

**Imperial College  
London**

# **Voronoi Analysis of Transport Networks on the Human Retina**



**Artur Donaldson**

**Supervisors:** Prof. Dimitri Vvedensky

& Simon Morgan

**Assessor:** Julia Sedgbeer

Department of Physics  
Imperial College London

This dissertation is submitted in partial fulfillment for the degree of  
*Master in Science*

**Word Count:** 9500 words

Project Code: CMTH-Vvedensky-1

April 2019



## **Declaration**

This project was a joint effort between my Shervin Sabhegi and myself. It was agreed to divide the work between data collection and data analysis. The former was responsible for the skeletonization of the tracings and labelling of branchpoints. I was responsible for the consideration of the geometry of the situation and the simulation of neovascularization. In addition, Voronoi cell areas on the placenta were obtained from Aytac Kilic and Farhan Chatha. Fits were calculated independently on the data. The code produced by the author is available at <https://github.com/tur-ium/voron-eye>. Unless otherwise specified diagrams and illustrations were produced by the author. Diagrams were produced using the matplotlib package in Python. Illustrations, unless otherwise specified, were produced using open-source image editing packages Paint.net and Inkscape.

Artur Donaldson  
April 2019



## Acknowledgements

I would like to thank my supervisors Dimitri Vvedensky and Simon Morgan. Professor Vvedensky has always offered sound guidance, and the careful selection of papers he has suggested have formed the basis of many parts of this project. I am thankful to Simon for his cross-examination of the mapping between coordinate systems in sec. 3.1 and for his advice on algorithm design. I would like to thank Geoffrey West whose book, *Scale*, inspired me to take a project on scaling theory at the culmination of my undergraduate studies and showed that Physics may be applied to understanding societal challenges on many scales. I am grateful to Shervin Sabhegi for having been a considerate partner in this project and my peers Aytac and Farhan for kindly sharing data on the placenta. I would also like to thank the curators of the DRIVE and HRF datasets, Bram van Ginneken and Jan Odstrčilík, for being swift to reply to my queries. Last, but certainly not least, I would like to thank my friends and family for having taken interest in my work and listened patiently to at least their fair share about Voronoi cell analysis.



## Abstract

Voronoi analysis is a powerful technique which has recently been used to find a correlation between the risk of autism and the structure of blood vessels on the placenta. Here we extend Voronoi analysis of blood vessel structure to the retina and find the normalized areas are given by a gamma distribution, which is changed by proliferative diabetic retinopathy (PDR). Unlike the vascular structure of other organs, such as the heart, the structure of the retinal vasculature does not scale across different species, but for humans the same normalized area emerges across individuals. The distribution of areas of core cells within three groups: healthy retinas and retinas with non-proliferative diabetic retinopathy and glaucoma were found to be fitted by a gamma with a maximum at a normalized area  $A^{\text{norm}} = 0.70$ . In contrast, the distribution for PDR peaks at a smaller normalized area of  $A^{\text{norm}} = 0.27$ . A model for the anisotropic addition of branchpoints is able to capture the effects of neovascularization on the Voronoi cell area distribution in PDR. Structural differences in PDR and healthy retinas imply that neovascularization most likely occurs by a different mechanism to healthy growth of blood vessels on the retina. Finally we compare area distributions of Voronoi cells on the retina and on the placenta and find that we can use Voronoi analysis to distinguish between the two organs suggesting blood vessel formation on the surfaces of different organs does not follow a universal scaling law.



# Table of contents

<b>List of figures</b>	<b>xi</b>
<b>List of tables</b>	<b>xiii</b>
<b>Nomenclature</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Voronoi Analysis, Motivation and Previous Work . . . . .	1
1.2 Research Objectives . . . . .	2
1.3 Outline . . . . .	3
<b>2 Biophysics of Transport Networks on the Eye and Applications of Retinal Imaging</b>	<b>5</b>
2.1 The Human Eye and Retinal Imaging Techniques . . . . .	5
2.2 The Animal Eye . . . . .	7
2.3 Growth of Blood Vessels . . . . .	8
2.4 Vascular Diseases of the Eye . . . . .	10
<b>3 Spherical Geometry and Uses of Statistical Distributions</b>	<b>13</b>
3.1 Mapping Fundus Images to Spherical Coordinates . . . . .	13
3.1.1 Mapping Lengths in the Image to Physical Distances on the Retina .	13
3.1.2 Mapping Areas on the Retina to Areas on the Image . . . . .	15
3.1.3 The Reduced Eye Model . . . . .	16
3.2 Fat-Tail Statistical Distributions . . . . .	18
3.3 Comparison of Statistical Distributions . . . . .	18
<b>4 Method</b>	<b>21</b>
4.1 Choice of Data . . . . .	21
4.2 Skeletonization . . . . .	22

4.3	Computer-Assisted Labelling of Branchpoints . . . . .	24
4.4	Voronoi Analysis in Curved Space . . . . .	25
4.4.1	Defining the Region of Interest . . . . .	26
4.4.2	Defining the Grid . . . . .	27
4.4.3	Normalization and Calculation of Areas of Voronoi Cells . . . . .	28
4.5	Data and Code Validation . . . . .	28
4.6	Calculation of Uncertainties in Voronoi Cell Areas . . . . .	29
<b>5</b>	<b>Results and Discussion</b>	<b>31</b>
5.1	Overview . . . . .	31
5.2	Data Collapse . . . . .	31
5.3	Comparison of DRIVE and HRF data . . . . .	34
5.4	Uncertainties in Area Measurements . . . . .	36
5.5	Effects of Curvature . . . . .	36
5.6	Effects of Neovascularization . . . . .	38
5.7	Comparison between the Retina and Placenta . . . . .	40
<b>6</b>	<b>Conclusion and Future Work</b>	<b>45</b>
6.1	Limitations and Future Work . . . . .	45
	<b>Bibliography</b>	<b>47</b>

# List of figures

1.1	Illustration of the construction of Voronoi Cell Diagrams . . . . .	1
2.1	Example of an image from the HRF dataset showing the macula, fovea and optic disc . . . . .	6
2.2	Layers of the wall of the eye showing the region around the fovea. Drawing by Dave Schumick, Cleveland Clinic, reproduced with permission via RightsLink. . . . .	7
2.3	Healthy retinal fundus images of the region for various species of animals .	9
2.4	Illustration of the mechanisms for sprouting angiogenesis . . . . .	10
2.5	A retina displaying diabetic retinopathy from the HRF dataset . . . . .	11
3.1	Area elements on a planar image and in a spherical model of the eye . . . .	14
3.2	A diagram of the reduced eye model . . . . .	16
4.1	Annotations of regions of skeleton produced by proposed skeletonization algorithm . . . . .	23
4.2	Screenshot of the GUI interface of computer-assisted labelling program .	25
4.3	Voronoi cell diagrams in 3D and 2D . . . . .	26
5.1	Distributions of normalized areas in the HRF dataset grouped by condition .	33
5.2	Comparison of the distributions for HRF and DRIVE . . . . .	35
5.3	Relative error in individual Voronoi cell areas . . . . .	36
5.4	A comparison of area distributions computed in spherical space and on the planar geometry of the image . . . . .	37
5.5	Neovascularization simulation Voronoi diagram . . . . .	40
5.6	Neovascularization simulation graph . . . . .	41
5.7	Distributions of normalized Voronoi cell areas grouped by condition . . .	42
5.8	Distribution of normalized areas on normal placentas compared with healthy retinas . . . . .	42

- 1 Distributions of normalized areas for each individual image and the best fit  
gamma to each in the HRF dataset grouped by condition . . . . . 49

# List of tables

4.1	Statistical summary of HRF and DRIVE . . . . .	22
5.1	Fit results for data from HRF . . . . .	32
5.2	Fit results for data from DRIVE . . . . .	35
5.3	Fit parameters calculated in the 2D plane of the image . . . . .	38



# **Nomenclature**

## **Acronyms / Abbreviations**

DR Diabetic Retinopathy

DRIVE Digital Retinal Images for Vessel Extraction

HRF High Resolution Fundus database

KS test Kolmogorov-Smirnoff Test

NPDR Non-Proliferative Diabetic Retinopathy



# Chapter 1

## Introduction

### 1.1 Voronoi Analysis, Motivation and Previous Work

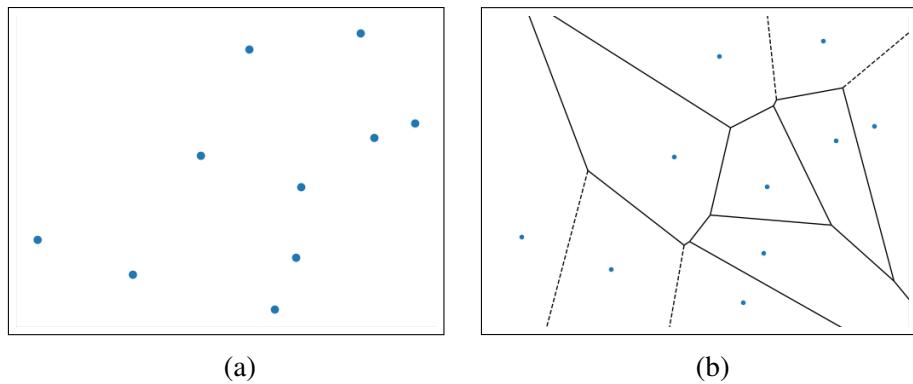


Fig. 1.1 The partitioning of 2D Euclidean space around a set of points (a) into a Voronoi diagram (b). Solid lines indicate the edges of closed regions, dashed lines indicate the edges of open regions.

**Voronoi analysis** is the partitioning of the space, as shown in Fig. 1.1, around a set of seed points  $S_i$  such that for each seed point  $S_i$  there is a corresponding convex polyhedron, **Voronoi cell**, enclosing the points in the space which are closer to  $S_i$  than any other point in the set. The set of cells is known as a **Voronoi cell diagram**.

Voronoi analysis has been used in analysing the structures of crystalline compounds [10, 20] and transport problems [3]. Recent work in biophysics has shown a correlation between risk of autism and the structure of the surface blood vessel system of the placenta using Voronoi cells centred on the midpoints of the branches [16]. In their work, Leonard et al. found scaling laws for the normalized areas of Voronoi cells on the placenta when normalized by dividing by the mean core cell area [16, 23]. Specifically, they found that

the sizes of Voronoi cells in a sample of placentas supplying children at an increased risk of autism during development were larger than those in population sample. Furthermore, when cells are labelled according to the number of branchpoints along the vessel between the vessel midpoint and the umbilical cord,  $G$ , the same scaling distribution emerged for each set of cells with the a given  $G$  label.

The discovery of scaling laws for blood vessel structure on the placenta and the surprising potential clinical implications leads to the question of whether Voronoi analysis may lead to new discoveries in other organs. The eye is an natural choice for Voronoi analysis since the blood vessels can be imaged non-invasively, rapidly and cheaply. Analysis of retinal blood vessel structure is already routine in diagnosis of diseases such as diabetic retinopathy, glaucoma and chronic kidney disease [6], as discussd in sec. 2.4, and studies have found correlations between fractal dimension and disease [8, 24]. Furthermore recent work using machine learning for segmenting blood vessels on retinal fundus images means that large numbers of datasets have become available with binary tracings [21, 28].

In this study we use the branchpoints of blood vessels as seed points for Voronoi cell analysis in order to capture the general structure of the blood vessels. The sizes of Voronoi cells within a given part of the retina gives a measure of how well the cells in that part are served by the vascular system. The fewer branchpoints, the larger the Voronoi regions, and as a result the greater the average distance of cells from the nearest blood vessel so exchange of oxygen via diffusion is limited.

## 1.2 Research Objectives

The questions this project has aimed to address are:

1. What is the distribution of areas of Voronoi cells formed using the branchpoints of retinal blood vessels as seed points?
2. When normalized, do the areas of Voronoi cells on human retinas follow the same scaling distribution? Can diseases cause significant differences between the distributions of areas?
3. Is there a significant effect on the analysis of the area distribution resulting from the curved geometry of the eye?
4. Is it possible to distinguish between the vascular networks of different organs using the distribution of areas?

## 1.3 Outline

This further sections of this thesis are structured as follows. In Ch. 2 we give an overview of the biophysics of the eye and retinal blood vessels used in later discussion. In Ch. 3 we explain the mapping of planar images to curved geometry and discuss the statistical distributions used in the analysis. In Ch. 4 we explain the methodology employed in the research, including the selection and processing of data, and Voronoi analysis in curved space. In Ch. 5 we discuss the results, considering each of the research questions above. We conclude in Ch. 6 with a summary of the findings and a discussion of further research which this project may lead to.



# Chapter 2

## Biophysics of Transport Networks on the Eye and Applications of Retinal Imaging

### 2.1 The Human Eye and Retinal Imaging Techniques

The transport networks responsible for the movement of oxygen, nutrients and metabolic waste products within the eye are blood vessels. Blood vessels emerge onto the retina as arteries from the **optic disc**, branch out into capillaries, and return to the optic disc as veins. Blood vessels grow on the surface of the retina which means that the vascular network can be imaged non-invasively. The network of retinal blood vessels, usually filtered automatically by the brain, can be seen if a small bright light is shone at the periphery of the field of view [15]. This image, known as the Purkinje tree, can be seen by the patient during examination using an ophthalmoscope.

**Fundus photography** is where a 2-dimensional representation of the retina is projected onto an imaging plane using reflected light, as shown in Fig. 2.1. It has the advantage that it selectively shows the retinal blood vessels and usually does not show blood vessels from deeper layers of the eye, unlike certain other techniques such as **fluorescein angiography**, whereby a fluorescent dye is injected into the subject's blood stream. In fluorescein angiography the dye can diffuse into deeper levels of the circulation which can be seen in the image [17]. Fluorescein images are helpful in the clinical diagnosis of diabetic retinopathy as they can show haemorrhages from damaged blood vessels [6]. However, fundus photographs are more accessible since they do not require the injection of a tracer and are taken during routine eye checkups thus they have greater ease of use in clinical contexts. In recent years, fundus imaging has become yet more accessible and images have improved in resolution thanks to the move from film to digital cameras [1].

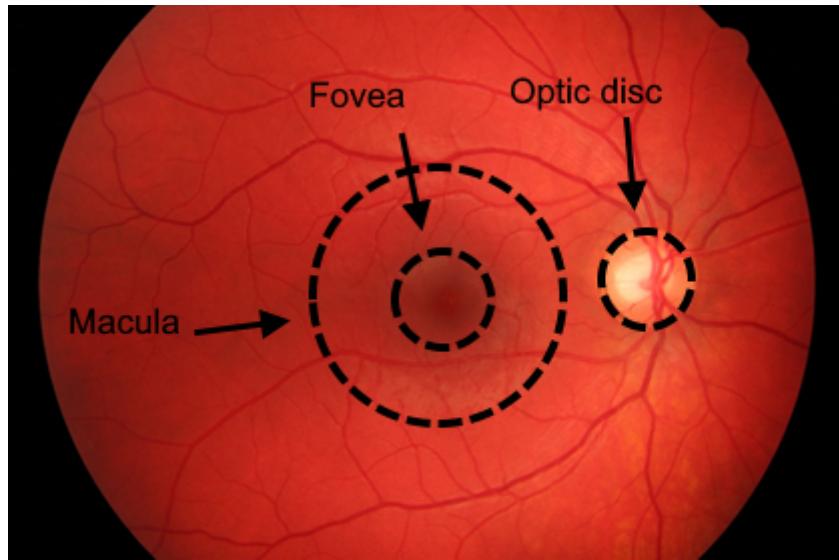


Fig. 2.1 An example of a fundus image from the HRF dataset showing the macula, fovea and optic disc.

**Tracing**, sometimes referred to as **segmentation**, of retinal images is the identification of pixels belonging to the retinal blood vessels [28]. Extensive work has been done into automated methods for segmentation of fundus images with promising results for clinical purposes [12, 21, 25, 28], however less work has been done in automated methods for separating the arterial and venous networks in the eye. The availability of tracings, and the need for research into automated methods for processing them further, has motivated the research presented here to use tracings to make new insights into the eye.

An example retinal fundus image used in this research is shown in Fig. 2.1 showing the optic disc, macula, fovea and the retinal blood vessels. The **macula** is the region of the retina responsible for high-resolution colour vision, and has a diameter of around 5.5 mm. The macula encompasses the **fovea**, a pigmented depression with a radius of 350 microns at its base [6, p. 570].

The wall of the eye is composed of a series of layers shown in Fig. 2.2. Starting from the outside, these layers are the sclera, choroid, photoreceptor layer, the deeper plexus and the primary (or superficial) plexus [4]. We are concerned in this research with the blood vessels in the primary and secondary plexuses, referred to as the **retinal blood vessels** as opposed to the choroidal blood vessels in the underlying choroid layer. In Fig. 2.2 we see the region around the fovea, the region of the retina with the highest sensitivity to light and is at the focal point of the lens of the eye when the eye is in focus. The retinal blood vessels are located on top of the photoreceptor layer, meaning that they partially obscure vision. The retinal blood vessels do not extend over the fovea, most likely because the evolutionary

advantage of greater acuity of central vision in humans was dominant over the limitation of supply of oxygen to the cells in this region.

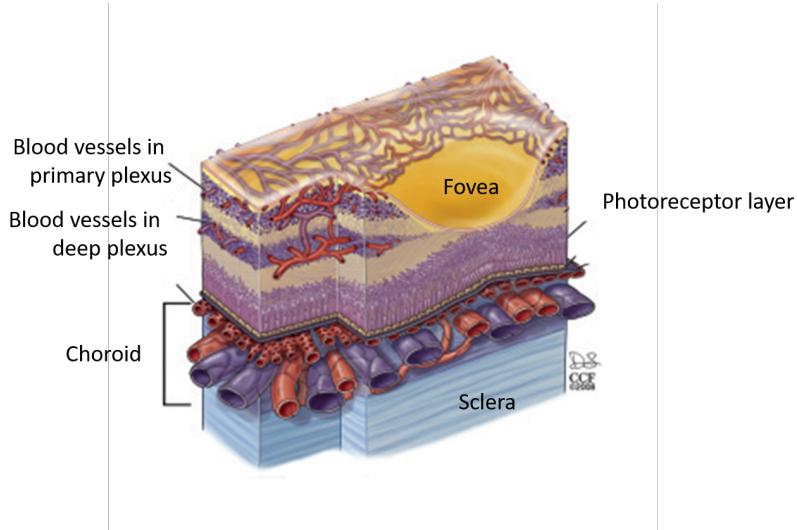


Fig. 2.2 Layers of the wall of the eye showing the region around the fovea. Drawing by Dave Schumick, Cleveland Clinic, reproduced with permission via RightsLink.

In this study we only consider the blood vessels in the primary plexus, so we remove the optic disc, where the vessels emerge from the optic nerve perpendicular to the retina, from analysis.

A natural question is whether methods for analysing the human retina can be extended to other species.

## 2.2 The Animal Eye

A remarkable observation attributed to the agricultural biologist, Max Kleiber, is that the metabolic rates of mammals scale with their body mass in a power law, specifically

$$R = M^{\frac{3}{4}} \quad (2.1)$$

where  $R$  is metabolic rate and  $M$  is body mass [29].

This tells us that the efficiency of metabolism is greater for larger mammals. Kleiber's law, as this observation is referred to, has been correlated with similar scaling laws observed in the number of other vascular factors such as heart rate and the number of generations of branches in vascular systems and indicates that there is a commonality in the mechanism by which mammals grow. For instance, the heart of a blue whale has the same structure as that

of a human heart, despite being tasked with pumping blood around a body over three orders of magnitude greater in mass [29]. In general, scaling laws can indicate whether structures arise by similar mechanisms. This brings to the fore the question whether the structure of the retinal blood vessels also scales with mass across mammals.

We considered three mammals for which retinal images are available: the equine, canine and feline eyes. Examples of healthy fundus images for these three species are shown in Fig. 2.3. In the equine eye retinal blood vessels extend only a short distance from the optic disc and there are no large blood vessels. The retina of the cat has much straighter blood vessels exiting the optic disc radially, rather than curving around the fovea as in humans. The canine retinal vascular system is most akin to the human vascular system but differs in that an opaque layer beneath the photoreceptors known as the **retinal pigmented epithelium** does not cover the entire retina, so vessels in deeper layers of the membrane are visible. From these images it is clear that the structure of the retinal vascular system varies between different mammals. Most likely since vision plays a greater or lesser part in the survival of different species the structure has been determined by evolutionary trial and error.

In summary, the structure of the blood vessel system in the eye varies in its underlying form between different species. The variety in the form of the eye between different species stands in contrast to other organs such as the heart which has the same form over a wide range of body masses. However this variety in form also means that it is not meaningful to try to find an analogous scaling law for the eye as it varies in size across various species.

## 2.3 Growth of Blood Vessels

Branches grow from existing blood vessels by two mechanisms: **sprouting angiogenesis** and **splitting** (or **intusseceptive**) **angiogenesis** as illustrated in Fig. 2.4. In the sprouting angiogenesis hypoxic cells emit Vascular Endothelial Growth Factor (VEGF), a transmitter molecule, which diffuses through the intercellular fluid to nearby blood vessels and stimulates the growth of a new branch. In splitting, or intusseceptive, angiogenesis the wall of the blood vessel grows inwards, splitting the blood vessel into two branches [2, p. 12].

In primates the initial vascular tree is formed at the fourth month of fetal development when cells known as angioblasts differentiate into the endothelial cells which line the walls of blood vessels. The retinal blood vessel network forms by sprouting angiogenesis from this network [4]. The adult retinal blood vessel pattern is established by the fifth month of development and does not usually change significantly in later life. The growth of new blood vessels in the eye which occur in adulthood are referred to as **neovascularization** and are

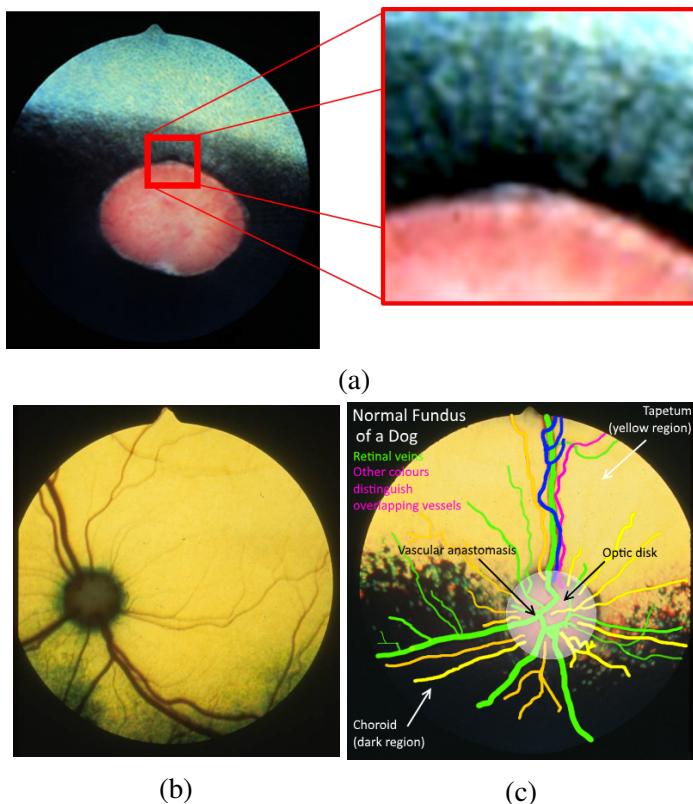


Fig. 2.3 Healthy retinal fundus images of the region around the optic disc of (a) horse, (b) cat and (c) dog. See text for a discussion of the variations in structure. Images reproduced from [22].

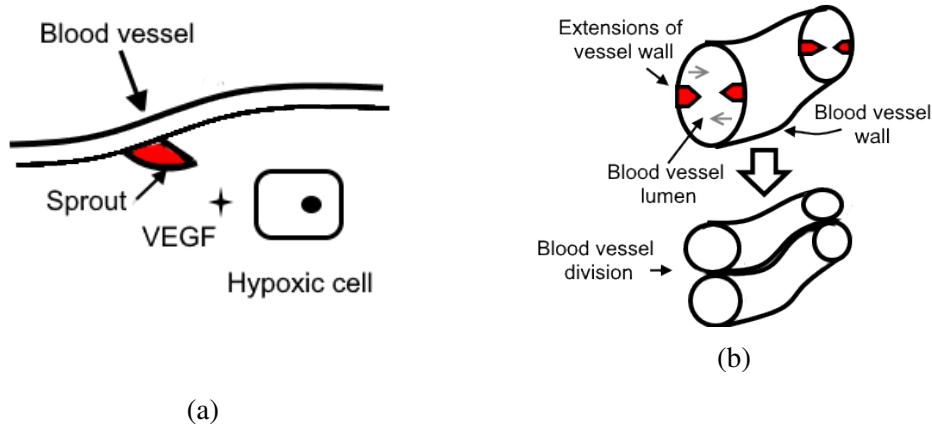


Fig. 2.4 The two mechanisms for angiogenesis. a) Sprouting angiogenesis b) Splitting (intussusceptive) angiogenesis. See the text in sec. 2.3 for an explanation of the process.

almost exclusively pathological as the blood vessels formed leak into the eye resulting in impairment or even loss of vision.

Age can result in degeneration of the pigmented layer between the retina and the choroid meaning that the choroidal blood vessels may become partially visible [6]. In the images analysed this was visible in some cases, however the choroidal vessels in fundus images appear much fainter and can be distinguished from the retinal vessels.

## 2.4 Vascular Diseases of the Eye

**Diabetic Retinopathy (DR)** is a condition which can result from diabetes wherein the smaller blood vessels on the retina become damaged by elevated blood sugar concentration. It is a severe condition which can result in partial or total loss of vision. Understanding the mechanism by which diabetic retinopathy occurs therefore has the potential to help in clinical treatment [5].

Diabetic retinopathy has two stages, in the first stage, **non-proliferative diabetic retinopathy**, blood vessels begin to leak fluids including blood cells and deposits of fats. The deposits, called **exudates**, are visible as yellow specks on the retina (Fig. 2.5 inset). The leakage of fluid into the eye causes the macula to swell and is the most common cause of loss of vision in diabetes. **Aneurysms**, localized bulging of blood vessels, which appear as red dots on fundus images can also develop. If untreated, **proliferative diabetic retinopathy (PDR)** can develop in which small blood vessels grow. These blood vessels lead to blood leakages, **haemorrhages**, producing floating objects in vision or in loss of vision. The process of new

blood vessel growth in PDR, referred to as neovascularization, is not well understood in diabetic retinopathy and there is debate over the process occurs by which it occurs [4, 19]. Diffusion-Limited Aggregation models produce patterns with similar fractal dimensions to those observed in neovascularization suggesting that it is likely limited by the diffusion of VEGF, whereas healthy growth forms in the embryo from embryonic precursor cells [8, 19].

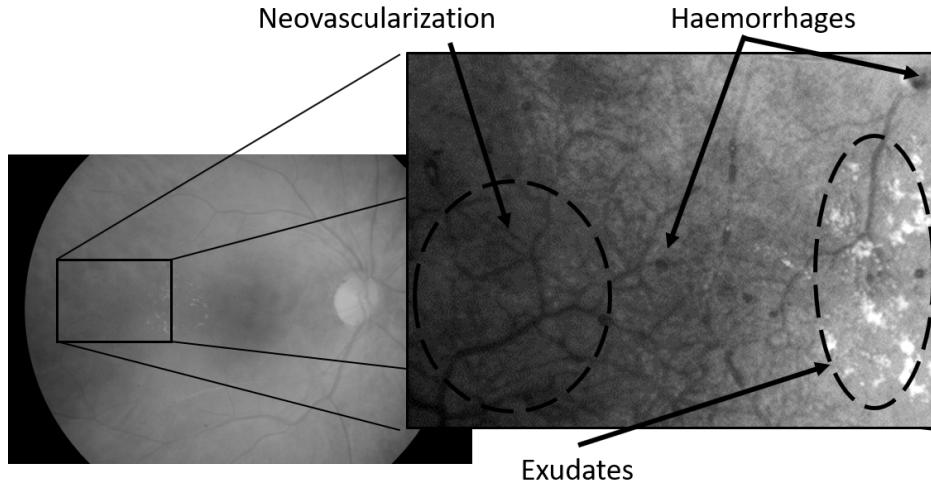


Fig. 2.5 A retina displaying diabetic retinopathy from the HRF dataset [21]. Inset: A magnified and contrast-adjusted section of the image shows haemorrhages (visible as dark spots), regions of neovascularization and exudates

**Glaucoma** is a condition in which there is damage to the optic nerve. Usually this is caused by extra pressure due to the build up of fluid in the eye. In later stages, glaucoma affects the optic disc, changing the shape of the lighter-coloured cup [17]. Diagnosis typically involves using fundus images to analyse changes in the optic disc. Typically glaucoma does not affect the structure of the retinal blood vessels, so we do not expect to see a difference in Voronoi cell structures when compared with the healthy retinal blood vessels.



# Chapter 3

## Spherical Geometry and Uses of Statistical Distributions

### 3.1 Mapping Fundus Images to Spherical Coordinates

This section gives a summary of the mathematics of mapping of length and areas between planar and spherical spaces. We introduce the reduced eye model in which the eye is modelled as being a spherical surface with a single lens and the camera is aligned so that the camera lens is on the axis of the eye. The mappings derived here are used in two parts of this research. Firstly, in performing the Voronoi analysis in spherical space in sec. 4.4 and secondly in analysing the errors in the areas as a result of the gridsize in sec. 5.4.

#### 3.1.1 Mapping Lengths in the Image to Physical Distances on the Retina

Fundus images taken using a camera are a projection of the curved surface of the retina onto a 2-dimensional plane. One problem that could arise from this is that distances toward the outside of the image are smaller than the physical distances on the retina.

When comparing lengths it is therefore necessary to make a mapping from the 2-dimensional  $(r_{\text{im}}, \phi_{\text{im}})$  plane of the image (Fig. 3.1a) to the  $(\phi, \lambda)$  surface of the eye (Fig. 3.1b).

Let positions on the image be defined by 2d polar coordinates  $(r_{\text{im}}, \phi_{\text{im}})$ , where  $r_{\text{im}}$  is the distance from the centre of the image and  $\phi_{\text{im}}$  is the angle anticlockwise from the x-axis as shown in Fig. 3.1a.

$$x = r_{\text{im}} \cos(\phi_{\text{im}}) \quad y = r_{\text{im}} \sin(\phi_{\text{im}}) \quad (3.1)$$

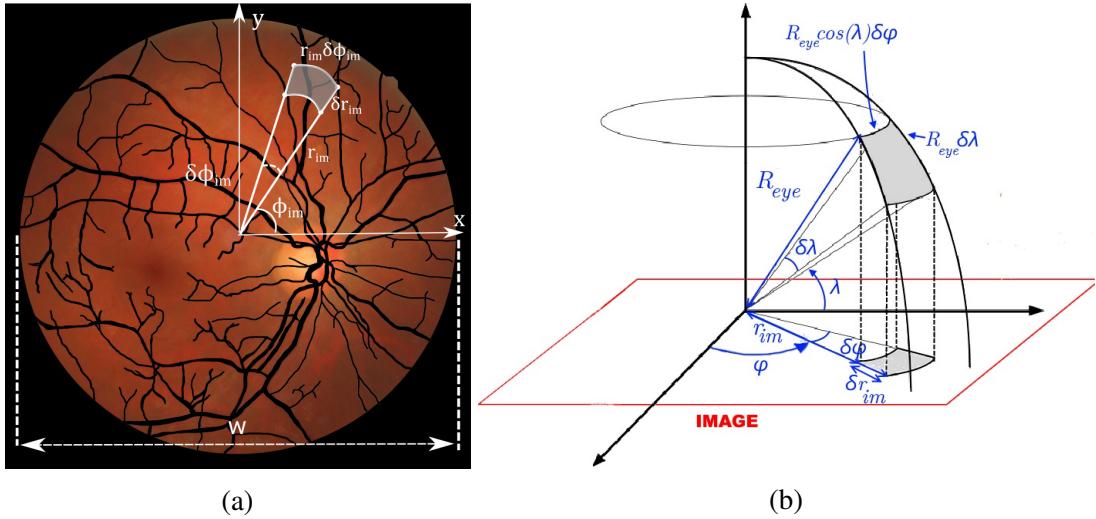


Fig. 3.1 Area elements (a) on a planar image, in a polar coordinate system, and (b) on a spherical model of the retina. Figure (a) is a tracing of an image from the FIRE dataset made by the author. Figure (b) is adapted from [astarmathsandphysics.com](http://astarmathsandphysics.com)

Let the surface of the retina be defined by the angular coordinates  $(\phi, \lambda)$  where  $\phi$  and  $\lambda$  are defined in Fig. 3.1b. Note that here  $\lambda$  is the *latitude* - the angle from the ‘equator’ of the eye as used in cartography not the angle from the axis of the eye so as to be consistent with the spherical law of cosines as in the form used later.

The surface of the eye can be approximated as a sphere<sup>1</sup> wherein the visible retinal blood vessels lie on the concave surface of the retina, meaning that all the points in image correspond to the same radial position,  $R_{\text{eye}}$ .

Referring to Fig. 3.1a, the radius of the eye  $R_{\text{eye}}$ , can be calculated from the Field of View, FOV, and the width of the photograph,  $w$ , by the following relation:

$$R_{\text{eye}} = \frac{w/2}{\sin(FOV/2)}. \quad (3.2)$$

This leaves only the two angular coordinates,  $\phi$  and  $\lambda$  (Fig. 3.1a) which are given by

$$\phi = \arctan(y/x) \quad (3.3)$$

$$\lambda = \arccos\left(\frac{r_{\text{im}}}{R_{\text{eye}}}\right). \quad (3.4)$$

In order to calculate the physical distance between two points  $x_1, y_1$  and  $x_2, y_2$  on the image we first find their positions in spherical coordinates  $(\phi_1, \lambda_1)$  and  $(\phi_2, \lambda_2)$  using eqs. 3.3 and

<sup>1</sup>Although the eye is not a perfect sphere and varies in shape between individuals, the error in length measurements on the retina resulting from this is less than 1% according to Doelemeyer et al [7].

3.4 and then use the law of spherical cosines to calculate the central angle,  $\alpha$ , between the two points [13, sec. 158]

$$\boxed{\alpha = \arccos(\sin(\lambda_1) \sin(\lambda_2) + \cos(\lambda_1) \cos(\lambda_2) \cos(|\phi_2 - \phi_1|))}. \quad (3.5)$$

The distance along the great circle between the two points is then

$$\boxed{d = R_{\text{eye}} \alpha}. \quad (3.6)$$

### 3.1.2 Mapping Areas on the Retina to Areas on the Image

We use a mapping from spherical polar coordinates on a 2-sphere to 2D polar coordinates on a plane. We use this mapping to estimate projection errors and understand the role of curvature on the analysis in sec. 5.5.

Assuming that the lens of the camera is aligned on the axis of the eye and that there is no rotational distortion to the image by the optics  $\phi = \phi_{\text{im}}$  and  $r_{\text{im}}$  is a function only of  $\lambda$ . The exact form of the function  $\lambda$  depends on the geometry of the camera and eye, which will be considered later in this section.

There are two boundary conditions

1. A point for which  $\lambda = \pi/2$  is directly above the centre of the image

$$r_{\text{im}}(\lambda = \pi/2) = 0 \quad (3.7)$$

where  $R_{\text{im}}$  is the radius of the image.

2. Any point on the retina at the edge of the field of view is at the edge of the image

$$r_{\text{im}}(\lambda = \pi/2 - FOV/2) = R_{\text{im}} \quad (3.8)$$

where  $FOV$  is the Field of View (as shown in Fig. 3.2).

The area of a surface element on a spherical surface (Fig. 3.1b) is given by

$$\begin{aligned} dA_{\text{sphere}} &= R_{\text{eye}} \cos(\lambda) \delta\phi \times R_{\text{eye}} \delta\lambda \\ &= R_{\text{eye}}^2 \cos(\lambda) \delta\phi \delta\lambda \end{aligned} \quad (3.9)$$

The area of a surface element on the image (Fig. 3.1a) is given by

$$dA_{\text{im}} = r_{\text{im}} \delta\phi_{\text{im}} \delta r_{\text{im}} = r_{\text{im}} \delta\phi \delta r_{\text{im}} \quad (3.10)$$

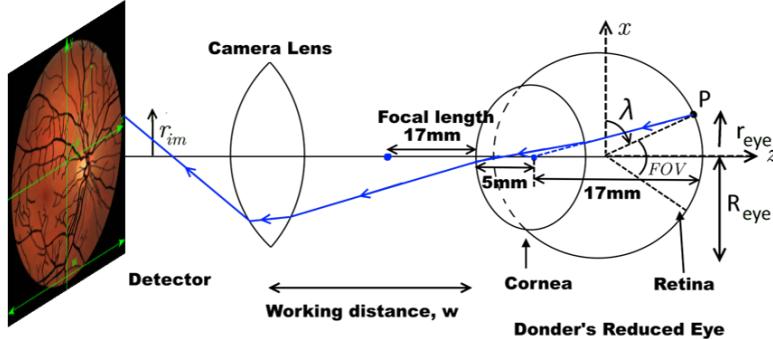


Fig. 3.2 A diagram of the optics and geometry of the reduced eye model

Thus the ratio of the surface elements is

$$\frac{dA_{sphere}}{dA_{im}} = \frac{R_{eye}^2}{r_{im}} \cos(\lambda) \left| \frac{d\lambda}{dr_{im}} \right| \quad (3.11)$$

giving the surface element in polar coordinates as

$$dA_{sphere} = R_{eye}^2 \cos(\lambda) \left| \frac{d\lambda}{dr_{im}} \right| dr_{im} d\phi_{im} \quad (3.12)$$

Hence to find the physical area on the retina corresponding to a region  $A_{im}$  on the image we perform the integral

$$A_{sphere} = \iint_{A_{im}} R_{eye}^2 \cos(\lambda(r_{im})) \left| \frac{d\lambda}{dr_{im}} \right| dr_{im} d\phi_{im} \quad (3.13)$$

The relationship between  $\lambda$  and  $r_{im}$  depends on the optical setup. We therefore build a model of the eye in order to derive a functional form for  $\lambda(r_{im})$ .

### 3.1.3 The Reduced Eye Model

In ophthalmology it is standard to use a model of the eye in which the eye consists of a spherical retina and a lens or set of lenses with empirically measured focal lengths. Such a model of the eye is referred to as a ‘reduced’ eye model because it reduces the complexity of the eye model to only its optical properties.

Here we use Donder’s Reduced Eye Model illustrated in Fig. 3.2, which consists of a single spherical refracting surface and we assume that the camera optics are such that the radial coordinate on the image,  $r_{im}$  are proportional to the distance from the optical axis.

In Donder's reduced eye model, the eye is modelled as being a sphere with a single refracting surface as shown in Fig. 3.2.

In addition we use the approximation that the distances on the image  $r_{\text{im}}$  are linearly proportional to the distances from the optical axis  $r_{\text{eye}}$  i.e.

$$r_{\text{im}} = M r_{\text{eye}} \quad (3.14)$$

where  $M$  is the magnification factor and  $r_{\text{eye}}$  is as defined in Fig. 3.2.

$$r_{\text{eye}} = R_{\text{eye}} \cos(\lambda) \quad (3.15)$$

hence

$$r_{\text{im}} = M R_{\text{eye}} \cos(\lambda) \quad (3.16)$$

So the  $\lambda$  coordinate on the surface of the retina can be expressed in terms of the distance from the centre of the image,  $r_{\text{im}}$ , as

$$\boxed{\lambda(r_{\text{im}}) = \arccos\left(\frac{r_{\text{im}}}{M R_{\text{eye}}}\right)} \quad (3.17)$$

The derivative with respect to  $r_{\text{im}}$  is

$$\begin{aligned} \frac{d\lambda}{dr_{\text{im}}} &= -\frac{1}{\sqrt{1 - \left(\frac{r_{\text{im}}}{M R_{\text{eye}}}\right)^2}} \frac{1}{M R_{\text{eye}}} \\ &= -\frac{1}{\sqrt{1 - (\cos\lambda)^2}} \frac{1}{M R_{\text{eye}}} \\ &= -\frac{1}{\sin(\lambda)} \frac{1}{M R_{\text{eye}}} \end{aligned} \quad (3.18)$$

Substituting into eq. 3.11,

$$\frac{dA_{\text{sphere}}}{dA_{\text{im}}} = \frac{1}{M} \frac{1}{\sin(\lambda)} \quad (3.19)$$

Since we only use normalized areas in the analysis, we use units wherein  $M = 1$ , giving the relationship between an spherical area element on the surface of the retina,  $dA_{\text{sphere}}$  and the corresponding region on the image in polar coordinates as

$$\boxed{\frac{dA_{\text{sphere}}}{dA_{\text{im}}} = \frac{1}{\sin(\lambda)}} \quad (3.20)$$

## 3.2 Fat-Tail Statistical Distributions

In this study we focus on two commonly occurring distributions for fitting the data: the gamma and lognormal distributions because of their ubiquity in biological sciences and Voronoi analysis. The gamma distribution emerges for the areas of Voronoi cells formed around a uniformly distributed set of points on a plane and also gives the best fit for Voronoi cells using midpoints of surface blood vessels of the placenta [16]. The lognormal distribution is another ubiquitous continuous fat-tail distribution with a similar shape to that of the gamma distribution arising frequently in the biological sciences [18].

The **gamma** distribution for a random variable  $x$  is given by

$$\text{Gamma}(y, a) = \frac{1}{\Gamma(a)} y^{a-1} e^{-y} \quad (3.21)$$

$$\text{Gamma}(x, a, b, c) = \text{Gamma}(y, a)/c$$

where  $y = (x - b)/c$ ,  $a > 0$  is the **shape** parameter,  $b$  is the **location** parameter and  $c > 0$  is the **scale** parameter. The peak probability density of the gamma distribution is located at  $x = c(a - 1)$ .

The **lognormal** distribution for a random variable  $x$  is given by

$$\text{Lognorm}(x, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{\ln(x)^2}{2\sigma^2}\right) \quad (3.22)$$

Again this can be re-expressed in the form  $\text{Lognorm}(x, a, b, c) = \text{Lognorm}(y, a)/c$  where  $y = (x - b)/c$ . The peak probability density of the gamma distribution is located at  $x = \exp(b - a^2)$  [26].

## 3.3 Comparison of Statistical Distributions

Since we are comparing non-Gaussian distributions, it is necessary to choose an appropriate measure of the goodness of fit to the data. In this project we use the Kolmogorov-Smirnoff (KS) test since it does not make any assumptions about the underlying distribution [14].

The two-sample Kolmogorov-Smirnoff test quantifies the difference between two non-parametric distributions giving two values: the K statistic and the p-value. The K value is a measure of the distance between the samples. Specifically for two sets of samples  $x_i, y_i$  the K value is the supremum of the difference between the cumulative distribution for each sample,  $C_x, C_y$  respectively,

$$K = \sup_i(|C_x - C_y|) \quad (3.23)$$

The p-value answers the question “if two samples were chosen at random from the same distribution what is the chance that the value of the Komogorov-Smirnov  $K$  statistic would be as large or larger than observed?” [11]. If  $p$  is larger than a significance level  $\theta$  we conclude that the samples come from the same distribution otherwise we conclude that the samples come from different distributions.

We use the two-sample KS test to compare the fits of the data using for a sample of  $10^7$  points sampled from the fit distribution and averaging the results over five iterations to ensure the p-values are consistent at a significance level of 5%.



# **Chapter 4**

## **Method**

### **4.1 Choice of Data**

For the purposes of this research, fundus images were chosen firstly since these images are mostly widely available as a result of active research on segmentation of vessels and secondly since images produced using techniques such as fluorescein imagery can include blood vessels from other layers of the eye such as the choroidal blood vessels. In addition, we considered images from other species to test whether scaling laws for retinal vasculature could be found across species, however it was concluded early on that the differences in structure between the eyes of mammals show this is unlikely. Two databases were selected for the purposes of our analysis: the High Resolution Fundus (HRF) and Digital Retinal Images for Vessel Extraction (DRIVE), which both provide expert tracings of the blood vessels [21, 25].

HRF consists of three sets of 15 high-resolution ( $3504 \times 2336$  pixel) images, making a total of 45 images. One set consists of images from healthy eyes, the second of images from patients with diabetic retinopathy and the third set from patients with advanced stages of glaucoma [21]. Three images in HRF show signs of neovascularization. The region in each image corresponds to a 60 degree field of view (FOV). The images were taken using a CANON CF-60 UVi equipped with CANON EOS-20D digital camera. Tracings of blood vessels made independently by three trained ophthalmologists are provided. HRF provides a sufficiently high resolution that the capillaries can be traced on the image. This make it ideal for analysis of the effects of diseases on the microvascular system.

DRIVE consists of 40 images selected at random from a population of 453 diabetic subjects between 25 and 90 years of age [25]. Images were taken with a  $45^\circ$  FOV at a resolution of  $768 \times 584$  pixels and tracings were made by three independent researchers who were instructed to mark pixels which they were at least 70% sure belonged to a vessel.

<b>HRF</b>	<b>FOV (°)</b>	<b>N</b>	$N_{core}$	$\langle N_{core} \rangle$
<b>Healthy</b>	60	15	2655	177
<b>Diabetic Retinopathy</b>	60	15	2403	160
- Non-Proliferative Diabetic Retinopathy	60	12	1748	146
- Proliferative Diabetic Retinopathy	60	3	661	220
- Neovascularized regions	N/A	3	258	86
<b>Glaucoma</b>	60	15	3150	210
<b>TOTAL</b>		<b>45</b>	<b>8208</b>	<b>182</b>
<b>DRIVE</b>	<b>FOV (°)</b>	<b>N</b>	$N_{core}$	$\langle N_{core} \rangle$
<b>Healthy</b>	45	33	3192	97
<b>Diabetic Retinopathy</b>	45	7	412	59
<b>TOTAL</b>		<b>40</b>	<b>2598</b>	<b>65</b>

Table 4.1 Statistical summary of the HRF and DRIVE datasets used, showing the field of view (FOV), number of images  $N$ , number of core cells  $N_{core}$ , and average number of core cells per image  $\langle N_{core} \rangle$ .

33 images show no signs of diabetic retinopathy and 7 show mild early stages of diabetic retinopathy. None of the images in DRIVE show symptoms of neovascularization [27].

The Fundus Image Registration dataset (FIRE) dataset was also used in preliminary analysis and manual tracing made by the author was used in testing the skeletonization algorithm. Images from the FIRE dataset are used in a number of diagrams in the text, such as in Fig. 3.1a [12].

A summary of the datasets is shown in Table. 4.1. The differences in the average number of branchpoints per image between the healthy and glaucomatous datasets in HRF appear may be a result of natural variabilities in tracings with some of the tracings showing finer blood vessels.

## 4.2 Skeletonization

In order to be able to uniquely mark the position of branchpoints we reduce the tracing to a single pixel path approximating the medial axis of each blood vessel, a process referred to as **skeletonization**. Standard skeletonization algorithms, such as used by MATLAB's `bwske1` were found to be poor at skeletonizing the tracings so a new algorithm was developed.

This skeletonization algorithm considers the eight nearest-neighbours to each gridpoint and applies a filter to determine if the pixel is on the edge of a filled region. It iterates through the image from top-to-bottom then bottom-to-top, left-to-right and right-to-left until there are no changes to the skeleton. Spurious features are removed by finding the diameter of each contiguous region of filled pixels and removing the region if the ratio of the filled to the

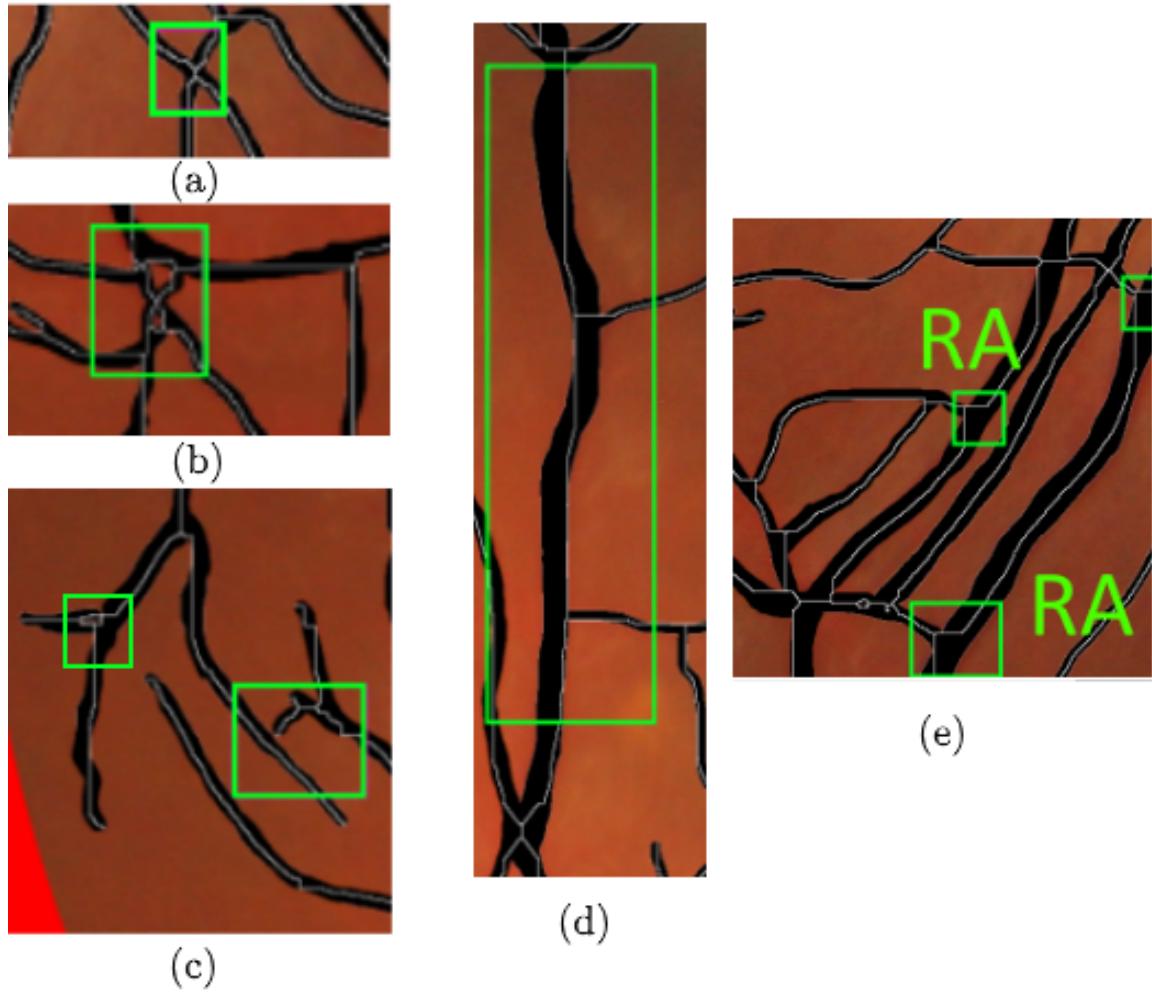


Fig. 4.1 Sections of a skeleton (white) produced by the proposed skeletonization algorithm overlaid with the tracing and original fundus image. See text on skeletonization (Sec. 4.2) for descriptions accompanying annotations.

unfilled pixels in the circular region is greater than a threshold. A threshold of 0.6 was found to be most effective.

The skeletonization algorithm developed is more effective than the standard MATLAB `bwskel` algorithm for the purposes of skeletonizing retinal blood vessels. The algorithm developed is suited for the purposes of Voronoi analysis, distinguishing between vessels in close proximity, successfully tracing complex regions with many self-loops, and accurately identifying the locations of points of intersection.

Magnified regions of a typical skeleton produced by the algorithm (white) overlaid onto the skeleton (black) and the original image are shown in Figs. 4.1a-e. Figure. 4.1a shows a

region in which the angle of intersection at a number of points (shown in red) was preserved by the skeletonization. However, as in Fig. 4.1e, the algorithm sometimes produces right-angled (RA) junctions at points of intersection particularly for thicker blood vessels, which may add a random shift to the positions of some branchpoints. In Fig. 4.1b we can see that the algorithm is effective in tracing complex loop structures. This property however lead to a drawback in that spurious single pixel holes in the tracings can lead to spurious features in the skeleton as in Fig. 4.1c. Also in Fig. 4.1c, we see the skeletonization is able to distinguish between vessels in close proximity which is important for being able to avoid falsely generating branchpoints.

There is some further work which could be undertaken on the algorithm as the skeletons have two drawbacks. Firstly, as in Fig. 4.1d, the skeleton is straightened out by the algorithm and so deviates significantly from the medial axis of the blood vessels, particularly for thick blood vessels. Secondly, as already discussed, close to points of intersection the algorithm often produces a skeleton with right angles. This means that the coordinates of the points of intersection of the skeleton can be a considerable distance from the actual point of intersection in the image. The distortion of the shape of the blood vessels means that branching angles cannot be obtained at present.

### 4.3 Computer-Assisted Labelling of Branchpoints

The skeletons are input to a program which identifies points of intersection in the image and assigns them to one of four categories: **definite overlap** (DO), **definite branchpoint** (DBP), **possible overlap** (PO) and **possible branchpoint** (PBP). The points labelled as potential overlaps and potential branchpoints are presented to a human classifier showing the corresponding region of the colour image and the tracing around the branchpoints using the GUI shown in Fig. 4.2. The human is asked to select which of the points of intersection within the skeleton are branchpoints. Large disconnected regions can be seen in some tracings as a result of small gaps in the tracing. These are corrected by connecting the end point to the closest point on the skeleton. The output of the program is a list of branchpoints which are used as the seed points for Voronoi analysis.

False negatives (mislabelling of branchpoints as points of overlap) and false positives (the reverse) can occur at two stages of the categorization. If the algorithm is too sensitive, it may assign a branchpoint to the category DO (or an overlap to the category DBP) and thus the point will not be verified manually. However, if the algorithm is not sufficiently sensitive the user is tasked with labelling a greater proportion of the points on each image, the efficiency of labelling decreases as labelling takes more time and the potential for human error increases.

Secondly, human classifier error can result from stochastic mistakes or ambiguities in the image, meaning the connectivity of the blood vessels can be interpreted in several equally likely ways.

This error was quantified by testing a sample of four images ('01\_h', '01\_dr', '01\_g' and '02\_g') from the HRF dataset by two independent classifiers. The classifiers differed in their labelling of 22 points, which is 3.4 % of the points in the sample selected. The error resulting from the classification of branchpoints, on the order of 5%, is significantly smaller than the error in areas as a result of resolution which is on the order of 11%.

Another, but far more difficult to quantify error, is the error in the tracings of the image. This error was reduced by using two datasets, HRF and DRIVE, with independently labelled images by experts working in the field of retinal image analysis.

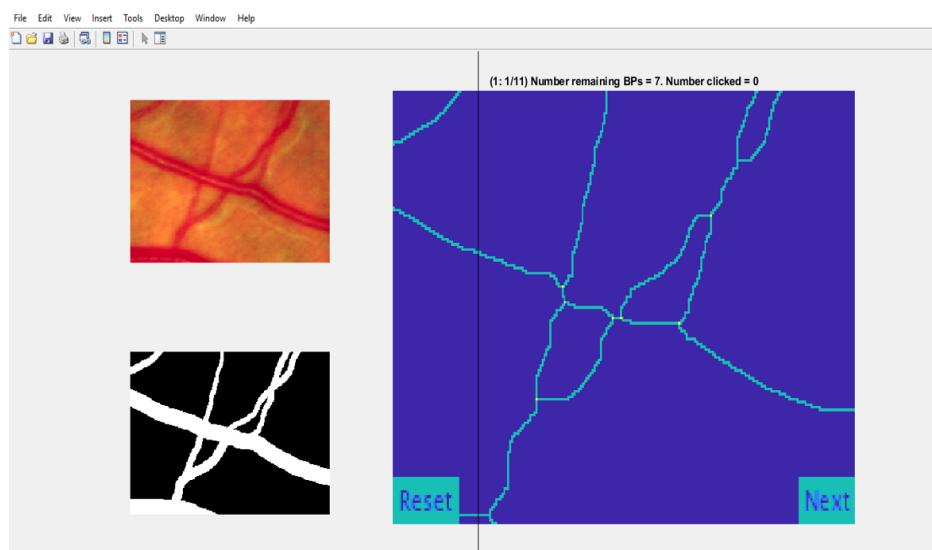


Fig. 4.2 The GUI for the computer-assisted labelling program produced as part of this research. Points of intersection identified by the computer are marked on the skeleton (right) in yellow. The human classifier is asked to select the branch points using the colour image and the tracings for that region (upper left and bottom left).

## 4.4 Voronoi Analysis in Curved Space

A fundus image is a projection of the blood vessels on the curved surface of the retina onto a flat array of pixels in a fundus camera. In order for the analysis of areas and distances on the retina to be meaningful, we must take care to represent the physical geometry of the object being imaged.

We approximate the retina as being a concave spherical surface and that the optical setup is such that there is a linear relationship between the position the image  $r_{im}$  and the distance from the optical axis,  $r_{eye}$ . Such a model is Donder's Reduced Eye illustrated in Fig. 3.2. In this model the eye is modelled as consisting of a spherical retina and a single refracting surface with a constant focal length.

Although the eye is not truly a sphere and varies in shape between individuals, the error in length measurements on the retina resulting from using a spherical rather than an elliptical eye model is less than 1% [7].

#### 4.4.1 Defining the Region of Interest

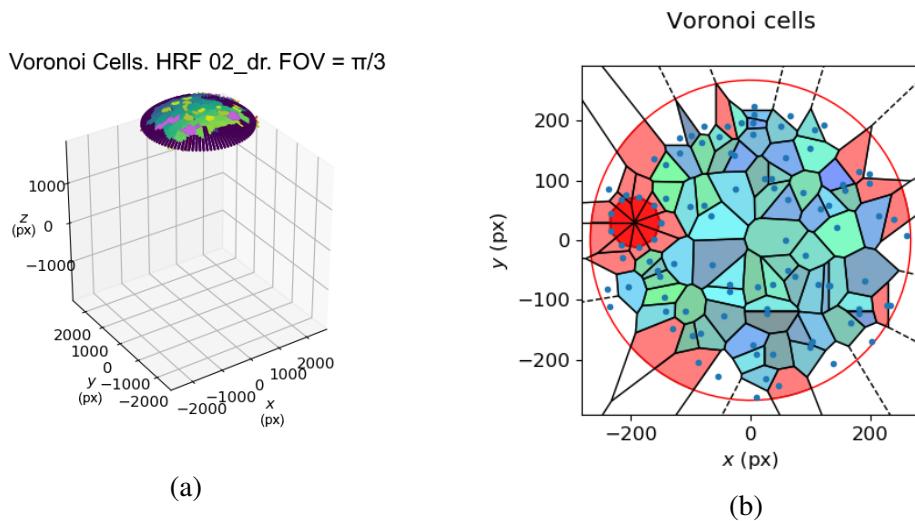


Fig. 4.3 A Voronoi cell diagram for an image from the HRF dataset (a) in 3D and (b) projected to a plane. In (a) colours indicate different Voronoi cells, with dark blue at the edge of image indicating cells beyond the FOV. In (b) colours indicate the optic disc (dark red), the core cells (blue and green), boundary cells (red), and cells beyond the field of view (white). Solid lines indicate the boundaries of closed Voronoi cells, dashed lines indicate boundaries of open regions. In this image the core cells cover 61% of area within the field of view.

For the purposes of understanding the structure of the retinal blood vessels, the region of interest is where the blood vessels run parallel to the retina and are contained within the Field of View (FOV) of the camera. We therefore exclude the optic disc, where one can only see the endpoints of the blood vessels in the optic nerve, and regions beyond the FOV. Branchpoints which are not visible because they are beyond the the field of view of the image would change the shapes and areas of Voronoi cells on the boundary of the image. For this reason we only consider Voronoi cells which do not intersect with the boundary of the image in the analysis of areas. We refer to these cells as **core Voronoi cells**.

The optic disc is the region of the retina where the optic nerve and the blood vessels serving the retina are attached. The blood vessels visible in the retinal images in the optic disc are the end points of a bundle of vessels running perpendicular to the surface being imaged, and are also points of origin for the blood vessels in the image. To reflect that the optic disc is an origin of the blood vessels but also a region with finite area with a different physical structure, we add a ring of Voronoi sites around the edge of the optic disc approximating the origin points of blood vessels and mask the optic disc area as shown in Fig. 4.3b.

#### 4.4.2 Defining the Grid

As explained in Sec. 3.1 in  $(\phi, \lambda)$  coordinates for the surface of the retina (shown in Fig. 3.1b), the angle between two points on the sphere is [13, Sec. 158]

$$\alpha = \arccos(\sin(\lambda_1)\sin(\lambda_2) + \cos(\lambda_1)\cos(\lambda_2)\cos(|\phi_1 - \phi_2|)) \quad (4.1)$$

The distance between two points on the retina is then

$$d = FR_{\text{eye}}\alpha \quad (4.2)$$

where  $F$  is a factor accounting for the geometry of the camera and lens,  $R_{\text{eye}}$  is the radius of the eye, and  $\alpha$  is the angle in radians. Since in our model this relationship is linear, we take this to be a constant factor. Since only relative areas and distances are considered, values of  $F$  and  $R_{\text{eye}}$  cancel when calculating Voronoi cell areas and determining boundaries, so we use units in which  $F = 1$  and  $R_{\text{eye}} = 1$ .

We divide the image into a uniformly-spaced  $\Phi \times \Lambda$  grid of coordinates in spherical polar coordinates

$$\phi = \{-\pi, -\pi + \delta\phi, \dots, \pi\}, \lambda = \{\lambda_{\min}, \lambda_{\min} + \delta\lambda, \dots, \pi/2\} \quad (4.3)$$

where  $\lambda_{\min} = \frac{\pi}{2} - \frac{FOV}{2}$ , and  $\delta\phi$  and  $\delta\lambda$  is the grid spacing.

For each gridpoint  $(\phi, \lambda)$  we calculate its distance from each of the seed points and assign it to the Voronoi cell of the seed point which gives the smallest distance in the metric. An example Voronoi diagram calculated using the metric of the 2-sphere above, in Fig. 4.3a, shows the masked regions in dark blue.

#### 4.4.3 Normalization and Calculation of Areas of Voronoi Cells

We remove cells outside of the region of interest and calculate the area of each Voronoi region  $A_r$  by taking the sum of the surface elements over all the grid points in the  $r^{th}$  region.

$$A_r = \sum_{\phi, \lambda} R_{\text{eye}}^2 \cos(\phi) \delta\phi \delta\lambda \quad (4.4)$$

The normalization should be such that it gives a unitless quantity, thus the normalization factor must have dimensions of area. Here we use the mean core Voronoi cell area on each image. This choice of normalization constant has previously been found to give a data collapse for areas of Voronoi cells formed using blood vessel midpoints on the human placenta [16]. Since blood vessel midpoints are simply the midpoint following the path of a blood vessel between two branchpoints, and so are equivalent measures of the network structure, the choice of the mean Voronoi cell area allows the results here to be compared with previous work.

Both the total area and the number of core cells vary between images as a result of the removal of the boundary cells which means that dividing by a factor which incorporates both of these parameters can be beneficial. Other parameters which vary between fundus photographs from different datasets are the resolution and the field of view.

We only consider core Voronoi cells since the boundary of the image is not a physical constraint on the blood vessels, and branchpoints outside of the field of view significantly affect the areas of cells on the boundary of the image. When areas of boundary cells within the field of view were included in the analysis a lognormal best fit was found to the area distributions.

Finally for each image we normalize the areas by dividing by the mean area of the non-zero cells  $\langle A \rangle$ , to produce the normalized area distribution

$$A^{\text{norm}} = A_i / \langle A \rangle \quad (4.5)$$

#### 4.5 Data and Code Validation

The code for Voronoi analysis was tested by calculating the Voronoi cell diagram for a planar random Voronoi cell diagram on a flat image where points are uniformly distributed across a plane, which is known to give a distribution of the areas given by the following Gamma distribution [9].

$$f(s) = \frac{343}{15} \sqrt{\frac{7}{2\pi}} x^{5/2} \exp\left(-\frac{7}{2}x\right) \quad (4.6)$$

In addition, visual inspection of Voronoi cells formed using a random set of test points generated by the algorithm used and the `SphericalVoronoi` package built into the Scipy library was used. The `SphericalVoronoi` library does not easily enable the use of masking and is not extensible to other geometries so was not used in this study.

The code for mapping areas was verified to check for numerical rounding errors, especially in the cosine function, by calculating the total surface area represented by a  $45^\circ$  field of view and comparing with the analytical result for a spherical cap,

$$A = \int_{\phi=0}^{2\pi} \int_{\lambda=\lambda_{\min}}^{\pi/2} \cos(\lambda) d\phi d\lambda \quad (4.7)$$

Numerical errors were not found to be significant for a  $500 \times 500$  grid and either field of view.

## 4.6 Calculation of Uncertainties in Voronoi Cell Areas

The error on the area of each Voronoi cell,  $A_i$  is defined to be half of the area of gridpoints on the boundary  $b_i$ .

$$\sigma_{A_i} = \frac{1}{2} \frac{b_i}{A_i} \quad (4.8)$$

A gridsquare with centre at  $(\phi, \lambda)$  on the sphere has an area

$$dA_{im} = R_{\text{eye}}^2 \cos(\lambda) \delta\lambda \delta\phi \quad (4.9)$$

Using the mapping from spherical to cartesian space in eq.3.11, the area in pixels is given by

$$dA_{im} = R_{\text{eye}}^2 \sin(\lambda) \cos(\lambda) \delta\lambda \delta\phi \quad (4.10)$$

so

$$A_i = \sum_i R_{\text{eye}}^2 \sin(\lambda) \cos(\lambda) \delta\lambda \delta\phi \quad (4.11)$$

Assume the probability density function for each area,  $A_i$ , is given by a Gaussian distribution

$$A_i \sim \frac{1}{b_i \sqrt{2\pi}} e^{-(\hat{A}_i - A_i)^2 / 2b_i^2} \quad (4.12)$$

where  $b_i$  is the area of the boundary,  $A_i$  is the actual area of the Voronoi cell, and  $\hat{A}_i$  is the measured area of the cell.

The total area of the cells on each image is not fixed by the field of view of the image because cells on the boundary, which have arbitrary areas, are excluded from analysis. For this reason, we can assume that the data in each bin are independent, therefore when we histogram the set  $A_i$  using  $M$  uniformly-sized bins, the probability that the  $\hat{A}_i$  falls in the  $j$ th bin of the histogram of  $A_i$  is given by

$$p_i(j) = \int_{l_j}^{u_j} \frac{1}{b_i \sqrt{2\pi}} e^{-(\hat{A}_i - A_i)^2 / 2b_i^2} d\hat{A}_i \quad (4.13)$$

where  $l_j$  and  $u_j$  are the lower and upper bounds of the bin. The mass of cells in bin  $j$  is a Bernoulli distribution

$$N(j) = \sum_i p_i(j) \quad (4.14)$$

where  $p_j$  is normalized such that  $\sum_j p_i(j) = 1$

The probability density for bin  $j$  is given by

$$p(j) = \frac{N(j)}{\sum_j N(j)} \quad (4.15)$$

The variance of the Bernoulli distribution is

$$Var(x) = \sum_i p_i(1 - p_i) \quad (4.16)$$

so the standard error on the probability density within each bin is

$$\sigma_j = \sqrt{\frac{\sum_i p_i(1 - p_i)}{N(j)}} \quad (4.17)$$

This was used to calculate error bars on distributions in the Results section.

# Chapter 5

## Results and Discussion

### 5.1 Overview

In this chapter we return to the research questions in sec. 1.2 which motivated this project. This section is structured as follows. In sec. 5.2 we begin by addressing the normalization used and whether the data collapse to a single distribution. Sec. 5.3 compares the HRF and DRIVE datasets. In sec. 5.4, we quantify the uncertainties in area measurements. Sec. 5.5 shows the results of curvature. In sec. 5.6 we consider the effects of proliferative diabetic retinopathy (PDR) further and compare with a simple model for neovascularization. Finally in sec. 5.7 we compare the results for the retina with the human placenta.

### 5.2 Data Collapse

The first two research questions asked what the distribution of areas of the Voronoi cells formed using the branchpoints of the retinal blood vessels takes, and whether, when normalized, the areas of the Voronoi cells follow the same scaling distribution.

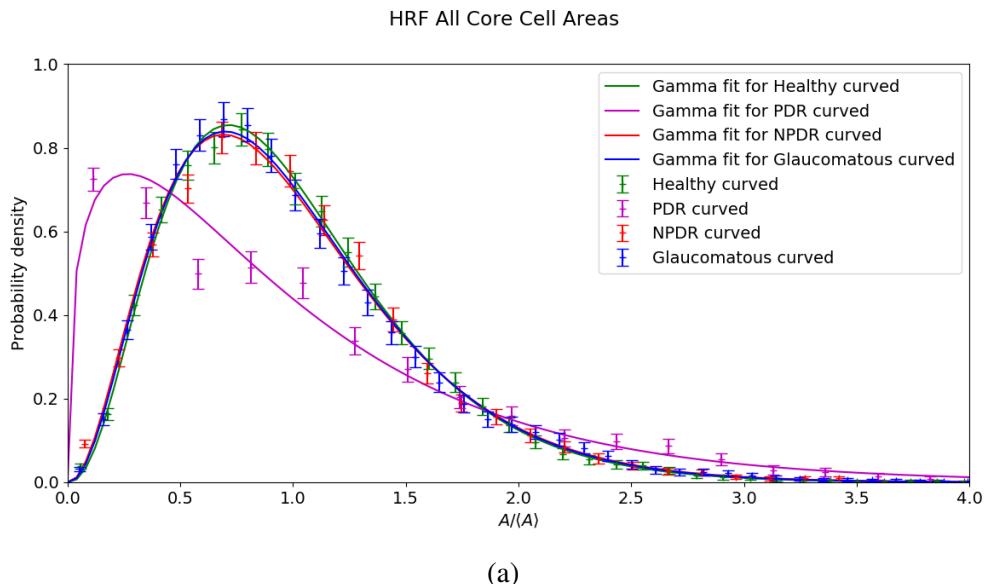
We normalize the areas by dividing by the average area of cells as described in sec. 4.4.3. We first consider the HRF dataset since HRF has the highest resolution, giving the most information about thinner blood vessels, and since it includes four conditions of retinas: healthy, glaucomatous, NPDR and PDR. We plot the distributions of normalized area distributions in Fig. 5.1 for each condition with uncertainty intervals. In appendix 1 we plot separate figures for each condition showing the distribution of areas on each image individually to check for outliers and to show the spread of data around the best fit. We found the fits for these four conditions were given by gamma distributions, with p-values from the Kolmogorov-Smirnov test (sec. 3.3) of 0.85, 0.23, 0.26 and 0.36 respectively (as shown in

<b>HRF</b>	<b>Fit parameters</b>				<b>Goodness of fit</b>	
	<b>shape</b>	<b>loc</b>	<b>scale</b>	<b>Peak</b>	<b>k value</b>	<b>p value</b>
<b>healthy</b>						
<b>Gamma</b>	3.504	0.000	0.285	0.71	0.012	0.85
<b>Lognorm</b>	0.588	0.000	0.861	0.71	0.049	0.00
<b>diabetic retinopathy</b>	<b>shape</b>	<b>loc</b>	<b>scale</b>	<b>peak</b>	<b>k value</b>	<b>p value</b>
<b>Gamma</b>	2.334	0.000	0.428	0.57	0.037	0.00
<b>Lognorm</b>	0.764	0.000	0.795	0.56	0.085	0.00
<b>NPDR</b>	<b>shape</b>	<b>loc</b>	<b>scale</b>	<b>peak</b>	<b>k value</b>	<b>p value</b>
<b>Gamma</b>	3.235	0.000	0.309	0.69	0.024	0.26
<b>Lognorm</b>	0.626	0.000	0.850	0.68	0.061	0.00
<b>PDR</b>	<b>shape</b>	<b>loc</b>	<b>scale</b>	<b>peak</b>	<b>k value</b>	<b>p value</b>
<b>Gamma</b>	1.363	0.000	0.734	0.27	0.036	0.36
<b>Lognorm</b>	1.029	0.000	0.664	0.35	0.093	0.00
<b>glaucomatous</b>	<b>shape</b>	<b>loc</b>	<b>scale</b>	<b>peak</b>	<b>k value</b>	<b>p value</b>
<b>Gamma</b>	3.321	0.000	0.301	0.70	0.018	0.23
<b>Lognorm</b>	0.6	0.000	0.854	0.70	0.036	0.00

Table 5.1 Fit results for data from HRF. p-values, calculated from the 2-sided KS test using a sample of  $10^7$  points chosen at random from the fit distribution, indicating a good fit (at the 5% significance level) are marked in bright yellow. The proximity of the peaks for healthy, NPDR and glaucomatous retinas is in sharp contrast with the much smaller peak for PDR. This is only seen by separating cases with diabetic retinopathy, which is poorly fitted by either the gamma or lognormal curves, into NPDR and PDR.

Table. 5.1). Since these larger than 0.05, at the 5% significance level we can conclude that normalizing by the average area on each image gives a good data collapse for each condition.

As can also be seen from the parameters of the gamma fits in Table. 5.1, the healthy and glaucomatous conditions follow a gamma distribution with a peak close to  $A/\langle A \rangle = 0.70$  but more significant for the glaucomatous as a result of inconsistencies in the tracing of finer blood vessels between different images which also affected the number of core cells (Table. 4.1). The spread is small for the healthy with no significant outliers. As glaucoma is a disease affecting the nervous system rather than the vascular system, we do not expect a difference in the distribution of areas of Voronoi cells formed using the branchpoints of the blood vessels as seeds.



(a)

Fig. 5.1 The distributions of normalized areas in the HRF dataset grouped by condition. Lines indicate the best fits to each and points show the results of histogramming across all retinas within each condition. Error bars are given by the standard error in each bin (sec. 5.4. Parameters and goodness of fits for each condition are shown in Table. 5.1.

For diabetic retinopathy the situation is more nuanced. As discussed in sec. 2.4, diabetic retinopathy comes in two stages. In non-proliferative diabetic retinopathy (NPDR), the first stage, the walls of capillaries are weakened resulting in microaneurysms, haemorrhages and exudates. In the second stage, proliferative diabetic retinopathy (PDR), neovascularization occurs. The distribution when all retinas with diabetic retinopathy is taken fits neither a gamma nor a lognormal distribution, however when we separate the retinas with PDR from those with NPDR, we find each individual set does follow a gamma distribution. Core cell areas on retinas with NPDR, as with glaucoma, follow the same distribution as the

healthy retinas with a peak at 0.69. However, what is striking from Fig. 5.1a is that the distributions for prolific diabetic differ significantly from that for healthy retinas, with a peak at  $A/\langle A \rangle = 0.27$ . This is a significant differences when compared even with the distribution for non-proliferative diabetic retinopathy. Voronoi analysis highlights that there is a fundamental shift in the structure of the blood vessels that occurs in the transition from non-proliferative to proliferative diabetic retinopathy.

This difference is discussed in greater depth in sec. 5.6 in which we model the effects of neovascularization to better understand what caused this shift.

### 5.3 Comparison of DRIVE and HRF data

To verify the results of the analysis on HRF we compared with the DRIVE dataset. The DRIVE dataset also includes validated manual tracings for all images and so was selected for this research.

Three parameters of DRIVE differentiate it from HRF. Firstly it consists entirely of fundus images from patients diagnosed with diabetes. 33 of the 40 images were labelled as showing no signs of diabetic retinopathy and 7 as showing NPDR. None show signs of PDR. Secondly, the resolution of the DRIVE dataset is lower and as a result the tracings do not distinguish smaller blood vessels. This also means that there are fewer branchpoints, and the number of core cells per image in DRIVE is much smaller: on average there are 65 core cells per image compared with 182 in HRF (Table. 4.1). Furthermore, the field of view for DRIVE is smaller (45 deg as opposed to 60 deg for HRF) meaning less of the peripheral retina is included (Fig. 2.1).

Fig. 5.2 shows the distributions for the healthy and NPDR conditions. As a result of there being fewer core cells with NPDR in DRIVE when compared with HRF (412 compared with 1748) there is a greater spread for the NPDR DRIVE data.

For the healthy data we can conclude that HRF and DRIVE, despite differences in the field of view do give the same distribution. The distribution for DRIVE healthy closely follows the distribution for retinas with NPDR in HRF, which we concluded give the same distribution as the healthy in the previous section.

The best fits to NPDR data differ significantly between HRF and DRIVE with a difference in the positions of the peaks of 0.06 (Table. 5.1 and Table. 5.2). This may partly be explained by the larger spread of data since a KS test between sets of NPDR cells in HRF and DRIVE gives a p-value of 0.38. As a result of the small sample of NPDR cells in DRIVE we cannot conclude that the data for the NPDR category agree.

The data are not sufficient to say whether the branching structure scales across the retina. An analysis of the network topology is required so as to number the cells by the number of the branchpoints along the blood vessel from the optic disc. Since the data only contain the positions of branchpoints, and not the relationships between them, further processing to distinguish connectivity of blood vessels is required to make a conclusion. Determining connectivity automatically is a more challenging problem than distinguishing between overlaps and branchpoints, and is an avenue for further work.

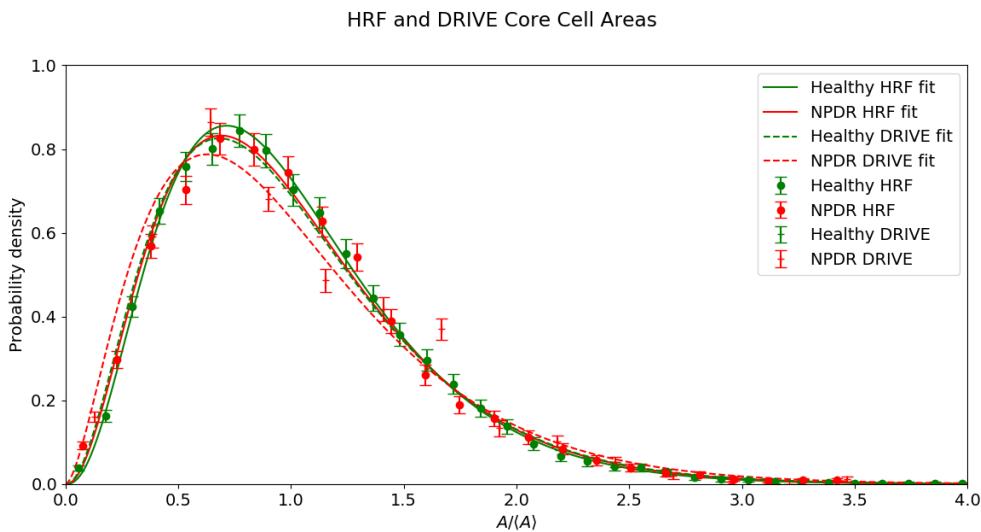


Fig. 5.2 The distributions of normalized areas in the HRF and DRIVE datasets grouped by condition with the best fit to each. The spread for the NPDR data in DRIVE is larger than for HRF as a result of a smaller sample of areas.

DRIVE	Fit parameters			peak	Goodness of fit	
	shape	loc	scale		k value	p value
<b>healthy</b>						
<b>Gamma</b>	3.151	0.000	0.317	0.68	0.021	0.30
<b>Lognorm</b>	0.651	0.000	0.846	0.65	0.058	0.00
<b>NPDR</b>	shape loc scale			peak	k value	p value
<b>Gamma</b>	2.692	0.000	0.371	0.63	0.042	0.45
<b>Lognorm</b>	0.746	0.000	0.821	0.57	0.081	0.01

Table 5.2 Fit results for data from DRIVE. p-values, calculated from the 2-sided KS test using a sample of  $10^7$  points chosen at random from the fit distribution, indicating a good fit (at the 5% significance level) are marked in bright yellow. See text in sec. 5.2 for details.

## 5.4 Uncertainties in Area Measurements

Using the method in sec. 4.6, we find that the distribution of the error on the area of each cell,  $\sigma_{A_i}$ , as a function of  $A^{norm}$  gives a fat-tail distribution with an average error for a  $500 \times 500$  grid of 11% as shown in Fig. 5.3. The only cells with relative errors of 1 (meaning all gridsquares were on the boundary of the Voronoi cell) were found in regions of neovascularization.

In order to see how this affects the significance of the differences between different datasets, we plot the probability density and the confidence interval, which we define to be one standard deviation (from eq. 4.17) of the value of each bin (eq. 4.15). We can see that the fits for healthy and glaucomatous lie within one standard deviation of the mean across the distribution (Fig. 5.4).

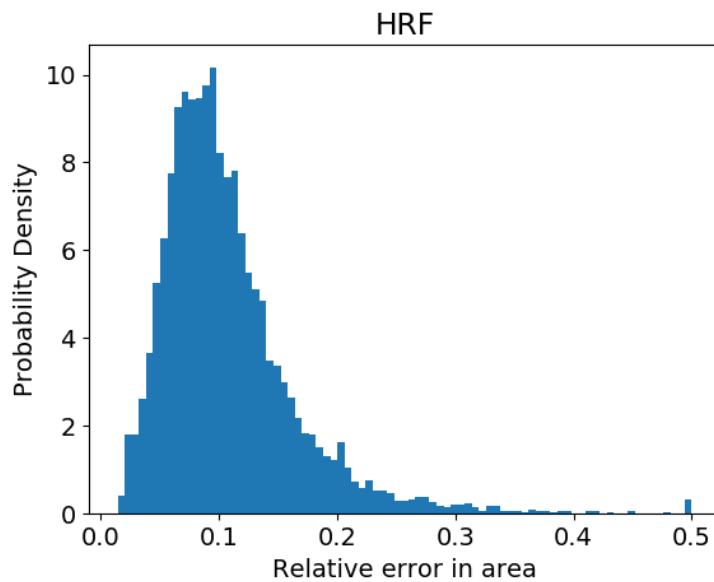


Fig. 5.3 The relative error in individual Voronoi cell areas across the HRF dataset at a grid size of  $500 \times 500$ . The mean relative error is 11%. The only cells with relative errors of 1 (meaning all gridsquares were on the boundary of the Voronoi cell) were found in regions of neovascularization.

## 5.5 Effects of Curvature

What effects does the geometry of the surface of the retina have on the analysis? Using eq. 3.11, at the edge of the field of view of the HRF dataset ( $\pi/3$ ), area elements on the image are reduced to  $\sin(\pi/2 - \pi/6) = 86.6\%$  of their original size as a result of the mapping. Thus

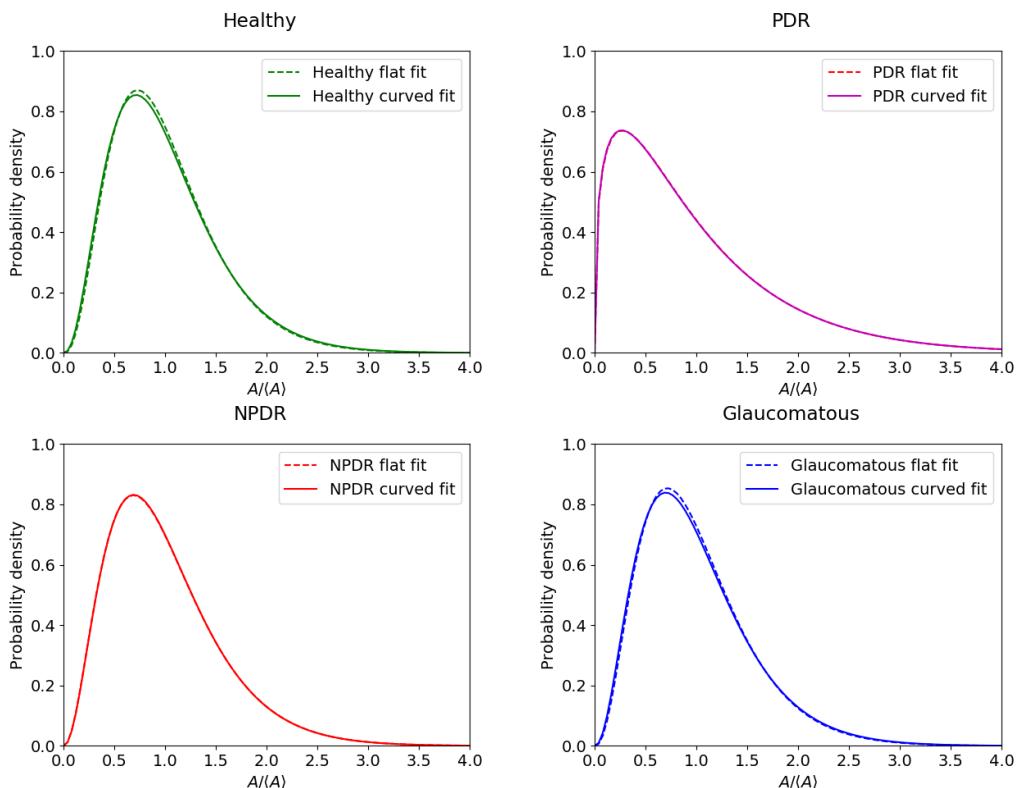


Fig. 5.4 Fits to the distributions of Voronoi cell areas for the healthy, PDR, NPDR and glaucomatous conditions of retinas in the HRF dataset when analysed in curved space (solid) and on the flat surface of the image (dashed). p-values for the fits are given in table 5.1.

HRF planar analysis	Fit parameters				Goodness of fit		HRF curved	
healthy	shape	loc	scale	peak	k value	p value	peak	p value
<b>Gamma</b>	3.245	0.00	0.308	0.69	0.012	0.77	0.71	0.85
<b>Lognorm</b>	0.593	0.00	0.851	0.77	0.035	0.00	0.71	0.00
PDR	shape	loc	scale	peak	k value	p value	peak	p value
<b>Gamma</b>	1.372	0.00	0.729	0.27	0.036	0.35	0.27	0.36
<b>Lognorm</b>	1.016	0.00	0.666	1.42	0.092	0.00	0.35	0.00
NPDR	shape	loc	scale	peak	k value	p value	peak	p value
<b>Gamma</b>	3.218	0.00	0.311	0.69	0.028	0.15	0.69	0.26
<b>Lognorm</b>	0.601	0.00	0.849	0.78	0.051	0.00	0.68	0.00
glaucomatous	shape	loc	scale	peak	k value	p value	peak	p value
<b>Gamma</b>	3.122	0.00	0.320	0.68	0.016	0.29	0.70	0.23
<b>Lognorm</b>	0.599	0.00	0.845	0.78	0.030	0.00	0.70	0.00

Table 5.3 Fit parameters calculated in the 2D plane of the image. p-values, from the 2-sided KS test indicating a good fit at the 5% significance level are marked in bright yellow. See text in sec. 5.2 for details. To the right are the peak and the p-values of the fit calculated when using areas in curved space. A plot of the best fits in curved and planar geometry are shown in Fig. 5.4.

we might expect the areas of cells toward the periphery of the image to be reduced as a result of the projection and hence a potentially significant effect on the area distribution.

When we compare the results of a simple analysis in the flat image plane with the spherical model of the eye described in sec. 4.4, we see from the fit parameters (Table. 5.3) that the curvature of the eye does cause a small shift to  $A^{norm}$ . However, plotting the two distributions for each condition on the same set of axes, as in Fig. 5.4, we see that this shift is relatively small when compared with the error margins. This is in line with the above heuristic argument that the peak is shifted to smaller normalized areas since the regions towards the periphery of the image are projected onto a smaller number of pixels, while the areas of cells close to the optical axis are unchanged by the projection.

Though this analysis is relatively simple for the retina because it is spherical, the surfaces of other organs have more complex geometry than the eye. This analysis could be extended to cover surface blood vessels in any geometry in which the metric can be calculated.

## 5.6 Effects of Neovascularization

As explained in sec. 5.2, the data for DR is not fit by either the lognormal or gamma distributions. When we separate the retinas with NPDR from those with PDR we find that the distribution for NPDR closely follows the healthy distribution both falling within the

same uncertainty interval, however, as can be seen in Fig. 5.4, the distribution for PDR is significantly different. 661 core Voronoi cells were used to produce a distribution for retinas with PDR (Fig. 5.7a) which peaks at a much smaller normalized area of 0.27 compared with 0.71 for healthy retinas. As shown in Table. 5.1, there is a strong goodness of fit between the data in the PDR dataset and a gamma distribution. This is a result of the formation of new branchpoints by the process of neovascularization. Neovascularization occurs in localized regions on the eye where blood vessels have become so damaged by DR that receptor cells have become hypoxic. As a result the number of smaller Voronoi cells increases from an average of 177 per image to 220 on average across the three neovascularized images, but these are added anisotropically across the image.

The effects of adding cells anisotropically to the retina on the Voronoi cell distribution are not immediately clear. The number of small cells may increase, but also the mean cell area must decrease which could act to reduce the change in the normalized cell area. In order to explain what may cause the observed shift in the peak we can use a simulation to explore the potential effects of adding branchpoints at random.

In the neovascularization simulation we begin with a Voronoi diagram with  $N_{init}$  sites, representing branchpoints, uniformly spaced on the image (fig.5.5a). We then add points uniformly at random and successively to one quarter of the image to represent the points at which new branches are formed in neovascularization and calculate the Voronoi cell distribution after each addition.  $S$  simulations are performed using different initial seed points to form an **ensemble** of images. Figure. 5.5 shows a simulation beginning with  $N_{init} = 100$  points, showing Voronoi cell diagrams for  $N_{sites} = (101, 150, 200, 250)$  points. The upper bound on the number of sites per image was 250 points to represent the most sever PDR.

The peak normalized area after adding  $t$  points was recorded  $A_{peak}^{norm} = (A/\langle A \rangle)_{peak}$  for  $t = 1, \dots, N$ . We then take the average peak area across the ensemble at each  $t$  point

$$\bar{A}_{peak}^{norm}(t) = \frac{1}{S} \sum_{m=1}^S A_{peak,m}^{norm} \quad (5.1)$$

where  $m$  labels the  $m^{th}$  simulation.

The uncertainty on  $\bar{A}_{peak}^{norm}$  is simply the standard deviation across the ensemble of simulations

$$\sigma(t) = \sqrt{\sum_{m=1}^S \frac{A_m^{norm}(t)^2}{S} - \left( \frac{\sum_{m=1}^S A_m^{norm}}{S} \right)^2} \quad (5.2)$$

where  $m$  denotes the  $m^{th}$  realization in the ensemble of Voronoi diagrams.

As neovascularization progresses a simulation using an ensemble of 100 realizations (Fig. 5.6) shows that the peak normalized area decreases. This simple model agrees with the observed decrease of the peak normalized area within neovascularized regions when compared with healthy retina or retina with NPDR and explains that the anisotropic addition of Voronoi cells across the image causes the shift to the distribution. Interestingly, the effect saturates when the number of sites added approaches the original number of sites.

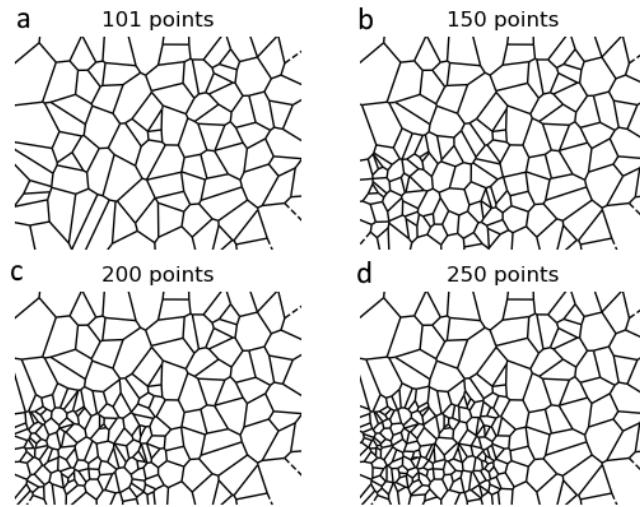


Fig. 5.5 An example of a simple computer simulation of neovascularization beginning with  $N_{init} = 100$  points showing Voronoi cell diagrams for  $N_{sites} = (101, 150, 200, 250)$

To investigate whether the center of the region of neovascularization in effect acts as another point of origin like the optic disc, we calculated the area distribution of cells contained within neovascularized regions. We included the boundary cells within this region, but removed the areas of cells which extend beyond the region of neovascularization.

The distribution for the neovascularized regions (Fig. 5.7b) is skewed even more strongly from the healthy distribution than the PDR. The very different structure produced, as quantified by the Voronoi area distribution, suggests that neovascularization relies upon a different mechanism to growth during ordinary development.

## 5.7 Comparison between the Retina and Placenta

Previous work by [16] used Voronoi cells formed around the midpoints of blood vessels to find a data collapse for the placenta. In their study they found a best fit to 3150 cells from a

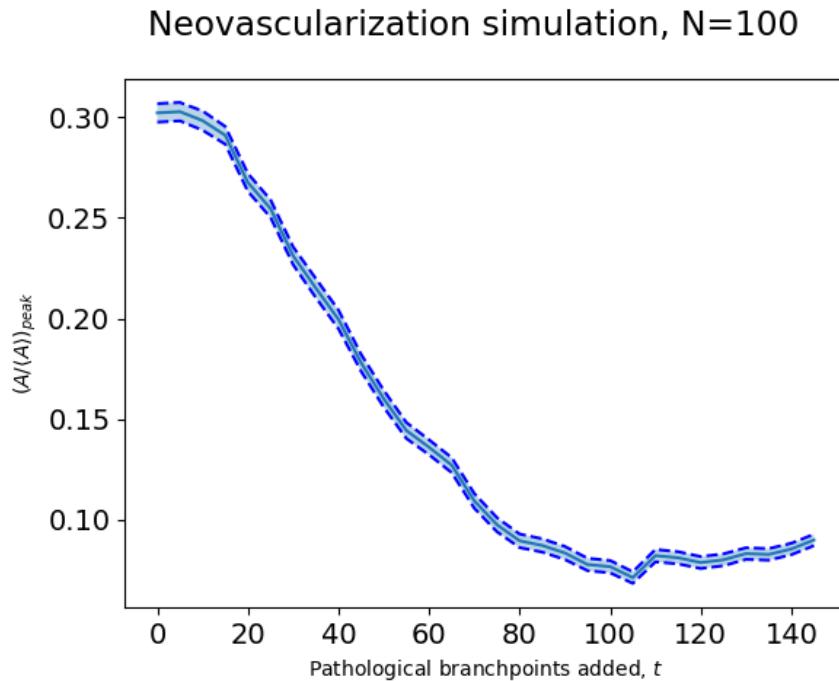


Fig. 5.6 A plot of peak normalized area as branchpoints are added in the neovascularization simulation presented in sec. 5.6 using an ensemble of  $S = 100$  realizations and starting with  $N_{init} = 100$  sites. The solid line shows the moving average using a window width of 5, dashed lines show the 68% confidence interval  $\widehat{\sigma}(t) = \frac{1}{\sqrt{5}} \sum_{j=t}^{t+5} \sigma(j)$ . Figure produced by code available in the online supplementary material.

non pre-selected population study of placentas was given by a gamma distribution peaking at a greater normalized area than a population identified as being at a greater risk of autism.

Using new data of normal (un-diseased) placentas provided by Aytac Kilic and Farhan Chatha and applying the constraint on the location parameter used for the retinal data, we find a distribution for the placentas peaking at  $A^{norm} = 0.67$ .

The Voronoi cell areas on the placenta and the retina are both fitted by gamma distributions with strong goodnesses of fit, see Fig. 5.8. However, the shape of the distributions for the retina differ from that of the placenta. The peak area is much smaller for the placenta than for the retina which means that the area served by a given blood vessel in the retina is larger in the eye than on the placenta. The placenta is a transfer organ which serves the purpose of trying to maximise the exchange of nutrients between the mother and the foetus whereas the retina is a sensory organ wherein the vascular system is constrained by the fact that blood vessels partially obstruct vision.

The following evolutionary argument may explain why the number of branchpoints in a given region should be smaller on the retina than on the placenta and hence the differences in

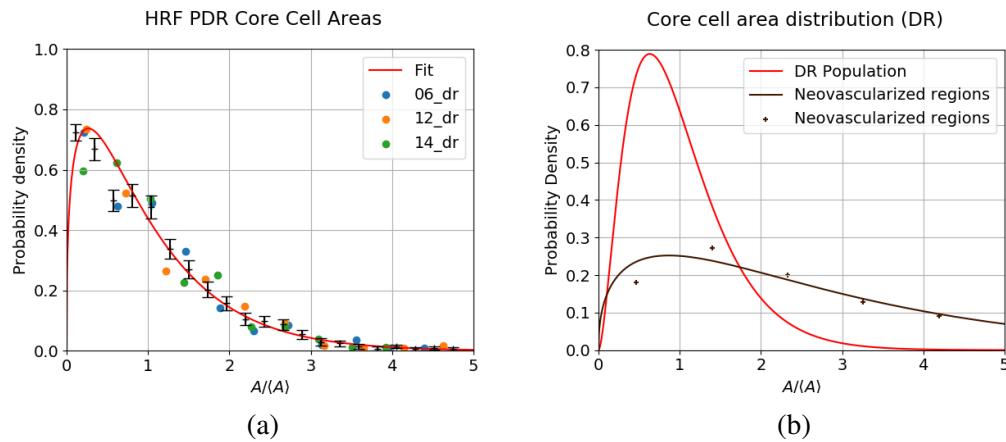


Fig. 5.7 The distributions of normalized Voronoi cell areas for (a) all cells within images showing symptoms of PDR, and (b) cells within the regions of neovascularization. Individual colours in (a) indicate different images. (b) also shows the best fit to the DR population. Error bars show the standard deviation in each bin computed using eq. 4.17. The y-axis is the probability density for each individual image. The legend shows image file names from the HRF dataset. The areas were calculated in curved geometry using a  $500 \times 500$  grid and normalized by dividing by the mean area of Voronoi cells on each image  $\langle A \rangle$ .

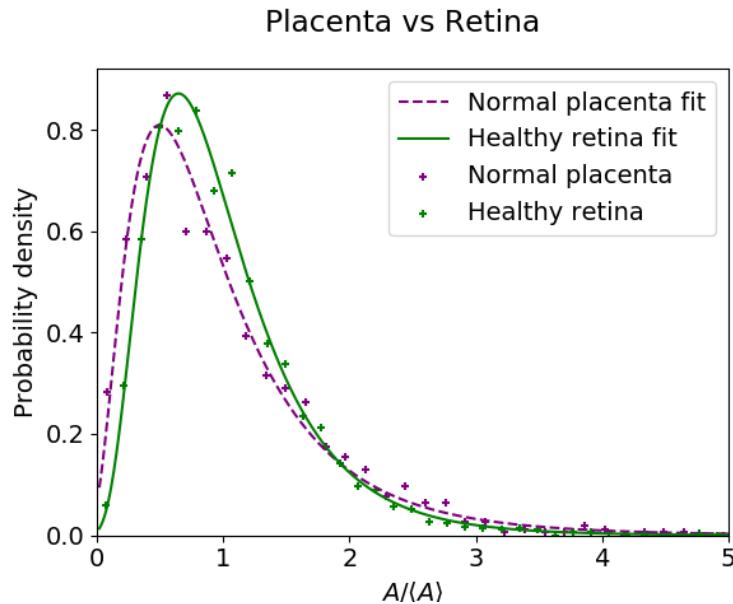


Fig. 5.8 The distributions of Voronoi cell areas on normal placentas and healthy retinas. Also shown are lognormal fits. Goodness of fit was calculated using the KS test giving  $p = 0.84$  and  $p = 0.52$  respectively. Data for placentas were provided by Aytac Kilic and Farhan Chatha.

the peak normalized areas. On the placenta, the more blood vessels the larger the surface area

for exchange of nutrients to the foetus and so the more likely the foetus will develop fully and healthily and pass on the genes of the mother. In contrast, for the eye more blood vessels may impede vision so there is a cost to building too many blood vessels. The difference in Voronoi cell structures reflects the distinction between the vascular systems of organs with different functions, and indicates that there is no universal scaling law for branching of surface vasculature that applies across different organs even within a given species. However it is possible that Voronoi analysis could reveal more subtle information about the functions of surface vasculature on different organs, and it would be useful to extend the analysis to other organs to determine how function and blood vessel branching structure are related.



# Chapter 6

## Conclusion and Future Work

The branchpoints of retinal blood vessels, identified using a novel machine-assisted labelling program developed as part of this project, were used as seed points for a spherical Voronoi cell analysis. Using two independent databases consisting of retinal fundus images from healthy, diabetic and glaucomatous patients, we were able to distinguish retinas with proliferative diabetic retinopathy (PDR) and found a strong shift in the peak of the distribution from the non-proliferative (NPDR) to the proliferative form of diabetic retinopathy.

The areas of the Voronoi cells were found to collapse to a gamma distribution for both healthy and diseased retinas with highly significant results from the KS test. The healthy, NPDR and glaucomatous conditions were found to all collapse to the same scaling distribution with a peak at  $A^{norm} = 0.70 \pm 0.01$  within one standard error of the mean. This compared with PDR which had a peak at a much smaller normalized area  $A^{norm} = 0.27$ . Curvature resulted in a small systematic shift in the positions of the distributions though the effect was not significant for the fields of views in the data.

A simple model for neovascularization in which we add branchpoints to a small region of the image reproduces the qualitative shift in the peak in proliferative diabetic retinopathy. When we compare the Voronoi cell area distributions for healthy retinas and placentas, we find that the peak normalized cell area is much larger for the retinas than for the placentas which indicates that differences in the functions of the two organs is reflected in the structure of the surface blood vessels.

### 6.1 Limitations and Future Work

Since visual inspection of retinal fundus images easily enables diagnosis of conditions, Voronoi analysis may not necessarily have immediate clinical applications, however the distribution of Voronoi cell areas may give information about the processes of pattern

formation, and hence may enable better understanding of angiogenesis in neovascularization. More advanced models of neovascularization using the connectivity of blood vessels and more images containing regions of neovascularization may make it possible to determine the mechanism by which neovascularization takes place in PDR.

Aside from further modelling of neovascularization, an exciting direction for further research is to investigate how the structure of blood vessels in the eye compares with other organs such as the brain using the methods for Voronoi analysis on non-planar geometries developed here. Determining the connectivity of blood vessels would enable study of whether there are scaling laws across different levels of vasculature on the retina. Furthermore, improving methods for skeletonization of retinal blood vessel images may open up research on whether scaling laws apply to other properties of the retinal vascular system, such as branching angles.

# Bibliography

- [1] Abramoff, M. D., Garvin, M. K., and Sonka, M. (2010). Retinal Imaging and Image Analysis. *IEEE Reviews in Biomedical Engineering*, 3:169–208.
- [2] Adair, T. H. and Montani, J.-P. (2010). *Overview of Angiogenesis*. Morgan & Claypool Life Sciences.
- [3] Amores, D. and Vasardani, M. (2010). Movement pattern analysis using Voronoi Diagrams.
- [4] Anand-Apte, B. and Hollyfield, J. (2010). Developmental Anatomy of the Retinal and Choroidal Vasculature. In *Encyclopedia of the Eye*, pages 9–15. Elsevier.
- [5] Boyd, K. (2018). What Is Diabetic Retinopathy?
- [6] Denniston, A. K. O. and Murray, P. I. (2018). *Medical retina*. Oxford University Press.
- [7] Doelemeyer, A. and Petrig, B. L. (2000). Estimation of projection errors in human ocular fundus imaging. *Ophthalmic and Physiological Optics*, 20(2):112–118.
- [8] Family, F., Masters, B. R., and Platt, D. E. (1989). Fractal pattern formation in human retinal vessels. *Physica D: Nonlinear Phenomena*, 38(1):98–103.
- [9] Ferenc, J.-S. and Néda, Z. (2007). On the size distribution of Poisson Voronoi cells. *Physica A: Statistical Mechanics and its Applications*, 385(2):518–526.
- [10] Fukunaga, T., Itoh, K., Otomo, T., Mori, K., Sugiyama, M., Kato, H., Hasegawa, M., Hirata, A., Hirotsu, Y., and Hannon, A. C. (2006). Voronoi analysis of the structure of Cu–Zr and Ni–Zr metallic glasses. *Intermetallics*, 14(8):893–897.
- [11] GraphPad (2010). Interpreting results: Kolmogorov-Smirnov test.
- [12] Hernandez-Matas, C., Zabulis, X., Triantafyllou, A., Anyfanti, P., Douma, S., and Argyros, A. A. (2017). FIRE: Fundus Image Registration dataset. *Journal for Modeling in Ophthalmology*, 1(4):16–28.
- [13] Kells, L. M., Kern, W. F., and Bland, J. R. (1940). *Plane And Spherical Trigonometry*. McGraw Hill Book Company, Inc.
- [14] Kirkman, T. W. (1996). Kolmogorov-Smirnoff Test.
- [15] Labib, B. A. (2017). Reflecting on Entoptic Phenomena. *Review of Optometry*.

- [16] Leonard, A. S., Lee, J., Schubert, D., Croen, L. A., Fallin, M. D., Newschaffer, C. J., Walker, C. K., Salafia, C. M., Morgan, S. P., and Vvedensky, D. D. (2017). Scaling of the surface vasculature on the human placenta. *Physical Review E*, 96(4):040401.
- [17] Lim, A. S. M., I Constable, and Tien Yin Wong (2004). Chapter 6: Ocular Manifestations of Systemic Diseases. In *Colour Atlas of Ophthalmology*. World Scientific, 5 edition.
- [18] Limpert, E., Stahel, W. A., and Abbt, M. (2001). Log-normal Distributions across the Sciences: Keys and Clues. *BioScience*, 51(5):341.
- [19] Masters, B. R. (2004). Fractal Analysis of the Vascular Tree in the Human Retina. *Annual Review of Biomedical Engineering*, 6(1):427–452.
- [20] Meijering, J. L. (1953). Interface area, edge length, and number of vertices in crystal aggregates with random nucleation. *Philips Res. Rep.*, 8:270–290.
- [21] Odstrcilik, J., Kolar, R., Budai, A., Hornegger, J., Jan, J., Gazarek, J., Kubena, T., Cernosek, P., Svoboda, O., and Angelopoulou, E. (2013). Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database. *IET Image Processing*, 7(4):373–383.
- [22] Schoster, J. (2006). Normal Fundus and Variations in the Dog, Cat and Horse.
- [23] Schubert, D. (2015). Machine Learning of Transport Networks in the Human Placenta. page 78.
- [24] Sng, C. C. A., Sabanayagam, C., Lamoureux, E. L., Liu, E., Lim, S. C., Hamzah, H., Lee, J., Tai, E. S., and Wong, T. Y. (2010). Fractal analysis of the retinal vasculature and chronic kidney disease. *Nephrology Dialysis Transplantation*, 25(7):2252–2258.
- [25] Staal, J. J., Abramoff, M. D., Niemeijer, M., Viergever, M. A., and Ginneken, B. v. (2004). Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509.
- [26] The\_Scipy\_community (2019). The `scipy.stats.lognorm` function.
- [27] van Ginneken, B. (2019). Diabetic Retinopathy images in the DRIVE dataset.
- [28] Vostatek, P., Claridge, E., Uusitalo, H., Hauta-Kasari, M., Fält, P., and Lensu, L. (2017). Performance comparison of publicly available retinal blood vessel segmentation methods. *Computerized Medical Imaging and Graphics*, 55:2–12.
- [29] West, G. (2017). *Scale: The Universal Laws of Life and Death in Organisms, Cities and Companies*. Weidenfeld and Nicolson.

# Appendix 1 : HRF Area Distributions

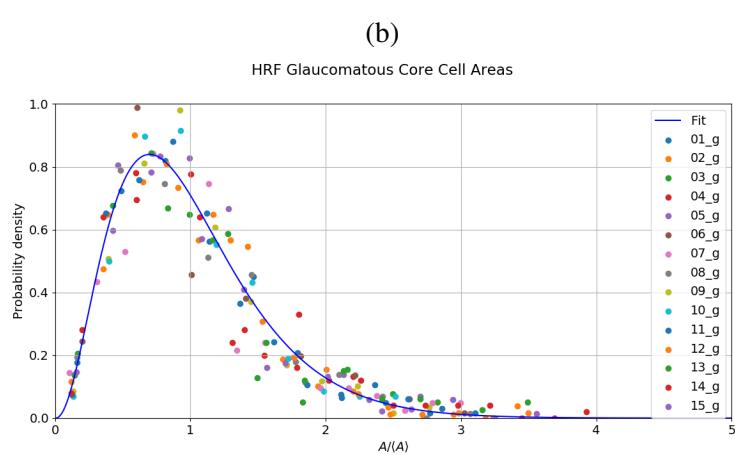
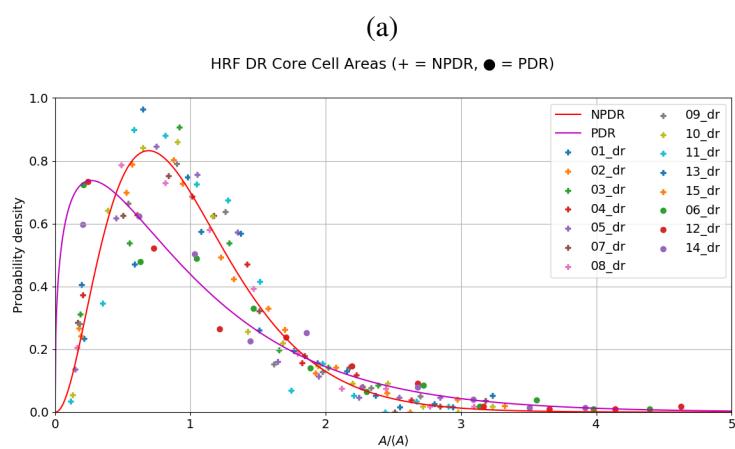
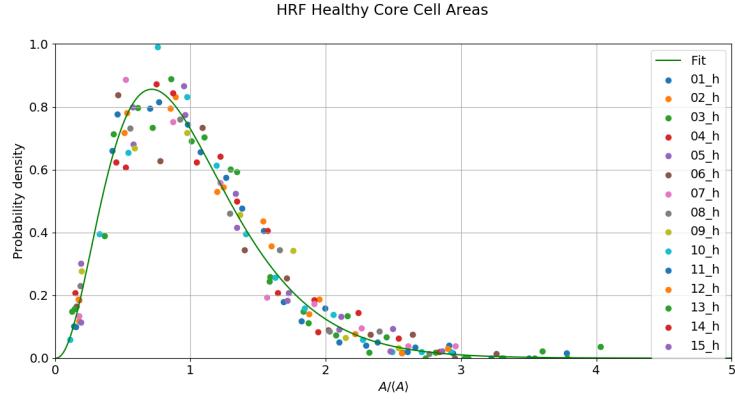


Fig. 1 The distributions of normalized areas showing the spread of data around the best fit gamma for each condition in the HRF dataset (a) healthy, (b) the two types of DR: NPDR and PDR, and (c) glaucomatous. Individual colours indicate different images and are labelled by image name from the HRF dataset.

