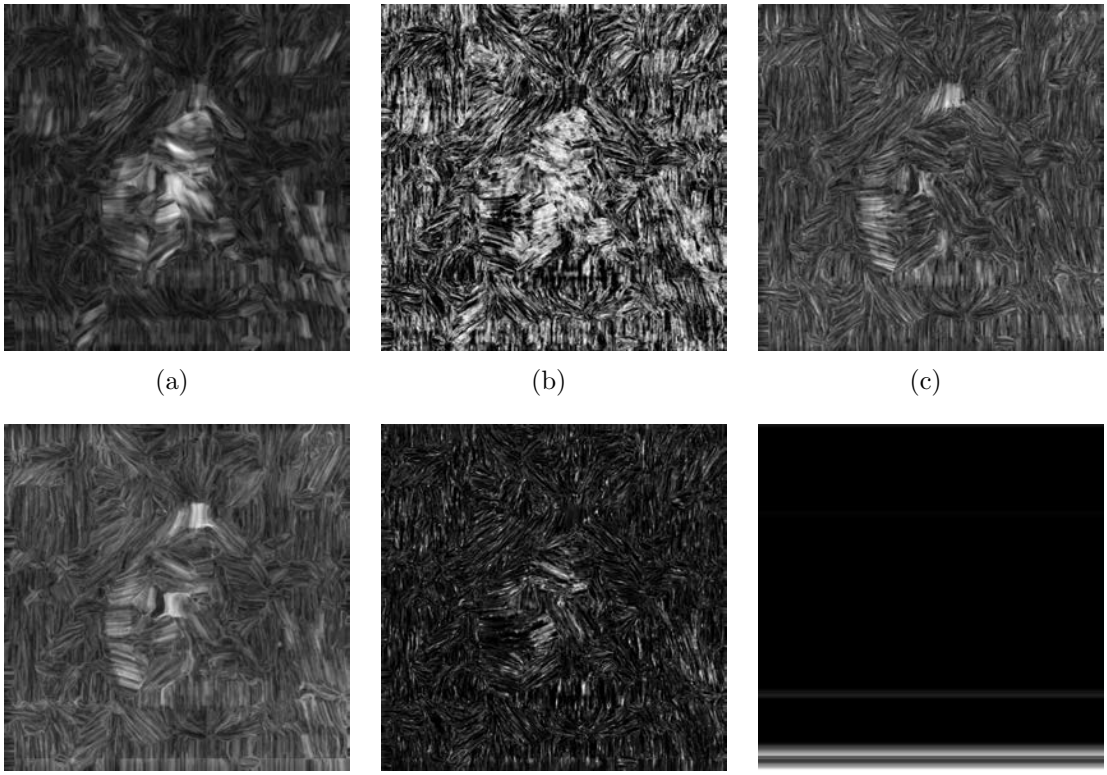


Figure 4.6: Some selected features (a) RMS error of fit sinusoid normal to direction (b) Peak count of signal normal to direction (c) Magnitude of fit to sinusoid of derivative (d) RMS error of sinusoid fit do derivative (e) Peak count for derivative of signal normal to direction (f) Cosine divided by standard deviation.

set, but were trained as separate classifiers. The features were extracted from the fold-enhanced and RPE complex images, and trained in three leave-one-patient out random forest regression analyses with 10,000 trees, each ten decisions deep, with a true set to 1 and a false set to 0. Once trained, the entire feature set from the left-out image was tested for each of the three classifiers, resulting in a pixel-level probability map. The results from the vessel and artifact classifications were used to mask (omit) positive results from the fold classification, while the vessel and artifact truths were used to mask the fold truth. The classifier was trained on 120 samples taken at random locations from each feature set of the 20 patients acquired from The University of Iowa Hospitals and Clinics (minus one left-out patient) to include an equal mix of folds, vessels, artifacts, and background.

For the updated version of this algorithm, only one classifier was trained, after masking out folds that overlapped vessel and artifact truth locations. There are still 120 samples taken per image for training, and they are still evenly distributed among fold, vessel, artifact, and vessel locations, but the difference is that there is only one classifier instead of three. The idea is that the features that resulted in a high probability of vessels and artifacts from the initial work will, instead, result in a low probability of folds if the vessel and artifact truths are used to mask the fold truth. While this carries a small risk of missing some folds that only overlap vessels, the main point of this work is not to track every fold but to determine a simple yes or no to the question of fold presence to distinguish between causes of optic disc edema.

4.6 Results

A total of 20 of the 88 usable SD-OCT images acquired from the University of Iowa Hospitals and Clinics (section 2.5) were marked for the presence of retinal folds. Since the results are to be a probability map, and the marking of truths is a somewhat tedious process, there is no need to use all 88 of the images, as there is effectively a sample for every pixel. Combined with the fact that the images are upsampled by 4

for greater resolution in marking the truth, this means that with 20 images there are already almost 13 million samples available for testing.

The initial pixel-based classification with a regionally marked truth was reported at MICCAI of 2017, with an AUC of 0.804 achieved [1]. Example probability maps can be viewed in Fig. 4.7. This result was obtained by simultaneously evaluating three thresholds at once. The fold, vessel, and artifact probability maps were each adjusted through their full range and a maximum true positive rate found for each false positive rate. The resulting ROC chart can be found in Fig. 4.9(a). With the second attempt, only the folds were predicted, with the vessels and artifacts serving as a mask (Fig. 4.8). However, three ROC charts were created from this. The first is the obvious one, which is a simple per-pixel threshold, which results in an AUC of 0.896 (Fig. 4.9(b)). The second and third both accounted for the fact that, despite the attempt to mark the folds precisely, there is still going to be a little bit of discrepancy, in that it is never going to be clear exactly how much of a fold ought to be marked. Essentially, the skeleton was marked in the truths, but if a probability map shows that a detected fold is slightly thinner than the marked skeleton, that shouldn't be counted as a problem - as the point is to detect for presence, not to track precisely. As such, an evaluation was made in which each fold marked in the truth was treated as a whole object, and when the probability map was thresholded, if any pixel in that object was found, the whole object was counted as a true positive. This still leaves the question of how to handle false positive results, and two approaches were taken for this. The first was to count the total number of objects detected that did not belong to a marked truth. The second is to treat the false positive just as it is normally treated - that is, pixel-based. These two methods can be summarized by noting that the coordinates of the ROC curve with the object based true positive rate and pixel-based false positive rate is represented by $(P(t)/P_{f'}, N(t)/N_f)$, where t is the threshold, $N(t)$ is the number of folds detected at each threshold, N_f is the

total number of fold objects in the dataset, $P(t)$ is the number of non-fold pixels detected for each threshold, and $P_{f'}$ is the total number of pixels that are not part of a fold object. A similar representation could be made for the case where the non-fold objects constitute the false positive rate, except instead of pixels for the x-coordinate, the total count of non-fold objects would be used. These two resulted in an AUC of 0.976 (Fig. 4.9(c)) and 0.964 (Fig. 4.9(d)), respectively.

4.7 Discussion

Despite the overall decent results, there are still some shortcomings that could be addressed. Particularly, the fact that there is no existing standard for the finding and marking of folds manually adds a fundamental ambiguity to the task of finding them automatically - no matter how reasonable a ground truth appears to be, unless it is vetted in some way, any classification based on that truth is going to be subject to justifiable scrutiny. That said, given the assumption that the truth was marked in a way that accurately captures both the presence and location of folds, the results here are quite good. The most notable problems seem to arise from exactly what one would expect: the vessels, artifacts, and the ONH. The vessels and artifacts are fairly well understood, though features were added in an attempt to alleviate that as much as possible. Indeed, this seemed to work fairly well, but qualitatively, it is far from ideal. Future work may consist of separate classifiers for both the vessels and artifacts, and using that to mask the probability map of the folds. This has the benefit of treating the vessel and fold problem as one that is already fairly well understood, thus allowing the focus of work to be strictly on fold features.

The ONH, however, is an entirely different matter. As shown in Fig. 4.10, the topology of the ONH is unpredictable with the method used in this doctoral work for fold enhancement. There could be any number of contours and patterns that present themselves, and indeed there are times when folds legitimately do appear on the ONH. For this doctoral work, the radius from the approximate center of the ONH

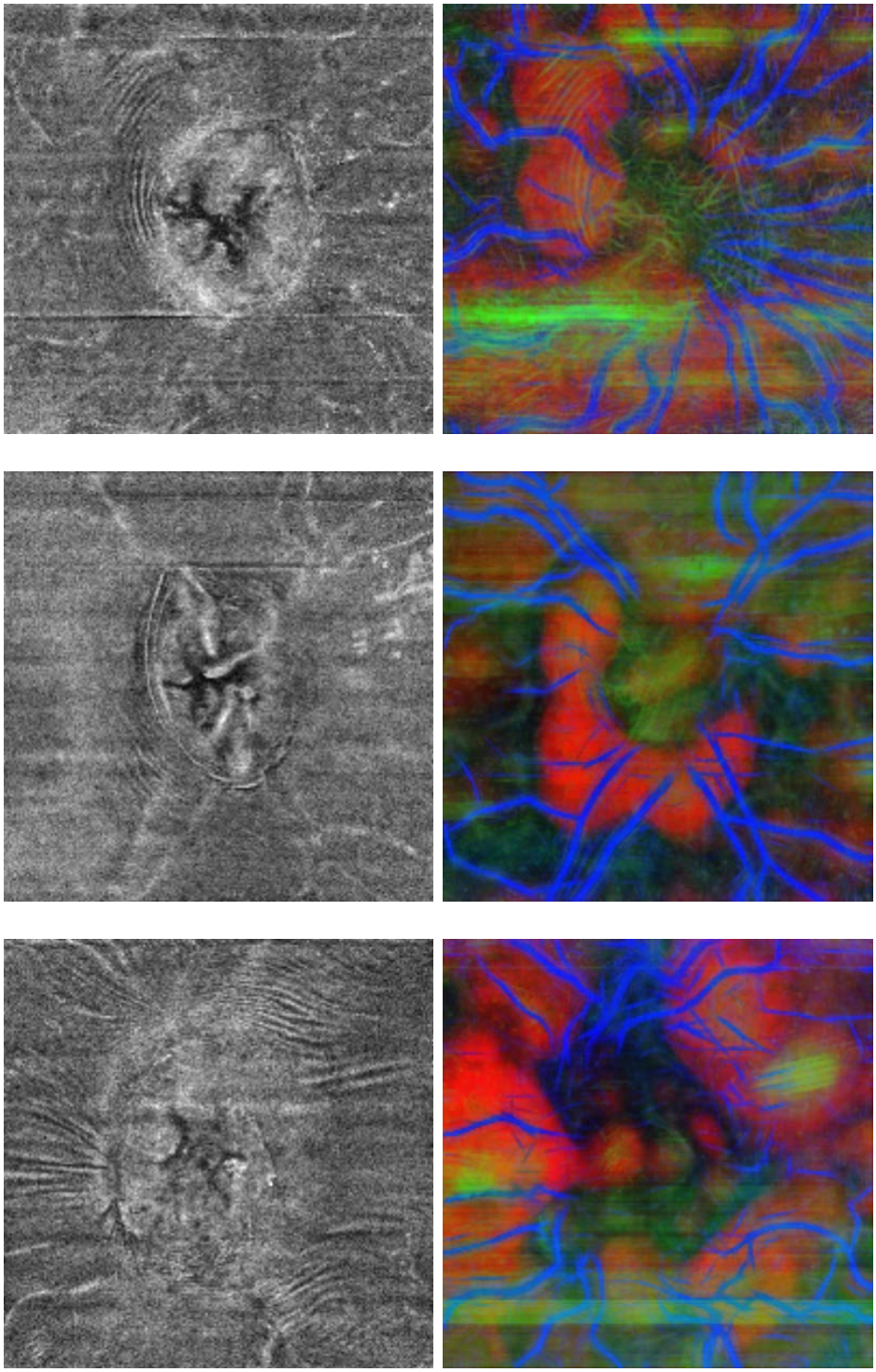


Figure 4.7: Fold enhanced image shown alongside resulting probability maps. Red corresponds to folds, blue to vessels, and green to artifacts.

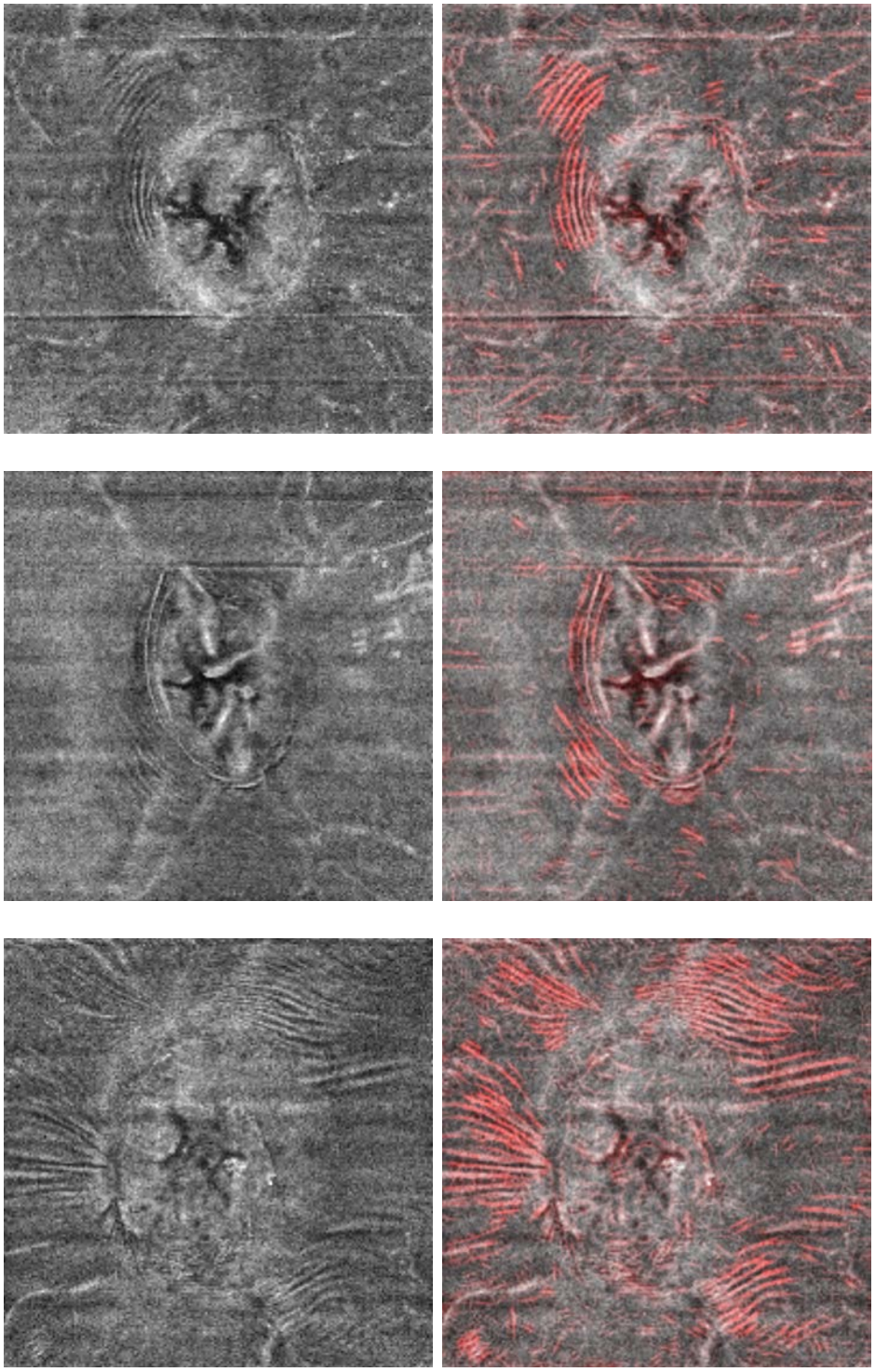


Figure 4.8: Fold enhanced image shown alongside resulting probability maps for the final iteration of this work.

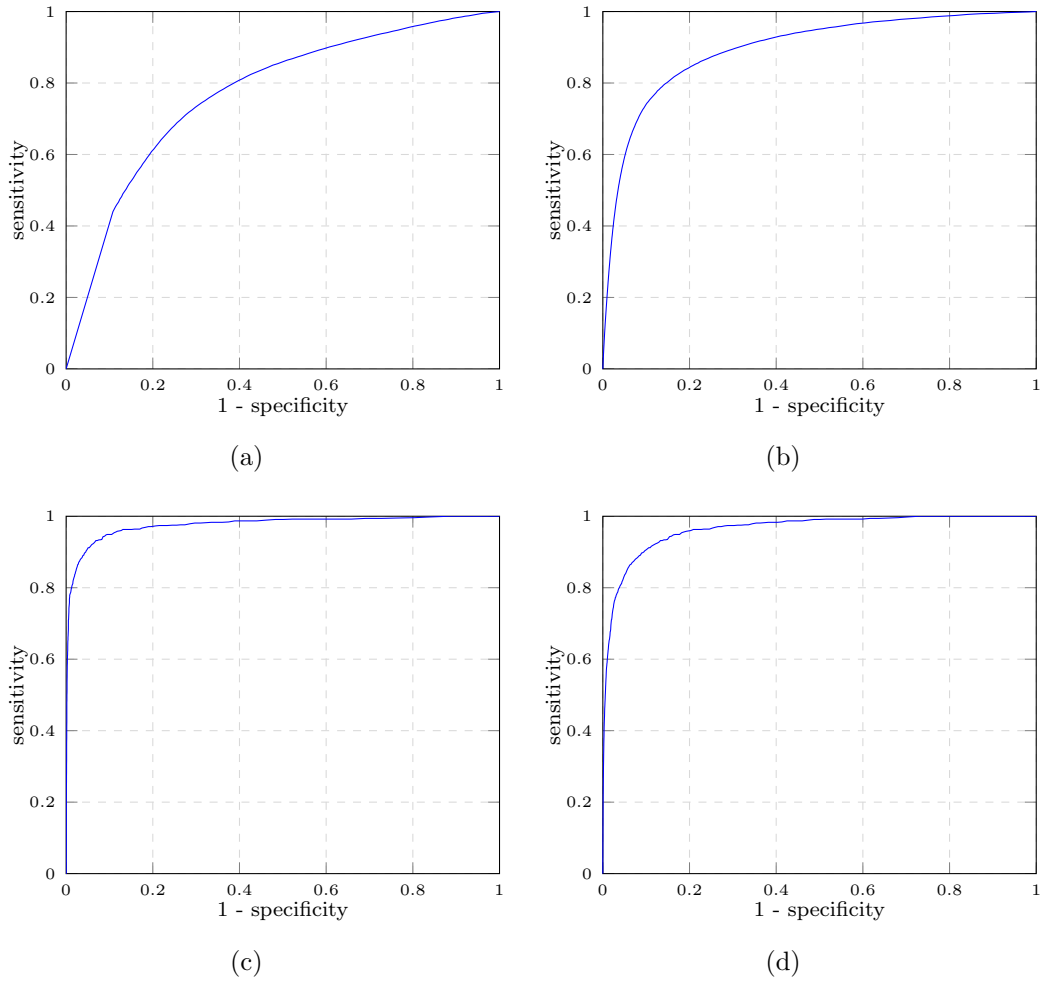
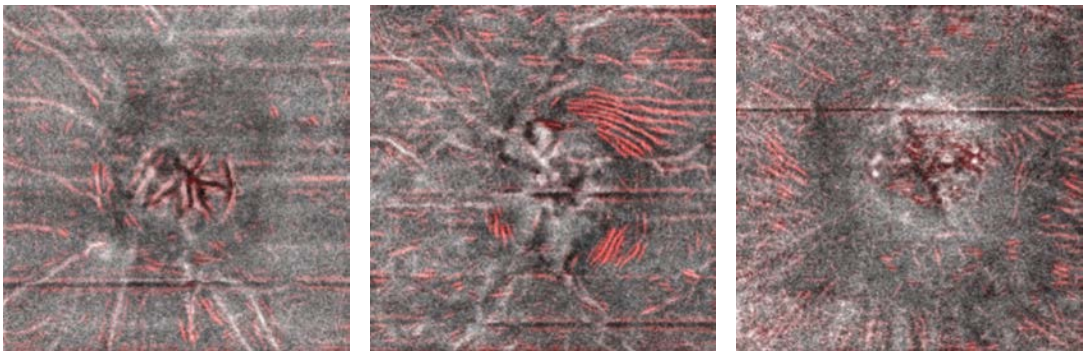


Figure 4.9: Resulting ROC curves of the initial classification technique (a) with an AUC of 0.804, reported at MICCAI in 2017 [1]. And ROC curves for the final classification technique, which includes traditional ROC generation with an AUC of 0.896 (b), ROC of objects with an AUC of 0.976 (c), ROC of objects for true positives and traditional for false positives with an AUC of 0.964 (d).

was used as a feature, and the ONH was effectively masked by simply not marking anything on the ONH as a fold. While this is reasonable in the sense that one cannot mark what one cannot see, it goes back to the problem of there being no standard way to visualize folds in the first place. However, this is somewhat less problematic than it may seem, as a reliable way to visualize folds on the ONH would necessarily include



(a) (b) (c)

Figure 4.10: Some additional probability maps. (a) and (c) had no folds marked in the truth, and despite a somewhat high presence of red markings, it can be seen that the red is not nearly as bright as the true folds in (b).

the side-effect of eliminating the unpredictable topology of the ONH in fold-enhanced images.

4.7.1 Categorization of Folds and Wrinkles

Ultimately, the main goal of this chapter is to determine if folds exist. However, there is clinical significance to categorizing folds, once found. Specifically, the orientation, which is discussed where it is assumed to be more useful (Section 5.2.3), is of interest to the determination of the cause of optic disc edema. It may be worthwhile to point out that orientation, spatial frequency, coherence, and other attributes are easier to obtain than the actual presence of folds, as some of these additional metrics are features used for fold classification. Clinically, folds are classified as either peripapillary wrinkles, retinal folds, or choroidal folds. Peripapillary wrinkles are defined as wrinkles within a half-disc diameter from the ONH center, whereas retinal folds are defined as folds greater than half a disc diameter from the disc. These categories, for example, could be assigned directly to any area found to have a fold by a slight alteration of features 2 and 107 from Table 4.2, and a segmentation of the opening to Bruch's membrane [32]. However, as it is explained in Section 5.2.3, it could be

expected that, apart from existence, orientation alone is to be the more important attribute for distinguishing the cause of edema.

4.7.2 Supplementary Topic: Fold Enhancement Techniques

In this subsection, the methods considered for creating the fold-enhanced image are discussed in detail. Since the automatic identification of folds is something of a largely unexplored area, in many ways the work in this thesis involved starting from scratch – though not truly as work has been done to identify folds manually [27] [26]. The first step is to enhance folds such that they can be more easily identified. There is no established method for doing so, and indeed no ground truth for such a thing to compare various methods for accuracy. The idea of creating an image before work can even begin is something of which its significance can be easily overlooked, as no matter how the images are produced or how reasonable they seem to look, there is always the chance that something is missing, and there is no way to tell without starting over. However, despite this, one needs to start somewhere, and the best option is to approach with reason and see where it leads. To that end, it only needs to be asked what folds are, and that answer is small fluctuations on a surface.

Each of the methods involved generating an *en face* image - as this deals with a surface it is both sufficient and far simpler to deal with a 2D image. The results are shown in Fig. 4.11, and are listed below:

(a) The ILM segmentation inherently includes some smoothing, which overlooks small fluctuations on the surface. For this image, the surface is iteratively relaxed by taking the vertical derivative and moving the surface at each pixel until a maximum derivative in the axial direction is found, with no smoothness constraints. The difference between the relaxed surface vertical location and the original smoothed surface vertical location is shown.

(b) The maximum difference between the relaxed and smoothed surfaces from (a) are used to create a range around the smoothed surface, and the average SD-OCT intensity in that range is shown.

(c) The smoothed surface is again relaxed but instead of finding the maximum derivative, it instead looks for the location for each pixel that results in a near-constant SD-OCT value throughout the surface. The difference between the smoothed and relaxed surface is shown as in (a).

(d) The maximum difference between the relaxed surface from (c) and smoothed surface are again used to create a range that is averaged as in (b).

(e) An average between 7 pixels of the ILM.

(f) Uses the surface of constant value (as found in (c)) to compute the average within 7 pixels.

(g) The same as in (f) but taking the median instead of the average.

(h) The SD-OCT image anterior to the ILM is set to zero and a summed cumulatively in the anterior-posterior direction. The surface of constant value is found within the cumulative sum image. The difference between the surface of constant value and the smoothed surface is shown.

(i) The relaxed surface is found after a local normalization and the average vector normal to the surface is found. A dot product of the normal at each location on the surface and the average is shown.

(j) The same as (i) except the SD-OCT is laterally upsampled by a factor of 2 using bicubic interpolation first.

(k) The same as in (j) but upsampled by 2 using bilinear interpolation.

(l) The same as (i) but using the dot product with a local normal instead of the normal for the whole image, and upsampled by 2.

The method in Fig. 4.11(e) was the one selected for this work. Again, one can only really base this decision on a qualitative assessment, and while Fig. 4.11 (g) and (h) both seem plausible alternatives, the folds in Fig. 4.11(e) seem to be the clearest. It could be noted, however, that the methods used to produce images Fig. 4.11(i,j,k) seem like a better way to visualize folds in the choroid, as choroidal folds tend to be of a wavelength that is too large for any practical surface smoothing to neglect its undulations.

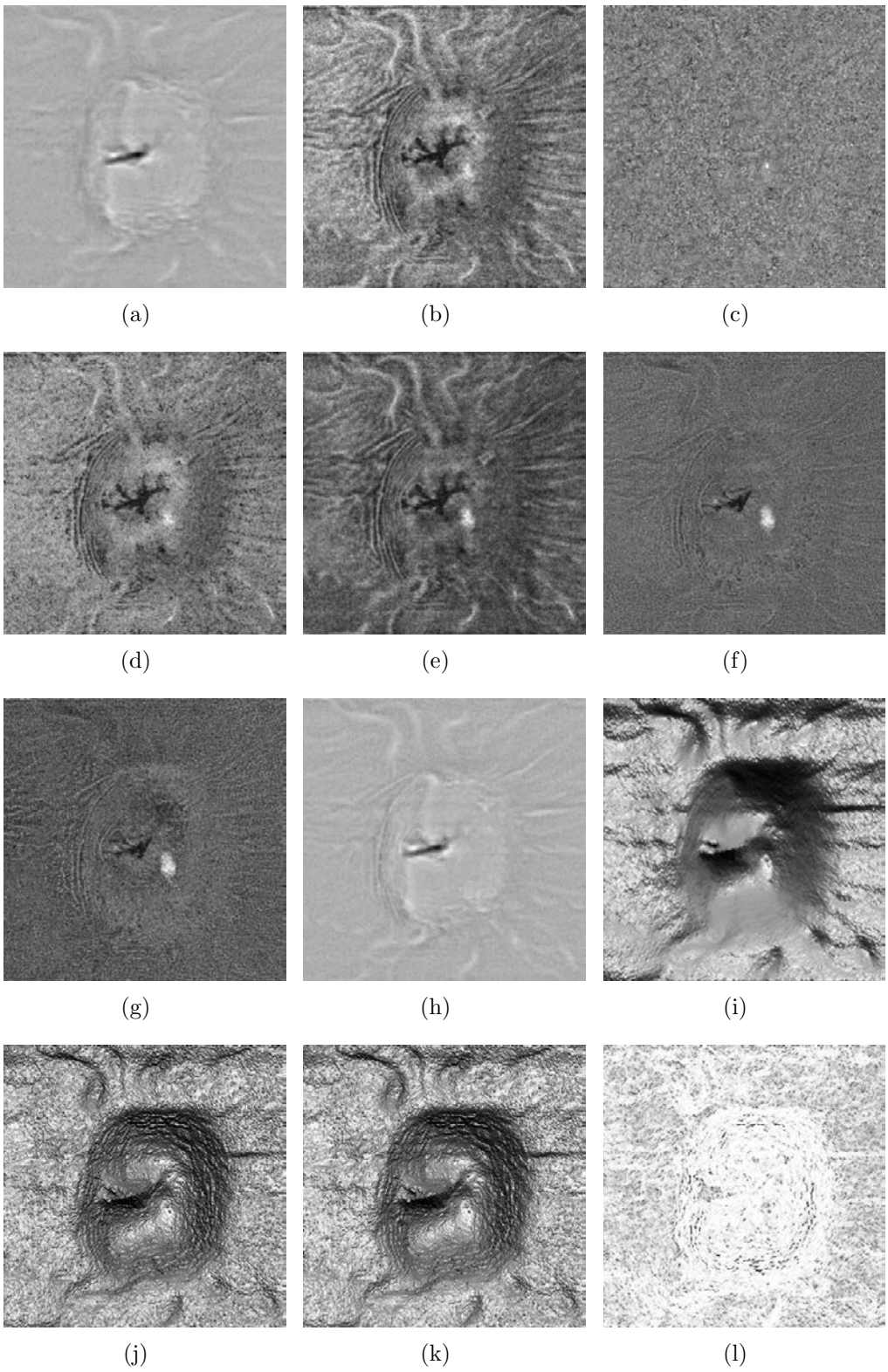


Figure 4.11: The resulting images from multiple attempts to enhance folds in SD-OCT images. The method resulting in figure (e) was the one ultimately chosen for the work in this thesis.

CHAPTER 5
DISTINGUISHING BETWEEN CAUSES OF OPTIC DISC EDEMA (AIM 2)

5.1 Introduction

In this chapter, an algorithm is developed that distinguishes between papilledema and other causes of optic disc edema. There are a number of methods that can be used to distinguish papilledema from other causes [33, 34], including invasive techniques such as lumbar punctures [35], and imaging techniques such as orbital ultrasonography [36] and fluorescein angiography [37], but ideally one would like to extract as much information as possible from a single imaging modality. The data acquired by The University of Iowa Hospitals and Clinics has 88 usable fundus/SD-OCT image pairs taken on the same day with multiple causes of optic disc edema, and since fundus images are the simplest of the two image types to obtain, it would be ideal if diagnostics could eventually be performed entirely with fundus images. While such a project may be a bit ambitious for one thesis, it is useful to work towards that ideal, and that is, in part, what is done here.

5.2 Features for Distinguishing Cause of Optic Disc Edema

Both SD-OCT and fundus features are used, such as RNFL thickness, BM shape, and fold-related features from Aim 1. The fundus features can be found in the next chapter. A summary feature list can be seen in Table 5.1. The intent is to identify features that are important to assessing both the cause and severity of optic disc edema.

5.2.1 Fundus Features

Much of this will be covered more thoroughly in the next chapter, but as the contents of the next chapter are exclusively about assessing the severity of papilledema, it seemed worthwhile first to show the algorithm that determines which patients actually

#	Image Source	Name
1	Fundus	Swollen ONH area
2	Fundus	Vessel Discontinuity Index
3	Fundus	MUCH Parameters
4	Fundus	Disc Margin Obscuration Ratios
5	Fundus	Disc Margin Obscuration Parameters
6	Fundus	Small Vessel Obscuration
7	Fundus	Vessel Border Obscuration
8	Fundus	Artery-Vein Ratio
9	SD-OCT	Fold Parameters
10	SD-OCT	TRV
11	SD-OCT	BMO Shape
12	SD-OCT	RNFL Thickness

Table 5.1: A summary list of features used for distinctions between causes of optic disc edema. Note that there are up to 12 variants of each of these, which consists of ratios, nonlinear transformations, and absolute differences. The full list can be seen in Fig. 5.6.

have it. Many of the fundus features covered are explicitly designed for determining severity. However, these features are expected to change in a way that is relatively well correlated to a change in severity of papilledema. Thus, it stands to reason that if it is not true papilledema then the correlation will have some discongruity. For example, if it is found that the area of the optic disc swelling (section 6.3.2) would indicate a high severity but the vessel tortuosity (section 6.3.1) suggests little or no severity, then this increases the likelihood that it is not true papilledema.

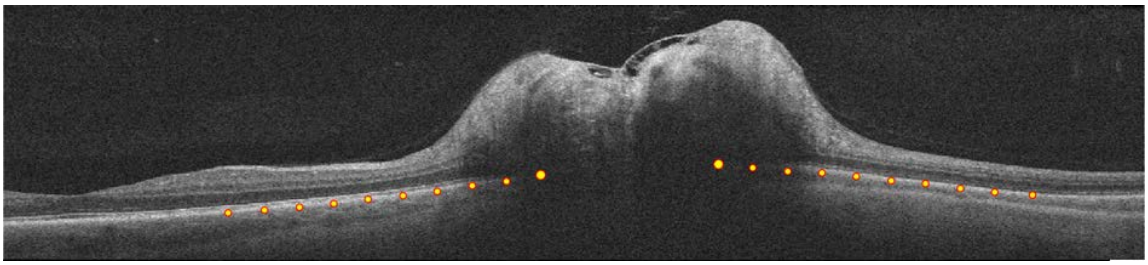
5.2.2 BM Shape

A relatively recent method for giving some insight as to the cause of optic disc swelling consists of an analysis based on Bruch's membrane (BM) [2, 9]. The idea is that when swelling is due to raised intracranial pressure, the increased pressure is relatively contained within the nervous system, and is the direct cause of the swelling visible at the optic disc. Since the optic nerve, which is posterior to the

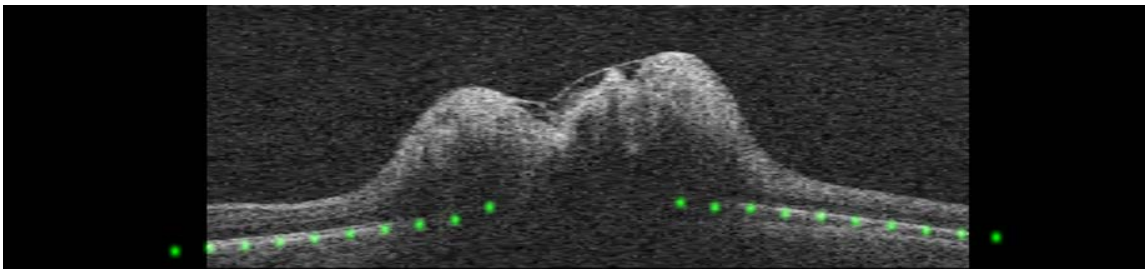
eye, is effectively pushing forward to cause the swelling, the back of the eye bends forward with it. Shape analysis of Bruch's membrane is performed using eigenvector decomposition and Procrustes analysis. Prior work on this topic (section 3.3) uses 5-line raster scans, which are a modified version of normal SD-OCT images, except with much higher resolution at the cost of less area covered. The 5-line raster scans represent 9mm of physical space in the lateral direction, whereas volumetric scans only represent 6mm. The algorithm is designed to assume a 2.5mm distance from the BMO marked and the final landmark placed on the Bruch's membrane layer. For the 5-line raster scans, as long as they are reasonably centered, an extra 2.5mm of space to work with on each side is a relatively safe assumption, but in volumetric scans it is markedly uncommon. While it may be true that the algorithm could be redesigned to consider a smaller area, for comparison, it is ideal that the shape parameter computations for the 5-line raster and volumetric scans to be as similar as possible, thus making a 2.5mm landmark length essential. To that end, the boundary is extrapolated by fitting 50 pixels on each side to a quadratic equation (Fig. 5.1). The conversion from raster scans to volumetric was reported at ARVO in 2016 [38], with a correlation of 0.967 (Fig. 5.2). Each patient was left out and eigenvectors computed for all patients remaining. The overall eigenvectors can be seen in Fig. 5.4 and all shape points can be seen in Fig. 5.3. The shape parameters for each left-out patient is computed for each image in the dataset acquired from the University of Iowa and used as a feature. Unlike the work in [2, 9], for this data it appeared that the first two eigenvectors showed some upward bending. As such, both were used as features, in addition to a simple addition of the two eigenvalues to accentuate upward bending.

5.2.3 Folds and Wrinkles

In the case of pseudopapilledema, there is no actual swelling. It is just the way a particular patient's optic disc looks naturally. As such, there is no mechanical stress,



(a)



(b)

Figure 5.1: A high-definition raster scan (a) and the central slice of a volumetric scan (b), with landmarks shown. Note the difference in image quality and the reduced size in the lateral direction necessitating an extrapolation of the bottom layer on the ends.

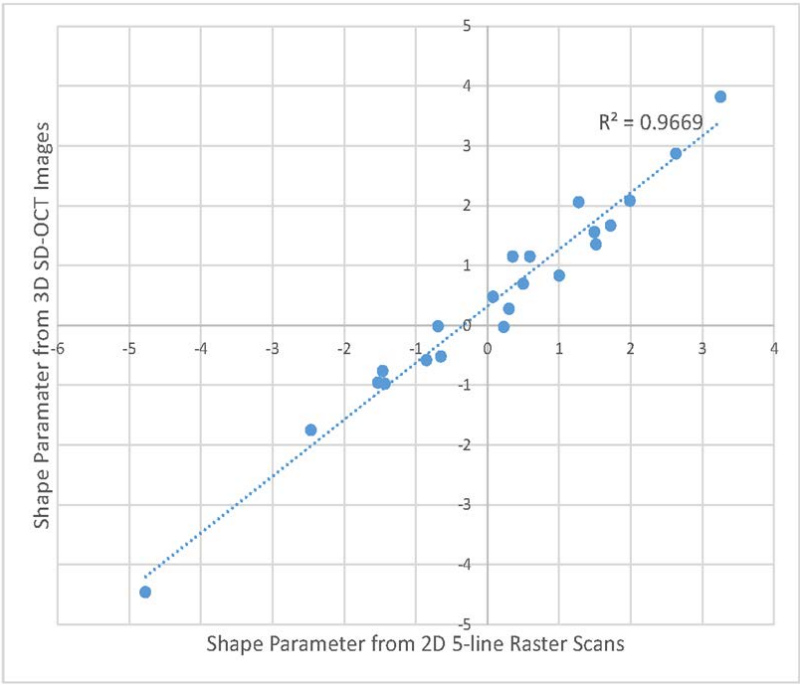
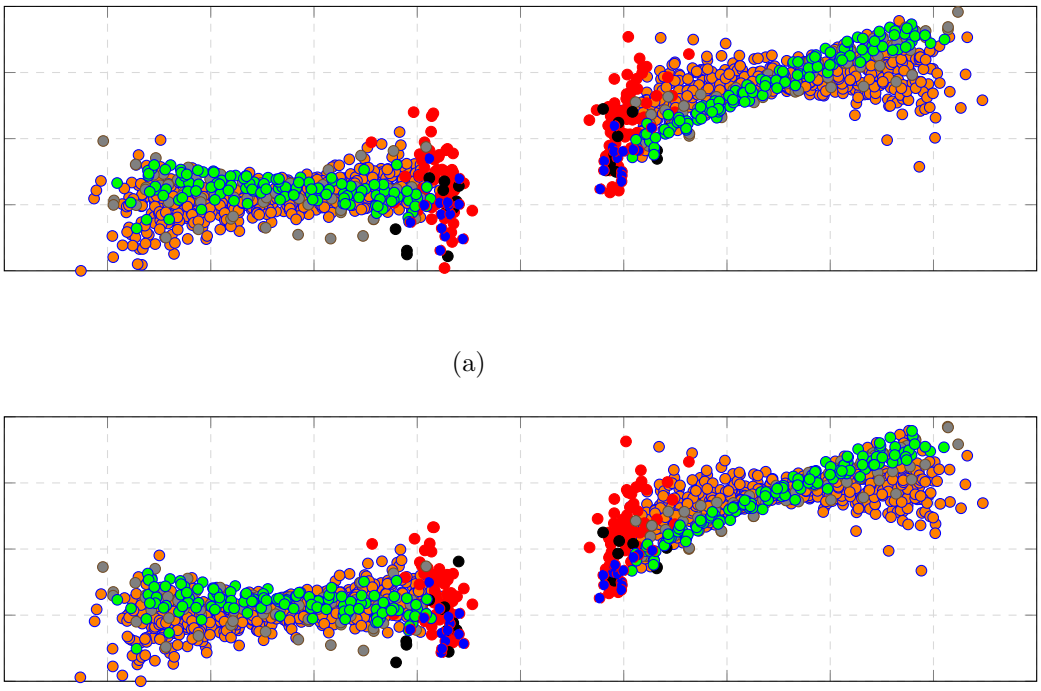


Figure 5.2: Correlation between shape parameter computed from raster scans and shape parameter computed from volumetric SD-OCT images.



(a)

(b)

Figure 5.3: Results of shape analysis in the University of Iowa dataset. (a) is all the plotted shapes, as they are, whereas (b) is the same shapes reproduced using only the first three eigenvectors. Green and blue are NAION, pink and red are papilledema, and gray and black are cases of optic disc edema due to other causes.

and we should not expect to see folds at all. For true papilledema, as mentioned in the previous section, the pressure originates posterior to the eye, pressing it forward. Just as pressing into a large rubber sheet will produce wrinkles that are aligned radially from the point of pressure, so too are we more likely to see this in true papilledema. For AION, which means anterior ION, the source of pressure is not outside the eye. This suggests a pushing from within the optic disc, which is likely to result in concentric folds, if any. From these observations, it seems prudent to extract the fold direction and detection as a feature.

The complete list of fold features can be seen in Table 5.2. Using an automatically

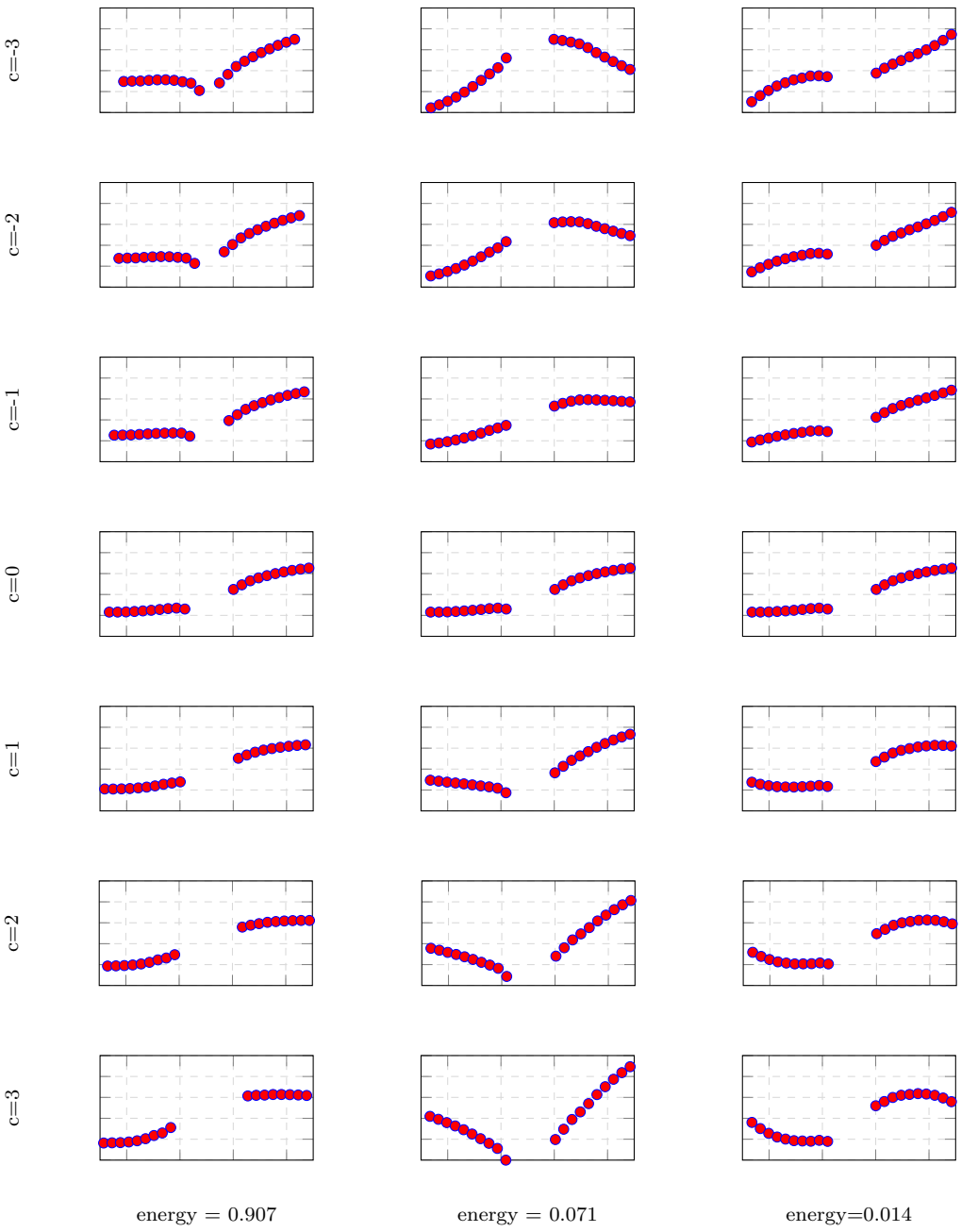


Figure 5.4: Eigenvectors of this dataset shown by increasing eigenvalues, c.

approximated center for the optic disc, a direction map is produced to determine which orientations should be labeled at either concentric or radial, relative to the optic disc center. Combined with the probability maps from the previous chapter, it

#	Feature Description
1	Max fold probability after 11×11 local minimum filter
2	Cosine of orientation angle at location corresponding to (1)
3	Cosine of max orientation angle of all pixels above threshold
4	Count of pixels in probability map above threshold
5	Count of pixels in largest object after thresholding
6	Cosine of max orientation angle of grouped objects using k-means clustering

Table 5.2: List of all fold features obtained from fold probability maps as created using the methods from chapter 4. Note that all angles listed (2, 3, and 6) use the cosine of the difference between the orientation angle at that location and the angle of radial direction with respect to the ONH center at that location. This causes values to range between 0 and 1 and a value of 0.15 is used when no folds are found, as that value is on the cusp between what could be defined as a radial fold or a concentric fold.

is a straightforward task to threshold and determine the presence and direction of folds in a given SD-OCT image. The direction of the folds is measured as it aligns with the radial direction. Thus, it serves to only extract the minimum value of fold alignment direction from an entire fold-enhanced image, as any radial fold suggests a higher probability of true papilledema. Whereas if a minimum fold alignment direction is still somewhat high, this suggests a concentric fold, which is only an indication that it is not pseudopapilledema.

Theoretically, in the event there are no folds, the random forest analysis should be able to handle that with a simple binary distinction, but it still raises the question of what value for fold direction should be used in the event there are no folds. Since there are to be two separate classifiers (Section 5.3), it makes sense that this answer may be different for each. The fold orientation is computed for each fold by taking the cosine of the difference between the orientation of a region and the radial direction with respect to the approximate ONH center. This results in high values for radial folds and low values for concentric folds. For the papilledema classifier, it is expected that if folds are present, a high (radial) value should give a high degree of certainty

for a true outcome, so a low value if no folds are found makes the best sense. For the NAION classifier, it is expected that low (concentric) values will give a more likely positive result, thus perhaps a high value is best. However, while all this may be a reasonable approach, there is something to be said for the simplicity in picking a value that is consistent for each. Perhaps the best idea, then, is to pick a value that is on the cusp between what would be identified as radial and what would be identified as concentric. Since the value of the orientation is determined as the cosine of the difference between the radial direction and the direction of a given fold, this would place the cusp at about 0.15, which is the value used in this work when no folds are found.

Using the optimal threshold from Fig. 4.9(c), fold probability maps were thresholded for each patient. A local minimum was found in an 11×11 window across the probability map, and the max for the filtered image used as a feature. Additionally, the orientation angle corresponding to the location of the maximum probability map value was used, in addition to the maximum orientation angle of all thresholded objects. The total number of true pixels in the thresholded image and the total number of pixels in the largest object were both used as features as well.

The final fold feature was found by applying a k-means clustering algorithm to the objects in the thresholded image. The dimensions in the k-means clustering are the orientation angle and x and y position of the centroid of each object. The cosine of the orientation angle is used to force it to vary between 0 (completely concentric) and 1 (completely radial), and the position coordinates are scaled to also range between -1 and 1, with respect to the approximated center of the opening to Bruch's membrane. The k of the k-means clustering algorithm is increased until the Euclidean distance between any point in a group and the centroid does not exceed 1/8. Once all of the k-means clusters meet this condition, any cluster with less than three members was eliminated (Fig. 5.5). The maximum average orientation angle of the remaining

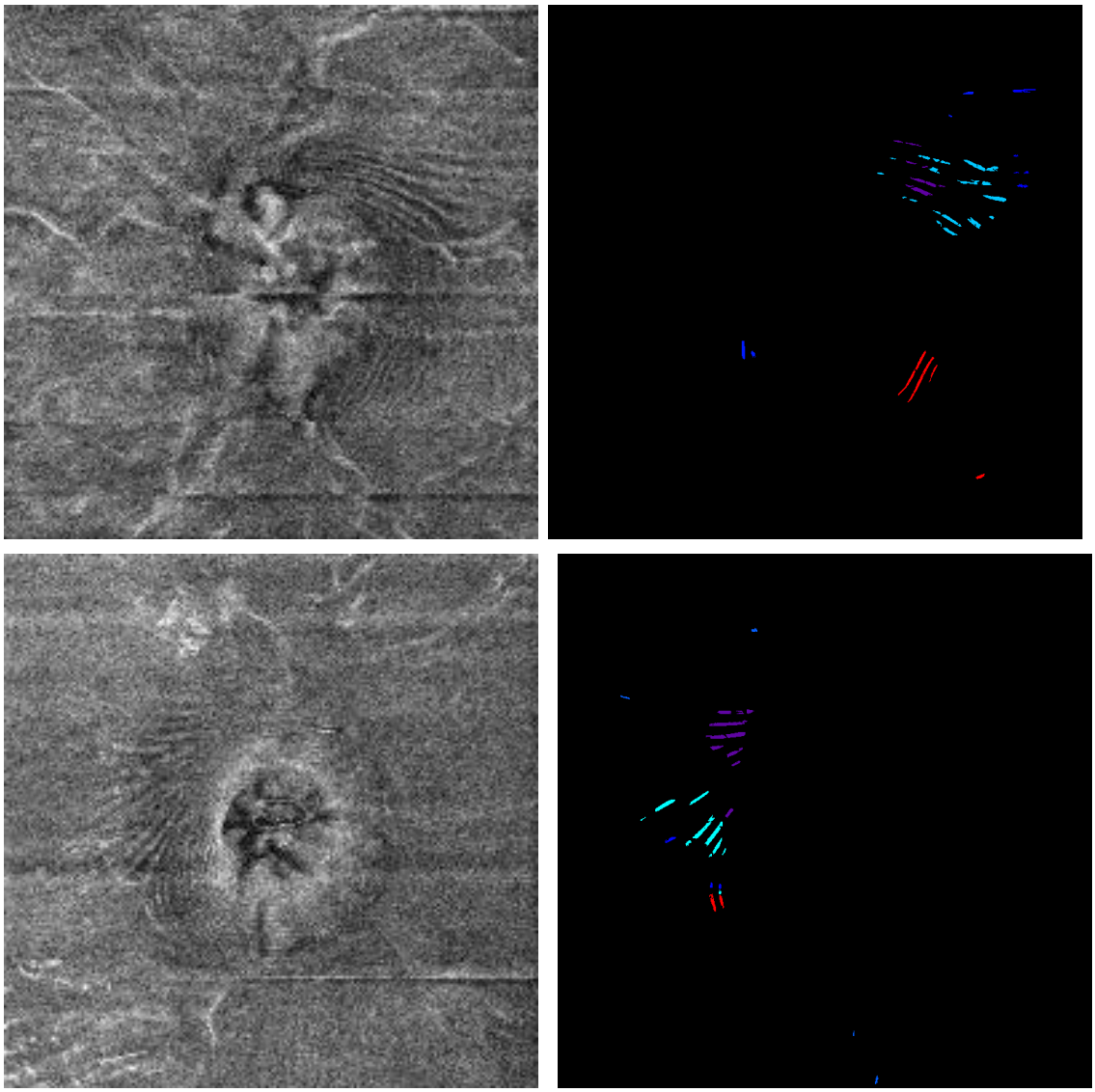
clusters is used as a feature. The maximum is taken to identify when folds are entirely concentric.

5.3 Classification

Classification is performed using a leave-one-patient-out random forest regression analysis with a positive condition set to 1 and a negative condition set to 0. The two main classifiers were performed on cases of papilledema volume matched to optic disc edema due to other causes, and cases of NAION volume-matched to optic disc edema due to other causes. For the papilledema classifier, there was an average papilledema volume of 13.9 mm^3 with a standard deviation of 3.22, while the average volume for other patients was 14.94 mm^3 with a standard deviation of 3.74. The patients were volume matched by use of a greedy algorithm that selected image pairs with the smallest difference in volumes until everything under a 4.5 mm^3 difference was matched. A total of 25 patient pairs were extracted from this process. This process was repeated for NAION patients, resulting in 14 pairs, with an average volume of NAION images of 13.56 mm^3 and standard deviation of 3.27, while other patients in this classifier had an average volume of 14.79 mm^3 with a standard deviation of 3.57. Other supplementary classifiers were also used, without volume matching and using only SD-OCT or fundus features.

5.4 Results

Of a total of 88 patients, 63 were diagnosed with true papilledema, and 14 were diagnosed with NAION. At the time of writing this, there were no viable images with pseudopapilledema, but 11 had optic disc edema due to other causes. When a classifier was trained and tested on papilledema patients volume matched to other causes of optic disc edema, using a leave-one-patient-out random regression analysis, an AUC was achieved of 0.89. When the same method was applied to volume matched NAION patients, an AUC of 0.847 was achieved. Unfortunately, as mentioned there was an



(a) (b)

Figure 5.5: Examples of fold-enhanced images (a) with the resulting k-means grouping (b).

insufficient number of pseudopapilledema patients for classification (Fig. 5.7). Additionally, a classifier was trained for both papilledema and NAION, without volume matching, for comparison purposes. An AUC of 0.861 for the papilledema classifier and 0.863 for the NAION classifier was achieved for these. This whole process was also repeated using only fundus features (Fig. 5.8).

Using leave-one-patient-out regression meant that the random forest importance could be computed with each iteration and summed across all iterations. The cumulative importance can be found in Fig. 5.6.The most important feature for the papilledema classifier was a fundus feature, which was the nasal/temporal border gradient ratio (item 13 in Table 6.1). Some other important features include the shape parameter (SD-OCT), the artery-to-vein ratio (Fundus), and max vessel width (Fundus). The remaining features were all relatively equal in importance, except for two that seemed to be largely unimportant. In contrast to the most important feature, the least important features were the nasal/temporal distance-to-ellipse fit ratios at 30 degree intervals (items 10 and 11 in Table 6.1). The lateral nasal/temporal distance-to-ellipse fit ratio (item 9), however, was of comparable importance to the rest.

For the NAION classifier, the most important feature was the shape parameter, followed closely by the nasal/temporal vessel border gradient ratio (item 3 in Table 6.1), max vessel width, and fold orientation. The remaining features were, again, all comparable in importance, except for the low importance features which were the 30 degree nasal/temporal distance-to-ellipse fit ratios and the count of fold pixels.

5.5 Discussion

Looking at Fig. 5.8, there seems to be something of an anomaly, in that the volume-matched regression analysis from using fundus features only actually performed better than the non-volume-matched. This is peculiar news. Granted, this was only an exercise in curiosity, and the classification based on fundus features alone

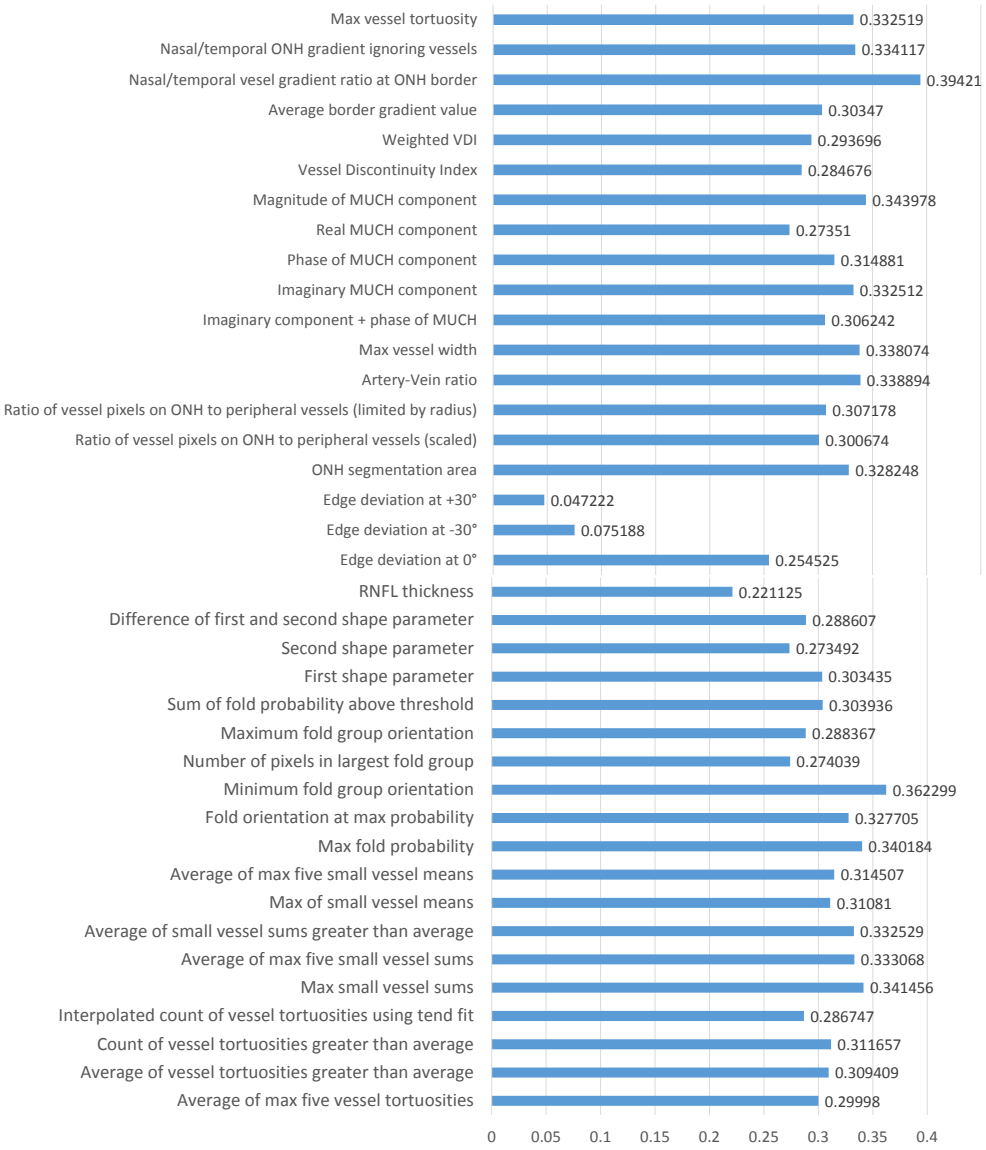


Figure 5.6: Cumulative importance of each feature, summed as each patient is left out and the regression analysis trained.

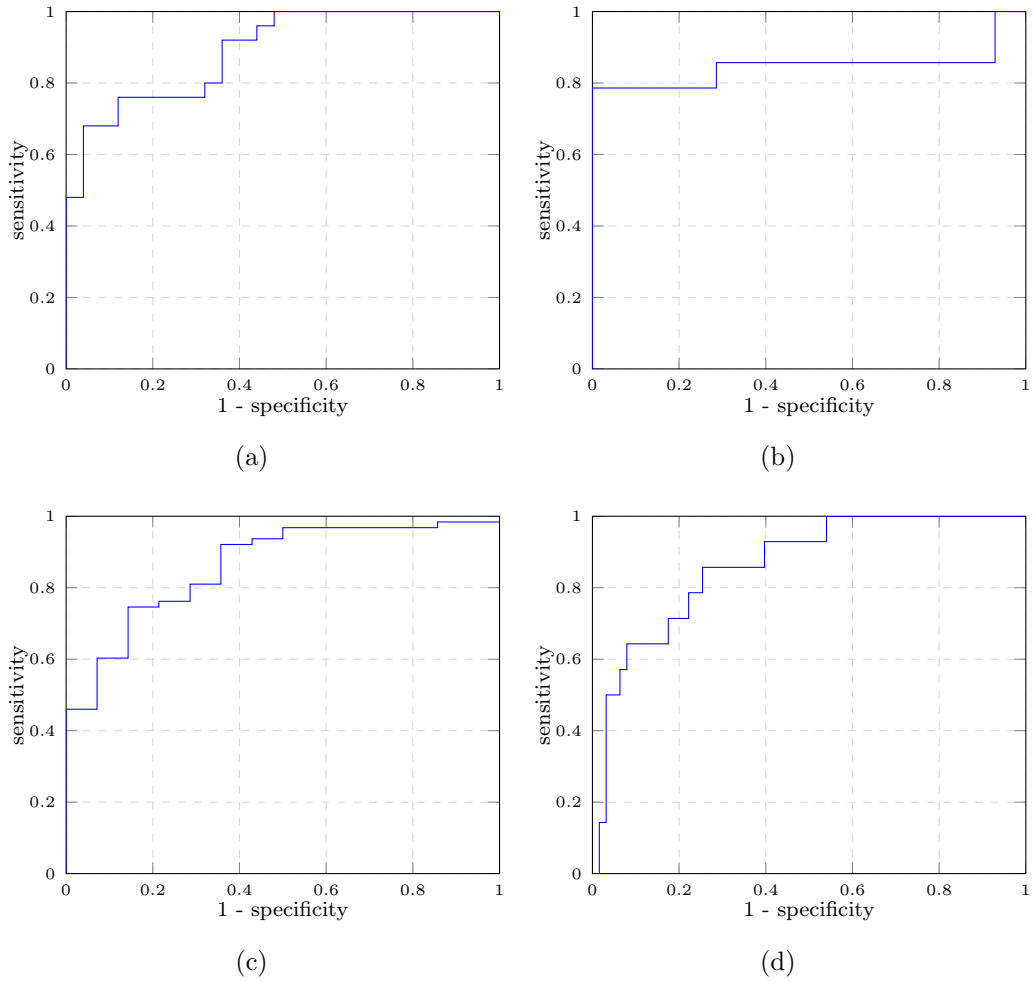


Figure 5.7: ROC charts from all features for (a) The volume-matched classifier between papilledema and all other causes of optic disc edema (b) the volume-matched classifier between NAION and other causes (c) papilledema classifier without volume matching, and (d) AION classifier without volume matching. Resulting AUCs were 0.89, 0.847, 0.861, and 0.863, respectively.

was not expected to perform well, but the fact that the results seem reasonable is unexpected. As to why this would not perform as well when the volumes are not matched invites much investigation, and although predicting cause from fundus features alone was not a goal in this doctoral work, I cannot help but speculate. Of the most important features in this fundus-only classification of cause was the nasal-

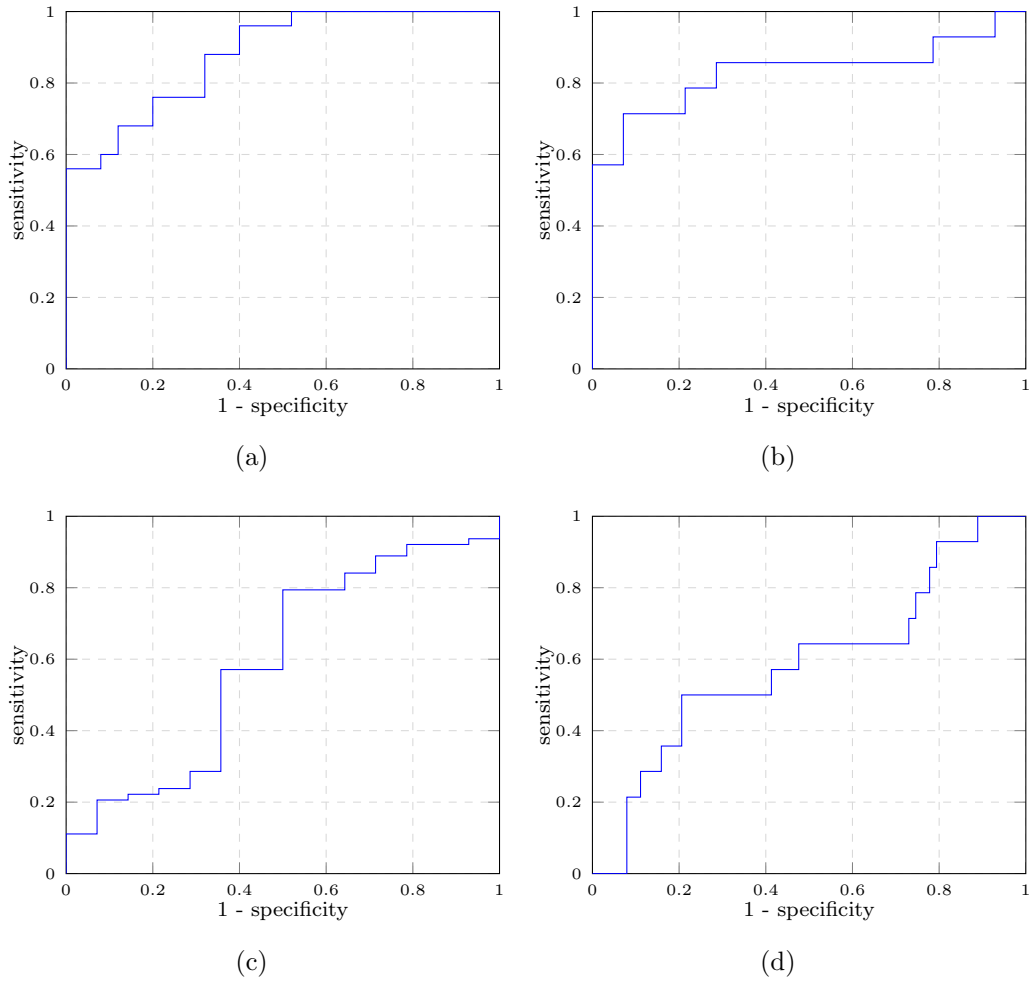


Figure 5.8: ROC charts from fundus features only for (a) The volume-matched classifier between papilledema and all other causes of optic disc edema (b) the volume-matched classifier between NAION and other causes (c) papilledema classifier without volume matching, and (d) AION classifier without volume matching., with AUCs of 0.88, 0.832, 0.593, and 0.59, respectively.

temporal gradient ratio. This makes sense as, because these images were volume matched, some of the high-severity images were left out. On the other hand, papilledema is partly identified through a 'c-shaped halo' of obscuration around the optic disc at the lower severity levels, whereas other causes of optic disc edema are not. Thus, we expect to see a different nasal/temporal gradient ratio for papilledema

cases than for other causes, at this severity level, but not at all severity levels.

In any event, the results in Fig. 5.7 are the main ones of interest in this doctoral work. Fig. 5.7(a) has an AUC of 0.89. While this does not qualify as a bad result, it leaves room for improvement, as is the case with the NAION classifier with an AUC of 0.863. It should be noted that the AUC of the NAION volume-matched cases using only SD-OCT features (Fig. 5.9) was 0.668, which seems rather low compared to the use of the full feature list and fundus-only features. It would appear odd that the full feature list should perform well at all, considering the fact that there are more features than data sets. This is likely due to the robustness of a random forest classifier in the face of redundant features. In an attempt to confirm this, the SD-OCT and fundus features were separately compressed by training to an intermediate number for each, and that number was used in place of the SD-OCT or fundus features in the classifier (Fig. 5.10). Indeed it doesn't seem to be the case that the number of features had a negative impact on the results, as the compressed feature sets performed poorly.

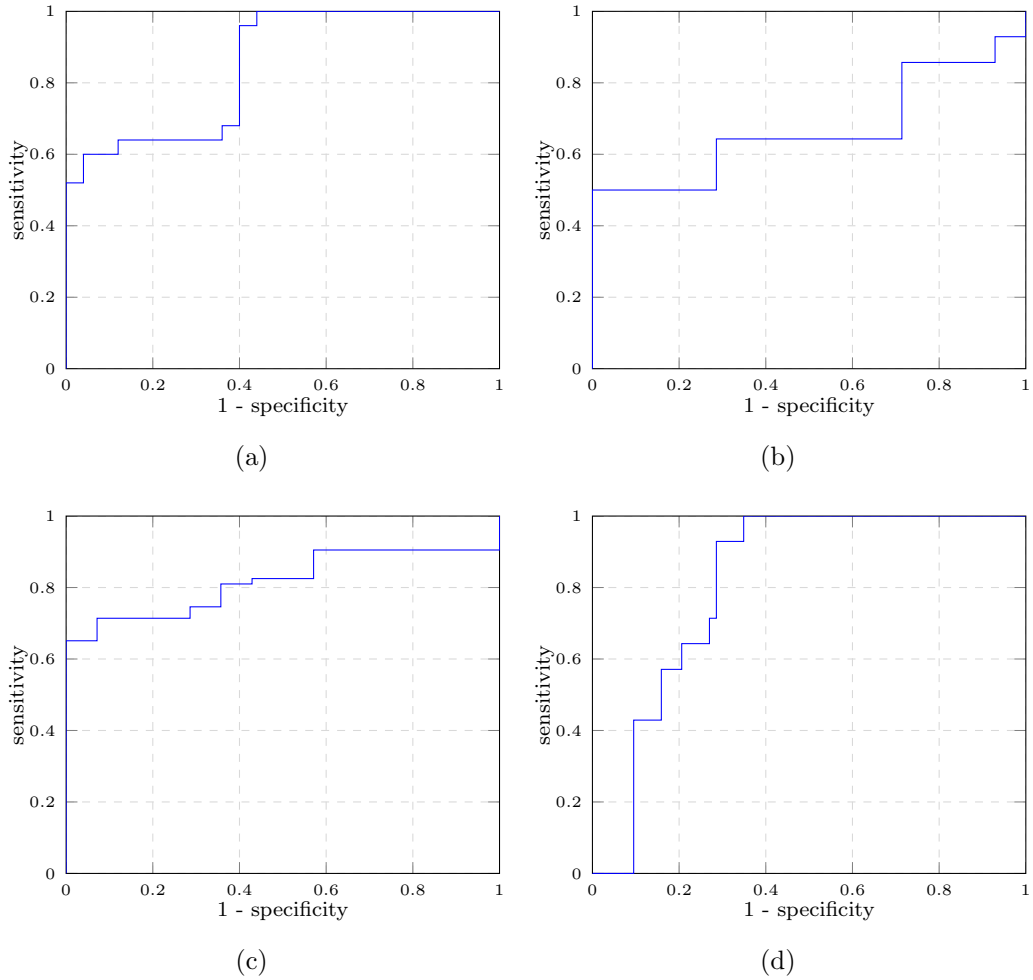


Figure 5.9: ROC charts from SD-OCT features only for (a) The volume-matched classifier between papilledema and all other causes of optic disc edema (b) the volume-matched classifier between NAION and other causes (c) papilledema classifier without volume matching, and (d) AION classifier without volume matching, with AUCs of 0.848, 0.668, 0.816, and 0.816, respectively.

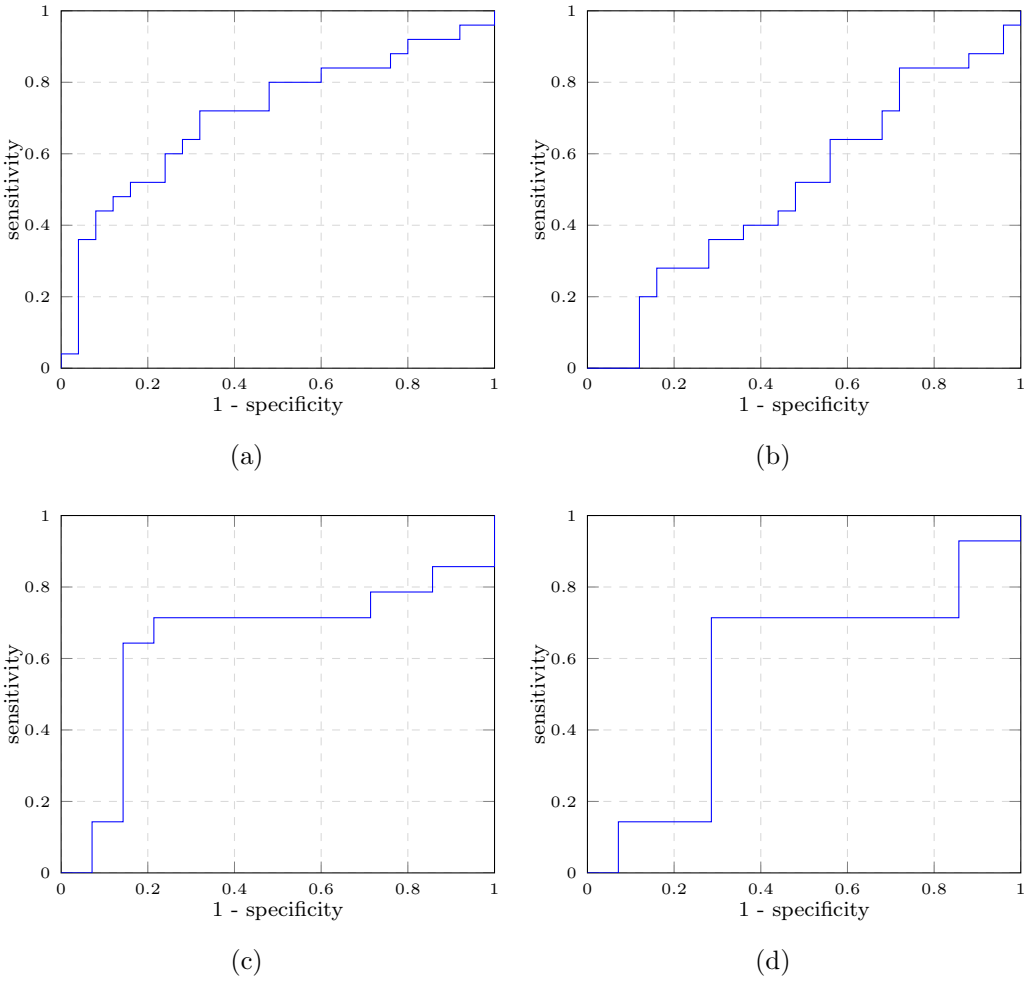


Figure 5.10: Using the volume-matched results from the fundus-only and SD-OCT-only features as features for each other. (a) Papilledema volume matched with fundus-only features used in a classifier, and the results of that classifier added to SD-OCT features and a classifier ran again. (b) The same thing except the first classifier was the SD-OCT-only features, and these results were added to fundus-only features. (c) The same as (a) but for NAION. (d) The same as (b) but for NAION. AUCs were 0.757, 0.566, 0.612, and 0.714, respectively.

CHAPTER 6
QUANTIFICATION OF OPTIC DISC EDEMA SEVERITY IN CASES OF PAPILLEDEMA (AIM 3)

6.1 Introduction

In the event optic disc edema is due to raised intracranial pressure (papilledema), the swelling is often measured by the Frisén scale [3], which is a 0-5 ordinal rating of severity (with 0 being normal). This rating requires clinical expertise and is subject to the interpretation of the measurer, making it prone to inconsistencies. With the introduction of spectral-domain optical coherence tomography (SD-OCT), an automated segmentation of the layers of the ONH has been developed, and the total retinal volume (Fig. 3.2) has been computed [28] in cases of papilledema, with a strong correlation found between the volume of the ONH and expert-defined Frisén scale grades. Given its continuous and repeatable nature, it is assumed that the SD-OCT-based ONH volume serves as a better representation for the severity of optic nerve edema, in papilledema cases. However, computing this volume requires one to capture the ONH with SD-OCT imagery, which requires the use of expensive equipment that is not commonly available in either an emergency room or telemedical setting. Having the ability to assess the severity of optic nerve edema in such settings is highly desirable, as many causes of optic nerve edema can indicate serious underlying conditions that need to be diagnosed early.

On the other hand, Echegaray et al. [5] proposed a classification approach to predict the Frisén-scale grades from color fundus features. In this thesis, given the limitations associated with Frisén-scale grades, instead of using Frisén-scale grades as the reference standard it is assumed SD-OCT-based 3D volumetric calculations serve as a more reliable reference standard for measuring the degree of optic nerve edema. In this chapter, a regression-based approach for predicting the volume computed from a 3D SD-OCT segmentation of the ONH from features extracted from 2D color fundus photographs is developed, thereby obtaining a continuous assessment of the severity

of optic nerve edema from fundus photographs alone.

6.2 Preprocessing Steps Prior to Feature Extraction

Some preprocessing steps were implemented prior to feature extraction. The first was determining the magnification, which was important as a first step as other preprocessing steps rely on it. The next step was the vessel segmentation, as discussed in section 3.4. The vessel segmentation was used twice for two different scales - one for the whole image and one restricted to only the ONH area. The vessel segmentation [39] used was robust enough that the magnification was not needed for the whole image, however for the ONH-vessel segmentation, this was required for proper scaling. The last preprocessing step was the generation of the cost image used to segment the ONH or the ONH swelling. For healthy eyes, segmentation is usually straightforward, but as the primary interest for this thesis is in images with optic disc edema, it is expected that there will be images that do not have a clear ONH boundary. Thus, there is a heavy reliance on the magnification for initial segmentations to be later refined, which is accounted for in the cost image generation.

6.2.1 Magnification

As the intent is to approximate a physical volume, the issue of scale is significant. However, among the datasets available, the individual magnifications and focal lengths of the fundus images were not recorded. To compensate for this, the magnification was approximated by manually registering the fundus image to the *en face* of the SD-OCT image (Fig. 6.1), using a three-point registration. While this is not an absolute magnification, it serves as a relative metric. This was computed by finding the pixel distance between each registration point, and generating ratios of fundus distances to SD-OCT distances. The median distance ratio was used as the magnification. The median was used instead of the mean because artifacts in the SD-OCT

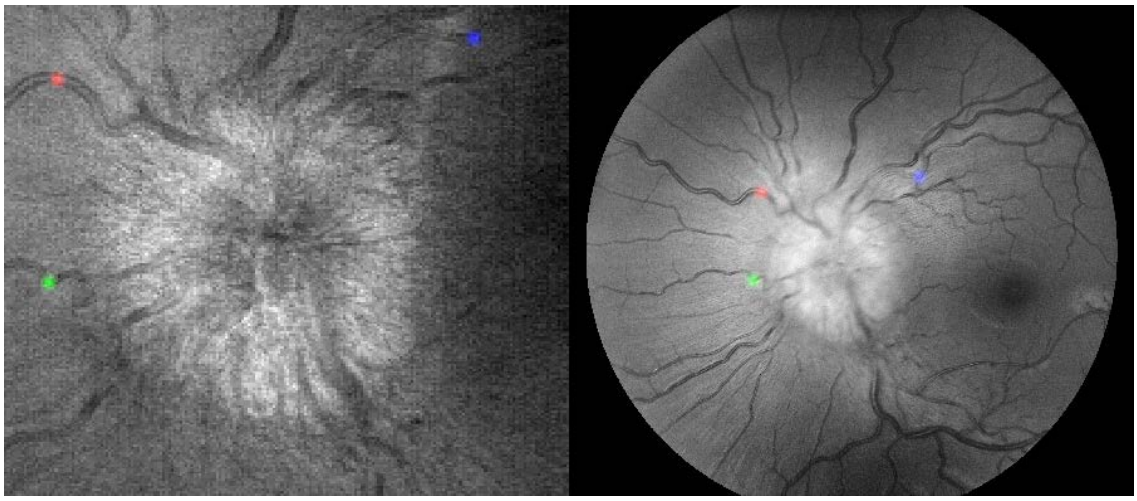
image provide for the possibility of outliers. Each resulting median ratio was normalized by dividing by the maximum of the entire data set (for example, the dataset acquired by the University of Iowa Hospitals and Clinics had a maximum scaling factor of 0.9376). This was done in an attempt to alleviate a possible degree of variation in other datasets affecting the algorithm unexpectedly; i.e. if the true magnifications and focal lengths were recorded, the calculated physical scale for each image could be similarly normalized, resulting in the same $(0, 1]$ range - despite the fact that the scaling factor is relative, this allows it to be treated as absolute.

6.2.2 Cost Function Generation for Segmenting the ONH Swelling Area

In swollen cases, the ONH boundary is often obscured, which complicates a segmentation of the ONH (in both 2D and 3D images) - which is partial motivation for segmenting the swollen area, instead of attempting to segment the disc, itself. Segmenting the ONH swelling area from the peripheral area is among the more important tasks in this thesis. This is done by unwrapping the image about a manually approximated center, generating a cost function, and using a graph-based algorithm [40] to find a single, continuous border.

The cost function is generated by first performing a vessel segmentation and using second-order interpolation to inpaint the vessels. This minimizes the effect the hard boundary of the vessels has on the segmentation. However, especially on the ONH, vessels can sometimes overlap and appear too large or unusual of shape to be marked as a vessel by a machine, preventing them from being detected by a vessel detection algorithm (Fig. 6.2). To circumvent this, a mask is applied near the center of the ONH. The mask is made using a simple 4-dimensional k-means segmentation. Three of the dimensions consisted of the RGB values, and the fourth was the radius from the approximated center. The inpainted image is segmented into three regions ($k = 3$), which was assumed to be the ONH, the peripheral area, and the background

(a)



(b)

Figure 6.1: The maximum (a) and minimum (b) magnified fundus images acquired by the University of Iowa Hospitals and Clinics. Shown is each *en face* of the SD-OCT image alongside the greyscale of the fundus image, and the three manually marked registration points shown as red, green, and blue dots. These image pairs correspond to a relative scaling value of 1 and 0.603, respectively.

(K_O, K_P, K_B).

However, the nature of the problem is that the ONH and the peripheral area are not easy to distinguish in cases of optic disc edema - not even manually. As such, a distinction is introduced by modifying the radius dimension. That is, instead of a linear radius, a Gaussian radius was used, with standard deviation scaled by magnification. This allows for the simple control over an expected optic disc radius by adjusting the Gaussian variance, and further allows uncertainty in the approximated center position to be considered by adjusting the Gaussian magnitude. More generally, if d is the vector of features used in the k-means classification, then $d_{1,2,3}$ are the RGB values, and

$$d_4 = \sqrt{t}e^{-r^2/(100^2 tM)}, \tag{6.1}$$

where M is the magnification and r is the radius in pixels. It is important to remember that the point of this k-means segmentation is not to come up with a final segmentation, but to error on the side of under-approximation of the ONH boundary - to mask out a region that can be said to be part of the ONH with high certainty, such that vessel anomalies on the ONH don't disrupt the segmentation. This is why a value of 100 pixels was chosen for the standard deviation - because a radius of 100 pixels is smaller than any ONH radius, at the magnifications of the images in the University of Iowa dataset. The purpose of t is to create a smooth boundary by checking the boundary between the ONH and peripheral area and iteratively increasing t if there are large jumps (i.e. greater than 20 pixels) in boundary radius between 0.5 degree increments (Fig. 6.3).

Using the unwrapped inpainted image, a median filter is applied in an attempt to diminish the impact any remaining smaller vessels or other anomalies may have on our cost function generation. Then, a small Gaussian is applied and a vertical derivative taken. It is relatively well known that the ONH area tends to have less color overall than the peripheral area, and that it is generally brighter. As such,

it is expected that the standard deviation of the vertical derivatives, the maximum of the vertical derivatives, and the mean of vertical derivatives, across each RGB frame, are each useful indicators of the boundary between the swollen ONH and the peripheral area. To compute the cost function, consider how each value is expected to have an impact. The expectations for what the mean and maximum produce are straightforward, but the information that the standard deviation provides may be slightly less obvious. A low standard deviation is indicative of each color frame changing in intensity at a similar rate, which suggests there is not a shift in color. In such a case, the maximum tells us nothing that the mean does not tell us. On the other hand, a high standard deviation indicates that at least one color frame is changing notably more than the other(s). In this case, the maximum becomes something worth paying careful attention to, because that is where all the change is going to be found. In an attempt to combine these in the most useful way possible, the maximum is weighed by the standard deviation and summed with the mean. More generally, our cost function can be defined as

$$C(i,j) = (\bar{I'_f}(i,j) + \sigma_{I'_f(i,j)} \times \max_{f=\{r,g,b\}} I'_f(i,j)) \times M, \begin{cases} M = 1, & (i,j) \in K_P \\ M = 0.5, & (i,j) \in K_O \\ M = 0, & (i,j) \in K_B \end{cases} \quad (6.2)$$

where I' is the radial derivative image (that is, vertical derivative for the unwrapped case). The intent of summing with the mean is to establish a threshold or minimum value for each pixel, which becomes relevant when the standard deviation is near zero. Once this cost function is constructed, a graph search approach is used to segment the ONH swelling from the peripapillary area, while preserving circularity constraints.

(a) (b)

Figure 6.2: A fundus image before (a) and after (b) inpainting. Note some of the overlapping vessels on the inferior region of the ONH that are still visible after inpainting. This could cause errors in the segmentation.

6.3 Feature Extraction

In this section, each feature is discussed in detail. The complete feature list can be found in Table 6.1. Because there is sometimes only a slight variation in computation method between one feature and another, when each feature is first described, it is annotated with angle brackets$^{\langle 0 \rangle}$ to indicate its place in Table 6.1.

6.3.1 Vessel Features

The Vessel Discontinuity Index (VDI) is the count of discontinuities in the vessel segmentation$^{\langle 1 \rangle}$ and was found by Echegaray et al. to be of particular significance in classifying the severity of ONH swelling. This is due to the vessel detection algorithm failing to detect sections of the vessels that are obscured by a swollen ONH. The VDI was found by unwrapping the vessel segmentation about the approximate center of the ONH, using local entropy thresholding to create a binary image [22], computing the skeleton, shrinking it for 30 iterations to remove spurious branches, and counting

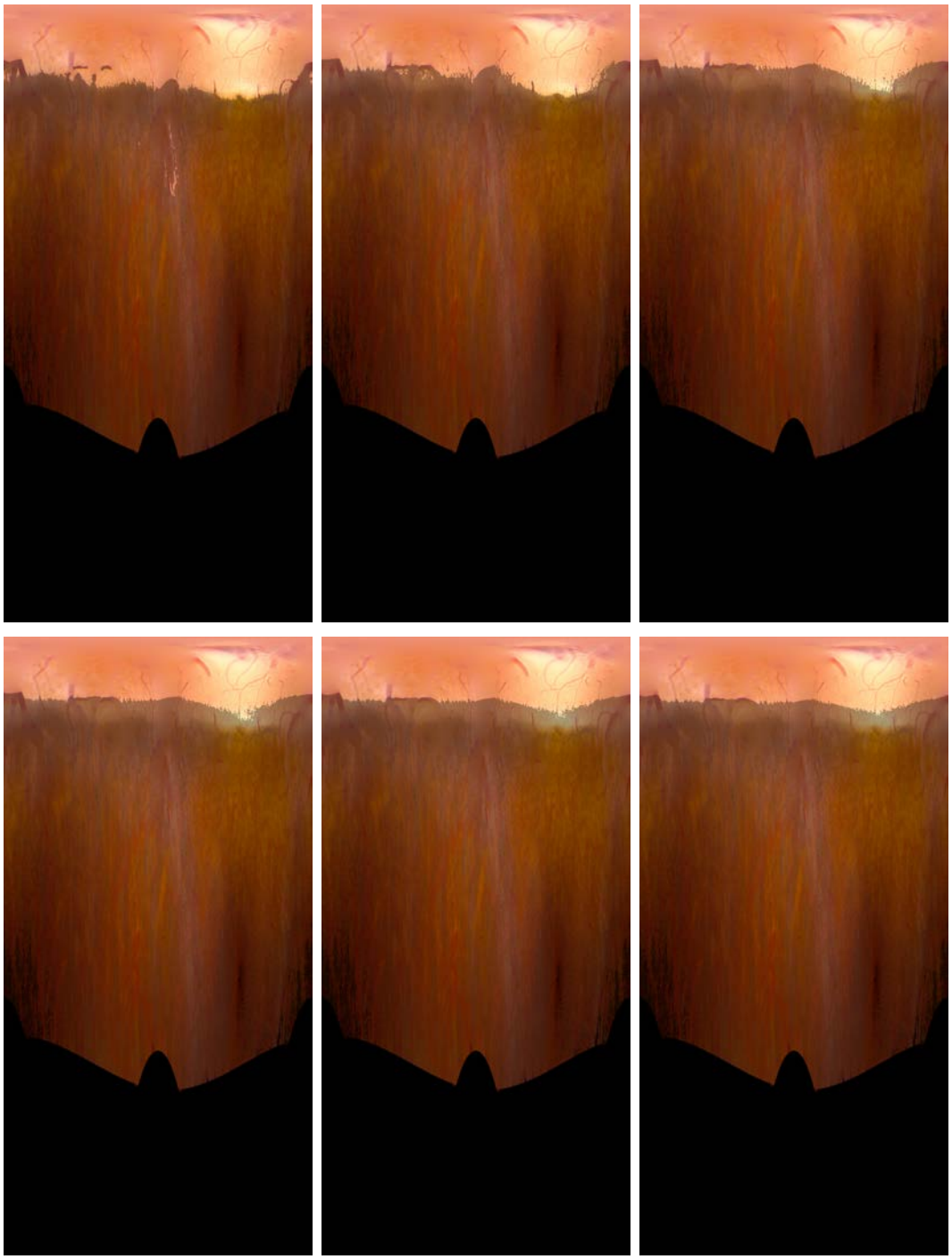


Figure 6.3: K-means results for increasing values of t. Each image is unwrapped at 0.5^o resolution - i.e. 720 pixels in the horizontal direction. The brighter, top, part in each figure is the k-means cluster in question (K_O). The boundary between the top cluster and the central cluster (K_P) is evaluated for smoothness, and t is increased in Eq. 6.1 until any jump in the border between the top and middle cluster does not exceed 20 pixels. For these images, $t = \{1, 4, 6, 9, 11, 13, 15\}$.

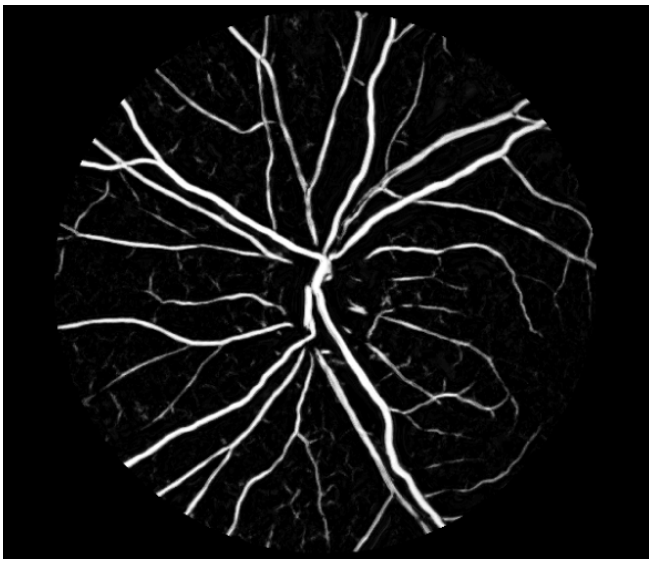
#	Description
1	Vessel Discontinuity Index (VDI)
2	VDI scaled by distance from center
3	Border Vessel nasal to temporal gradient ratio
4	Ratio of ONH vessels to Peripheral scaled
5	Ratio of ONH vessels to Peripheral limited
6	Artery-to-Vein ratio
7	Max vessel width
8	ONH swelling area
9	Ratio of central nasal to temporal distance to fit ellipse edge
10	Ratio of superior nasal to inferior temporal distance to fit ellipse edge
11	Ratio of inferior nasal to superior temporal distance to fit ellipse edge
12	Mean ONH swelling border gradient value
13	ONH temporal to nasal border gradient ratio
14	MUCH components (four total)
15	dMUCH imaginary component minus phase
16	Maximum of small vessel sums
17	Mean of max 5 small vessel sums
18	Maximum of small vessel means
19	Mean of max 5 small vessel means
20	Maximum vessel curvature
21	Mean of five highest vessel curvatures
22	Mean of greater-than-average vessel curvatures
23	ONH vessels to Peripheral scaled
24	Mean of small vessel means greater than the mean
25	Max vessel segment tortuosity
26	Location of tortuosity trend derivative equal to 1
27	Tortuosity trend parameter

Table 6.1: Table of all fundus features.

the number of upper end-points (Fig. 6.4). Additionally, as it is expected that the count of vessel discontinuities nearest the ONH is a more reliable indicator for the swelling, a scaled version of the VDI$^{\langle 2 \rangle}$ is also calculated, with each discontinuity weighted exponentially by its distance in pixels (scaled by magnification) from the approximate center.

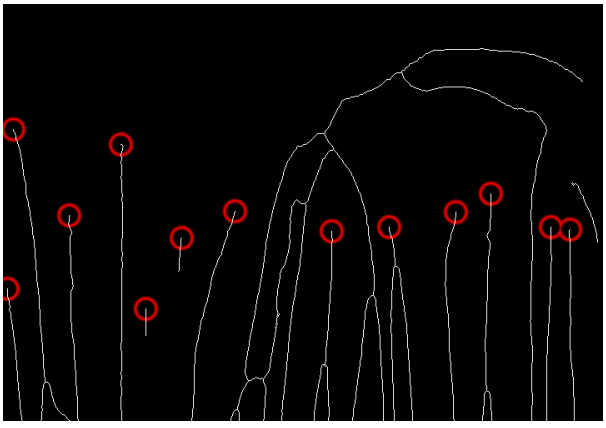
In addition to quantifying the errors caused in the vessel detection due to vessel obscuration, the vessel segments that were detected are also looked at in an attempt to assess some measure of how close they were to becoming errors. An evaluation of the vessel gradient at the borders of the vessels is a straightforward way of evaluating the obscuration of the vessels. However, the vessels are too randomly distributed in both location and numbers to evaluate them all in any consistent way. As such, it would seem prudent to focus on the location with the most clinical significance - and that is the vessels as they leave the surface of the ONH. Using the border found as described in section 6.2.2, the gradients of the detected vessels along the ONH borderline are evaluated. Specifically, in cases of papilledema it is common for the nasal region to exhibit signs of swelling before the temporal region. Thus, a ratio of temporal vessel border gradients to nasal vessel border gradients is used for a feature$^{\langle 3 \rangle}$.

In severe cases of papilledema, even the large vessels on the ONH become obscured. As such, a simple count of vessel pixels on the ONH, and off, was used to create another ratio. However, despite this being unitless, magnification can have a significant influence due to coverage. For example, if the magnification is such that the ONH takes up more than half the image, one would expect this ratio to be much larger than at smaller magnifications. To address this, the region outside the ONH is limited to 50 pixels in radius, scaled by magnification$^{\langle 4 \rangle}$, and scaled by magnification squared$^{\langle 23 \rangle}$. However, this 50-pixel radius still depends on the accuracy of our ONH segmentation, which is intentionally not accurate because the intent is to segment the swelling of the ONH, not the ONH itself (unless there is no swelling). So, another



(a)

(b)



(c)

Figure 6.4: For features $\langle 1, 2 \rangle$: (a) Vessel segmentation. (b) The unwrapped vessel segmentation. Image cropped vertically at 500 pixels to show most relevant region. (c) The found vessel discontinuities, circled in red.

ratio is introduced consisting of the total vessel pixels inside the ONH swelling area to those outside, scaled by the magnification⟨5⟩.

Swelling of the optic disc tends to restrict blood flow out of the eye, which can cause retinal veins to expand [12]. As such, a measure of the maximum vessel width could be used as an additional indicator of the severity of the swelling. For this, the binary vessel segmentation was eroded completely, counting the number of pixels removed at each iteration. From analyzing the erosions of several images outside of our dataset, it was noticed that the resulting count, when plotted against the iteration number, was approximately Gaussian. This information is used by noting that if there is no significant difference between artery and vein widths, they will both erode at the same rate and approach zero pixels at the same rate, effectively resulting in matching Gaussians. However, if the veins are notably wider than the arteries, the arteries will erode sooner, and the veins will erode after that, effectively resulting in a disjointed Gaussian.

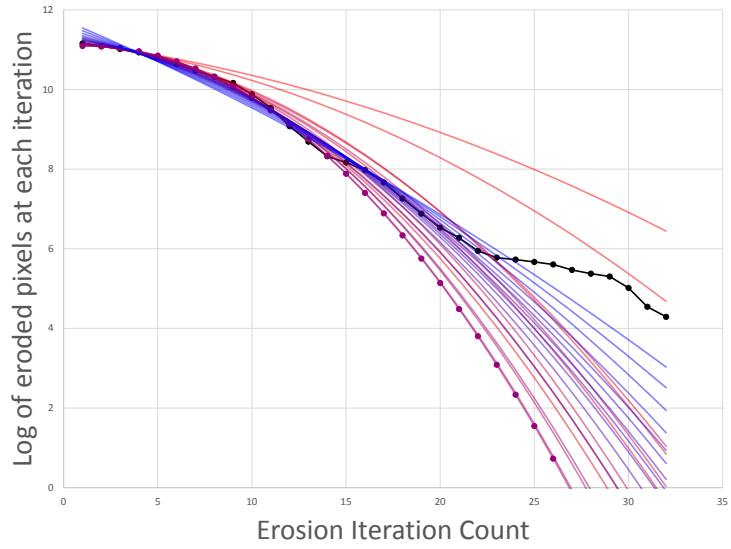
This disjoint is quantified by taking the log of the count of pixels eroded at each iteration (Fig. 6.5) and fitting the trend to a parabola, starting with the first three iterations and adding each iteration to the fit. The intercept with 0 for each parabola fit is computed, and the minimum intercept is assumed to represent the parabolic parameters that correspond to the maximum artery width. The remaining iterations are split up into groups of three, with each group also fit to a parabola. The quadratic and linear parameters of each fit are compared to the artery fit, and the closest match is assumed to be the curve corresponding to the veins, with its intercept with 0 assumed to represent the maximum vein width. From this, an artery-to-vein ratio (AVR)⟨6⟩ is computed as a ratio of the intercepts. Since this is a unitless ratio, it is not dependent on magnification. However, the intersection of 0 with the venous parabola, scaled by magnification, is used for a representation of the maximum vessel width⟨7⟩.

The last group of vessel features are an attempt to quantify vessel curvature. As the ONH swells, it restricts blood flow out of the eye. As mentioned above, this restriction can cause the veins to swell, but also sometimes to twist. This tortuosity of the vessels is computed by breaking the skeletonized vessel segmentation into objects with only two endpoints. Any skeletal object with more than two endpoints is broken up at its intersections. The area of the convex hull of each object is computed, along with the length of the skeleton. The area-to-length ratio is computed for each object. The max$^{\langle 20 \rangle}$ and mean of the max five$^{\langle 21 \rangle}$ were computed. Additionally, as most lower-value ratios are expected to correspond to noise objects (and, more importantly, most noise objects to have low ratios), the average ratio of all ratios greater than the average$^{\langle 22 \rangle}$ is also taken.

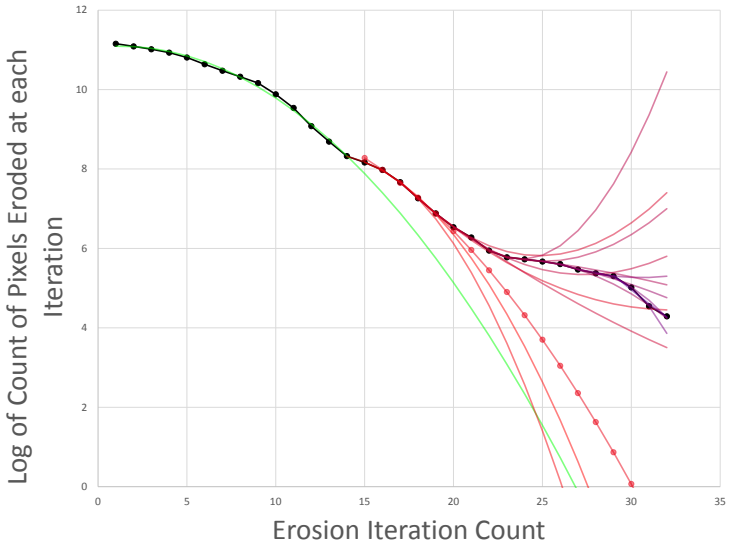
6.3.2 Optic Disc Features

Using the segmentation found from section 6.2.2, the most straightforward and obvious feature to use is the cross sectional area of the swollen ONH$^{\langle 8 \rangle}$. This is computed by a simple count of the pixels enclosed by the ONH swelling border, and scaled by magnification. Ratios of the deviations in the ONH swelling edges$^{\langle 9,10,11 \rangle}$, with respect to an ellipse fit, are used as features as in Echegaray et al [5]. The distance, in pixels, between the edge of a fit ellipse and a point on the segmentation at 30 degree intervals is first found. Then the ratio of the temporal distance to nasal distance was computed by the transformation $r = (1 + d_t)/(1 + d_n)$ to eliminate the possibility of errors due to a nasal distance of zero.

These edge ratios are assumed relevant because the ONH border becomes obscured with increasing severity. Thus, a greater uncertainty in ONH edge location is expected for swollen cases, which results in different ratios than would be found from healthy patients. However, the theory behind these ratios only holds for mild-to-moderate cases of swelling (also identified as Frisén grades 1 or 2), in which the ONH border is not circumferentially obscured. For anything beyond that, the temporal edge is



(a)



(b)

Figure 6.5: For features $\langle 6, 7 \rangle$: The artery (a) and vein (b) fits. The black dots are the log of erosion counts at each iteration. For (a) the brightest red indicates a fit to the first three points, and as each point is added the shown curve approaches blue - i.e. the second-brightest red indicates a fit to the first four point, the third-brightest the first five points, and so on. The fit with the lowest intersection with 1 is chosen to represent the arteric width. This selected fit is shown in (b) in green and annotated with dots in (a). For (b) three points are selected at a time, starting with the last point used for the arteric fit. The group of three that has the quadratic and linear parameters most closely matching the arteric fit is chosen for the venous fit. The selected venous fit is annotated with dots in (b). The intersections with 1 for the arteric and venous fits are used for the AVR $\langle 6 \rangle$ and the venous intersect with 1 is scaled by magnification and used for the maximum vessel width $\langle 7 \rangle$.

expected to be just as uncertain as the nasal edge, and these ratios could easily go the other way. For such cases, some measure of the obscuration of the ONH border is still a good idea. Using the radial gradient image, a magnification-scaled Gaussian filter is applied to achieve a pseudo-normalization of sharpness across magnifications. Masking out the vessels, the mean gradient strength along the ONH border is used as a feature$\langle 12 \rangle$.

As mentioned, the Gaussian filter is applied in an attempt to achieve a pseudo-normalization for various magnifications - the same edges are naturally sharper at lower magnifications, and the effects magnification has on decorrelating this feature with the truth needs to be minimized. However, there are other factors besides magnification that could interfere with the reliability of this feature, such as image artifacts, contrast inconsistency, and of course the image simply being out of focus. In an attempt to address this, another ratio of the mean vessel border gradient strength on the nasal side to the mean vessel border gradient strength on the temporal side$\langle 13 \rangle$ is created. While it is expected that this will be limited in the same way as the elliptical fit error ratios $\langle 9, 10, 11 \rangle$, for the same reasons, it is also expected that this will be a more stable feature, as there is no need to offset the means to account for the variance in relative image clarity between images.

The in-painted greyscale image is unwrapped about the approximate center and vertically cropped to 800 pixels[1], scaled by magnification. This image is then histogram equalized to achieve a quasi-normalization, and the horizontal mean is taken. In the ideal case, the Mean of the Unwrapped Cropped Histogram-equalized image (MUCH) would resemble a half-wavelength sinusoid (Fig. 6.6). As such, a measure of how sinusoidal the MUCH is could be indicative of how ideal the overall appearance of the ONH is. The imaginary component, phase, magnitude, and real component$\langle 14 \rangle$

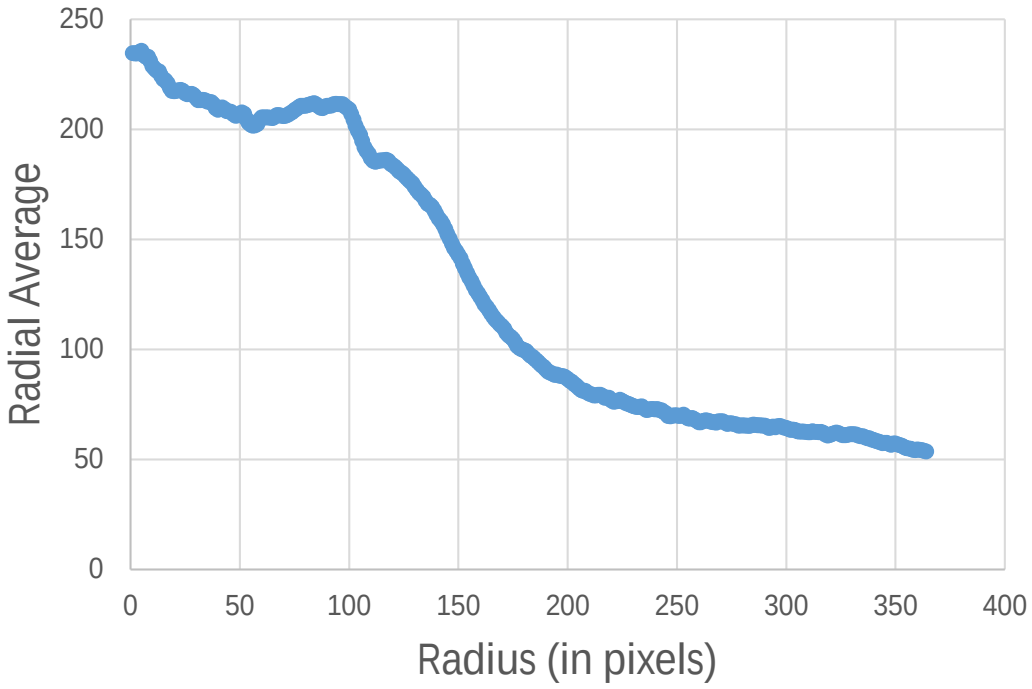
[1]800 pixels was experimentally determined on a separate set of images to be enough to include the ONH area, but exclude the edges of the fundus image. This assumes the ONH is reasonably centered in the fundus image.

of the first coefficient of the FFT of the MUCH are used as features - the phase is shifted to range between $(0, 2\pi)$. To minimize the effects of non-sinusoidal elements, instead of using the MUCH directly, the derivative of the MUCH is used. This shifts the expected high value, in the case of healthy eyes, from the cosine (real) component to the sine (imaginary) component. In normal eyes the phase is ideally expected to be zero, or near zero, and the imaginary component to be high. To accentuate this expectation, the difference between the imaginary component and the phase is also as a feature⟨15⟩.

With the ONH segmented, there was also a focus on the small vessels overlaying the ONH. In swollen cases, these vessels tend to become harder to distinguish, and it would be prudent to quantify that. To do this, first an image of only the swollen ONH area is created, scaled, and upsampled by 4. The upsampling is done because the small vessels on the ONH are usually too small to be detected using the method in section 6.3.1. However, the vessel segmentation results from 6.3.1 is used to mask the vessel segmentation here, to prevent the larger vessels from interfering with this evaluation.

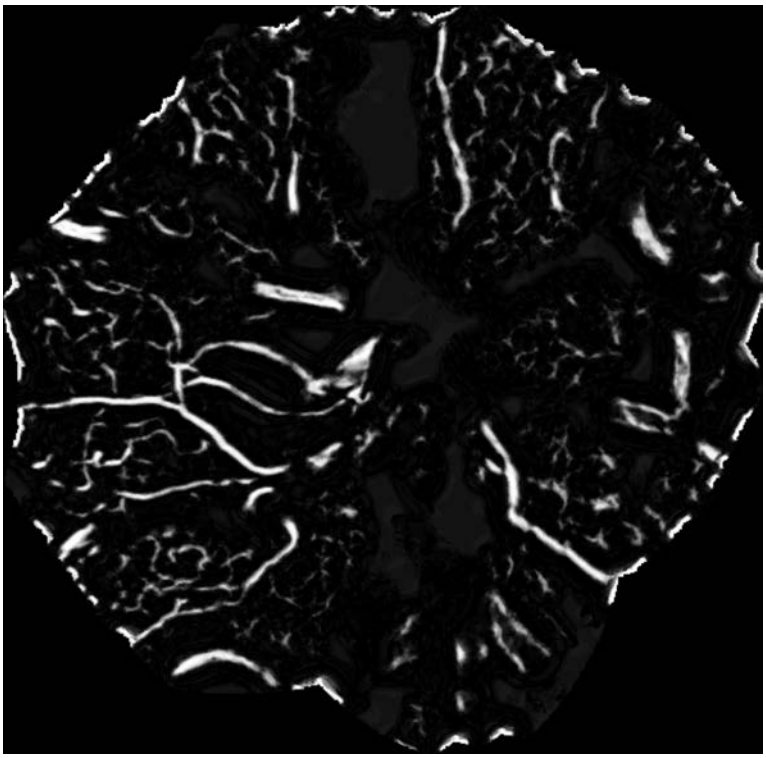
The small-vessel segmentation (Fig. 6.7) is thresholded, resulting in an object image. Each object is then skeletonized, shrunk, and cleaned. The distance from each skeletal pixel to the nearest dark point in the object image is computed. Each object is evaluated for the consistency of distance values its skeleton contains. Specifically, the standard deviation and sum of skeletal distance values are stored for each object. The sums are computed because the longer a detected vessel is, the more it can be said that the vessels on the surface of the ONH are visible. Each object sum is also weighted by the standard deviation of the skeletal distances, as larger detected objects with high sums may not be long thin vessel-like objects. The maximum weighted sum⟨16⟩, and the average of the top five⟨17⟩ are each used as features. These features are also repeated using means instead of sums ⟨18,19⟩.

(a)

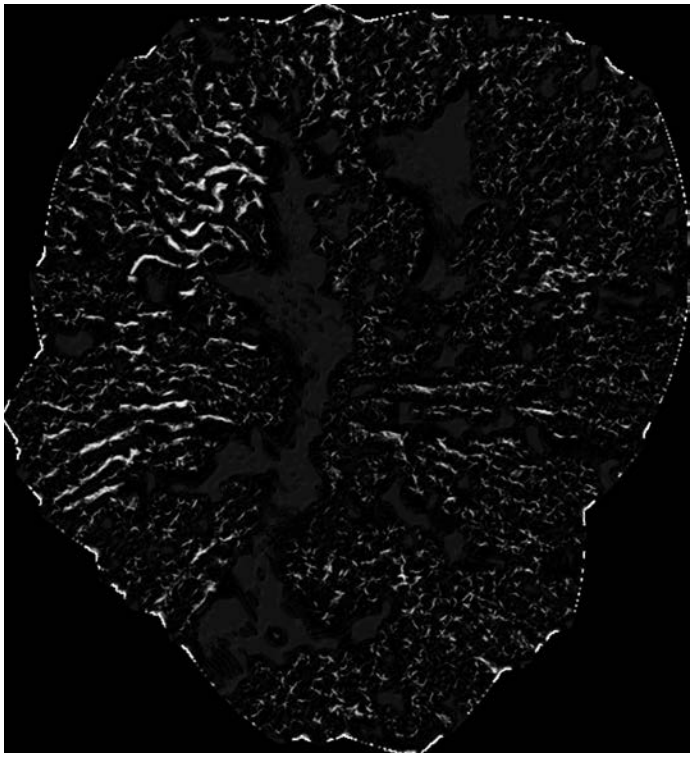


(b)

Figure 6.6: For features $\langle 19, 20, 21, 22, 23 \rangle$: (a) The unwrapped, cropped, and histogram-equalized image. (b) The horizontal mean (MUCH).



(a)



(b)

Figure 6.7: Small vessel segmentations corresponding to total retinal volumes of (a) 12.57mm^3 and (b) 26.56mm^3.

6.3.3 Feature Selection

With the 24 features extracted, the next step is to determine which of these are the most useful for predicting volume. Additionally, it is also true that some of these features are likely to be of unequal use for varying volumes. The presence of small vessels on the ONH, for example, is expected to carry significant weight for moderate to high swelling, but deviations in the results from quantifying small vessel presence for mild swelling should be comparable to that of normal patients. While this is precisely the sort of thing that a random forest regression analysis is intended to take into account, it may prove prudent to ensure the results include as much of the range of the truth values as possible. In so many words, when the truth is plotted against the results and fit linearly, a high correlation is of primary interest, but obtaining slope as close to 1 as can be achieved may also prove worthwhile. Intuitively, a slope of exactly zero or a slope of exactly one would necessarily have a high correlation, but it is also known that, practically speaking, such ideal values are unlikely. To get a sense of what happens when either the slope or correlation is prioritized, a cascaded random forest regression analysis is implemented using the following steps:

1. Select one patient to leave out. This results in two datasets - one entirely consisting of the left-out patient (P) and the other consisting of all other patients (P').

2. Train the random forest classifier on all of P' to generate importance values for each feature.

3. Run a leave-10%-out random forest regression analysis on P'.

   (a) Order features by importance.

   (b) First four features are included.

   (c) Remaining features are added or removed four at a time. If a reduction in

RMS is found, each of the four is swapped out one by one. If no reduction in RMS is found, the next four in the list are added.

(d) Repeat 3c until a feature selection that results in the lowest RMS is found for P'.

4. Using the results of the regression analysis, P' is further divided into four groups $P''_1, P''_2, P''_3, P''_4$ for each 25^{th} volumetric percentile of P'. Step 3 is repeated, for each P''_n. However, because the dataset used to develop this algorithm is limited in size, adjacent $P''_{n\pm1}$ are added to the group that is to be left out and iterated over. This results in five total regression analysis random forest models, each with their own set of features.

5. The left-out patient P is then tested on the model that results from (3).

6. Depending on which percentile the results from 5 falls in, P is then tested on the corresponding model from 4, to obtain a final volume prediction for P.

7. Steps 1 through 6 are repeated for each patient.

Naturally, to save processing time, it serves to only train the model from step 4 that is to be used, rather than all four. The intent of this is to use a standard leave-one-patient-out random forest regression analysis to prioritize the highest possible correlation, at the possible cost of a slope further from one, but to compare these results with this cascaded random forest regression analysis to attain a slope closer to one at the cost of correlation.

6.4 Experimental Methods and Results

Multiple fundus/SD-OCT image pairs were taken for each of 44 patients with papilledema at the University of Iowa. The SD-OCT-derived ONH volume corresponding to each color fundus photograph was computed using the methodology described in [28] from the SD-OCT images. One fundus/SD-OCT image pair was selected for

further analysis from each of the 44 patients by choosing the image pair with the largest amount of swelling from the ONH volume computation of the SD-OCT image. When more than one fundus image was available for a given SD-OCT volume, the fundus image which had the ONH area most in focus was chosen for use in this algorithm.

When comparing the volumes computed with from the SD-OCT images and the volumes predicted through this method, initial results were reported at an SPIE conference in 2015 [41], with a mean square error of 2.27 mm^3 and correlation of $R = 0.788$. Since then, after refining features, a mean square error of 1.93 mm^3 was achieved with a correlation of $R = 0.887$. After feature selection, an RMS error of 1.93 mm^3 was achieved, with a correlation coefficient of $R = 0.902$, and using feature selection in a cascaded random forest regression analysis resulted in an RMS error of 1.87 mm^3 with a correlation coefficient of $R = 0.888$.

An importance metric for each feature was also computed by running the classifier on the entire data set and extrapolating the increase in accuracy associated with trees that include each feature (Fig. 6.8). The predicted and actual values are shown in Fig. 6.9. The most important features dealt with the ONH. The area enclosed by the swollen ONH boundary$^{\langle 8 \rangle}$ and its greyscale entropy$^{\langle 13 \rangle}$ were of particular importance. Additionally, the first imaginary component of the FFT of the MUCH$^{\langle 19,20 \rangle}$ turned out to be of notable importance, along with the mean derivative of the swollen ONH boundary$^{\langle 12 \rangle}$. This suggests that, for the purpose of estimating the swelling of the ONH, features derived from the ONH are the best indicators.

6.5 Discussion

The determination of severity covered in this chapter applies only to papilledema cases. While useful, it may also be prudent to determine if there is a clinical use for the volume-defined severity for other causes of optic disc edema. While the features extracted in this doctoral work were primarily intended to follow the progression of

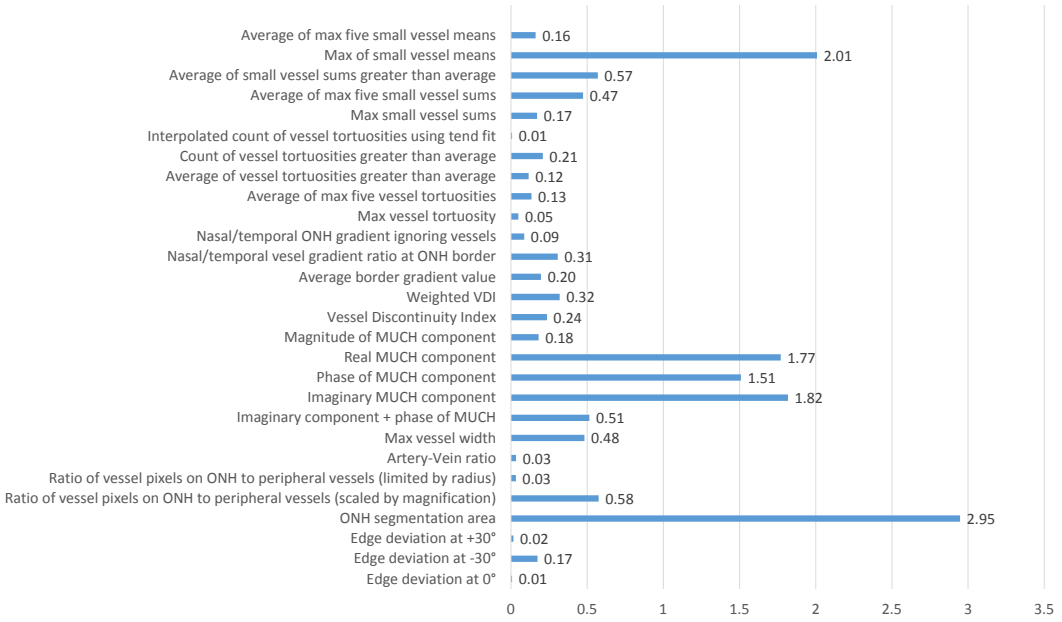


Figure 6.8: Average importance of each feature used for prediction of volume from fundus features.

the Frisén scale, swelling due to any cause is going to have similarities to swelling due to any other cause. As such, these same features may perform just as well when trained on other causes of swelling.

In any event, these results are overall decent. There is clearly room for improvement, but it is unclear as to how much improvement is even possible. The idea of predicting a 3D attribute from 2D data is inherently limited. This is underscored by noting that one of the most important features was that of the ONH swelling area, which was constructed for the sole purpose of obtaining a direct cross-section of what the TRV represents. Indeed, perhaps the most prevalent reason for most of the outliers is that an accurate segmentation of the swollen ONH area depends mainly on lighting, focus, contrast, and any number of imaging nuances that can become a problem. In Fig. 6.10, for example, we can see that an outlier in this prediction had an ONH swelling segmentation far inside the area that one might consider swollen.

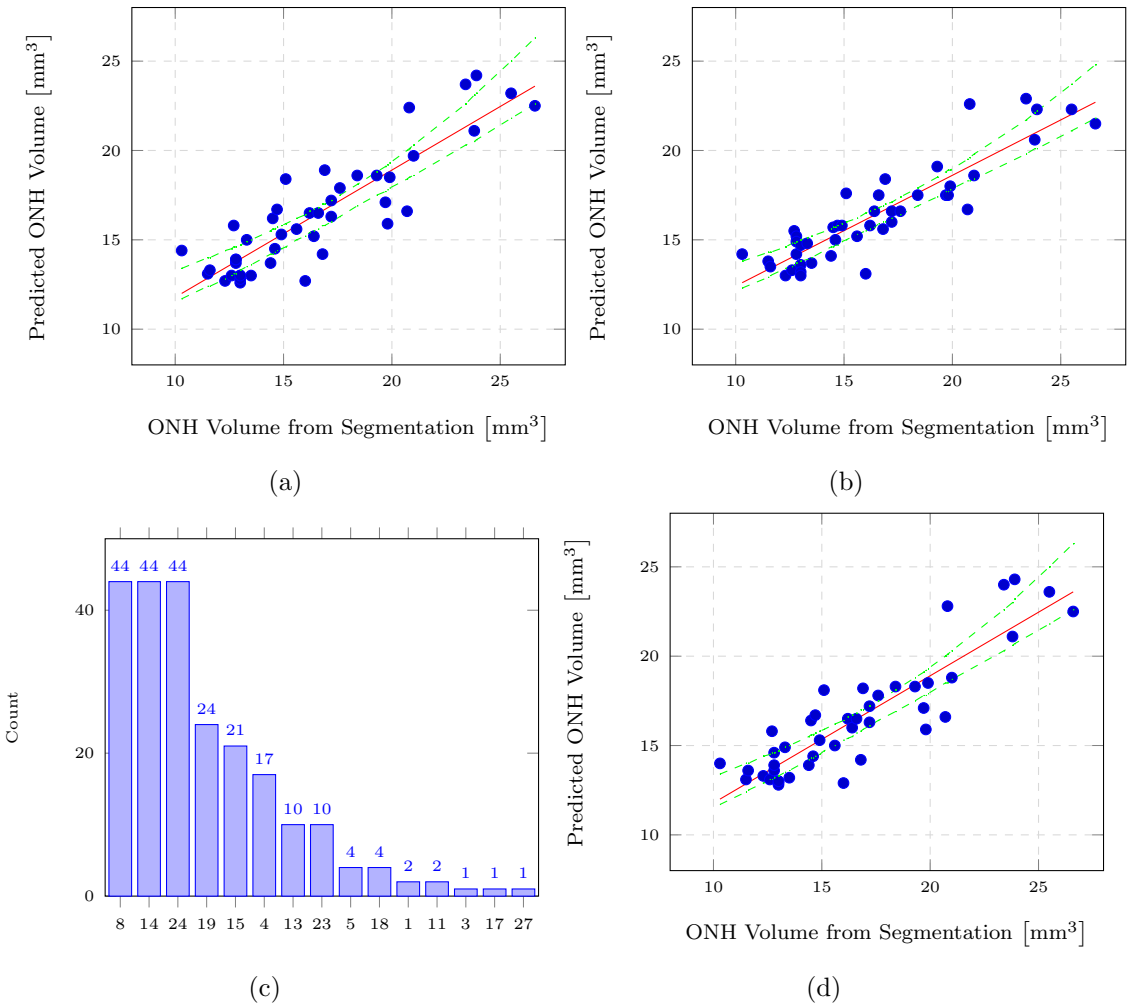


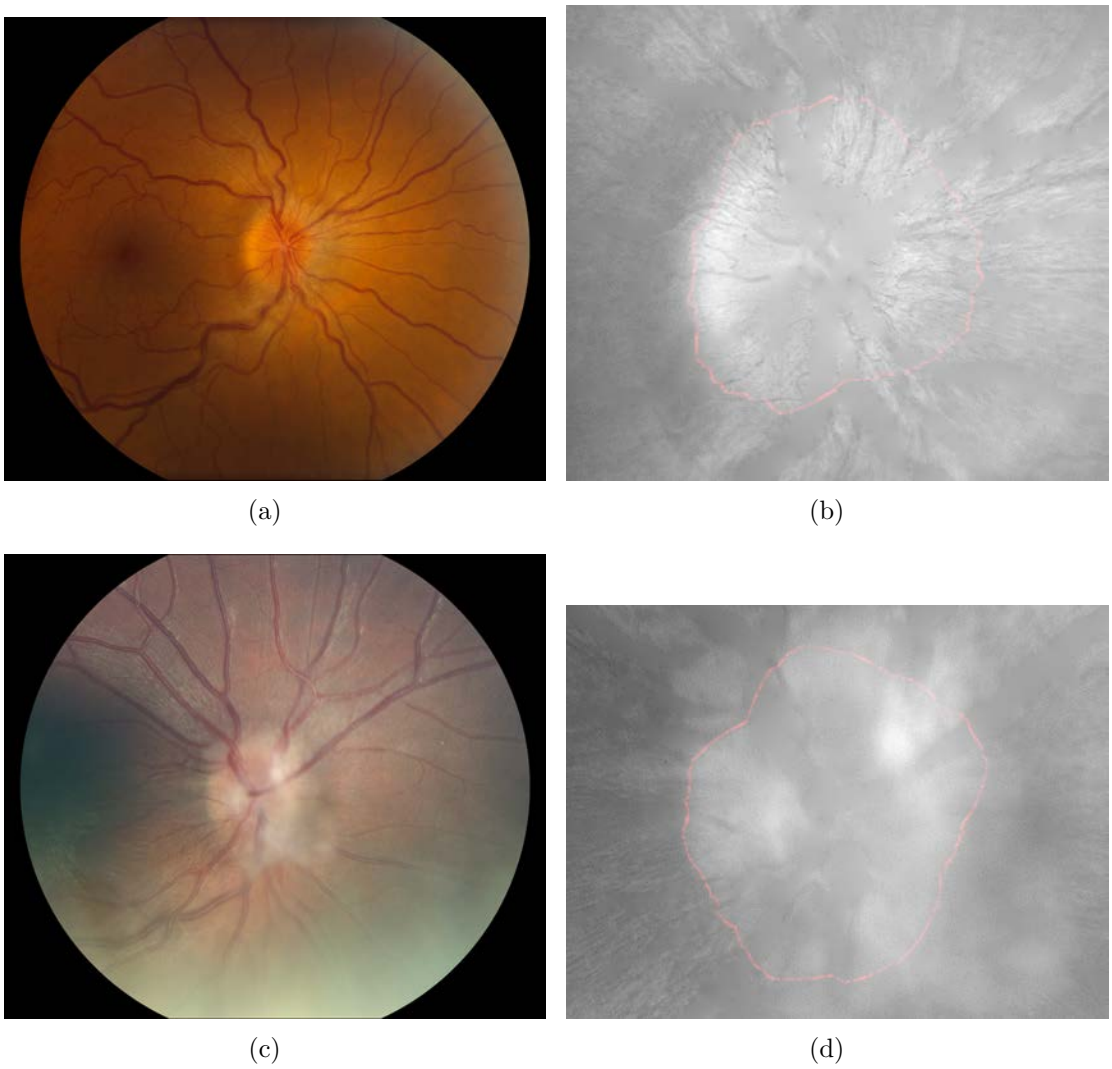
Figure 6.9: Plot of predicted vs. actual volume. The first plot (a) is using all features in a leave-one-out random forest regression. This resulted in a correlation of $R = 0.887$ and an RMS of 1.87 mm^3. The slope of the trend line was 0.717. The second plot (b) was after a non-cascaded feature selection, using up to steps 3d in section 6.3.3 in an attempt to get the highest correlation. This had a resulting correlation of $R = 0.902$ and an RMS of 1.93 mm^3, with a trend slope of 0.62. The final plot (d) used all of the steps in section 6.3.3 in an attempt to get the highest slope and had a resulting $R = 0.888$ and an RMS of 1.87 mm^3 with a trend slope of 0.714. The feature occurrence (c) shows how many times each feature was used, with the feature labels corresponding to Table 6.1. Confidence bands (95%) are shown in green.

This is mainly due to the dull color of the swelling around the ONH overlapping the ONH itself, which creates difficulty in finding an appropriate segmentation, even by

hand. Furthermore, in Fig. 6.11, we see that the small vessel calculations on the ONH were more clear than one might reasonably expect, as compared to the small vessel calculation for the nearest image in volume, which was not only obscured by a cataract, but also magnified enough for the original vessel segmentation to find the small vessels on the ONH. This plus the venous dilation inherent in ONH swelling images caused some of the small vessels to be masked before computing the small vessel attributes.

Another example in Fig. 6.12 shows hemorrhaging being treated as detected vessels (Fig. 6.13), which again a false border for the ONH and creates false small vessels on the ONH. On top of all of this, another key feature was the MUCH components, which were effectively a way to determine a radial intensity profile between the ONH and the surrounding area after local histogram equalization. In Fig. 6.14 we can see the shadows in (a) and the hemorrhaging in (c) set the groundwork for giving a high contrast between the ONH and the surrounding area in a radial intensity profile, thus contributing to the outlier status. Furthermore, the reliability of the MUCH depends heavily on a good segmentation of the ONH swelling area, as the MUCH is centered on that area. If the segmentation is clipped, as in Fig. 6.10(b) or Fig. 6.12(b), it will offset the center to capture the best contrast in radial intensity, which is falsely inflated by shadows and hemorrhaging.

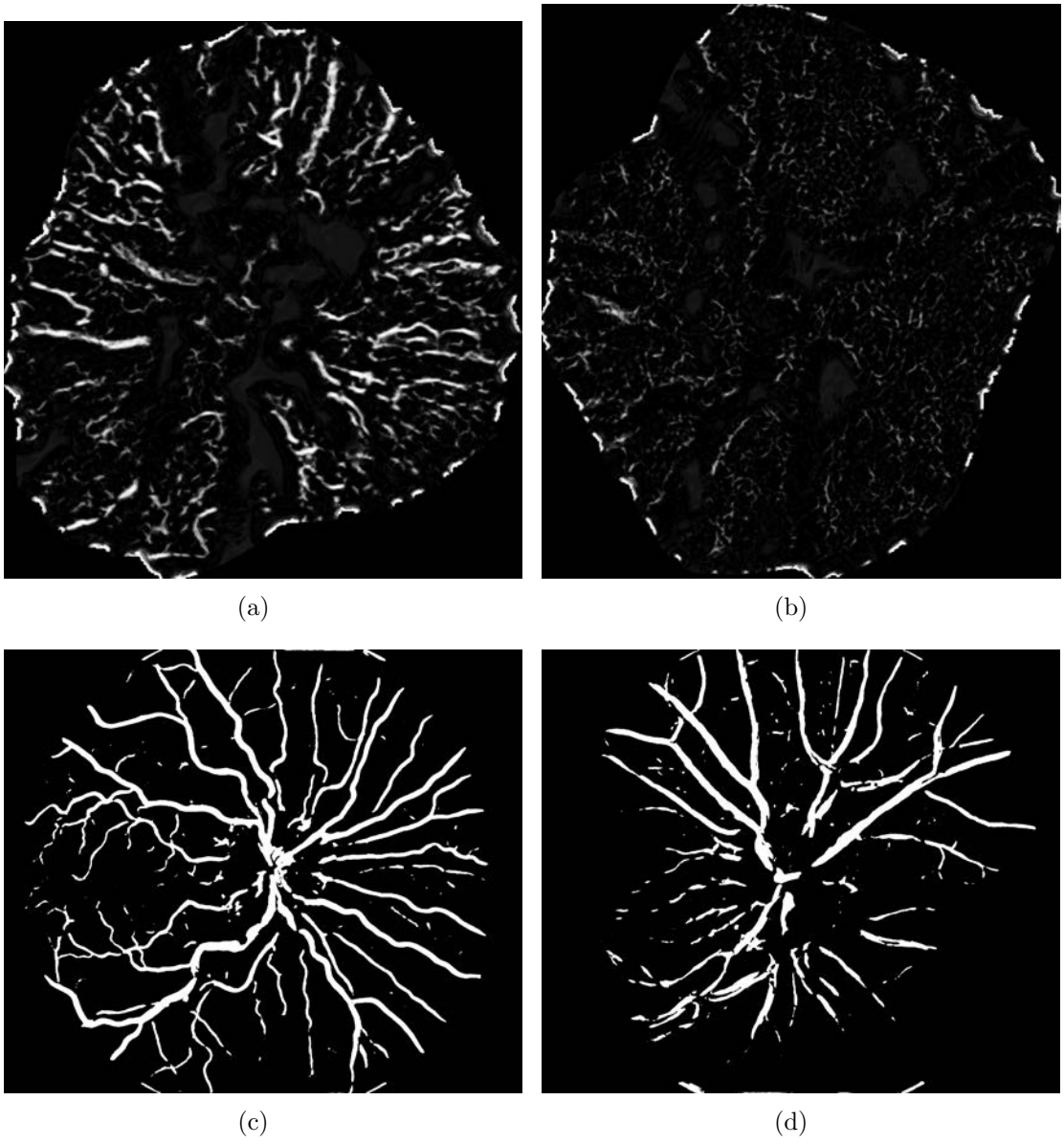
In future work, it would be useful to detect and inpaint hemorrhaging separately, in addition to implementing some form of image correction in the event of cataracts. While this seems to be a bit daunting, especially the latter, it could be possible given that one can view a fundus image with a cataract and still know what they are looking at. As long as the expectations are known, it could be possible to adjust an image in such a way that allows a machine to work within the framework of those expectations, instead of attempting to train around the anomalies.



(a) (b)

(c) (d)

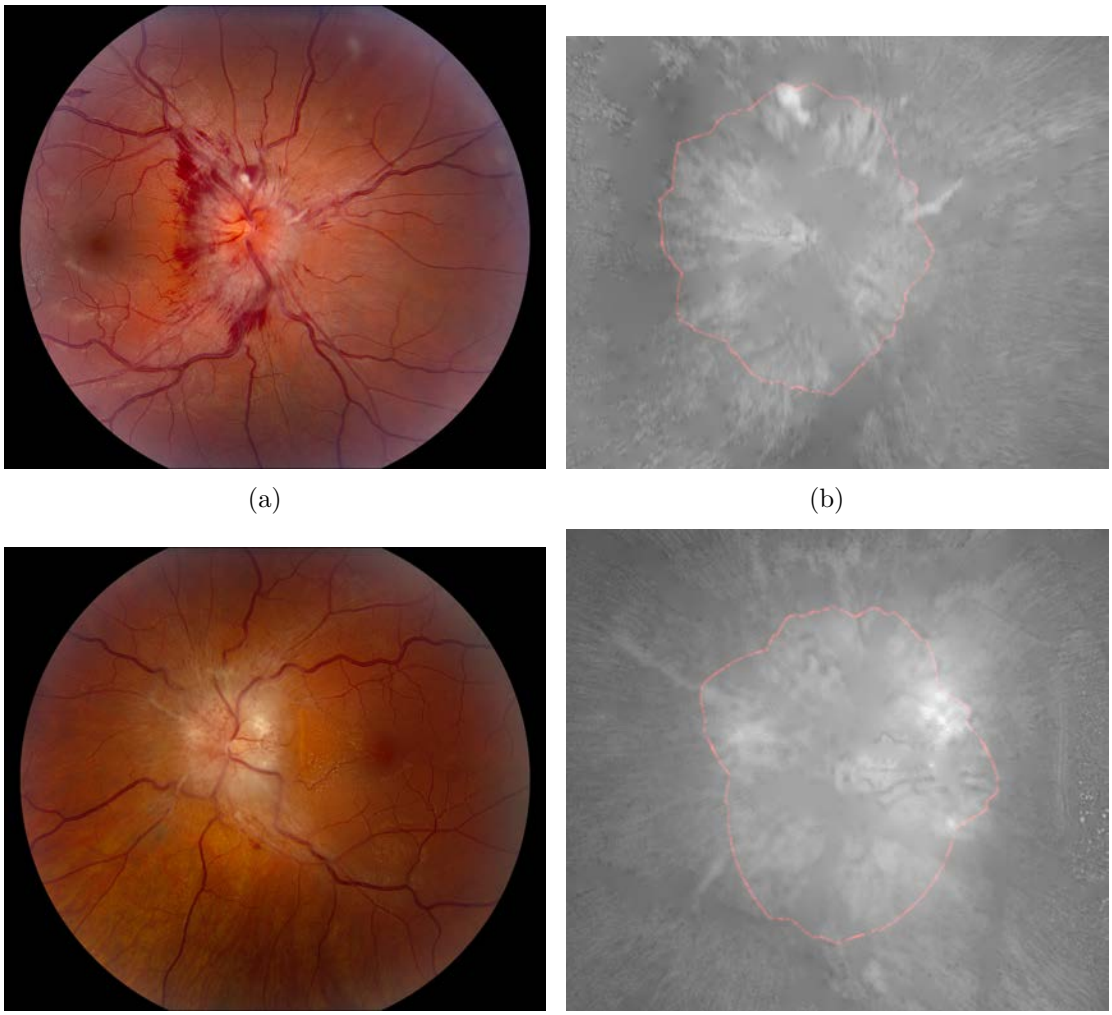
Figure 6.10: Original fundus image (a) and ONH swelling segmentation (b) of an outlier. This had a predicted volume of 13.1 mm^3, but an actual volume of 16.0 mm^3. Compared to the next-closest image in volume (c and d) of actual volume 16.2 mm^3, with a predicted volume of 15.8 mm^3. The swelling area would normally be outside the dull area in (a), but due to the contrast between the ONH and the surrounding area, it is instead inside the dull area.



(a) (b)

(c) (d)

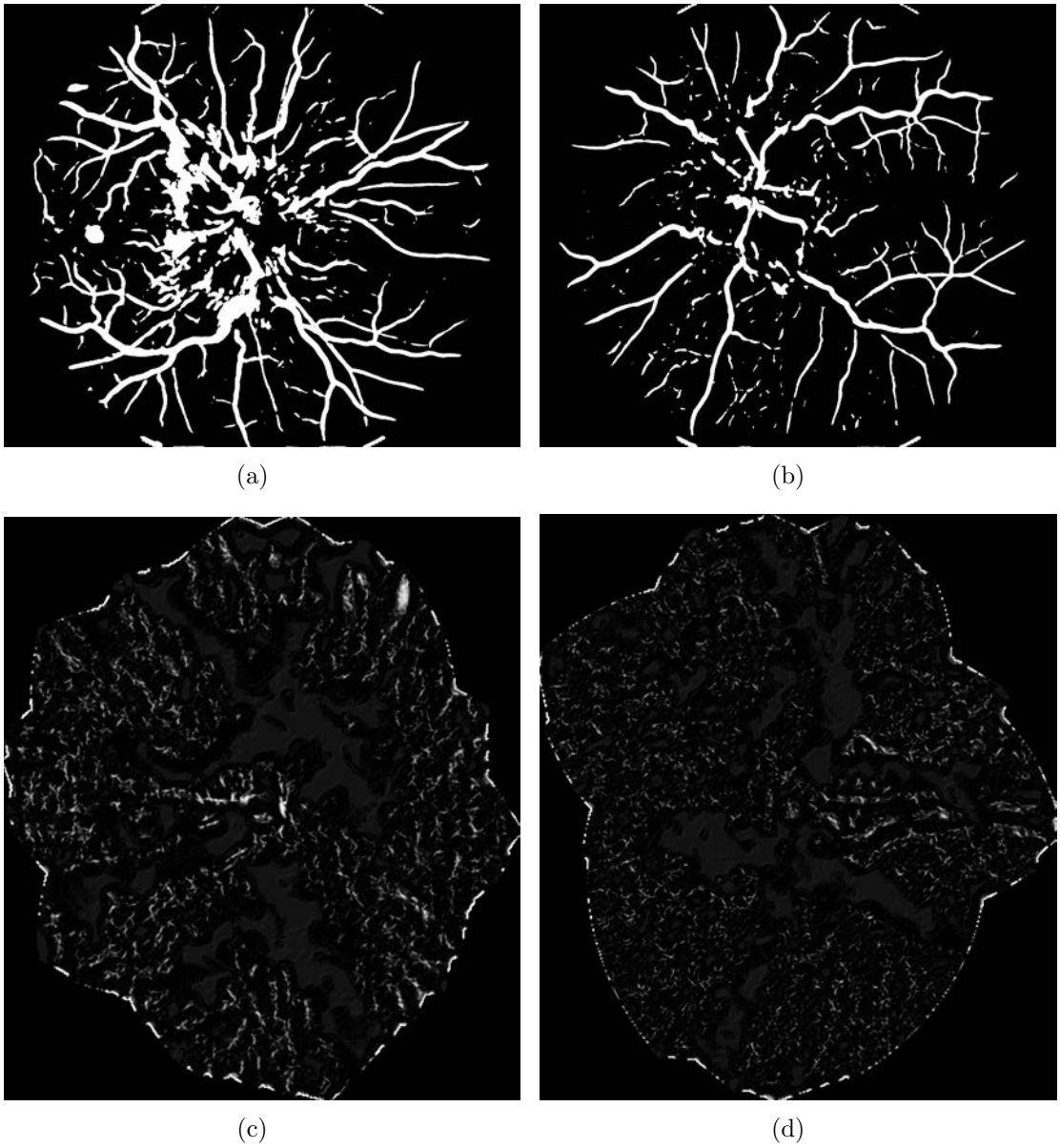
Figure 6.11: The small vessel and vessel segmentation (VDI) comparisons for an outlier (a and c) with an image of similar volume (b and d) that was not an outlier. These are additional features continued from Fig. 6.10. This is an example of an image with notably better quality than others of comparable volume resulting in a prediction of much lower volume.



(a) (b)

(c) (d)

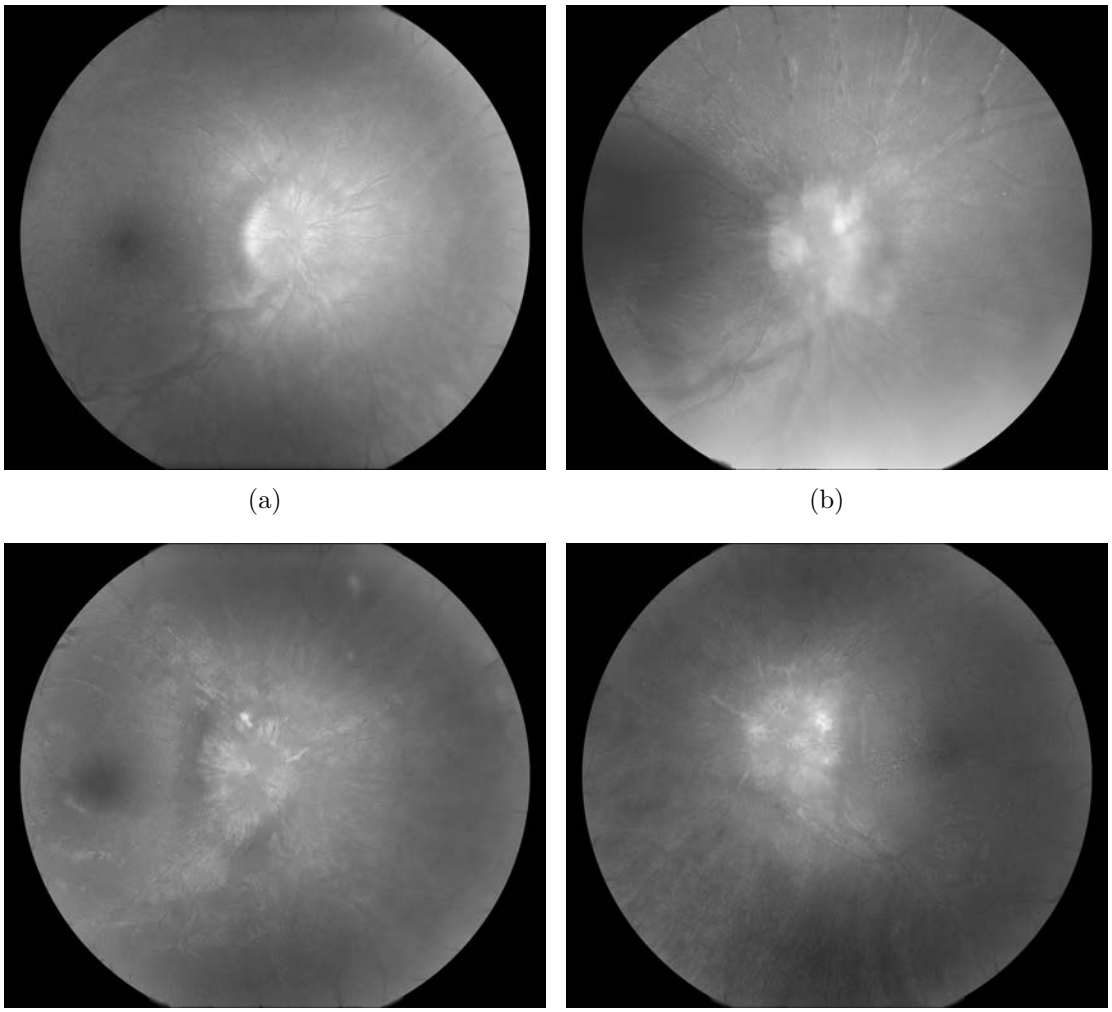
Figure 6.12: Original fundus image (a) and ONH swelling segmentation (b) of another outlier. This had a predicted volume of 16.7 mm^3, but an actual volume of 20.7 mm^3. Compared to the next-closest image (c and d) of actual volume 20.8 mm^3, with a predicted volume of 22.6 mm^3. The brightness of the area on the ONH compared to the dull color of the swelling around it contributed to a false, smaller, segmentation that resulted in a low prediction.



(a) (b)

(c) (d)

Figure 6.13: The small vessel and vessel segmentation (VDI) comparisons for an outlier (a and c) with an image of similar volume (b and d) that was not an outlier. These are additional features from Fig. 6.12. The hemorrhaging in (a) results in a number of false vessels which affect the VDI, but also add a number of false vessels to the small vessel attributes.



(a) (b)

(c) (d)

Figure 6.14: Greyscale inpainted image the MUCH components were extracted from. Actual (predicted) volumes were (a) 16.0 mm^3 (13.1 mm^3) (b) 16.2 mm^3 (15.8 mm^3) (c) 20.7 mm^3 (16.7 mm^3) (d) 20.8 mm^3 (22.6 mm^3). The shadows in (a) and (c) around the ONH accounts for a low prediction.

CHAPTER 7
CONCLUSIONS

Swelling of the optic disc may be somewhat rare, both in general and compared to other problems that may be found with the optic disc, but when it is a problem it tends to be one of particular significance. With the exception of pseudopapilledema, when an optic disc is swollen there is at least the threat of abrupt loss of eyesight. The ease with which the back of the eye can be observed to check for this is awfully convenient, but there is still the possibility of observer uncertainty. Due to the rarity of optic disc edema, there tends to not be extensive training on the topic for emergency or any personnel outside of a neuro-ophthalmology clinic. It is, thus, an understatement to say that a device that can automatically assess a risk, and its severity, would be useful.

Pseudopapilledema tends to be the ultimate confounder in this field, as an optic disc that looks swollen naturally is easy to mistake for one that is swollen for risky reasons. There are a number of ways that experienced ophthalmologists can tell the difference, but it is not something the general population of medical practitioners tends to know. Somewhat recently, folds in the retina have been discovered that are caused due to the pressure a swollen optic disc applied to its surrounding area. These folds are perhaps most easily visualized in SD-OCT images, but can also usually be seen in fundus images, and are a bit of a giveaway when it comes to determining if an optic disc is swollen due to some problem. In Chapter 4 folds were evaluated for their presence and attributes. In addition to the presence of folds ruling out the likelihood of pseudopapilledema, it has been theorized that different types of optic disc edema can result in different types of folds. As the detection and quantification of folds is a largely unexplored area, this doctoral work first started with a discussion on how to best visualize them. Truths were marked, and features extracted to train a pixel-based classifier to produce a probability map in the service of detecting folds.

The attributes of the folds are, indeed, among the features used, which simplifies the results as part of a classification of the cause.

In Chapter 5, the heart of the issue was addressed, and that is the distinction of the cause of optic disc edema. As mentioned, the ultimate goal is to determine the cause and severity of optic disc edema using only fundus images, completely automatically. However, this is also a largely undeveloped area and as such it is important to give it a useful starting point. Along that line of logic, both SD-OCT and fundus images were used. Among other things, the shape of Bruch's membrane is believed to be an important factor in distinguishing between causes for which the optic disc is swollen due to problems posterior to the optic disc or anterior. Specifically, it is theorized that the opening to Bruch's membrane is likely to bend towards the eye in the event an optic disc is swollen due to raised intracranial pressure (papilledema). Shape modeling has been implemented to quantify the bending of Bruch's membrane, but ordinarily requires special SD-OCT imaging variations to make it work reliably. Part of the work in this thesis was to extend the same shape modeling theory to standard SD-OCT images, to consolidate diagnosis to a more widely used imaging technique.

Additionally, many fundus features were used to determine the cause of optic disc edema. Many of them were used in Chapter 6 to predict volume and were designed to each correlate to the criteria by which papilledema is ordinarily assessed. The idea being that if one or more of such features correlated poorly, while others correlated well, a possible explanation is that the cause of optic disc edema was due to something other than papilledema.

The fundus features in question were additionally used to predict volume from fundus images. The volume of an optic disc (also called total retinal volume or TRV) was taken from SD-OCT images and is believed to be a more reliable standard for measuring papilledema severity than the current Frisén grading system, which is ordinal and subjective. These features consisted of vessel features and optic disc

features. As it is the optic disc that is swollen, it is expected that features related to the optic disc would be the most substantial in an algorithm designed to predict the severity of the swelling. This proved to be correct, and the cross sectional area of the swelling along with various texture attributes related to the optic disc turned out to be the most important in the classifier. However, the vessel features were not insignificant. The swelling of the optic disc limits blood flow out of the eye by constricting the vessels at the optic disc. This results in several different problems that can be, and were, detected, measured, and used as features. Future work could include the detection of folds from fundus images. This is not a simple task, as folds do not show up easily in fundus images but when they are visible they can be definitive. Additional future work may also include different sorts of cameras. For example, there is a camera that provides a sweep of focal points in an image capture. It may prove useful to get a sense of the topology of the optic disc from this sweep [42] and combined with some implementation of shape modeling, predictions of the cause of optic disc edema may be more accurate. Infrared cameras could also provide an image of sub-retinal layers, giving a better image of the optic disc even in the event of swelling. Stereoscopic imaging may be used as a possibly faster way to compute topology. There are a variety of possibilities when one considers the source of the images to work with, each with their limitations and benefits.

To summarize, this doctoral work consisted of the development of several machine-learning approaches to determine both the cause and, in some cases, the severity of optic disc edema based on features found in both SD-OCT and fundus images. The ultimate goal to this work is of progress towards the development of an algorithm that completely diagnosis a cause of optic disc edema, in addition to some estimate of its urgency, using only fundus images.

REFERENCES

[1] J. Agne, J. Wang, R. H. Kardon, and M. K. Garvin, "Automatic detection of folds and wrinkles due to swelling of the optic disc," in *Fetal, Infant and Ophthalmic Medical Image Analysis - International Workshop, FIFI 2017, and 4th International Workshop, OMIA 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, 2017, Proceedings*, 2017, pp. 235–242.

[2] J.-K. Wang, P. A. Sibony, R. H. Kardon, M. J. Kupersmith, and M. K. Garvin, "Semi-automated 2D bruch's membrane shape analysis in papilledema using spectral-domain optical coherence tomography," *SPIE Proceedings: Biomedical Applications in molecular, structural, and functional imaging*, vol. 9417.

[3] L. Frisén, "Swelling of the optic nerve head: a staging scheme," *J Neurol Neurosurg Psychiatry*, vol. 45, pp. 13–18, 1982.

[4] C. J. Scott, R. H. Kardon, A. G. Lee, L. Frisén, and M. Wall, "Diagnosis and grading of papilledema in patients with raised intracranial pressure using optical coherence tomography vs clinical expert assessment using a clinical staging scale," *Arch Ophthalmol.*, vol. 128, no. 6, pp. 705–711, 2010.

[5] S. Echegaray, G. Zamora, H. Yu, W. Luo, P. Soliz, and R. Kardon, "Automated analysis of optic nerve images for detection and staging of papilledema," *IOVS*, vol. 52, no. 10, pp. 7470–7478, 2011. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/21862651

[6] D. Huang, E. Swanson, C. P. Lin, J. schuman, W. Stinson, W. Chang, M. Hee, T. Flotte, and K.Gregory, "Optical coherence tomography," *Science*, vol. 254, no. 5035, pp. 1178–1181, 1991.

[7] J.-K. Wang, R. H. Kardon, M. J. Kupersmith, and M. K. Garvin, "Automated quantification of volumetric optic disc swelling in papilledema using spectral-domain optical coherence tomography," *IOVS*, vol. 53, no. 7, pp. 4069–4075, 2012.

[8] P. A. Sibony, M. J. Kupersmith, S. E. Feldon, J.-K. Wang, M. Garvin, and OCT Substudy Group for the NORDIC Idiopathic Intracranial Hypertension Treatment Trial, "Retinal and choroidal folds in papilledema," *IOVS*, vol. 56, no. 10, pp. 5670 – 5680, 2015.

[9] M. J. Kupersmith, P. Sibony, G. Mandel, M. Durbin, and R. H. Kardon, "Optical coherence tomography of the swollen optic nerve head: Deformation of the peripapillary retinal pigment epithelium layer in papilledema," *IOVS*, vol. 52, no. 9, pp. 6558 – 6564, 2011.

[10] S. A. Mayer and J. Y. Chong, "Critical care managment of increased intracranial pressure," *J Intensive Care Med*, vol. 17, no. 2, pp. 55 – 67, 2002.

[11] A. J. Sinclair, M. A. Burdon, P. G. Nightingale, T. D. Matthews, A. Jacks, M. Lawden, A. Sivaguru, B. J. Gaskin, S. Rauz, C. E. Clarke, and A. K. Ball, "Rating papilloedema: an evaluation of the frisén classification in idiopathic intracranial hypertension," *J Neurol*, vol. 259, no. 7, pp. 1406 – 1412, 2012.

[12] H. E. Moss, G. Treadwell, J. Wanek, S. DeLeon, and M. Shahidi, "Retinal vessel diameter assessment in papilledema by semi-automated analysis of SLO images: Feasibility and reliability," *IOVS*, vol. 55, no. 4, pp. 2049–2054, 2014.

[13] S. S. Hayreh, "Pathogenesis of optic disc edema in raised intracranial pressure," *Progress in Retinal and Eye Research*, vol. 50, pp. 108–144, 2016.

[14] J. Chiang, E. Wong, A. Whatham, M. Hennessy, M. Kalloniatis, and B. Zangerl, "The usefulness of multimodal imaging for differentiating pseudopapilloedema and true swelling of the optic nerve head: a review and case series," *Clin Exp Optom*, vol. 98, no. 1, pp. 12 – 24, 2015.

[15] B. L. Lam, C. G. M. Jr., and J. Pasol, "Drusen of the optic disc," *Current Neurology and Neuroscience Reports*, vol. 8, pp. 404 – 408, 2008.

[16] S. S. Hayreh, "Ischemic optic neuropathy," *Prog Retin Eye Res.*, vol. 28, no. 1, pp. 34 – 62, 2009.

[17] ——, "Anterior ischemic optic neuropathy," *Jama Neurology*, vol. 38, no. 11, pp. 675 – 678, 1981.

[18] ——, "Posterior ischemic optic neuropathy," *Ophthalmologica*, vol. 182, no. 1, pp. 29 – 41, 1982.

[19] P. Sibony, M. J. Kupersmith, and F. J. Rohlf, "Shape analysis of the peripapillary rpe layer in papilledema and ischemic optic neuropathy," *IOVS*, vol. 52, pp. 7989 – 7995, 2011.

[20] M. Wall, M. J. Kupersmith, K. D. Kieburtz, J. J. Corbett, S. E. Feldon, and D. I. Friedman, "The idiopathic intracranial hypertension treatment trial: Clinical profile at baseline," *JAMA Neurol*, vol. 71, no. 6, pp. 693–701, 2014.

[21] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," *Computer*, no. 1496, pp. 532 – 572, 1998.

[22] T. Chanwimaluang and G. Fan, "An efficient blood vessel detection algorithm for retinal images using local entropy thresholding," *ISCAS*, vol. 2, no. 5, pp. 21 – 24, 2003.

[23] K. Li, X. Wu, D. Z. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images - a graph-theoretic approach," *IEEE Trans On Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 119 – 134, 2006.

[24] M. K. Garvin, A. MD, and X. Wu, "Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomograhpy images," *IEEE Transaction on Medical Imaging*, vol. 28, pp. 1436 – 1447, 2009.

[25] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Transaction on Medical Imaging*, vol. 23, no. 4, pp. 501 – 509, 2004.

[26] M. J. Kupersmith, P. A. Sibony, S. E. Feldon, J.-K. Wang, M. Garvin, R. Kardon, and The OCT Substudy Group for the Nordic IIHTT, "The effect of treatment of idiopathic intracranial hypertension on prevalence of retinal and choroidal folds," *American Journal of Ophthalmology*, vol. 176, pp. 77–86, 2017.

[27] P. A. Sibony, M. J. Kupersmith, S. E. Feldon, J.-K. Wang, M. Garvin, and The OCT Substudy Group for the NORDIC IIHTT, "Retinal and choroidal folds in papilledema," *IOVS*, vol. 56, pp. 5670–5680, 2015.

[28] J.-K. Wang, R. H. Kardon, M. J. Kupersmith, and M. K. Garvin, "Automated quantification of volumetric optic disc swelling in papilledema using spectral-domain optical coherence tomography," *IOVS*, vol. 53, no. 7, pp. 4069 – 4075, June 2012.

[29] N. Batool and R. Chellappa, "Fast detection of facial wrinkles based on Gabor features using image morphology and geometric constraints," *Pattern Recognition*, vol. 48, no. 3, pp. 642–658, 2015.

[30] Z. Püspöki, M. Storath, D. Sage, and M. Unser, "Transforms and operators for directional bioimage analysis: A survey," in *Focus on Bio-Image Informatics*, ser. Advances in Anatomy, Embryology and Cell Biology, W. De Vos, S. Munck, and J.-P. Timmermans, Eds. Springer International Publishing, 2016, vol. 219, ch. 3, pp. 69–93.

[31] R. Rezakhaniha, A. Agianniotis, J. Schrauwen, A. Griffa, D. Sage, C. Bouten, F. van de Vosse, M. Unser, and N. Stergiopulos, "Experimental investigation of collagen waviness and orientation in the arterial adventitia using confocal laser scanning microscopy," *Biomechanics and Modeling in Mechanobiology*, vol. 11, no. 3-4, pp. 461–73, 2012.

[32] J.-K. Wang, R. H. Kardon, and M. K. Garvin, "Automated bruchs membrane opening segmentation in cases of optic disc swelling in combined 2D and 3D sd-oct images using shape-prior and texture information," *Proceedings of the Ophthalmic Medical Image Analysis Second International Workshop, OMIA MICCAI, Munich, Germany*, pp. 33 – 40, October 2015.

[33] K. L. Malaika and M. Kurz-Levin, "A comparison of imaging techniques for diagnosing drusen of the optic nerve head," *Arch Ophthalmol.*, vol. 117, pp. 1045 – 1049, August 1999.

[34] M. Mehrpour, F. O. Torshizi, S. Esmaeeli, S. Taghipour, and S. Abdollahi, "Optic nerve sonography in the diagnostic evaluation of pseudopapilledema and raised intracranial pressure: a cross-sectional study," *Neurology Research International*, vol. 2015, p. 146059, January 2015.

[35] S. A. Mayar and J. Y. Chong, "Critical care management of increased intracranial pressure," *Journal of Intensive Care Medicine*, vol. 17, no. 2, pp. 55–67, 2002.

[36] S. B. Carter, M. Pistilli, K. G. Livingston, D. R. Gold, N. J. Volpe, K. S. Shindler, G. T. Liu, and M. a Tamhankar, "The role of orbital ultrasonography in distinguishing papilledema from pseudopapilledema," *Eye*, vol. 28, no. 12, pp. 1425 – 30, December 2014.

[37] G. Lascaratos, S. Ahmed, and S. A. Madill, "Spontaneous venous pulsation and its role in differentiating papilledema from pseudopapilledema," *Neurology*, vol. 75, no. 13, pp. 53 – 55, September 2010.

[38] J. Agne, J.-K. Wang, R. H. Kardon, and M. K. Garvin, "Monitoring intracranial pressure by a semi-automated determination of bruch's membrane deformation from routine sd-oct volume scans of the optic disc," *IOVS E-abstract 4544*, May 2016.

[39] Z. Hu, M. Neimeijer, M. D. Abramoff, and M. K. Garvin, "Multimodal retinal vessel segmentation from spectral-domain optical coherence tomography and fundus photography," *IEEE Transaction on Medical Imaging*, vol. 31, no. 10, pp. 1900–11, 2012.

[40] Z. Hu, M. D. Abramoff, Y. H. Kwon, K. Lee, and M. K. Garvin, "Automated segmentation of neural canal opening and optic cup in 3D spectral optical coherence tomography volumes of the optic nerve head," *IOVS*, vol. 51, pp. 5708 – 5717, 2017.

[41] J. Agne, J.-K. Wang, R. H. Kardon, and M. K. Garvin, "Determining degree of optic nerve edema from color fundus photography," *SPIE Proceedings, Medical Imaging: Computer-Aided Diagnosis*, March 2015.

[42] S. Perutz, D. Puig, and M. Angel, "Analysis of focus measure operators for shape-from-focus," *Pattern Recognition*, vol. 46, pp. 1415 – 1432, 2013.