FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Estimation of vessel caliber in retinal images

M.Sc. Dissertation

Teresa Manuel Sá Finisterra Araújo

Mestrado Integrado em Bioengenharia

Supervisor: Aurélio Campilho, Full Professor, FEUP

Co-supervisor: Ana Maria Mendonça, Associate Professor, FEUP

July 11, 2016

Estimation of vessel caliber in retinal images

M.Sc. Dissertation

Teresa Manuel Sá Finisterra Araújo

Mestrado Integrado em Bioengenharia

July 11, 2016

Resumo

Os vasos da retina são a única porção diretamente observável da rede vascular humana. Estes vasos constituem uma janela única para a avaliação não invasiva da saúde da vasculatura sistémica, e podem ser avaliados através de imagens coloridas de alta resolução do fundo ocular. Alterações no calibre dos vasos da retina estão associadas com uma variedade de doenças de grande prevalência, nomeadamente diabetes, hipertensão e aterosclerose. O estreitamento arteriolar da retina é um marcador de hipertensão e de retinopatia diabética, sendo que as alterações de diâmetro dos vasos podem ser usadas tanto para prevenção como diagnóstico destas doenças. Contudo, a avaliação clínica dos vasos da retina é frequentemente cansativa e proposensa a erros, particularmente no contexto de rastreio. Assim, há a necessidade de desenvolvimento de métodos automáticos para a medição objetiva e precisa do calibre dos vasos da retina. A variabilidade da aparência dos vasos sanguíneos e da qualidade e resolução das imagens dificulta este processo. Ademais, há falta de dados e critérios padrão, necessários para realizar comparações significativas em grande escala. Apesar de terem sido propostos alguns métodos automáticos de medição do calibre de vasos da retina, a maioria revela limitações, apresentando um desempenho inferior em imagens de baixa resolução, estimando mal os diâmetros em vasos mais finos ou sendo suscetível à presença de artefactos, patologias, etc.

Propõe-se um método de estimação do calibre dos vasos em imagens da retina, baseado no ajuste de modelos aos perfis de intensidade dos vasos. Inicialmente, os vasos da retina são segmentados. De seguida, é detetada a linha central e são extraídos segmentos de vasos individuais através da remoção de junções. Os segmentos são suavizados através da aproximação por *splines* e os perfis de intensidade transversais aos vasos são extraídos nos diferentes pontos do segmento. Após determinação do comprimento dos perfis, estes são cortados e suavizados. Cada perfil é ajustado usando um novo modelo paramétrico com base numa Diferença-de-Gaussianos (DoG) modificada através da multiplicação por uma linha com inclinação variável que é capaz de descrever a assimetria dos perfis. Este modelo permite um bom ajuste aos perfis mais difíceis, nomeadamente os que apresentam reflexo central. Por fim, os parâmetros do modelo que melhor aproxima cada perfil são usados para determinar a largura do vaso nesse ponto com base num método de regressão que usa múltiplas árvores de regressão agregadas por *bootstrap* com seleção aleatória de características.

Os resultados são avaliados numa base de dados pública, REVIEW, que contém anotações de diâmetro dos vasos. Os dois novos modelos DoG com 7 e 8 parâmetros, DoG-L7 e DoG-L8, respetivamente, são comparados com um modelo Hermite de 6 parâmetros. Para avaliar os resultados, são usadas validação cruzada e validação por exclusão de um segmento do conjunto de treino, quer para cada um dos subconjuntos da base de dados quer para o conjunto completo de imagens. Os resultados da validação cruzada apresentam uma precisão próxima dos resultados das anotações manuais. Tal como esperado, os resultados nos subconjuntos mais desafiantes são piores do que os obtidos para os subconjuntos mais simples. Os erros são ligeiramente superiores quando a validação é feita usando a base de dados completa, uma vez que a diversidade dos vasos

é maior. O modelo com melhor desempenho é o DoG-L7, embora o DoG-L8 dê origem a um melhor ajuste aos perfis dos vasos. Para além disso, o desempenho é independente do verdadeiro diâmetro dos vasos. Os resultados obtidos usando validação por exclusão de um segmento são ligeiramente inferiores aos da validação cruzada, devido à remoção de perfis representativos do conjunto de treino.

O método proposto neste trabalho apresenta resultados semelhantes ou mesmo superiores aos métodos do estado da arte. O nosso método apresenta a melhor precisão (menor desvio-padrão dos erros) quando a validação cruzada com 10 conjuntos é usada. Neste caso, reduz-se aproximadamente para metade o desvio-padrão dos erros reportado por outros autores para alguns dos subconjuntos da base de dados. Os novos modelos DoG-L7 e DoG-L8 ajustam melhor os perfis dos vasos do que os modelos conhecidos, especialmente os perfis mais difíceis. A utilização de várias etapas de processamento previamente ao ajuste do modelo também contribui de forma significativa para a melhoria dos resultados. A precisão do nosso método é uma garantia da robustez do algoritmo e valida a sua utilização para a medição do diâmetro dos vasos da retina e a sua possível integração num sistema para avaliação da rede vascular da retina.

Abstract

The retinal vasculature is the only portion of the human vascular network that is directly observable. It is a unique window for non-invasive assessment of the health of the systemic vasculature, and it can be evaluated using high-resolution eye fundus color images. Changes in the retinal vessel caliber are associated with a variety of major diseases, namely diabetes, hypertension and atherosclerosis. Retinal arteriolar narrowing is a marker of hypertension and diabetic retinopathy. Consequently, vessel width alterations can be used both for prevention and diagnosis of these diseases. However, the clinical assessment of the retinal vessels is most of the times tiresome and prone to errors, particularly in the context of large-scale studies. Therefore, there is the need for the development of automatic methods for objective and precise retinal vessel caliber measurement. The variability of blood vessel appearance and of image quality and resolution hinders the process. Further, there is lack of standardized data and criteria, needed for significant comparisons in large scale. Although some automatic methods have been proposed to measure the caliber of retinal vessels, most of them present limitations, either showing poor performance in lower resolution images, returning bad estimates on thinner vessels or being susceptible to the presence of artifacts, pathologies, etc.

A method based on vessel intensity profile model fitting for the estimation of vessel caliber in retinal images is proposed. First, vessels are segmented from the fundus image. Then the centerline is detected and individual vessel segments are extracted through junction removal. Segment smoothing is performed by spline approximation and the cross-sectional intensity profiles at the segment points are extracted. After a profile length determination, profiles are cut and smoothed. Parametric model fitting is then applied to the profiles. The fitting is performed using a novel parametric model based on a Difference-of-Gaussians (DoG) model, modified through a multiplying line with varying inclination which is able to describe profile asymmetry. This model allows a good adjustment to the most difficult profiles, namely those showing a central light reflex. Finally, the parameters of the best-fit model are used to determine the vessel width through regression using ensembles of bagged regression trees with random feature selection.

Results are evaluated on REVIEW, a publicly available dataset with width measurement annotations. Different models are tested: two modified DoG models with 7 and 8 parameters newly introduced in this work , DoG-L7 and DoG-L8, respectively, and an Hermite model with 6 parameters. To evaluate the results, cross-validation and leave-one-segment-out validation are performed, both for each dataset and the whole REVIEW. Cross-validation results show a precision close to the observers' results. As expected, results on more challenging datasets are worse than in easier ones. Errors are slightly higher when validating in the whole REVIEW, since there is a greater variety of vessels. The best performing model is the DoG-L7, although DoG-L8 fits better the vessel profiles. Further, the performance of the algorithm appears not to depend on the true diameters. There is a decrease in performance when leave-one-segment-out validation is performed, due to the removal of representative profiles from the training set.

The method presented in this work is close and even outperforms other state-of-the art methods. Our method retrieves the highest precision (lowest standard deviation of the errors) when 10-fold cross-validation is performed, approximately halving the standard deviation of the errors reported by other, for some datasets. The novel DoG-L7 and DoG-L8 models fit the vessel profiles better than the known models, specially the most difficult profiles. The use of several processing steps prior to model fitting also contributes significantly to the improvement of the results. The precision of our method supports the robustness of the algorithm and validate its use in retinal vessel width measurement and possible integration in a broader system for retinal vasculature assessment.

Acknowledgments

This work is financed by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme, and by National Funds through the FCT – Fundação para a Ciência e a Tecnologia within project CMUP-ERI/TIC/0028/2014

Contents

List of Figures

List of Tables

Abbreviations and Symbols

2D	Two-dimensional
3D	Three-dimensional
AVR	Arteriolar-to-venular ratio
CAD	Computer-aided diagnosis
CLR	Central light reflex
CLRIS	The central light reflex image set
DoG	Difference of Gaussians
DoG-L7	Modified Difference of Gaussian model with 7 parameters
DoG-L8	Modified Difference of Gaussian model with 8 parameters
DoOG	Difference of offset Gaussians
DRIVE	Digital Retinal Images for Vessel Extraction
ESP	Extraction of segment profiles
FOV	Field of view
FP	False positive
HRIS	The high resolution image set
KPIS	The kick point image set
LDA	Linear Discriminant Analysis
LM	Levenberg-Marquardt
MAP	Maximum a posteriori probability
MoM	Measure of Match
REVIEW	Retinal Vessel Image set for Estimation of Widths
RMS	Root-mean-square
ROI	Region of interest
OD	Optic disc
SIRIUS	System for the Integration of Retinal Image Understanding Services
SLRF	Sliding linear regression filter
SP	Seed point
SR	Success rate
TRR	Trust-Region-Reflective
ULDM	Unsupervised Linear Discriminant Analysis
VAMPIRE	Vessel assessment and measurement platform for images of the Retina
VDIS	The vascular disease image set

Chapter 1

Introduction

1.1 Motivation

1.1.1 Retina and retinal blood vessels

The retina is a thin transparent tissue that constitutes the third and most internal layer of the eye. It is placed between the vitreous (interiorly) and the choroid layer (externally) (Fig. 1.1). The retina is a light-sensitive tissue that converts the incoming light into neural signals that can be understood by the brain. With the adequate techniques, the retina is visible from the outside, making it accessible for non-invasive imaging. Retinal vessels are the part of the human vasculature that is most accessible to study, as they are the only portion of the circulation that is directly observable (Abramoff *et al.* (2010)).

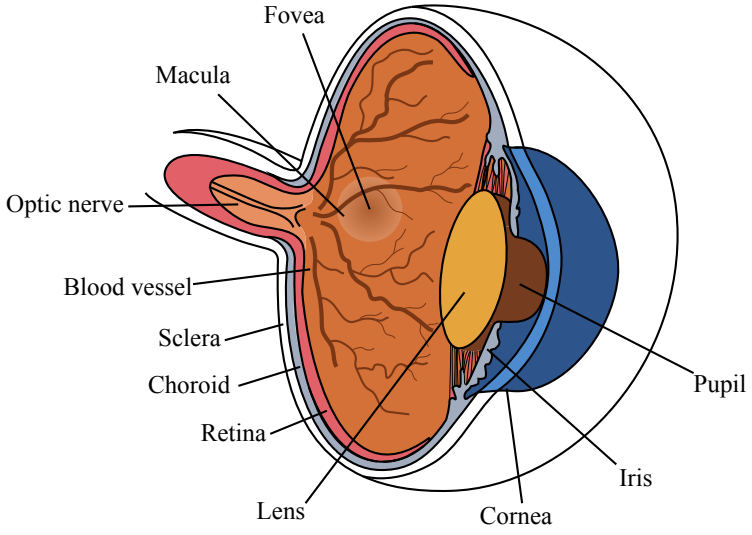


Figure 1.1: Anatomy of the eye. Image adapted from WebMD (2015).

1.1.2 Eye fundus images

The retinal vessels can be evaluated using high-resolution eye fundus color images (Fig. 1.2a). Fundus photography is a well established technique for retinal imaging, widely used for screening and detection of several diseases. This technique consists in capturing a photograph of the back of eye (i.e., fundus). In these images, the intensity represents the amount of reflected light of a specific waveband (Abramoff *et al.* (2010)). Fundus cameras are used for capturing these images. These cameras are described by the angle of view, i..e, field of view (FOV). The normal FOV is 30°. Wide angle fundus cameras capture images between 40° and 140°, providing proportionately less retinal magnification (Saine and Marshall (2001)).

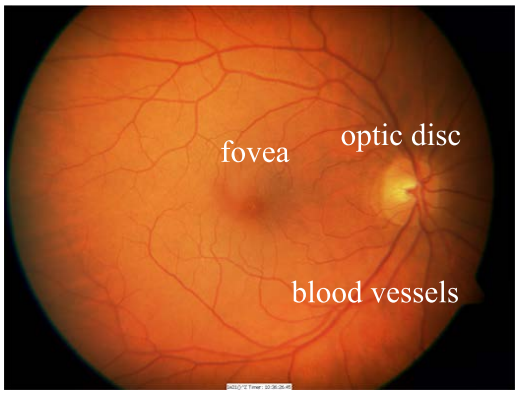
In eye fundus images one can distinguish the following main structures (Fig. 1.2a): 1) *fovea*, that appears as a reddish spot, with no vessels, in the centre of the macula; 2) *optic disc* (OD), located in the center of the retina; it is the beginning of the optic nerve and the entry for vessels that supply the retina; 3) *blood vessels*, that radiate from the OD center; normally, arteries appear lighter and thinner than veins.

1.1.3 Diseases affecting the retina

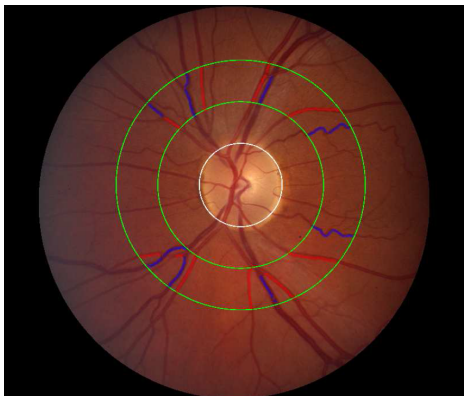
The most prevalent causes of blindness in the world include macular degeneration, diabetic retinopathy and glaucoma (Abramoff *et al.* (2010)). Macular degeneration occurs when the central portion of the retina (i.e., macula) deteriorates, and is highly correlated with aging. It is the most common cause of visual loss in the USA. Glaucoma is also an ocular disease, characterized by damages on the optic nerve, which is often related to elevated intraocular pressure. Diabetic retinopathy appears as a complication of diabetes, and involves changes in the retinal blood vessels. These changes include abnormalities at early stages and vessel diameter modifications. Diabetic retinopathy is the second cause of blindness in the USA and the most relevant cause in the working age population (Abramoff *et al.* (2010)).

Retinal vessels are a window to the systemic vasculature, since they are affected by the factors that influence the body vasculature in general (Nguyen *et al.* (2007)). Many relevant systemic diseases have manifestations in the retinal vasculature. These diseases can originate in the brain, cardiovascular system, among others. Examples are hypertension and atherosclerosis, that can cause changes in the relation between artery and vein diameters.

Retinal vessel caliber changes The changes on the retinal vasculature affect several morphometric indices, such as vessel length, vessel caliber, tortuosity and branching pattern. The quantification of these features can lead to the identification of markers (i.e., indicators) associated to the diseases and, subsequently, be useful for diagnosis. Changes in retinal vessel caliber are an important sign of diabetes mellitus, hypertension, arteriosclerosis and cardiovascular diseases. Further, vessel caliber changes can also be associated to pre-diabetes and pre-hypertension (Nguyen *et al.* (2007), Ikram *et al.* (2013)). Consequently, vessel width alterations can be used for prevention and



(a) Retinal fundus image taken from the VDIS dataset of the REVIEW dataset (image VDIS007).



(b) AVR computation (green: region of measurement; red: arteries; blue: veins). Image from Dashtbozorg (2015).

Figure 1.2: Eye fundus images.

diagnosis. Hypertension has long been associated with generalized or localized arteriolar narrowing. Diabetic retinopathy was related to arteriolar narrowing in some studies, to venular dilation in others, and also to wider arteries (Ikram *et al.* (2013)).

The Arteriolar-to-venular ratio (AVR) is a parameter derived from the vessel caliber measurements (performed at a specific region, normally centered in the OD), that is commonly used as a descriptor of general arteriolar narrowing (Fig. 1.2b). A reduced AVR is an indicator of diabetes and hypertension, as well as a predictor of other cardiovascular diseases in adults.

1.1.4 CAD systems and diameter measurement

The early diagnosis of the above mentioned diseases is crucial to prevent and reduce health damages. The new imaging technologies have allowed more precise detection of subtle changes in the retinal microvasculature. Image analysis can be used as a tool for diagnosis, since it allows the assessment of vascular changes. Automated segmentation and measurement of the vessels is desirable, in order to improve the efficiency, reliability and reproducibility of the results. Besides, automatic methods are useful in wide screening programs for vascular conditions, since the human analysis of a large number of images allied with the complexity of the retinal vascular network can be an extremely heavy task. Consequently, there is the need for developing computer-aided diagnosis (CAD) systems to help in the quantification of retinal structures and biomarkers assessment.

CAD systems should incorporate methods for objective and precise measurement of retinal vessel calibers. Precise caliber measurements are needed if the diameters are to be used directly for diagnosis or for the computation of the AVR. Several authors stress that the vessel width measurement stage is sufficiently critical to be individually and carefully studied (Lowell *et al.* (2004), Kumar *et al.* (2012)). The development of automated methods for width measurement is a demanding process, considering: 1) the variability of the appearance of blood vessels; 2) the

variability of image quality and resolution and 3) the lack of standardized data and criteria for comparing algorithms, preventing significant comparisons in large scale (Lupascu *et al.* (2013)).

Although some automatic methods have been proposed to measure the caliber of retinal vessels, most of them present limitations, either showing poor performance in lower resolution images, returning bad estimates on thinner vessels or being susceptible to the presence of artifacts, pathologies, among others.

1.2 Objectives

The main objective of this M.Sc. dissertation is the development of an automatic method for robust estimation of vessel caliber in fundus images of the retina. This dissertation is thus aimed at:

1. the development of a robust and automatic method for measurement of retinal blood vessel calibers in eye fundus images, in order to allow the improvement of the current state-of-the-art results, particularly in the most difficult cases of small blood vessels and images with pathologies and artifacts;

2. the evaluation of the method on a publicly available dataset designed for vessel width measurement evaluation (REVIEW) and the comparison with the state-of-the-art methods using adequate performance metrics.

1.3 Contributions

An automatic method for vessel caliber measurement is proposed in this M.Sc. dissertation. The main contributions of this work are:

1. a novel parametric model for vessel intensity profile fitting. This model is a modified Difference-of-Gaussians model where a multiplying line with varying inclination modulates the asymmetry of the vessel edges, thus allowing the adjustment of a large variety of vessel profiles. The performance of this model is quantitatively evaluated and proves to outperform other known models for fitting of vessel intensity profiles;

2. a top-level performing model fitting-based approach for retinal vessel width estimation. The method combines model fitting with multiple preprocessing steps, estimating the vessel diameter using ensembles of bagged regression trees with random feature selection. The combination of different approaches makes the algorithm robust and reliable for width estimation in images with pathologies and artifacts, with performance independent of the true vessel widths. The results are close to the gold standard and often outperform the state-of-the-art methods.

1.4 M.Sc. dissertation overview

This document is divided in 5 chapters. The current chapter presents the motivation, objectives and contributions of this work. In chapter 2, the methods for retinal vessel width measurement are described. The state-of-the-art algorithms are divided in two types of approaches and explained in separate sections. A theoretical explanation of the most pertinent concepts is also presented. Finally, the two works considered more relevant to the present study are further detailed. Chapter 3 describes the novel method herein proposed for retinal vessel width estimation. The different steps of the method developed in this dissertation for retinal vessel width estimation are detailed and motivated. Chapter 4 presents and discusses the experimental results. The publicly available dataset of fundus images commonly used to evaluate the methods is described, as well as the evaluation metrics generally used for assessing the algorithms' performance. Further, the results of the evaluation of some of the works from chapter 2 are compared and discussed. Finally, in chapter 5 the main conclusions of this work are presented, as well as suggestions of future work to improve the method.

Chapter 2

State-of-the art methods for retinal vessel width measurement

Several methods have been proposed for retinal vessel analysis. Depending on the study, these methods can target segmentation, landmark detection (such as bifurcation and crossover points) or vessel feature quantification (such as diameter and tortuosity). As the focus of this work is diameter (i.e., width) estimation, this chapter presents the main algorithms described in the literature for retinal blood vessel diameter measurement.

In order to estimate vessel widths in retinal images, a step of vessel detection is needed prior to the measurement. This step allows to determine the vessels' location, leading to the definition of the positions where the width should be measured. For a given detected vessel, each point in the vessel centerline will have a correspondent vessel width. However, if the focus of a work is exclusively the correct diameter estimation, one can use specific vessel locations, known from the ground truth, for example, and measure the width of the vessel solely on those points.

The majority of the methods for vessel diameter estimation fit in two major broad schemes (Lupascu *et al.* (2013)): 1) width measurement from parametric intensity profile model fitting and 2) width measurement from contour detection. In the following sections the two types of diameter estimation methods are discussed and the most relevant algorithms are briefly described. In each section the method considered to be more relevant to the current study is further detailed.

2.1 Width measurement from parametric intensity profile model fitting

These methods fit a parametric model of the intensity profile across the vessel and estimate the width based on the best-fit profile. In Fig. 2.1 the general approach of this type of method is presented. Although the majority of the algorithms follow this broad scheme, the particular methods used in each step and other details differ from study to study. Further, some methods do not perform vessel detection and skip the first steps, starting by the profile determination at a given point of interest. A brief explanation of each of the steps is given in the next subsections.

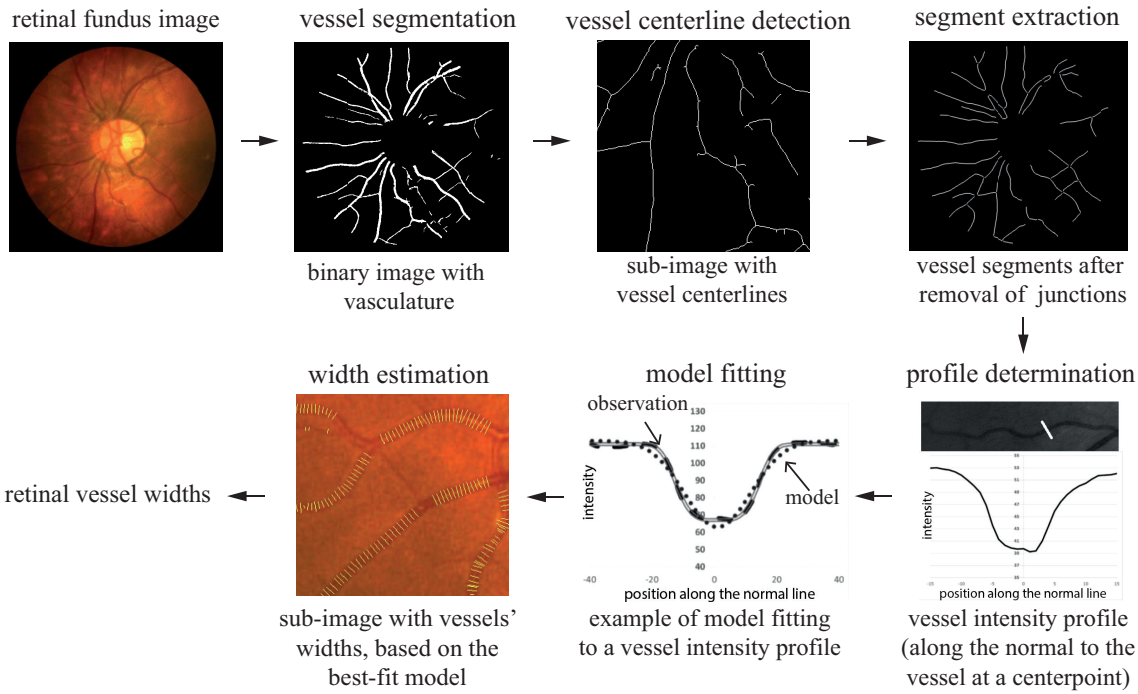


Figure 2.1: General approach for retinal vessel diameter estimation using parametric intensity profile model fitting. First, vessels are segmented from the fundus image, resulting in a binary image with the vasculature. Then, vessel centerlines are detected, followed by the removal of junctions, leading to the extraction of vessel segments. For each point in a vessel segment, the intensity profile normal to the segment at that point is determined. After profile determination, model fitting is performed, being the parametric intensity model adjusted to the observed intensity profile. Once the best-fit model is determined, its parameters are used for determining the vessel width. Images adapted from Fraz *et al.* (2012) and Girard *et al.* (2014).

2.1.1 Vessel segmentation

The first step usually consists in obtaining an initial segmentation of the vasculature. This is needed for posterior centerline detection. For this reason, and because the diameter measurement is not performed directly on the segmented image, the segmentation does not need to provide extremely accurate results, contrarily to what happens in contour-based methods. The main requirement is that the segmentation detects the majority of vessels, in order to allow the measurement in all the existent vessel points.

Several algorithms that perform this step are available in the literature. The method to develop in the context of this work is intended to take advantage of a vessel segmentation algorithm previously developed in the C-BER group, consisting in a version of the method proposed by Mendonça and Campilho (2006), adapted for the segmentation of high resolution images (Mendonça *et al.* (2014)), combining vessel centerline detection and region growing.

2.1.2 Vessel centerline detection

Vessel centerlines can now be detected from the vessel image. The most commonly used approach for this purpose is the skeletonization of the binary image. The skeleton represents the structural connectivity of the objects in the image with one-pixel width.

Thinning The method generally used for producing the skeleton is thinning (Lam *et al.* (1992)). Thinning algorithms can be classified in two broad categories: 1) iterative thinning (i.e., pixel-based) or 2) non-iterative thinning (i.e., non pixel-based). Iterative thinning methods perform pixel-by-pixel operations until a skeleton is obtained. Iterative methods can be divided in sequential and parallel thinning (Kumar and Kaur (2011)). In parallel thinning (Zhang and Suen (1984), Guo and Hall (1989)) the decision for pixel deletion at a given iteration is based on the values of that pixel and of the neighbors at the previous iteration. Hence, all pixels can be processed simultaneously. In sequential thinning, the points are examined for deletion in a predetermined order. The produced skeleton may depend on the visiting order of the pixels. Non-iterative thinning methods are not based on individual pixel examination. Examples of non-iterative thinning include medial axis transforms and distance transforms (Hastings (1992)).

An example of a thinning algorithm commonly used for skeletonization (Gonzalez and Woods (2002)) is illustrated in Fig. 3.2f. It consists in a binary morphological operation that removes pixels on the object's boundaries, while maintaining the connectivity scheme. For this a 3×3 neighborhood is defined around the pixel. When applied to a segmented vessel, the method iteratively removes exterior pixels from the vessel resulting in a image containing the connected centerlines.

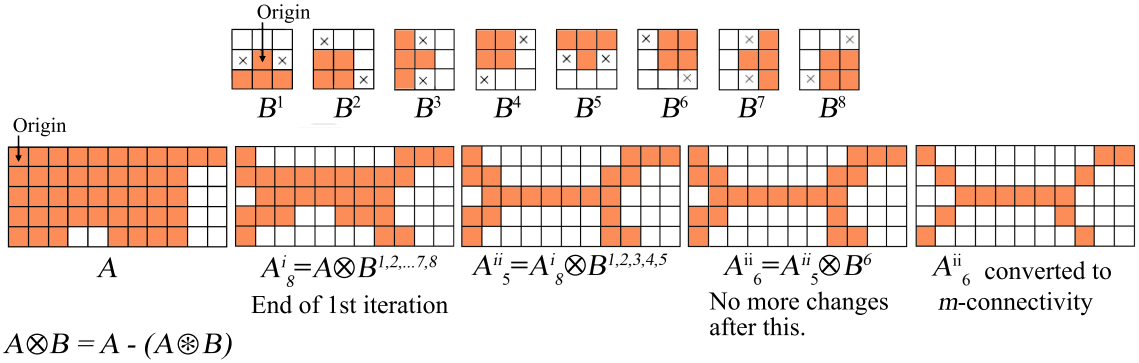


Figure 2.2: Example of thinning operation. The 3×3 structuring elements B^n ($n \in \{1,2,\ldots,8\}$) are iteratively applied to the image A ($\times$ indicate don't care pixels, $\circledast$ represents the hit-and-miss transform and $\otimes$ the thinning operation). A^i_8 corresponds to A after one iteration, i.e., application of all B^n. As no changes occurred from A^{ii}_5 to A^{ii}_6, the process ends with A^{ii}_6, i.e., after the second pass of A with B^6. Extra pixels are removed by converting A^{ii}_6 to m-connectivity, eliminating multiple paths. Note that some of the intermediate steps are not shown. Image adapted from Gonzalez and Woods (2002).

Scale-space skeletonization Other methods can be used in order to obtain the skeleton, such as the scale-space skeletonization (da Fontoura Costa and Estrozi (1999)). It consists in a multi resolution shape representation without border shifting. The methods is based on exact dilations to determine the Euclidean distance transform and propagate labels to the pixels in an object. An advantage of this technique compared to thinning is that it avoids the appearance of small branches along the skeleton of the vessel.

Other approaches In model fitting methods it is also possible to skip the vessel segmentation part and directly find vessel centerlines using methods for this specific purpose. Zhou *et al.* (1994) used a centerline searching strategy that uses tracking based on Gaussian fitting.

Crease detection algorithms can also be applied for finding vessel centerlines. Here, the image is considered a topographic surface, being the elevation given by the pixel intensity values. The crease is a set of continuous points that link the highest or the lowest elevation levels. An example of crease detector is the presented in Lopez *et al.* (2000).

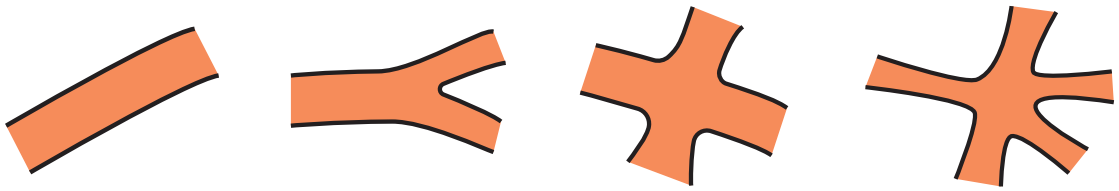
The tramline algorithm, described in Hunter *et al.* (2005), can also be applied to find centerlines. It uses two oriented elements: a line oriented along the direction of the vessel (inner element) and a pair of tramlines lying to either side of the line (outer element). The filter is rotated to various angles and applied to image. If the inner and the outer elements lie inside and outside the vessel, respectively, then the inner element pixels should be darker than the outer element pixels. When an oriented filter response is greater than or equal to a threshold, the pixel belongs to a vessel. The resulting vessel map is morphologically cleaned. This not only detects the centerlines, producing very few false positives (FPs), but also avoids junctions.

2.1.3 Vessel segments extraction (junctions removal)

After vessel centerline location, bifurcation and crossover points are excluded since they would require special treatment. First, diameters are not well-defined at junctions. Further, diameters measured before and after a junction are not directly comparable due to blood flow differences (Bankhead *et al.* (2012)).

As can be seen in Fig. 3.4b, in bifurcation points there are 3 connected neighbors whereas in crossover points there are 4 or more neighbors (8-connection scheme). As so, to remove bifurcations and crossovers: first, the vessel centerline image is scanned and the number of neighbors for each of the centerline pixels is counted; afterwards, centerline points with three or more neighbors are removed. The removal of these points divides the vessel centerlines into individual segments.

It should be noted that the described methods (both for vessel centerline detection and junctions removal) are not exclusively relevant for model fitting-based approaches, but are also commonly used in contour-based approaches, as described in section 2.2.



(a) Single vessel point.

(b) Bifurcation vessel point.

(c) Crossover point (2 vessels).

(d) Crossover point (vessel and bifurcation).

Figure 2.3: Possible vessel point configurations. Bifurcation points occur when one vessel gives origin to two vessels. Crossover points occur when two or more vessels intersect or when vessels meet bifurcations.

2.1.4 Curve fitting

Curve fitting is the process of constructing a curve that best fits a series of data points, possibly under certain constraints (Arlinghaus *et al.* (1994)). Fitting a model to an intensity profile consists in finding the parameters of that model curve that lead to the best adjustment between the model and the observed profile. Depending on the function used for fitting the profile, a different set of parameters are considered. Further, several methods exist for the optimization of these parameters.

2.1.4.1 Models for vessel profile fitting

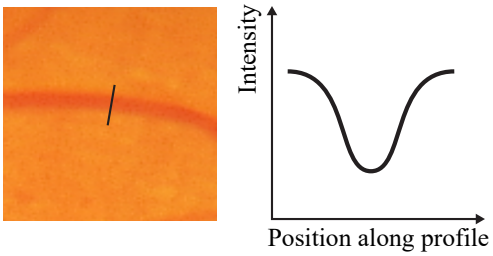
For vessel intensity profile fitting, the fitting function often has a Gaussian-like shape, based on the general appearance of the vessel profile (Fig. 2.4a). As can be observed, vessels are darker than the background and generally have their minimum intensity along the central line.

1D models 1D-Gaussian models are frequently used for fitting the vessel profile (Zhou *et al.* (1994)). A 1D-Gaussian function that can be used for modelling the vessel is defined as:

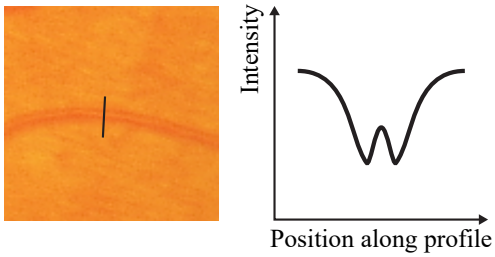
$$g(x) = t - h \times e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.1}$$

where x represents the coordinate along the vessel profile, t is the maximum of the function, h is the height of the Gaussian, μ the location of the center and σ the spread (standard deviation) of the Gaussian. This equation has 4 parameters to be optimized in order to fit the model to a vessel profile. Other Gaussian-based functions can also be used, such as n-order Gaussian derivatives. One example is the second-order differential Gaussian filter used by Gang *et al.* (2002). The combination of multiple Gaussian functions have also been proposed (Aliahmad *et al.* (2012)). Generalized Gaussians, i.e., Gaussian functions of higher order, can also be used (Aliahmad *et al.* (2012), Girard *et al.* (2014)). These functions are defined by:

$$g(x) = t - h \times e^{-\frac{(x-\mu)^n}{2\sigma^n}} \tag{2.2}$$



(a) Vessel without CLR and a profile taken from the vessel.



(b) Vessel width CLR and a profile taken from the vessel.

Figure 2.4: Vessel profiles with and without central light reflex (CLR) in a fundus image. Vessel profiles commonly present a Gaussian-like shape, with the minimum intensity occurring approximately in the vessel center region. Large vessels can present CLR, that appears as an increase of intensity in the center of the vessel, i.e., an elevation in the center of the vessel profile.

where n represents the order of the generalized Gaussian and the other parameters have equivalent meaning as in Eq. 2.1. When using generalized Gaussians models to fit the vessel profile, one extra parameter has to be optimized compared to the Gaussian model.

Vessel profiles do not always present the shape illustrated in Fig. 2.4a. In some cases, the intensity profile of some large vessels is not Gaussian due to the central light reflex (CLR). This phenomena is caused by specular reflection and happens more frequently in digital fundus images than in fluorescein images (Gao *et al.* (2001), Martinez-Perez (2000)). As seen in Fig. 2.4b, this artifact appears as a local intensity elevation approximately in the center of the vessel profile. As the Gaussian profile assumption can not handle the central light reflex artifact, other functions should be used for modelling the vessel intensity profile. Consequently, more general models, considering the CLR, were introduced in the literature such as the twin-Gaussian functions (Gao *et al.* (2001)) and piecewise Gaussian model (Li *et al.* (2003)). The twin-Gaussian function consists in the difference of two different Gaussian functions ($g_1(x) - g_2(x)$), with individual parameters.

Other functions that also consider the CLR have been proposed, as the Hermite polynomial presented in Wang *et al.* (2007). A 1D-Hermite model, depending on 5 parameters, is defined by:

$$h(x) = t - h \times (1 + \beta \times ((x-\mu)^2 - 1)) \times e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.3}$$

where β is an adaptive parameters controlling the depth of the concavity that models the CLR and the other parameters are equivalent to those in Eq. 2.1.

2D-models The use of a 2D model can prevent problems arising from noisy data, such as small branches in the vessels, by introducing smoothing in the process. In this case, instead of fitting a 1D model to a 1D cross-sectional profile of the vessel, a 2D model is fitted to a 2D cross-sectional region of the vessel corresponding to multiple neighboring 1D profiles. Examples of 2D-models used in the literature are the 2D-Gaussian, 2D-Difference-of-Gaussians and 2D-Hermite models.

The vessel intensity can be modelled by a y-independent 2D-Gaussian, resulting from the extrusion of the 1D-model from Eq. 2.1, defined by:

$$g(x,y) = t - h \times e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.4}$$

where x is the coordinate along the vessel cross-section, y is the coordinate along the perpendicular direction, t is the maximum of the function, h is the height of the Gaussian, μ is the center location and σ is the spread of the Gaussian.

A 2D-Difference-of-Gaussians model can also be used, since it takes in account the CLR (Lowell *et al.* (2004)). Considering the Eq. 2.4, this function is given by:

$$g(x,y) = t - (h_1 \times e^{-\frac{(x-\mu)^2}{2\sigma_1^2}} - h_2 \times e^{-\frac{(x-\mu)^2}{2\sigma_2^2}}) \tag{2.5}$$

where the parameters have equivalent meaning as in Eq. 2.4.

A 2D-Hermite polynomial function has also been proposed for vessel fitting (Lupascu *et al.* (2013)). A 2D-Hermite model consisting of an extrusion of the 1D model from Eq. 2.3 to a cylindric surface and allowing asymmetry can be given by:

$$h(x,y) = t - h \times (1 + \beta \times ((x-\mu-\delta)^2 - 1)) \times e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.6}$$

where x is the coordinate along the vessel cross-section, y is the coordinate along the perpendicular direction, δ is a parameter that controls the asymmetry of the model and the other parameters mean the same as in Eq. 2.3. Note that the multi-resolution Hermite model is independent of y. When modelling the vessel profile using this function, 6 parameters need to be optimized.

Therefore, a relevant aspect that should be considered for vessel profiling is the asymmetry of the profile, since vessels can present an asymmetrical shape, corresponding to different intensity profiles on each side of the centerline. Some authors addressed this problem, fitting models that consider non-symmetrical shape of the vessels, either by using a model that includes an asymmetry-related parameter (Lupascu *et al.* (2013), Eq. 2.6) or by fitting two different Gaussian-based models on each of the vessel edges (Aliahmad *et al.* (2012)).

2.1.4.2 Curve fitting methods

Several techniques can be used for finding the parameters of the best-fit model to a given vessel intensity profile. The vessel profile fitting problem can be seen as a least squares problem: given a set of m empirical data pairs of variables (x_i, y_i), optimize the parameters β of the model curve $f(x,\beta)$ so that the sum of the squared differences, $S(\beta)$, given by Eq. 2.7 becomes minimal:

$$S(\beta) = \sum_{i=1}^{m} [y_i - f(x_i,\beta)]^2 \tag{2.7}$$

Given the nature of the models used for vessel fitting, the least-squares problem is non-linear, since the functions are non-linear in the parameters. The methods to solve this problem involve an interactive procedure to improve the parameter values, β, in order to reduce $S(\beta)$.

Levenberg-Marquardt method The non-linear Levenberg-Marquardt (LM) algorithm, also known by damped least-squares method, is commonly used for this purpose (Press *et al.* (2007)). The LM method consists in a combination of two minimization methods: the gradient-descent and the Gauss-Newton methods. The gradient-descent method reduces the sum of the errors by updating the parameters in the direction of the steepest-descent direction. In the Gauss-Newton method, the minimum of the least squares function is found by assuming the function to be locally quadratic (for the parameters near the solution) and finding the minimum of the quadratic (Björck (1996)). Typically, the Gaussian-Newton method converges faster than the gradient-descent method. The LM algorithm adaptively varies the way the parameters are updated between the two methods. In the end, LM behaves closely to the gradient-descent method when the parameters are far from the solution and more similarly to Gauss-Newton method if otherwise. To start the minimization, an initial guess of β is required. In each iteration step, β is replaced by a new estimate, $\beta + \delta$. The LM algorithm (Levenberg (1944), Marquardt (1963)) is given by:

$$(J^T J + \lambda diag(J^T J))\delta = J^T [y - f(\beta)] \tag{2.8}$$

where J is the Jacobian matrix, with element J_i defined by $J_i = \frac{\partial f(x_i,\beta)}{\partial \beta}$, λ is the damping factor (updated in each iteration), $diag(J^T J)$ is the diagonal matrix consisting of the diagonal elements of $J^T J$, δ is the increment to the estimated parameter vector (β) and f and y are vectors with i^{th} component equal to $f(x_i, \beta)$ and y_i, respectively (as defined in Eq. 2.7). This set of equations can be solved for δ. As referred before, if an iteration results in a worse approximation, λ is increased, approximating the LM to the gradient-descent method; otherwise, λ decreases, approximating the LM to the Gauss-Newton method.

Quasi-Newton method Quasi-Newton (variable metric) method can also be used for parameter optimization (Bishop (1995)). These methods are used for finding zeros or local maxima and minima of functions. Quasi-Newton methods rely on Newton' s method to find the point with zero gradient (i.e., $fx' = 0$), using the Hessian matrix of second derivatives of the function to be minimized (f). This function can be equal to the function S from Eq. 2.7 (least squares problem). The Newton method for minimization of a function f with multiple variables is given by:

$$x_{k+1} = x_k - [H(x_k)]^{-1} \nabla f(x_k) \tag{2.9}$$

where k is the step (i.e., iteration) of the Newton method, x is the vector of variables of the f, $\nabla f(x)$ is the gradient of $f(x)$ and $H(x)$ is the Hessian matrix of $f(x)$. Geometrically, in each iteration the function f_x is approximated by a quadratic function around x_k, and a step is taken

towards the minimum of that quadratic function. As usual, an initial guess for x is needed in the beginning of the algorithm.

Differently from the Newton method, in Quasi-Newton methods the Hessian matrix does not need to be computed, being approximated based on the gradient values from previous steps (i.e., Hessian is approximated by finite differences of the gradients). Consequently, Quasi-Newton methods are frequently used when the second derivatives are not available.

Needer-Mead method Other methods have been used in the literature for parameter optimization, such as the Needer-Mead optimizer, also known as downhill simplex method (Nelder and Mead (1965)). It is one of the best known algorithms for multidimensional optimization without derivatives. The algorithm uses only the values of the function (f) at some points in $\mathbb{R}^n$ and does not perform any gradient approximation at these points. A simplex in $\mathbb{R}^n$ is the convex hull of $n+1$ vertices $x_0, \ldots, x_n \in \mathbb{R}^n$ (a simplex in $\mathbb{R}^2$ is a triangle and in $\mathbb{R}^3$ is a tetrahedron). The method starts with the set of $n+1$ vertices (in $\mathbb{R}^n$) of the simplex S and their function values ($f(x_j)$, for vertex j). Then, a series of transformations of S are performed targeting the decrease of the function values at the vertices. The transformations are determined by the comparison of the function values at certain test points with those at the vertices. These transformations can involve reflexion, expansion, contraction or shrinkage (Lagarias *et al.* (1998)). One of the benefits of the Nelder-Mead algorithm is the low number of required function evaluations, which is particularly important when each evaluation is time-consuming (Wright (1995)).

There are aspects to consider when performing model fitting, relevant for all the fitting methods. One of them is the range of the parameters. This restricts the values of the variables to lie within certain intervals. The optimum range of values for each parameter can be defined using, for example, the observation of a large number of vessel intensity profiles, as was performed by Lupascu *et al.* (2013). Besides this, and as mentioned above, one also has to consider the initialization of the parameters. A better initialization can greatly reduce the steps and time needed to find the optimal solution. Finally, termination conditions have to be defined. These conditions determine when to stop the iterations, i.e., when the solution is close enough to the objective.

2.1.5 Width estimation

Once the best-fit model to the vessel profile is found, there is the need to find the relationship between its parameters and the vessel width. Depending on the functions used for modelling the profile, different parameters are used for estimating the vessel width. Even for the same model, several relationships have been proposed in the literature to relate the parameters to the widths.

Single constant scaling factor For the Gaussian model, some authors proposed constant scaling factors to relate the spread (σ) of the best-fit model with the vessel width. Examples presented in the literature estimate the width as $2 \times 1.96 * \sigma$ (Zhou *et al.* (1994)), $2.33\sqrt{2}\sigma$ (Chapman (2001)), or $2\sqrt{2}\sigma$ (Lowell *et al.* (2004), Girard *et al.* (2014)). Others found these factors through

experiments and calibration procedures, determining widths as $2.4\sqrt{2}\sigma$ (Gao *et al.* (2001)) or as $2.03\sigma + 0.99$ (Gang *et al.* (2002)). This procedure uses the measurements of experts to determine the relationship between σ and the width.

Multiple scaling factors It has been proposed in the literature that a scaling factor should be used for each individual edge from a profile, in order to match different blurriness level at the vessel boundaries (Aliahmad *et al.* (2012)). The factor for a given edge depends on the spread (σ) of the Gaussian that was fitted to each edge. Also, a process of learning the best scaling factors for each profile can be applied. Supervised learning has already been used in the literature to learn the mapping between the 6 parameters of the best-fit 2D-Hermite model to the vessel width (Lupascu *et al.* (2013)). The training set is composed by the values of the parameters associated to each intensity surface fitting and the corresponding widths manually measured by an observer.

2.1.6 Most relevant state-of-the-art methods

Several models have been proposed in the literature to determine the vessel width by fitting retinal vessel intensity profiles. In this section the most relevant methods for this work are described, either due to the method itself or by the availability of results on known databases.

Simple rules applied to profiles Brinchmann-Hansen and Heier (2009) proposed the half-height-full-width (HHFW) algorithm for vessel width estimation. The vessel width is estimated as the distance between the points on the cross-section intensity profile at which the function reaches 50% of its maximum value to each side of the estimated centerline. Gregson *et al.* (1995) fitted a rectangle to the vessel cross-section intensity profile by adjusting its width so that the area under the profile is equal to the area of the rectangle. The height of the rectangle is fixed to the difference between the minimum and maximum profile intensity values. Vessel diameter is considered equal to the width of the fitted rectangular profile.

1D approaches A matched filtering approach coupled with *a priori* knowledge about retinal blood vessel properties was proposed by Zhou *et al.* (1994) for vessel boundary detection, centerline tracking and parameter extraction, including vessel width. The vessel profile is modelled using 1D-Gaussians. Vessel edges are located by cutting the best-fit Gaussian model on both sides at the 95% point, meaning that the vessel walls are at 1.96 standard deviations from the mean.

1D approaches considering the CLR and profile asymmetry Aliahmad *et al.* (2012) presented a multi-step regression method for diameter measurement. The technique is based on Gaussian modelling, adaptively combining second and higher order Gaussians to model the vessel profile. The method comprises the following steps: 1) combinations of multiple second order Gaussians are used for modelling the central light reflex of the profile; higher order Gaussians are used for reproducing the edges, being one model fitted to each edge; 2) the orders of the models and other parameters are estimated through regression analysis: the non-linear Levenberg-Marquardt

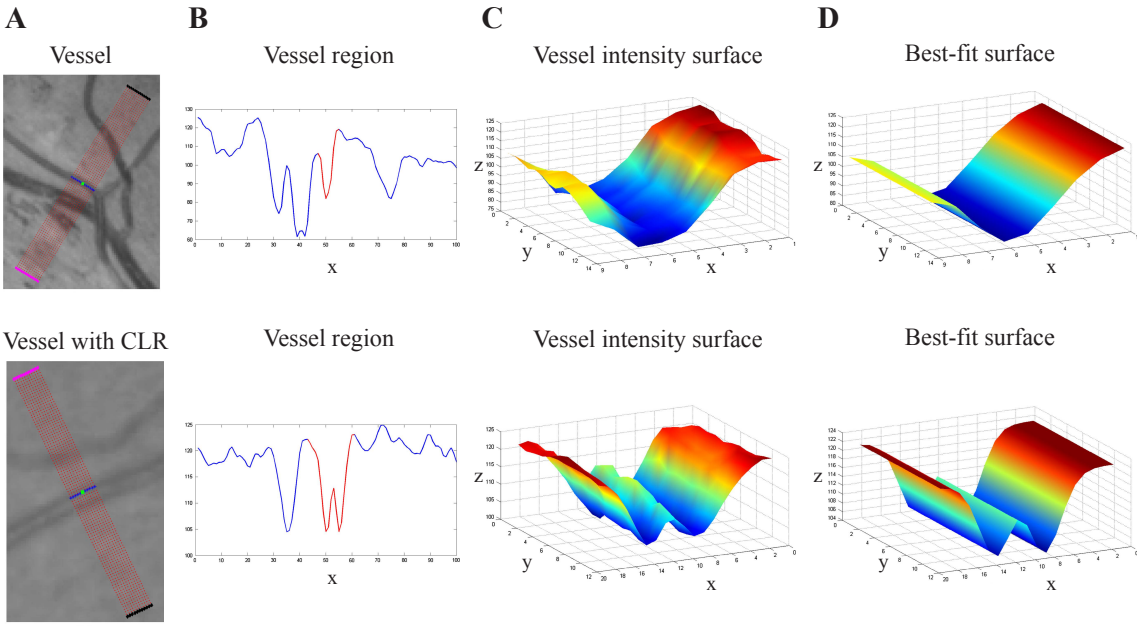


Figure 2.5: Method of Lupascu *et al.* (2013) for retinal vessel width determination. **A** - Initial images of the vessels along with the cross-sectional profiles (red dots) selected around the targeting locations (green squares). Top: vessel without central light reflex; bottom: vessel width central light reflex. **B** - Mean intensity profiles around the target locations, with the vessel regions highlighted (red). **C** - Intensity surfaces, taken from the 2D regions (profiles) represented in A by the red dots. **D** - Best-fit surfaces obtained through parameter optimization of the Hermite model using the Nelder-Mead method. Images adapted from Lupascu *et al.* (2013).

method is used for fitting the blurriness level at the vessel contours; 3) the sigma parameters and the distance between the minima of the best fitted Generalized Gaussians to the boundaries are used for estimating the diameter. The method, contrarily to some of the previous methods, doesn't assume the vessel profile as being symmetrical (with similar blurriness level at both edges).

2D approaches considering CLR In the work of Lowell *et al.* (2004), diameter is measured based on 2D-Gaussian or 2D-Difference-of-Gaussians (DoG) models, optimized to fit a 2D intensity vessel segment. The method has special attention to the possible light reflex. In the proposed algorithm, the following steps are applied: 1) a tramline algorithm is used in order to find vessel centerlines and remove junctions (see subsection 2.1.2); 2) the model is fitted on the green channel of the original image; 3) variable-metric optimization is used for finding a good set of parameters for the models; 4) both models (2D-G and 2D-DoG) are fitted and the best-fit is selected: 2D-Gaussian profile is used taking into account vessels without reflex, whereas 2D-DoG profile is used for vessels with reflex; 5) parameters of the best-fit model are used for estimating the width.

2D approaches considering CLR and profile asymmetry Lupascu *et al.* (2013) used bagged decision trees and an extended multiresolution Hermite model to estimate the width of retinal vessels. The model is a surface obtained by extruding a 6-parameter Hermite cross-sectional

model, which is a generalization of Wang *et al.* (2007). This work is particularly interesting, as it is robust and generic, allowing a great variety of vessel intensity profiles. Also, the method shows one of the best performances, as shown in Chapter 4.

The following steps are performed to assess the vessel width:

1. *Extraction of vessel local intensity surfaces*: Initially, a point for width estimation is selected inside a vessel. Then, 5 parallel cross-sections are chosen on each side of the target location (for a total of 11 cross-sections) and their intensity profiles are extracted (see Fig 2.5-A). This is performed using bilinear interpolation of the intensities at non-integer locations. The orientation of vessels is given by the eigenvectors of the Hessian matrix. The length of the profiles is fixed to 100 pixels, in order to be large enough to cover the largest vessel.

   As there is the possibility that the profile includes more than one vessel, there is the need to identify the region containing only the vessel of interest (Fig 2.5-B). For that, the mean profile of the vessel along the 11 sections is computed. Then, the maximum with lowest intensity (expected to be the central reflex) is found, along with the two adjacent minimum and maximum in each side. If this region does not contain the target point, it is removed from the profile and the process is iterated. If no CLR vessel is found, the process restarts, this time seeking for the absolute minimum and the adjacent maximum on each side, i..e, normal vessel configuration. This process is done assuming that vessel intensity maxima are always lower than background maxima. Finally, a 3D surface is formed by the support regions of the vessel cross-sections, constituting the data for the surface fit (Fig 2.5-C);

2. *Surface fitting using an Hermite model*: A parametric surface model of the cross-sectional intensities of the vessels is then fitted to the surface. The model used is the Hermite model with 6 parameters, as defined in Eq. 2.6. The 6 parameters allow to control 1) the depth of the CLR of the profile; 2) the asymmetry of the model; 3) the mean of the Gaussian (shifting the function along the x-axis); 4) the standard deviation of the Gaussian; 5) the height of the model and 6) the maximum of the model (shifting the function along the z-axis).

   For surface fitting, the sum of the squared differences between the model and the vessel data is minimized using a Needer-Mead optimizer, described in section 2.1.4. The best-fit model to the vessel intensity surface is found (Fig 2.5-D). The initialization surface is determined by finding the peaks of mean profile of the 11-cross-sections, determining the parameters of a 1D-model based on the peaks and extruding the 1D-model to a surface. The optimum range for each parameter was estimated by observation of a large number of profiles.

3. *Width estimation*: A supervised learning method consisting of an ensemble of bagged decision trees is used in order to estimate the width from the parameters of the best-fit surface (Breiman (1996)). This classifier learns the mapping from a point in the 6D parameter space to the vessel width. The training set is composed by the association of the values of the 6 parameters of the fitted-model with the manual width estimates of an observer.

Alternative approach Delibasis *et al.* (2010) proposed a method for the segmentation of the whole vessel network and for diameter estimation consisting in a model-based vessel tracing algorithm. The algorithm uses a generic parametric vessel model (that can assume arbitrarily complex shape), consisting in a "stripe" that exploits geometric properties to define the model parameters. A Measure of Match (MoM) between the model, i.e., the stripe, and the image is used, that quantifies the similarity between the vessel model and the given image. SPs that lie sufficiently close to vessels centerlines are automatically determined. The vessel tracking proceeds by identifying the stripe that better matches the vessel, using the SP, stripe orientation, stripe width and MoM information. After the current vessel cannot be traced any longer, the algorithm actively seeks bifurcations in order to avoid user intervention. To determine the vessel diameter, the geometric model is used, which parameters were already determined during the vessel tracking. A set of train images are selected to discover the relation between the fitted model "diameter", i.e., stripe's width, and the vessel diameter, assuming a linear dependency between them. The derived diameters are subsequently used for segmenting the vessels.

2.2 Width measurement from contour detection

Width estimation algorithms that fit in this category use contour detection methods to directly find vessels boundaries and determine the vessels' width based on these. The majority of the vessel segmentation methods uses the contrast between the blood vessel and background. However, structures in the image (OD, pathological signs, etc.) can hamper the segmentation. Additionally, narrows vessels show low contrast, hindering the segmentation. Further, the CLR in larger vessels can lead to the identification of two vessels instead of one.

The methods available in the literature use a wide variety of contour delineation methods, such as active contours, graph-based approaches, wavelets, classification or tracking-based methods. The approaches that fit in the category of width measurement from contour detection can be further divided in methods that: 1) perform vessel segmentation and measure the width based on the segmentation or 2) perform vessel segmentation but use it only as a starting point for the width measurement. The most relevant approaches in each of these sub-categories are described in the current section.

2.2.1 Width measurement directly from segmentation

Active contour-based approach Al-Diri *et al.* (2009) presented a method for segmentation and width measurement of retinal vessels, the "extraction of segment profiles" (ESP) algorithm, based on the "Ribbon of Twins" (ROT) active contour model. The ESP uses two pairs of contours (i.e., a twin) to capture each vessel edge (one inside and one outside the vessel). In total, each ROT contains four linked active contours: the two internal contours move outwards toward the vessel edge and are linked together to maintain width consistency; the two outside contours move in the opposite direction, towards the internal contour. The major steps of the ESP algorithm are: 1) *tramline algorithm:* an initial set of potential vessel segment centerline pixels is located

(see subsection 2.1.2); 2) *segment growing algorithm*: segments (consisting of series of profiles) are produced, while discarding false positive (FP) pixels; the ROT active contour model is used for exploring the vessel segment and add profiles; 3) *junction resolution algorithm*: the discrete segments are extended, and the crossings, junctions and joinings resolved by using an implicit neural cost function to determine the network topology.

Graph-based approach A graph-based approach for vessel boundary delineation and diameter estimation was proposed by Xu *et al.* (2011a) and Xu *et al.* (2011b). In this algorithm, the two boundary segmentation is converted into an optimal surface problem on a two slice 3D graph. Consequently, the two boundaries of the vessel are simultaneously segmented. The main steps of the algorithm are the following: 1) images are preprocessed: an initial segmentation is generated from a vessel probability map (Niemeijer *et al.* (2004)), sequential thinning is applied to find vessels centerlines and bifurcations and crossing points are excluded, cutting the vessels in segments; 2) a 3D graph is constructed based on each vessel segment; the weights of the profile nodes in the graph are obtained using steerable first order Gaussian derivatives; 3) vessel widths are computed as the distance between the two corresponding nodes on the optimal surface from the two slices.

Classification-based approach A diameter estimation algorithm that relies on Unsupervised Linear Discriminant Analysis (ULDM) was proposed by Kumar *et al.* (2012). The algorithm can be divided in the following parts: 1) the grader identifies the region of interest (ROI); 2) *initialization:* the vessel boundaries are estimated, as described in Aliahmad *et al.* (2011), and the intensity cross-section profiles are obtained; 3) *Linear Discriminant Analysis (LDA)* (Wijaya *et al.* (2011)): the training data is automatically generated using the LDA classifier; the classifier is trained in order to separate the profile in three sections, one corresponding to the vessel surface and the other two to the background at either side of the vessel; 4) *diameter measurement:* the trained LDA identifies the three sections; the width of the first is the vessel diameter.

Tracking-based approach An algorithm for vessel segmentation and diameter estimation was proposed by Yin *et al.* (2013), consisting of an automatic tracking-based method using probabilistic formulation. After automatic selection of seed points (SPs), the tracking, for one SP, proceeds as follows: 1) *initialization:* initial parameters (vessel center point and direction) are retrieved from the SP; 2) *iteration:* vessel edge points are detected iteratively; in each iteration, a dynamic search window is defined (based on the local vessel parameters); then, statistical sampling is used for selecting new candidate edge points; after, the criterion of Bayesian method with maximum *a posteriori* probability (MAP) is used in order to select edge points among the candidates; finally, vessel parameters are updated for the next iteration (further, in each iteration, a branch prediction scheme is applied to detect vessel branches in the area); 3) *end:* the process stops when the end conditions are satisfied. Once all the SPs are processed, the whole vascular tree is obtained. Based on the detected vessel edges it is possible to deternine the vessel diameter at a given point.

Alternative approach In the work of Yin *et al.* (2014) a method for retinal vasculature segmentation is presented, being the vessel diameter estimated on the segmented images. The segmentation starts by graph partitioning, using spectral clustering based on morphological features to pre-group the vessels. The eigen values of Hessian matrix are used for estimating the local curvature, in order to enhance the vessels. This image is combined with a binarized one, obtained using a threshold that maximizes the entropy. The binarization is repeated on the original image using a different threshold that takes into account vessels with CLR. The final result is the combination of the two segmentations (with and without CLR). For diameter estimation, 1) morphological thinning followed by skeletonization are applied for location of the one-pixel thin centerline; 2) edge points are detected with 3x3 windows on the vessel centerline: each window contains three connected and non-colinear pixels, aligned along one of the 14 different possible configurations/orientations; the selection of such aligned pixels as candidates allows to reject branch pixels for edge detection; 3) by rotating 90° the set of three centerline pixels around the central point (in order to obtain an approximation of its perpendicular) and extending the segments until they intersect the closest vessel boundary (in the segmented image), the vessel diameter corresponding to the central point in study is computed as the Euclidean distance between the intersection points.

2.2.1.1 Methods integrated in frameworks for retinal image analysis

Enhancement of binary vessel masks for width measurement A constrained spline fitting algorithm has been proposed, that does not perform vessel contour detection from the original image, but refines binary vessel masks. The algorithm of Trucco *et al.* (2013) is integrated in **VAMPIRE** ("Vessel assessment and measurement platform for images of the Retina"), which is a software application for quantification of retinal vessel properties, including detection of landmarks, determination of vessel width, tortuosity and others (Perez-Rovira *et al.* (2011)). The method, which is applied to binary images with the vasculature, can be divided in the following parts: 1) the skeleton is obtained through morphological thinning; then, branching points are removed; finally, a natural cubic spline is fitted to the centerline; 2) two coupled cubic splines are fitted to the vessel contours by solving a linear system, overconstrained by a parallel-tangent constraint coupling the two splines and penalizing locally non-parallel contours; 3) the vessel width at each point lying on the spline-smoothed centerline is found by determining the Euclidean distance between the two refined contours, lying on the segment orthogonal to the centerline at the point in study.

Width measurement for AVR determination In the work of Vazquez *et al.* (2013), a deformable model-based algorithm was presented for retinal vessel width measurement. The proposed methodology integrates the **SIRIUS** ("System for the Integration of Retinal Image Understanding Services") framework, a web-application for retinal image analysis arteriolar-to-venular diameter ratio (AVR). In this framework, vessel widths are measured at equidistant points from the optic disk (OD) (4-6 radii). The major steps of the algorithm are the following: 1) vessel location is performed using vessel centerline extraction and tracking: vessel centerlines are detected (for

each analysis radius) using a crease detector algorithm, the Multilocal Level Set Extrinsic Curvature enhanced by the Structure Tensor (Lopez *et al.* (2000)); centerline tracking is performed, in order to avoid FPs and obtain a segment for the measurement; 2) a specific snake is used for measuring each vessel (the initial snake has two chains of nodes initialized in the previously located centerline); after energy minimization, the snake should be a parallelogram and the nodes should be in the vessel edges; 3) the vessel width is determined having in account a certain number of profiles in the snake parallelogram, equally spaced and perpendicular to the centerline.

2.2.2 Width measurement partially dependent on segmentation

These methods usually perform a segmentation of the vasculature that allows to locate the vessels obtain and/or to obtain a coarse first estimation of the vessel width. Then, the vessel width is measured in a non-binary image, considering the information obtained from the initial segmentation.

The most relevant work where this strategy is applied is that of Bankhead *et al.* (2012), who proposed a method for vessel detection and diameter measurement relying on wavelets and edge refinement. This method is considered of particular interest for this study, since the segmentation is used only as an initial step for width determination, being followed by edge refinement, i.e., the method takes advantage of a prior segmentation but does not heavily depends on it for width measurement. Further, results presented in Chapter 4 show the robustness of the method.

In the method of Bankhead *et al.* (2012), segmentation is performed by thresholding wavelet coefficients, obtained through Isotropic Undecimated Wavelet Transform (IUWT). The main steps of the width measurement algorithm, following the vessel segmentation, are:

1. *Centerline computation*: morphological thinning is applied to the segmented image, resulting in vessel centerlines. Then, branch pixels are removed (>2 neighbors), dividing the vascular tree into individual segments. Next, the segments are processed, being the short segments (<10 pixels) removed. A coarse estimation of the width is calculated as follows: 1) the distance transform of the inverted segmented image is computed; 2) the diameter at each center point is computed by doubling the distance transform results; 3) the maximum of the diameters from a given segment is taken as the width estimation for that segment. Centerline segments that have fewer pixels than its estimated diameter are not considered.

2. *Centerline refinement*: a least-squares cubic spline is fitted to each centerline segment in order to smooth it and allow accurate derivative computation, leading to accurate vessel orientation determination. A parametric spline curve is used, with the centripetal parameterization scheme described in Lee (1989).

3. *Image profile generation*: Intensity profiles are determined from the original grayscale image, perpendicularly to the spline, at every point, using linear interpolation. The length of the profiles is computed by doubling the diameters calculated from the distance transform, to ensure that large vessels are included. The image profiles are aligned side by side in straightened vessel images, in which each row corresponds to a profile (Fig. 2.6-A).

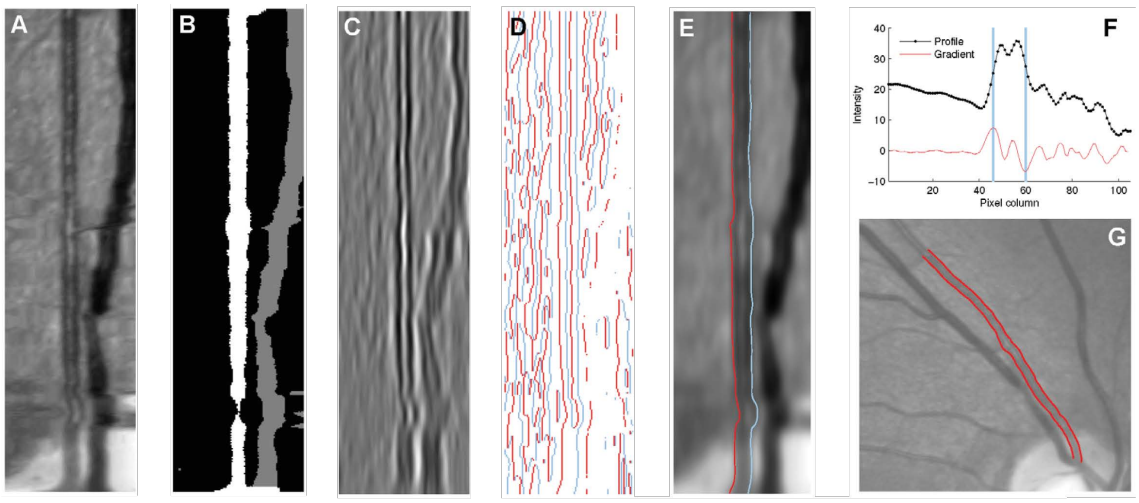


Figure 2.6: Determination of vessel edges in the method of Bankhead *et al.* (2012). **A** - "Straightened" vessel image (stacked vessel profiles). **B** - Correspondent profiles obtained from the segmented image. White: possible vessel pixels; gray: pixels from other vessels or outside the FOV. A provisional width is estimated as the median of the sums of the white pixels in each profile, and is then refined (see **F**): 1) the average of all vessel profiles is computed (black curve), 2) the maximum and minimum gradient to the left and right of the centerline, respectively, are located (blue lines) 3) the distance between the two positions is computed. **C** - Smoothing of the profiles in A by an anisotropic Gaussian filter, followed by a filter to determine the second derivative perpendicularly to the vessel, i.e., horizontally. The size of the Gaussian filters is dependent on the initial with estimation. Most of the vessel has negative values, except in the CLR. **D** - Pixels from C where sign changes occur. Red: positive to negative transitions; blue: negative to positive transitions. (Transitions belonging to the other vessel (gray in B) were removed). **E** - Located edges, shown in the straightened vessel. These edges are determined as follows: the length of each connected component in D is computed and only the longest lines that fall close to the locations determined in F (blue lines) are kept, one in each side of the vessel center. The diameter can be computed as the Euclidean distance between the edges. **G** - Located edges, shown in the original image (from REVIEW). Images from Bankhead *et al.* (2012).

Two corresponding binary images are also generated by applying this procedure to the segmented image (and using nearest neighbor interpolation): 1) one image only with "vessel" pixels (connected region that overlaps the centerline); this constitutes an initial estimation of the vessel location (Fig. 2.6-B); 2) another image containing pixels outside the FOV and other vessel pixels not overlapping the centerline (i.e., belonging to neighbor vessels); this will be used later to define regions in the profile images where vessel edges cannot be found (Fig. 2.6-C);.

4. *Vessel edge identification*: The edge is defined, in this work, as occurring at a local gradient maximum or minimum, that can be identified by performing a search for zero-crossings of the second derivative on the direction perpendicular to the vessel. The description of the vessel edge identification process is available in Fig. 2.6. In the end, the diameter is extracted from the detected edges.

2.3 Concluding remarks

In this chapter, the most relevant methods presented in the literature for retinal vessel caliber estimation are described. These algorithms fit in one of two major schemes: those that measure the vessel widths from parametric intensity model fitting to the vessel intensity profiles, and those that measure the widths from vessel contour detection.

Methods that fit parametric models to the vessel intensity profiles use the parameters of the best-fit model to determine the vessel width. The general approach involves vessel segmentation, centerline detection, removal of junctions, generation of the intensity profiles, fitting of a parametric intensity model to the vessel intensity profile and determination of the vessel widths based on the parameters of the best-fit profile. Some of the works do not perform vessel detection and measure the width at specific points. Model fitting-based methods rely on the general behavior of the vessel profile, which is similar to a Gaussian shape. However, some vessels present CLR, having a lighter intensity along their central region. Further, vessels can present an asymmetrical shape. Both 1D and 2D models have been used for modelling vessel profiles, being that 2D models can reduce problems arising from noisy data. Some of the most recent works address the CLR and profile asymmetry, but only a few methods consider both aspects. Regarding the estimation of the vessel width from the model parameters, the majority of the works consider a fixed relation between the parameters and the width, independently of the vessel and image in study. However, some of the most recent methods consider multiple scaling factors.

Methods that fit in the category of width measurement from contour detection perform vessel segmentation and measure the diameter from the vessel contours. These methods rely on the contrast between vessel and the background. However, other anatomical structures that are present in the image, low contrast and the CLR in the vessels can hinder the segmentation. A wide variety of techniques has been proposed for this purpose, such as active contours, graph-based approaches, classification and tracking-based methods. Although the majority of these methods measures the width directly from the segmentation, others use the segmentation as a starting point for width measurement, using it only for vessel detection and/or to obtain an initial estimation of the widths. The latter methods are of greater interest to this work, since the method to develop in this M.Sc. dissertation can take advantage of a segmentation algorithm previously developed in the C-BER research group.

Chapter 3

Method for retinal vessel width estimation

3.1 Method overview

The proposed method for retinal vessel width estimation follows a model fitting-based approach. It includes several profile preprocessing steps before model fitting, and uses ensembles of bagged regression trees to estimate the vessel diameters from the best-fit model parameters. As referred in chapter 2, the method presented in this work includes some steps adapted from the algorithms of Bankhead *et al.* (2012) and Lupascu *et al.* (2013). An overview of the different phases involved in the algorithm is shown in Fig. 3.1. First, vessels are segmented from the eye fundus image, resulting in a binary vessel image. Then, vessel centerlines are obtained through thinning. Afterwards, junctions are removed, leading to the extraction of vessel segments. These vessel segments are then smoothed through spline approximation and, for each segment point, the intensity profile normal to the segment at that point is extracted. After this, the approximate length of the profiles is determined and the profiles are smoothed. Afterwards, model fitting is performed, being the parametric model fitted to the observed intensity profiles. The parameters of the best-fit model are determined and used for the estimation of the vessel width through regression.

3.2 Vessel segmentation

The first step of the algorithm consists in obtaining an initial segmentation of the vasculature. The main purpose of this step is the detection of the retinal vessels for posterior centerline detection. In this work we apply a vessel segmentation algorithm previously developed in the C-BER group: the method proposed by Mendonça and Campilho (2006), adapted for the segmentation of high resolution images (Mendonça *et al.* (2014)). This method combines centerline detection and region growing for the segmentation of retinal blood vessels. Roughly, the method has three major phases: 1) *preprocessing:* a) the image intensity is normalized by subtraction of the estimated background; b) thin vessels are enhanced by applying four directional line filters; the highest filter

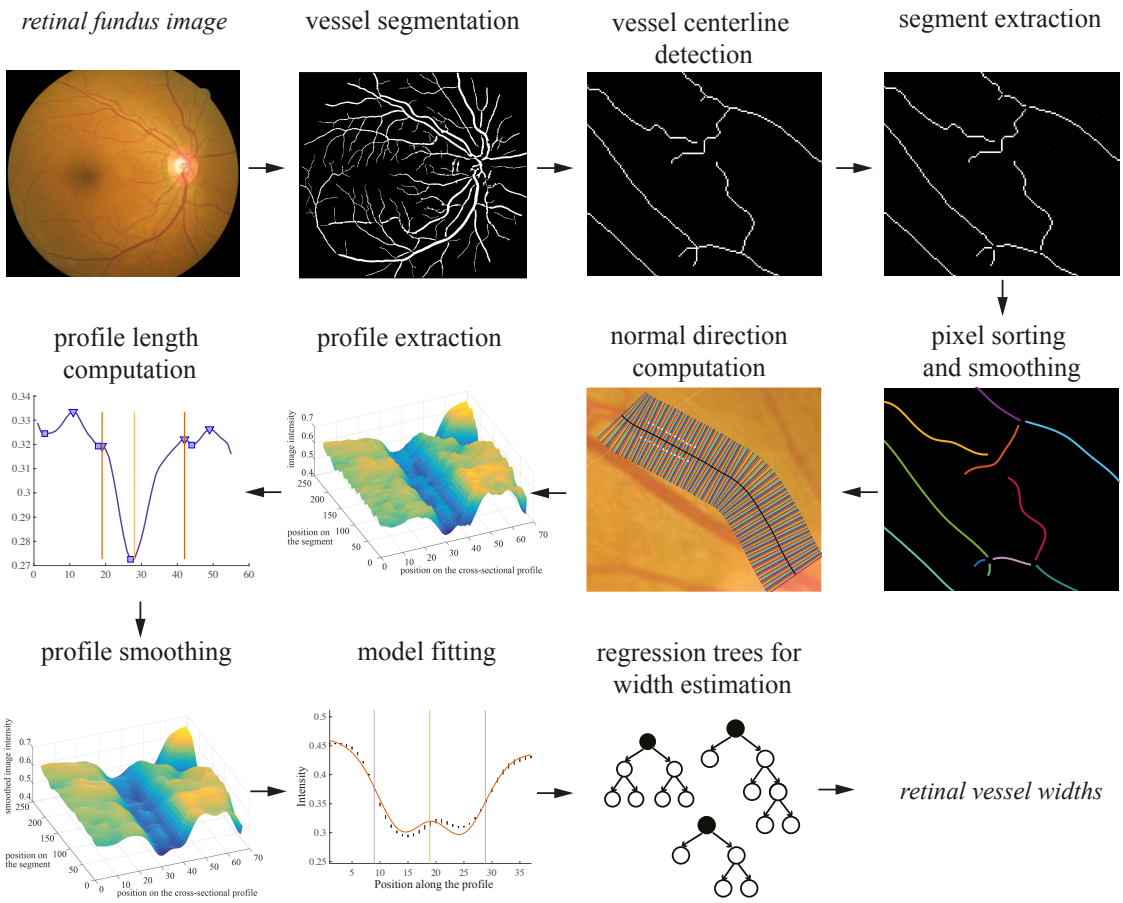
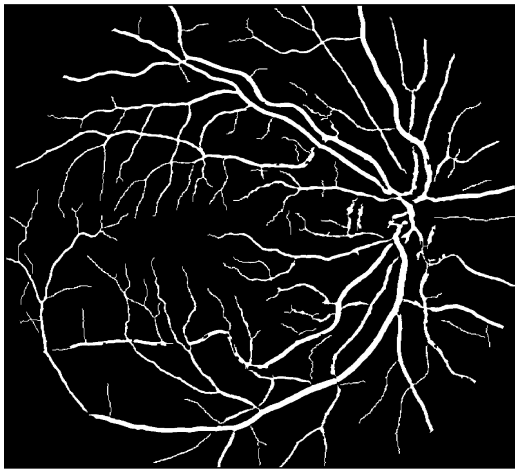


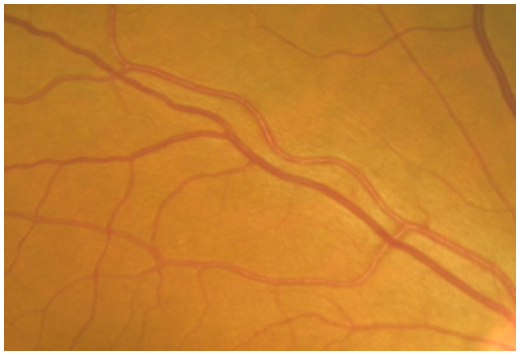
Figure 3.1: Method overview. First, vessels are segmented from the retinal image, followed by vessel centerline detection. Then, junctions are removed, resulting in vessel segments that are afterwards smoothed. Intensity profiles are extracted perpendicularly to the vessel, the length of the profiles is determined and the profiles are cut. Then, model fitting is performed on smoothed profiles. Based on the best-fit model parameters, vessel width is determined through regression.

response is selected and added to the normalized image; 2) *vessel centerline location:* a) the outputs of a set of four directional difference of offset Gaussian (DoOG) filters are processed to detect centerline candidates; segments are validated based on their intensity and length; b) region growing is applied to the candidate points, connecting them into segments; c) centerline segments are validated based on their intensity and length; 3) *vessel segmentation:* a) multi-scale morphological vessel enhancement is performed by using a modified top-hat transform with variable-sized structuring elements; b) multi-scale morphological vessel reconstruction is used for generating binary vessel maps at four scales; c) region growing is applied for vessel filling (SPs are the centerline pixels); in each step, the output image of the previous iteration is used as seed; the process is successively applied to the four scales. Fig. 3.2 shows an example of the application of the segmentation algorithm to an eye fundus image from the REVIEW dataset.

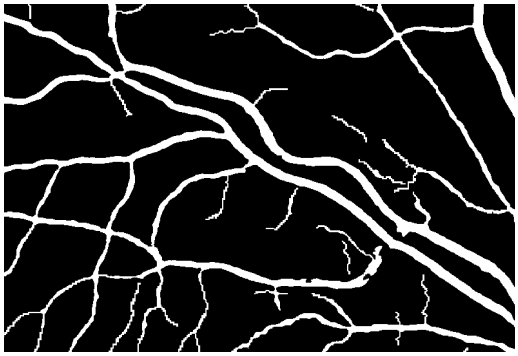
(a) Initial image from REVIEW dataset (CLRIS001) (cropped in width).



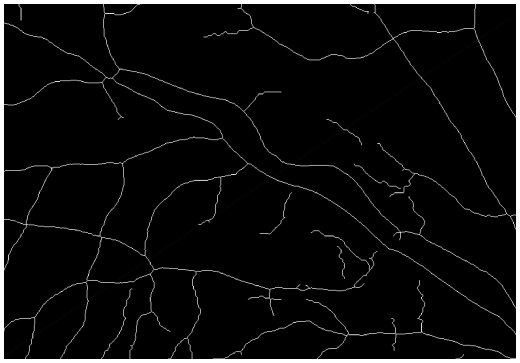
(b) Segmented image, by application of the algorithm described in Mendonça *et al.* (2014).



(c) Region selected from 3.2a.



(d) Region selected from 3.2b.



(e) Vessel centerlines obtained through thinning of 3.2d.

(f) Thinned vessels on the original vessel image. Black: vessel contours; white: centerlines.

Figure 3.2: Example of the application of the vessel segmentation and thinning algorithms.

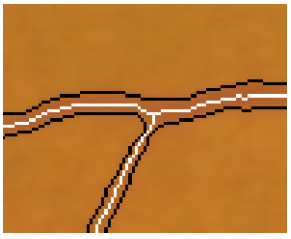
3.3 Vessel centerline detection

Vessel centerlines are detected from the binary vessel image. For this purpose, the skeletonization of the binary image is performed. The skeleton represents the structural connectivity of the objects in the image with one-pixel width. In this work we use a thinning technique to obtain the skeleton of the vessels (Lam *et al.* (1992)). In Fig. 3.2 an example of application of the thinning algorithm to a binary image with the vasculature is shown.

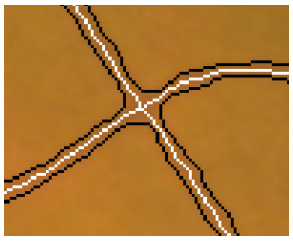
3.4 Vessel segment extraction (junctions removal)

After vessel centerline detection, bifurcation and crossover points are excluded, since they would require special treatment. Bifurcation points are characterized by having 3 connected neighbors (3.3b) whereas crossover points have 4 or more neighbors (3.3c), considering a 8-connection scheme. As so, in order to exclude these points from the vessel centerline image, pixels with three or more neighbors are removed. The removal of these points divides the vessel centerlines into individual vessel segments. These segments correspond to a pixel-thin group of connected pixels limited by two end points, i.e., by pixels with a single neighbour. An example of removal of vessel junctions in a retinal image is shown in Fig. 3.4.

(a) Single vessel point with 2 neighbors.



(b) Bifurcation vessel point (3 neighbors).



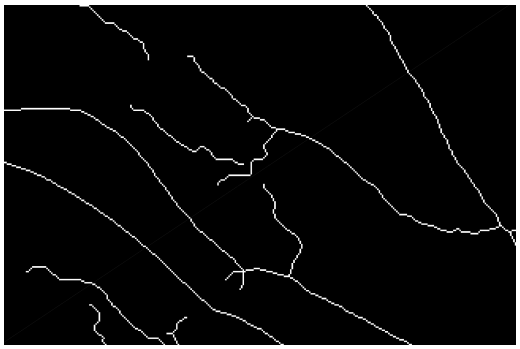
(c) Crossover vessel point with 4 neighbors.

Figure 3.3: Examples of vessel point configurations (regions selected from Fig. 3.2f).

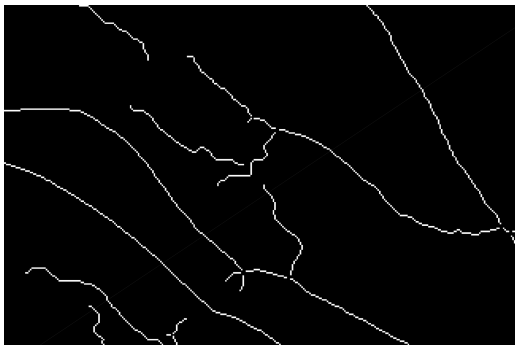
The obtained segments are then refined by removing short spurs that may have resulted from the thinning process. Similarly to Bankhead *et al.* (2012), all terminal segments, i.e., that contain end points, shorter than 10 pixels are removed. These structures usually result from irregularities on the vessel segmentation and thus are not of interest. Longer structures are kept because they may correspond to short vessel segments.

3.5 Vessel intensity profile extraction

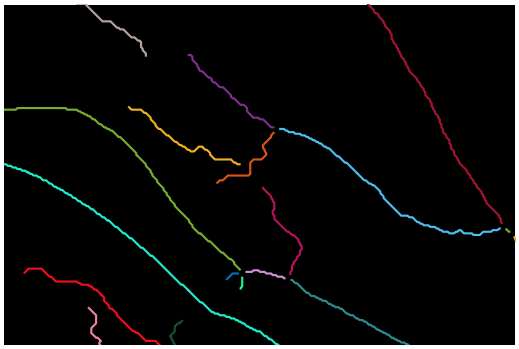
Once vessel segments are found, intensity profiles are determined perpendicularly to the vessel centerlines. For a given vessel segment, at each center point, the intensity values along the normal to the centerline at that point are obtained. In order to do this, we use a process similar to the one presented in Bankhead *et al.* (2012). First, for each object, i.e., segment, in the image resulting



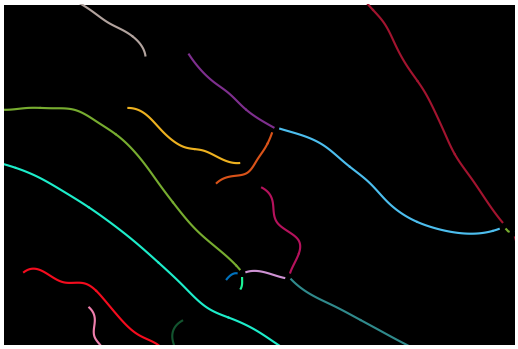
(a) Thinned vessels obtained for a region of the CLRIS001 image (REVIEW).



(b) Vessel segments, after junction removal from 3.4a.



(c) Vessel segments of 3.4b, labeled with different colors.



(d) Vessel segments of 3.4c after spline approximation.

Figure 3.4: Example of removal of junctions from a thinned vessel image and vessel segment smoothing through spline fitting. Colors are used to better distinguish between vessel segments.

from the junction removal algorithm, the pixels are ordered starting from an end pixel to the other. This is done to obtain an ordered sequence of segment pixels to which the spline approximation will be applied to smooth the vessel segments.

3.5.1 Vessel segment smoothing

The extraction of the intensity profiles requires knowledge of the vessel orientation. A simple approach would be to compute the derivatives at the pixel-discrete vessel segments. However, this process may retrieve inaccurate results because very abrupt changes can occur between one pixel and its neighbor if the segment is not properly smoothed. In this work, spline fitting is applied to each segment to smooth the vessel and avoid this problem, as in Bankhead *et al.* (2012).

Splines are piecewise-polynomial real functions, and are one of the most successful approximating functions for practical applications (Rice (1969)). Consider a spline defined in an interval $[a,b]$, composed of n subintervals. For each of the n intervals the spline is defined by a polynomial function. The highest order, k, of all the polynomials (one per interval) determines the order of the spline. Furthermore, consider the points x_0, x_1, ..., x_n, such that $x_0 = a$, $x_n = b$ and $x_0 < x_1 < ... < x_n$. Then, a spline function of degree k with knots x_0, $x_1, ..., x_n$, S_k, is a function

that: 1) is a polynomial of degree $\leq k$ in each interval $[x_{i-1}, x_i]$; 2) has continuous derivatives of all orders up to $k-1$ on $[a,b]$. For instance, a cubic spline, $S_3(x)$ is defined by:

$$S_3(x) = \begin{cases} a_1 + b_1 + c_1 x^2 + d_1 x^3, \ x \in [x_0, x_1] \\ a_2 + b_2 + c_2 x^2 + d_2 x^3, \ x \in [x_1, x_2] \\ \ldots \\ a_n + b_n + c_n x^2 + d_n x^3, \ x \in [x_{n-1}, x_n] \end{cases} \tag{3.1}$$

$S_3(x)$ is continuous, and has continuous first and second derivatives. The main requirement in this spline approximation is to correctly choose the polynomials in order to originate a sufficiently smooth spline. By adjusting the space of the intervals, i.e., the number of polynomial pieces, one can control the degree of smoothness in detriment of the ability to fit complex shapes.

In this work, least-squares cubic spline approximation is performed. Lee's centripetal scheme is used for parametrization (Lee (1989)). This approach uses the accumulated square root of the chord length. In this case, the chord length is defined as the euclidean distance between two consecutive segment pixels, thus having as value 1 or $\sqrt{2}$. Consider a vessel segment that has N points. The coordinates of the point j given by $p(j) = (c(j), r(j))$, with $r(j)$ and $c(j)$ denoting the row and the column coordinates of the point, respectively. A spline is approximated to each of the coordinates, c and r, separately. The spline approximation problem consists in minimizing

$$\sum_{j=1}^{N} |S_{3_c}(x(j)) - c(j)|^2 \tag{3.2}$$

to obtain the first spline, S_{3_c}, that smooths the column coordinates of the segment, and

$$\sum_{j=1}^{N} |S_{3_r}(x(j)) - r(j)|^2 \tag{3.3}$$

to obtain the second spline, S_{3_r}, that smooths the row coordinates of the segment. As referred, x is found using the centripetal parametrization scheme:

$$x(j) = \sum_{i=1}^{j} \sqrt{\|p(i+1) - p(i)\|_2} \tag{3.4}$$

We determine the number of polynomial pieces of the spline based on the length of the segment in the image, i.e., $\sum_{i=1}^{N} \|p(i+1) - p(i)\|_2$. The number of pieces is obtained by dividing the length of the segment by 20 pixels. This 20 pixel spacing was found to retrieve an acceptable smoothing for the tested images, which was assessed by visual inspection of the results. In Fig. 3.4 one can see the effect of the spline fitting on the vessel centerlines. An example of a full image with the corresponding smoothed vessel segments is shown in Fig. 3.5.

Once the splines are fitted to the vessel segments, the new smoothed segment points are retrieved and the first derivatives of the splines at these points are computed. These points' coordinates correspond to the values of S_{3_c} and S_{3_r} at the integer values of the spline domain. From

the direction of the vessel at a given point, the normal at that vessel point can be determined. The normal has a slope of $-1/m$, where m is the slope of the tangent to the vessel.

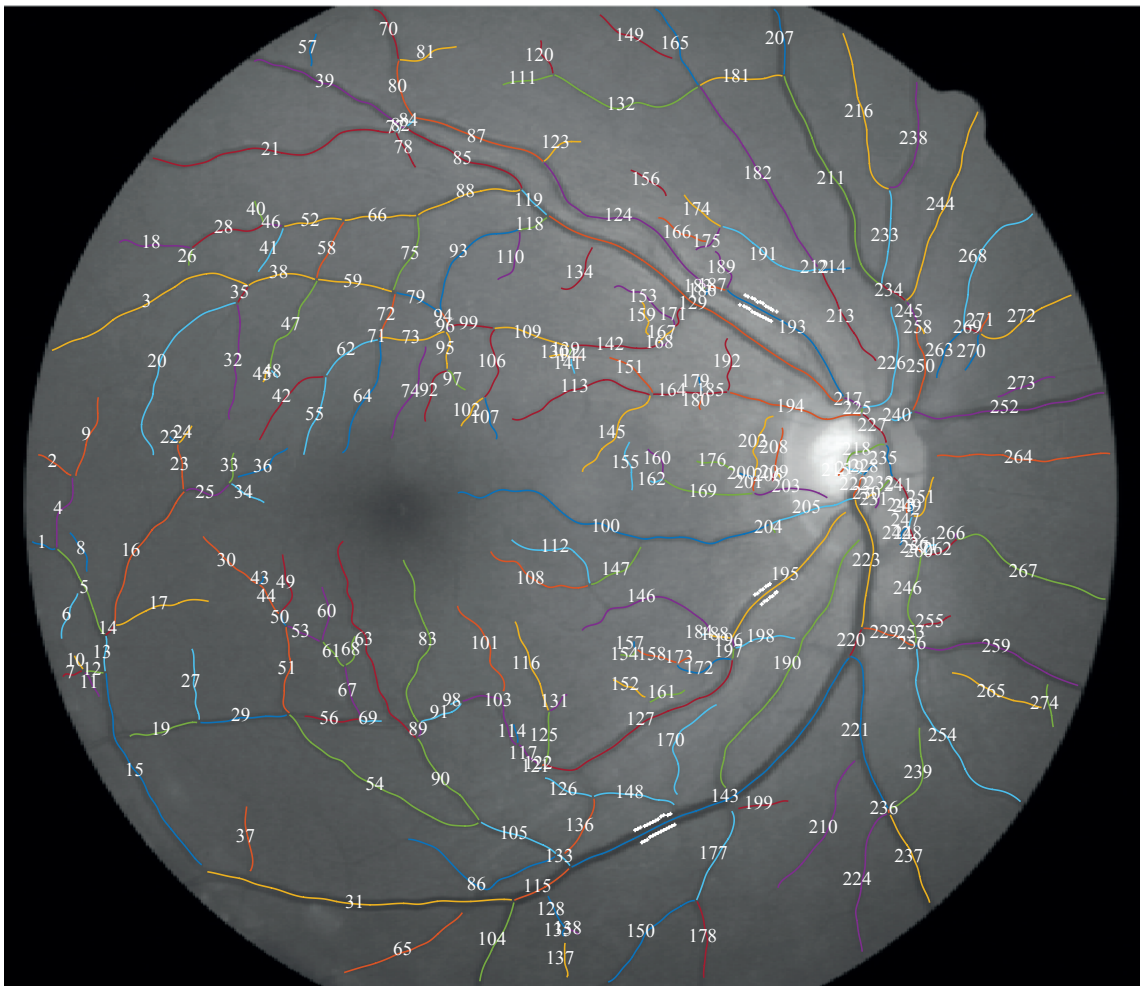


Figure 3.5: Smoothed vessel segments obtained for CLRIS001 image (REVIEW dataset). Segments are numbered and overlapped with the green channel of the original RGB image. Colors are used to help distinguish the segments (note that different segments may be represented in the same color). White marks along some vessel edges represent ground truth points marked by observers.

3.5.2 Profile extraction

Finally, the intensity profiles along the normals to the segments are determined (Fig. 3.6b). The green channel of the RGB image is used, since it is known to present higher contrast between vessel structures and the background. The intensities along the normal to the vessel are obtained with 1 pixel spacing and by applying bilinear interpolation of the intensities at non-integer locations.

The length of the profiles for a given image is determined based on the binary mask containing the segmented vessels. The method used guarantees that this length is larger than the largest vessel present in the image: 1) the minimum distance between each vessel segment point and the vessel contour is computed; the preliminary diameter at each point is considered to be $2\times$ the

computed distance; 2) the preliminary profile length for a given image is computed as being $3\times$ the maximum of the computed diameters for that image. This ensures that a good margin is kept, being the profile length for an image large enough to fully contain the largest vessel in that image as well as possibly useful background information. Fig. 3.6c shows a surface constituted by the 1D profiles extracted from the segment of Fig. 3.6b, stacked together in parallel to each other, aligned by their center points. In Fig. 3.6d the top view of the surface is shown, along with the ground truth marked by the observers. This image is a straightened vessel image, where the all the profiles of the segment lay horizontally.

3.5.3 Determination of profile lengths

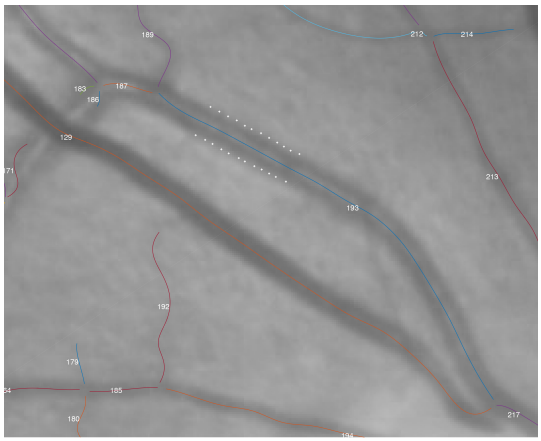
After determining the intensity profiles for each vessel center point, the profile lengths of each segment are more accurately determined. This is important since unnecessary information should not be included in the profiles when model fitting is performed. Note that a profile may include more than one vessel. Consequently, there is the need to identify the region containing only the vessel of interest. For that, we use a method based on peak search on the profiles, that is described bellow. A method based on the same main principles has been described in Lupascu *et al.* (2013).

3.5.3.1 Preparation of the profiles

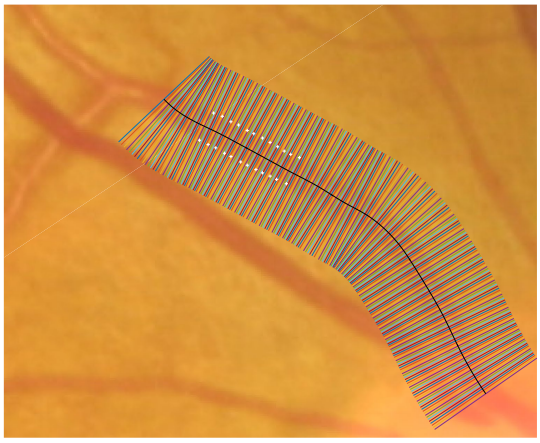
The profile length determination algorithm is applied on the mean of the vessel profiles along 11 adjacent sections, 5 for each side of the profile in study. This is performed since the averaged profiles offer less noise than the individual one. Further, the mean profile is smoothed using Savitzky-Golay filtering (Savitzky and Golay (1964), Orfanidis (1996)). This filter minimizes the least-squares error by fitting a polynomial to the noisy data and allows to increase the signal-to-noise ratio without excessively distorting the signal. The amount of smoothing is controled by the polynomial order and filter size. The polynomial order is determined based on the initial diameter estimation (see subsection 3.5.2): the larger the profile, the lower the order, the higher the amount of smoothing. This ensures that images with larger vessels have their profiles more smoothed than images with thinner vessels, thus guaranteeing that the information of thinner vessels is maintained. Then, the minima and maxima in the smoothed mean profile are detected. A profile point is considered a local extreme if it is smaller or larger than its two neighboring points.

3.5.3.2 Search for CLR regions

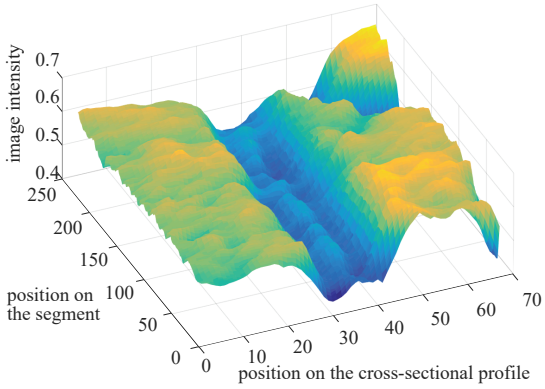
After smoothing, a search for the vessel region contained in the profile is performed. However, a simple search for a typical vessel region, i.e., defined by the positions of two maxima adjacent to a minimum, does not account for the possible existence of CLR, resulting in wrong detections on vessels with this characteristic. The profile of a vessel with CLR can be characterized as a region containing a maximum with one adjacent minimum and maximum on each side, as shown in Fig. 3.7a. A search for CLR regions in the smoothed profile is performed. First, the CLR center, corresponding to a maximum in the center region of the vessel, is detected. Assuming that



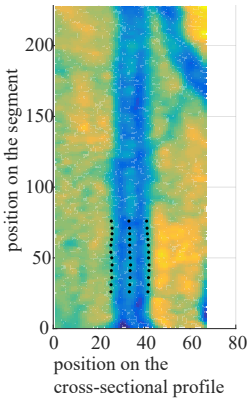
(a) Segments from CLRIS001 (REVIEW) (region of Fig. 3.5); white marks: ground truth.



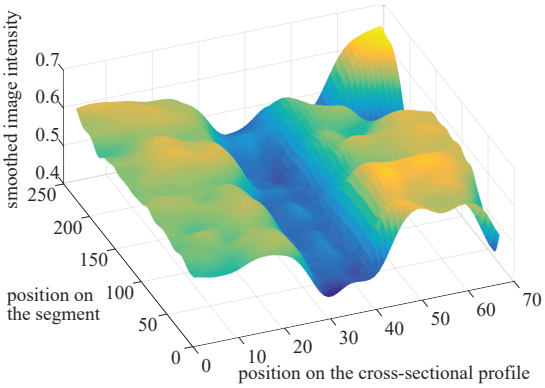
(b) Profile directions obtained for the segment in 3.6a; white marks:ground truth.



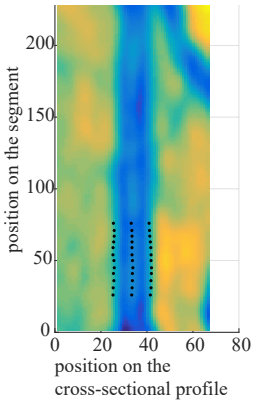
(c) Intensity profiles for all points in the segment in 3.6b, stacked in parallel to each other.



(d) Top view of the surface from 3.6c; black marks: ground truth.

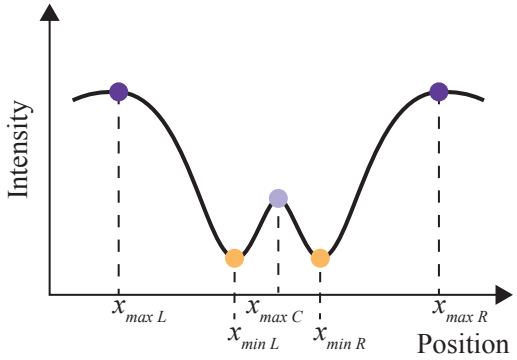


(e) Smoothed intensity profiles of the segment in 3.6b, using anisotropic Gaussian filtering.

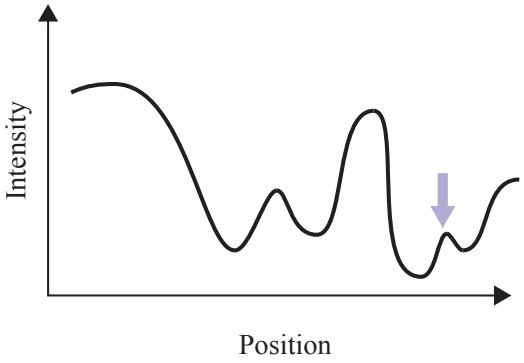


(f) Top view of the surface from 3.6e; black marks: ground truth.

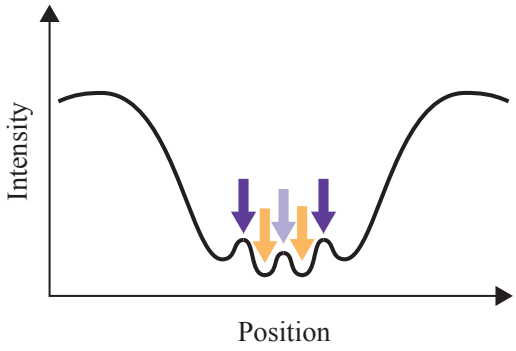
Figure 3.6: Example of vessel intensity profile determination and smoothing of the extracted profiles. Colors in the plots are representative of the intensity values: warmer colors represent higher intensity whilst cooler colors represent lower intensity.



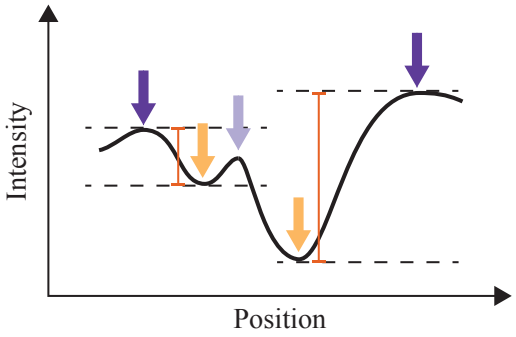
(a) Typical shape of a vessel with CLR, along with its extreme point positions.



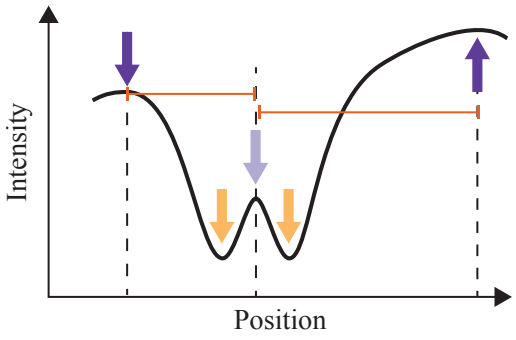
(b) The lowest maximum of the profile is not the correct CLR center.



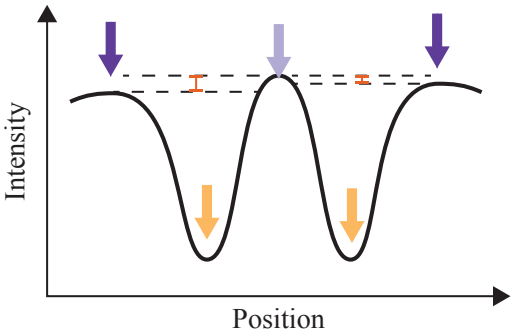
(c) The found minima positions are too close to each other to constitute a CLR region.



(d) The two bumps of the CLR have an intensity difference larger than the acceptable.



(e) The distances between the maxima and the vessel center are too different.



(f) The elevation in the CLR center has larger intensity than the vessel limits.

Figure 3.7: Typical shape of a vessel with CLR and examples of cases where the conditions imposed for CLR detection are violated. The arrows indicate the locations of the peaks that would define the CLR region, if one of the conditions had not been violated.

all vessel maxima are always lower than background maxima, one can simply detect the maximum with lowest intensity, and consider its position to be the CLR center. However, other vessels may be present in the profile (see Fig. 3.7b), possibly leading to a wrong maximum detection. To avoid this, the maximum closest to the center of the profile is detected, and chosen instead of the lowest profile maximum if it is not too far from the profile center and it is close in value to the lowest maximum.

Then, the two adjacent minima to the CLR center are detected, one to the left and one to the right of that maximum (Fig. 3.7a). Finally, the two adjacent maxima are detected, one to the right of the right minimum and one to the left of the left maxima. These maxima positions are considered the limits of the vessel. A set of conditions is established to avoid the recognition of false CLR regions. First, the locations of the two minima should have a minimum distance. This prevents misclassifications as CLR if the peaks are too close. (Fig. 3.7c). The depths of the two bumps of the CLR should not differ too much, in order to avoid large intensity differences between the two sides of the CLR (Fig. 3.7d). Then, the distances between the vessel center and the two maxima should not differ more than a established value (Fig. 3.7e). Besides, the maxima should not be too far from the vessel center. If only one of the maxima is too far away from the center, what happens in the other side of the vessel is replicated, symmetrically to the vessel center. These conditions avoid a big asymmetry between the two sides of the CLR. Additionally, the elevation of the CLR center should not surpass the limits of the vessel (Fig. 3.7f). If no such region containing the vessel center is found, the algorithm proceeds by searching for non-CLR vessels. Otherwise, the length of the profile is computed as follows: $prof_{length} = 2 \times mean(|x_{maxL} - x_{maxC}|, |x_{maxR} - x_{maxC}|)$, where x_{maxL} and x_{maxR} correspond to the positions of the maxima that limit the vessel region and x_{maxC} to the position of the central minimum (see Fig. 3.7a). The parameters used in these rules were obtained by experimentation, having achieved good results in the tested images.

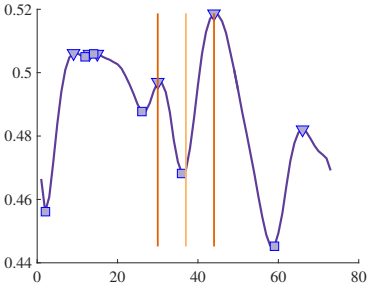
3.5.3.3 Search for non-CLR regions

If no CLR vessel is found, the process restarts, this time seeking for vessel regions without CLR. The minimum closest to the vessel center and the adjacent maximum on each side are searched. These maxima locations correspond to the limits of the profile. The found peaks are analysed to verify the validity of the limits. If there is no maximum to the left or to the right of the minimum, what occurs in the the other side is replicated. The two maxima must not be too close to the vessel center: if only one of the maxima is too close to the center, what happens in the opposite side is replicated; if both are too close, an iterative search for other maxima to the left and to the right of the vessel is performed until suitable maxima are found. If the above conditions are met, the length of the profile is computed as follows: $prof_{length} = 2 \times mean(|x_{maxL} - x_{minC}|, |x_{maxR} - x_{minC}|)$, where x_{minC}, x_{maxR} and x_{maxL} represent the positions of the central minimum, right maximum and left maximum, respectively. Otherwise, the profile length is considered to be equal to the initial profile length.

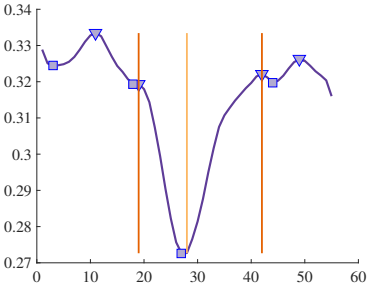
In Fig. 3.8 examples of difficult cases in which the algorithm succeeds and others in which it fails are shown. For example, in Fig. 3.8a and 3.8b, the established conditions allowed to not

wrongly detect a CLR. However, in 3.8d and 3.8e the conditions were not restrict enough and so a CLR was detected in a non-CLR vessel. Fig. 3.8c shows a case where the conditions led to the replication of the right side of the vessel since the left limit was too far away from the center. Fig. 3.8f is an example of a common problem, that is the lack of peaks near the vessel limits. This leads to an overestimation of the profile length.

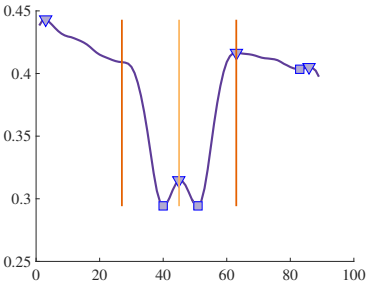
Note that, despite the cases where the profile length is overestimated, the length actually considered in the next steps is the median of the profiles for a given segment, and not the individual profile lengths. This means that the final length is less affected by the overestimation. Further, the overestimation, although not desirable, is preferable to the underestimation, which would lead to the loss of vessel profile information. Before model fitting, the profiles are cut, symmetrically relatively to the centerline, to the determined profile length for that segment, i.e., all profiles belonging to the a given segment have the same length.



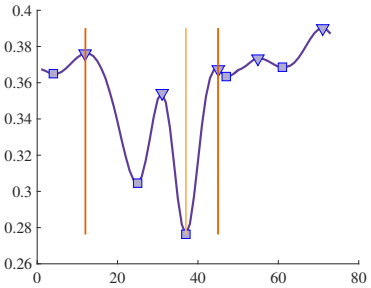
(a) CLR correctly rejected due to the big difference in the depth of the two depressions.



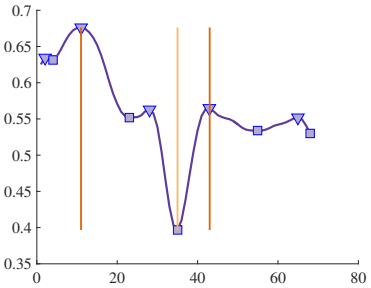
(b) CLR correctly rejected due to the big difference in the depth of the two CRL depressions.



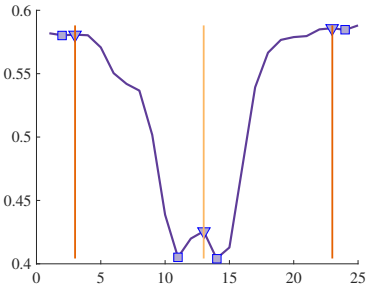
(c) Left limit symmetric to the right, since it was too far away from the center.



(d) CLR wrongly detected (conditions not restricted enough).



(e) CLR wrongly detected (conditions not restricted enough).



(f) Length overestimated (lack of peaks near the vessel limits).

Figure 3.8: Examples of profile length determination results. Top row: successful cases; bottom row: non-successful cases, for which the conditions were not restrict enough. Curve: smoothed mean intensity profile; triangular marks: detected maxima; square marks: detected minima; orange vertical lines: detected vessel limits; yellow vertical line: center of the vessel profile.

3.5.4 Profile smoothing

The obtained vessel profiles can be quite noisy, depending on the image quality, existence of pathologies, etc. To overcome this a smoothing filter is applied to the profiles in order to reduce that noise. The smoothing is applied to the image with the segment profiles stacked together in

parallel to each other. A colormapped version of this straightened image is shown in Fig. 3.6d. In this image, the vessel center points are all aligned vertically in the center column, being that the profiles are horizontal, one per row. As in Bankhead *et al.* (2012), the filter applied in this work is the anisotropic Gaussian filter, as it has different standard deviations along the two dimensions. The anisotropic Gaussian function can be written as follows:

$$g(x,y) = h \times e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}} \tag{3.5}$$

where x is the coordinate along the vessel cross-section, y the coordinate along the perpendicular direction, h the height of the Gaussian, μ_x and μ_y are the center locations of the Gaussian along x and y, respectively, and σ_x and σ_y are the spreads of the Gaussian along x and y, respectively.

This is of special interest for this application, since the straightened image has the profiles oriented horizontally, being possible to apply different degrees of smoothing in the direction of the profiles and in the direction of the vessel, perpendicularly to the vessel cross-section profiles. This is performed by adjusting σ_x and σ_y. Since there is interest in reducing noise but not excessively blurry the edges of the vessels, a larger amount of smoothing is applied on the direction of the vessel than in the direction of the profiles. The values used are computed as function of the estimated profile length for the segment, so that vessels with larger diameters are more smoothed than thinner vessels (Bankhead *et al.* (2012)). In our work, $\sigma_x = \sqrt{0.1 \times prof\_length_{segment}}$ and $\sigma_y = \sqrt{1.5 \times prof\_length_{segment}}$ are used, where $prof\_length_{segment}$ is the median of the individual profile lengths belonging to that segment, computed as described in subsection 3.5.3. These values showed to retrieve good results for the tested images, smoothing the vessels without blurring too much the edges. In 3.6 one can see the effect of the anisotropic Gaussian smoothing filter on a vessel segment.

3.6 Model fitting

The vessel intensity profiles are then approximated by a model by finding the model parameters that lead to the best adjustment between the model curve and the observed profile. These parameters will then be used to estimate the vessel widths. In this work, different models are tested, as it is described in the current section. Further, fitting is tested for both 1D profiles and stacks of adjacent 1D profiles, shortly named as 2D profiles.

3.6.1 Model for vessel profile fitting

Only models with CLR-fitting capability are tested in this work. The Hermite model is one of them, due to its good performance on previous works (Lupascu *et al.* (2013)). Two other models are proposed, based on a Difference-of-Gaussians model.

3.6.1.1 Hermite model (6 parameters)

The first is an adapted Hermite model with 6 parameters presented in Lupascu *et al.* (2013). This model is defined, in 1D, as:

$$h(x) = t + h \times (1 + \beta \times ((x - \mu - \delta)^2 - 1)) \times \frac{1}{\sqrt{2\pi \times \sigma^2}} \times e^{-(\frac{x-\mu}{\sqrt{2}\times\sigma})^2} \tag{3.6}$$

where x is the coordinate along the vessel cross-section, t is the maximum of the function, h is the height of the Gaussian, μ the location of the center, σ the standard deviation of the Gaussian, β is an adaptive parameter controlling the depth of the concavity of the CLR and δ is a parameter that controls the asymmetry of the model.

The same model, in 2D, is defined as:

$$h(x,y) = t + h \times (1 + \beta \times ((x - \mu - \delta)^2 - 1)) \times \frac{1}{\sqrt{2\pi \times \sigma^2}} \times e^{-(\frac{x-\mu}{\sqrt{2}\times\sigma})^2} \tag{3.7}$$

where y is the coordinate along the perpendicular direction and the other parameters mean the same as in Eq. 3.6. The 2D model can be seen as a stack of 1D models along the y axis direction and, consequently, the expression is independent of y. Another way to express the model, by separating its terms, is as follows:

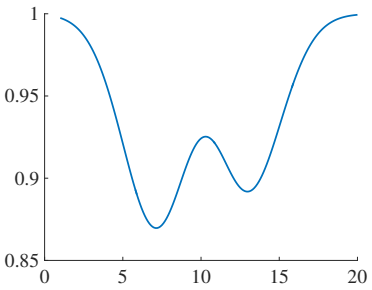
$$\begin{aligned} h(x,y) = t + h \times & \frac{1}{\sqrt{2\pi \times \sigma^2}} \times e^{-(\frac{x-\mu}{\sqrt{2}\times\sigma})^2} - h \times \beta \times \frac{1}{\sqrt{2\pi \times \sigma^2}} \times e^{-(\frac{x-\mu}{\sqrt{2}\times\sigma})^2} + \\ & + h \times \beta \times (x - \mu - \delta)^2 \times \frac{1}{\sqrt{2\pi \times \sigma^2}} \times e^{-(\frac{x-\mu}{\sqrt{2}\times\sigma})^2} \end{aligned} \tag{3.8}$$

where the first term is the main Gaussian, that models the overall vessel shape, the second term is the second Gaussian, that is subtracted to the first Gaussian and models the CLR, and the last is a Gaussian multiplied by a parabola, displaced in x, which controls the model asymmetry. Note that all the Gaussians have the same center and spread. The Gaussian which modulates the CLR has an independent amplitude. However, the third Gaussian, which is multiplied by the parabola, is also multiplied by the amplitude parameter of the CLR Gaussian. The effect of the values of the model parameters in the model shape is shown in Fig. 3.9.

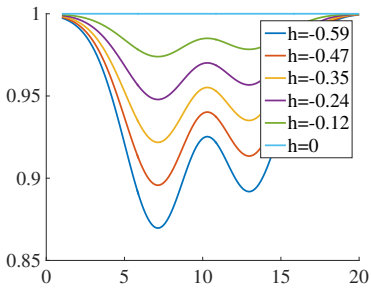
One of the problems of this model is the fact that the second Gaussian (CLR) has the same spread as the the main Gaussian and as the third one. In real vessel profiles this is not necessarily true and often the spread of the second Gaussian is smaller than the spread of the main Gaussian (see Fig.2.4b). Further, the fact that the two last Gaussian are multiplied by the same parameter leads to restrictions in the fitting: if the CLR Gaussian has larger amplitude, the third Gaussian amplitude will also be affected. An example of the fitting result to a vessel profile using the Hermite model is shown in Fig. 3.12a and 3.12b. Although the model fits well simpler profiles, vessels with CLR are often worse fitted by this model, as it is the case in the figure.

As referred, one can fit a 1D model or a 2D one. One advantage of fitting a 2D model is the robustness in the presence of noisy data, introducing some smoothing in the fitting. In this case, instead of fitting a 1D model to a 1D cross-sectional profile of the vessel, we fit a 2D model to a

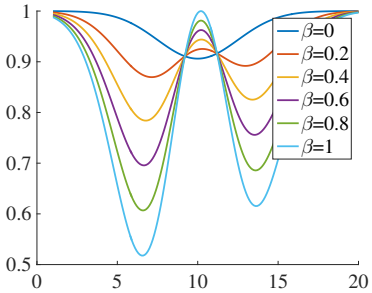
2D cross-sectional region of the vessel composed by multiple neighboring 1D profiles. Note that it is identical to use a 2D model and fit it to the 2D stack of profiles or to consider the points of all the neighboring profiles in the same plane and fit the 1D model, since the equation of the 2D model is independent of y. In this work we do the latter, as seen in Fig. 3.12.



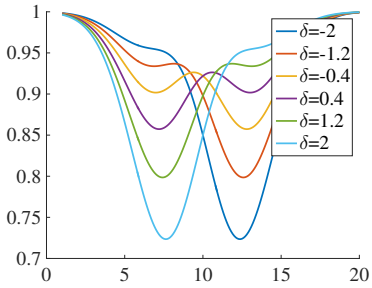
(a) Initial profile (t=1, h=−0.588, β=0.2, μ=10, δ=0.2, σ=2.5).



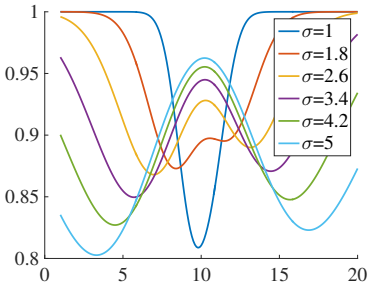
(b) Effect of the amplitude of the main Gaussian (h).



(c) Effect of the amplitude of the two other Gaussians (β).



(d) Effect of the CLR asymmetry (δ).



(e) Effect of the spread of the Gaussians (σ).

Figure 3.9: Examples of 1D curves of the Hermite model, as defined in Eq. 3.6. A profile length of 20 pixels was set to the vessel. In each plot a parameter is varied at a time, with the remaining parameters fixed, in order to evaluate the influence of that parameter in the overall model shape.

3.6.1.2 Modified DoG model with 7 parameters (DoG-L7)

A new model consisting in a modified Difference-of-Gaussians (DoG) model is proposed. Although the common DoG model takes into account the CLR, it does not allow asymmetry between the vessel edges. This assymmetry is in fact present in some vessel profiles, which can lead to a poor fitting. This new model takes an adapted DoG (constrained in some parameters) and multiplies it by a line, in order to achieve the desired asymmetry in the vessel edges. This model is define as:

$$m_1(x,y) = (t + h_1 \times e^{-(\frac{x-\mu}{\sqrt{2}\times\sigma_1})^2} - h_2 \times e^{-(\frac{x-\mu}{\sqrt{2}\times\sigma_2})^2}) \times (\lambda \times (x-\mu) + t) \tag{3.9}$$

where x is the coordinate along the vessel cross-section, t is the maximum of the function, h_1 is the height of the first (main) Gaussian, μ the location of the center, σ_1 the spread of the first Gaussian, h_2 the height of the second Gaussian, σ_2 the spread of the second (CLR) Gaussian and λ

is the slope of the multiplying line. As can be seen, the means of the two Gaussians are the same, centering the light reflex in the center of vessel. The effect of the values of the model parameters in the overall shape of the curve is shown in Fig. 3.10. An example of the fitting result to a vessel profile using the DoG-L7 model is shown in Fig. 3.12c and 3.12d.

3.6.1.3 Modified DoG model with 8 parameters (DoG-L8)

Although the modified DoG model from Eq. 3.9 behaves fairly well on the tested profiles, it does not allow to control in a good extent the asymmetry in the CLR. This means that in cases such as the one shown in Fig. 3.12 the central part of the vessel is not very well fitted by the model. To overcome this, another parameter is added to this model. In this work we chose to allow the CLR Gaussian to have a different mean from the main Gaussian.

$$m_2(x,y) = (t + h_1 \times e^{-(\frac{x-\mu_1}{\sqrt{2}\times\sigma_1})^2} - h_2 \times e^{-(\frac{x-\mu_2}{\sqrt{2}\times\sigma_2})^2}) \times (\lambda \times (x-\mu_1) + t) \tag{3.10}$$

where μ_1 the center (mean) of the first Gaussian, μ_2 the height of the second Gaussian and the other parameters have the same meaning as in Eq. 3.9.

This new parameter allows the displacement of the CLR Gaussian relatively to the main one, being able to model the desired asymmetry. The influence of this new parameter in the shape of the model is shown in Fig. 3.11. In Fig. 3.12e and 3.12f one can see the result of the fitting a profile using the DoG-L8 model. It is visible that the central region is now being very appropriately adjusted by the model curve.

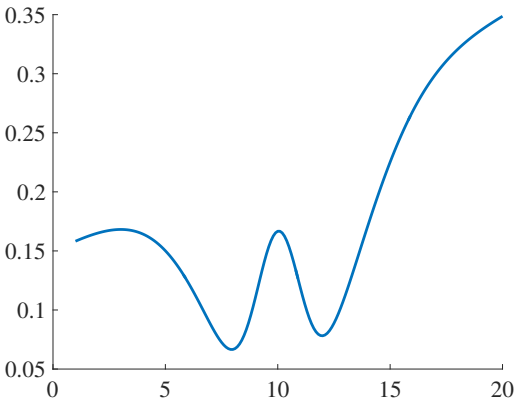
3.6.1.4 Parameter ranges and initialization

Before fitting, values for parameter initialization and parameter ranges should be defined. For the Hermite model, the same parameter ranges and initialization of Lupascu *et al.* (2013) are used:

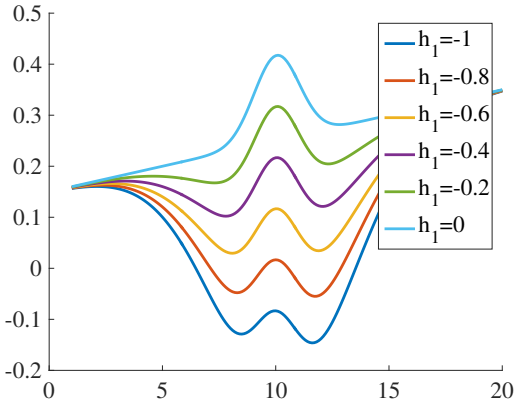
- $t \in [0, 1]$, t_{init}=max($mean\_profs$);
- $h \in [-0.588, 0]$, h_{init}=$-$(max($mean\_profs$)$-$min($mean\_profs$));
- $\beta \in [-1, 1]$, β_{init}=0;
- $\mu \in [0, prof\_length]$, μ_{init}=$center\_prof$;
- $\delta \in [-2, 2]$, δ_{init}=0.2;
- $\sigma \in [1, 15]$, σ_{init}=$prof\_length$/std($mean\_profs$),

where $mean\_profs$ is the mean of the adjacent profiles used in the 2D fitting, $center\_prof$ is the estimated center of the profile and $prof\_length$ is the estimated length of the profile.

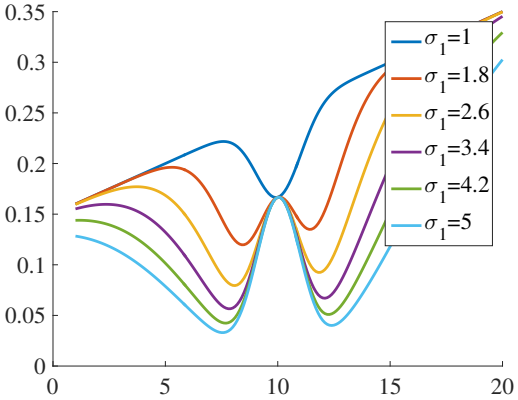
Regarding the proposed models, the optimum range of values for each parameter is defined based on the known relationship between some of the parameters and the common vessel profiles. To confirm that the ranges were broad enough, the observation of a large number of vessel intensity



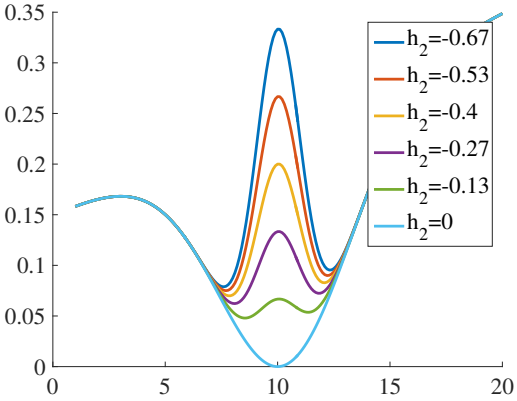
(a) Initial profile (t=0.5, h_1=−0.5, μ =10, σ_1=3, h_2=−0.33, σ_2=1, λ=0.02).



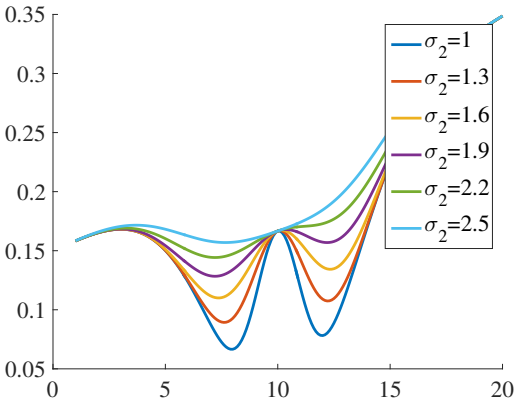
(b) Effect of the amplitude of the 1^{st} Gaussian (h_1 parameter).



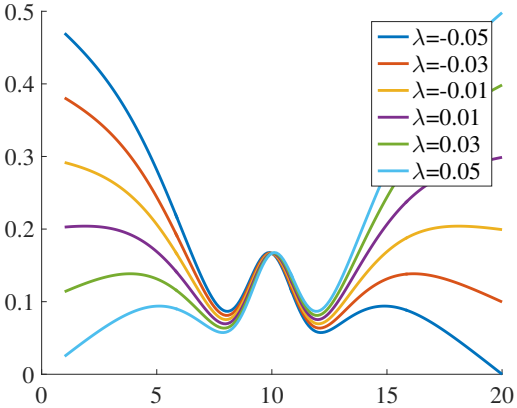
(c) Effect of the spread of the 1^{st} Gaussian (σ_1 parameter).



(d) Effect of the amplitude of the 2^{nd} Gaussian (h_2 parameter).



(e) Effect of the spread of the 2^{nd} Gaussian (σ_2 parameter).



(f) Effect of the slope of the multiplying line (δ parameter).

Figure 3.10: Examples of curves of the modified DoG model with 7 parameters (DoG-L7), as defined in Eq. 3.9 but in 1D. A profile length of 20 pixels was set to the vessel. In each plot a parameter is varied (inside a established range) at a time, with the remaining parameters fixed, in order to evaluate the influence of that parameter in the overall model shape.

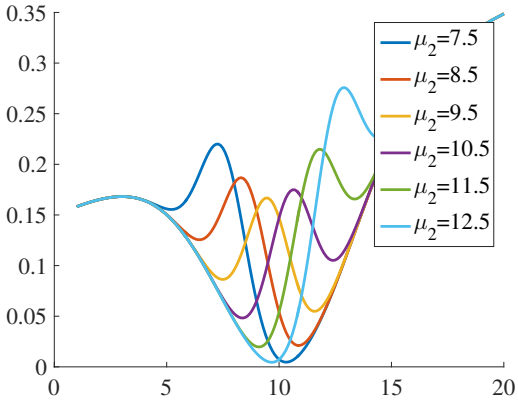


Figure 3.11: Example of curves of the modified DoG model of 8 parameters (DoG-L8), as defined in Eq. 3.10 but in 1D. The plot shows the effect of the variation of the μ_2 parameter. The remaining parameters are fixed: t=0.5, h_1=−0.5, μ_1 =10, σ_1=3, h_2=−0.33, σ_2=1, λ=0.02. The effect of the variation of the other parameters can be seen in Fig. 3.10.

profiles was performed. The chosen parameter ranges and initialization for the modified DoG model with 7 parameters are the following:

- $t \in [0, 1]$, t_{init}=max($mean\_profs$);
- $h_1 \in [-1, 0]$, $h_{1_{init}}$=−(max($mean\_profs$)−min($mean\_profs$));
- $\mu \in [0, prof\_length]$, μ_{init}=$center\_prof$;
- $h_2 \in [-1/1.5, 0]$, $h_{2_{init}}$=$-1/3$;
- $\sigma_1 \in [1, prof\_length/4]$, $\sigma_{1_{init}}$=(1+$prof\_length$/4)/2;
- $\sigma_2 \in [1, prof\_length/8]$, $\sigma_{2_{init}}$=$\sigma_{1_{init}}$/3;
- $\lambda \in [-1/prof\_length, 1/prof\_length]$, λ_{init}=0,

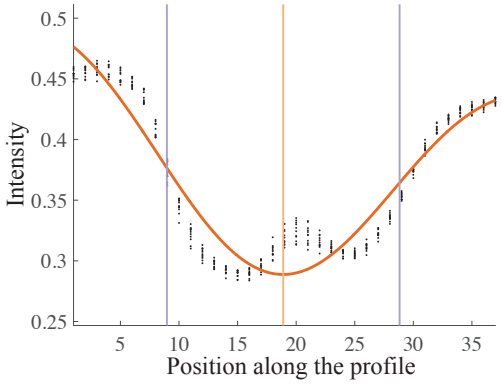
where $mean\_profs$ is the mean of the adjacent profiles used in the 2D fitting, $center\_prof$ is the estimated center of the profile and $prof\_length$ is the estimated length of the profile. For the 8 parameter model, the parameter ranges are the same as for the other DoG model, and the additional parameter, μ_2, has the same range and initialization as μ.

3.6.2 Model parameter optimization

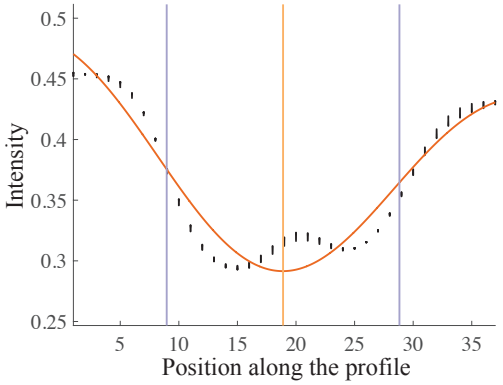
In order to find the parameters of the best-fit model to the vessel profiles, one can solve a non-linear least squares problem. The solution consists in the set of parameters that minimize the sum of the squared differences, defined as:

$$s(\boldsymbol{\beta}) = \sum_{i=1}^{m} [y_i - f(x_i, \boldsymbol{\beta})]^2 \tag{3.11}$$

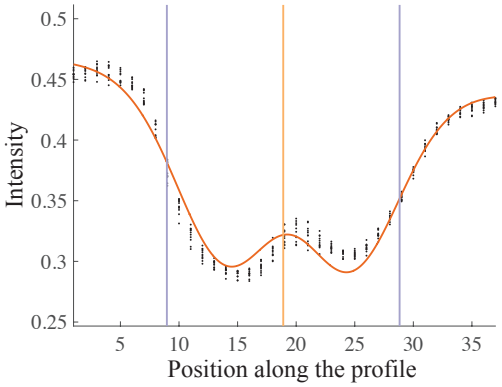
where (x_i, y_i) are empirical data pairs, m is the number of points, $f(x, \boldsymbol{\beta})$ is the model curve and $\boldsymbol{\beta}$ are parameters of the model curve.



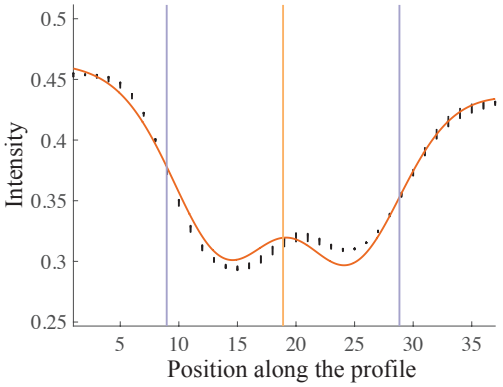
(a) Fitting of the original data from 11 adjacent profiles using the Hermite model (6 parameters) from Eq. 3.7.



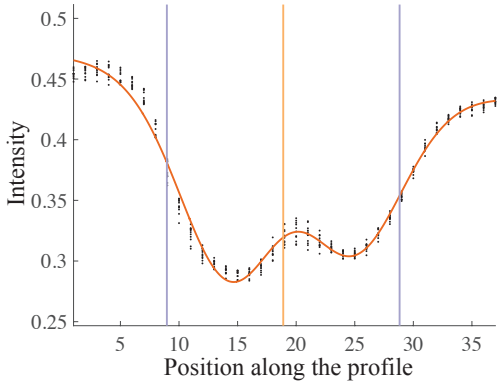
(b) Fitting of the smoothed data from 11 adjacent profiles using the Hermite model (6 parameters) from Eq. 3.7.



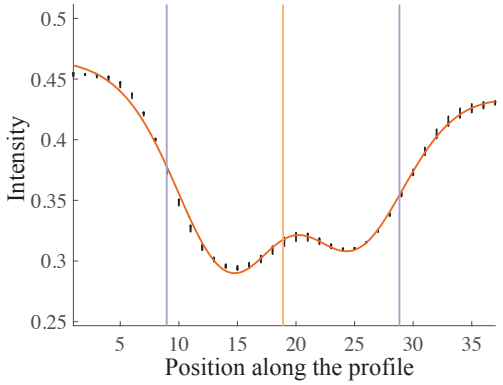
(c) Fitting of the original data from 11 adjacent profiles using the DoG-L7 model from Eq. 3.9.



(d) Fitting of the smoothed data from 11 adjacent profiles using the DoG-L7 model from Eq. 3.9.



(e) Fitting of the original data from 11 adjacent profiles using the DoG-L8 model from Eq. 3.10.



(f) Fitting of the smoothed data from 11 adjacent profiles using the DoG-L8 model from Eq. 3.10.

Figure 3.12: Examples of fittings of the three referred models to a vessel profile from CLRIS002 image (REVIEW), using the parameter ranges and initialization referred in subsubsection 3.6.1.4. Both profile fitting with and without profile smoothing are shown. Black dots: profile data points; orange curve: fitted curve through Trust-Region-Reflective method (see subsection 3.6.2); vertical yellow line: center of the profile; vertical purple lines: ground truth edge positions.

There are several algorithms to solve the least squares problem (see subsection 2.1.4). Examples of methods commonly used for this purpose are the Gauss-Newton, Levenberg-Marquardt (LM) and the Trust-Region-Reflective (TRR). Although other algorithms would be suitable to find the best-fit model to the profile points, such as the LM algorithm, the method used in this work to find the parameter values is the Trust-Region-Reflective (Coleman *et al.* (1999)). One of the main reasons for the choice is that region algorithms are robust, reliable and have very strong convergence (Yuan (2000)). They have good performance, retrieving accurate results and being suitable for solving difficult nonlinear problems more efficiently than other algorithms.

The TRR approach is based on the notion that, for minimizing a function $s(\boldsymbol{\beta})$, i.e., searching for $\boldsymbol{\beta}_{i+1}$ that has a smaller function value than the current estimation, $\boldsymbol{\beta}_i$, s can be approximated by a simpler quadratic function, q in the neighborhood N of $\boldsymbol{\beta}_i$. The neighborhood N is the trust region. Both $\boldsymbol{\beta}_i$ and $\boldsymbol{\beta}_{i+1}$ must be inside it. The trust-region subproblem consists in finding the step, $\boldsymbol{u}_i$, that solves

$$min(q_i(\boldsymbol{u})),\ \boldsymbol{u} \in\ N \tag{3.12}$$

$\boldsymbol{\beta}_i$ is updated to $\boldsymbol{\beta}_{i+1} = \boldsymbol{\beta}_i + \boldsymbol{u}$ only if $s(\boldsymbol{\beta}_{i+1}) < s(\boldsymbol{\beta}_i)$. Otherwise, $\boldsymbol{\beta}_i$ is not updated, and the trust region is shrunk. The quadratic function is approximated by its Taylor expansion, typically its first two terms, around $\boldsymbol{\beta}_i$ (Moré and Sorensen (1983)):

$$q_i(\boldsymbol{u}) = \boldsymbol{g}^T\ \boldsymbol{u} + \frac{1}{2}\boldsymbol{u}^T\ \boldsymbol{H}\boldsymbol{u} \tag{3.13}$$

where $\boldsymbol{g}$ is the gradient of s evaluated at $\boldsymbol{\beta}_i$ and $\boldsymbol{H}$ is the Hessian matrix of s at $\boldsymbol{\beta}_i$. The way used to solve Eq. 3.13 in an efficient manner was to restrict N to a 2D subspace M (Branch *et al.* (1999)). In this work, M is considered to be the linear space spanned by m_1 and m_2, where m_1 is the vector with the direction of the gradient of s at $\boldsymbol{\beta}_i$, i.e., g, and m_2 is either an approximate Gauss-Newton direction, found by solving

$$\boldsymbol{H} \cdot\ m_2 = -\boldsymbol{g} \tag{3.14}$$

or a direction of negative curvature, verifying

$$m_2^T \cdot \boldsymbol{H} \cdot\ m_2 < 0 \tag{3.15}$$

The determination of M using this process aims to induce fast local convergence (through the Newton step, Eq. 3.14) and global convergence (through the steepest descent direction, Eq. 3.15) (Coleman *et al.* (1999)).

3.7 Width estimation

Once the best-fit model to the vessel profile is found, the relationship between its parameters and the vessel width is determined. This relation can be found using supervised methods for regression. One example is Support Vector Machine Regression (Vapnik (1995)). Neural Networks can also be applied in regression, being Generalized Regression Neural Network commonly used for function

approximation (Specht (1991)). Decision trees and ensembles of trees are also commonly used methods for regression. These trees have several advantages over other existing models. First, they do not require special prior assumptions on the data distribution (Ikonomovska (2012)). Second, they are less dependent on their parameter values than other methods, not requiring as much tuning of parameters, as it is the case of the support vector machines, for instance. Further, no heavy computation and long training are required. As importantly, a decision tree can be seen as a set of rules describing the relations between the input variables and the target variable. They are easily interpretable, contrarily to other methods, as it is the case of neural networks. However, these trees may overfit data. This motivates the use of ensembles, which increase the predictive power of the regression trees and reduce variance.

In this work we use a supervised learning method consisting of an ensemble of bagged decision trees in order to estimate the width from the parameters of the best-fit surface. Besides the attractive features of the method, it has proven to retrieve good results for this particular application (Lupascu *et al.* (2013)). This classifier learns the mapping from a point in the nD parameter space to the vessel width, where n is the number of parameters of the model. The training set has the values of the n parameters of the best-fit model with the correspondent ground truth diameters.

3.7.1 Ensembles of bagged regression trees

A regression tree (Breiman *et al.* (1984)) is a decision tree where the target variable takes continuous values, typically real values. It is used to approximate real-valued functions, instead of being used for classification, as is the case of classification trees. Regression trees have a single numerical output and can have multiple input variables. The most common trees are the binary regression trees. The process of building this tree is based on binary recursive partitioning. This consists in the iterative partition of the sample space into smaller and smaller groups, i.e., branches or splits (Fig. 3.13). Initially, all training data samples belong to the same partition. Then, the data is placed into the first two branches by answering a binary question. From all possible splits, the algorithm selects the one that minimizes the sum of the squared deviations from the mean of the two partitions. Note that there are alternatives to this greedy-search approach, but this is the most common strategy. This rule is then applied to each of the new branches. Generally, at each step, only one predictor, i.e., independent variable, is used. The process continues until a previously defined stopping rule is verified, as, for example, the decrease in the error being less than a given threshold or one of the nodes having less than a given number of points. The sum of the squared errors for a tree T is given by

$$S = \sum_{c \in \, leaves(T)} \sum_{i \in \, c} (y_i - m_c)^2 \tag{3.16}$$

where y_i are the response variable data points and $m_c = \frac{1}{n_c} \sum_{i \in \, c} y_i$ is the value predicted at leaf c. As referred, by minimizing S at each iteration, one finds the best split.

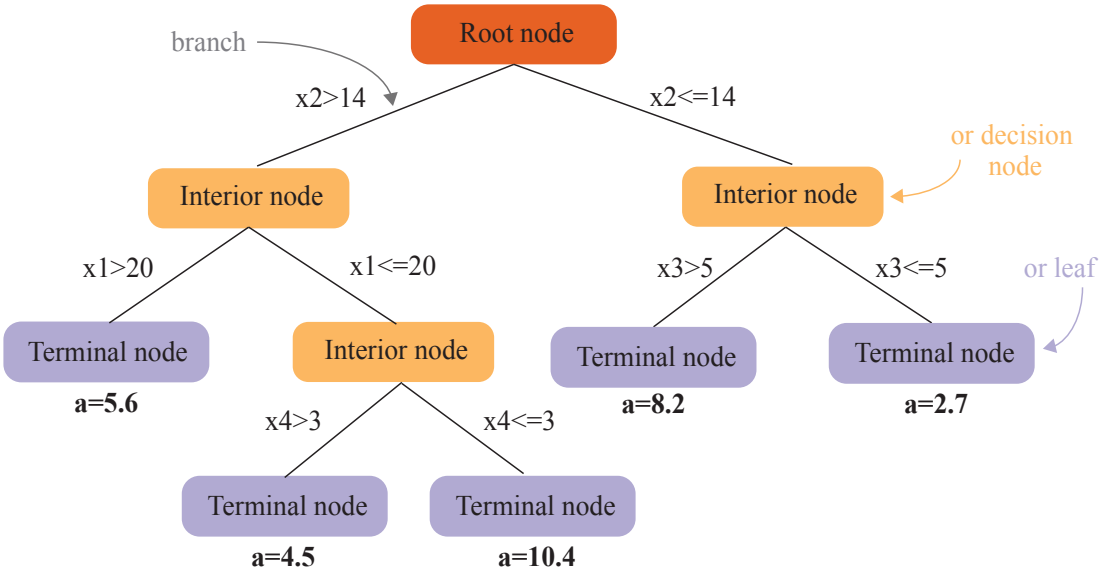


Figure 3.13: Example of a regression tree structure. The main components are labelled. The predictors, i.e., features, are x_1, x_2, etc., and the target variable is a. To train a regression tree, we start at the root node, and ask questions about the predictors. The interior nodes make questions, dictating the splits, and the branches between the nodes contain the answers. The terminal nodes or leafs contain the output values attributed to a.

Ensemble methods allow to combine multiple weak regression trees, that together form a more accurate and robust regression tree (Breiman (1996)). These methods create several regression trees, by extracting different data points from the training set, and in the end combine their outputs. The resulting strong regression tree has a reduced variance compared to the weak trees. One of the most used ensemble methods is bagging, i.e., bootstrap aggregation (Fig. 3.14). This algorithm, despite its simplicity, is very effective. It generates multiple training sets using random sampling with replacement, i.e., generates many bootstrap replicas of the original dataset. Each of the replicas is obtained by randomly selecting N observations from the N observations of the dataset, with replacement. This process omits an average of 37% of the observations for each regression tree ("out-of-bag" observations). A regression tree is found for each of the replicas. In the end, the predictions for the new data are computed as the average of the responses of the individual tree models, being the prediction at a point x, $\hat{f}_{bag}(x)$ given by

$$\hat{f}_{bag}(x) = \frac{1}{B}\sum_{b=1}^{B}\hat{f}_b(x) \tag{3.17}$$

where B represents the number of trees used in the ensemble and $\hat{f}_b(x)$ is the prediction for x of the b^{th} tree.

Regression tree ensemble methods are very powerful, generally showing a better predictive performance than a single regression tree. Additionally, in this work, each tree in the ensemble can randomly select predictors, i.e., features, for the decision splits, a technique known to improve

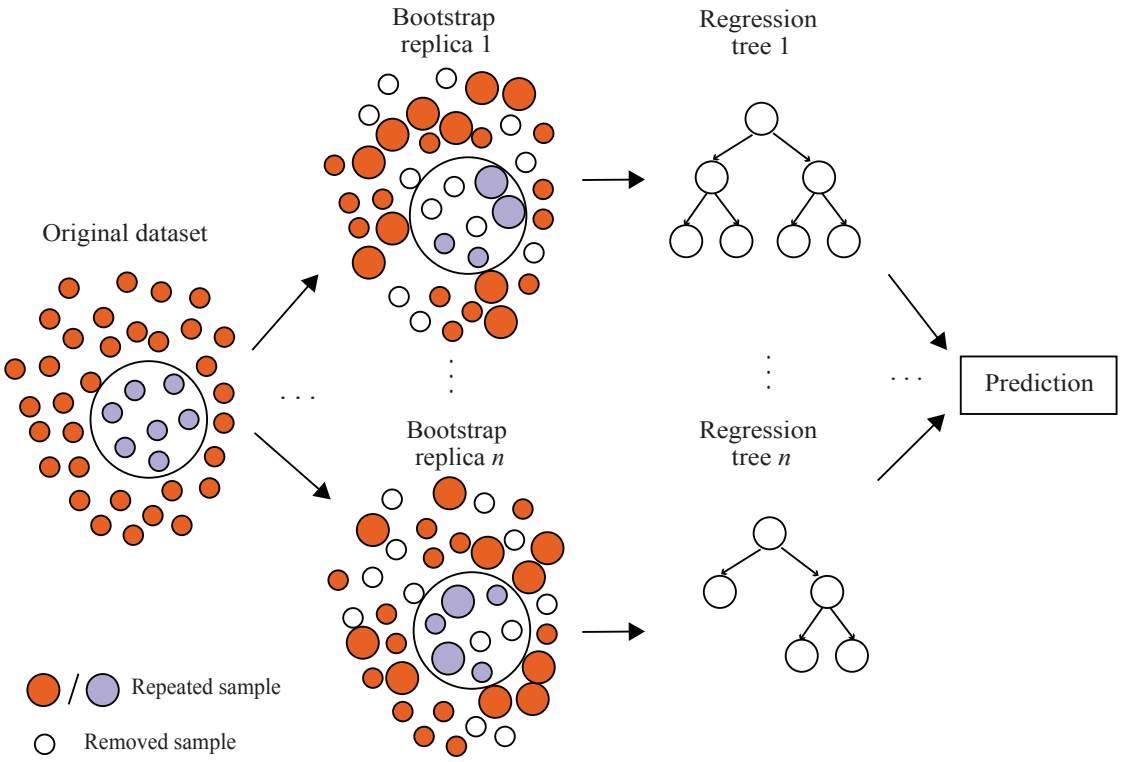


Figure 3.14: Bootstrap aggregation scheme. Multiple bootstrap replicas are drawn from the original dataset. Considering this dataset has N samples, each of the replicas has N samples randomly selected, with replacement. Each of the replicas will be used to train a regression tree. In the end, the results of the individual trees are combined to obtain the prediction (for representation purposes only two classes are considered, represented with different colors.)

the accuracy of the predictions. This method, known as random forests, was used by (Breiman (2001)) along with bagging. It adds additional randomness to bagging, originating a method that's very robust against overfitting and is usually not very sensitive to the parameter values (Liaw and Wiener (2002)). The generalization error of these forests converge to a limit when the number of trees is large (Breiman (2001)).

Several parameters can be adjusted when training bags of trees with random feature selection. In this work, the minimal leaf size is set to 5. This setting grows deep trees, being suitable for the predictive power of a bagging ensemble. Nevertheless, a larger leaf size could generally be used without sacrificing too much the predictive power of the trees and the ensemble. This could be useful to reduce the computational cost of the training and prediction process. Another parameter is the number of predictors selected at random for every decision split. In this work, the number of predictors randomly selected at the splits is set to $1/3$ of total number of predictors, as in Liaw and Wiener (2002). The number of trees used in the ensemble is set to 100. Although ensembles with more trees would still be suitable, the addition of more trees does not significantly influence the error.

3.8 Concluding remarks

The method for retinal vessel width estimation herein presented combines a model fitting-based approach with vessel intensity profile preprocessing steps, and estimates the vessel diameters from the best-fit-model parameters using ensembles of bagged regression trees with random feature selection. The method includes a novel model based on Difference-of-Gaussians for vessel intensity profile fitting, whilst applying some methodologies adapted from the algorithms of Bankhead *et al.* (2012) and Lupascu *et al.* (2013).

First, vessels are segmented from the eye fundus image, resulting in a binary image with the vasculature. This is performed using the algorithm of Mendonça *et al.* (2014). This step detects the vessels where the diameter will be measured, being the main requirement the detection of the majority of the vessels. Then, vessel centerlines are obtained through thinning, followed by junction removal. This leads to the extraction of vessel segments, which will be processed independently. These segments are smoothed using least-squares cubic spline approximation in order to allow a more accurate determination of the vessel direction. Afterwards, for each point in the segment, the cross-sectional vessel intensity profile, i.e, normal to the segment, at that point is extracted. Then, the length of the profiles is determined based on peak detection and established conditions to avoid wrong vessel region detection. Based on the determined profile lengths, the profiles are cut in order to remove unnecessary information, and then smoothed using anisotropic Gaussian filtering. The level of smoothing is dependent on the profile length.

Finally, model fitting is performed, being the parametric model fitted to the observed intensity profiles. In this work, 2D model fitting is applied to the set of points retrieved from neighboring vessel cross-sectional intensity profiles. This is performed since 2D approach is more robust than the 1D one and introduces some smoothness in the fitting. Three models are tested: an Hermite model with 6 parameters and two novel models, consisting in adapted DoG models, multiplied by a line, that are able to model the asymmetry in the vessel edges (7 parameter model) and also the asymmetry in the center of the vessel (8 parameter model). The best fit model parameters are found by solving the non-linear least squares problem, using the Trust-Region-Reflective algorithm, constrained by predetermined parameter initialization and parameter ranges. Once the best-fit model is found, its parameters are used for determining the vessel width. This is performed with a supervised learning method consisting in ensemble of bagged regression trees, combined with random feature selection (random forests). These ensembles preserve the attractive properties of regression trees, such as fast training, little parameter tuning and easy interpretability of the results, whilst reducing variance and overfitting, being very robust.

Chapter 4

Results of the width measurement algorithms

The conceived methodologies described in the previous chapter are evaluated in a publicly available dataset of annotated images. The experimental methodology for evaluating our approach is herein detailed. The results of our method as well as from other state-of-the-art algorithms are presented, discussed and compared in this chapter.

4.1 Evaluation of the results

There are several publicly available datasets of retinal fundus images that contain the ground truth for vessel segmentation, such as DRIVE (Niemeijer *et al.* (2004)) and STARE (Hoover *et al.* (2000)). However, to the best of our knowledge, REVIEW (Al-Diri *et al.* (2008)) is the only one dataset that contains vessel width measurements performed by three observers.

4.1.1 Publicly available retinal image dataset for width measurement evaluation

The REVIEW (Retinal Vessel Image set for Estimation of Widths) dataset (Al-Diri *et al.* (2008)) is the only publicly available dataset that includes retinal vessel width measurements. The dataset is marked by three independent observers, using a specially designed tool for marking edge points and extracting the vessel widths from the marked points. Note that only some randomly selected vessel segments are marked. The dataset is composed by 4 image sets, for a total of 16 images (color fundus photographs) with 193 vessel segments and 5066 cross-sectional blood vessel profiles. These sets have images presenting a variety of pathologies and artifacts. The ground truth for the evaluation is considered to be the mean of the measurements of the three observers in the majority of the works that evaluate width measurement algorithms.

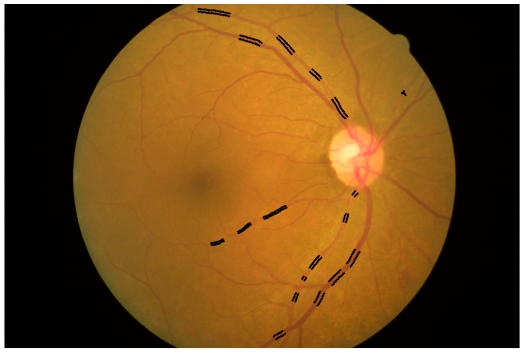
HRIS - The high resolution image set This set contains images (60° FOV) with different grades of diabetic retinopathy. The abnormalities that appear near vessels are one of the challenges for diameter measurement. It contains 4 images (2438×3584 pixels) and 90 segments with 2368

marked profiles. Before being used to evaluate algorithms, these images are usually sampled by a factor of 4, being the vessel widths known to an accuracy of ± 0.25 pixels.

VDIS - The vascular disease image set Contains images (50° FOV) of normal and diseased retina, including diabetic and arteriosclerotic retinopathies. Due to its variety, diameter measurement is challenging. It contains 8 noisy images (1024×1360 pixels), where 6 present diabetic retinopathy, for a total of 79 segments and 2249 profiles.

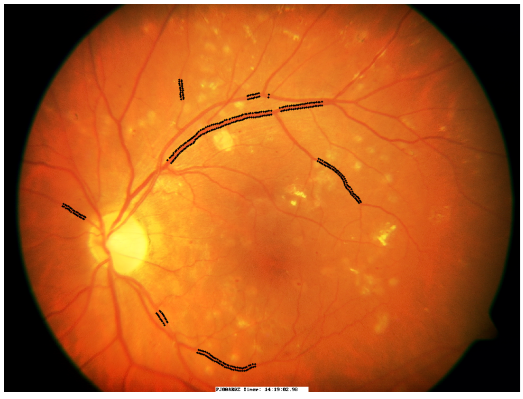
CLIRS - The central light reflex image set It is composed by images (50° FOV) with early atherosclerotic changes with an exaggerated vascular light reflex. It contains 2 images (1440×2160 pixels) with 21 segments and 285 cross-sections.

KPIS - The kick point image set This set is composed by 2 images (60° FOV), of size 288×119 and 170×192 pixels. The images have good quality, as they were retrieved from full retinal 2600×3330 pixel images, and contain 3 segments and 164 cross-sectional profiles that were taken from clean, large and non tortuous vessel segments.

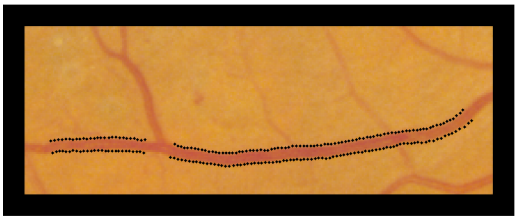


(a) CLRIS dataset (image CLRIS002).

(b) HRIS dataset (image HRIS001).



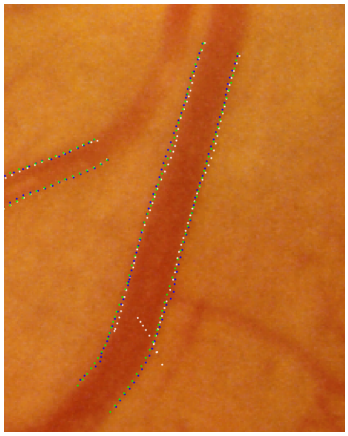
(c) VDIS dataset (image VDIS005).



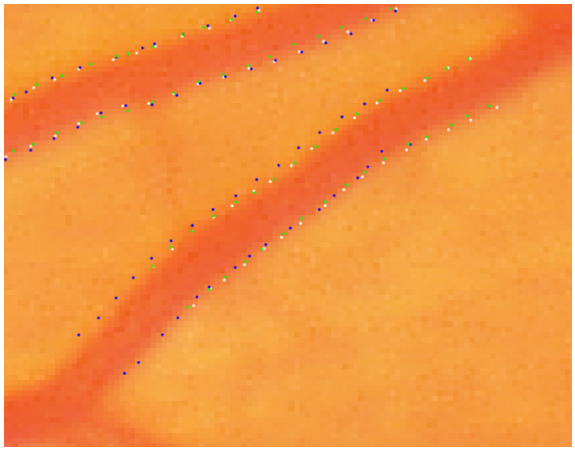
(d) KPIS dataset (image KPIS001).

Figure 4.1: Retinal fundus images selected from the REVIEW dataset. Black marks represent the ground truth edge points defining the vessel width (mean of the annotations from the 3 observers).

Dubious markings on the images There are cases of ground truth marks that look dubious since they are not close to the edges of the vessels. Fig. 4.2 shows examples of the cases we found where the ground truth does not seem correct. These images show the marks, at different colors, of the three observers. As shown in Fig. 4.2b, in some cases the marks of one of the observers are displaced relatively to the other two, leading to a wrong ground truth since the mean of the three does not return points on the edges. Besides, in one of the cases the marks of one observer are not placed along the vessel at some points (Fig. 4.2a). However, to our knowledge, there are no references regarding this problem in other works.



(a) HRIS001 (REVIEW).



(b) VDIS006 (REVIEW).

Figure 4.2: Examples of dubious markings on images from the REVIEW dataset. The images shown correspond to the regions of the original images where the erroneous markings were detected. The marks represent the three observers' annotations, distinguishable by color.

4.1.2 Evaluation metrics

For the current application, it is more relevant that the algorithms retrieve precise results, i.e., with a low standard deviation of the width errors, than accurate, i.e., low mean of the width errors (Lowell *et al.* (2004)). Consistency between the measurements is more significant than absolute diameter: any consistent bias can be compensated, whereas no compensation is possible for fluctuating bias. Consequently, the standard deviation of the point-by-point differences between the measured and the ground truth diameters should be used to evaluate the performance of the algorithms. (Al-Diri *et al.* (2008)). This difference, at given vessel profile i, is given by

$$\chi_i = \omega_i - \psi_i \tag{4.1}$$

where ω_i is the estimated width and ψ_i is the correspondent ground truth. The standard deviation of the width differences is given by

$$\sigma_{error} = \sqrt{\frac{1}{n_p}\sum_{i=1}^{n_p}(\chi_i - \mu_{error})^2} \tag{4.2}$$

where μ_{error} represents the mean of the width differences and is given by $\mu_{error} = \frac{1}{n_p}\sum_{i=1}^{n_p}\chi_i$, being n_p the number of vessel cross-sections, i.e., profiles, evaluated.

The success rate (SR) is commonly used as a measure of stability (Lupascu *et al.* (2013)). It is usually defined as the ratio between the meaningful measurements returned by the algorithm and the total number of measurements. Note that, dependently on the nature of the methods, the works may slightly differ in their definition of SR. For instance, methods that measure the width at a given center point (retrieved from the ground truth, for example), do not return meaningful measurements only when their algorithms do not converge. Differently, methods such as the ones from Al-Diri *et al.* (2009), Xu *et al.* (2011a) and Bankhead *et al.* (2012), perform vessel detection and so, if a vessel is not detected, no measurements are performed and the SR is reduced. Several authors also present the mean and the standard deviation of the width measurements.

In order to evaluate the performance of the algorithm, a correspondence has to be established between each ground truth center point (i.e., center point of the marked edge points) and a point in the detected centerline. Here, we associate each ground truth center point with the closest detected center point, as long as they are within less than 5 pixels from each other and that no other ground truth center point is closer to that detected point. This value is chosen since it is smaller than most of the diameters on the dataset and ensures some margin to account for possible mislocation of the detected center point. This leads to a unique match between ground truth and detected center points, ensuring that each center point is only used once for measurement. In the work of Bankhead *et al.* (2012) a similar scheme is used, but a larger tolerance is given when it comes to the maximum distance between the two points (it has to be less than the true vessel diameter at that point). As referred by the authors, this strict criteria of unique matching between a ground truth and a detected point can lead to a decrease in the SR in cases where the ground truth points have a distance of less than 1 pixel from each other.

4.1.2.1 Further evaluation of the algorithm's performance

Results per range of diameters Other ways of evaluating the results can be of interest to fully understand the behaviour of the algorithm. One of them is the analysis performance of the algorithm for different ranges of diameters. Ideally, the behaviour should be independent of the real vessel diameter, but some algorithms tend to retrieve worse results for a given range of diameters.

One simple way to initially assess the performance of the algorithm in terms of diameter ranges is to compare the distributions of the measured and ground truth diameters. However, this only allows a general idea of the results per diameter and does not retrieve information regarding the error for each of the diameters.

Consequently, evaluation can be performed using Bland-Altman plots of the results, by plotting the differences between the measured and the ground truth widths (χ_i) as a function of the mean of those differences.

Goodness-of-fit Specifically for model fitting-based methods, there is interest in analysing the goodness-of-fit of the model curves to the intensity cross-sectional profiles. This is true since the

determination of the diameters is performed based solely on the model parameters. Consequently, these should represent the profile as accurately as possible, without compromising the performance of the regressor. Different metrics can be used, being the following the most common:

- sum of squares due to error (SSE): it is also called sum of square of residuals and represents the deviation of the data points from the fitted curve. It is given by:

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{4.3}$$

  where n is the number of points in the profile, $\hat{y}_i$ the predicted, i.e., the model, value at point i and y_i the observation value. A smaller value, i.e., closer to zero, means that the model has a smaller random error, being more useful for prediction;

- R-square (R^2): measures how well the fit explains the variation of the data, being given by:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \tag{4.4}$$

  where SSR is the ratio of the sum of squares of the regression, $SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$, SST is the total sum of squares, $SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$, verifying $SST = SST + SSE$, with $\bar{y}$ being the mean of the observations. It is also called the square of the correlation between the observation and the predicted values. It ranges from 0 to 1, with higher values indicating that the model accounts for a greater proportion of variance. Note that if the number of model coefficients increases, the R-square increases without the fitting necessarily improving; to avoid this, the number of degrees of freedom should be accounted for (adjusted R-square);

- adjusted R-square (R^2_{adj}): is based on the the R-square, resulting from an adjustment based on the degrees of freedom. It is given by:

$$R^2_{adj} = 1 - \frac{SSE(n-1)}{SST(v)} \tag{4.5}$$

  where v is the number of residual degrees of freedom, $v = n - m$, with n being the number of data points and m the number of fitting coefficients. This metric can have any value smaller or equal to 1, being that values closer to 1 are indicative of a better fit.

- root mean squared error (RMSE): it is also called fit standard error, and is defined as:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{SSE}{v}} \tag{4.6}$$

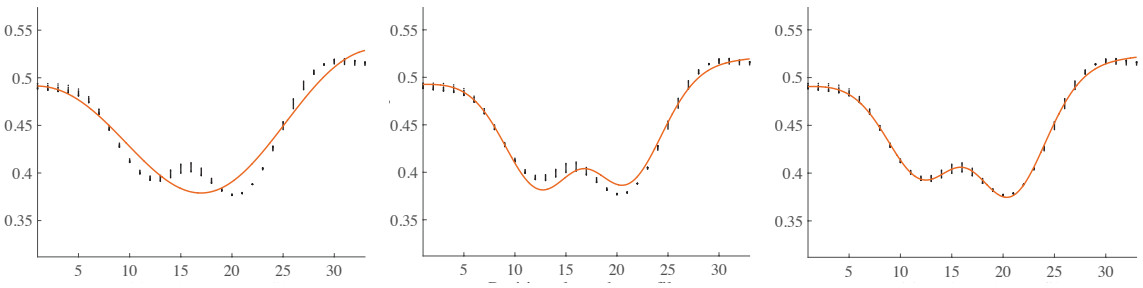
  RMSE values closer to 0 indicate a fit more useful for prediction, as happens with the SSE.

4.2 Experimental results

The results of the proposed method on REVIEW are herein presented and discussed. The models presented in Chapter 3 are evaluated, both in terms of quality of the fitting and of the width measurements obtained through regression. Different validation schemes are used to assess the results. Results of the state-of-the-art algorithms are also reported and compared with our method.

4.2.1 Model fitting results

The goodness-of-fit is evaluated for the three tested models: DoG×line model with 7 parameters (DoG-L7), DoG×line model with 8 parameters (DoG-L8) and the Hermite model with 6 parameters. The results of the goodness-of-fit (gof) metrics for each dataset of REVIEW are in Table 4.1. For each metric and dataset, it shows the mean of the metrics for all the profiles in that dataset. In Fig. 4.3, examples of intensity profiles and their best-fit models are shown, along with the computed gof metrics. As it can be seen, the DoG-L8 model consistently returns better gof metrics than the other two models, followed by the DoG-L7. In fact, the 8 parameter model has always the lowest *SSE* and *RMSE* and the highest R^2 and R^2_{adj}, which suggests it is the model with smaller random error and the one that better explains the variation of the data. This is specially noticeable for the CLRIS dataset, where the DoG-L8 model shows the largest improvement relatively to DoG-L7. This is expected since the new introduced parameter allows to model the asymmetry in the CLR, which is frequent in the CLRIS images (see subsection 3.6.1). The Hermite model with 6 parameters is the worst fitting model, retrieving worse values for all the gof metrics. Again, this is most prominent for the CLRIS dataset. As shown in subsection 3.6.1, the fitting of this model (considering the referred parameter ranges) to the profiles of this dataset is relatively poor.



(a) Hermite model fitting. $SSE = 0,0586$; $R^2 = 0,9329$; $R^2_{adj} = 0,9320$; $RMSE = 0,0128$.

(b) DoG-L7 model fitting. $SSE = 0,01524$; $R^2 = 0,9825$; $R^2_{adj} = 0,9822$; $RMSE = 0,0065$.

(c) DoG-L8 model fitting. $SSE = 0,0030$; $R^2 = 0,9965$; $R^2_{adj} = 0,9964$; $RMSE = 0,0029$.

Figure 4.3: Example of fitting of the tested models to a vessel intensity profile, along with the goodness-of-fit metrics. Black dots are the profile points and the orange curve the best-fit model.

4.2.2 Regression tree results

The results of the method proposed in this work are evaluated in four different ways:

1) k-fold cross-validation, with k=10, inside each dataset (CLRIS, HRIS, KPIS and VDIS);

Table 4.1: Goodness-of-fit metrics obtained for the tested models: DoG-L7, DoG-L8 and Hermite models. The results shown for each metric and for each dataset are the mean of the values from all the profiles of that dataset. The best results for each dataset are highlighted.

Model	Dataset	*SSE*	R^2	R^2_{adj}	*RMSE*
DoG-L7	CLRIS	0,0072	0,9774	0,9769	0,0043
	HRIS	0,0075	0,9733	0,9721	0,0062
	KPIS	0,0055	0,9935	0,9932	0,0053
	VDIS	0,0072	0,9726	0,9718	0,0045
DoG-L8	CLRIS	**0,0053**	**0,9817**	**0,9813**	**0,0037**
	HRIS	**0,0071**	**0,9750**	**0,9736**	**0,0060**
	KPIS	**0,00500**	**0,9942**	**0,9940**	**0,0050**
	VDIS	**0,0064**	**0,9744**	**0,9735**	**0,0043**
Hermite	CLRIS	0,0412	0,9208	0,9197	0,0097
	HRIS	0,0090	0,9694	0,9683	0,0069
	KPIS	0,0110	0,9873	0,9870	0,0078
	VDIS	0,0114	0,9683	0,9675	0,0054

2) k-fold cross-validation, with k=10, in the whole REVIEW dataset;
3) leave-one-segment-out validation, inside each dataset;
4) leave-one-segment-out validation, in the whole REVIEW.

Table. 4.2 shows the results of the different tested models, for each of the datasets of REVIEW, using each of the 4 evaluation schemes referred. The metrics presented in the table are the success rate of the algorithm, the mean and standard deviation of the measurements and the mean and standard deviation of the measurement errors. For comparison, results from the observers and the ground truth values are also shown. One should note that the goal here is not to achieve zero standard deviation of the errors, but instead to be close to the observers' results. This means that for datasets in which the observer's measurements have larger variance, a larger variance is expected from the measurement algorithms. The results of our method will be further detailed and discussed in this subsection.

4.2.2.1 k-fold cross-validation results

In k-fold cross-validation, the original dataset is randomly partitioned in k subsets. At each time, one of the subsets is used for testing and the remaining $k-1$ subsets for training. This is repeated k times, so that each subset is used exactly one time for testing. This ensures that each profile enter exactly once for testing. In this work we use 10 folds, similarly to Lupascu *et al.* (2013).

In the work from Lupascu *et al.* (2013), the authors perform cross-validation inside each dataset. This means that for each of the 4 datasets in REVIEW, the cross-validation scheme is applied, independently of the other datasets. However, we also perform cross-validation in the whole REVIEW. This allows to assess the robustness of the regression method, evaluating if it is able to return good results even when dealing with a large variety of images, both in terms of size, resolution, contrast, presence of pathologies, etc.

Table 4.2: Results of the proposed method for retinal vessel width estimation, for the three tested models: DoG× line model with 7 parameters (DoG-L7), DoG× line model with 8 parameters (DoG-L8), and Hermite model with 6 parameters. For each of the methods, 4 evaluation schemes are presented (shown in different colors): cross-validation inside each dataset (Cv_d) and in the whole REVIEW (Cv_R) and leave-one-segment-out inside each dataset (Lso_d) and in the whole REVIEW (Lso_R). The results of the measurements from the three observers (O1, O2 and O3) who marked the REVIEW database are also shown, as well as the ground truth, i.e., mean of the three observations (G.T). Results are shown for each of the datasets of REVIEW (CLRIS, HRIS, KPIS and VDIS). In terms of evaluation metrics, SR stands for Sucess Rate, μ_{meas} and σ_{meas} are the mean and standard deviation of the width measurements, respectively, whereas μ_{error} and σ_{error} are the mean and standard deviation of the measurement errors. The * marks next to the SR values indicate they are negatively influenced by errors found in the observers' marks of that dataset.

Method		CLRIS					HRIS					KPIS					VDIS				
		SR (%)	μ_{meas} (px)	σ_{meas} (px)	μ_{error} (px)	σ_{error} (px)	SR (%)	μ_{meas} (px)	σ_{meas} (px)	μ_{error} (px)	σ_{error} (px)	SR (%)	μ_{meas} (px)	σ_{meas} (px)	μ_{error} (px)	σ_{error} (px)	SR (%)	μ_{meas} (px)	σ_{meas} (px)	μ_{error} (px)	σ_{error} (px)
O1		100	13.19	4.01	-0.61	0.566	100	4.12	1.25	-0.23	0.288	100	7.97	0.47	0.45	0.233	100	8.50	2.54	-0.35	0.543
O2		100	13.69	4.22	-0.11	0.698	100	4.35	1.35	0.002	0.256	100	7.60	0.42	0.08	0.213	100	8.91	2.69	0.06	0.621
O3		100	14.52	4.26	0.72	0.566	100	4.58	1.26	0.23	0.285	100	7.00	0.52	-0.53	0.234	100	9.15	2.67	0.30	0.669
G.T.		100	13.80	4.12	-	-	100	4.35	1.26	-	-	100	7.52	0.42	-	-	100	8.85	2.57	-	-
DoG-L7	Cv_d	97.5*	13.78	3.90	0.01	0.563	98.4*	4.30	1.21	0.002	0.217	98.8	7.40	0.26	-0.001	0.298	99.7*	8.78	2.45	0.007	0.690
	Cv_R		13.61	3.92	-0.16	0.685		4.36	1.22	0.06	0.302		7.34	0.29	-0.06	0.299		8.75	2.53	-0.02	0.721
	Lso_d		13.56	3.64	-0.20	1.236		4.30	1.10	-0.004	0.539		7.37	0.14	-0.03	0.384		8.69	2.25	-0.08	1.048
	Lso_R		13.12	3.51	-0.64	1.156		4.53	1.32	0.23	0.829		7.16	0.27	-0.24	0.365		8.68	2.44	-0.09	1.092
DoG-L8	Cv_d	97.5*	13.78	3.93	0.01	0.592	98.4*	4.31	1.19	0.002	0.248	98.8	7.41	0.23	0.005	0.304	99.7*	8.79	2.39	0.016	0.780
	Cv_R		13.43	3.90	-0.34	0.796		4.37	1.23	0.062	0.320		7.33	0.28	-0.08	0.327		8.78	2.49	0.008	0.840
	Lso_d		13.63	3.68	-0.14	1.297		4.28	1.06	-0.03	0.598		7.39	0.15	-0.001	0.434		8.62	2.10	-0.155	1.234
	Lso_R		12.64	3.29	-1.12	1.519		4.49	1.35	0.19	0.901		7.12	0.26	-0.29	0.408		8.66	2.36	-0.11	1.270
Hermite	Cv_d	97.5*	13.78	3.62	0.02	1.152	98.4*	4.30	1.21	0.002	0.221	98.8	7.40	0.26	0.002	0.283	99.7*	8.77	2.41	0.001	0.726
	Cv_R		13.52	3.23	-0.24	1.391		4.38	1.22	0.07	0.287		7.33	0.27	-0.07	0.292		8.72	2.51	-0.05	0.785
	Lso_d		13.70	3.24	-0.06	2.306		4.32	1.12	0.02	0.488		7.42	0.17	0.02	0.359		8.66	2.24	-0.11	1.099
	Lso_R		12.92	2.59	-0.85	2.752		4.60	1.29	0.29	0.718		7.18	0.22	-0.22	0.384		8.58	2.44	-0.19	1.139

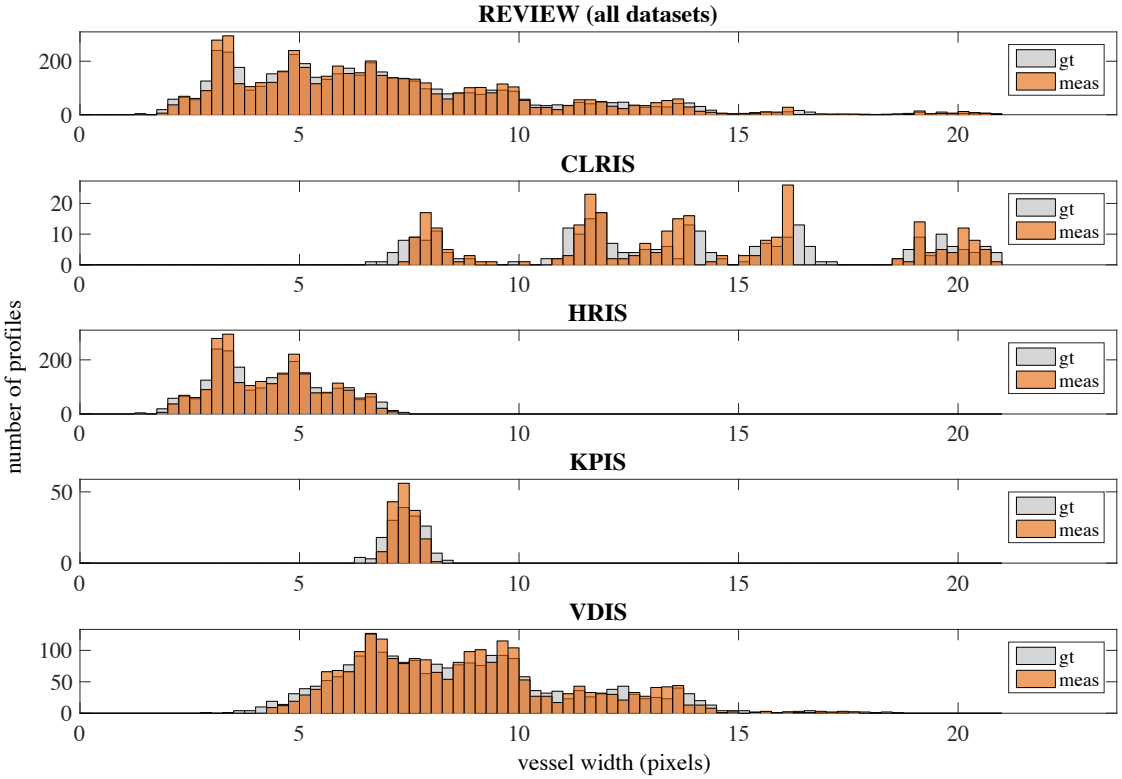
The results of the proposed method using the 3 tested models, evaluated through 10-fold cross-validation inside each dataset and in the whole REVIEW, are shown in Table. 4.2. Results show to be close to the observers' in terms of precision, as it is the goal. The standard deviation of the errors, σ_{error} is consistently higher for CLRIS and VDIS datasets comparing to the HRIS and KPIS, which is coherent with the observers' values. This is true across all the tested models, and for both cross-validation schemes. It is known that CLRIS is a difficult dataset due to the presence of accentuated CLR, and VDIS has a large variety of images, both normal and diseased, representing a greater challenge for diameter measurement.

Further, one notices that σ_{error} is generally slightly higher when performing cross-validation in the whole REVIEW than inside each dataset. This effect is expected, since, despite the increase on the training set size that occurs when all datasets are considered, the variability of vessels properties also increases, as referred above. For instance, the range of diameters when performing the cross-validation in the whole REVIEW is larger than when each dataset is considered separately.

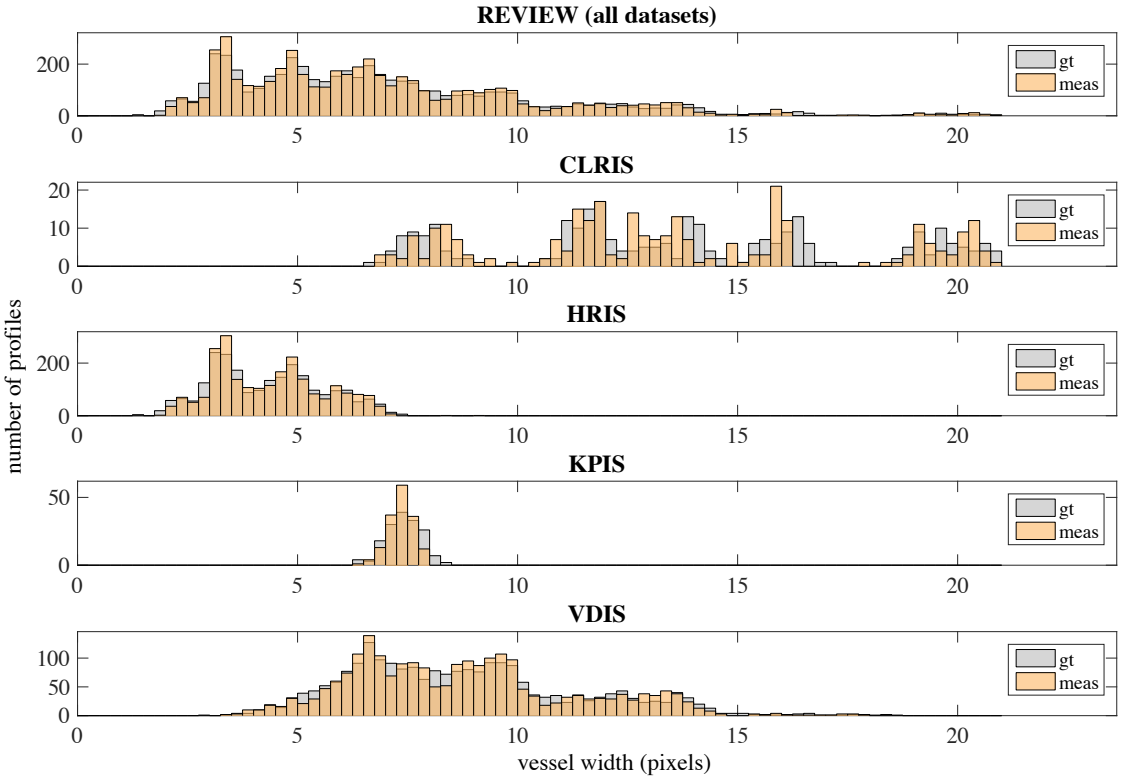
From the tested models, DoG-L7 shows better results. Although in the majority of the datasets the results of the three models are relatively close, and close to the observers', for the CLRIS dataset the Hermite model behaves considerably worse, doubling σ_{error} of the other models. As referred in subsection 4.2.2, the fitting of the CLRIS profiles by this model is relatively poor, which negatively affects the results. Although the DoG-L8 model fits slightly better the CLR vessels (subsection 4.2.2), this improvement seems not to add much relevant information for regression. As can be seen, σ_{error} is similar for both DoG$\times$line models in CLRIS, but for VDIS, for instance, it is higher for the 8 parameter model. This suggests that the addition of the parameter introduces extra information not relevant for the ensembles of bagged regression trees, which can even constitute noise that hinders the regression.

Regarding the success rates (SR), the major reason for the less than 100% SR of proposed algorithm is the fact that no association is found between the dubious ground truth points (Fig. 4.2) and the detected center points, due to the misplacement of the ground truth and the strict association criteria (subsection 4.1.2). In the case of KPIS, where no dubious ground truth marks were found, the profiles not measured correspond to junction regions. These junction points are removed in the preprocessing phase of the algorithm. It appears that some junctions were not avoided when marking the ground truth, probably due to the intersections with thin vessels that were not accounted for. However, as our segmentation detects even the thinnest vessels, these junctions are detected. Similar cases may occur in HRIS and VDIS datasets. Since our association criteria leads to a unique match between ground truth and detected center points, no association is performed at bifurcations and crossings. Despite this, considering the whole REVIEW, 99% of the ground truth vessel profiles are measured by the algorithm.

In Fig, 4.4 the distributions of the diameters measured using the modified DoG model with 7 parameters are shown, along with the ground truth diameter distributions. Both results from cross-validation inside each dataset and in the whole REVIEW are presented. It is clear that the measured and ground truth diameters follow a similar distribution, being that CLRIS is the dataset with the measurement distribution farther from the reference one. We see that the range



(a) Diameter distribution using cross-validation inside each dataset (CLRIS, HRIS, KPIS and VDIS).



(b) Diameter distribution using cross-validation inside the whole REVIEW.

Figure 4.4: Distribution of the ground truth and measured diameters, using the DoG-L7 model for fitting and 10-fold cross-validation, both inside each dataset and in the whole REVIEW.

of diameters present in the CLRIS and VDIS datasets are significantly larger than the ones from HRIS and KPIS, being that KPIS has the narrowest diameter range. Further, CLRIS is practically the only dataset that contains diameters largest than 15 pixels. HRIS also contains vessels with very small diameters that poorly represented in other datasets. This helps corroborating the above mentioned results, since the CLRIS and VDIS higher variance errors obtained through in-dataset cross-validation may also be due to their broader diameter ranges. In Lupascu *et al.* (2013), the authors suggest that two criteria should be met when constructing the datasets for the ensembles of bagged regression trees: 1) there should be an approximately uniform diameter distribution in the training set and 2) the range of widths in the training and testing sets should not be very wide.

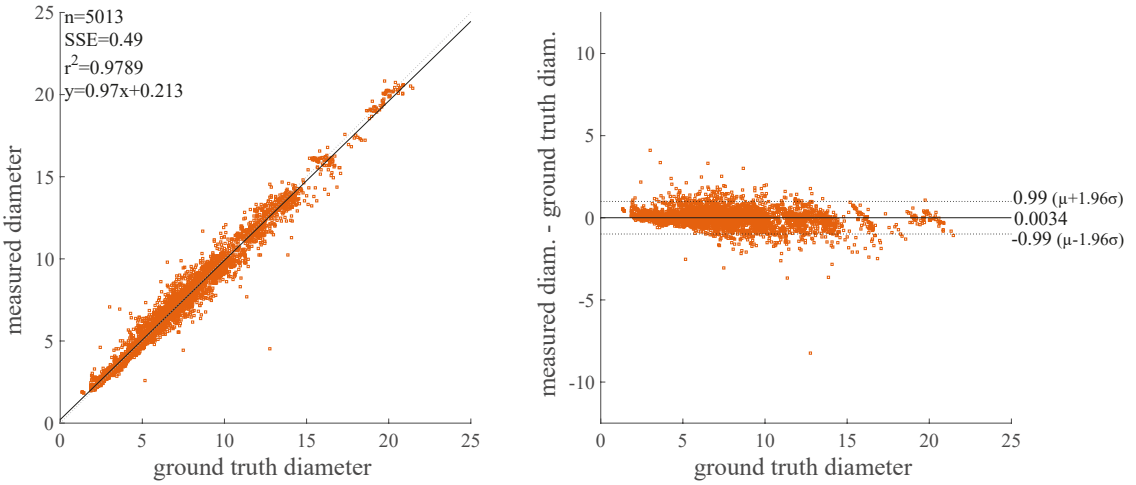
Fig. 4.5 shows the correlation and the Bland-Altman plots of the ground truth and the measured diameters from REVIEW, using the DoG-L7 model and cross-validation inside each dataset and in the whole REVIEW. In the correlation plots, each measured profile is represented by a point with coordinates (x_p,y_p), where x_p is the ground truth diameter and y_p the measured diameter. In the BA plot, each point has coordinates (x_p,$y_p - x_p$). By analysing the plots, one sees that there is little dispersion of the points, indicating a small variance of the errors. As the points are close to the $y = x$ line, in the correlation plot, and to the $y = 0$, in the BA plot, a low measurement error is verified. Consequently, the measurements are both precise and accurate. The standard deviation of the errors for REVIEW is 0.51 pixels when the cross-validation is performed in each dataset, and 0.56 pixels when it is performed in the whole REVIEW, being the mean error close to zero. Further, the errors do not seem to depend on the range of diameters, since the points appear to be distributed in a similar manner regardless of the true diameter.

4.2.2.2 Leave-one-segment-out validation results

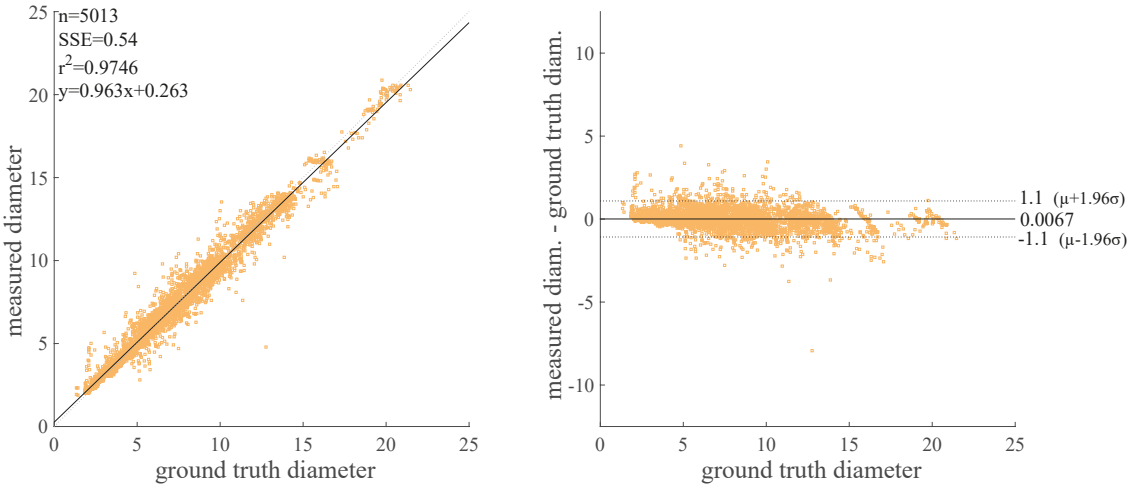
The algorithm is also evaluated performing leave-one-segment-out validation. This consists in leaving at each time one vessel segment for testing, and training in the remaining segments. The procedure is repeated n times, where n is the number of segments in the dataset. We evaluate our method inside each dataset, as well as in the whole REVIEW. This way of evaluating the results ensures that similar neighboring profiles are not considered both in training and testing, which is not secured by cross-validation scheme. It is known that neighboring profiles can be very similar. Further, since 2D model fitting is performed, i.e., 11 neighboring profiles are considered, two consecutive 2D profiles have 1D profiles in common.

The results of the proposed method using the 3 tested models and this evaluation scheme are shown in table. 4.2. In general, the errors follow a tendency similar to the described for cross-validation, being verified higher σ_{error} for the CLRIS and VDIS datasets, for the three models and for the two validations (in-dataset and in the whole REVIEW). Similarly to what happened for cross-validation, DoG-L7 shows better or similar results to the ones from the other models, being the biggest improvement verifier for CLRIS and when comparing with the Hermite model.

We see there is an increase of σ_{error} when using this validation instead of cross-validation, for all the datasets, which is expected. Removing an entire segment from the training set may significantly reduce or eliminate the presence of similar profiles to the tested ones. This effect is



(a) Bland-Altman plots using cross-validation inside each dataset (CLRIS, HRIS, KPIS and VDIS).

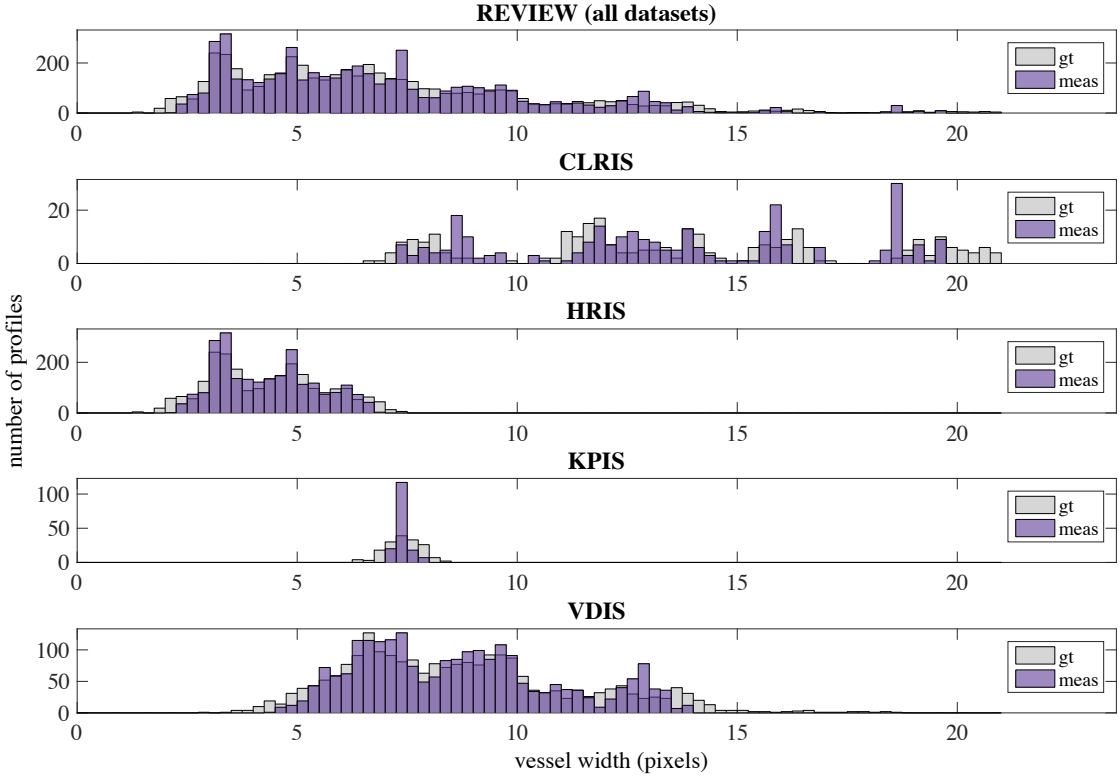


(b) Bland-Altman plots using cross-validation inside the whole REVIEW.

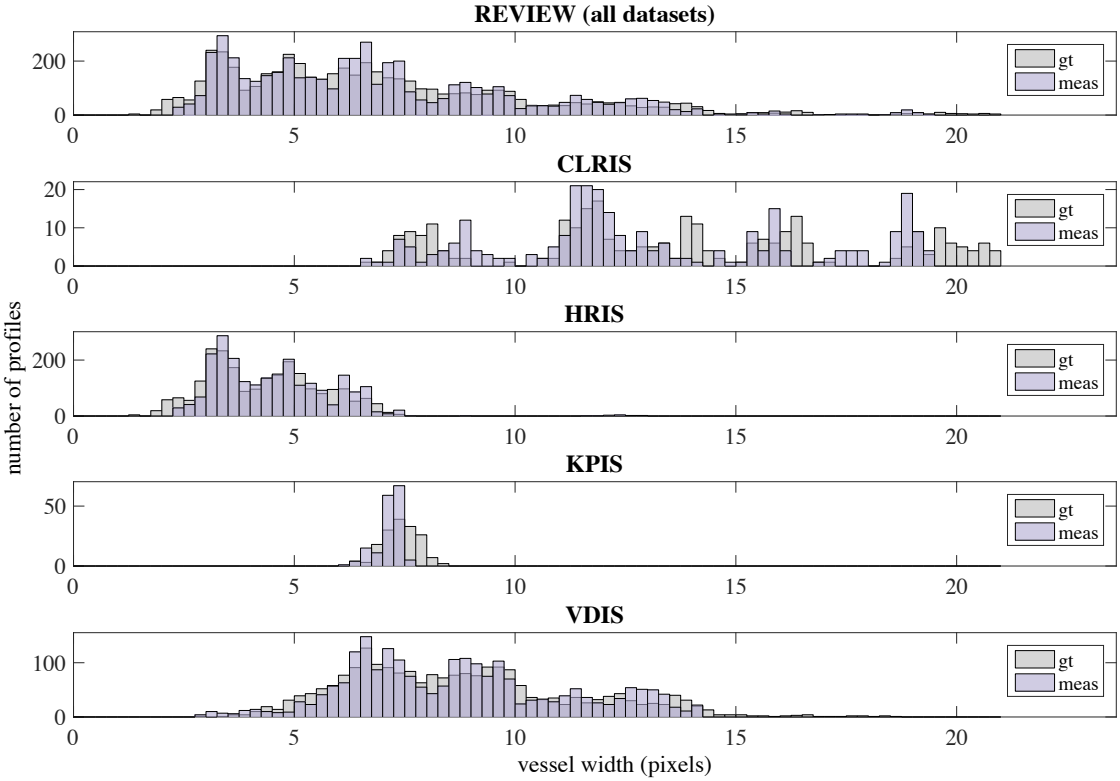
Figure 4.5: Correlation and Bland-Altman plots of the ground truth and measured diameters, using the DoG-L7 model for fitting and 10-fold cross-validation, inside each dataset and in the whole REVIEW (in the *xx* axis of the Bland-Altman plot we show the ground truth diameters instead of the mean between the ground truth and measured diameters).

most prominent for the CLRIS dataset. As known, CLRIS contains segments with strong CLR, and has a very wide range of diameters. Further, it contains the majority of the vessels with diameters >20 pixels (see Fig. 4.6). Additionally, it only contains 20 segments. Since the diameter range is very wide, few segments are available for each diameter. Consequently, the removal of one segment from the training set can largely affect the ensembles of bagged regression trees since it is probably a representative segment. In Fig. 4.6 one sees that the measured diameter distribution does not have any profile with >20 pixels.

Further, σ_{error} does not vary in the same manner for the different datasets when comparing both leave-one-segment-out validations. For instance, for HRIS, σ_{error} is higher (for all three



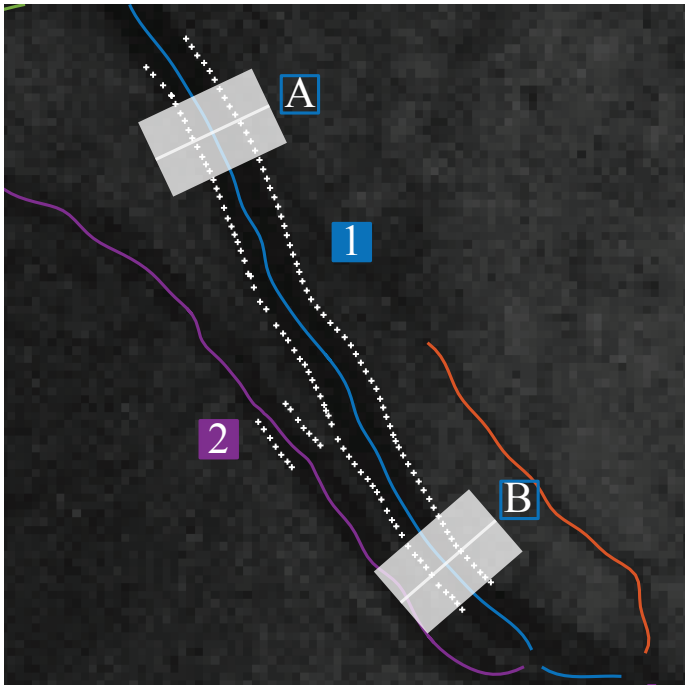
(a) Diameter distribution using leave-one-segment-out validation inside each dataset.



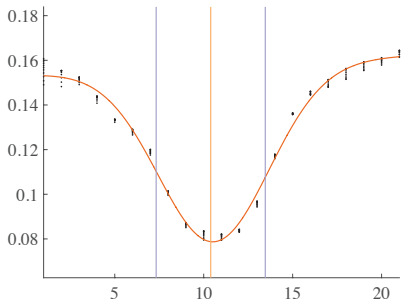
(b) Diameter distribution using leave-one-segment-out validation inside the whole REVIEW.

Figure 4.6: Distribution of the ground truth and measured diameters, using the DoG-L7 model for fitting and leave-one-segment-out validation, both inside each dataset and in the whole REVIEW.

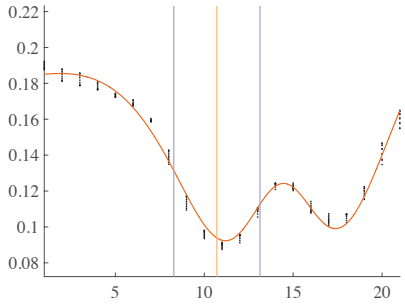
models) when performing the validation in the whole REVIEW, whereas for the other datasets the differences are negligible. For the HRIS dataset, the diameter range is relatively narrow, and there are 90 segments in total (Fig. 4.6). The fact that σ_{error} increases significantly when validating in the whole REVIEW is considered to be due to the introduction of noise by other datasets. Fig. 4.7 shows a vessel profile from HRIS, where the measured diameter was, for certain profiles, the double of the real one, due to the similarity of the vessel profile with a profile of a vessel with CLR, caused by the inclusion of a near vessel in the profile. This identification of CLR only happened because of the presence of CLRIS in the training.



(a) Segment from HRIS003 image (REVIEW).



(b) Intensity profile from region A.



(c) Intensity profile from region B.

Figure 4.7: Example of a segment with poor width measurements due to a false CLR detection. Image (a) shows an HRIS vessel segment (labelled as 1), that at a given point runs next to another segment (labeled as 2); (b) shows the smoothed profiles extracted from region A; (c) shows the smoothed profiles extracted from region B, where the presence of another vessel close to the main vessel simulates the presence of CLR. Profiles as the one shown in (c) were wrongly measured by our algorithm, when leave-one-segment-out validation is performed in the whole REVIEW, being retrieved a diameter that is approximately two times the real diameter. (Black points in (b) and (c) represent the intensity profiles, the orange curves are the best-fit models, the yellow vertical lines are the centers of the profiles and the purple vertical lines represent the ground truth locations. White marks in (a) represent the ground truth points.)

Nevertheless, one should note that, even when a segment is left out of the training procedure, the obtained results are still very satisfactory, never surpassing approximately 1 pixel of σ_{error} (except for CLIRS, where σ_{error} is of 1.236 pixels). Finally, μ_{error} is considerably higher when for the leave-one-segment-out validation in the whole REVIEW than for the other three validation schemes. Despite this, they are still close to the values of the observers.

An additional analysis is performed in order to assess the similarity between the segments measured in each dataset, aiming to corroborate the results of the leave-one-segment-out validation inside each dataset. If the similarity between segments in a dataset is low, it is expected that the removal of a segment from the training set significantly affects the results. Obviously, this depends also on the number of segments in the dataset. To assess the segment similarity, for each dataset, the correlation between each pair of segments is computed. For that, the profile for a given segment is taken as the mean of the profiles from that segment. For each segment pair, the two profiles are aligned by their maximum value. Results are normalized by dividing by the maximum of the autocorrelations of the two profiles. Results are shown in Table 4.3.

Table 4.3: Correlation between the segments of each dataset. μ_{corr} and σ_{corrr} are the mean and standard deviation of the correlations of all pairs of segments, max_{corr} is the maximum correlation and $\#comb$ is the number of combinations of 2 segments found in the dataset.

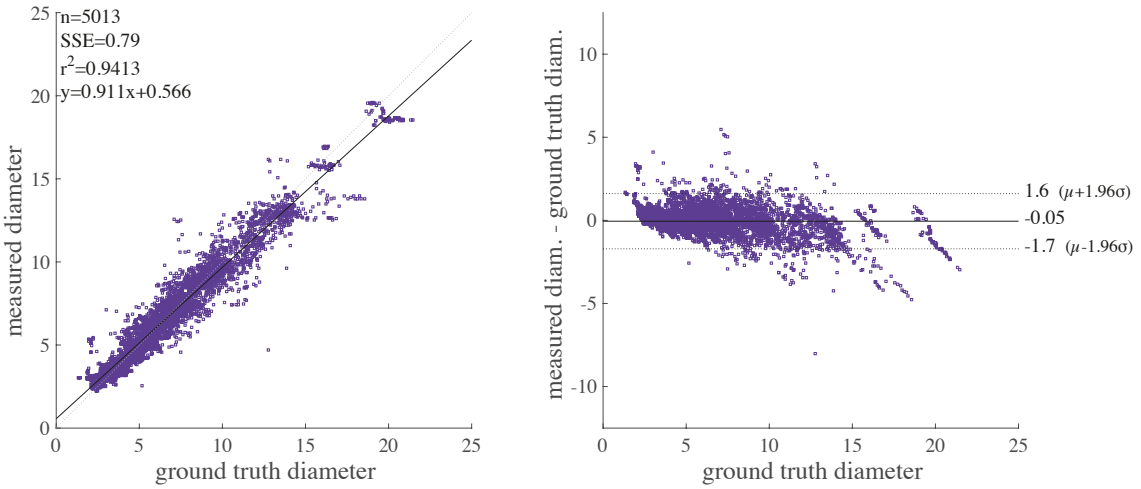
Dataset	μ_{corr}	σ_{corr}	max_{corr}	$\#comb$
CLRIS	0.510	0.252	0.982	190
HRIS	0.612	0.217	0.996	3160
KPIS	0.804	0.103	0.960	15
VDIS	0.454	0.258	0.997	6670

HRIS is the dataset that shows higher correlation between its segments (higher μ_{corr} and smaller σ_{corr}), which is coherent with the the evolution of σ_{error} when a segment is left out of the training, which is not very significant. CLRIS dataset shows one of the lowest correlations, wich corroborates the big effect of the removal of a segment from the training set. Although VDIS has lower correlation values, the largest number of segments (79 vs 20) justifies the smaller increase in σ_{error} when comparing to CLRIS.

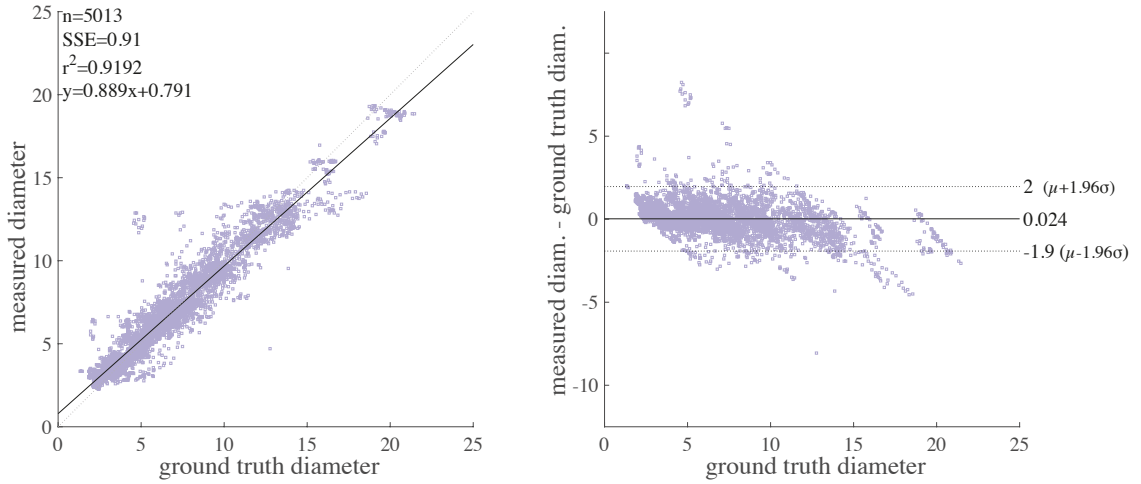
Fig. 4.8 shows the correlation and the Bland-Altman plots of the ground truth and the measured diameters from REVIEW, using the DoG-L7 model and leave-one-segment-out inside each dataset and in the whole REVIEW. Although there is little dispersion of the points, indicating a small variance of the errors, the dispersion is larger than that of cross-validation results. The standard deviation of the errors for REVIEW is 0.84 pixels when the cross-validation is performed in each dataset, and 0.99 pixels when it is performed in the whole REVIEW, being the mean error close to zero. Similarly to the cross-validation results, the errors do not seem to depend on the range of diameters, although a slight tendency to underestimate the widths for larger vessels can be detected.

4.3 Comparison with the state-of-the-art results

Together with our DoG-L7-based method, Fig. 4.9a depicts the performance of some of the methods presented in Chapter 2 on the REVIEW dataset. The computed metrics are, in most of the



(a) Bland-Altman plots using leave-one-segment-out validation inside each dataset (CLRIS, HRIS, KPIS and VDIS).



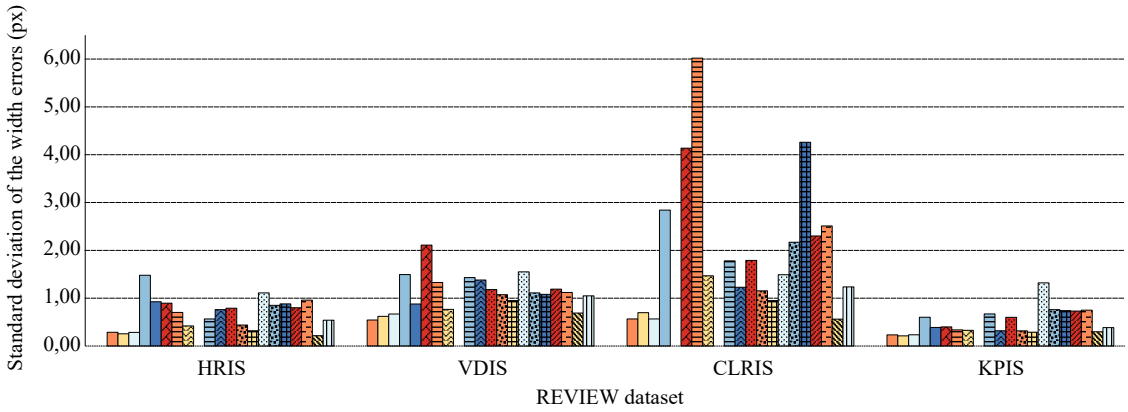
(b) Bland-Altman plots using leave-one-segment-out validation inside the whole REVIEW.

Figure 4.8: Correlation and Bland-Altman plots of the ground truth and measured diameters, using the DoG-L7 model for fitting and leave-one-segment-out validation, inside each dataset and in the whole REVIEW (in the *xx* axis of the Bland-Altman plot we show the ground truth diameters instead of the mean between the ground truth and measured diameters).

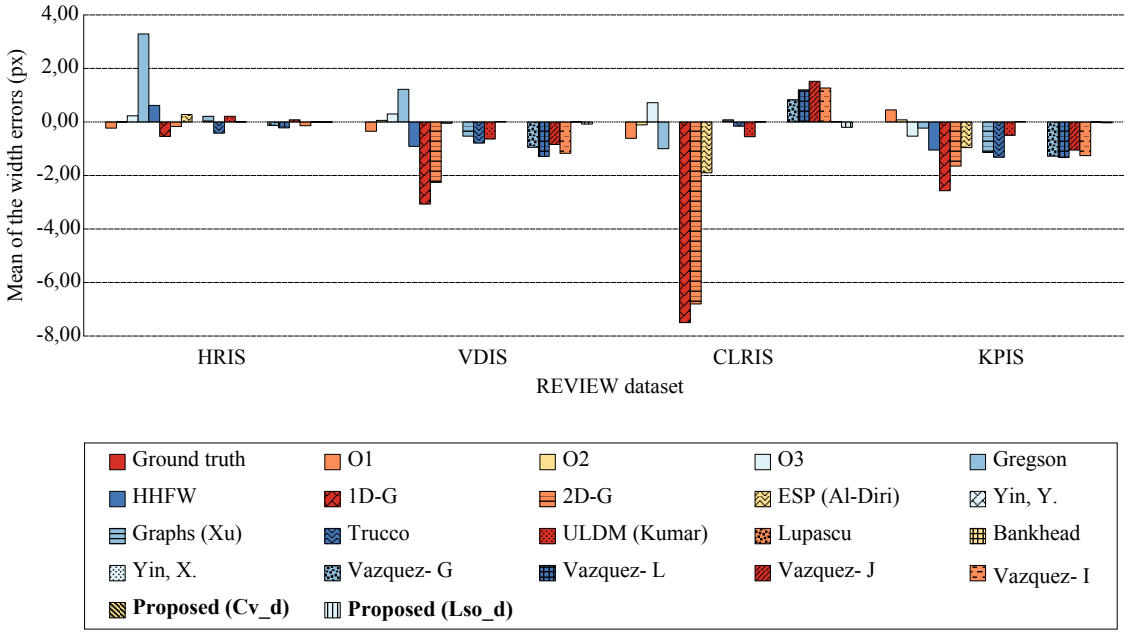
cases, the referred in subsection 4.1.2. Note that the evaluation of the methods that are not compared in the chart is not available in the literature. In Appendix A the results of Fig. 4.9a are presented in a table-format, along with additional results of the proposed algorithm.

4.3.1 Model fitting-based methods

Earliest methods show, in general, poor results in comparison with the most recent ones. The simple HHFW algorithm proposed by Brinchmann-Hansen and Heier (2009), although retrieving relatively good precision, i.e., low standard deviation of the width differences, for some of the



(a) Standard deviation of the differences between the measured diameters and the corresponding ground truth diameters.



(b) Mean of the differences between the measured diameters and the corresponding ground truth diameters.

Figure 4.9: Performance of vessel width measurement methods on the REVIEW dataset, for each of the four datasets: CLRIS, HRIS, KPIS and VDIS. The results of the three observers (O1, O2 and O3) and the ground truth (mean of the three observations) are also shown. The methods compared are that of Gregson *et al.* (1995), Brinchmann-Hansen and Heier (2009) (HHFW), Zhou *et al.* (1994) (1D-G), Lowell *et al.* (2004) (2D-G), Al-Diri *et al.* (2009) (ESP), Yin *et al.* (2013), Xu *et al.* (2011a) (Graphs), Trucco *et al.* (2013), Kumar *et al.* (2012) (ULDM), Lupascu *et al.* (2013), Bankhead *et al.* (2012), Yin *et al.* (2014) and Vazquez *et al.* (2013), the last evaluated in four types of images: green channel of the RGB color model (G), lightness from the Lab color space (L), and two different gray transformations: light gray scale image (I) and darker image (J) and, finally, the proposed method using the DoG-L7 model and evaluated both with cross-validation and leave-one-segment-out inside each dataset, denoted by Cv_d and Lso_d, respectively. "x" marks indicate that no values are available.

datasets, sometimes shows a poor SR (for CLRIS dataset the SR is 0%). Results from the method of Gregson *et al.* (1995) are significantly worse than that of HHFW when it comes to precision but has a SR of 100%. The 1D-Gaussian fitting-based method proposed by Zhou *et al.* (1994) does not improve (and sometimes worsens) the results of the two latter methods, although presenting good SR. Particularly in the CLRIS dataset, the precision and accuracy are poor (reaching 4 and -7.5 pixels, respectively). The 2D-Gaussian and DoG fitting of Lowell *et al.* (2004) shows in general better results, except for CLRIS, where the standard deviation of the errors is the highest from all the methods (6 pixels). This is not expected since the method considered the central light reflex present in the vessels, by using a DoG model.

Although not present in Fig. 4.9a, Aliahmad *et al.* (2012) evaluated the method on CLRIS and VDIS, but only on 580 randomly selected vessel cross-section and considering the first observer measurements as the ground truth instead of the mean of the 3 observers. The proposed algorithm shows the best results, in terms of precision and accuracy, when compared with the methods of Chapman (2001) (sliding linear regression filter) and Gao *et al.* (2001) (1D-twin Gaussian model fitting). The latter shows the worst performance. One advantage is the use of multiple scaling factors to relate the model σ parameter and the vessel diameter, being a scaling factor computed for each edge. This takes in account the possibility of different blurriness levels and illumination at the edges. Contrarily, previous methods consider a single fixed scaling factor for all images.

The method of Lupascu *et al.* (2013) shows a mean of the width differences very close to the observers in all datasets, being close to 0 pixel (very good accuracy). The method shows considerably higher accuracies and precisions than the majority of the methods, even comparing to the method of Al-Diri *et al.* (2009), one of the leading contour-based algorithms. For the CLRIS dataset this is specially noticeable, being the standard deviation of the errors considerably lower than most of the methods. The correlation between the measurements and the ground truth was also analysed, and showed that the measurements were slightly smaller for thin vessel and larger for large vessels. By analysing the performance of the method on a partitioned VDIS dataset (divided in subsets with widths with low inter-observer variability), the authors concluded that the method behaved worse in the subsets with wider diameter range and with a non uniform diameter distribution, giving varying results depending on the training set characteristics. However, if criteria are met to avoid the referred situations, results should be consistent.

4.3.2 Contour-based methods

The ESP algorithm of Al-Diri *et al.* (2009) shows very good performance, returning results close to the observers for all the datasets except for the CLRIS, due to the blurred edges and the vessel CLR. The algorithm shows an high level performance on fine vessels (from 3 to 4 pixels). However, it may not detect the vessel width at a given point, i.e., it fails to converge (SR $\neq$ 100%). It may fail to identify overlapping or closely parallel vessels and in cases with a steep background intensity gradient, obscuring one edge of the vessel. These limitations can be due to the need of ESP to clearly identify two parallel edges.

The graph-based approach of Xu *et al.* (2011a) is robust, retrieving small σ_{error} on HRIS and KPIS. For the remaining datasets, the standard deviation of the errors are more distant from the observers. The correlation between the measurements and the ground truth in HRIS show that the method provides smaller measurements for fine vessels and larger measurements for large vessels. An advantage of the method is that the simultaneous two-boundary detection allows to detect both boundaries even if one is noisy or has low contrast. The strong dependence of the results on the initial segmentation is a limitation, as well as the presence of false positives or disconnected vessels, sometimes leading to incorrect measurements.

The ULDM method of Kumar *et al.* (2012) shows a high success rate and low deviation from the ground truth measurements. The authors also computed the cross-correlation between the vessel widths predicted using ULDM, the manual measurements from each observer and the ground truth, for each of the 4 datasets. The correlation between ULDM and the ground truth is similar to that of the 3 manual measurements for all datasets, except for KPIS (slightly lower).

The method of Bankhead *et al.* (2012) presents very good precision, with one of the lowest (most of the times the lowest) standard deviation of the errors. The most distinct improvement is in the CLRIS dataset, which probably arises from the algorithm being able to deal with the CLR. For VDIS, the results do not show much improvement, caused by the noisy nature of the dataset. However, even in the worst case, the standard deviation of the errors is ≤ 1 pixel. The major limitation of the method is that if the image contrast decreases the zero crossings may not be found for all locations along the vessel.

The algorithm presented by Vazquez *et al.* (2013) was tested on images with four different color-spaces (see subsection 2.2.1). In all datasets except CLRIS, σ_{error} is small, being ≤ 1.2 pixels. SR are in general lower than that of the other methods. Generally, the measurements are little dependent on the used color space. Limitations of the method are the treatment of images with strong central light reflex, which can lead to the detection of two different vessels.

The method of Trucco *et al.* (2013) shows a performance comparable to that of sophisticated algorithms, but has lower SR. For KPIS and CLRIS it is one of the best performing methods (smaller σ_{error}). As was the goal of this work, the method improved the accuracy of width estimation taken directly from unprocessed binary vessel masks.

The algorithm of Yin *et al.* (2013) shows good performance, specially in the CLRIS dataset and has a SR of 100%. However, the only metric provided is the mean of the measurements. The measured and reference widths on the REVIEW dataset follow a similar distribution.

Yin *et al.* (2014) present consistent results for all images and a SR of 100%. The precision is relatively good, but always $\geq$ 1pixel, being higher than that of the most successful methods, specially for HRIS and KPIS. However, only 4 images of the REVIEW (HRIS001, CLRIS002, VDIS006 and KPIS001) were used for evaluation, being the results not as representative as the other methods.

4.3.3 Comparison with the proposed method

The performance of our method is close and even outperforms the best-performing state-of-the-art algorithms. The most robust state-of-the-art algorithms are the approaches from Al-Diri *et al.* (2009), Bankhead *et al.* (2012) and Lupascu *et al.* (2013), being those that return consistently low standard deviation of the errors for all 4 datasets. Although having common concepts with the last two methods, our method shows, in general, better performance. The preprocessing steps are inspired in the work of Bankhead *et al.* (2012). The diameter measurement is conceptually similar to the reported in Lupascu *et al.* (2013). Our method shows lower σ_{error} than all these methods when it is evaluated using cross-validation, both inside each dataset and in the whole REVIEW. Even considering other validation schemes, our method is still among the top-performing approaches. Considering a score defined as the mean of the standard deviation of the errors for all datasets, our method presents scores of 0.442 pixels and 0.502 pixels when using cross-validation in each dataset and in the whole REVIEW, respectively. The work of Bankhead *et al.* (2012) has the third best score (0.628 pixels). When leave-one-segment-out is performed, our method is still among the three best scored-works (see Appendix A).

Regarding the method of Lupascu *et al.* (2013), which is the closest to out proposal, the authors evaluated results through cross-validation inside each dataset. Using the same evaluation scheme, our method achieves almost an half of the σ_{error} for the majority of the datasets. The improvement of the results can be attributable to the use of a model that fits best the vessel profiles, specially those with CLR, and to the preparation of the profiles before model fitting using several preprocessing steps that improved the subsequent steps. This is true since the results of our method (see Table 4.2) but using the Hermite model with 6 parameters of Lupascu *et al.* (2013) are still superior to those of that work for the majority of the datasets.

4.4 Concluding remarks

Three models are tested in this work: DoG-L7, DoG-L8 and 6 parameter-Hermite model. DoG-L8 is the model that returns better goodness-of-fit metrics, which is specially noticeable for CLRIS. This is expected since the new parameter allows to model the asymmetry in the CLR. The Hermite model with 6 parameters is the worst fitting model.

Four validation schemes are used to evaluate the results of our method in terms of width measurement: 10-fold cross-validation inside each dataset and in the whole REVIEW, and leave-one-segment-out validation inside each dataset and in the whole REVIEW. Results from cross-validation are close to the observers' in terms of precision, as it is the goal. The standard deviation of the errors is consistently higher for the CLRIS and VDIS datasets comparing to the HRIS and KPIS, which is expected since the first two are challenging datasets. Further, errors are generally slightly higher when validating in the whole REVIEW than inside each dataset. This is due to the introduction of a large variety of vessel properties when the whole REVIEW is considered. The DoG-L7 model retrieves the best results, specially for CLRIS. For this dataset, the Hermite model

returns the double of the standard deviation of the errors when compared to the first model. Although DoG-L8 fits better the CLR vessels, this seems not to be much relevant for regression. The correlation and the Bland-Altman plots of the ground truth and the measured diameters suggest that the errors of the method do not depend on the range of the diameters. In general, the errors evaluated using the leave-one-out scheme vary similarly to the cross-validation ones. Nevertheless, there is an increase of σ_{error} when using this validation scheme, specially for CLRIS, since removing an entire segment from the training set may significantly reduce and even eliminate representative profiles.

From the analysis of the state-of-the-art results, the majority of the methods tend to underestimate the widths (mean of the measurement errors ≤ 0). The best results, considering both accuracy and precision, usually occur for the HRIS and KPIS datasets. Some of the methods, as the earlier methods and that of Vazquez *et al.* (2013), see their performance reduced in CLRIS dataset, namely in terms of precision and success rate. Generally, the most recent methods (from 2009 onwards) have more promising results, showing higher precision and accuracy. Some of the methods distinguish themselves from the others by showing good consistence among all the datasets, low errors and high success rates. Results of the model fitting-based method of Lupascu *et al.* (2013) show almost always the best standard deviation of the measurement errors (and a close to zero mean of the measurement errors). In turn, the algorithm proposed by Bankhead *et al.* (2012), which is a contour-based method, presents high SR (always between 99% and 100%) and very low standard deviation of the errors, even lower than that of Lupascu *et al.* (2013). For these reasons, these works seem to be the most promising. However, methods such as the ESP from Al-Diri *et al.* (2009) also present very good results and should not be dismissed.

The performance of the method proposed in this work is similar or better than the best-performing state-of-the-art algorithms. The most robust state-of-the-art algorithms are the ones from Al-Diri *et al.* (2009), Bankhead *et al.* (2012) and Lupascu *et al.* (2013). Our method shows the lowest σ_{error} when cross-validation is performed, both inside each dataset and in the whole REVIEW. Even with other validation schemes, our results are among the best found in the literature. Our method almost halves the σ_{error} reported by Lupascu *et al.* (2013). The model introduced in this work uses a model that fits better to the vessel profiles, specially the ones with CLR. Further, it uses several preprocessing steps, inspired in Bankhead *et al.* (2012), that improve the vessel profiles before the model fitting is performed.

Chapter 5

Conclusions and Future work

Retinal blood vessels are very useful for non-invasive diagnosis of diseases that affect the systemic vasculature, since they are the only part of the vasculature that is directly observable. These vessels can be assessed using eye fundus color images. Retinal vessel width changes are related with several major diseases, such as diabetes, hypertension and atherosclerosis. Automatic methods are needed for objective and precise determination of vessel widths, since the manual measurement would be tiresome and prone to errors. The development of these methods is demanding, due to the great variability in image resolution and appearance of the vessels. Although automatic methods with good performance have been proposed, most of them have limitations, such as poor performance in lower resolution images, on thinner vessels or susceptibility to artifacts and pathologies.

The method herein presented combines model fitting with several preprocessing steps, and estimates the widths based on the best-fit-model parameters using ensembles of bagged regression trees with random feature selection. It uses a novel parametric model based on a Difference-of-Gaussians (DoG) model, modified through a multiplying line with varying inclination which is able to describe profile asymmetry. Additionally, it applies some methodologies adapted from other top-performing state-of-the-art algorithms. The algorithm starts by segmenting the vessels from the eye fundus image. Afterwards, vessel centerlines are detected through skeletonization of the vasculature. Then, individual segments are extracted through junction removal. Segments are then smoothed by spline approximation and the cross-sectional intensity profiles at the segment points are extracted. The length of the profiles is then determined and used for profile cut and smoothing. Parametric model fitting is applied to the intensity profiles. Based on best-fit model parameters the vessel width is determined through regression.

The algorithm is evaluated on REVIEW, the only publicly available dataset that contains width measurement annotations. This dataset is composed of 4 subdatasets with very different characteristics (resolution, presence of pathologies, vessel appearance, etc.). The tested models are two modified DoG models, newly introducted in this work, DoG-L7 and DoG-L8, that deal with the asymmetry of the profiles, and an Hermite model with 6 parameters. To evaluate the results, 10-fold cross-validation and leave-one-segment-out validation are performed, both inside each dataset

and in the whole REVIEW. Cross-validation results are close in precision to that of the observers, being the errors consistently higher for the most challenging datasets (CLRIS and VDIS) than for the easier ones. The performance is slightly worse when validating in the whole REVIEW, which can be due to the presence of a greater variety of vessel properties in comparison to each individual dataset. The best results are obtained with the DoG-L7 model, specially for CLRIS, for which it shows twice the performance of the Hermite model. Although the 8 parameter model provides a better profile fitting, specially for the CLRIS dataset, it does not improve the regression ensembles performance. Further, the performance of the method does not seem to depend on the true diameters. Generally, leave-one-segment-out validation results vary similarly to the cross-validation ones. Nevertheless, the performance is worse comparing to cross-validation, specially for CLRIS, since removing a segment from the training set eliminates representative profiles.

The method presented in this M.Sc. dissertation often shows better results than the top-performing state-of-the art algorithms. It has consistently the higher precision (lowest standard deviation of the errors) when cross-validation is performed. When a segment is left out, our results are still among the best found in the literature. Our method practically halves the σ_{error} reported by Lupascu *et al.* (2013). The novel DoG-L7 and DoG-L8 models fit best the vessel profiles, specially the most challenging ones, such as the ones with CLR. Results are further improved due to the use of several preprocessing steps before model fitting.

Despite all this, there is still room for improvement and adaptations. The algorithm was designed independently of any framework. However, one of the thought applications for the method is the integration in a system for AVR computation, for instance (Dashtbozorg (2015)). Depending on the type of application, it could be of interest to adapt the width measurement algorithm to be autonomous from the vessel centerlines. In a framework where segmentation is not performed, a method that does not depend on the centerlines may be preferable. With this adaptation the vessel width could be measured at given point (placed inside the vessel). The algorithm would detect the vessel center point that is closer to that point, and measure the width at that specific center point. However, additional steps would be required to determine the vessel direction at the target point and detect its centerline neighbors so that the stack of 1D neighboring profiles could be constructed.

Further, considering the nature of the method used for width estimation, which relies on supervised learning, the definition of the training dataset is a key factor. However, the REVIEW dataset, as well as its subdatasets, does not contain an uniform distribution of the diameters. This is not desirable and can influence the ensembles' performance, being that a more balanced dataset would be of much more interest. For that, observers should mark the majority of the vessels in the image instead of specific segments, and the dataset should contain more images. The increase in the quality of the training set is expected to improve the results and, perhaps more importantly, increase their representativity and reliability.

Moreover, despite the method's good performance, additional efforts in optimizing certain steps of the algorithm would probably improve the results. For instance, the determination of the profile lengths prior to model fitting could be further refined, since it tends to overestimate the

results. Although not critical, a better initial width estimation could improve the model fitting. This could start by a better adjustment of the parameters of the smoothing filter applied before peak detection, so that subtle peaks on the vessel edges are not missed. Again, although the parameter choice is not very determinant to the ensembles of trees' performance, the influence of the parameters, such as the number of trees and the number of predictors selected at random at each split, could be further assessed. Additionally, other regression methods could be tested for results comparison, such as Support Vector Machine or Neural Networks for regression.

Our retinal vessel width measurement method has a performance that is close or outperforms the top-performing state-of-the-art methods. The method shows to retrieve precise results, close to that of the observers, as was the goal. This shows the robustness of our method and its great potential to be used directly for measurement of retinal vessel widths and/or to be integrated in a framework for retinal vascular assessment.

Appendix A

Results of the algorithms

Table A.1: Standard deviation of the width errors, i.e., point-by-point differences between the ground truth and the width measurements (pixels). Cv_d, Cv_R, Lso_d and Lso_R stand for cross-validation in the dataset, cross-validation in the whole REVIEW, leave-one-segment-out validation in the dataset and leave-one-segment-out in the whole REVIEW, respectively. The score corresponds to the mean of the values of all datasets. The three best scores at highlighted.

Method	HRIS	VDIS	CLRIS	KPIS	Score
O1	0.288	0.543	0.567	0.233	0.408
O2	0.256	0.621	0.698	0.213	0.447
O3	0.285	0.669	0.566	0.234	0.439
Gregson	1.479	1.494	2.841	0.602	1.604
HHFW	0.926	0.879	-	0.389	-
Zhou (1D-G)	0.896	2.110	4.137	0.399	1.886
Lowell (2D-G)	0.703	1.328	6.019	0.337	2.097
Al-Diri (ESP)	0.420	0.766	1.469	0.328	0.746
Yin, Y.	-	-	-	-	-
Xu (Graphs)	0.567	1.43	1.78	0.67	1.112
Trucco	0.760	1.381	1.229	0.319	0.922
Kumar (ULDM)	0.79	1.18	1.79	0.60	1.090
Lupascu	0.438	1.073	1.154	0.318	0.746
Bankhead	0.32	0.95	0.95	0.29	**0.628**
Yin, X.	1.11	1.55	1.49	1.32	1.368
Vazquez- G	0.85	1.11	2.17	0.76	1.223
Vazquez- L	0.88	1.08	4.26	0.74	1.740
Vazquez- J	0.80	1.19	2.30	0.73	1.255
Vazquez- I	0.96	1.12	2.51	0.75	1.335
Proposed, DoG-L7 (Cv_d)	0.217	0.690	0.563	0.298	**0.442**
Proposed, DoG-L7 (Cv_R)	0.302	0.721	0.685	0.299	**0.502**
Proposed, DoG-L7 (Lso_d)	0.539	1.048	1.236	0.384	0.802
Proposed, DoG-L7 (Lso_R)	0.829	1.092	1.156	0.365	0.860

Table A.2: Mean of the width errors, i.e., point-by-point differences between the ground truth and the obtained width measurements (pixels). Cv_d, Cv_R, Lso_d and Lso_R stand for cross-validation in the dataset, cross-validation in the whole REVIEW, leave-one-segment-out validation in the dataset and leave-one-segment-out in the whole REVIEW, respectively.

Method	**HRIS**	**VDIS**	**CLRIS**	**KPIS**
O1	−0.23	−0.35	−0.61	0.45
O2	0.002	0.06	−0.11	0.08
O3	0.23	0.30	0.72	−0.53
Gregson	3.29	1.22	−1.0	−0.23
HHFW	0.62	−0.91	-	−1.05
Zhou (1D-G)	−0.54	−3.07	−7.5	−2.57
Lowell (2D-G)	−0.17	−2.26	−6.8	−1.65
Al-Diri (ESP)	0.28	−0.05	−1.90	−0.96
Yin, Y. (2013)	-	-	-	-
Xu (Graphs)	0.21	−0.53	0.08	−1.14
Trucco	−0.42	−0.79	−0.16	−1.32
Kumar (ULDM)	0.21	−0.64	−0.55	−0.50
Lupascu	0.004	0.015	0.006	0.015
Bankhead	-	-	-	-
Yin, X. (2014)	-	-	-	-
Vazquez- G	−0.13	−0.95	0.83	−1.28
Vazquez- L	−0.22	−1.29	1.20	−1.33
Vazquez- J	0.08	−0.84	1.52	−1.05
Vazquez- I	−0.14	−1.17	1.27	−1.26
Proposed method, DoG-L7 (Cv_d)	0.002	0.007	0.01	−0.001
Proposed method, DoG-L7 (Cv_R)	0.06	−0.02	−0.16	−0.06
Proposed method, DoG-L7 (Lso_d)	−0.004	−0.08	−0.20	−0.03
Proposed method, DoG-L7 (Lso_R)	0.23	−0.09	−0.64	−0.24

Table A.3: Success rate (%).

Method	HRIS	VDIS	CLRIS	KPIS
O1	100	100	100	100
O2	100	100	100	100
O3	100	100	100	100
Gregson	100	100	100	100
HHFW	88.3	78.4	0	96.3
Zhou (1D-G)	99.6	99.9	98.6	100
Lowell (2D-G)	98.9	77.2	26.7	100
Al-Diri (ESP)	99.7	99.6	93.0	100
Yin, Y.	100	100	100	100
Xu (Graphs)	100	96.0	94.1	99.4
Trucco	95.7	92.1	90.2	93.9
Kumar (ULDM)	99.6	96.3	98.2	100
Lupascu	100	100	100	100
Bankhead	99.5	99.0	100	100
Yin, X.	100	100	100	100
Vazquez- G	78.89	78.70	91.58	100
Vazquez- L	81.71	69.68	74.39	100
Vazquez- J	73.86	57.80	80.35	100
Vazquez- I	83.36	74.34	75.79	100
Proposed method	98.4	99.7	97.5	98.8

References

Abramoff, M. D., Garvin, M. K., and Sonka, M. Retinal Imaging and Image Analysis. *IEEE Transactions on Medical Imaging*, 2010, 3:169–208. doi: 10.1109/RBME.2010.2084567.Retinal.

Al-Diri, B., Hunter, A., Steel, D., et al. REVIEW - a reference data set for retinal vessel profiles. In *Conference proceedings : 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society*, 2008, pages 2262–2265. doi: 10.1109/IEMBS.2008.4649647.

Al-Diri, B., Hunter, A., and Steel, D. An active contour model for segmenting and measuring retinal vessels. *IEEE Transactions on Medical Imaging*, 2009, 28(9):1488–1497. doi: 10.1109/TMI.2009.2017941.

Aliahmad, B., Kumar, D. K., Janghorban, S., et al. Automatic retinal vessel profiling using multi-step regression method. In *Conference proceedings : 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2011, pages 2606–2609. doi: 10.1109/IEMBS.2011.6090719.

Aliahmad, B., Kumar, D. K., Janghorban, S., et al. Retinal vessel diameter measurement using multi-step regression method. In *2012 ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC)*, pages 1–4. 2012, IEEE. doi: 10.1109/BRC.2012.6222171.

Arlinghaus, S. L., Arlinghaus, W. C., Drake, W. D., et al. *Practical Handbook of Curve Fitting*. 1994, CRC Press.

Bankhead, P., Scholfield, C. N., McGeown, J. G., et al. Fast retinal vessel detection and measurement using wavelets and edge location refinement. *PLoS ONE*, 2012, 7(3):1–25. doi: 10.1371/journal.pone.0032435.

Bishop, C. M. *Neural Networks for Pattern Recognition*. Oxford University Press, Inc., 1995, New York, NY, USA.

Björck, Å. *Numerical Methods for Least Squares Problems*, volume 11. Society for Industrial and Applied Mathematics, 1996, Philadelphia. doi: 10.1137/1.9781611971484.

Branch, M., Coleman, F., and Li, Y. A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems. *SIAM Journal on Scientific Computing*, 1999, 21(1):1–23.

Breiman, L. Bagging predictors. *Machine Learning*, 1996, 24(2):123–140. doi: 10.1007/BF00058655.

Breiman, L. Random Forests. *Machine learning*, 2001, 45(1):5–32. doi: 10.1017/CBO9781107415324.004.

Breiman, L., Friedman, J., Stone, C. J., et al. *Classification and Regression Trees*. 1984, CRC Press.

Brinchmann-Hansen, O. and Heier, H. Theoretical relations between light streak characteristics and optical properties of retinal vessels. *Acta Ophthalmologica*, 2009, 64(S179):33–37. doi: 10.1111/j.1755-3768.1986.tb00701.x.

Chapman, N. Computer algorithms for the automated measurement of retinal arteriolar diameters. *British Journal of Ophthalmology*, 2001, 85(1):74–79. doi: 10.1136/bjo.85.1.74.

Coleman, T., Branch, M. A., and Grace, A. *Optimization Toolbox For Use with M ATLAB*. 1999, Matlab The Mathworks Inc.

da Fontoura Costa, L. and Estrozi, L. Multiresolution shape representation without border shifting. *Electronics Letters*, 1999, 35(21):1829. doi: 10.1049/el:19991262.

Dashtbozorg, B. *Advanced Image Analysis for the Assessment of Retinal Vascular Changes*. Doctoral program in electrical and computer engineering,

Delibasis, K. K., Kechriniotis, A. I., Tsonos, C., et al. Automatic model-based tracing algorithm for vessel segmentation and diameter estimation. *Computer Methods and Programs in Biomedicine*, 2010, 100(2):108–122. doi: 10.1016/j.cmpb.2010.03.004.

Fraz, M. M., Remagnino, P., Hoppe, A., et al. A Model Based Approach for Vessel Caliber Measurement in Retinal Images. In *2012 Eighth International Conference on Signal Image Technology and Internet Based Systems*, pages 129–136. 2012, IEEE. doi: 10.1109/SITIS.2012.29.

Gang, L., Chutatape, O., and Krishnan, S. Detection and measurement of retinal vessels in fundus images using amplitude modified second-order Gaussian filter. *IEEE Transactions on Biomedical Engineering*, 2002, 49(2):168–172. doi: 10.1109/10.979356.

Gao, X., Bharath, A., Stanton, A., et al. A method of vessel tracking for vessel diameter measurement on retinal images. In *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, volume 2, pages 881–884. 2001, IEEE. doi: 10.1109/ICIP.2001.958635.

Girard, F., Kavalec, C., Tahar, H. B., et al. Uncertainty Assessment of Vessels Width Measurement from Intensity Profile Model Fitting in Fundus Images. In *Proceedings of the Ophthalmic Medical Image Analysis First International Workshop, OMIA 2014*, 2014, pages 41–48.

Gonzalez, R. C. and Woods, R. E. *Digital Image Processing*. 2002, Robbins, Tom.

Gregson, P., Shen, Z., Scott, R., et al. Automated Grading of Venous Beading. *Computers and Biomedical Research*, 1995, 28(4):291–304. doi: 10.1006/cbmr.1995.1020.

Guo, Z. and Hall, R. W. Parallel thinning with two-subiteration algorithms. *Communications of the ACM*, 1989, 32(3):359–373. doi: 10.1145/62065.62074.

Hastings, E. A Survey of Thinning Methodologies. *Pattern analysis and Machine Intelligence, IEEE Transactions*, 1992, 4(9):869–885.

Hoover, A., Kouznetsova, V., and Goldbaum, M. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 2000, 19(3):203–210. doi: 10.1109/42.845178.

Hunter, A., Lowell, J., Ryder, R., et al. Tram-Line filtering for retinal vessel segmentation. *Proceedings of the 3rd European Medical and Biological Engineering Conference*, 2005, 11(1): 3–6.

Ikonomovska, E. *Algorithms for Learning Regression Trees and Ensembles on Evolving Data Streams*. Doctoral dissertation,

Ikram, M. K., Ong, Y. T., Cheung, C. Y., et al. Retinal Vascular Caliber Measurements: Clinical Significance, Current Knowledge and Future Perspectives. *Ophthalmologica*, 2013, 229(3): 125–136. doi: 10.1159/000342158.

Kumar, D. K., Aliahmad, B., and Hao, H. Retinal Vessel Diameter Measurement Using Unsupervised Linear Discriminant Analysis. *ISRN Ophthalmology*, 2012, 2012:1–7. doi: 10.5402/2012/151369.

Kumar, H. and Kaur, P. A Comparative Study of Iterative Thinning Algorithms for BMP Images. *International Journal of Computer Science and Information Technologies*, 2011, 2(5):2375–2379.

Lagarias, J. C., Reeds, J. A., Wright, M. H., et al. Convergence Properties of the Nelder–Mead Simplex Method in Low Dimensions. *SIAM Journal on Optimization*, 1998, 9(1):112–147. doi: 10.1137/S1052623496303470.

Lam, L., Lee, S.-W., and Suen, C. Thinning methodologies-a comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, 14(9):869–885. doi: 10.1109/34.161346.

Lee, E. Choosing nodes in parametric curve interpolation. *Computer-Aided Design*, 1989, 21(6): 363–370. doi: 10.1016/0010-4485(89)90003-1.

Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Quarterly Journal of Applied Mathmatics*, 1944, 2:164–168.

Li, H., Hsu, W., Lee, M., et al. A piecewise Gaussian model for profiling and differentiating retinal vessels. In *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)*, volume 1, pages I–1069–72. 2003, IEEE. doi: 10.1109/ICIP.2003.1247151.

Liaw, A. and Wiener, M. Classification and Regression by randomForest. *R news*, 2002, 2(3): 18–22. doi: 10.1177/154405910408300516.

Lopez, A. M., Lloret, D., Serrat, J., et al. Multilocal Creaseness Based on the Level-Set Extrinsic Curvature. *Computer Vision and Image Understanding*, 2000, 77(2):111–144. doi: 10.1006/cviu.1999.0812.

Lowell, J., Hunter, A., Steel, D., et al. Measurement of Retinal Vessel Widths From Fundus Images Based on 2-D Modeling. *IEEE Transactions on Medical Imaging*, 2004, 23(10):1196–1204. doi: 10.1109/TMI.2004.830524.

Lupascu, C. A., Tegolo, D., and Trucco, E. Accurate estimation of retinal vessel width using bagged decision trees and an extended multiresolution Hermite model. *Medical Image Analysis*, 2013, 17(8):1164–1180. doi: 10.1016/j.media.2013.07.006.

Marquardt, D. W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 1963, 11(2):431–441. doi: 10.1137/0111030.

Martinez-Perez, M. E. *Computer analysis of the geometry of the retinal vasculature*. Dissertation thesis,

Mendonça, A. M. and Campilho, A. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE Transactions on Medical Imaging*, 2006, 25(9):1200–1213. doi: 10.1109/TMI.2006.879955.

Mendonça, A. M., Dashtbozorg, B., and Campilho, A. Segmentation of the Vascular Network of the Retina. In *Image Analysis and Modeling in Ophthalmology*, 2014. pages 85–110. doi: 10.1201/b16510-6.

Moré, J. and Sorensen, D. Computing a Trust Region Step. *SIAM Journal on Scientific and Statistical Computing*, 1983, 3:553–572.

Nelder, J. A. and Mead, R. A Simplex Method for Function Minimization. *The Computer Journal*, 1965, 7(4):308–313. doi: 10.1093/comjnl/7.4.308.

Nguyen, T. T., Wang, J. J., and Wong, T. Y. Retinal Vascular Changes in Pre-Diabetes and Prehypertension. *Diabetes Care*, 2007, 30(10):2708–2715. doi: 10.2337/dc07-0732.Abbreviations.

Niemeijer, M., Staal, J., Ginneken, B., et al. 2004, DRIVE: digital retinal images for vessel extraction. URL `http://www.isi.uu.nl/Research/Databases/DRIVE`.

Orfanidis, S. J. *Introduction to Signal Processing*. 1996, Prentice Hall, Inc.

Perez-Rovira, A., MacGillivray, T., Trucco, E., et al. VAMPIRE: Vessel assessment and measurement platform for images of the Retina. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2011, pages 3391–3394. doi: 10.1109/IEMBS.2011.6090918.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., et al. *Numerical recipes: the art of scientific computing*. 2007, Cambridge University Press.

Rice, J. R. *The approximation of functions, Vol. 2: Nonlinear and multivariate theory*. Addison-Wesly, 1969, Reading Massachusetts.

Saine, P. J. and Marshall, E. *Ophthalmic Photography: Retinal Photography, Angiography, and Electronic Imaging*. Butterworth-Heinemann, 2001, 2nd edition.

Savitzky, A. and Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares procedures. *Anal. Chem.*, 1964, 36(8):1627–1639. doi: 10.1021/ac60214a047.

Specht, D. F. A general regression neural network. *IEEE Transactions on Neural Networks*, 1991, 2(6):568–576.

Trucco, E., Ballerini, L., Relan, D., et al. Novel VAMPIRE algorithms for quantitative analysis of the retinal vasculature. In *2013 ISSNIP Biosignals and Biorobotics Conference*, 2013, pages 1–4. doi: 10.1109/BRC.2013.6487552.

Vapnik, V. *The Nature of Statistical Learning Theory*. Springer, 1995, New York.

Vazquez, S. G., Barreira, N., Penedo, M. G., et al. Evaluation of SIRIUS retinal vessel width measurement in REVIEW dataset. In *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, pages 71–76. 2013, IEEE. doi: 10.1109/CBMS.2013.6627767.

Wang, L., Bhalerao, A., and Wilson, R. Analysis of retinal vasculature using a multiresolution hermite model. *IEEE Transactions on Medical Imaging*, 2007, 26(2):137–152. doi: 10.1109/TMI.2006.889732.

WebMD. 2015, Picture of the Eyes. URL `http://www.webmd.com/eye-health/picture-of-the-eyes`.

Wijaya, I. G. P. S., Uchimura, K., and Koutaki, G. Human face security system using alternative linear discriminant analysis based classifier. In *2011 17th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV)*, pages 1–6. 2011, IEEE. doi: 10.1109/FCV.2011.5739706.

Wright, M. Direct Search Methods: Once Scorned, Now Respectable. In D.F. Griffiths and G.A. Watson, editor, *Numerical Analysis 1995, Proceedings of the 1995 Dundee Biennial Conference in Numerical Analysis*, 1995, pages 191–208.

Xu, X., Niemeijer, M., Song, Q., et al. Retinal vessel width measurements based on a graph-theoretic method. In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 641–644. 2011, IEEE. doi: 10.1109/ISBI.2011.5872488.

Xu, X., Niemeijer, M., Song, Q., et al. Vessel boundary delineation on fundus images using graph-based approach. *IEEE Transactions on Medical Imaging Med Imaging*, 2011, 30(6):1184–1191. doi: 10.1109/TMI.2010.2103566.Vessel.

Yin, X., Ng, B. W.-H., He, J., et al. Accurate Image Analysis of the Retina Using Hessian Matrix and Binarisation of Thresholded Entropy with Application of Texture Mapping. *PLoS ONE*, 2014, 9(4). doi: 10.1371/journal.pone.0095943.

Yin, Y., Adel, M., and Bourennane, S. Automatic Segmentation and Measurement of Vasculature in Retinal Fundus Images Using Probabilistic Formulation. *Computational and Mathematical Methods in Medicine*, 2013, 2013:1–16. doi: 10.1155/2013/260410.

Yuan, Y.-x. A review of trust region algorithms for optimization. *ICIAM*, 2000, 99(271-282).

Zhang, T. Y. and Suen, C. Y. A fast parallel algorithm for thinning digital patterns. *Communications of the ACM*, 1984, 27(3):236–239. doi: 10.1145/357994.358023.

Zhou, L., Rzeszotarski, M., Singerman, L., et al. The detection and quantification of retinopathy using digital angiograms. *IEEE Transactions on Medical Imaging*, 1994, 13(4):619–626. doi: 10.1109/42.363106.