Spring 2018

Differentiation between causes of optic disc swelling using retinal layer shape features

John William Miller
University of Iowa

DIFFERENTIATION BETWEEN CAUSES OF OPTIC DISC SWELLING USING RETINAL LAYER SHAPE FEATURES

by

John William Miller

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

May 2018

Thesis Supervisor: Associate Professor Mona K. Garvin

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

MASTER'S THESIS

This is to certify that the Master's thesis of

John William Miller

has been approved by the Examining Committee for the thesis requirement for the Master of Science degree in Electrical and Computer Engineering at the May 2018 graduation.

Thesis Committee: _____
Mona K. Garvin, Thesis Supervisor

_____
Mathews Jacob

_____
Hans Johnson

ACKNOWLEDGEMENTS

I'd like to begin by thanking my advisor, Professor Mona Garvin. Your enthusiasm for research and guidance while I worked in your lab and developed my thesis motivated me and kept me focused on my goal. To my friends in the Garvin lab, Ray, Victor, Jason, Will, Shafkat, Yashila, Sam, and Qingyang, your collaboration, camaraderie, and sympathetic commiserating were a key part of what helped me through my two years of grad school. I'll miss our afternoons spent recuperating in the Seamans Center nap pods after our repeated overindulgence at India Cafe. Or was that just me? I especially want to thank Ray for his support – both technical and emotional – during my time in the lab and in particular for his deeply insightful comments as I struggled through drafts of this thesis. I would not be where I am now were it not for you, Ray. This work was supported, in part, by the following research grants: I01 RX001786, R01 EY023279, U10 EY017281-01A1, U10 EY017387-01A1, 3U10EY01728101A1S1.

Thank you also to Professors Jacob and Johnson for serving on my thesis committee and taking the time to have a number of insightful, engaging conversations with me while I've been in grad school. Professor Jacob, the courses I've had with you have been superb. You've inspired in me a new enthusiasm for and confidence in exploring mathematical descriptions of the world. To Dr. Kardon, I'd like to extend my enthusiastic thanks for offering your expert clinical advice and research perspective as I worked on this thesis. Cathy Kern, thank you for helping keep me sane whenever I would panic over whether I'd actually taken all the courses needed for graduation. I have, right?

I want to also thank Dr. Chandan Reddy and the members of the Howard Lab, where I worked for two years before starting my masters. You threw me into the deep end of the research pool and gave me just enough support to learn to swim on my own. I would not have been nearly as well-equipped for this program as I was had I

not spent the time working with you all.

I'd like to express my particular thanks and gratitude to Dr. Scott Liu and the incredible staff of the University Counseling Service. The UCS provides invaluable, enriching services that deserve tremendous recognition and support.

Shouts out are due to my Monday night Mosley's trivia team, which, to maintain a semblance of decency, will remain unnamed. You've made me feel so welcomed, and I cherish the time I've spent with you all. To Sketch House and my wider Goshen community, thank you for your support! The countless Skypes, conversations, and random Snapchats helped more than I can adequately express. Jonathan, my housemate and conversationalist, thank you for putting up with my dirty dishes in the kitchen and broken bike in the living room. Maybe now I'll finally get around to fixing that tire.

Finally, and most importantly, I want to acknowledge the unending love and support I've received from my family. My Fisher and Miller grandparents have offered their encouragement, praise, and sincere interest in my education for as long as I've been in school and likely longer than I can remember. Mom, Dad, Peter, Leah, and Chris, thank you for your generous advice, perspectives, and humor that has kept me afloat during grad school. And a special thanks to James for your unending supply of cute baby pictures. I can't wait until you're old enough to tell me what you think of this thesis!

ABSTRACT

The optic disc is the region of the retina where the optic nerve exits the back of the eye. A number of conditions can cause the optic disc to swell. Papilledema, optic disc swelling caused by raised intracranial pressure (ICP), and nonarteritic anterior ischemic optic neuropathy (NAION), swelling caused by reduced blood flow to the back of the eye, are two such conditions. Rapid, accurate diagnosis of the cause of disc swelling is important, as with papilledema the underlying cause of raised ICP could potentially be life-threatening and may require immediate intervention.

The current clinical standard for diagnosing and assessing papilledema is a subjective measure based on qualitative inferences drawn from fundus images. Even with the expert training required to properly perform the assessment, measurements and results can vary significantly between clinicians. As such, the need for a rapid, accurate diagnostic tool for optic disc swelling is clear.

Shape analysis of the structures of the retina has emerged as a promising quantitative tool for distinguishing between causes of optic disc swelling. Optic disc swelling can cause the retinal surfaces to distort, taking on shapes that differ from their normal arrangement. Recent work has examined how changes in the shape of one of these surfaces, Bruch's membrane (BM), varies between different types of optic disc swelling, containing clinically-relevant information.

The inner limiting membrane (ILM), the most anterior retinal surface and furthest from BM, can take on shapes that are distinct from the more posterior layers when the optic disc becomes swollen. These unique shape characteristics have yet to be explored for their potential clinical utility. This thesis develops new shape models of the ILM.

The ultimate goal of this work is to develop noninvasive, automated diagnostic tools for clinical use. To that end, a necessary first step in establishing clinical relevance is demonstrating the utility of retinal shape information in a machine learning

classifier. Retinal layer shape information and regional volume measurements acquired from spectral-domain optical coherence tomography scans from 78 patients (39 papilledema, 39 NAION) was used to train random forest classifiers to distinguish between cases of papilledema and NAION.

On average, the classifiers were able to correctly distinguish between papilledema and NAION $85.7 \pm 2.0\%$ of the time, confirming the usefulness of retinal layer shapes for determining the cause of optic disc swelling. The results of this experiment are encouraging for future studies that will include more patients and attempt to differentiate between additional causes of optic disc edema.

PUBLIC ABSTRACT

The optic disc is the region of the retina where the optic nerve exits the back of the eye. A number of conditions can cause the optic disc to swell. Papilledema, optic disc swelling caused by raised intracranial pressure (ICP), and nonarteritic anterior ischemic optic neuropathy (NAION), swelling caused by reduced blood flow to the back of the eye, are two such conditions. Rapid, accurate diagnosis of the cause of disc swelling is important, as with papilledema the underlying cause of raised ICP could potentially be life-threatening and may require immediate intervention.

Current clinical methods for determining the cause of optic disc swelling rely on techniques that are either subjective and vary between clinicians or are invasive and painful for the patient. A noninvasive, computer-based method for determining the cause of optic disc edema could save doctors and patients time, minimize patient discomfort, and provide an objective measure of confidence when deciding on possible treatments.

Shape analysis of the retinal layers has emerged as a promising quantitative tool for distinguishing between types of optic disc swelling. Optic disc swelling can cause the retinal layers to distort, taking on shapes that differ from their normal arrangement. Recent work has examined how changes in the shape of one of these layers, Bruch's membrane, varies between different types of optic disc swelling, containing useful, clinically-relevant information. This thesis introduces new models of a different retinal layer, the inner limiting membrane, that has previously been unexplored for clinically relevant characteristics.

Retinal layer shape information acquired from spectral-domain optical coherence tomography (SD-OCT) scans from 78 patients (39 papilledema, 39 NAION) was used to train random forest classifiers to distinguish between papilledema and NAION. On average, the classifiers were able to correctly identify a patient as having papilledema or NAION $85.7 \pm 2.0\%$ of the time, confirming the usefulness of retinal layer shapes for determining the cause of optic disc swelling. The results of this experiment are

encouraging for future studies that will include more patients and attempt to differentiate between additional causes of optic disc edema.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

CHAPTER 1
INTRODUCTION

The optic disc is the region of the retina where the optic nerve exits the back of the eye. There are various conditions that can cause the optic disc to swell, referred to generally as optic disc edema. One type of optic disc edema is papilledema, defined to be any swelling of the optic disc due to raised intracranial pressure (ICP), a condition which can be indicative of a serious underlying neurological disorder [2]. Another type of optic disc edema is nonarteritic ischemic optic neuropathy (NAION), in which swelling is caused by reduced blood flow to the anterior portion of the optic nerve [3].

Currently, clinicians have limited options for diagnosing optic disc edema. The most common technique ophthalmologists use for assessing papilledema is the Frisén scale, a subjective, ordinal measurement based on inferences drawn from fundus images [4]. Even amongst neuro-ophthalmologists, clinicians with the expert training required to perform the exam, there can be considerable variation in its results [5]. When a more immediate or definitive diagnosis is needed, clinicians turn to invasive methods (i.e. lumbar puncture) to measure ICP and/or expensive imaging tests such as magnetic resonance imaging (MRI) or magnetic resonance venography (MRV) to check for common causes of raised ICP (e.g. tumor) [6, 7].

The invasive procedures are uncomfortable for the patient, time-intensive for the physician, and carry the risk of more serious complications [8]. A noninvasive, automated diagnostic tool could minimize patient discomfort and increase physician efficiency as well as provide a more robust method for determining the cause of optic disc swelling. Although not life-threatening, NAION almost always leads to permanent vision loss [9]. If raised ICP is the true cause of edema (i.e. papilledema) and goes unnoticed (or the swelling is incorrectly diagnosed as something other than papilledema), the patient may be at risk for developing more serious neurological conditions or even death if the underlying cause of raised ICP is not immediately

addressed [10]. As determining the true cause of disc edema is necessary for proper clinical management – and in some cases is highly time-sensitive – it is important to develop algorithmic techniques for differentiation that are both fast and reliable. The need for an accurate and rapid diagnosis of optic disc edema is clear.

Retinal layer shape analysis is a relatively new area of research that shows promise as a quantitative diagnostic tool. The retina is divided into several layers whose shape can become distorted in the presence of optic disc edema. Recent research has shown that these changes in shape can in fact be characteristic to the disorder causing the swelling [11–14]. Shape analysis could potentially provide an extra dimension of information for clinicians trying to determine the cause of optic disc swelling: while both papilledema and NAION cause the optic disc to swell and result in larger-than-normal volumetric measurements, there are differences in retinal layer shape between the two cases that cannot be accounted for through volume measurements alone [11, 12].

For example, several recent studies have demonstrated that for subjects with papilledema, when observed through optical coherence tomography (OCT) images, Bruch's membrane (BM), the bottommost surface of the retinal pigment epithelium (RPE), often distorts towards the vitreous in an inverted-U shape [11–14]. Conversely, while volume measures of the optic disc are also elevated (i.e. the disc is swollen) with NAION, the shape of BM remains more similar to normal, unswollen optic discs [11, 12, 15]. Clearly, shape information can be useful for differentiating between causes of optic disc edema.

Up to this point, shape analysis has only been performed on BM. In cases of extreme swelling, the inner limiting membrane (ILM), the retinal structure closest to vitreous, can distort into shapes distinct from BM and may contain useful information currently unexplored by researchers and inaccessible to clinicians. For example, with NAION, ophthalmologists are aware of a particular pattern of regional swelling and

collapsing of the optic disc that changes over time as the disease progresses [16]. While BM is likely too posterior to reflect these changes, tracking the ILM shape over time could potentially allow clinicians to determine how far the disease has progressed. Furthermore, there may be distortions in the ILM shape in cases of papilledema that, like the inverted-U observed with BM, may help distinguish the condition from other types of optic disc edema. To our knowledge, this thesis represents the first development of shape models of the ILM for use in distinguishing between different types of optic disc edema.

These novel ILM shape models, along with previously-developed models of BM and regional volumetric measures, will be used as features in a machine learning classifier trained to distinguish between causes of optic disc swelling. By training a classifier on shape information and then checking which features are important to its performance, it is possible to gain a more quantitative understanding of how different shapes may be more or less clinically relevant. The demonstration of this classifier is a necessary first step in the development of automated ophthalmic diagnostic tools.

In summary, the main goal of this work is to develop an automated diagnostic tool trained on newly-developed shape models of the ILM along with previously-established BM models and regional volumetric measures capable of reliably distinguishing between different causes of optic disc swelling.

1.1 Main contributions

The two main contributions of this thesis are the development of shape models of the inner limiting membrane and the use of retinal layer shape features in a machine learning classifier trained to distinguish between two types of optic disc edema.

1.2 Thesis overview

The remainder of this thesis is organized as follows:

- **Chapter 2** provides a clinical background on optic disc edema and summarizes

relevant prior technical work in the field of ophthalmic imaging and analysis.

- **Chapter 3** describes the theory and technical methods involved in the implementation of the inner limiting membrane shape model.
- **Chapter 4** describes the theory and implementation of the classifier used to differentiate between causes of optic disc edema.
- **Chapter 5** offers a discussion on the material presented here and describes potential avenues for future work.

CHAPTER 2
CLINICAL AND TECHNICAL BACKGROUND

2.1 Clinical background

2.1.1 Terminology

There are several key anatomical terms used throughout this thesis that are listed for reference here:

- Optic nerve: transmits nerve impulses from the retina to the brain.
- Optic nerve head (ONH): region where the optic nerve exits the back of the eye
- Vitreous: clear gel-filled space separating the eye's lens and retina
- Retinal pigment epithelium (RPE): outermost (posterior) structure of retina, comprised of multiple layers of pigmented hexagonal cells. Highlighted in green in Fig. 2.1 C.
- Bruch's membrane (BM): lower bounding surface of the RPE, most posterior layer of the retina.
- Inner limiting membrane (ILM): innermost (anterior) membrane of retina, separates the retina and vitreous. Highlighted in red in Fig. 2.1 C.

This thesis focuses primarily on the ILM and RPE/BM. It is not uncommon for the terms RPE and BM to be used interchangeably, although they are not synonymous. The RPE is comprised of multiple layers, of which BM is the lower bounding surface (i.e. most posterior). This thesis will primarily refer to BM, only using RPE when intentionally referring to the multilayer structure. Fig. 2.1 provides an overview of the layers of the retina.

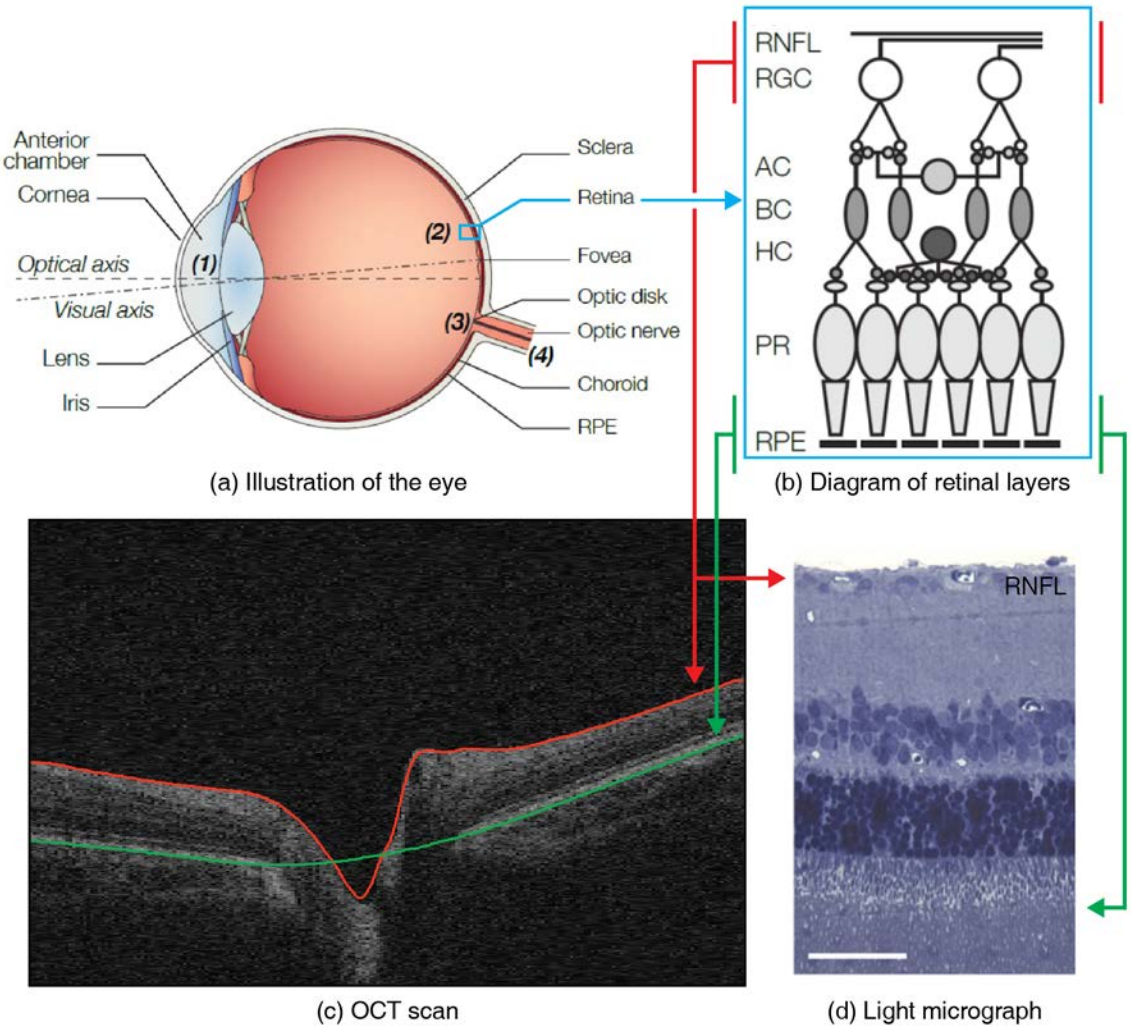


Figure 2.1: Layers of the retina. (a) The eye. (b) Diagram of the retinal layer cells. (c) OCT view of the optic nerve head. Highlighted surfaces are the ILM (red) and BM (green). (d) Light micrograph of the human retina. (Panels a, b, and d modified with permission from Daniel Palanker [1].)

2.1.2 Papilledema

Papilledema is swelling of the optic disc caused by raised intracranial pressure (ICP) [2]. Because the back of the eye is in a pressure equilibrium with the subarachnoid space [17], as ICP increases, the pressure against the back of the eye also increases, potentially distorting the optic disc [2, 17].

Any condition causing raised ICP, such as space-occupying lesions (e.g. tumors), cerebral edema, restricted cerebrospinal fluid (CSF) flow, or increased CSF production, can result in papilledema [2, 18]. By definition, papilledema cannot be present without raised ICP, but a patient may have raised ICP without papilledema. Additionally, the cause of raised ICP may be unknown – a condition referred to as idiopathic intracranial hypertension (IIH). IIH occurs in 1 in 100,000 patients for the general population, but occurs most often in obese women of child-bearing age, with a rate of 19.3 per 100,000 [19].

Papilledema typically presents along with other common symptoms of raised ICP including headaches, nausea, and pulsatile tinnitus [19, 20]. Some amount of vision loss is seen in 86% of patients, with 10% experiencing severe visual impairment [19]. Two common imaging techniques ophthalmologists employ for observing papilledema are fundus photography and optical coherence tomography (OCT) imaging. OCT is described in detail in Section 2.2.1, but for now it is sufficient to understand OCT as a noninvasive imaging modality that produces high-definition 3D scans of the retina, enabling clinicians to view details of the retina (e.g. retinal layers) in much greater detail than with an ophthalmoscope or fundus photograph. Fig. 2.2(b) shows an OCT image from a papilledema patient with a noticeably swollen optic disc.

As papilledema is a symptom of raised ICP, treatment for papilledema depends entirely on the underlying cause of increased pressure. In cases of IIH, change of diet and weight loss combined with a daily diuretic (e.g. acetazolemide) has been shown to be effective [18, 21]. While most papilledema patients do not experience long-

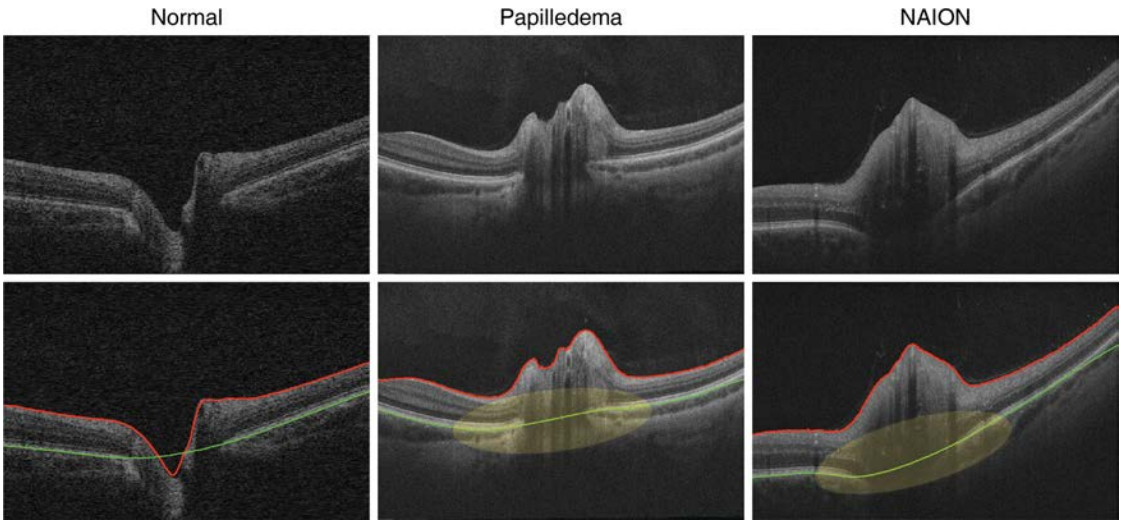


Figure 2.2: Comparison of optic disc swelling in papilledema and NAION. Highlighted surfaces are the ILM (red) and BM (green). Yellow regions highlight the difference in BM shape between the two types of swelling. A normal subject with a non-swollen optic disc is shown at left for comparison.

term severe vision loss, those who do may undergo an optic nerve sheath fenestration surgery to protect the optic nerve from further damage [18]. Cases of raised ICP due to a space-occupying mass such as a tumor will likely require surgical intervention [18].

As papilledema is a result of increased ICP, a lumbar puncture (i.e. spinal tap) to measure CSF pressure is the definitive means for diagnosing papilledema and ruling out other types of optic disc edema [6, 17]. A lumbar CSF opening pressure greater than 250 mm of water is considered raised ICP and is required for a diagnosis of IIH [22]. Due to the patient discomfort, length of time, and potential for further complications associated with lumbar punctures [8], noninvasive techniques for measuring ICP – and more specifically assessing papilledema – are desired. Because of the potential seriousness of conditions causing raised ICP, it is important to quickly and reliably diagnose papilledema.

The current standard for assessing papilledema severity is the Frisén scale, an ordinal measurement based on descriptive features of fundus images [23]. However,

the method requires expert training (i.e. neuro-ophthalmologist), is limited in its descriptive ability by the nature of its ordinal scale [24], and can have high variability between clinicians [4, 5, 25].

These limitations of the Frisén scale combined with a desire for a more robust, objective method, coinciding with the commercial availability of OCT instruments, has led to the development of new quantitative methods for the assessment of papilledema severity. Previous OCT studies have shown thickening of the retinal nerve fiber layer (RNFL) in patients with papilledema [4,6,26,27]. Additionally, as layer segmentation techniques have become more capable of handling cases of extreme optic disc swelling, researchers have been able to show a roughly linear relationship between volumetric OCT measurements [24] and a modified Frisén scale [4], an encouraging step towards more objective measurements of papilledema.

2.1.3 Nonarteritic anterior ischemic optic neuropathy

Anterior ischemic optic neuropathy (AION) is neuropathy resulting from reduced blood flow (i.e. ischemia) to the anterior portion of the optic nerve [3, 28, 29]. Of the two types of AION – arteritic (AAION) and non-arteritic (NAION) – the non-arteritic form is most common, accounting for 95% of all reported cases of AION and is the most common form of acute optic neuropathy for people over the age of 50 [30]. Estimates of incident rates range from approximately 2 to 10 individuals per 100,000 [28, 31].

Symptoms of NAION include disc edema and painless vision loss with an onset of hours to days [29]. While NAION can affect patients of any age, it more often occurs in patients over the age of 50, with an approximate median age of incidence of 62 years [29, 31]. NAION most often presents unilaterally, although in some patients it does progress to the other eye [32]. Bilateral cases of NAION present the most difficulty in distinguishing from papilledema. Fig. 2.2(c) shows an OCT image from an NAION patient with a swollen optic disc.

As its pathophysiology is still not fully understood, there is no generally agreed upon treatment for NAION [3]. Current medical therapies include blood thinners (e.g. aspirin) to address thrombosis and vasodilators [3, 29]. The efficacy of surgical interventions such as optic nerve decompression is questionable [29]. As summarized by Atkins et al. in 2009: "NAION remains frustrating for clinicians and often devastating for patients. The pathophysiology remains unclear, and it is uncertain whether any treatment will be effective for NAION" [29].

2.2 Technical background

2.2.1 Optical coherence tomography

Optical coherence tomography (OCT) is a noninvasive imaging modality developed by Huang et al. in 1991 [33] that captures high-resolution 3D images of the back of the eye, enabling clinicians and researchers to view retinal structures that were previously inaccessible. OCT operates through low-coherence interferometry, emitting broadband light into the eye and measuring the reflected response [33].

The first implementation of OCT worked in the time domain (TD-OCT) [33] and was commercially available beginning in 2003. OCT has more recently been adapted to work in the spectral domain (SD-OCT), allowing for higher resolution images and faster sampling times [34], becoming commercially available around 2007. The imaging protocols employed by OCT machines vary by manufacturer. The OCT images described in this work are all SD-OCT images collected using the Cirrus OCT device from Carl Zeiss Meditec, Inc. Dublin, CA). Two types of scanning protocols, volumetric and five-line raster (5LR), were followed for the data described in this work.

The first, volumetric scanning, takes 200 slices, each with a dimension of 200×1024 pixels to produce a 3D image of the time spanning a physical volume of (6 mm $\times$ 6 mm $\times$ 2 mm). The second protocol, high-definition five-line raster (HD-5LR), samples only five higher-resolution slices to produce a 2D image of the eye. Fig. 2.3 details

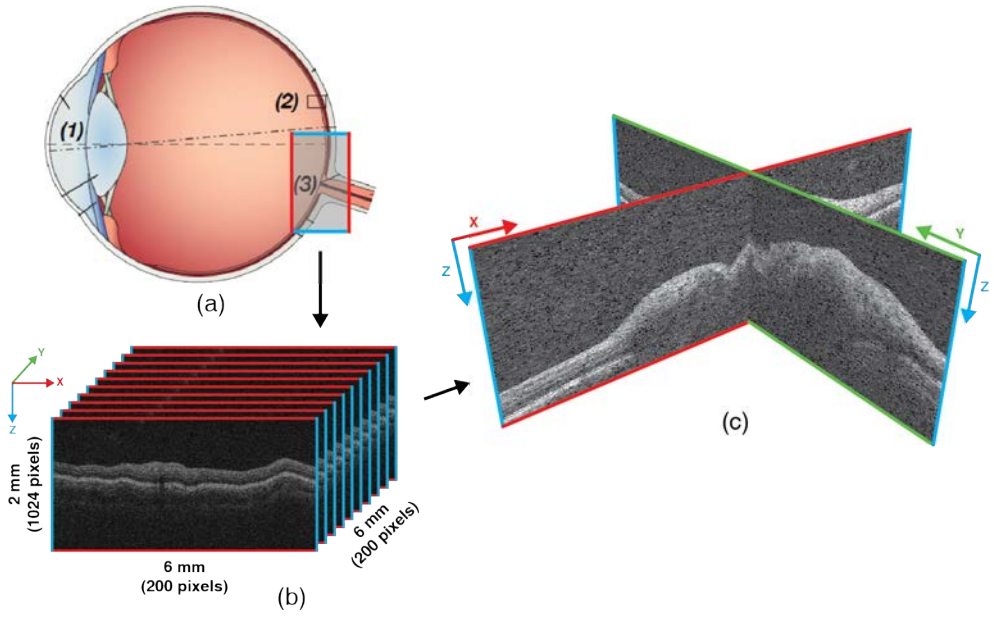


Figure 2.3: Orientation and dimensions of SD-OCT images. (a) Diagram of the eye with approximate scan region highlighted (modified with permission from Palanker [1]). (b) A 3D image compiled from multiple OCT image slices. (c) Two 2D perpendicular scans, oriented in 3D space from a patient with papilledema.

the orientation of OCT scans in relation to the back of the eye (a) and details the slice arrangement in a typical volumetric scanning protocol (b, c).

2.2.2 Graph-based segmentation of retinal layers

The main contribution of this thesis (shape analysis of the ILM) is made possible through graph-based segmentation of the retinal layers. The segmentation methods employed in this thesis originated with Li et al.'s development of a fully 3D, optimal surface segmentation technique in 2006 [35]. Their work was then adapted by Garvin et al. in 2009 to optimally segment retinal layers in 3D OCT scans [36]. The technique was adapted by Lee et al. in 2010 to include a multiresolution approach that increased efficiency [37]. In 2012, Wang et al. adapted the technique to handle images of swollen optic discs by designing cost images specific to scans of papilledema [24]. The group further developed their methods in 2013 to calculate regional volume information [25]. The bottom panels of Fig. 2.2 show examples of segmentations of the ILM (red) and BM (green).

2.2.3 Shape analysis

Accurate segmentation of retinal layers has made possible the quantitative analysis of retinal layer shapes, an area of study which in recent years has shown to be useful for differentiating causes of optic disc swelling [11–15]. The shape analysis techniques used in this thesis are derived from methods originally developed by Cootes et al. in 1995 called Active Shape Models (ASMs) [38]. The ASM technique derives a statistical model of class-specific (e.g. faces, hands, etc.) shapes from images containing manually-labeled examples from that class. The statistical shape model is then used to develop an understanding of typical modes of variation seen within that class of objects. For example, the first mode of variation of a statistical shape model trained on faces may reflect the scale of the mouth, whereas the second mode may represent the angle of the eyebrows [39]. Full details on the theory and implementation of shape

models will be given in Chapter 3.

Some of the first work to examine the effects of papilledema and elevated ICP on retinal structure [11] and shape [12] using high-resolution SD-OCT images was published in 2011. Both studies employed a fully manual process and focused on the angulation of BM. Both studies showed that BM bowed in towards vitreous in cases of raised ICP (i.e. papilledema) and remained angled posteriorly in NAION and normal cases [11, 12]. The yellow ovals in the bottom images of Fig. 2.2 (b)-(c) highlight the difference in BM shape between papilledema and NAION patients. Fig. 2.2 (a) displays a normal patient scan without any swelling of the optic disc.

Addressing the limitations of a fully manual process and building on these first two studies, in 2015 our group developed a semi-automatic approach that makes possible the analysis of large data sets whose size would otherwise be prohibitive to work with [13]. Examples of shape variations of BM extracted from papilledema images using this semi-automatic approach are shown in Fig. 2.4. The second principal component demonstrates angulation of BM towards (-) and away from (+) vitreous, one of the key findings of retinal layer shape analysis to date [11–13].

More recently, retinal layer shape analysis was applied to data from the Idiopathic Intracranial Hypertension Treatment Trial (IIHTT), a longitudinal study of IIH [20], and used to quantify the affects of medical treatment for IIH on layer shapes over time [14]. Retinal layer shape has also been shown to correlate with ICP and change within an hour after lumbar puncture [40], providing further evidence for the relationship between ICP and retinal shapes.

2.2.4 Random forest classifier

The ultimate goal of this thesis is to use retinal layer shape information to distinguish between causes of optic disc edema. Given the limited amount of data available in this work, a random forest classifier was selected in order to minimize overfitting. Additionally, random forests provide an intuitive measure of feature importance, a

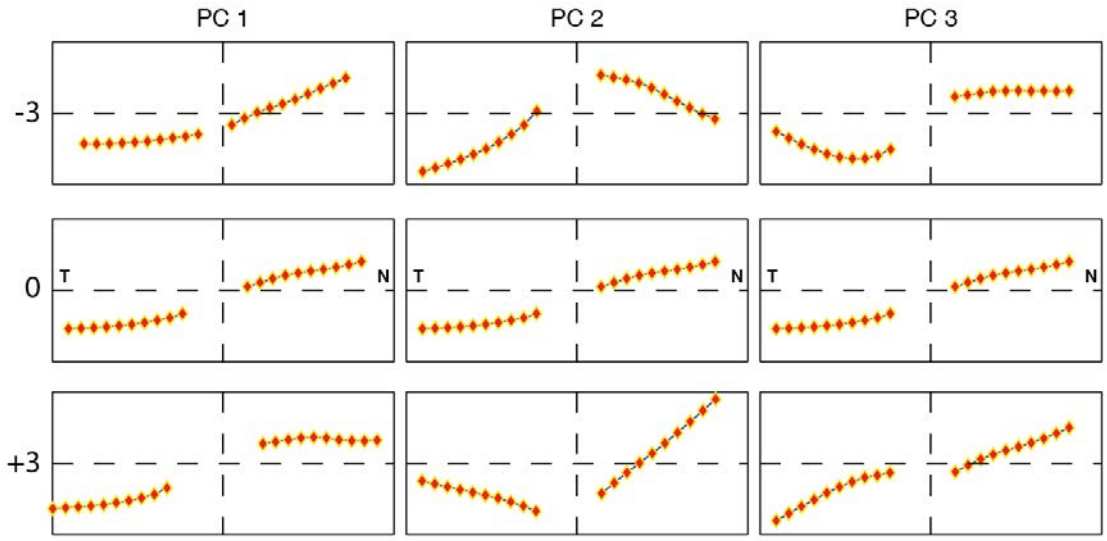


Figure 2.4: Example of variations in shape of BM. The second principal component (PC) demonstrates a bending of BM towards (-) and away from (+) vitreous. The -3 and +3 correspond to weightings of the eigenvectors added to the mean shape (middle row), the details of which are explained in Chapter 3. Indicated directions are temporal (T) and nasal (N).

desirable trait when evaluating the utility of retinal layer shapes for clinical relevance. The full details of the random forest classifiers used in this work will be given in Chapter 4.

The use of random forests in ophthalmic image analysis is well-established. The technique has been used to generate cost images for layer segmentation [41], predict optic nerve head (ONH) volume from fundus photographs [42], and estimate papilledema severity using retinal regional volumes and layer shapes [43]. This thesis demonstrates one of the first [44] uses of retinal layer shapes as features in a classifier for determining the cause of optic disc swelling.

CHAPTER 3
SHAPE ANALYSIS OF THE INNER LIMITING MEMBRANE

This chapter provides a detailed mathematical explanation of shape models, describes the implementation of shape models of the inner limiting membrane (ILM), and details the characteristics of the trained ILM models. The novel contribution of this thesis is the application of shape analysis to the ILM. Prior to this work, retinal shape models had only ever been built from the retinal pigment epithelium/Bruch's membrane (RPE/BM) [12–14], the posterior-most portion of the retina. The ILM is the retinal structure closest to vitreous, and in cases of optic disc swelling, can distort into shapes different from the more posterior layers, which if analyzed, may provide currently unused clinical information.

3.1 Active shape model theory

The goal of shape analysis is to create a statistical model that is capable of describing the typical variations in shape seen across a training set of images for a single class of objects. For example, the shape models trained on retinal layers reflect the ways in which the layers bend and contort throughout various stages of swelling.

The process of creating a shape model can be split into three main components: assigning landmarks to objects in images, aligning shapes with Procrustes analysis, and reducing the dimensionality of the aligned shapes using principal component analysis (PCA). The notation in this section follows the original work of Cootes et al. [38].

3.1.1 Assigning landmarks

Shape analysis starts with either manual or automatic labeling of objects within a training set of images using user-defined landmark points. Landmarks are typically selected as key points that are common across objects of the same class. For example, in a training set of faces, landmark points would be assigned along the eyes, nose, lips,

etc. In retinal layer shape analysis, landmarks are placed along the retinal surface.

Shapes are represented as a sequence of these landmark points. In 2D, a given shape, $\mathbf{q}_i$, is represented as a column vector consisting of (x, y) pairs for each of the n landmark points:

$$\mathbf{q}_i = (x_{i0}, y_{i0}, x_{i1}, y_{i1}, \ldots, x_{in-1}, y_{in-1})^T \tag{3.1}$$

where $\mathbf{q}_i$ is the ith shape in a training set of N images. The math for 3D shapes is essentially the same, adding a z value where appropriate.

3.1.2 Aligning shapes

After each shape in the training set of images is labeled with n landmarks, the shapes must be aligned to ensure consistency across points when creating the statistical model described in Section 3.1.3. The typical method used for shape alignment is Procrustes analysis. Procrustes analysis is an iterative technique that typically uses translation, rotation, and scaling to minimize the sum of squared error of the distance between landmark points shared across a set of shapes [38]. Procrustes analysis can also be performed without all three of the transformations (e.g. leaving out scaling, as was done in this thesis).

Procrustes analysis essentially removes differences in translation, rotation, and scaling (referred to as "pose parameters") from images in the training set, aligning each shape to the group's mean shape. By removing the pose parameters (which are linear in the image space) from the training shapes before calculating the statistical shape model, the eventual model is able to more effectively represent changes in the object's shape, which, in the image space, are typically nonlinear.

3.1.3 Creating the statistical shape model

The third and final step in the process of creating a shape model is to use PCA to reduce the dimensionality from $2n$-dimensional shape space to a t-dimensional model space, where $t \leq 2n$ and is determined by the number of principal components (PCs) used in the final shape model. In practice, t is chosen to be just large enough to accurately approximate the original shapes and is typically much less than $2n$ [38]. As PCA is a standard statistical technique for which numerous resources are readily available [45, 46], the technique will be described here only in brief, focusing on its relevance to shape models.

PCA identifies the dimensions of most variation in a dataset, which, in shape analysis, correspond to variations in shape that are shared between images in the training data. Returning to the example of face shapes from before, PCA may identify the angle of the eyebrows or length of the ears as variations in shape shared across the training images [39]. Or, as is seen in shape analysis of the retina, shape variations can correspond to the angle of the retinal layers or width of the Bruch's membrane opening.

PCA begins with calculating the mean shape, $\bar{\mathbf{x}}$, of all N aligned shapes, $\mathbf{x}_i \ldots \mathbf{x}_N$:

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i. \tag{3.2}$$

The $2n \times 2n$ covariance matrix, $\mathbf{S}$, is calculated as

$$\mathbf{S} = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T. \tag{3.3}$$

When plotted in shape space, where each landmark is an axis, the training shapes are assumed to form a Gaussian hyperellipsoid. The eigenvectors of the covariance matrix (i.e. PCs), $\mathbf{p}_k$, correspond to the major axes of this distribution. The PCs are

calculated as

$$\mathbf{S}\mathbf{p}_k = \lambda_k \mathbf{p}_k \tag{3.4}$$

where λ_k is the kth eigenvalue of S, k ranges from 1 to $2n$, and $\lambda_k \geq \lambda_{k+1}$.

3.1.4 Representing new shapes

After calculating the statistical shape model, new shapes can be represented as weighted linear combinations of the mean shape with the first t PCs:

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b} \tag{3.5}$$

where $\mathbf{P}$ is comprised of the first t eigenvectors

$$\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_t)$$

and b is a vector of t weights

$$\mathbf{b} = (b_1, b_2, \ldots, b_t)^T \tag{3.6}$$

specific to each new shape.

To ensure that shapes remain within a reasonable range of deformation, each weight is constrained to a range of values determined by the corresponding PC's eigenvalue:

$$-3\sqrt{\lambda_k} \leq b_k \leq 3\sqrt{\lambda_k}. \tag{3.7}$$

These shape weight coefficients typically range from about -3 to 3, as most shapes within a training set will be contained within three standard deviations of the mean.

Once the shape model has been trained, new shapes can be represented as weighted linear combinations of the model's PCs. The model weights for a new shape $\mathbf{q}_i$ can be determined by rearranging Eq. 3.5:

$$\mathbf{b}_i = \mathbf{P}^T(\mathbf{x}_i - \bar{\mathbf{x}}) \tag{3.8}$$

where $\mathbf{P}^T = \mathbf{P}^{-1}$, since the columns of $\mathbf{P}$ are orthonormal. Finding the model weights for a new shape is equivalent to projecting the shape from the image space into the shape model space defined by $\mathbf{P}$.

3.2 Implementation of the inner limiting membrane and Bruch's membrane shape models

This section describes the implementation of statistical shape models of the ILM and BM following the theory presented in the previous section. In general, the methods described here closely follow the techniques described in [13, 14] for the semi-automatic construction of BM shape models, adapted for the ILM. A brief overview of the shape model construction process is provided here.

Eight shape models were produced as part of this work, six of which were developed for the first time for this thesis. Three of the models consisted of landmarks assigned solely to the ILM, three were constructed from landmarks on BM, and the remaining two were combination models of both ILM and BM landmarks. Fig. 3.1 provides an overview of the models trained for this work.

Shape models are generally constructed in either two or three dimensions. Of the shape models in this work, six were 2D and two were 3D. The 2D models consisted of landmarks assigned to slices oriented in either the temporal/nasal or superior/inferior direction (Fig. 3.2b). The 3D models were comprised of landmarks in a radial arrangement spanning the entire retinal surface (Fig. 3.2a).

Two of the BM models described above were previously established elsewhere.

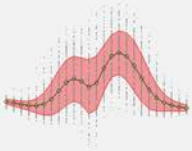
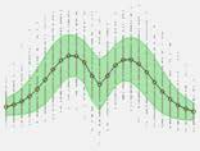
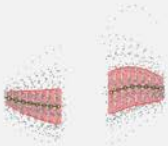
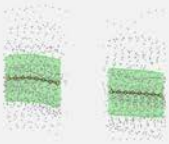
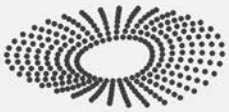
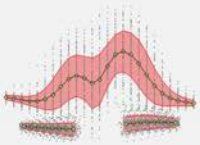
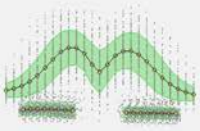
Landmarks	Retinal layer	Orientation	Number of landmarks	Number of subjects
	ILM	temporal/ nasal	25	116
	ILM	superior/ inferior	25	116
	ILM	3D	432	116
	BM	temporal/ nasal	16	116
	BM	superior/ inferior	16	116
	BM	3D	288	116
	ILM+BM	temporal/ nasal	41	116
	ILM+BM	superior/ inferior	41	116

Figure 3.1: Overview of the shape models created for this thesis. The three ILM models, the superior/inferior BM model, and the ILM+BM models are unique to this thesis. Orientation of model indicated in red (temporal/nasal) and green (superior/inferior). Mean shape shown in black. Transparent regions indicated one standard deviation in shape. Gray dots indicate landmarks from each shape. All models were built from the 116 IIHTT baseline OD scans.

The 2D BM model in the temporal/nasal orientation is the original retinal shape model [12], adapted for volumetric scans in 2016 [47]. The 3D BM model was recreated from [48].

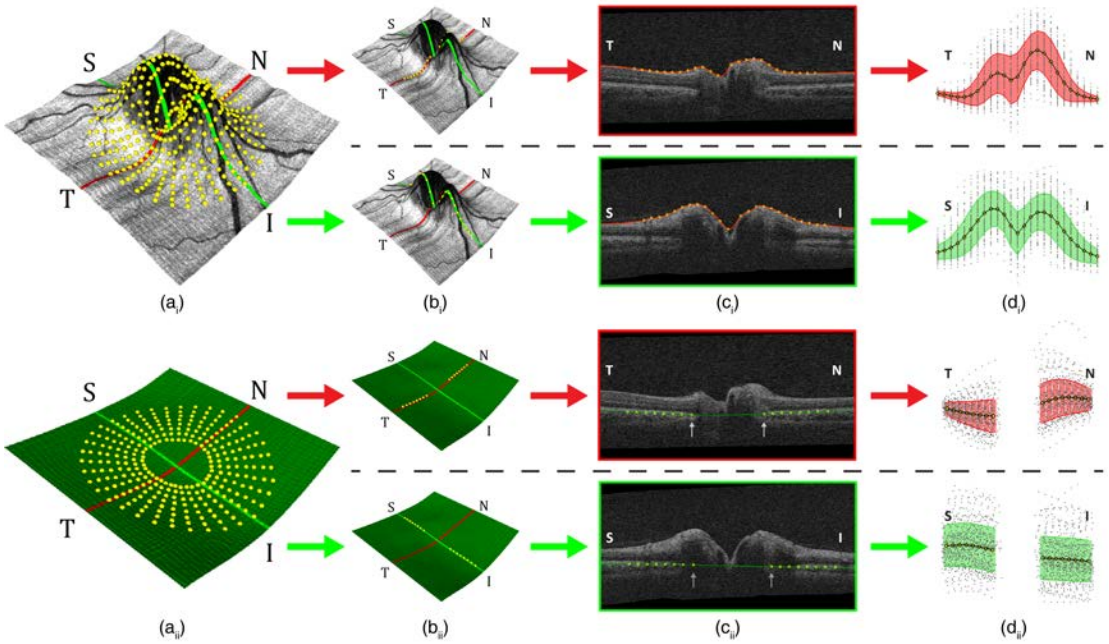


Figure 3.2: Construction flowchart for the ILM (i) and BM (ii) shape models. (a) Landmarks arranged in 3D on the retinal surface. (b) Landmarks isolated to the temporal/nasal (red) and superior/inferior (green) orientations. (c) Landmarks placed on central slices from an example SD-OCT volume. (ii) White arrows indicate BMO points (d) Layer shapes from all 116 subjects. Mean shape shown in black. Shaded areas depict one standard deviation.

3.2.1 Data acquisition: SD-OCT images

The data used to create the shape models in this work were drawn from the Idiopathic Intracranial Hypertension Treatment Trial (IIHTT), a multi-site, double-blinded trial to study the effectiveness of medical treatment for IIH [49]. HD-5LR and volumetric OCT scans were collected from 116 patients in the IIHTT as part of the OCT substudy [20].

All eight of the shape models were constructed from landmarks placed on volumetric SD-OCT scans. However, the segmentation process involved manually placing two points on an HD-5LR scan and transferring those points to the subject's volumetric scan, as will be described in Section 3.2.2. Each of the 116 scans were of the patients' right eyes, centered on the optic nerve head (ONH), and performed at their first clinical visit (i.e. baseline). The volumetric OCT scans had a scan dimension of $200 \times 200 \times 1024$ voxels, corresponding to a physical dimension of $6 \times 6 \times 2$ mm^3 [20] (Fig. 2.3). The HD-5LR scans consisted of five high-resolution slices (1024×1024 pixels, 9×2 mm^3) spaced at 0.25 mm intervals across the eye, of which the central (i.e. 3^{rd}) scan was used.

3.2.2 Segmentation

Both the HD-5LR and volumetric OCT scans were segmented using the graph-based techniques described in Section 2.2.2. The segmentation approach used here was a multistage process that began with a segmentation of the retinal nerve fiber layer (RNFL) and the RPE complex using a graph-algorithmic approach [36]. A multiresolution approach [35] was employed such that rough details were first segmented at low resolution and a finer level of detail was achieved at each higher resolution.

After the retinal layers were segmented, the Bruch's membrane opening (BMO) center was estimated using the semi-automatic approach presented in [50] and described here in brief. The GNU Image Manipulation Program (GIMP, version 2.8.10) was used to manually place two landmarks at the BMO points on the BM surface in the central B-scan of each of the 116 HD-5LR scans. The BMO points were then automatically transferred from the 5LR to the volumetric OCT scans by registering the two images together using an intensity-based registration. The BMO points are indicated with white arrows in Fig. 3.2c$_{ii}$. The BMO center was defined as the halfway point between the two manually-assigned landmarks in the central 5LR scan. The manually-labeled BMO points were then set as very low cost points in the radially-

unwrapped *en-face* projection image in order to aid the segmentation of the BMO. If an HD-5LR scan was not available for a subject, the BMO points were estimated from and placed directly on the volumetric SD-OCT image.

Variations in the image acquisition procedure can on occasion result in a tilting of the retinal image in the XY plane. To address this tilting, a standard procedure in OCT image processing to shift the vertical columns (A-scans) was followed so that the end points of BM both sat at a pre-assigned Z-value.

3.2.3 Landmark placement

Whether a model depicts the ILM or BM, both cases begin with the segmentation of an SD-OCT volumetric image as described in the previous section. In a basic sense, landmark placement is the process of deciding which points from the 200×200 surface produced from the segmentation will be included in the shape model. The 3D shape models consisted of landmark points placed in a radial arrangement spanning the entire retinal surface (Fig. 3.2a). The 2D models were built from landmarks placed only on the central temporal/nasal or superior/inferior slices of the surface (Fig. 3.2b). The next four subsections explain the landmark assignment in detail for each type of model.

3.2.3.1 Landmark placement for 3D models

The landmarks for the 3D BM model were assigned following the procedure described in [48]. 36 segments of eight landmarks were assigned to the BM surface in a radial arrangement, spaced evenly every 10° (Fig. 3.2a$_i i$). The first landmark in each segment was placed on the segmented BMO contour (Section 3.2.2) and the remaining seven landmarks were spaced evenly in the outward radial direction, not measuring distance along the retinal surface, spanning a total distance of 1.5 mm. A total of 288 landmarks were assigned to the BM surface for the 3D model.

A similar procedure was followed for the ILM. The 3D ILM model included 432

total landmarks, consisting of 36 radial segments separated by 10° each comprised of 12 landmarks. The 36 segments formed a circle centered at the BMO center point (Fig. 3.2a$_i$). The 12 landmarks of each segment spanned a distance of 2.5 mm and were spaced evenly in the radial direction, not measuring distance along the retinal surface.

3.2.3.2 Landmark placement for 2D ILM models

The temporal/nasal and superior/inferior 2D ILM models consisted of 25 landmarks spanning 6 mm, centered on the BMO center point (Fig. 3.2b$_i$). The landmarks were spaced evenly in the X (temporal/nasal) or Y (superior/inferior) orientation, as opposed to measuring distance along the retinal surface.

For some subjects (typically due to the 200×200 image not being perfectly centered on the ONH) the set of 25 landmarks would not fit within the width of the 6 mm scan. The approach of [47] was adapted to handle these cases; a thin plate spline was used to extend the ILM by 50 pixels (1.5 mm) at the temporal/nasal and superior/inferior edges to accommodate the landmarks.

3.2.3.3 Landmark placement for 2D BM models

The temporal/nasal and superior/inferior BM models were both extracted from the full 3D model of BM. The 2D temporal/nasal model consisted of the 16 landmarks from the 1^{st} and 19^{th} radial segments of the 3D BM model, and the superior/inferior model used the 16 from the 10^{th} and 28^{th} segments (Fig. 3.2b$_{ii}$).

3.2.3.4 Landmark placement for 2D ILM+BM models

The ILM+BM combination models were created by concatenating the 25 landmarks from the ILM and 16 landmarks from the BM into a single shape vector. The landmarks for the ILM and BM were assigned as described in the previous two sections, and then the vector containing the BM landmarks was concatenated onto the

end of the ILM vector.

3.2.4 Shape alignment and statistical model calculation

The same shape alignment procedure was followed for each of the eight shape models. The landmarks for each scan were translated so that the BMO center point was at the center of the 200×200 surface, i.e. shifted to $(100, 100)$ in the XY plane. As described in Section 3.2.2, the segmentation procedure for the volumetric OCT scans included a rotation adjustment to account for tilting of the retina. Therefore, it was not necessary to rotate the shapes after landmarks were placed. No changes in scaling were applied to the layer shapes. As the translation and rotation were taken care of through different means and no scaling adjustments were needed, Procrustes analysis was not performed on the retinal layer shapes. The statistical shape models were calculated using PCA implemented in MATLAB (r2014a).

3.3 Results: shape model variations

Six of the eight shape models described in the previous section are unique to this thesis. Five of the models included landmarks placed on the ILM, and the sixth was a previously undeveloped 2D model of BM, oriented in the superior/inferior direction. The results of the shape model implementations described in Section 3.2 are presented here in terms of the variations of their PCs.

Figs. 3.3 - 3.8 all follow the same format: columns correspond to the model's PCs, sorted in order of decreasing energy; rows display the shape variations corresponding to positive and negative sums of the PC; The middle row of each figure displays the model's mean shape. The individual and cumulative energy contributions of each PC are shown at the bottom of each column. Following Eq. 3.5, the shape variations (rows) in each figure are produced by adding or subtracting integer multiples $(1 - 3)$ of the PCs, $\mathbf{P}$, to the mean shape, $\bar{\mathbf{x}}$.

The rows of each column display shape variations three standard deviations from the mean shape (determined by Eq. 3.7). The first four PCs are shown for the 2D shape models. Fig. 3.5 shows the shape variations corresponding to the first three PCs of the 3D ILM model. Many of the shape variations are too complex for intuitive description. A textual description of each PC for each model would ultimately be of limited use. As such, only PCs with clear physical interpretations or utility as features in the classifier covered in Chapter 4's are described here. The reader is encouraged to examine the PC variations in each figure for themselves.

The first PC of the 2D ILM temporal/nasal model (Fig. 3.3) reflects a bulging of the ILM towards (-) and away from (+) vitreous while displaying a slightly asymmetry in the nasal direction. The second, third, and fourth PCs each represent the presence or absence and location of smaller bulges in the center of the surface. The first PC of the 2D ILM superior/inferior model (Fig. 3.4) also demonstrates a distortion of the surface towards and away from vitreous while maintaining the two central bulges.

The first PC of the 3D ILM model (Fig. 3.5) largely reflects the presence or absence of swelling of the optic disc. The second PC demonstrates a more central swelling (-) and dimpling (+) that appears to more heavily affect the temporal/inferior quadrant. The third PC demonstrates an asymmetry in the temporal/nasal directions.

The first PC of the 2D temporal/nasal ILM+BM model (Fig. 3.7) clearly reflects an asymmetry in the swelling of the ILM in the temporal/nasal orientation. Additionally, the first PC demonstrates an angulation of BM towards (-) and away from (+) vitreous, the common result of [11–13]. The third PC of the same model shows the same angulation of BM along with a clear asymmetry in the temporal/nasal direction.

The first PC of the 2D superior/inferior ILM+BM model (Fig. 3.8) reflects a general increase (-) and decrease (+) in swelling of the ILM while maintaining two bulges. The second PC maintains a consistent ILM height while shifting between one (-) and two (+) bulges in the surface.

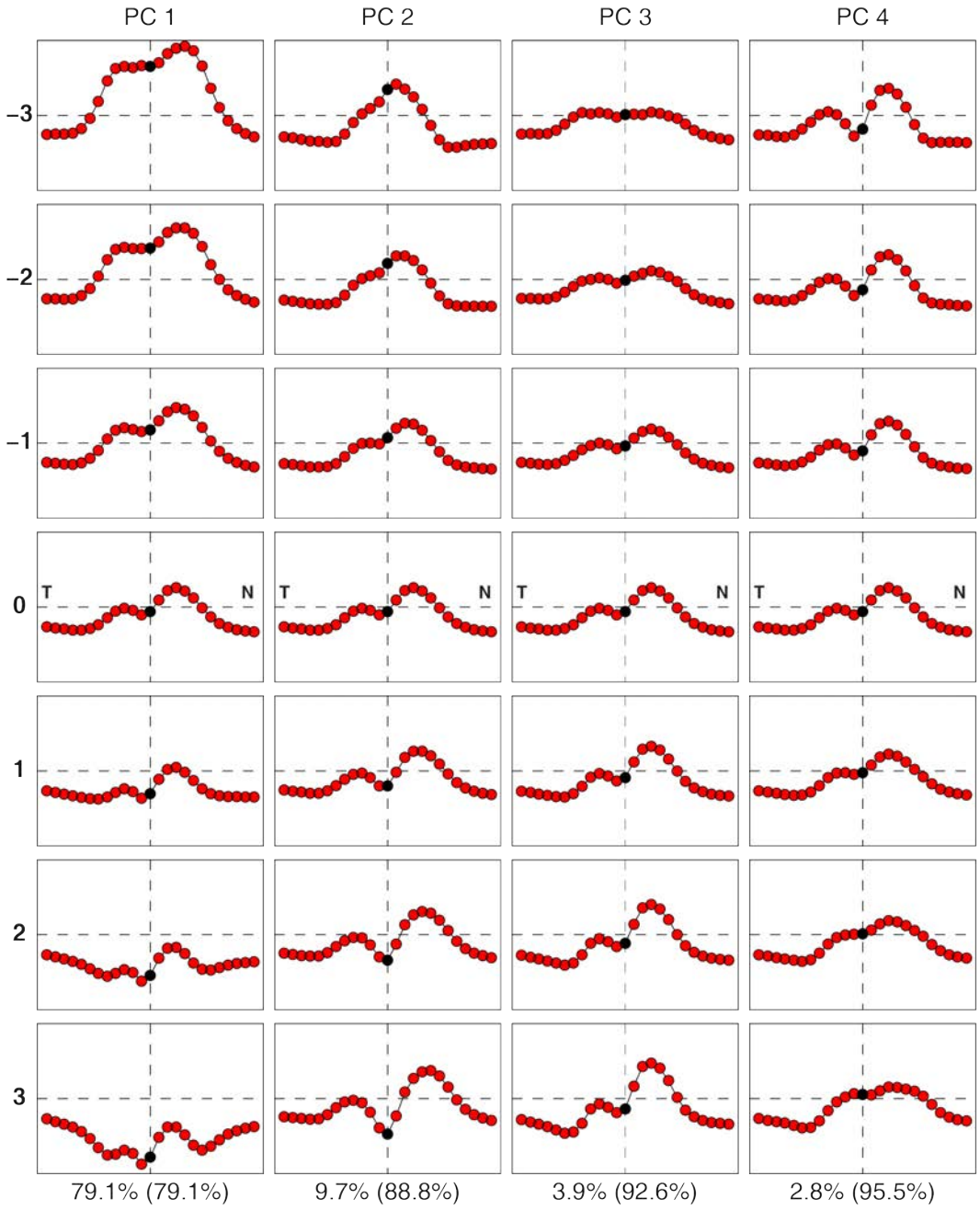


Figure 3.3: Principal component variations of the 2D ILM temporal (T)/nasal (N) shape model. Center landmark indicated in black. Shape coefficient weights shown at left. Rows demonstrate the effect of adding and subtracting PCs to the mean shape. Individual and cumulative energy of each PC shown at bottom.

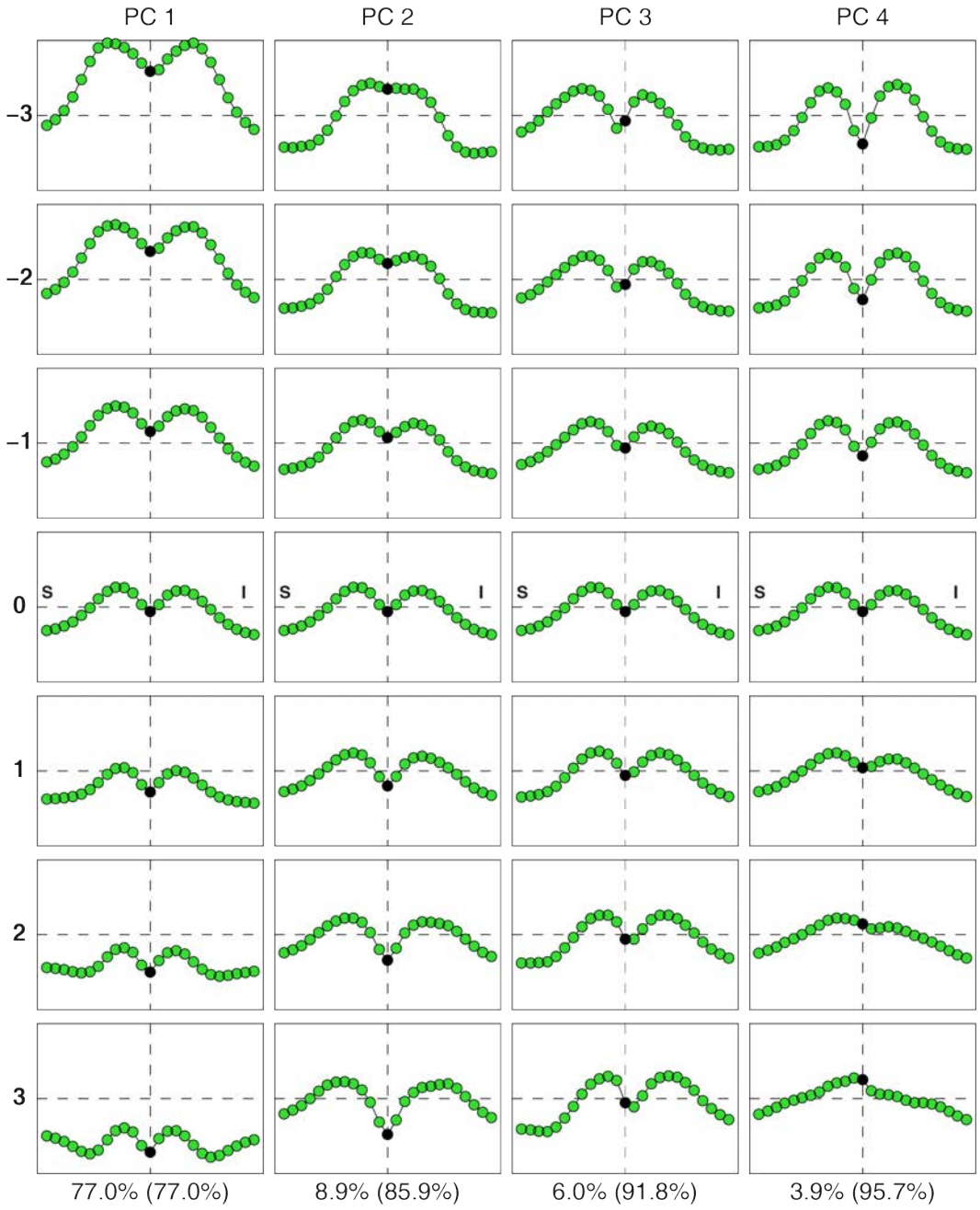


Figure 3.4: Principal component variations of the 2D ILM superior (S)/inferior (I) shape model. Center landmark indicated in black. Shape coefficient weights shown at left. Rows demonstrate the effect of adding and subtracting PCs to the mean shape. Individual and cumulative energy of each PC shown at bottom.

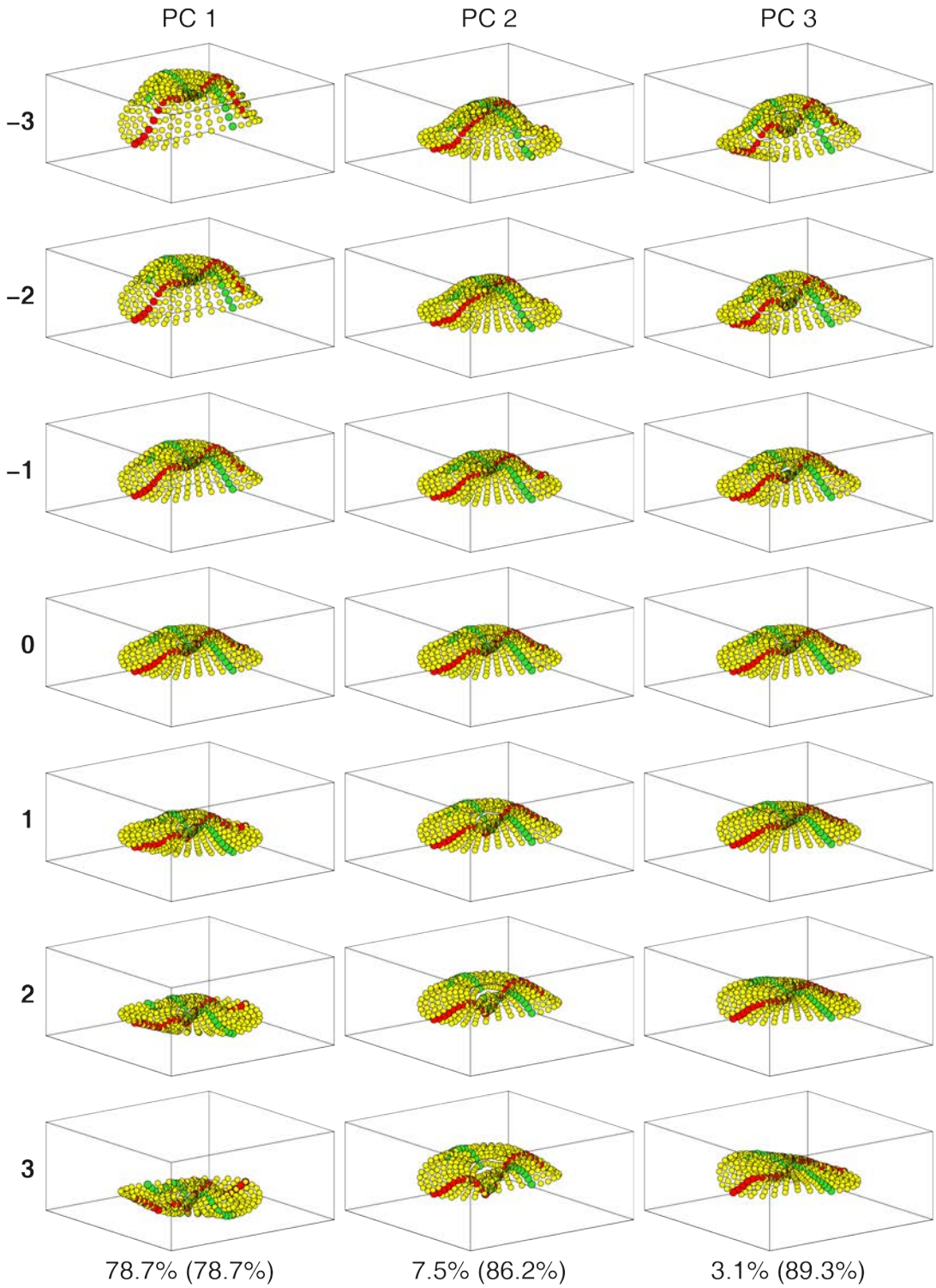


Figure 3.5: Principal component variations of the 3D ILM shape model. Highlighted segments are temporal/nasal (red) and superior/inferior (green). Shape coefficient weights shown at left. Rows demonstrate the effect of adding and subtracting PCs to the mean shape. Individual and cumulative energy of each PC shown at bottom.

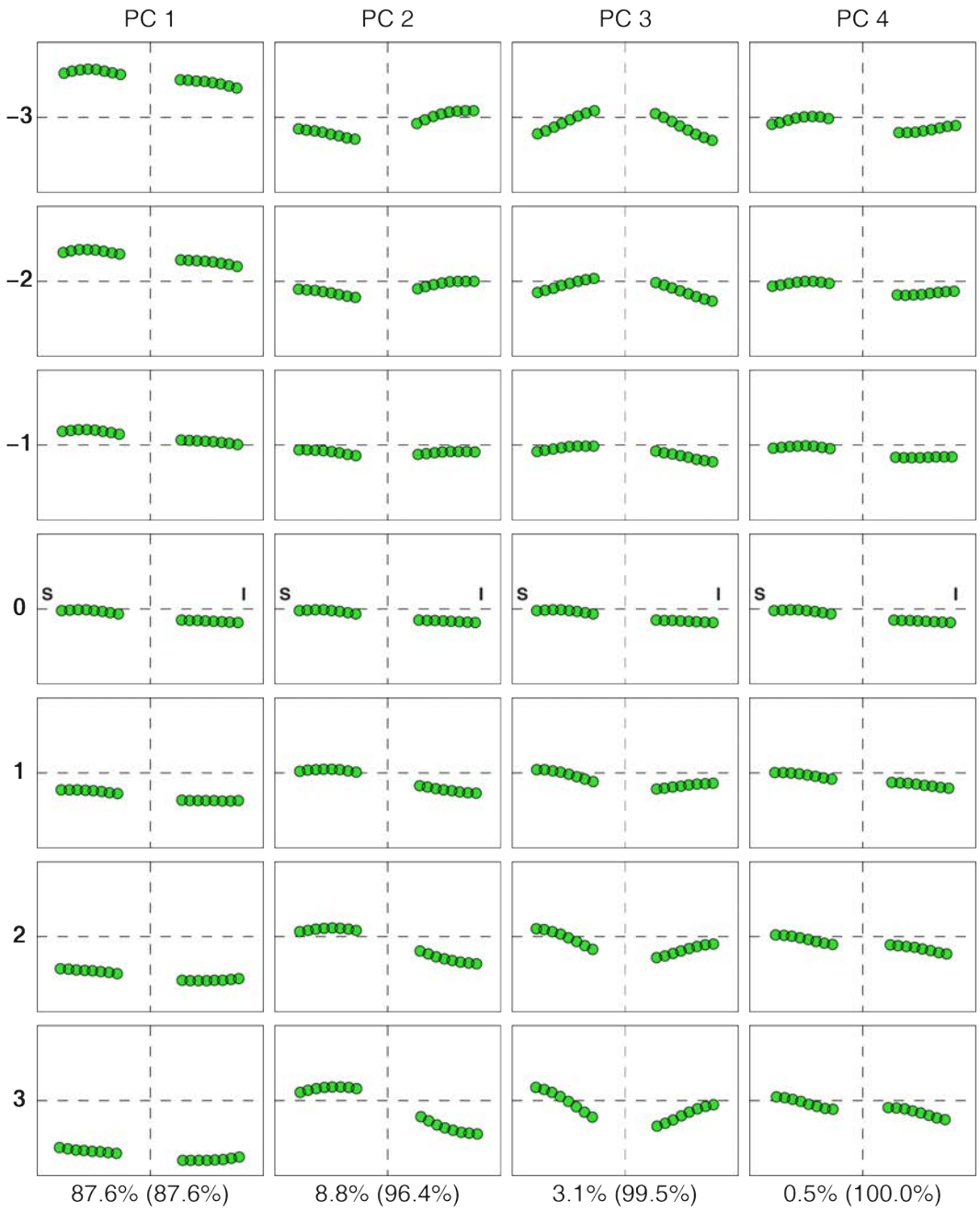


Figure 3.6: Principal component variations of the 2D BM superior (S)/inferior (I) shape model. Center landmark indicated in black. Shape coefficient weights shown at left. Rows demonstrate the effect of adding and subtracting PCs to the mean shape. Individual and cumulative energy of each PC shown at bottom.

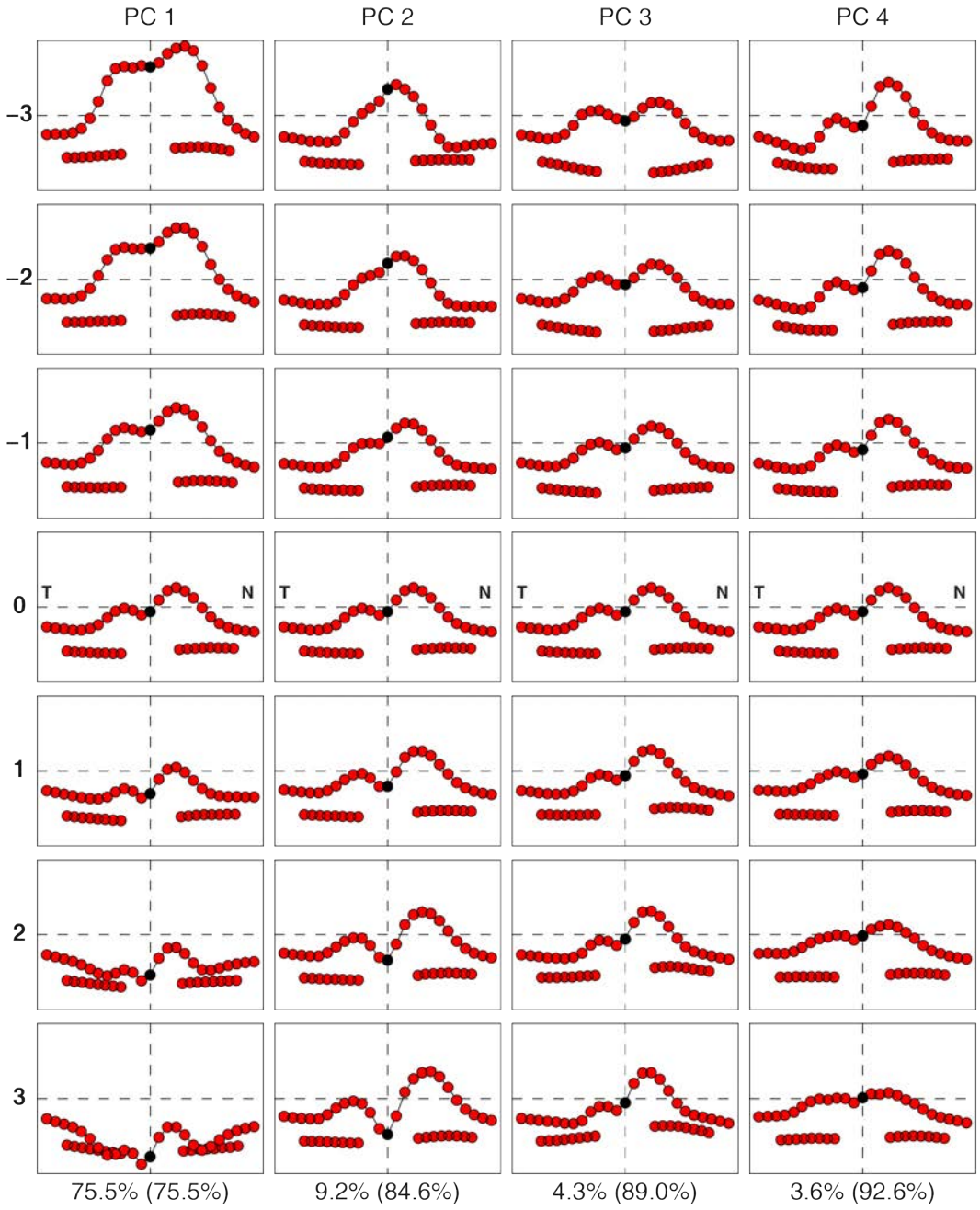


Figure 3.7: Principal component variations of the 2D ILM+BM temporal (T)/inferior (I) shape model. Center ILM landmark indicated in black. Shape coefficient weights shown at left. Rows demonstrate the effect of adding and subtracting PCs to the mean shape. Individual and cumulative energy of each PC shown at bottom.

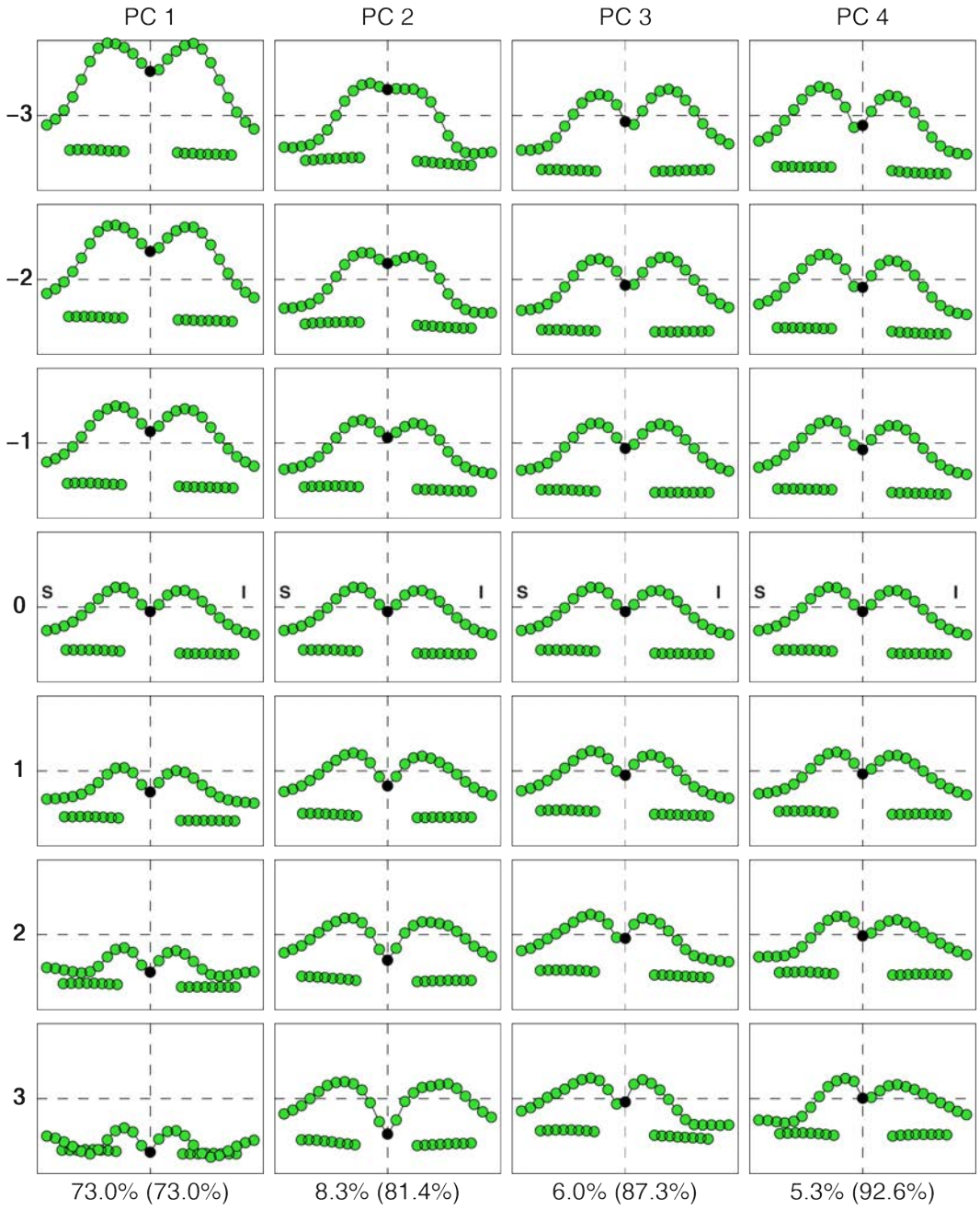


Figure 3.8: Principal component variations of the 2D ILM+BM superior (S)/inferior (I) shape model. Center ILM landmark indicated in black. Shape coefficient weights shown at left. Rows demonstrate the effect of adding and subtracting PCs to the mean shape. Individual and cumulative energy of each PC shown at bottom.

3.4 Discussion

The previous section describes the PC variations of six new retinal shape models. There is a clear difficulty in describing the shape variations of these models. The second PC of the original BM model could be adequately explained as angulation towards or away from the vitreous, but the shape variations of the ILM models are much more complex. The ILM+BM model, for example, is comprised of 41 landmarks and represents simultaneous variations in the ILM and BM structures. The 3D ILM model consists of 432 landmarks and its PCs represent simultaneous changes in swelling across all directions.

There are no theoretical constraints on the models that suggest they will produce shape variations with intuitive interpretations. As the shape models grow in complexity, it will be important to develop more systematic methods for describing their behavior. One option for a more quantitative evaluation of the relevance of the PCs of each model is explored in Chapter 4, where the shape model weights are used as features in random forest classifiers. If variations in shape can be described quantitatively in terms of their role in a classifier, it may make the descriptions of models' traits more manageable.

The ILM presents unique challenges for shape analysis that were not present with the original models of the RPE/BM. Unlike the BM, which has two physical landmarks consistent across all subjects (the BMO points), the ILM does not include physical landmark features that provide continuity from subject to subject. Shape analysis assumes some amount of physical continuity between the landmarks from different training samples.

Whereas the BM surface mostly varies in its angulation towards or away from vitreous and maintains its general shape, the ILM in a papilledema subject with extreme swelling will take on a vastly different shape from an NAION subject with minimal swelling. This greater complexity in shape of the ILM is reflected in the

energy contributions of each PC; the first PC of the 2D ILM temporal/nasal model contributes 79% of the total energy (Fig. 3.3) whereas the first PC of the original BM model accounted for approximately 90% [13].

This wide variation in shape of the ILM was addressed in part by assigning landmarks evenly in the horizontal plane, rather than measuring distance along the retinal surface as was done with the original BM models. Placing landmarks evenly along the ILM surface could potentially reflect the shapes of individual scans more accurately, but landmark placement would vary widely between scans depending on the curvature of the ILM.

An additional challenge when constructing shape models from volumetric SD-OCT scans of swollen optic discs is precisely locating the center of the BMO. This uncertainty in the BMO location further increases the physical landmark inconsistencies described above. Our lab is currently developing automatic methods for more accurately segmenting the BMO contour.

Another potential limitation of this study – that will be addressed at length in Chapter 5 – is using a shape model training dataset consisting entirely of subjects with papilledema. While the 116 papilledema scans used to train the shape models show a wide variety of swelling, it's possible that the shape models are biased by their training data and may more accurately represent papilledema shapes than NAION.

CHAPTER 4
DIFFERENTIATING BETWEEN CAUSES OF OPTIC DISC EDEMA

The ultimate goal of this work is to develop a noninvasive, semi-automated diagnostic tool for differentiating between papilledema and NAION, two types of optic disc edema. Current clinical techniques for diagnosing swelling of the optic disc require expert training and rely on subjective inferences drawn from fundus photographs. A rapid, accurate diagnosis of the type of swelling is essential, as in cases of papilledema, the underlying cause of raised ICP may require immediate intervention. An algorithmic, objective technique for diagnosing optic disc edema could save doctors and patients time while providing a more robust metric. This chapter describes the training and evaluation of random forest classifiers as progress towards that goal.

Retinal layer shape has emerged as a metric that is useful for differentiating between types of optic disc edema [11, 12]. Shape information is particularly useful for differentiation in cases where the optic disc is not severely swollen and the optic disc volume measurement could be found in cases of both NAION or papilledema [12]. While retinal layer shape (in particular the angulation of BM) has been shown to differ between types of optic disc edema, a necessary step towards clinical relevance is demonstrating the utility of layer shapes in the context of a machine learning classifier.

The random forest classifiers described in this chapter were trained on 78 volumetric SD-OCT scans of papilledema and NAION patients collected at the University of Iowa. Retinal layer shapes were the primary features extracted from the SD-OCT scans. Regional measures of the optic disc volume were also calculated from the scans and used as additional features.

This chapter first describes the theory behind random forest classifiers, then details the specific implementation of a classifier trained on retinal layer shapes derived from volumetric SD-OCT images. Finally, the results of the classifier are presented

along with a brief discussion of the work.

4.1 Random forest theory

Random forest classifiers are an ensemble learning method built from decision trees. Decision trees often follow the Classification and Regression Tree (CART) framework, and, as the name implies, can be used for either classification or regression tasks [51]. This thesis focuses solely on binary classification trees and will use the general term decision trees to refer specifically to binary classification trees. Random forests, along with other improvements which will be detailed in Section 4.1.2, combine the outputs of many decision tree models in order to produce an output that is both more accurate and less sensitive to noise than traditional decision trees.

4.1.1 Decision trees

Decision trees operate with a simple objective of partitioning a dataset into two categories using a sequence of binary "Yes" or "No" splits. Decision trees achieve this objective by sequentially identifying features on which to split the data into subsets that are more homogeneous than the current subset. For example, when diagnosing a fever, a decision tree may classify a patient as having or not having a fever by answering a series of Yes or No questions, such as "is the temperature greater than 100F?", "are the palms sweaty?", etc, and arriving at a classification of having or not having a fever. The process of training a decision tree determines which Yes or No questions to ask (i.e. which features to split on) and at what values to set each Yes/No threshold.

There are a few parameters to select while training the binary tree: which features do we split on? How do we determine the threshold for splitting? When do we stop splitting the data (i.e. stop adding nodes to the tree)? The first two questions can be addressed through the use of an impurity metric.

Impurity can be thought of as a measure of the heterogeneity of dataset; a set

of data with an equal distribution of two classes would be highly impure. The Gini impurity index, $I(t)$, is a common measure of impurity in random forests [52–54]:

$$I(t) = \sum_{m=1}^{M} P(\omega_m|t)(1 - P(\omega_m|t)) \tag{4.1}$$

where $P(\omega_m|t)$ is the probability of finding a point, ω_m, from class m in a subset X_t and M is the total number of classes. X_t is the current set of points at node t, before the data are split into two purer subsets. The probability is approximated as

$$P(\omega_m|t) = \frac{N_t^m}{N_t}, m = 1, 2, \ldots, M$$

where N_t^m is the number of points from class m in a potential new subset X_{tY} and N_t is the total number of points in X_t.

The Gini impurity is maximal when $P(\omega_m|t)$ is uniform. Decision trees identify the optimal feature and threshold at each split by iterating through each of the features, testing different threshold values, and picking the combination of feature and threshold that results in the biggest decrease in $I(t)$. By minimizing the Gini impurity, decision trees identify subsets of the data that will be maximally pure.

The third question, determining when to stop splitting the data, is typically addressed through a heuristic approach known as pruning [53]. It is difficult to stop training a tree at precisely the right point; stopping early would result in poor classification results and stopping too late may result in overfitting. One common practice is to allow the tree to grow to a large size – potentially overfitting to the data – and then removing nodes to balance prediction error and model complexity [53].

Decision trees provide a number of advantages over other standard classification techniques. If a tree remains small, its results are easily interpretable. Decision trees can be adapted for regression or classification. And in general, decision trees are computationally simple and quick to fit.

However, there are also a number of substantial disadvantages associated with decision trees. Compared to support vector machines and other ensemble techniques, decision trees have low accuracy and are highly sensitive to small changes in their training data [53]. Additionally, decision trees are only able to split data parallel to the feature axes and are forced to approximate diagonal lines as many horizontal and vertical splits [52]. Random forests improve substantially on both of the issues of accuracy and instability while making justifiable concessions in interpretability of the model.

4.1.2 Improvements of random forests over classical decision trees

Random forests were introduced in 2001 by Breiman to address a number of shortcomings associated with traditional decision trees, such as low accuracy and high sensitivity to variations in training data [55]. Random forests use a bootstrapped ensemble of decision trees along with a random sampling of features during training to achieve good accuracy while remaining robust to noise [55]. Random forests are now widely used in a variety of disciplines and generally applicable to a number of problems [56, 57].

Bagging (bootstrap aggregating) refers to the practice of training multiple trees on random subsets (with replacement) of the training data with the intent of reducing variance in the final model [58]. The final output of the ensemble of classifiers is then typically determined through an averaging or winner-takes-all vote of the output of each tree [53].

Traditional bagging trains trees on distributions that are identically distributed but not necessarily independent. The variance of N identically distributed random variables with variance σ is $\rho\sigma^2 + \frac{1-\rho}{N}\sigma^2$, whereas the variance of an average of N independent identically distributed random variables is $\frac{1}{N}\sigma^2$ [56]. Random forests attempt to achieve the smaller predictor variance by producing trees that are inde-

pendent as well as identically distributed. This is achieved by randomly selecting a subset of features at each split of the data (i.e. tree node).

As the trees comprising a random forest are trained using bagging, which selects a random subset of data for each tree, it is possible to identify samples that were not included in the training of a given tree. These samples are referred to as out-of-bag (OOB) and can be treated like an internal test set to evaluate the model [56]. The OOB error estimate is the mean prediction error of a sample x_i calculated from the trees which did not include x_i in their subset of training data [55, 56].

Random forests (and decision trees is general) lend themselves to straightforward calculation of feature importance. A common approach for calculating feature importance in random forests is to simply rank features based on their total reduction in the Gini impurity at each split node accumulated across each of the trees used to build the forest [56]. It is common practice to normalize the importance of all features to range from 0 to 1 or 0 to 100 [54].

4.2 Implementation

The random forest classifiers in this work were trained on 78 volume-matched SD-OCT scans (39 Papilledema, 39 NAION) collected from subjects with papilledema and NAION at the University of Iowa, the details of which will be described in Section 4.2.1. The goal of the classifiers was to determine as accurately as possible whether a given scan was from a patient with papilledema or NAION. To address the limited amount of available data, 100 random forest classifiers were trained, each using leave-one-out cross-validation, and their accuracies were averaged together. Each random forest consisted of 50 trees trained on random subsets of $\log_2 N$ features, where $N = 76$. The classifiers were implemented in Python using the *scikit-learn library* [54] on a 64-bit Windows 7 PC with 32 GB of RAM and an Intel Xeon 3.60 GHz E5-1630 CPU. Fig. 4.1 outlines the classifier training and evaluation procedure.

The random forest classifiers were trained on a combination of retinal layer shapes

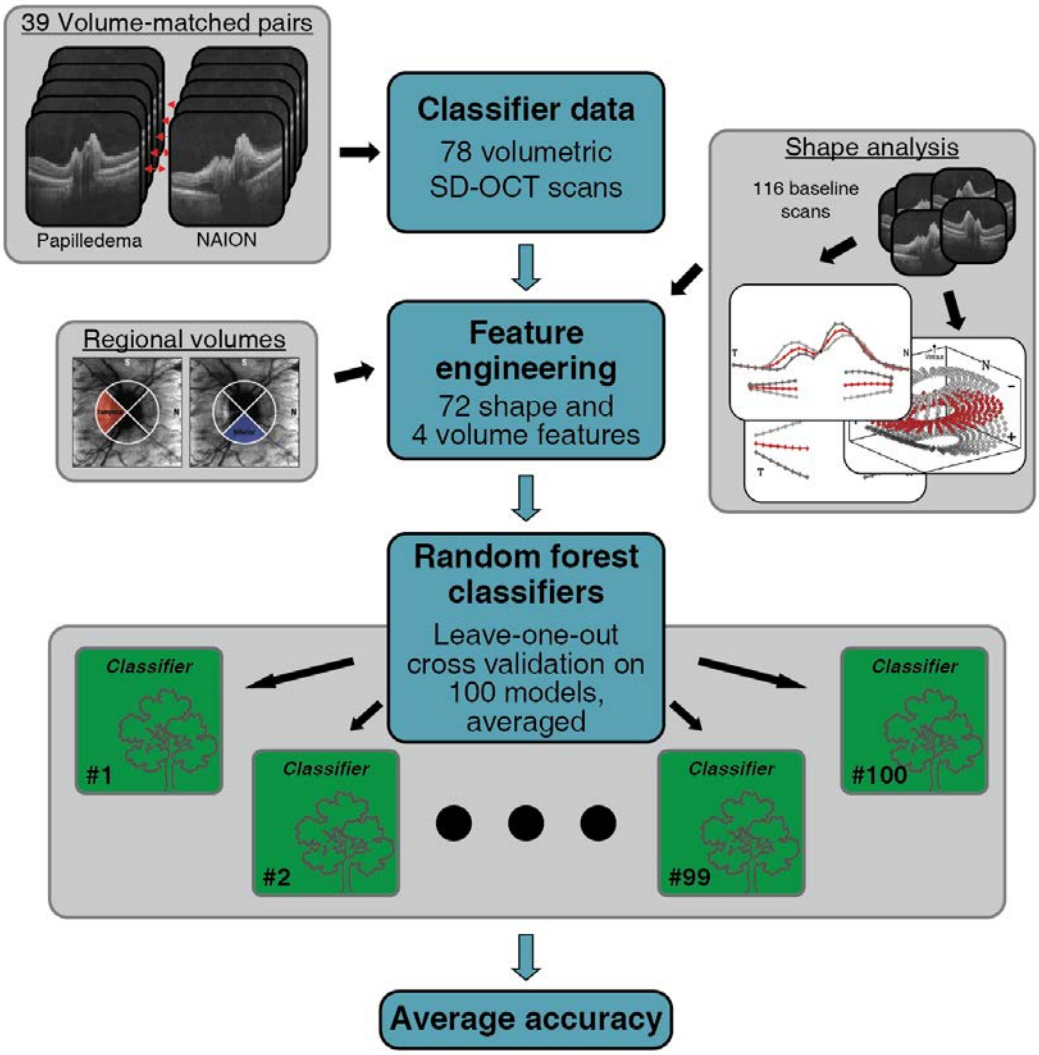


Figure 4.1: Overview of the random forest model training and evaluation procedure.

and regional volume measurements, both of which were derived from volumetric SD-OCT scans. The classifiers were initially trained and evaluated on all 76 shape and regional volume features, details of which are given in Section 4.2.3. Then, after calculating the feature importance for each classifier, a new set of 100 classifiers were trained and evaluated on only the top 15 most important features. Feature importance was calculated using the Gini impurity metric as described in Section 4.1.1 and averaged across all classifiers. The classifiers were evaluated and are presented

here in terms of average accuracy and confusion tables.

4.2.1 Scan selection

The SD-OCT scans collected at the University of Iowa consisted of 61 subjects with papilledema and 60 subjects with NAION. All of the papilledema subjects had both HD-5LR and volumetric scans available. Of the 60 NAION subjects, 12 had both types of scan available. For the subjects with only volumetric scans available, the BMO points for segmentation were placed directly on central slices from volumetric scan as described in Section 3.2.2.

A number of the papilledema and NAION subjects had to be excluded from the analysis for a variety of reasons: scans without both the ILM and BM visible were excluded; the segmentation had to correctly identify BM and the ILM; each scan was required to have a matching scan from the opposite class with a difference in total optic disc volume below a certain threshold (Section 4.2.2). Additionally, it was desired to have an equal number of papilledema and NAION subjects in the dataset. Therefore, the total number of available scans for each class was 39, determined by the number of papilledema subjects available.

4.2.2 Volume matching

Volume of the optic disc is typically measured as the total volume contained between the ILM and BM surfaces, which can be computed after segmenting a volumetric OCT image. As was previously established in Chapter 2, it's possible for OCT scans from two different types of optic disc edema, (e.g. one papilledema and the other NAION), to have equal volume measurements while having different shape characteristics.

Clinically, the volume of the optic disc is a useful feature for distinguishing between papilledema and NAION, as the distributions of volume seen between the two causes of swelling differ. However, the tails of the distributions do overlap, and it is in these

cases where volume is no longer as useful for determining the cause of swelling. This thesis aims to focus on these "hard" cases and does not use total volume as a feature for differentiation. It is for this reason, to control for differences in total disc volume when evaluating the efficacy of layer shapes for differentiation, that the data were comprised of volume-matched papilledema and NAION pairs. In other words, unlike in the general population where the distributions are fairly separate, the distributions of total optic disc volume for the NAION and papilledema subjects used to train and evaluate the classifiers overlap.

A greedy algorithm was used to match papilledema and NAION scans, starting with the lowest volume NAION scan and moving in order to the highest volume scan. If a papilledema scan whose total volume differed from the NAION scan by less than $0.5\,\mathrm{mm}^3$, the scan was left out of the analysis. Fig. 4.2 demonstrates the overlap in distributions of total optic disc volume for the two classes.

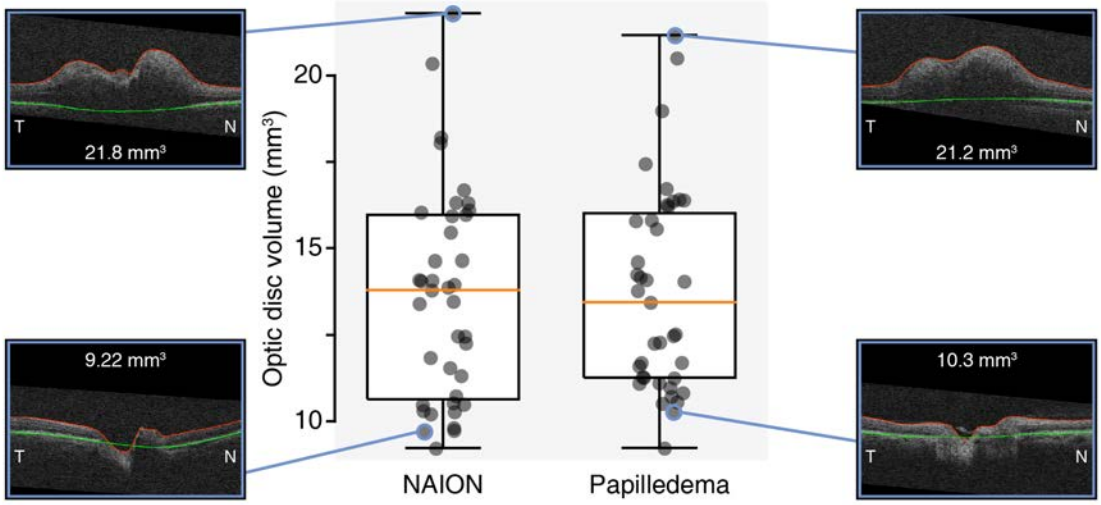


Figure 4.2: Distributions of total optic disc volume for 78 NAION and papilledema scans. The highlighted SD-OCT scans demonstrate the difference in retinal layer shape between scans with some of the most and least swollen optic discs. Orange bars indicate medians. Boxes extend to lower and upper quartiles of data.

4.2.3 Feature extraction

Of the 76 features used to train the classifiers, 72 of them were shape coefficients derived using the shape analysis techniques described in Chapter 3: 8 features from the 2D temporal/nasal ILM model, 8 from the 2D superior/inferior ILM model, 20 from the 3D ILM model, 4 from the 2D temporal/nasal BM model, 4 from the 2D superior/inferior model, 8 from the 3D BM model, and 10 from both ILM+BM models.

From a machine learning perspective, the purpose of Chapter 3 can be thought of as feature design for use in the random forest classifiers. After the 78 scans were segmented and labeled with landmarks as described in Section 3.2.3 of Chapter 3, the layer shapes were transferred from image space to shape model space using Eq. 3.8, reproduced here for reference:

$$\mathbf{b}_i = \mathbf{P}^T(\mathbf{q}_i - \bar{\mathbf{x}})$$

where $\mathbf{b}_i$ is a column vector of length t, the number of principal components used in the reconstruction, and $\mathbf{q}_i$ is the vector of landmarks (i.e. shape) for scan i.

The remaining four features were the normalized regional volumes within the peripapillary circle, shown in Fig. 4.3. The regional volumes were calculated using the volumetric segmentation technique described in [25]. Each scan's total volume was computed as the volume contained between the ILM and BM surfaces. The regional volumes were then computed as quadrants evenly spaced in a circle surrounding the BMO center point (Fig. 4.3). The regional volumes for each subject were normalized based on the scan's total volume.

4.3 Classifier results

When trained on all 76 features, the random forest classifiers determined the correct diagnosis with an average accuracy of 85.1±2.2%. After retraining on only the

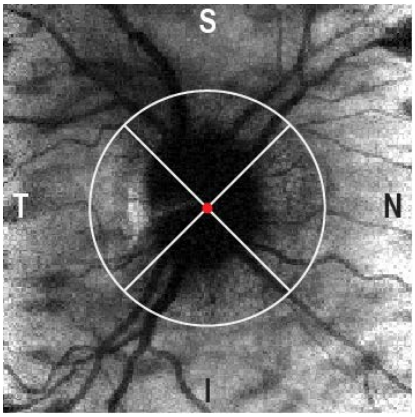


Figure 4.3: Depiction of the temporal, superior, nasal, and inferior regions from which the regional volume features were computed. BMO center indicated in red.

top 15 most important features, the average accuracy increased to $85.7 \pm 2.0\%$. Fig. 4.4(a) shows the importance of each of the 76 features used to train the random forest classifiers. Fig. 4.4(b) shows all features, sorted in order of decreasing importance.

Fig. 4.5 illustrates the characteristics of the four most important features across the 100 classifiers. In order of decreasing importance, the top four features were the third PC of the 2D ILM+BM TN model, first PC of the 2D BM TN model, first PC of the 3D BM model, and the first PC of the 2D BM SI model. All of the four most important features were drawn from either a BM model or an ILM+BM model (Fig. 4.4(a)).

Fig. 4.6 provides a more detailed view of the discriminating ability of the top 15 features. A wider separation between the NAION and papilledema values in the box plots corresponds to a greater ability of that feature to separate the two classes.

The confusion table for the classifiers trained on the top 15 most important features is shown in Fig. 4.7. The total number of predictions was 7800: 100 classifiers $\times$ 78 $\frac{\text{predictions}}{\text{classifier}}$. The table also includes percentage values, indicating the percent of total predictions a single category (i.e. NAION classified as NAION, NAION classified as papilledema, etc.) accounts for.

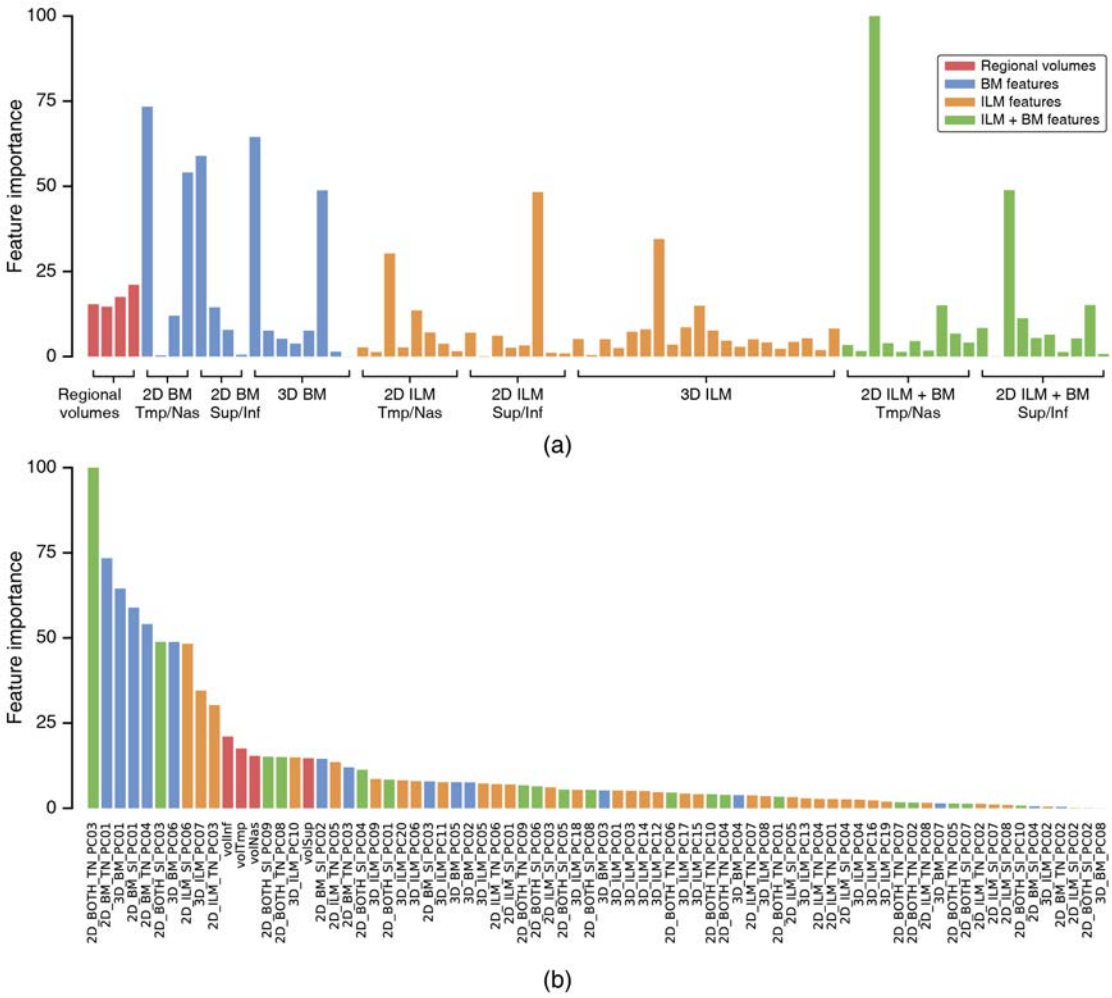


Figure 4.4: Importance of the 76 features used to train the random forest classifiers, scaled to percentage of the most important feature. (a) All features, organized by feature type. (b) All features, sorted in order of decreasing importance. Abbreviated feature names listed beneath each column. TN = temporal/nasal; SI = superior/inferior.

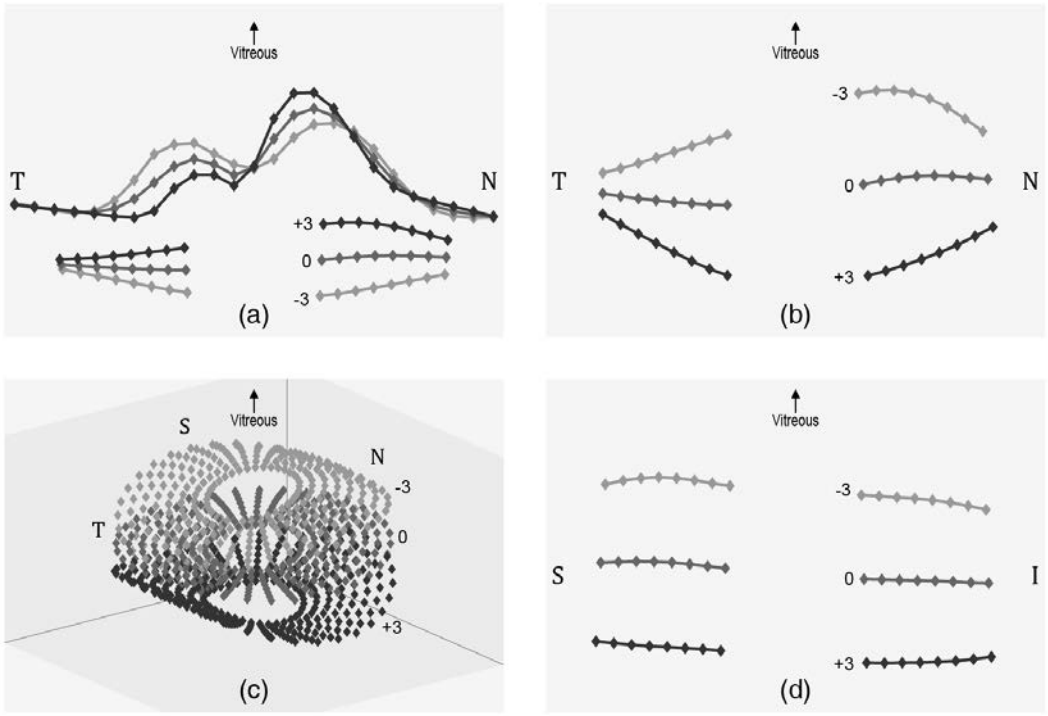


Figure 4.5: Four most important features across all classifiers. (a) Third PC of the 2D ILM+BM TN model. (b) First PC of the 2D BM TN model. (c) First PC of the 3D BM model. (d) First PC of the 2D BM SI model. Mean shape indicated by 0, standard deviations from the mean shape indicated by -3 and $+3$.

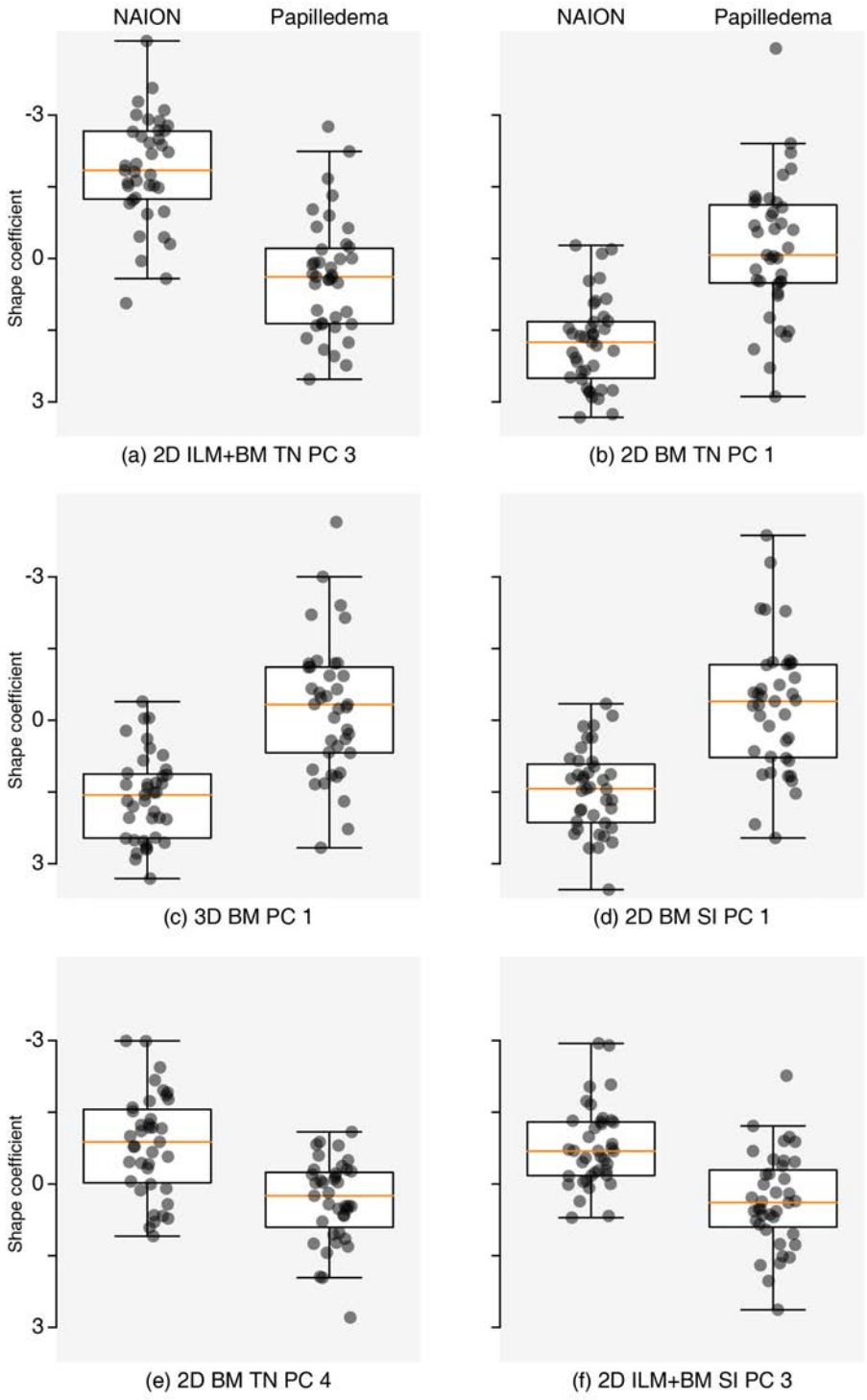


Figure 4.6: Discriminating ability of the six most important features. Further separation between boxes indicates greater discriminating ability. Orange bars indicate medians. Boxes extend to lower and upper quartiles of data.

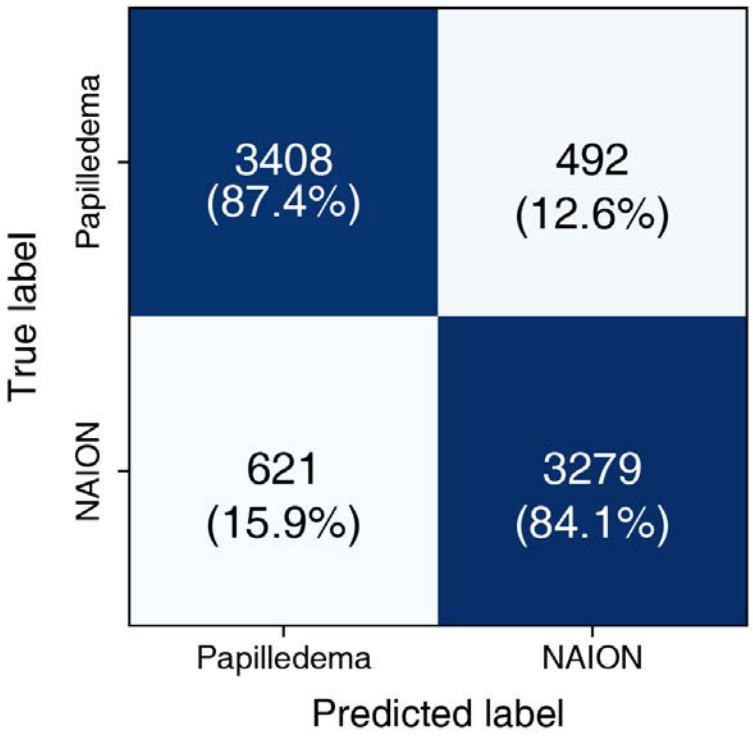


Figure 4.7: Confusion table for the classifiers trained on the top 15 most important features. The integer values sum to the total number of predictions across all 100 classifiers. Parenthetical terms sum to 100% for each row, indicating the percentage of predictions within a true class.

4.4 Discussion

While working with a relatively small dataset, it was important to avoid overfitting. As such, the potential parameter space to explore for fine-tuning the classifiers (e.g. number of trees, bagging size, random feature selection) was kept small. Random forests were chosen in part for their tendency towards minimal overfitting [55,56]. A larger dataset would of course allow for further exploration while reducing the risk of overfitting.

As expected based on previous work [11, 13, 59], angulation of BM towards and away from vitreous was one of the most useful features for differentiating between types of optic disc edema. The BM angulation is represented in some form in each of the first three most important features. One area to explore more in the future is further refinement of the classifier features in order to reduce redundancy in the feature space. For example, performing PCA on the feature space would likely identify combinations of features that all represent the same general shape trends (e.g. identifying multiple shape models each reflecting angulation of BM). This reduced feature space could then be used to train a new classifier.

It is also important to point out the role ILM shape plays in the most important feature. This feature suggests that asymmetry of the ILM surface in the temporal/nasal direction is useful for differentiating between NAION and papilledema, an insight that could not have been drawn from models of BM alone. If ILM shape were not useful for differentiation, the classifier would have likely preferred any number of the features from BM models that only showed angulation of BM.

It is worth noting which scans were most misclassified. A cursory examination of the misclassified scans reinforces the importance of the BM angulation. The scans that were most often misclassified were papilledema subjects whose BM angulation was directed away from vitreous, a shape characteristic that is more common in normal or NAION subjects. As more data are added to the study, it will be interesting to

see what additional systematic trends in misclassification may arise.

The confusion table in Fig. 4.7 shows that NAION subjects were slightly more likely to be misclassified (15.9% error) than scans from papilledema subjects (12.6% error). However, more data are required to draw any strong conclusions about a bias towards classifying one condition over the other.

CHAPTER 5
DISCUSSION AND FUTURE WORK

5.1 Discussion and future work

The two key contributions of this thesis are the development of shape models of the inner limiting membrane (ILM) and the use of ILM and retinal pigment epithelium (RPE) shapes in a random forest classifier trained to distinguish between two causes of optic disc edema: papilledema and nonarteritic anterior ischemic optic neuropathy (NAION).

Along with a conference abstract published in 2018 [60], this thesis includes the first development of shape models of the ILM and the use of retinal shape information as features in a machine learning classifier trained to distinguish between types of optic disc edema. The temporal dynamics of ILM shape remain uninvestigated. Clinical evidence suggests cases of NAION may present with an asymmetry of the optic disc in the superior/inferior orientation [7] that changes over time and which shape models of the ILM may be able to help study.

A clear limitation of this work is the limited amount of data available for the training and evaluation of the random forest classifiers. The risk of overfitting on the small dataset limited the amount of hyperparameter tuning (e.g. number of trees and number of random features) that could be done with the classifiers. Additionally, with testing limited to repeated leave-one-out cross validation, it is difficult to predict how the accuracy of the classifiers may change when evaluated on a larger test set of data.

One potential means for increasing the data available for training and evaluating the classifiers is currently under development. As described in Section 4.2.1, a number of the SD-OCT scans collected from NAION subjects at the University of Iowa were improperly segmented due to the extreme swelling of the ILM (Section 4.2.1). The graph-based segmentation approaches don't always follow the steep curves sometimes

present in images of extreme optic disc edema. Gradient vector flow fields have previously been combined with graph-based techniques to segment the steep curvature present in SD-OCT images from glaucoma patients that can sometimes cause traditional graph algorithmic approaches to fail [41]. These techniques are currently being adapted for cases of disc edema and could make additional NAION scans available in future work.

Another potential limitation of this study is the baseline data used to train the preliminary shape models (which were later used to generate the shape features for the classifier). The shape models were trained using the baseline scans of the right eyes of 116 patients from the Idiopathic Intracranial Hypertension Treatment Trial (IIHTT), i.e. trained only on patients with papilledema. It is possible that this model doesn't reflect all of the potential shape variations seen with NAION.

However, the baseline scans from the IIHTT cover a wide range of degrees of swelling, many of which fall in the same range of swelling as typically seen in patients with NAION. Additionally, preliminary examinations using model error – the ability of the shape model to accurately represent both the NAION and papilledema – as a feature in the classifiers did not provide further discriminating ability, an indication that the model does recreate papilledema shapes any better (or worse) than shapes from NAION patients. The ideal shape model training set would include numerous examples from papilledema and NAION subjects along with normal scans.

Envisioning automated diagnostics at use in an actual clinical setting, it would likely be useful to adopt more of a Bayesian perspective when considering the problem of diagnosing optic disc edema. Incorporating the incidence rates (i.e. priors) for NAION and papilledema along with other relevant clinical information (e.g. sex, age, weight, visual acuity) could certainly improve the performance of the classifier.

This work has clear applications to the field of telemedicine. Expert clinicians already of course incorporate their prior knowledge when diagnosing patients, but an

algorithmic system which has the relevant priors built-in could provide physicians who do not have the necessary extensive training in neuro-ophthalmology a useful – and potentially time-critical in cases of papilledema – point of reference when diagnosing patients.

5.2 Conclusions

The two main contributions of this thesis are the development of shape models of the inner limiting membrane (ILM) and the use of retinal layer shapes in a machine learning classifier. The theory and implementation of retinal layer shape models was described in Chapter 3. The training and evaluation of random forest classifiers for diagnosing papilledema and NAION from retinal shape features was covered in Chapter 4. The most important feature for differentiating between NAION and papilledema was drawn from a combination shape model of both the ILM and BM surfaces. The asymmetry of the ILM displayed in this feature has not previously been reported in the context of retinal layer shape models.

As more meaningful shape features are developed and more data are used to train and evaluate the classifiers, the estimates of classification accuracy and robustness will improve, continuing towards the ultimate goal of developing reliable and accurate noninvasive tools for diagnosing optic disc edema.

REFERENCES

[1] Palanker, "Diagram of the eye and placement of the retinal implants," 2016, wikimedia Commons. [Online]. Available: https://commons.wikimedia.org

[2] G. Van Stavern, "Optic Disc Edema," *Seminars in Neurology*, vol. 27, no. 3, pp. 233–243, Jul 2007.

[3] M. Fontal, J. Kerrison, R. Garcia, and V. Oria, "Ischemic Optic Neuropathy," *Seminars in Neurology*, vol. 27, no. 3, pp. 221–232, Jul 2007.

[4] C. Scott, R. H. Kardon, 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," *Archives of Ophthalmology*, vol. 128, no. 6, pp. 705–711, 2010.

[5] 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," *Journal of Neurology*, vol. 259, no. 7, pp. 1406–1412, 2012.

[6] A. Ophir, M. Karatas, J. Ramirez, and R. Inzelberg, "OCT and Chronic Papilledema," *Ophthalmology*, vol. 112, no. 12, p. 2238, 2005.

[7] M. J. Kupersmith, R. Kardon, M. Durbin, M. Horne, and J. Shulman, "Scanning laser polarimetry reveals status of RNFL integrity in eyes with optic nerve head swelling by OCT." *Investigative Ophthalmology & Visual Science*, vol. 53, no. 4, pp. 1962–1970, 2012.

[8] V. L. Migdal, W. K. Wu, D. Long, C. D. McNaughton, M. J. Ward, and W. H. Self, "Risk-benefit analysis of lumbar puncture to evaluate for nontraumatic subarachnoid hemorrhage in adult ED patients," *The American Journal of Emergency Medicine*, vol. 33, no. 11, pp. 1597–1601, Nov 2015. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0735675715005343

[9] M. J. Kupersmith, P. A. Sibony, and S. Dave, "Nonarteritic anterior ischemic optic neuropathy induced retinal folds and deformations," *Investigative Ophthalmology & Visual Science*, vol. 58, no. 10, pp. 4286–4291, 2017.

[10] J. Sahuquillo and F. Arikan, "Decompressive craniectomy for the treatment of refractory high intracranial pressure in traumatic brain injury." *Cochrane database of systematic reviews (Online)*, vol. 3, no. 4, p. CD003983, 2006.

[11] 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," *Investigative Ophthalmology & Visual Science*, vol. 52, no. 9, pp. 6558–6564, 2011.

[12] P. Sibony, M. J. Kupersmith, and F. James Rohlf, "Shape analysis of the peripapillary RPE layer in papilledema and ischemic optic neuropathy," *Investigative Ophthalmology & Visual Science*, vol. 52, no. 11, pp. 7987–7995, 2011.

[13] 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 9417, Medical Imaging 2015: Biomedical Applications in Molecular, Structural, and Functional Imaging,*, vol. 9417, p. 941721, 2015.

[14] J.-K. Wang, R. H. Kardon, J. Ledolter, P. A. Sibony, M. J. Kupersmith, and M. K. Garvin, "Peripapillary retinal pigment epithelium layer shape changes from acetazolamide treatment in the Idiopathic Intracranial Hypertension Treatment Trial," *Investigative Opthalmology and Visual Science*, vol. 58, no. 5, p. 2554, 2017.

[15] P. A. Sibony, "Gaze evoked deformations of the peripapillary retina in papilledema and ischemic optic neuropathy," *Investigative Ophthalmology & Visual Science*, vol. 57, no. 11, pp. 4979–4987, 2016.

[16] S. S. Hayreh and M. B. Zimmerman, "Optic disc edema in non-arteritic anterior ischemic optic neuropathy," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 245, no. 8, pp. 1107–1121, 2007.

[17] H. E. Killer, G. P. Jaggi, and N. R. Miller, "Papilledema revisited: Is its pathophysiology really understood?" *Clinical and Experimental Ophthalmology*, vol. 37, no. 5, pp. 444–447, 2009.

[18] M. Rigi, S. J. Almarzouqi, M. L. Morgan, and A. G. Lee, "Papilledema : epidemiology , etiology , and clinical management," *Journal of Eye and Brain*, vol. 2015, no. 7, pp. 47–57, 2015.

[19] M. Wall and D. George, "Idiopathic Intracranial Hypertension," *Brain*, vol. 114, pp. 155–180, 1991.

[20] OCT Sub-Study Committee for the NORDIC Idiopathic Intracranial Hypertension Study Group, "Baseline OCT measurements in the Idiopathic Intracranial Hypertension Treatment Trial, Part I: Quality control, comparisons, and variability," *Investigative Ophthalmology & Visual Science*, vol. 55, no. 12, p. 8180, 2014.

[21] M. Wall, M. P. McDermott, K. D. Kieburtz, J. J. Corbett, S. E. Feldon, D. I. Friedman, D. M. Katz, J. L. Keltner, E. B. Schron, and M. J. Kupersmith, "Effect of acetazolamide on visual function in patients with idiopathic intracranial hypertension and mild visual loss: the idiopathic intracranial hypertension treatment trial." *JAMA : the journal of the American Medical Association*, vol. 311, no. 16, pp. 1641–51, 2014.

[22] D. I. Friedman and D. M. Jacobson, "Diagnostic criteria for idiopathic intracranial hypertension," *Neurology*, vol. 59, no. 10, pp. 1492–1495, 2002.

[23] L. Frisen, "Swelling of the optic nerve head: A staging scheme," *Journal of Neurology Neurosurgery and Psychiatry*, vol. 45, no. 1, pp. 13–18, 1982.

[24] 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," *Investigative Ophthalmology & Visual Science*, vol. 53, no. 7, pp. 4069–4075, 2012.

[25] J.-K. Wang, M. S. Miri, R. H. Kardon, and M. K. Garvin, "Automated 3D region-based volumetric estimation of optic disc swelling in papilledema using spectral-domain optical coherence tomography," *Medical Imaging 2013: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 8672, p. 867214, 2013.

[26] E. Z. Karam and T. R. Hedges, "Optical coherence tomography of the retinal nerve fibre layer in mild papilloedema and pseudopapilloedema," *British Journal of Ophthalmology*, vol. 89, pp. 294–298, 2005.

[27] G. Rebolleda and F. J. Muñoz-Negrete, "Follow-up of mild papilledema in idiopathic intracranial hypertension with optical coherence tomography," *Investigative Ophthalmology & Visual Science*, vol. 50, no. 11, pp. 5197–5200, 2009.

[28] D. R. Boghen and J. S. Glaser, "Ischemic optic neuropathy: the clinical profile and natural history," *Brain*, vol. 98, pp. 689–708, 1975.

[29] E. J. Atkins, B. B. Bruce, N. J. Newman, and V. Biousse, "Treatment of nonarteritic anterior ischemic optic neuropathy," *Survey of Ophthalmology*, vol. 55, no. 1, pp. 47–63, 2009.

[30] M. G. Hattenhauer, J. A. Leavitt, D. O. Hodge, R. Grill, and D. T. Gray, "Incidence of Nonarteritic Anterior Ischemic Optic Neuropathy," *American Journal of Ophthalmology*, vol. 123, no. 1, pp. 103–107, 1997.

[31] L. Johnson and A. C. Arnold, "Incidence of nonarteritic and arteritic anterior ischemic optic neuropathy. Population-based study in the state of Missouri and Los Angeles County, California." *Journal of Neuro-Ophthalmology*, vol. 14, no. 1, pp. 38–44, 1994.

[32] J. F. Rizzo and S. Lessell, "Optic neuritis and ischemic optic neuropathy. Overlapping clinical profiles." *Archives of Ophthalmology*, vol. 109, no. 12, pp. 1668–72, 1991.

[33] D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and J. G. Fujimoto, "Optical Coherence Tomography," *Science*, vol. 22, no. 2545035, pp. 1178–1181, 1991.

[34] M. Wojtkowski, R. Leitgeb, A. Kowalczyk, T. Bajraszewski, and A. F. Fercher, "In vivo human retinal imaging by Fourier domain optical coherence tomography," *Journal of Biomedical Optics*, vol. 7, no. 3, p. 457, 2002.

[35] K. Li, W. Xiaodong, D. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images-a graph-theoretic approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 119–134, Jan 2006.

[36] M. K. Garvin, M. D. Abràmoff, X. Wu, S. Member, S. R. Russell, T. L. Burns, and M. Sonka, "Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images." *IEEE Transactions on Medical Imaging*, vol. 28, no. 9, pp. 1436–1447, 2009.

[37] K. Lee, M. Niemeijer, M. Garvin, Y. Kwon, M. Sonka, and M. Abramoff, "Segmentation of the optic disc in 3-D OCT scans of the optic nerve head," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 159–168, Jan 2010.

[38] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38 – 59, 1995.

[39] S. Milborrow and F. Nicolls, "Locating facial features with an extended active shape model," in *Computer Vision ECCV 2008*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 504–513. [Online]. Available: http://link.springer.com/10.1007/978-3-540-88693-8_37

[40] A. Gampa, G. Vangipuram, Z. Shirazi, and H. E. Moss, "Quantitative association between peripapillary Bruch's membrane shape and intracranial pressure," *Investigative Ophthalmology & Visual Science*, vol. 58, no. 5, pp. 2739–2745, 2017.

[41] M. S. Miri, M. D. Abràmoff, K. Lee, M. Niemeijer, J.-K. Wang, Y. H. Kwon, and M. K. Garvin, "Multimodal segmentation of optic disc and cup from SD-OCT and color fundus photographs using a machine-learning graph-based approach," *IEEE Transactions on Medical Imaging*, vol. 34, no. 9, pp. 1854–1866, 2015.

[42] J. Agne, J.-K. Wang, R. H. Kardon, and M. K. Garvin, "Determining degree of optic nerve edema from color fundus photography," *SPIE 9414, Medical Imaging 2015: Computer-Aided Diagnosis*, vol. 9414, pp. 1–9, 2015.

[43] J.-K. Wang, "Estimation of papilledema severity using spectral-domain optical coherence tomography," Ph.D. dissertation, University of Iowa, 2016.

[44] J. Agne, "Automated image-based estimation of severity and cause of optic disc edema," Ph.D. dissertation, University of Iowa, 2017.

[45] I. T. Jolliffe, "Principal Component Analysis, Second Edition," *Encyclopedia of Statistics in Behavioral Science*, vol. 30, no. 3, p. 487, 2002.

[46] J. Shlens, "A tutorial on principal component analysis," *CoRR*, vol. abs/1404.1100, 2014. [Online]. Available: http://arxiv.org/abs/1404.1100

[47] 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," *Investigative Ophthalmology & Visual Science*, vol. 57, no. 12, p. 4544, 2016.

[48] J.-K. Wang, S. Johnson, R. H. Kardon, P. A. Sibony, M. J. Kupersmith, and M. K. Garvin, "Three-dimensional Bruch's membrane shape model in cases of papilledema," *Investigative Ophthalmology & Visual Science*, vol. 57, no. 12, 2016.

[49] D. I. Friedman, M. P. McDermott, K. Kieburtz, M. Kupersmith, A. Stoutenburg, J. L. Keltner, S. E. Feldon, E. Schron, J. J. Corbett, and M. Wall, "The idiopathic intracranial hypertension treatment trial: Design considerations and methods," *Journal of Neuro-Ophthalmology*, vol. 34, no. 2, pp. 107–117, 2014.

[50] J.-K. Wang, R. H. Kardon, and M. K. Garvin, "Combined use of high-density and volumetric optical coherence tomography for the segmentation of neural canal opening in cases of optic nerve edema," *SPIE 9413, Medical Imaging 2015: Image Processing,*, vol. 9413, 2015.

[51] L. Breiman, *Classification and Regression Trees*, ser. Wadsworth statistics/probability series. Wadsworth International Group, 1984.

[52] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[53] S. Theodoridis, *Machine Learning: A Bayesian and Optimization Perspective*, ser. .NET Developers Series. Elsevier Science, 2015.

[54] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[55] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[56] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning, 12^{th} printing*, ser. Springer Series in Statistics. New York, NY, USA: Springer New York Inc., 2017.

[57] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognition in parts from single depth images," *Cvpr 2011*, pp. 1297–1304, 2011.

[58] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[59] P. Sibony, M. J. Kupersmith, R. Honkanen, F. J. Rohlf, and A. Torab-Parhiz, "Effects of lowering cerebrospinal fluid pressure on the shape of the peripapillary retina in intracranial hypertension," *Investigative Ophthalmology & Visual Science*, vol. 55, no. 12, pp. 8223–8231, 2014.

[60] J. W. Miller, J.-K. Wang, M. Thurtell, R. H. Kardon, and M. K. Garvin, "Differentiation between papilledema and nonarteritic anterior ischemic optic neuropathy using retinal layer shape and regional volume features in spectral-domain optical coherence tomography," *Investigative Ophthalmology & Visual Science*, vol. 59, ARVO E-Abstract, 2018.