Analysis of Retinal Images

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Medical and Human Sciences

2014

Reyhaneh Sadegh Zadeh

School of Medicine

Table of Contents

Word Count: 45, 210

List of Figures

List of Acronyms

AMD – Age-related Macular Degeneration

ANN – Artificial Neural Networks

CART – Classification and Regression Tree

CHC – Circular Harmonic Component

CSP – Complex Steerable Pyramid

CSR – Convex Set Region

CTA – Computed Tomography Angiography

CWT – Complex Wavelet Transform

DRIVE – Digital Retinal Images for Vessel Extraction

DT-CWT – Dual-Tree Complex Wavelet Transform

DWT – Discrete Wavelet Transform

FIR – Finite Impulse Response

FOV – Field of View

HOAC – Higher-Order Active Contour

HSV – Hue Saturation Value

ICP – Inter-Coefficient Product

JSE- Junction Segment Ends

KNN- K-Nearest Neighbour

LIBSVM – Library for Support Vector Machines

MAA – Maximum Average Accuracy

MATLAB – Matrix Laboratory

NMS – Non-Maximal Suppression

NN – Nearest Neighbour

OCT – Optical Coherence Tomography

OOB- Out of Bag

PFM – Pratt's Figure of Merit

PHT – Parabolic Hough Transform

RF- Random Forest

RGB – Red Green Blue

RISAR – Rotation Invariant Simultaneous Autoregressive

RMS – Root-Mean-Square

ROC – Receiver Operating Characteristic

ROI – Region of Interest

SAR – Simultaneous Autoregressive

SE – Segment Ends

SIFT – Scale-Invariant Feature Transform

SNR – Signal to Noise Ratio

STARE – Structured Analysis of the Retina

SVM – Support Vector Machine

VRP – Value of Radial Projection

Abstract

Diabetic retinopathy is an eye disease that causes blindness amongst many people with diabetes if left untreated. When the eye is affected, various changes in the blood vessels occur. To be able to observe these changes over time, images of the back of the eye called retinograms are acquired. Automated analysis of retinograms becomes important as the number of people afflicted with diabetes increases worldwide. A diabetic retinopathy detection system should be able to analyse retinograms, interpret them and record the changes over time.

The methods associated with blood vessel detection in retinograms have a common drawback which is that small vessels that have low contrast are normally missed out during detection. Evaluation of these methods is commonly carried out using a technique which has the drawback of having a bias towards the detection of thick vessels.

In this thesis a retinal vessel detection method is proposed which is capable of detecting blood vessels of different width, length and orientation in the back of the eye. The state-of-the-art method proves to be comparable to the existing methods applied to the same images used in this project. When applied for segmenting the whole vessel network, it achieves an area under the ROC curve (A_Z) of 0.960(±0.0021) compared to the best result of 0.972 obtained via another method.

When applied for vessel centreline detection, it achieves an A_Z of 0.977(±0.0013) which is higher than the best method with an A_Z of 0.967(±0.0017). These results are obtained using an evaluation method proposed in this project that eliminates the drawback of current evaluation methods hence removes the bias towards the detection of thick vessels.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

Acknowledgements

"I can do all things through Christ, who strengthens me"

The above bible verse has been my driving force for completing this PhD. I had lost motivation to continue but the Lord saved me and showed me that faith in Him is all I need. ALL glory to Him.

About the Author

Reyhaneh Sadegh Zadeh received a first class bachelor's degree in electronic engineering from Bradford University, graduating in June 2007. She was awarded the best final year project award. In September 2008 she obtained a Masters in Engineering with merit from the same university. In 2008 she began the research submitted to thesis in the Division of Imaging Science and Biomedical Engineering, at the University of Manchester. During the course of this research, she published the following papers related to this thesis.

- R Sadegh Zadeh, M Berks, S M Astley, C J Taylor, (2011) "Evaluation of Blood vessel detection methods" *Medical Imaging 2011: Image Processing. Proceedings of the SPIE,Orlando,***7962**:79623U-79623U-12

- R Sadegh Zadeh, M Berks, S M Astley, C J Taylor, (2010) "Detection of Retinal Blood Vessels Using Complex Wavelet Transforms and Random Forest Classification". *Medical Image Understanding and Analysis; Warwick. BMVA*: 127-132.

CHAPTER1: Introduction

1.1 Introduction

The aim of the work described in this thesis is to detect blood vessels of different width, length and orientation in the back of the eye. Images of the back of the eye are called retinograms. Retinograms are digitally recorded and are of particular interest in this project because modelling them could provide a basis for the monitoring of early eye disease – particularly diabetic retinopathy which is a major cause of adult blindness [1]. Figure 1 shows an example of a Retinogram. Figure 2 shows the blood vessels of the retinogram in Figure 1.

Figure 4. A retinogram [2].

Image of the back of the eye. The vasculature can be observed along with the optic disc through which the retinal vessels enter the eye.

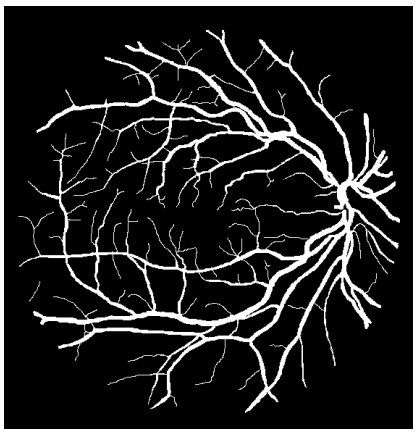


Figure 5. Blood vessel network of the retinogram in Figure 1 [2].

This is called the ground truth of the retinogram.

Diabetic retinopathy causes changes in the blood vessel network and vessels continually grow until blindness is caused. To be able to detect these changes at an early stage before blindness occurs is important for effective treatment. Many developed countries have now introduced a retinal screening programme, creating a need for automated analysis to reduce workload [3].

Current retinal vessel detection methods give less than perfect results, typically because the images are intrinsically noisy and the finer vessels have low contrast [4].

The state-of-the-art method developed in this thesis is based on two local detection methods: Random forest classifiers [5] and the Dual-tree complex wavelet transform [6]. Random forest classifiers prove better than other classifiers since they require less training data and take less time to train (Refer to chapter 4). The Dual-tree complex wavelet transform provides a rich description of local structure amongst other local detection methods and are more suitable for the detection of vessels with varying size and orientation due to their scale, rotation and shift invariant properties (Refer to chapter 3).

Effective methods of evaluating performance are important both for selecting the best detection method for a given application, and for driving the development of new detection methods [4]. Existing common evaluation methods are biased towards the detection of coarse vessels and rate methods that have detected coarse vessels better regardless of whether they have detected the very finer vessels or not.

A vessel detection evaluation method is proposed which removes this bias and rates a detection method to be accurate when all vessels in terms of their length have been detected which includes both coarse and fine vessels.

1.2 Aims and Objectives

The hypothesis of the project is that a retinal vessel detection method can be developed which is capable of capturing local structure with the following properties:

1. Adaptable to a range of scales so that it can capture both the fine and coarse vessels.

2. Shift invariant – capturing all vessels regardless of their position in the image.

3. Rotation invariant –capturing the vessels regardless of their orientation in the image.

As we will show later, the dual-tree complex wavelet transform [6] provides a promising approach. Also, an evaluation method can be developed to assess whether both fine and coarse vessels with varying width, orientation and length have been detected accurately.

1.3 Overview of Thesis

Chapter 2 gives a description of the structure of the eye and the most common diseases that manifest in the retina. An explanation of different retinal imaging techniques is also given followed by a detailed description of diabetic retinopathy and how it affects the eye. An overview of retinal image databases is given including the Digital Retinal Images for Vessel Extraction (DRIVE) database which has been used in this thesis.

Chapter 3 describes methods that can be used to extract local structure. Amongst these methods is the dual-tree complex wavelet transform (DT-CWT) which has been described in detail. The reason why it has been chosen is because it provides a rich description of local structure and provides advantages over other local methods as described in the chapter. An experiment is carried out which characterises the responses of the DT-CWT and shows how these responses extract blood vessels with varying size and orientation in retinograms.

Chapter 4 describes machine learning techniques that can be used in order to characterise local structure and classify pixels within a retinal image into vessel and non-vessel pixels. The random forest classifier and the support vector machine are identified as the best classification approaches suitable for this thesis.

Chapter 5 describes previously developed vessel detection methods. These methods are divided into low level and high level methods. The low level methods detect vessels on a pixel basis whereas the high level methods are used for connecting the discontinuities between detected blood vessels.

Chapter 6 describes the proposed method which combines the DT-CWT with random forest classifiers and how the results show to be better than combining the DT-CWT with

support vector machines. A few pre-processing methods are also applied as part of optimising the approach.

Chapter 7 proposes a few post processing methods in order to improve the results obtained in chapter 6.

Chapter 8 describes evaluation methods including the standard evaluation method applied to existing methods on retinograms of the DRIVE database. The drawback of this method is that it is biased towards detecting thick vessels in retinograms. An evaluation method is proposed which removes this bias and evaluates the detection of vessels regardless of their width.

Chapter 9 draws conclusions from the thesis and the presented work. The future work is also presented which is to develop a method that can detect the global structure within a retinogram and to combine it with the local detection method proposed in this thesis.

CHAPTER 2: Digital Retinography

2.1 Introduction

This chapter focuses on the structure of the eye, how the eye is imaged, various imaging techniques, diseases of the eye along with a description of diabetic retinopathy. There is a particular focus on Retinograms which are digital representations of the back of the eye and are used in order to monitor the state of the retina. If the retina is not in a good condition and it is diseased, the retinogram will contain various abnormalities including changes to the blood vessel network. The most common diseases that manifest in the retina are diabetic retinopathy (pathology of the retina due to diabetes), macular degeneration, glaucoma and cardiovascular disease [7].

2.2 Structure of the eye

A typical retina is shown in Figure 3. It consists of light-sensing cells, an optical nerve head and a network of blood vessels. Light enters through the pupil of the eye and is then focused on the retina. The lens helps to capture images from different distances and the iris controls the amount of light entering the eye. If the light is bright the iris closes and when it is dim it opens [8]. The retina has a thickness of about 0.5mm and lines the inside of the eyeball and has the task of converting light that passes through the eyes into a neural signal that is further processed by the brain.

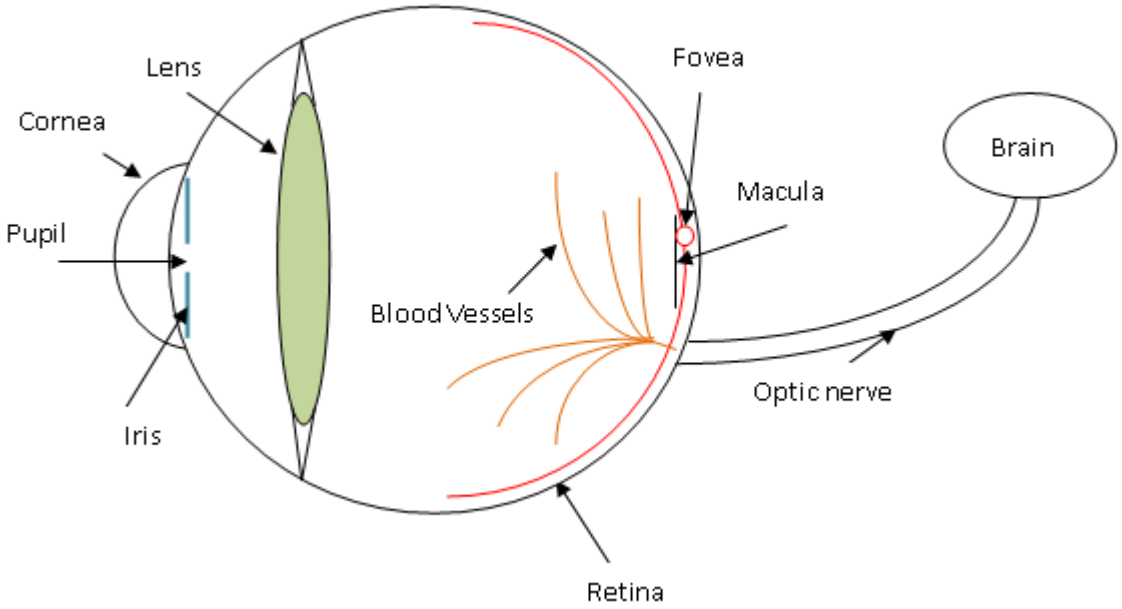


Figure 6. Location of the retina.

The retina lines the back of the eye. It is the red circle behind the lens in Figure 3.

Blood vessels transport blood throughout the body and in the retina look like thin elongated structures with different widths and lengths. They vary in thickness in the range 50-200µm with a median of 60 µm [9]. The optic nerve head is shown in Figure 3 and has a mean vertical and horizontal diameter of 1.88 and 1.77 mm, respectively [10]. The vessels emanate from the optic nerve head and branch out to cover most of the retina. Nerves that connect to the brain are all gathered in the optic nerve. The other side of the optic nerve head leads to the brain hence the retina is considered a part of the brain.
The macula is located at the centre of the retina and is considered the area with the highest visual acuteness (Figure 3) [11]. The fovea as shown in Figure 3 is located in the centre of the macula area and is the point where light rays are focussed which produces fine details for sharp vision [12, 13].

2.3 Diseases in the eye

The most common diseases that manifest in the eye are diabetic retinopathy which causes the growth of new blood vessels, age-related macular degeneration, glaucoma and

cardiovascular disease. All involve damage to the retina which can be observed in retinograms.

Figure 4 shows examples of normal and abnormal retinograms. The left image shows a normal retinogram where no damage to the blood vessel network or any other structure is present, whilst the right one shows an abnormal one where the optic disc is in the form of an extended yellow patch.

Figure 7. Example of retinograms [2].

Left: Normal retinogram with no signs of disease. Right: Abnormal retinogram. Damage has been caused on the optic disc which is shown as the expanded yellow patch and differs from the normal optic disc shown in the left retinogram. The difference in colour between these two images is a normal variation.

Diabetic Retinopathy affects 80% of people with diabetes worldwide. It is one of the leading causes of blindness in the United States [14, 15] and the second leading cause of blindness in the western world [15].

Diabetic Retinopathy has three stages: 1. Pre-proliferative diabetic retinopathy, 2. Proliferative diabetic retinopathy and 3.Severe diabetic retinopathy.

During pre-proliferative diabetic retinopathy the retinal arteries are weakened and start to leak and form haemorrhages [16]. Microaneurysms (Figure 5) which are small red dots are also formed along with dot-haemorrhages. These red dots initiate from the fovea area and expand to different areas of the macular region. They are produced from weakening of vessel walls of the smallest vessels called capillaries. When microaneurysms burst, haemorrhages (Figure 6) are caused. Haemorrhages differ from microaneurysms as they become larger and become flame shaped and expand towards other anatomical structures in their surroundings [17].

Exudates are also formed which are yellow patches with sharp edges. They are plasma protein precipitation [17] or lipoprotein deposits [15] that have been caused by retinal vessel bleeding and appear as white lesions (dots) that have the same size as microaneurysms but can become larger by merging with their neighbouring exudates. At this stage, swelling or Oedema is caused in the retina which impairs vision [8].

Proliferative diabetic retinopathy causes severe bleeding and leakage of blood and an increase of exudates around the macula [16]. This stage is very dangerous for the central vision since retinal exudation expands to the fovea region and causes arterial and capillary closing. White patches called cotton wool spots are also formed. Some areas of the retina become oxygen deprived or ischemic and whilst the circulatory system attempts to maintain adequate oxygen levels, new vessels start to grow which causes changes in vessel diameter leading to changes in the blood vessel network. These new vessels are called neovascularisations [18].

During the final stage, an increase of new blood vessels occurs that causes blindness by either a number of haemorrhages occurring or by retinal detachment from connective tissues of new vessels [15, 19]. Figure 6 shows diabetic retinopathy progression at a severe level [20].

Age-related Macular Degeneration (AMD) has two forms: I) Dry AMD where visual acuity is slowly lost and ii) Wet AMD where a growth of vascular structure into the macula is caused. The increase in vascular permeability causes fluid to be gathered underneath the retina which leads to permanent visual loss [7].

Glaucoma involves raised pressure within the eye, causing damage to the optic nerve [7]. Cardiovascular disease which causes deep white spots in the retina, widening of retinal veins and thinning of retinal arteries is another form of disease which manifests in the retina [7].

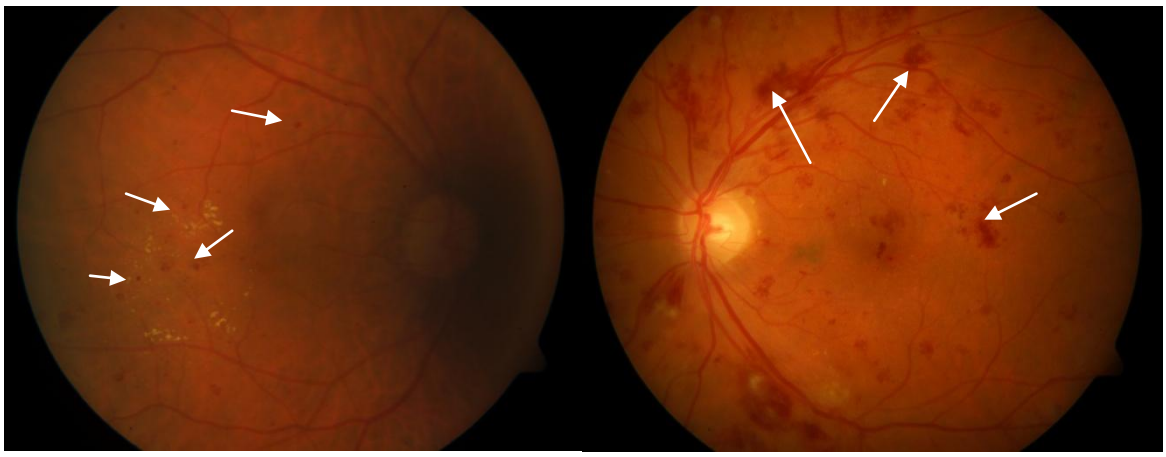


Figure 8. Left) Microaneurysms shown with arrows. Right) Haemorrhages shown with arrows [20].

Figure 9. Severe diabetic retinopathy [20].

Untreated diabetic retinopathy will cause blindness, therefore monitoring the disease progression via population screening will allow for timely and effective treatment. Screening programs are now taking place in various countries and early detection is only possible through automated analysis since there is too much workload for manual detection [7, 11]. The machine analyses the retinal images of patients and if abnormalities are found, a retinal specialist carries out a detailed diagnosis [18]. Screening should be performed at regular intervals such as every six months to a year [21].

When diabetic retinopathy is identified from screening, successful treatment can be performed by laser surgery and strict glucose control. Blindness cannot be reversed once it has occurred due to diabetic retinopathy but further damage to the retina can be avoided. The laser treatment is called photocoagulation where the laser is directly applied to the leaking microaneurysms so that further haemorrhaging is avoided. It can also be applied to

a larger part of the retina to reduce the need for oxygen and further damage. Photocoagulation reduces the risk of blindness before blindness has occurred [18].

2.4 Imaging the eye

In order to be able to analyse the retinal structure and its changes over time, the retina should be imaged effectively. Blood vessels need to be regularly detected so that early signs of disease can be identified and the retina can be treated in advance. Vessel segmentation is not trivial due to similarity between vascular and background intensities of finer vessels and a difference in the vessel parameters such as their width and tortuosity.

To be able to analyse them automatically is also extremely important since the retinal images of patients are now acquired routinely, this produces significant workload and it becomes an impossible and inaccurate task for humans to carry out.

Retinal imaging techniques are developed in order to capture the retina morphology effectively and to model its changes over time. Fundus photography, Optical coherence tomography (OCT) and Fluorescein angiography are retinal imaging techniques that are widely used for the detection of diabetic retinopathy. Below gives a brief description of each technique.

2.4.1 Fundus Photography

The retina can be directly examined using digital photographs. A digital fundus camera is used to take images of the eye (Figure 7). It contains special optics that are used to produce photographs of the retina [18]. An array of light sensitive diodes converts the light signals into electrical signals. The number of pixels that are created from the array from the analogue image will define the spatial resolution of the image [22]. The camera is described by its angle of view which is the angle of its lens. The normal angle is 30º. Magnification depends on the angle of view, the focal length of the lens and the distance between the object and the lens [23].

Fundus images produced by a fundus camera are 2D images that represent retinal semi-transparent tissues projected on the image plane obtained by using reflected light. The quantity of the reflected light of a particular waveband is recorded as the image intensities. For colour fundus images, the quantity of the reflected Red, Green and Blue (RGB) wavebands are used for the intensities. An RGB image is an M x N x3 array where each pixel is a triplet corresponding to the 3 colours red, green and blue. An RGB image can have different classes which determine its range.

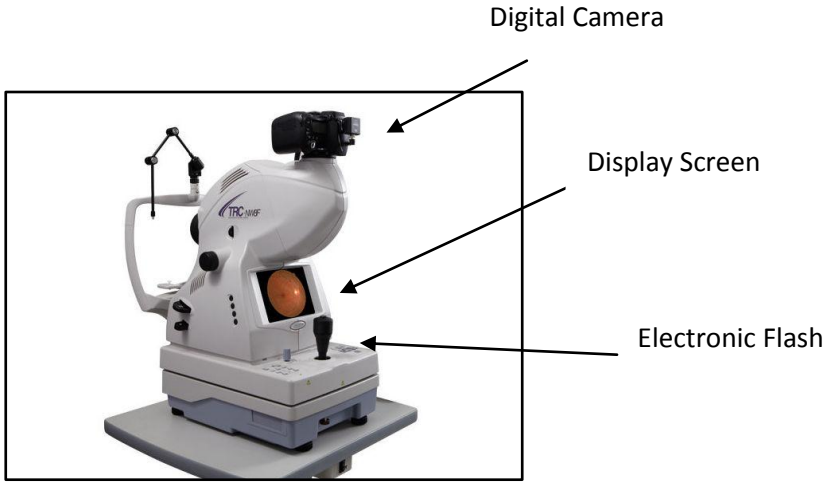


Figure 10. A fundus camera. Taken from Topcon medical [24] with permission.

The display screen shows the captured retinogram. The electronic flash generates and focuses light.

Retinograms that are used in this project are fundus photographs. The reason why fundus images are used in this project is that the publicly available databases have used fundus photographs of patients. Also, fundus photography is more useful for population based and large scale detection of retinopathy [7]. The images can be accessed anywhere as they can be sent via internet to a central server. A trained reader specialized in diabetic retinopathy can log into the server and grade the images. It also allows for quick storage and immediate examination of the retina and the interpretation stages [18].

2.4.2 Fluorescein Angiography

Fluorescein angiography photographs a sodium fluorescein dye as it travels through the blood vessels of the eye. The dye is injected into the hand or arm. When a blue light is shone on the dye, it turns into a yellow-green colour and an image of the vessel structure is

captured by the camera. The image intensities represent the amount of emitted photons from the dye. The resulting images are called angiograms. A fundus digital camera can be used to obtain an angiogram.

Figure 8 shows an example of a fluorescein angiogram:

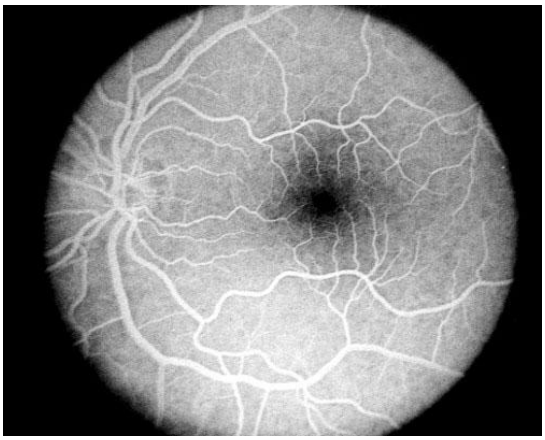


Figure 11 . A Fluorescein Angiogram. Taken from Robert Scharf [25] with permission.

Maximal fluorescence of the vessels and enhanced resolution of vessel detail has occurred.

2.4.3 Optical Coherence Tomography

Optical Coherence Tomography (OCT) produces high-resolution cross-sectional images of the internal micro-structure of biological tissue [22] and works by measuring back reflected light in order to image internal microstructures with a resolution of 1-15 μm. OCT produces three-dimensional images and does not require any subject preparation like injection. Figure 9 shows a standard Michelson interferometer which is used to obtain images via OCT. The light is divided into a reference path and a sample path which is then combined into an interference signal after it is back reflected from the reference mirror. When the path lengths of the sample section and the reference mirror are equal to the coherence length of the light, the interference signal is created which carries the sample information. The outgoing light is reflected from the mirror and exits the photo detector to produce the photo.

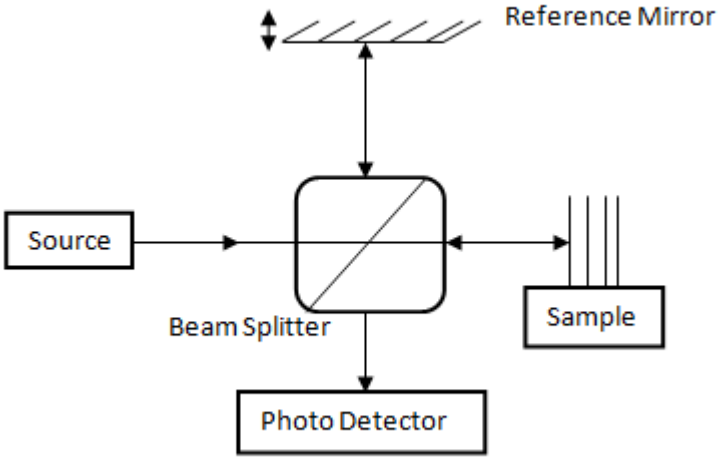


Figure 12 . Optical Coherence System [23].

The disadvantage is its limited imaging depth and that it may only produce an image of around 1mm sample depth. It can only provide for high resolution cross sectional images of tissues up to 1-2mm deep and any light penetrated further than that will be too scattered and an image cannot be created. Figure 10 shows an OCT scan containing different colour-coded layers of the retina.

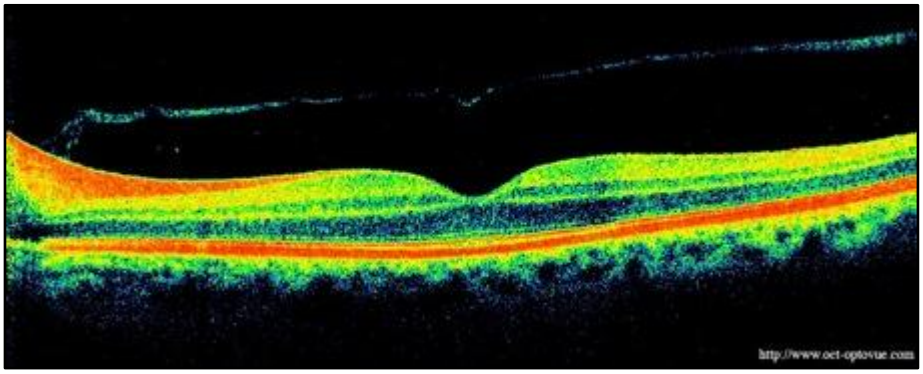


Figure 13 . An OCT Scan. Taken from Ophthalmology and Optovue with permission [22].

Different colour-coded layers of the retina are displayed.

2.5 Retinographic Image Data

The DRIVE and Structured Analysis of the Retina (STARE) databases are publically available retinogram databases which contain retinal images. They are widely used for retinal vessel detection. A brief description of them is given below:

2.5.1 DRIVE Database

Figure 11 shows an image, typical of those obtained in the screening programme, obtained from the publicly available DRIVE (Digital Retinal Images for Vessel Extraction) database which we have used in this project [2]. The images are obtained from a screening program in the Netherlands. The screening population consists of 400 subjects with diabetes aged between 25-90 years. 40 retinograms have been randomly selected, 33 are normal whilst 7 show mild diabetic retinopathy..

A Canon CR5 non-mydriatic camera was used with a 45 degree field of view (FOV). Each image has a size of 584×565 pixels. For each retinogram, a mask version is also provided which shows the FOV area. The FOV has a diameter of about 540 pixels.

A non-mydriatic camera is a camera that does not require dilation of the pupils therefore eliminating the need for dilating drops and bright lights. Therefore, it provides for faster examination as it takes less time for the pupils to adjust after they see a flash of light. A single non-mydriatic digital photograph is equivalent to standard color photography and provides for more sensitivity than mydriatic cameras in the detection of diabetic retinopathy [27].

This digital camera system allows immediate view of the fundus photos and the unique feature of the Canon system is a movable internal fixation light that can be enlarged up to four times bigger than the original size. The advantage of this movable light is that a full scan of the eye becomes viable allowing for other pathology to be found. Also the larger size is advantageous for low vision patients for finding pathology that was not previously detected. The resulting images can be electronically transmitted which allows for faster referral and appointments of diabetic patients to be scheduled quickly [27].

The digital resolution is lower than current systems that are 1365×1000 pixels as stated by the national screening committee in the UK. The resolution in current systems is ideal for the detection of very small structures specifically microaneurysms [27].

The national screening committee have specified an image with 45 degree field of view as a requirement for an image with good quality [28].

The set of 40 images are divided into a test and training set both containing 20 images. Each retinogram also has a ground truth version which is the manual segmentation of the image (see Figure 12). They are segmented by two observers. The first and second author of the database [2].

Figure 14 . A Retinogram from the DRIVE database [2].

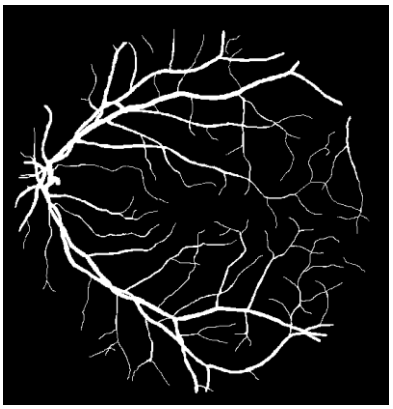


Figure 15 . The segmented ground truth of the retinogram in Figure 11 from the DRIVE database [2].

2.5.2 STARE Database

The STARE (Structured analysis of the retina) database also known as the Hoover database consists of 20 images, 10 of them are normal and contain no pathology whilst the other ten are abnormal. The images are captured by a TopCon TRV-50 fundus camera with 35º field of view. Each image has a size of 605 × 700 pixels with 24-bits per pixel. Hand labelled segmentations of each image have been produced by two experts [29].

The disadvantage of this database is that the images are not screening images. They are scanned from film therefore the quality of the images are lower than what they would have been if acquired using digital fundus cameras. Also, the variability between the manual segmentation from the first and second observer is large [18].

Based on the above, we have chosen the DRIVE database for this project. It contains more images than the STARE database. The images are screening images and are a random selection therefore they contain adequate variability. Since this project focuses on blood vessel detection which includes local information and not abnormality detection, the 20 images used for testing are sufficient as one image contains many blood vessels. Also it is the only database that has been widely used by other methods therefore comparison can be made between obtained results and other results.

2.6 Summary

In this chapter, the structure of the eye was explained and how the retina is digitally recorded and analysed in order to observe whether it is normal or abnormal. The most common diseases that manifest in the eye were also described. Diabetic Retinopathy which is a common cause of blindness in diabetic patients is of particular interest in this project and was explained in detail in section 2.3. When the retina is affected by this disease, the blood vessel network changes as new vessels continually grow in order to supply the retina with oxygen. Therefore it is important for an automated system to monitor these changes and to provide for early diagnosis.

Several automated vessel segmentation methods have been tested on the DRIVE database to date, a few of which are described in chapter 5. All methods use different approaches to segment vessel and non-vessel pixels in retinograms.

In the next chapter local image representation techniques that can be used to extract structure from retinograms at a pixel level are described. Amongst these, the best local image representation method is chosen for this project.

CHAPTER 3: Representation of local Image Structure

3.1 Introduction

There are various techniques that can be applied to an image in order to extract local information. This local information, called Texture, which forms repeated patterns in an image, can be segmented via texture segmentation methods. In retinograms, the blood vessels, branches, and crossings, are examples of such repeated patterns.

A good texture method should be rotation invariant and be able to capture texture in an image regardless of orientation. It should be shift invariant and capture texture regardless of a change of position within an image and, in terms of scale, it should be able to deal with a range of structure sizes.

This chapter describes and compares a number of local representation/texture methods that can be used to extract local structure in retinograms.

3.2 Texture Models

Texture models can segment textural patterns and characterise the two-dimensional variations in the intensities present in the image. Texture models can be divided into statistical, model based, filter-based and structural methods [30].

Statistical methods use statistical features to define the spatial distribution of grey values. The statistical features are broadly classified into first-order, second order statistics and higher-order statistics depending on the number of pixels involved. First-order statistics define properties such as average and variance of pixels and do not take into account the spatial interaction between pixels. Second-order and higher order statistics consider pixels at specific locations relative to each other and estimate the properties between them.

To construct models with the ability to identify and synthesise texture is the basis of model based methods. These methods model an image in terms of a probability model or as a linear combination of a set of basis functions [30]. The resulting coefficients are used to describe the texture within the image. The drawback of these methods is the difficulty in

choosing the right model for the required texture and how to estimate the coefficients of the models.

Filter-based methods analyse the textures in the spatial and frequency domains. Convolutions of the original image are obtained and parameters such as energy are computed over the neighbourhood of each pixel. The common problem with these methods is the edge effects that are caused where significant errors arise along the boundaries of different textures.

Structural methods assume that textures are composed of texture primitives. The arrangement of the primitives according to placement rules and their extraction forms the basis of these methods. The primitive can be in the form of a single pixel with a grey value or usually a collection of pixels. These methods provide good symbolic descriptions of images which is ideal for synthesis. The drawback with these methods is that the resulting descriptions are ill defined for natural textures due to the variability of primitives and spatial arrangements between them and no clear distinction.

3.2.1 Statistical Methods

The most common statistical methods that are used for modelling texture are described below.

3.2.1.1 Circular Harmonic Decomposition

Circular harmonic decomposition is a method that identifies and describes 2D patterns that are rotation invariant by applying the circular harmonic expansion [31] which is closely related to the Fourier expansion. The circular harmonic expansion of an object is a 1-D Fourier expansion of a 2-D object function $f(\mathbf{r},\theta)$ with respect to the radial variable **r** and the angular variable θ in polar coordinates. An image is decomposed into a set of circular harmonic components (CHCs) in its polar form. It is then projected onto the harmonic operators whose convolution with textures creates complex outputs with magnitude and phase values. The magnitudes of the resulting coefficients give the texture features which are invariant to rotation and hence patterns are recognized. The disadvantage of this method is that it is only scale and rotation invariant [31, 30].

3.2.1.2 Moment Invariants

An image can be characterized by its 2D pattern features called moments [32].Orthogonal complex Zernike moments are obtained through projection of the function $g(x,y)$ onto Zernike polynomials which are orthogonal over the unit circle. The Zernike moments have magnitudes which are rotation invariant [33] and thus the features described by them can be constructed to an arbitrary higher order which can be used for reconstruction [32]. This method works very well when the same object can be present with different scale, and orientation characteristics. The drawback is that it is very sensitive to the background around the wanted object (since this affects the calculation of the moments) [30].

3.2.1.3 Feature Distribution Method

This method aims at discriminating between textures by learning from the shape of distributions of a number of existing rotationally invariant measures (e.g. the methods described above) [34]. Most approaches of texture classification quantify texture measures by single values such as mean and variance but this will cause most of the information contained in the whole distribution of feature values to be lost. Therefore a log-likelihood pseudo-metric, the G statistic, is used for comparing feature distributions. If the value is low, a good similarity between two distributions is observed, but if it is high, it means that the two are not identical. This method has the drawback of not capturing enough information of texture with strong global directionality since the features are based on local window statistics. This degrades the texture classification performance [30].

3.2.1.4 Scale Invariant Feature Transform

The Scale-invariant feature transform (SIFT) is an algorithm that is used to model the local features of images [35]. The features detected by SIFT are based on the appearance of an object at different locations in the image and are rotation and scale invariant. The initial step in the algorithm is to detect interest points in the original image. This is done by

convolving the image with a series of Gaussian filters at different scales and by computing the difference between the resulting blurred images successively. The maxima and minima points are identified by comparing each pixel with its 26 neighbours in 3×3 regions in its current and adjacent scales. The points obtained are called key-points. An orientation is assigned to each of them. This is done on the basis of local image gradient directions.

The magnitude and orientation calculations are done for every pixel in the neighbouring region around the key-point in the Gaussian blurred image. The rotation, scale and orientation invariance properties have already been achieved from the previous steps. The pixels around each key-point are sampled whilst the region is separated into sub-regions. A 4×4 gradient window is created for each key-point and a histogram with 8 bins for each window. This leads to a SIFT feature vector of 4×4×8 elements which is normalized to enhance illumination invariance.

The advantage of this method is that the key point descriptors are distinctive, which allows a single feature to find its correct match with good probability in a large database of features. The disadvantage of this method is that false matches are found in a cluttered image because many features from the background will not have a correct match in the database containing the features extracted [30].

3.2.1.5 Optical Digital Method

The image is characterized by the 2D Fourier spectrum of the texture [30]. The magnitude of the spectrum is obtained by measuring the distribution of energy over the different range of frequencies. In the Fourier spectrum, a contour line is defined by selecting the points whose energy is equal to a given percentage of the central maximum, hence contours of constant energy are extracted from the Fourier spectrum, and rotation invariant features that describe the shapes of the resulting 2D curve are derived. The major drawback of this method is the difficulty in locating the contour line in the spectrum domain [30, 36].

3.2.2 Model Based Methods

The most common model-based method that is used for modelling texture is the Markov Random Field Model which is described below.

3.2.2.1 Markov Random Field Model

In Markov Random Field (MRF) models, images and their textures are modelled using Gaussian random fields, providing for both scale and rotation invariant results. This means that textures are described by relationships between neighbouring pixels. The assumption is that if a small neighbourhood of pixels is chosen the value of intensity of the centre pixel can be predicted from its neighbours.

Simultaneous autoregressive model (SAR) and Rotation invariant SAR (RISAR) models are examples of MRF models. In SAR [37, 38] the centre pixel is represented as the weighted average of its neighbouring pixels. If s is a position of a pixel in an image and $g(s)$ is the intensity at position s, the second-order neighbourhood of the pixel is identified and a SAR model is constructed from:

$$g(s) = \mu + \sum_{r \in \mathbf{D}} \theta(\mathbf{r}) g(s + \mathbf{r}) + \varepsilon(s) \tag{1}$$

where μ is the bias which is dependent on the mean grey value of the image, $\theta(\mathbf{r})$ are the model parameters where **r** is a point in set **D,** characterising the dependence of a pixel to its neighbours, **D** is the set of neighbours of pixels at site S and $\varepsilon(s)$ is an independent Gaussian random variable with zero mean and variance σ^2.

$\theta(\mathbf{r})$ is learnt from training data. Textures can be generated from a SAR model by initialising with, say, a noise image and iteratively adjusting pixel intensities until they are consistent with the model.

RISAR [37, 38] is a rotation invariant version of SAR where the sampling is circularly symmetric which means that the average intensity around circles are taken as sampling points. A set of measurements are derived from the circle which can describe the texture.

The problem with SAR and RISAR is that they are computationally expensive and it is difficult to choose an appropriate neighbourhood size in which pixels are regarded as independent [30, 39, and 40].

3.2.3 Filter Based Methods

The most common filter-based methods that are used for modelling texture are described below.

3.2.3.1 Multichannel Gabor Filter

Gabor was one of the first to propose an approach based on filter banks [41]. The approach involves oriented band-pass filters that can analyse local structure in terms of orientation and scale. If quadrature pairs of Gabor filters are used, magnitude and phase information can be extracted. An important property of Gabor filters is that they have optimal joint localization, or resolution, in both the spatial and the spatial-frequency domains [42]. The filters can be adjusted in terms of their orientation, frequency bandwidths and central frequencies, thus a central issue in applying these filters to texture segmentation is the determination of these parameters [30, 43].

3.2.3.2 Wavelet Based Methods

Wavelet based methods are a special case of filter-based methods that will be discussed in more detail in section 3.3.

3.2.3.3 Steerable Pyramid

Like the Gabor approach, a steerable oriented pyramid comprises a set of oriented band-pass filters that are tuned to different orientations and scales [44, 45]. A key feature is that a finite set of filters is designed so that the response at any angle can be computed. This

means that the dominant orientation can be estimated at each point and, hence, orientation invariant features can be computed. Although this method is reasonably computationally efficient, it does not guarantee perfect reconstruction and thus important detail may be lost. Portilla et al. [46] present a modelling and synthesis approach based on steerable pyramid decomposition, which is described in more detail in section 3.2.5.

3.2.4 Structural Methods

Structural methods classify textures in terms of their elements (texels) and how they are arranged with respect to each other. Texels have certain properties such as their tone, shape, directionality and area which are invariant to rotation and can be used to distinguish between textures. Many methods such as the invariant histogram [47], texture descriptors [48], perimeter contribution [49,50,51,52] and morphological decomposition [53] have been implemented which are scale and rotation invariant but have a common drawback which is difficulty in extracting texels from textures. Structural methods are best suited to textures composed of regularly repeating patterns, where it is relatively simple to define the texels [30].

3.2.5 Texture Model Examples

Below describes an example of a texture model and how the model extracts texture and models local structure in an image:

Portilla et al. [46] develop a method which works by learning a set of statistical constraints from training images. There are two main steps in this method which are shown in Figure 13:

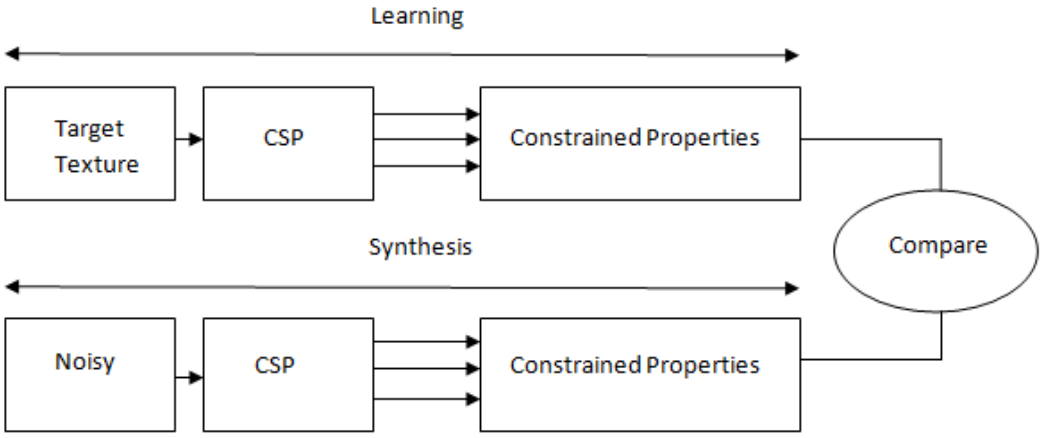


Figure 16 . The two main steps of texture classification by Portilla [46].

It consists of a learning and synthesis stage where the outputs are compared to each other.

Learning the constraints: In this process, the original images are decomposed into complex steerable pyramids (CSP) each consisting of a low pass residual image and a set of oriented sub-bands at a series of scales. This transform has the benefit of providing for translation invariance and rotation invariance properties as well as a good reconstruction of the image. Statistical properties of the CSP coefficients are measured which form a set of constraints consistent with the training images. Each constraint C_i consists of a property P_i and a constant value K_i corresponding to that property.

In total, 710 constraint properties were identified for textures. The necessity of these properties was examined by generating synthetic images ignoring each constraint in turn, and examining the quality of the result. From this, the constraints which captured important characteristics were identified and integrated in the synthesis algorithm.

Synthesis: A noise image is taken as an initial guess and is then decomposed using the steerable pyramid. The constraint properties are measured and compared with the constraint properties measured for the training images. Each of the properties is compared to the training values one at a time and if they are all identical, a legal texture is declared. If a property is not the same, it needs to be modified in a way so as to minimise the change

to other properties. This is done using an algorithm that approximates orthogonal projection.

If the image is represented as a set of values (N pixels) which together form a vector in N dimensional space, a manifold which is the set of images that satisfy all of the constraints identified during training is assumed to exist in that space. When a property P_i is found to require adjustment it is moved towards the manifold in the direction of maximum gradient which is determined by the derivative of the property with respect to the image ($\frac{\delta P_i}{\delta x}$), where δx is the difference between the image which satisfies all constraints and the noise image. The ideal situation would be to move in a way so as to change the property as rapidly as possible whilst changing the other properties and the image as little as possible, but this is not guaranteed. Once the point is as close as possible to the manifold, the image satisfies that particular constraint. The other properties are then adjusted in turn using the same process and hence a new different image that is similar to the original is produced. This process is shown below in Figure 14.

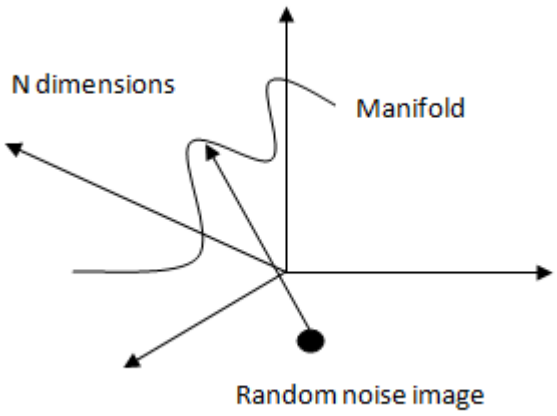


Figure 17. The manifold in N dimensions.

A random noise image is moved towards the manifold to satisfy the appropriate constraints and properties.

The method described above is very simple and, in practice, typically reaches convergence after 50 iterations. The model works very well on periodic textures. For samples that have patterns placed at non–repeated positions, it also produces reasonable results.

However, the paper shows examples where the model fails. For example:

Straight and curved contours in textures cannot be distinguished and if they are at different orientations, the original straight lines will convert to curve ones. Structures that have different grey-scale values will result in a final image without the low value pixels because the marginal statistics only work efficiently on high contrast structures. Closed contours will not be fully reconstructed, although the size and thickness will be the same. Edges cannot be distinguished from lines due to the phase of the edges which are calculated during the steerable pyramid process and are seen to be positive or negative which results in them being averaged to zero and hence step edges are displayed as dark or bright lines.

Therefore the main limitation of texture models is that they cannot capture the global organization.

3.3 Wavelets

Wavelet analysis provides a powerful basis for capturing local structure. Wavelet transforms are described in detail below, because they have the capability of extracting detailed structure efficiently. The dual-tree complex wavelet transform is shown to have properties that make it attractive as a basis for further work.

3.3.1 Definition

Wavelets (specialised filters) provide information on the local frequency structure of images. This is achieved by convolving the image with wavelets that are localised in both frequency and space. The output of the filter responses are the wavelet coefficients.

The advantages of using wavelets are that data operations can be carried out by using only the linear series of coefficients that are obtained from analysing the wavelets. They also provide for a good representation of signals that contain discontinuities.

There are many kinds of wavelets such as Daubechies[54], Symlets[55], Biothorgonal[55] and many more. The Haar wavelet [55] is the first known example of a wavelet and has been used here for illustrative purposes.

The mother Haar wavelet is shown in Figure 15 and is represented as a square function (equation 2).

$$\psi(t) = \begin{cases} 1, & 0 \le t < 1/2, \\ -1, & 1/2 \le t < 1, \\ 0 & otherwise \end{cases} \quad (2)$$

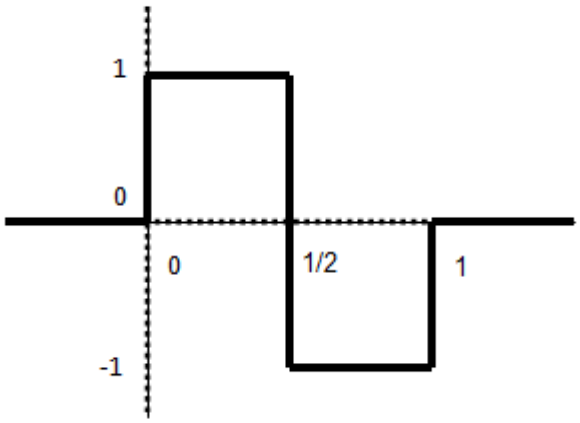


Figure 18 . The Haar wavelet which is represented by a square function.

When the mother wavelet is scaled and translated in space, continuous wavelets at different scales and positions are produced, as shown in Figure 16.

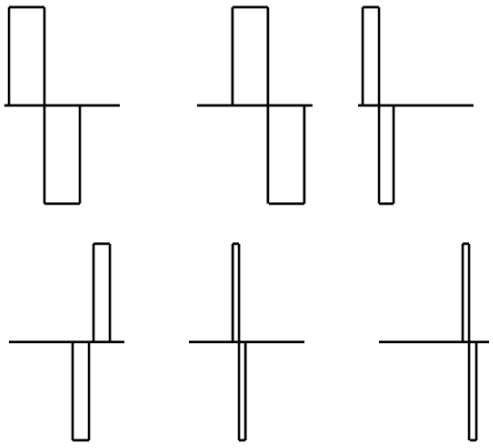


Figure 19 . Mother Haar wavelet.

It is scaled and shifted in the time domain. Top left: Original. Top middle: Haar shifted to the right. Top right: Haar scaled. Bottom row: Haar shifted and scaled.

3.3.2 Wavelet Transforms

This is the process where the wavelets are applied to an image. There are a few variations of wavelet transforms namely the discrete, continuous and complex wavelet transforms. Discrete wavelet transforms and, in particular, the dual-tree complex wavelet transform [6] are relevant to this project and are described in detail below. The most important advantage of wavelet transforms is their ability to characterise textures at different scales [56]. They apply the band-pass filtered approach and provide for a complete and non-redundant description of an image. Complex wavelets provide for oriented responses and are computationally extremely efficient. The steps in the analysis of a texture would be to initially decompose it using the transform, then to model the probability density function of the coefficients and finally extract parameters from the model to classify the texture [57].

3.3.3 Discrete Wavelet Transform

In the discrete wavelet transform (DWT), the mother wavelet is shifted and scaled in discrete steps to give a set of wavelets that can be discretely sampled.

Figure 17 shows an image that is transformed discretely. The left colour image is the original image. High-pass and low-pass filters are applied to it which separates the high frequency (images with a red frame) and low frequency components. The high frequency components are the details of the original image and the low frequency components give a low resolution version of the original image. These components are called wavelet coefficients (sub-bands). The filters are repeatedly applied to the down-sampled images (low resolution images) until a very blurred image is obtained and no more details can be extracted. The blurred image is shown as the bottom right image. The low-pass image is down-sampled by a factor of two at each level which means that the image becomes half of its original size in terms of dimensions. The detail coefficients and the final compressed image are required for reconstruction.

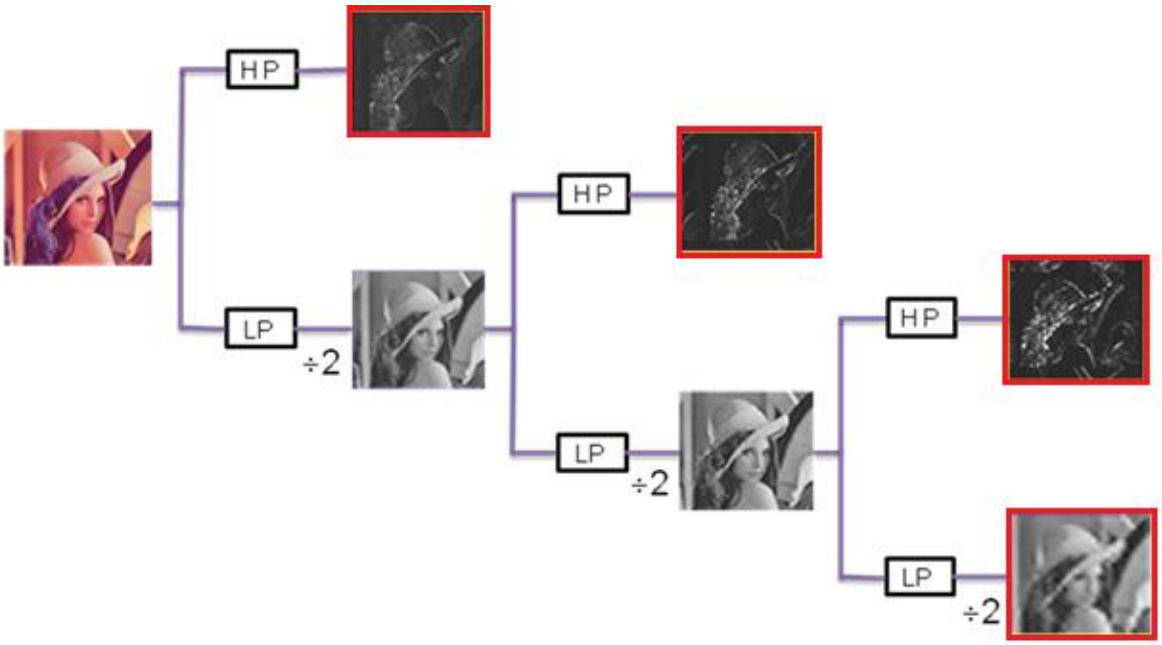


Figure 20 . The original image is shown on the top left and its compressed image is shown on the bottom right.

At each level the image is high pass and low passed to give the high and low frequency components. The filters give a response at each pixel and at each scale.

One of the drawbacks of the discrete wavelet transform is its shift dependence property which means that when the input signal is shifted relative to the sampling grid, the values of the wavelet coefficients can change, even though the image can still be perfectly reconstructed.

Another disadvantage of the standard 2D DWT is that it provides very limited information on the orientation of image features. This is due to the fact that the filters are real and half-banded [6].

3.3.4 Dual Tree Complex Wavelet Transform

To overcome the problems of the DWT, complex wavelets can be used, producing complex wavelet coefficients. A complex wavelet consists of two real wavelets that have a $90°$ phase difference (even and odd) producing a real and imaginary response from which magnitude and phase can be computed.

It is desirable for wavelets to be spatially localised and have finite support so DWT can be implemented with Finite Impulse Response (FIR) filters. The problem is that a finitely supported function will not be analytic. An analytic wavelet is a wavelet with no negative frequency components. Hence approximately analytic wavelets need to be chosen and a complex wavelet transform (CWT) that is approximately magnitude/phase, free from aliasing and most importantly shift invariant.

For phase to be used there is a limited choice of basis functions. The scaling function and wavelets need to be designed so that they are complex and approximately analytic. Designing such wavelets is more difficult than designing real valued wavelet basis. The wavelets should be 90 degrees out of phase with each other, give perfect reconstruction and be approximately analytic. There are certain constraints that need to be satisfied in order for the dilations and translations of a single function ψ (t) to constitute a basis for signal expansion whilst providing for perfect reconstruction [6]. So there is a trade-off between reconstruction ability and being analytic. We used filters already designed by Kingsbury et al [6] based on the above properties.

One form of complex wavelet transform is the dual-tree complex wavelet transform (DT-CWT) which is approximately shift invariant [6] and is based on a real discrete wavelet transform with two high pass filters instead of one. It works by convolving the image with even and odd filters. The even and odd filters are applied at different angles providing information on orientation at each pixel and scale.

The filters are oriented at : $\pm 15, \pm 45, \pm 75$. These orientations are achieved by combining one dimensional real and imaginary high-pass and low-pass filters along with their conjugates and choosing the combinations that provide for the most detail (Appendix 4).

In Figure 18, the directionality of the coefficients is shown, separated into real and imaginary parts. The top row shows the even filters which are symmetrical and the lower row shows the odd filters which are anti-symmetrical.

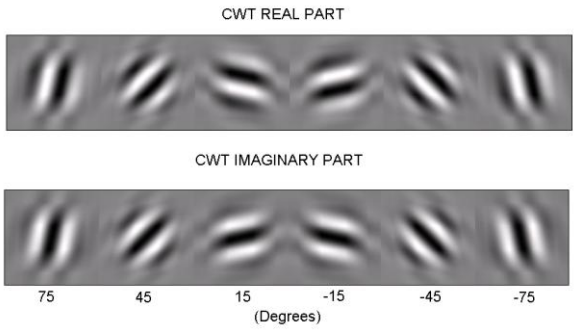


Figure 21 . The real and imaginary coefficients of the Dual-tree wavelet transform.

Each filter is applied to the image at different orientations.

The even filters give a real response and the odd filters give an imaginary response. This is at each pixel, orientation and scale.

$$\mathrm{R}(x,\theta,s)+\mathrm{il}(x,\theta,s) \tag{3}$$

Where R is the real part and il is the imaginary part at each pixel ***x*,** at each scale *s* and each orientation θ.

From each of the responses a magnitude and phase can be extracted. The magnitude is obtained from equation 4. The phase $\angle \mathrm{c}$ is obtained from equation 5.

$$M(x,\theta,s)=\sqrt{\mathrm{R}^2+\mathrm{I}^2} \tag{4}$$

$$\angle \mathrm{c} = \arg \mathrm{R}(x,\theta,s)+\mathrm{il}(x,\theta,s) \tag{5}$$

The magnitude of an object is invariant to translation so when the object is shifted in the image, its magnitude stays the same as its phase varies.

The phase indicates the positioning of an object in an image; it varies as you move across an object. For example, the phase in the centre of a symmetrical object is zero and varies in other points within the object.

Figure 19 shows the sub-band directions of the DT-CWT. A detailed explanation of how the sub-band orientations arise is given in Appendix 4:

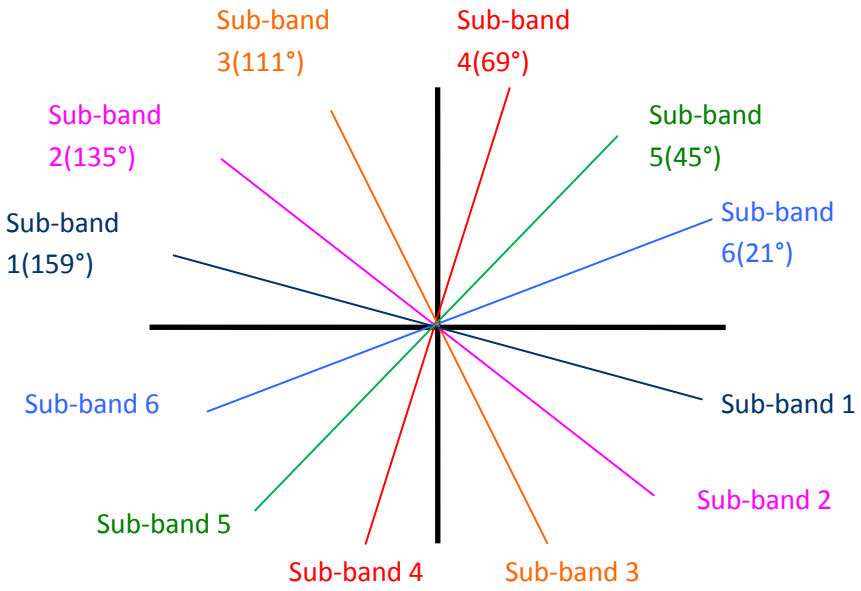


Figure 22 . The dual-tree complex wavelet transform sub-band directions.

Figures 20 and 21 show an example of applying the DT-CWT at 3 levels to a binary version of a retinogram.

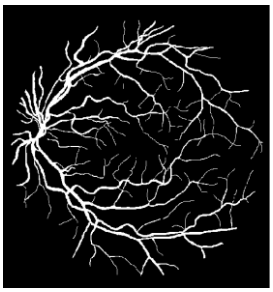


Figure 23 . Original Image

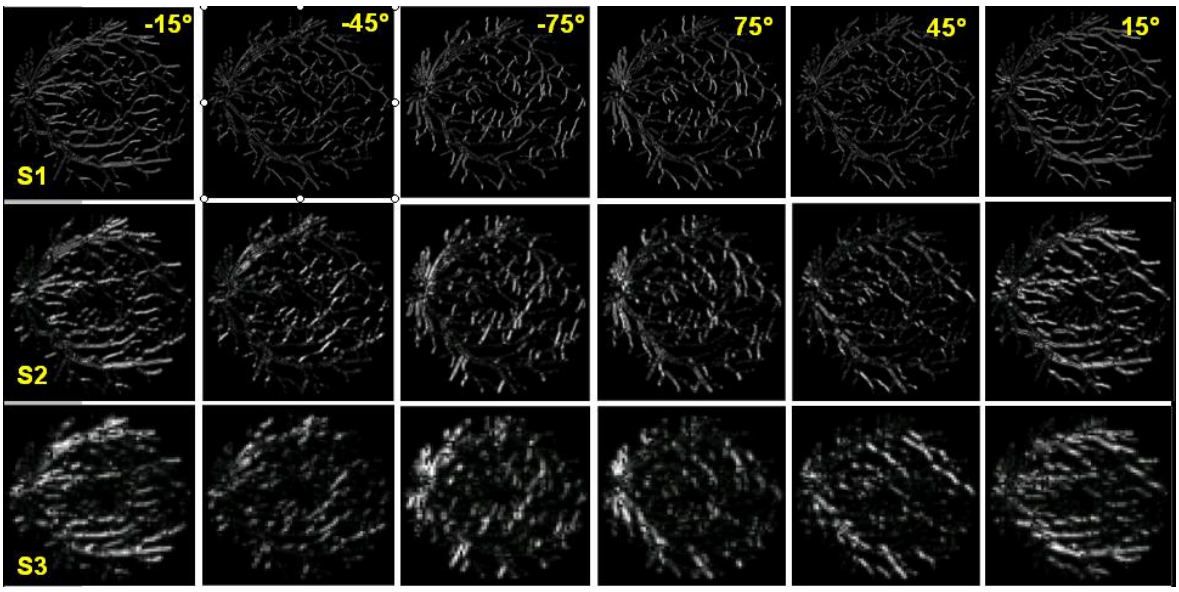


Figure 24 . Magnitude images from the DT-CWT.

The transform is applied to the original image (Figure 20) at 3 levels. The rows show the scales of the transform whilst the columns show the sub-bands.

DT-CWT has proved to provide better results than the real discrete wavelet transform, since the real discrete wavelet transform is not shift invariant [6].

The maximum magnitude over all six orientations is of interest because it is the strongest response and shows that the filter is lined up with the dominant feature. From the maximum magnitude the corresponding orientation can be extracted. This orientation is the orientation of the dominant feature. The maximum magnitude is expressed as $M_{\max}$ and the corresponding orientation is $\theta_{\max}$ shown in equation 7.

$$M_{\max}(x,s) = \max[M(x,\theta,s)]_{\theta} \tag{6}$$

$$\theta_{\max}(x,s) = \arg\max[M(x,\theta,s)]_{\theta} \tag{7}$$

Figure 23 is an example of an orientation map obtained by applying the DT-CWT to a binary version (Figure 22) of a retinogram using Matrix Laboratory (MATLAB) code obtained from Kingsbury et al. [6].

The image is decomposed into three levels; the left image is the first level decomposition which displays the finest details. The vessels are shown using a Hue, Saturation and Value (HSV) colour-map where $M_{\max}$ determines the intensity in the image whilst $\theta_{\max}$ determines the hue as shown on the colour-map displayed on the right of Figure 23. The

orientation with the maximum response is found via equation 7 and then the colour map is used for displaying the orientations with different colours. For example the red vessels have a common orientation. The middle image is the second level decomposition which displays the coarser details and the third image shows the coarsest details with the fine ones less obvious. The orientations of all vessels are displayed.

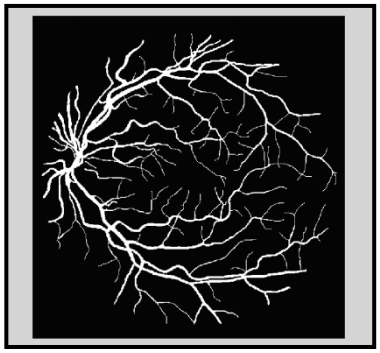


Figure 25 . A binary version of a retinogram also known as the ground truth

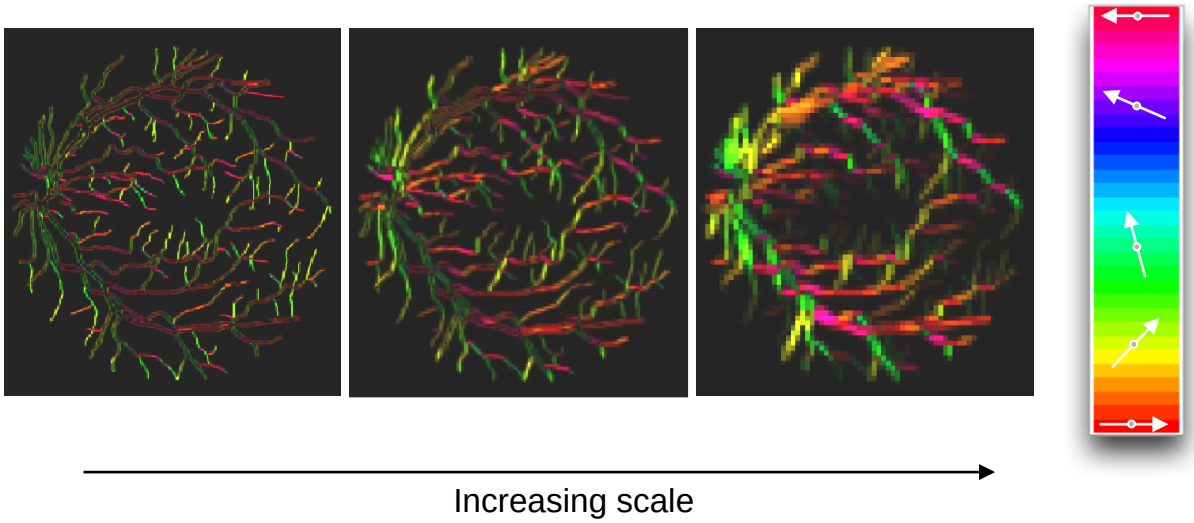


Figure 26 . A three level decomposition of a retinogram.

Each vessel has an orientation which can be obtained from its colour and the colormap.

Figure 24 shows the three levels combined by choosing the largest magnitude response at each pixel over all three scales in order to display both the fine and coarse vessels with their corresponding orientations.

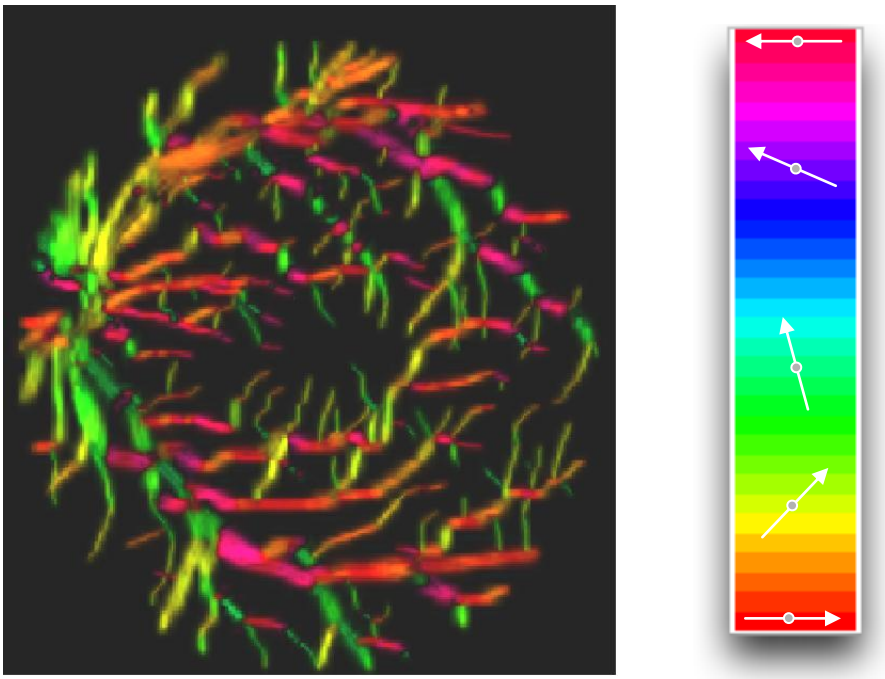


Figure 27 . The three levels of vessels combined together with their corresponding orientations.

3.3.5 The Inter Coefficient Product

The Inter-Coefficient Product (ICP) [58] is a multi-scale representation for 2D images. It consists of a pyramid of complex values, where the phases of its coefficients correspond to the angles of dominant directional features. It is suitable for large dataset analysis since it is relatively simple to calculate. ICP provides for representations of multi-scale features such as edges and ridges [58].

As mentioned earlier, the DT-CWT produces complex wavelets with complex coefficients containing real and imaginary parts in six sub-bands. ICP finds the dominant orientation at each pixel by first finding the sub-band that gives the maximum response and then observing how the phases of adjacent wavelet coefficients change relative to the sub-band direction [60]. The task is to relate the adjacent phases of coefficients to the angles of the desired edges.

The ICP is a constant-phase complex value where its magnitude indicates the strength of a feature and its phase gives the orientation [58].

A more detailed description is given in Appendix 5.

3.4 Characterising Dual Tree Complex Wavelet Responses

This section describes a set of experiments that were carried out in order to characterise the behaviour of the dual-tree complex wavelet transform and to investigate systematically how the transform responds to vessel-like structures of different scales and orientations.

The DT-CWT was applied to synthetic images of simple structures described in section 3.4.1. The experiments were carried out using Interpolation (section 3.4.2).

3.4.1 Synthetic Data

We have created a set of Gaussian bar images. The equation of a Gaussian bar is given below:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} \tag{8}$$

where x is a random variable with mean $\bar{x}$, and σ is the standard deviation which measures the width of the Gaussian. The bars have widths which span a range of σ =1-65 pixels and each of them rotate from θ=0-179 degrees in steps of 1°. Figure 25 shows example images of Gaussian bars which have different widths and are rotated at different angles.

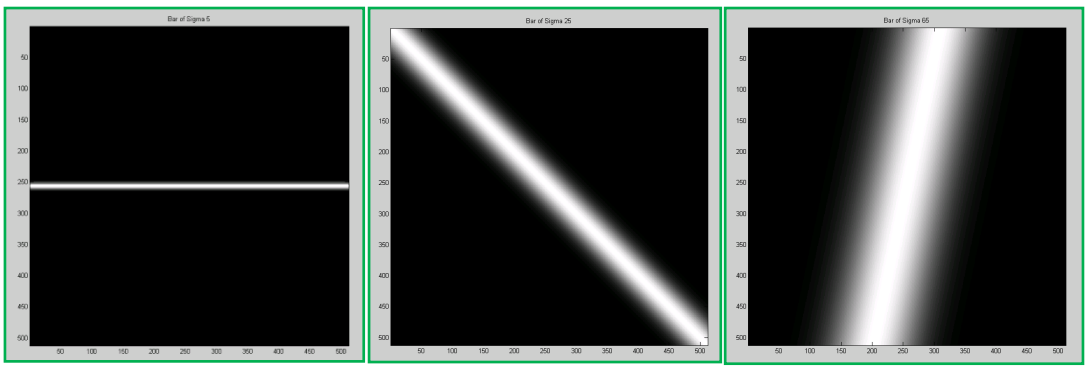


Figure 28 . Gaussian bars.

From left: Bars of width (sigma) 5, 25 & 65 rotated at angles 0, 135 and 65 degrees respectively.

3.4.2 Characterisation Using Interpolation

If the original image is m × n, after applying the dual-tree complex wavelet transform, in level 1 each sub-band has m/2 × n/2 coefficients which are 2 pixels apart in each axis. In level 2 each sub-band has m/4 × n/4 coefficients which are 4 pixels apart in each axis. Therefore, a sub-band in level L has m/2^L × n/2^L coefficients which are 2^L pixels apart in each axis. To sample from the equivalent region at each level, the coefficients need to lie at the same points. Because level 1 has more coefficients than other levels, if for example, the coefficients of level 4 are taken which are m/16 × n/16 and 16 pixels apart in each axis, they need to be compared to the coefficients of level 1 and the gaps between the level 4 coefficients need to be filled. This is called interpolation where each sample is obtained from its neighbouring points. So, at level 4, from the level 4 coefficients that are 16 pixels apart in each axis, the sampling points that correspond to the coefficients at level 1 which are 2 pixels apart in each axis need to be calculated.

If the coefficients were just real numbers, this would be a simple 2D interpolation of the level 4 coefficients. However, the coefficients are complex numbers with dependencies between the real and imaginary parts so the following steps are carried out (Kingsbury et al. [59]):

- Perform phase unwrapping to remove the dependencies between the real and imaginary parts of the level 4 coefficients, which we can now separate – so there are 2 sets of m/16 x m/16 numbers (that are both still spaced 16x16 pixels apart with respect to the image)
- Perform a standard 2D interpolation on each of these sets of numbers separately. So now there are m/2 x n/2 numbers in each set, and the spacing is 2x2 pixels apart. The 2 sets of interpolated numbers can then be combined back into complex numbers
- Rewrap the phase of the new complex numbers to undo the effects of step 1), resulting in a final set of m/2 x n/2 interpolated coefficients for level 4.

Now there are a set of level 4 coefficients the same size and spacing (i.e. with sample points at the same locations with respect to the original image) as the level 1 coefficients, so they can be compared directly. The same process is applied to every level.

In more detail, band pass interpolation is chosen since the information contained in a given directional complex sub-band and level is band-limited to a particular region of the 2D frequency space, which has a centre frequency $\{\omega_1, \omega_2\}$. Band pass interpolation can be performed by:

- Shift by $\{-\omega_1, -\omega_2\}$ down to zero frequency (i.e. multiply by $e^{-j(\omega_1 x_1 + \omega_2 x_2)}$ at each point $\{x_1, x_2\}$);
- Low pass interpolate to each new point (spline/bi-cubic/bi-linear);
- Shift up by $\{\omega_1, \omega_2\}$ (multiply by $e^{j(\omega_1 y_1 + \omega_2 y_2)}$ at each new point $\{y_1, y_2\}$).

Once interpolation has been carried out samples can be taken directly from the centres of each bar.

The DT-CWT was applied to the synthetic Gaussian bar images described in section 3.4.1 and interpolation was used for sampling. The maximum response at each level and for each sub-band have been extracted with their corresponding bar width and angles. Table 25 (Appendix 1) shows the recorded values.

Figure 26 shows the response of the dual-tree filters with bar angle when applied to the Gaussian bar images containing interpolated samples. The response is shown for the optimal width and level 6 of the transform.

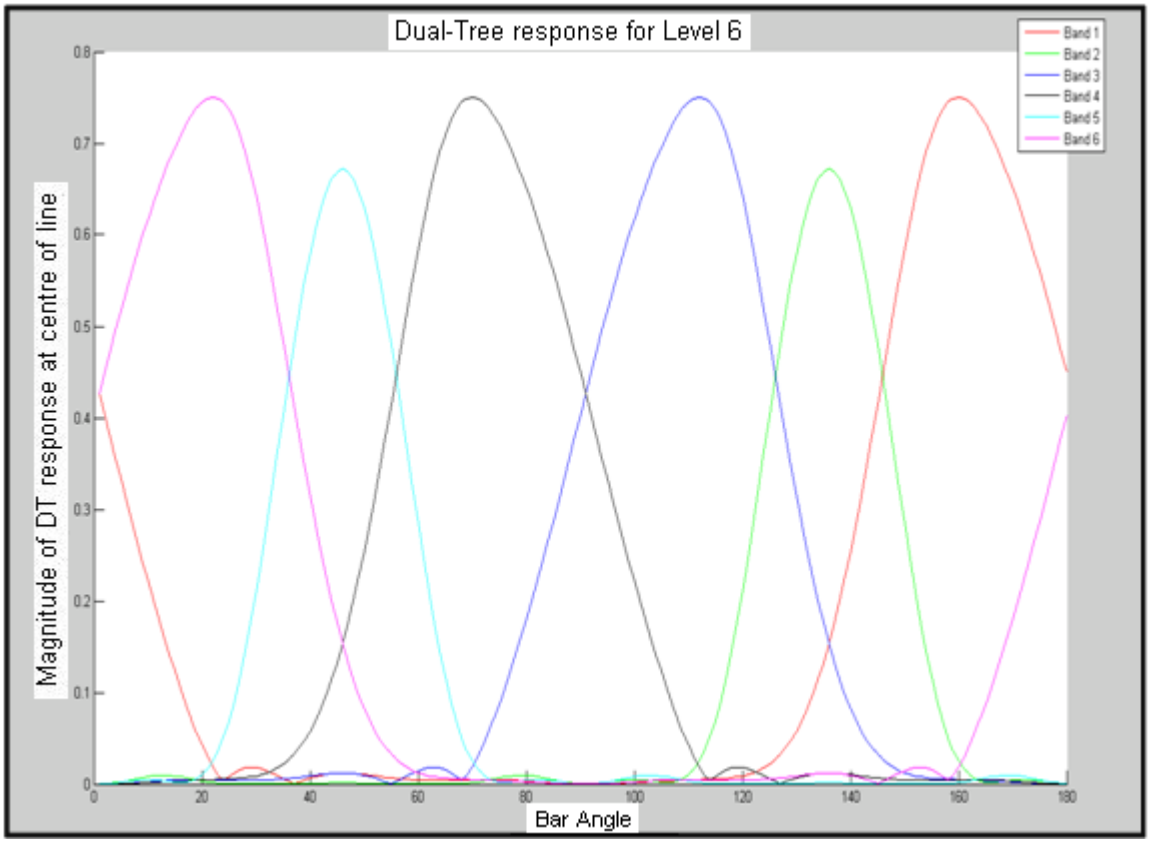


Figure 29 . Response of the dual-tree filters with angle (level 6, optimal width).

The curves show the magnitude response of each sub-band and what bar angles they can detect. Bands 2 (green) and 5 (light blue) have lower responses and are symmetrical whilst the other bands are asymmetrical.

As can be seen from Figure 26, optimal responses are achieved at angles 21 for sub-band 6, angle 45 for sub-band 5, angle 69 for sub-band 4, angle 111 for sub-band 3, angle 135 for sub-band 2 and angle 159 for sub-band 1. (Table 25- Appendix 1). Sub-bands 2 and 5 have lower responses than the other sub-bands and detect slightly smaller objects.

The responses as a function of bar-width have been plotted in order to observe the relationship between levels, bar widths and sub-bands.

Figure 27 shows the response of the dual-tree filters with bar angle when applied to the Gaussian bar images containing interpolated samples. The response is shown for sub-band 6 and the optimal angle (which in this case would be 21º - refer to Table 25).

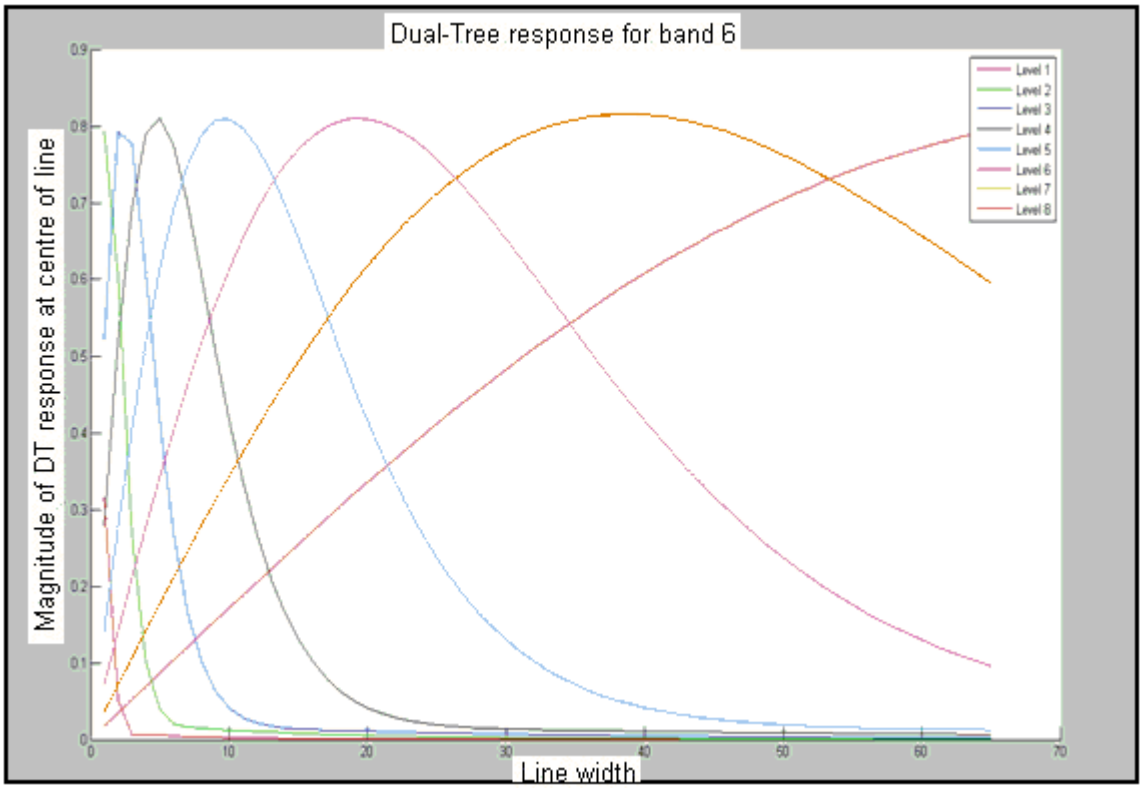


Figure 30 . Response of the dual-tree filters with width (sub-band 6, optimal angle).

The curves show the magnitude response of each level and the bar widths each can detect.

Each level detects objects that are twice the size of the objects in the previous level. In this graph, the curves for 8 levels are shown.

There is different scale tuning of filters at angles 45º and 135º vs other angles for this reason bands 2 & 5 detect slightly smaller bars (Figure 28).

Referring to Table 1 (Appendix A), only the level 1 angles are different and the level 2 angles are equivalent to the angles in levels 3, 4, 5 & 6

Level one always shows slightly different responses and angles which is due to the fact that the filters at that level are different (Figure 27).

3.5 Summary

The various types of texture models have been described and compared with each other. Most methods are only rotation invariant and some rotation and scale invariant but none of them are shift dependant.

Wavelet transforms characterise textures at different scales. The discrete wavelet transform is shift dependant and gives limited information on orientation hence complex wavelets are more advantageous since they are shift invariant. The wavelet transforms used in this project are the dual-tree complex wavelet transforms which are shift, rotation and scale invariant therefore more information can be extracted at a pixel level hence a better representation is achieved. This information is useful in classifying the pixels of an image.

The DT-CWT is therefore chosen as the best local structure representation method to be applied to retinograms in order to classify vessel and non-vessel pixels. The next chapter explains the classification process and how the information obtained from applying the DT-CWT can be used to predict the class of pixels in retinograms.

CHAPTER 4: Machine Learning

4.1 Introduction

Machine learning is a process in which a machine learns from past experience in order to extract useful information and then to apply that information for prediction and planning with new data [60]. During the learning phase, the machine is trained to learn the best decision criteria and to make full use of background knowledge in order to efficiently predict unseen data. A model with a number of parameters is defined and at the learning stage, the best parameters for that model are identified for predicting. The more capability the machine has to effectively use the training data to generalize to new examples, the better the performance of the machine.

There are three types of machine learning:

1. Supervised learning, which is when mapping of the input data to the output is done by a supervisor who supplies the correct values. So the labels associated to the outputs are known [61].
2. Un-supervised learning, also often referred to as clustering which is when there is no supervisor and only the input data is known, therefore unlabelled examples are used for learning. The structure of the data is not known in advance.
3. Reinforcement learning, where the supervisor provides training data to the system in a different form. The form is a scalar that gives a measure of the system performance [61].

This chapter gives an explanation of machine learning techniques that can be used to model local structure. These techniques are the most common used for machine learning. They can deal with non-linear models.

Classification which is a supervised learning method and is used to characterise whether pixels in an image belong to one class or another is explained and used in this project. Linear [62] and Quadratic classifiers [62] can be used for classification. Non-linear classifiers are more suitable for non-linear data and have been compared and described in detail below.

4.2 Classification

Classification is the process of classifying a set of objects into different groups based on their attributes. The attributes are called features and are assembled together into a feature vector **x**. Each object is assigned to a class y based on some chosen decision or forecast. A small proportion of the whole data is taken as training data and used to construct a classifier that will assign each to the correct class. The classifier is then ready to predict the class of the remaining feature vectors of the test data and to assign them to appropriate classes. During training a decision boundary is constructed in a way which separates the feature vectors such that the vectors of the same category lie on one side of the boundary and the vectors with the other category lie on the other side of it [60].

4.2.1 Example

If patients that want to know whether they have a high risk of diabetic retinopathy are considered, their general health, age, images of their eye and whether symptoms such as abnormalities and blood vessel growth are present are combined to form training data and based on a general rule that determines the association between these patient attributes and his/her risk, the patients are divided into two groups (classes): high-risk and low-risk patients. The general rule could be for instance:

IF age> θ_1 AND number of blood vessels> θ_2 THEN high-risk ELSE low-risk

Figure 28 shows a hypothetical training set where the patients have been divided into low-risk and high-risk patients based on their age and number of blood vessels in their eye. The negative circles indicate low-risk patients whilst the positive circles indicate high-risk patients [60].

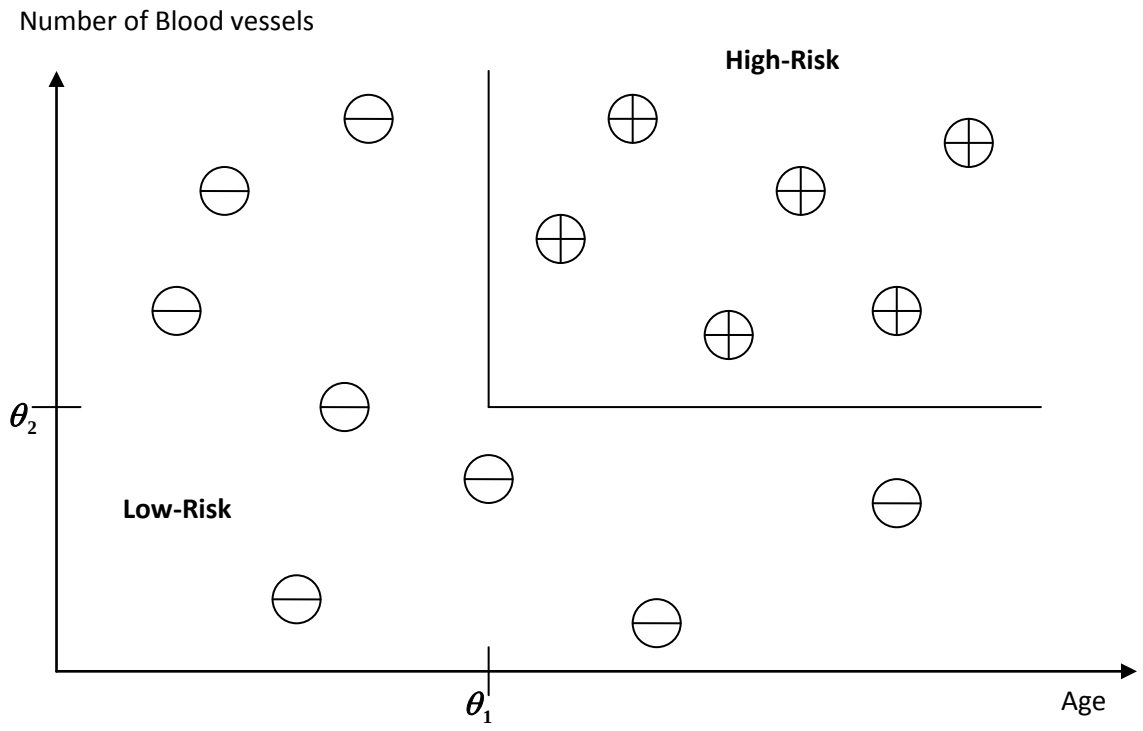


Figure 31 . A training data set used for classification [60].

The circles represent data instances (patients) and their signs indicate their class. The inputs are the patients age and number of blood vessels in the eye. θ_1&θ_2 are limits from which a decision boundary is constructed. A new data instance (patient) is assigned to the low-risk or high-risk group based on a general rule.

4.2.2 Feature Selection

The classifier's ability to generalize well is affected by choosing the appropriate features. The appropriate features of the same data will vary depending on the classification task. It is computationally inefficient to include all possible features. Feature selection algorithms search through the subsets of features and attempt to find the best one amongst all subsets according to some evaluation function. This method is costly and practically impossible when trying to find the best feature. Other methods attempt to reduce the computational complexity which compromises performance. For example, in the Heuristic feature selection method, generation of subsets is incremental (either increasing or decreasing). In the random feature selection approach, less number of subsets is searched for by setting a maximum number of iterations possible.

4.3 Artificial Neural Networks

An Artificial neural network (ANN) is an effective way of processing and representing information. It is used to derive patterns and complex relations between data which is difficult for humans to see [63].

Artificial neurons consist of inputs (synapses) which are multiplied by weights that correspond to the strength of signals. A mathematical function then determines the activation of the neuron. A high weight will increase the strength of the input it is multiplied by. The neurons are connected to each other via synaptic connections which transfer data.

Artificial neurons are organised in layers. The artificial network contains an input layer where the inputs are received by the neurons and an output layer which contains the outputs from neurons. It also could contain one or more intermediate hidden layers. The hidden layers provide for representations for the inputs.

Each neural network has a different interconnection pattern between its layers, a different learning process which updates the weights of the interconnections and a different activation function which converts a neuron's weighted input to its output activation [64]. Figure 29 shows an example of a three layer model.

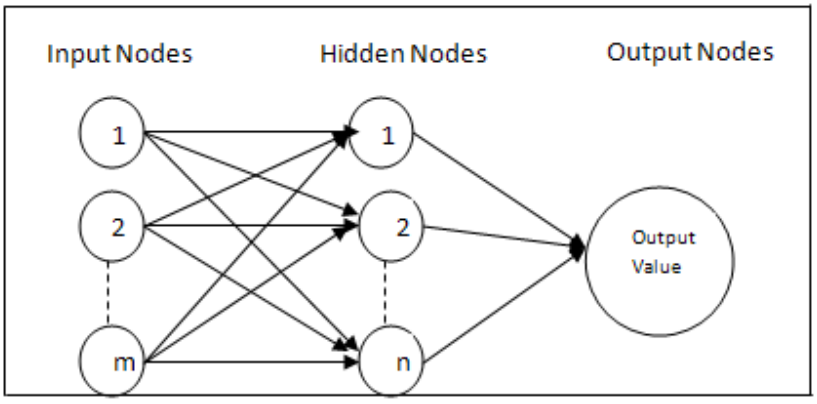


Figure 32 . A three layer neural network model [65].

ANNs have the advantage of detecting complex relationships between variables. They are relatively simple to train and can handle high-dimensional non-linear data. But they are computationally expensive and the training is very slow. Also, if the ANN is complex, it is

not easy to identify the decision criteria. There is a tendency to overfit the training data. The model is a black box since the connection of the nodes are not developed analytically [65].

4.4 K nearest neighbour networks

K-nearest neighbour network (KNN) is a classification method where the class of a sample is identified based on the class of its k nearest neighbours.

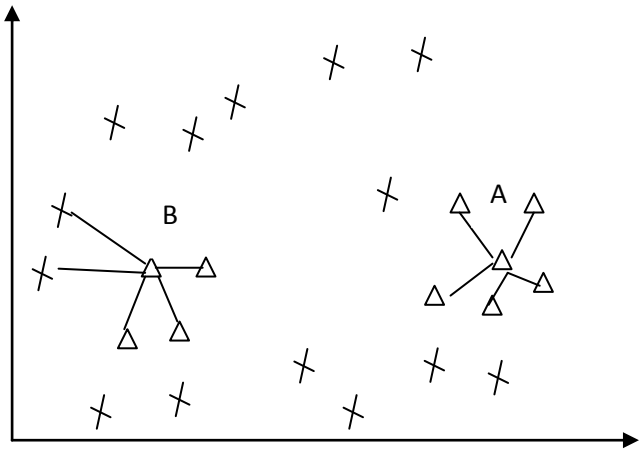


Figure 33 . An example of a 5 nearest neighbour classifier.

Figure 30 illustrates the process in a 2 dimensional feature space. There are two classes: triangle and cross. The task is for the classifier to determine the class of A and B.

In case A, all 5 neighbours have the same class, therefore A will take the same class as that of its neighbours therefore it takes the form of a triangle as shown in Figure 30. In case B, there are 3 triangles and 2 crosses, therefore the class of B will be determined either based on the majority class of its neighbours or based on a distance voting algorithm where more weight is assigned to the nearest neighbours and the class is determined by the summation of the inverse distances of each neighbour [66].

It is advantageous because of its easy implementation. Also, it is less sensitive to noise since there are techniques designed to reduce noise sensitivity for KNN networks. The

disadvantages are that it requires very large training sets and is also very sensitive to irrelevant features [66].

4.5 Support Vector Machines

Support vector machines (SVMs) [67] are a family of learning algorithms for classifying objects into two or more classes. When new data is fed to the classifier, the class of each feature vector is predicted.

To introduce the basic ideas we deal first with the case of data that is linearly separable- that is, a hyper-plane y(**x**) = **w.x** + b = 0 exists such that,

$$y(\mathbf{x}) = \text{sign}(\mathbf{w.x} + b) \tag{9}$$

$$\mathbf{w.x} + b > 0 \text{ if } y = +1 \tag{10}$$

$$\mathbf{w.x} + b < 0 \text{ if } y = -1 \tag{11}$$

where **x** is a n dimensional feature vector and the sign of the decision function indicates which label should be assigned to each vector. A value of +1 indicates one class and a value of -1 indicates the other class, b is the intercept term and $\vec{w}$ is the decision hyper plane normal vector commonly referred to as the weight vector.

This is shown in Figure 31, where squares and triangles represent different feature vectors and their symbol indicates the class. $\mathbf{x} = (x_1, x_2)$ is the input of the classifier from which the feature vectors are formed.

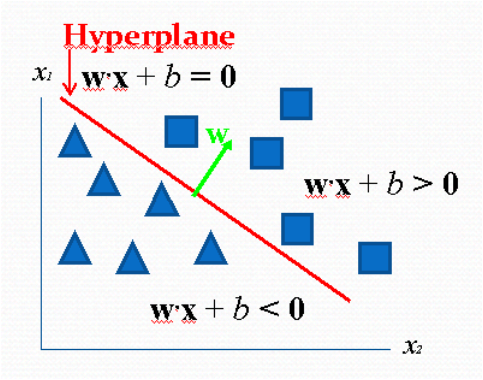


Figure 34 . A support vector machine constructs a hyper-plane (red line) to separate the data vectors of different categories.

The squares are in one category and have positive labels as they lie above the hyper-plane whilst the triangles are in a separate category and have negative labels therefore lie below the hyper-plane.

In order to classify the vectors efficiently, the task is to determine the best hyper-plane. As can be seen in Figure 32, there are numerous hyper-planes that can be constructed for separating the data.

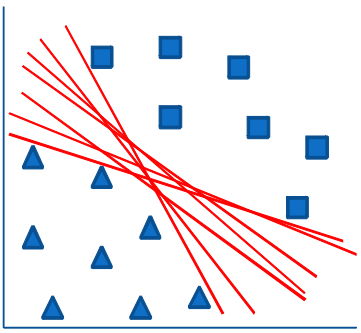


Figure 35 . The hyper-plane that separates the data vectors can have many widths and orientations.

Choosing the right hyper-plane to classify data is important.

The solution according to Vapnik [68, 69] is to obtain the largest margin, which means the hyper-plane with maximum width between the two sets of data vectors– see Figure 33. This is due to the fact that a large margin is better for the generalization ability of the SVM – since the distance between the two sets has been maximized misclassification is unlikely to occur.

Figure 33 shows Vapnik's margin for the problem above.

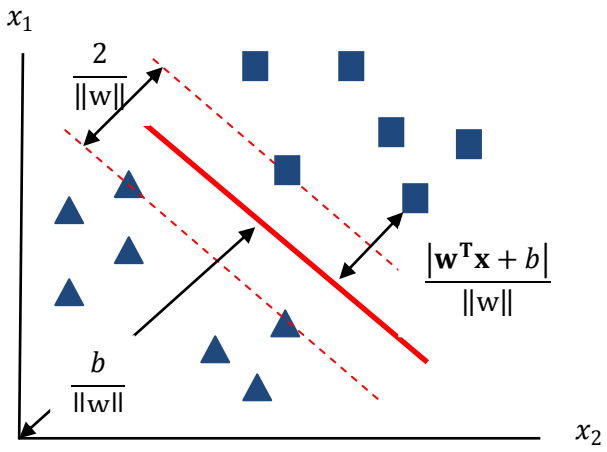


Figure 36 . The best hyper-plane is the one with the maximum width between the different classes of data according to Vapnik.

This is due to the fact that the larger the width, the less flexible the margin is in the space and therefore the less complex the classifier.

From trigonometry, the distance between a point x and a plane ($\mathbf{w}$, b) is

$$\frac{\mathbf{w}^T\mathbf{x}+b}{\|w\|} \tag{12}$$

The solution is chosen for which the discriminant function becomes one for the training examples closest to the boundary

$|\mathbf{w}^T\mathbf{x_i} + b| = 1$ where $\mathbf{x_i}$ is the example closest to the boundary.

Therefore, the distance from the closest example to the boundary is

$$\frac{\mathbf{w}^T\mathbf{x}+b}{\|w\|} = \frac{1}{\|w\|} \tag{13}$$

And therefore the margin becomes:

$$m = \frac{2}{\|w\|} \tag{14}$$

where m is the width of the margin. This is demonstrated in Figure 33.

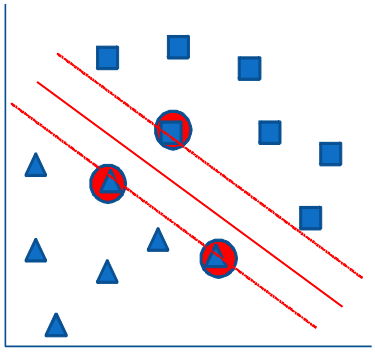


Figure 37 . The support vectors are shown as red circles.

They are the data vectors that lie on the margin and separate the two classes from each other. The more support vectors, the better the classification.

In Figure 34, the support vectors are shown as the red circles that lie on the margin and separate the data vectors. The optimization problem can be solved using standard methods from quadratic programming. Once the problem is solved, the support vector machines can be identified and the data can be classified.

4.5.1 Non linear SVM

Non-linear data is shown in Figure 35, where a linear hyper-plane cannot separate the two classes.

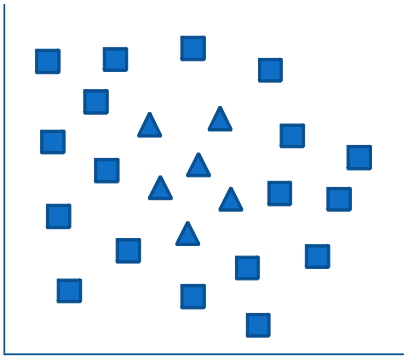


Figure 38 . Non-linear data is shown where a hyper-plane cannot classify it.

The solution in such cases is to map the non-linear data from the low-dimensional feature space to a higher-dimensional feature space in which it can be separated by a hyper-plane. This is demonstrated in Figure 36.

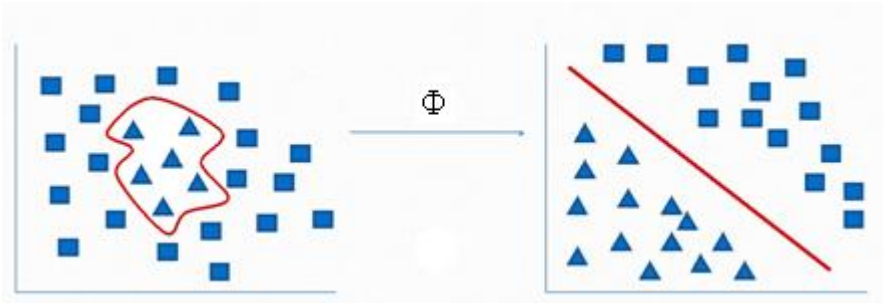


Figure 39 . To classify non-linear data and for simplicity, the data is transferred to an N-dimensional feature space where it becomes linear and a hyper-plane can then classify it.

In the training problem the data points appear in the form of inner products $\mathbf{x}_i.\mathbf{x}_j$ which is a measure of distance between them. In the testing phase they appear as $\mathbf{x}_i.\mathbf{x}$ where $\mathbf{x}$ is a new feature vector.

For a given mapping Φ from the space of objects $\mathbf{x}$ to a feature space, the inner product of two objects $\mathbf{x}$ and $\mathbf{x}'$ cannot in general be calculated since the feature space is too large, but it can be replaced by a kernel.

$$\forall \mathbf{x},\mathbf{x}' \in X, K(\mathbf{x},\mathbf{x}') = \Phi(\mathbf{x}).\Phi(\mathbf{x}') \tag{15}$$

There are many kernels that can be used for this purpose. The following are the most common:

Gaussian Radial Basis Function

$$K(\mathbf{x},\mathbf{x}') = -\exp\left(\frac{\left\| \mathbf{x}-\mathbf{x}' \right\|^2}{2\sigma^2}\right) \tag{16}$$

Polynomial

$$K(\mathbf{x},\mathbf{x}') = (\mathbf{x}.\mathbf{x}' + 1)^d \tag{17}$$

Sigmoid

$$K(\mathbf{x},\mathbf{x}') = \tanh(k\mathbf{x}.\mathbf{x}' + \theta) \tag{18}$$

Figures 37-40 show examples of classifying data using Gaussian and Polynomial kernels. The data was provided in the "SVM and Kernel Methods" MATLAB toolbox.

Figure 37 shows the untransformed data before applying a classifier.

Figure 38 shows a polynomial kernel classifying the data into two groups where one class of the data is represented as green circles and the other half by red circles.

Figure 39 shows the data being classified into two groups by a Gaussian kernel with σ =0.5 and C=1. The red region indicates one group while the blue region indicates another group. The circles shown are the support vectors. Figure 40 shows the same data classified by a Gaussian kernel with σ =1 and C=10.

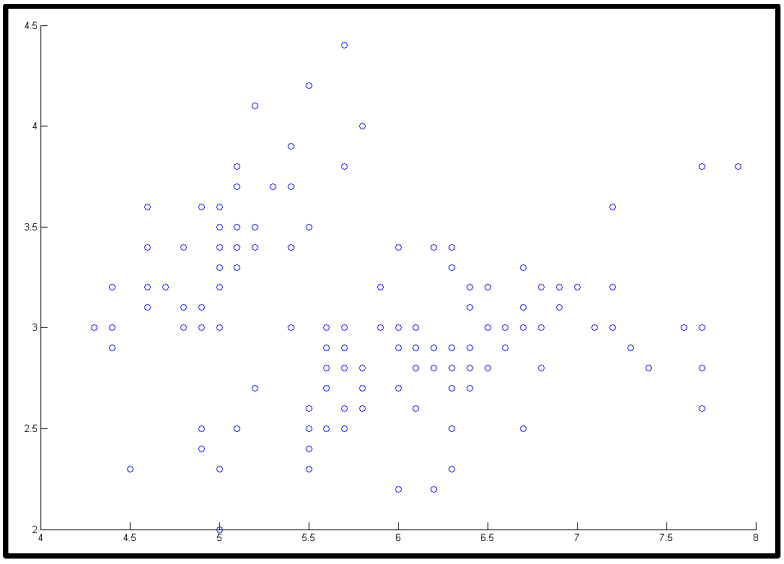


Figure 40 . Untransformed data before applying a classifier.

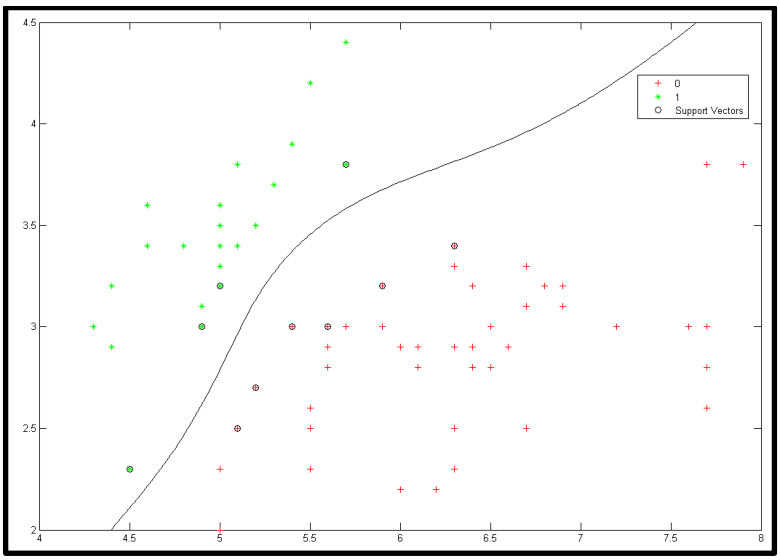


Figure 41 . A Polynomial kernel with d=1 has been used to classify the data.

The support vectors are shown as circles. The red dots indicate one class whilst the green dots indicate another class.

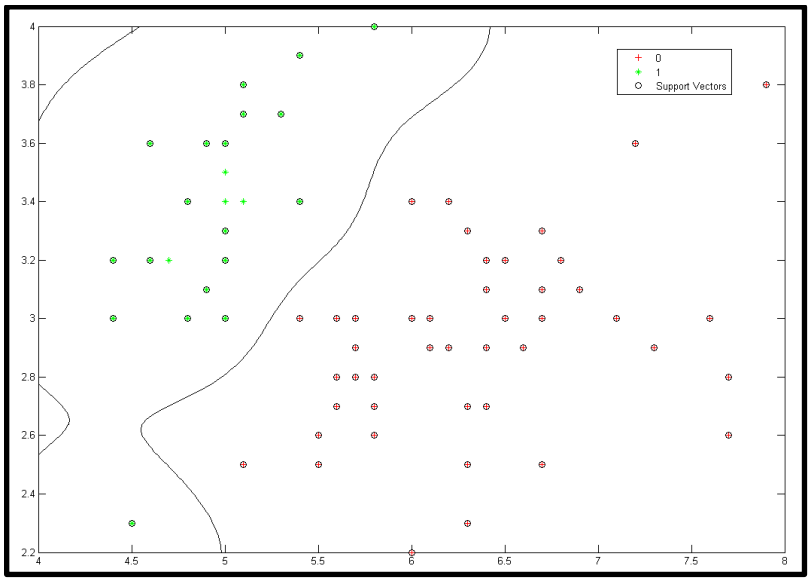


Figure 42 . A Gaussian kernel with σ=0.5 and C=1 has been used to classify the data.

The support vectors are shown as circles. The red dots indicate one class whilst the green dots indicate another class.

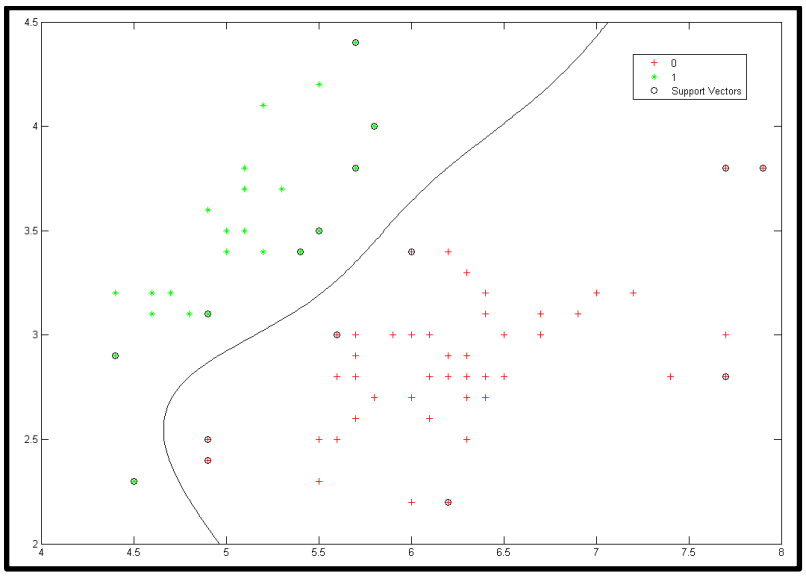


Figure 43 . A Gaussian kernel with σ=1 and c=10 has been used to classify the data.

The support vectors are shown as circles. The red dots indicate one class whilst the green dots indicate another class.

The advantages of SVMs are that it scales relatively well to high dimensional data and has good generalization performance. Data capacity is controlled by optimizing the margin. It can find the optimal separation plane. The disadvantages are that there is a need to choose a "good" kernel function and find the best parameters therefore training is time consuming. They are limited in speed and size for training and testing. They are sensitive to noisy data and hence prone to over fitting. They are also not convenient for discrete data.

4.6 Decision Trees

A decision tree [70] uses training data in order to construct a model that can predict a target value based on a set of binary rules. An example of a decision tree is given in Figure 41.

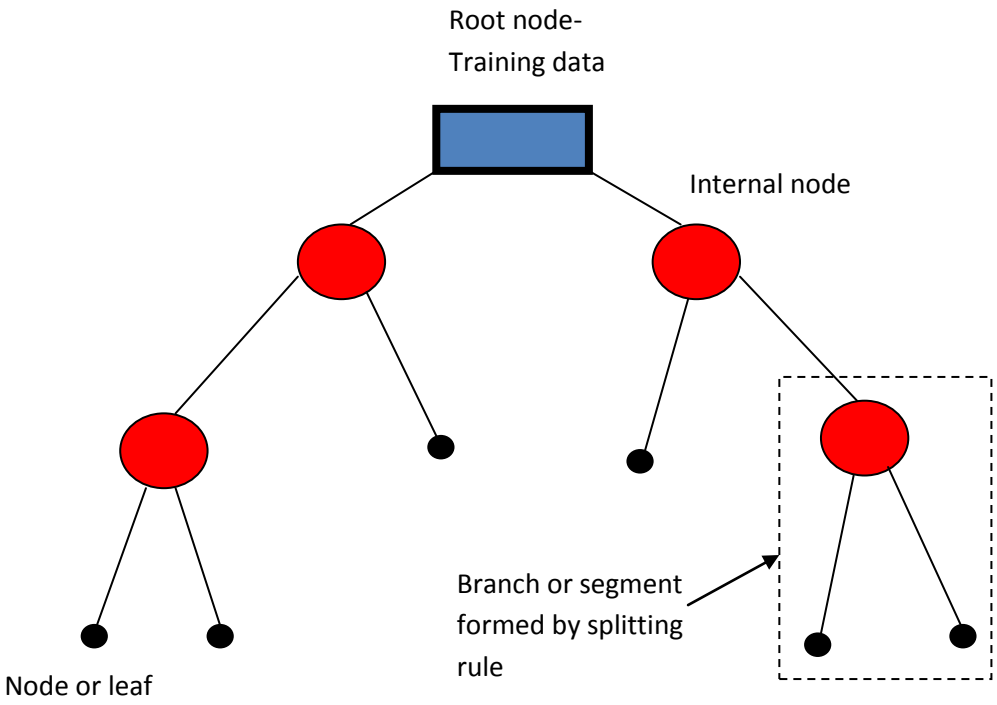


Figure 44 . A decision tree diagram.

The blue box shows the initial training data, the red circles show the internal nodes and the black circles show the leaf or terminal nodes. At each node the variable that is used to split the data is determined along with the value of the split. The lines show the outcome of the decisions made at each internal node and the terminal nodes are assigned to each class once the decision to stop and not split again has been made.

Decision trees are easy to interpret the decision rules within them. They provide for a means to visually examine and analyse the relationships between the input and target values. The decision rules can predict the values of unseen observations.

4.7 Random Forest

4.7.1 Introduction

Random forest classifier [5] is an approach that is well-suited to non-linear classification in a high-dimensional space. A random forest comprises a set of tree predictors (decision trees). Below gives an explanation of how they work.

4.7.2 Random forest classification

The training data consists of N samples each of which is a D-dimensional feature vector labelled as belonging to one of C classes. Bootstrap sampling is applied to the training data where a set of N samples are chosen randomly, with replacement, from the original data. Then a tree in the forest is built based on a bootstrap sample. The trees are built using a standard classification and regression tree (CART) algorithm. CART is a model that describes the conditional distribution of y given x. If the response y is discrete, the model is a classification tree and if it is continuous, it is a regression tree.

The binary trees partition the predictor space into distinct homogenous regions via different nodes. At each node, rather than assessing all D dimensions for the optimal split, only a random subset of $d < D$ dimensions are considered. Splitting rules associated with the internal nodes are determined for the partitioning of data. The trees are built to full size (i.e. until a leaf is reached containing samples from only one class) and are not pruned.

Each tree in the forest casts a unit class vote, and unseen feature vectors are classified independently by each tree. Hence the most popular class is assigned to the input vector. Alternatively, the proportion of votes assigned to each class can be used to provide a probabilistic labelling of the input vector [5].

Below provides a step by step algorithm:

1. N bootstrap samples are drawn from the original dataset.
2. For each bootstrap sample, an unpruned classification or regression tree is grown where at each node, a random sample of predictors is taken and the best split is chosen among the variables.

3. N tree predictions are combined and used for prediction of new data [5].

Figure 42 shows a random forest classifier. The training set can have the same size as the original data set. Many trees are built and each tree votes for the class of the data. The trees are different from each other by giving each one different expertise and decision rules and a different training set and also a different subset of features which are chosen at random.

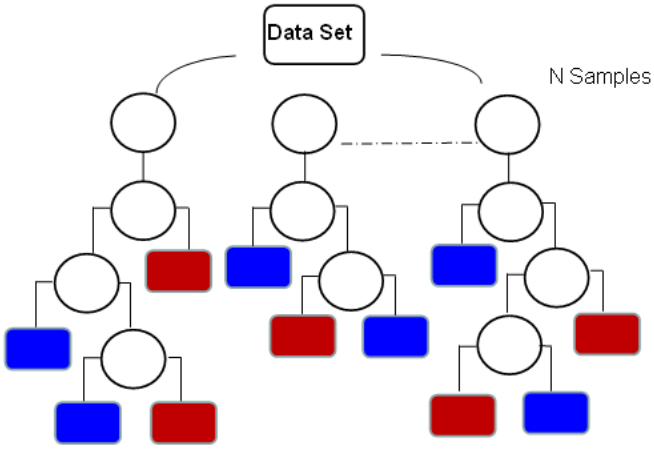


Figure 45 . A random forest classifier with three trees.

Each tree votes for the class of the data. The red boxes show one class and the blue ones show the other class. N samples which are the top circles are drawn at random from the data set. The data is then classified into classes at each node which are the bottom rectangles.

The parameters of the random forest can be tuned in order to get the best results. These parameters are the number of trees, the training dataset sample size for building each decision tree and the number of random variables which determine the partitioning of the training dataset at each node.

4.7.3 Out of Bag

Out of Bag (OoB) refers to the fact that bootstrap sampling will result in roughly one third of the training data not being included in the sample for the construction of the k^{th} tree. OoB samples are used to obtain an unbiased estimate of the classification error and variable importance as trees are added to the forest. Therefore cross validation is not

required. The Out of bag error estimate is calculated by initially obtaining a classification for each case after running the OoB data down the k^{th} tree. Therefore in one third of the trees, a test set classification is produced for each case. Every time case n is OoB, j is the class that got most votes. When j is not equal to the case of n, it is counted and accumulated and averaged over all cases [5].

4.7.4 Advantages and Disadvantages

Random forest classifiers provide for high classification accuracy and good generalisation ability for non-linear data. They are adaptable to large data sets and maintain efficiency and can handle a large number of input variables. They can also maintain their accuracy when data is missing or corrupted. The randomness of data and variable selection causes computational efficiency. They give useful internal estimates of error and variable importance. Parameter selection is easy and the training process is fast. They also have the ability to handle large numbers of variables whilst maintaining efficiency. The data does not need to be transformed, rescaled or modified [5]. Random forests show to overfit noisy data. Regression can't predict beyond range in the training data. The more trees are added, the more time it takes for training and testing [5].

4.7.5 Applications

Random forests have various applications. Below are a few examples:

They are considered for classification of multisource remote sensing and geographic data [71]. They are commonly used for land cover classification [72], gene selection and microarrays [73], classification in ecology [74], language modelling [75], and verbal autopsy analysis [76].

4.8 Comparison between classification methods

The above machine learning techniques are compared below.

For ANN networks, choosing the network topology, learning rule, performance function, training phase and stopping criteria makes it difficult to refine the solution.

KNN classifiers are slow because they need access to all the training data for classification. Also, they are computationally expensive and require large memory because at each prediction, the algorithm computes the distance between the k closest neighbours and the new instance from the training data and sorts all the training data at each prediction.

From the classification methods described, SVMs and Random forests have been chosen for this project. SVM has a straightforward classification algorithm which is adaptable to all kinds of data therefore has very good generalisation performance. Random forests provide for high accuracy which is the main target required in choosing a good classifier.

Random forests are capable of working with large data sets. SVMs cannot handle large training sets and therefore they take far more time to train and test compared to the random forests. Also parameter selection is difficult. Random forests have very good computational scalability in comparison to SVMs.

Based on the above, random forests and SVMs have been used as the main classification methods in this project.

4.9 Summary

In this chapter, classification has been described with several approaches. A qualitative comparison has been made between these methods and the best ones have been chosen to be used in this project. Support vector machines and random forests are the best two classifiers amongst the others described previously. Random forests are more applicable for this project since they allow for large training sets and the training and testing time are much faster. Hence they are more computationally efficient. Also the only parameters required for training are the number of trees and therefore less parameter tuning is required.

In the next chapter we have applied both classifiers to retinograms and compared their performance on the same data.

CHAPTER 5: Computer-aided Retinography

5.1 Introduction

Many techniques have been developed that detect vessels in medical images. In this chapter a literature review on these methods and a comparison between them has been carried out. They have been divided into two groups: Low-level methods and Intermediate-level methods. Low-level methods operate on a pixel by pixel basis and are described in section 5.2. Intermediate-level methods use prior knowledge of structure to extract whole vessels or vessel trees which are described in section 5.3. Some of these methods have been applied to the databases described in chapter 2. These are highlighted in green.

This is followed by a discussion on evaluating vessel detection methods including a summary of published results for the DRIVE database (Table 1). The chapter concludes, for completeness, with a summary of other work on analysing retinograms (Table 2).

The below methods have been chosen based on the fact that they are the most commonly used techniques as shown in the tables of Appendix 2 and 3.

5.2 Low level analysis

Low-level methods detect local structure by analysing the image around each pixel.. A drawback of this pixel classification approach is that because they use very lowly level information and don't have access to global properties, they can be more prone to fail with unexpected patterns of data values. Several types of common low-level methods have been applied for the detection of blood vessels. The following section describes each one.

5.2.1 Classification based Methods

The purpose of classification methods is to classify data into suitable classes by using local features obtained from the training data. At the training phase, the algorithm learns to

correctly classify the pixels in the training images and at the test phase the algorithm is able to classify pixels in unseen images. For example, vessels or non-vessel classes can be assigned to pixels in a retinal image. Therefore using several vessel properties as possible at the training phase can improve results. Multi-feature vectors can be calculated at each pixel and the image can therefore be transformed into the feature space hence the pixels are classified based on their positions in the feature space.

There are two types of classifications: hard classification and soft classification. Hard classification is used to assign categories such as vessel/non -vessel labels to each pixel whereas soft classification is used to assign probabilities.

There are several types of classifiers such as support vector machines, random forests, KNN classifiers, linear [62] and quadratic [62] classifiers. Below describes a few approaches that use the above classification methods. A detailed description of classification methods is given in chapter 4.

5.2.1.1 Methods using support vector machines

Ricci et al [77] use line operators [78] to detect blood vessels in retinograms of the DRIVE and STARE databases using an unsupervised binary pixel classification method. Lines of fixed length passing through a target pixel at different orientations are used in order to obtain the average intensity. 12 angles are considered and a 15º of angular resolution. The average intensity is computed for each line and the line with the largest average intensity is then found and for each pixel, the line strength is evaluated by subtracting the largest average intensity from the average intensity of the square window containing all lines. Therefore the line strength will be large if the winning line is aligned with a vessel and low if it is not.

When this method is applied to retinograms, because of the very bright pixels surrounding the field of view (FOV), the line strength of the border pixels is biased by this region therefore to overcome this problem, a neighbourhood region around the target pixel is considered and the average intensity pixel of this region is used for the target pixel. The above basic line detector is used to obtain the line strength as one of the features that are fed to a linear SVM classifier. The pixels close to large and bright vessels are comparable to the line strength of pixels inside darker and thinner vessels. To overcome this problem, a

line of 3 pixels is centred on the midpoint of the main line and orthogonal to it. It is 3 pixels so that the centre pixel will be placed on the midpoint and each side of the point will have a pixel belonging to the line which would make the line have minimum length. The average intensity of this line is denoted with $L_o(i,j)$. Its strength is found by subtracting the average intensity in the 15×15 square from its average intensity $S_0(i,j) = L_0(i,j) - N(i,j)$. It will be a 15×15 square so that the line of length 15 will fit within this window. The strength of this line will be high for pixels inside vessels and very small for the pixels outside vessels. This is used as the second feature. For the third feature, the intensity of each pixel is considered and fed to the classifier. 1000 pixels from each training image were sampled to train the classifier before it was applied to all 20 test images. Experiments were carried out with both linear and nonlinear SVMs. Nonlinear SVMs presented similar results which have not been demonstrated in this paper. Also the linear SVM is less computationally demanding [77]. The advantages of this method are the simplicity of the line detector, the extraction of features is computationally simple and only a few features are needed for training. Also robust detection is achieved in the presence of vessels with different size. Good edge localization is achieved and from a visual inspection, vessels have a few gaps. The disadvantages are that false positives are found around the border of the optic disk and thin vessels with low contrast are not detected. They are only detected with the expense of a high false positive rate which introduces a lot of noise.

5.2.1.2 Methods using neural networks

Sinthanayothin et al. [79] apply a neural network (NN) to classify each pixel of a retinogram into vessel or non-vessel class. The edge strength and first principal component of a sub image of 10 × 10 pixels localised on the pixel being classified is taken as the inputs of the NN. They can consist of 20 hidden nodes, two output nodes and 200 input nodes as a three layer perceptron. This method is applied to 112 retinal images. The size of each retinogram is 570× 550. For each retinogram, 73 sample patches of size 20 × 20 pixels are taken from the image and the sensitivity and specificity are computed. The overall sensitivity (fraction of pixels correctly classified as vessel pixels) of detecting blood vessels is 80% ± 17% and the specificity (fraction of pixels incorrectly classified as vessel pixels) is 91% ± 5% The pre-processing and post-processing of the above method

reduces the reliance upon the NN and improves computational efficiency. The fine vessels are not detected.

5.2.1.3 Methods using KNN classifier

Niemeijer et al. [2] compare the performance of five vessel segmentation methods on the DRIVE database which they construct. Four of the methods are previously published and the fifth one, which is the only supervised method, proposed in this paper. This method uses a pixel classification approach to classify vessel/non-vessel pixels. A feature vector is assigned to each pixel where features are extracted from the green channel of the images. The green channel displays the best vessel/background contrast while the red and blue channels tend to be very noisy. For feature selection, the outputs of a Gaussian filter are used.

A linear classifier, a KNN classifier and a quadratic classifier have been compared where the KNN classifier is chosen as the best classifier due to the results of initial experiments. The filter features which consist of the Gaussian and its derivatives up to order 2 at scales s =1, 2, 4, 8, 16 are fed to the KNN classifier with k=30. Each feature is normalized to zero mean and unit variance before classification. Soft classification is applied where a probability map is constructed and the probability of each pixel belonging to a vessel is identified. In soft classification, the class conditional probabilities are estimated and the classification is performed based on estimated probabilities. From the probability map obtained, a binary segmentation of the vasculature is obtained through thresholding at an optimal threshold determined on the training set of the database. From the resulting thresholded images, the accuracy per image is determined by obtaining the ratio of the sum of correctly classified vessels and non-vessel pixels over the total number of pixels in the field of view. The field of view is the image area which contains the object of interest that is being measured. In this case it would be the area containing the blood vessel network of the retina. The maximum average accuracy (MAA) is then calculated. The maximum average accuracy (MAA) is calculated by establishing an optimal threshold on a training set and applying that threshold to a test set, then by calculating the average accuracy and obtaining the maximum value. The ROC curves are also plotted for the images and results are compared to previous methods. The Kappa value which is a measure of agreement between two observers is calculated for all methods. The advantage of the proposed

method is that it is a supervised method based on pixel classification compared to other methods which are non-supervised (not trained with examples). A disadvantage is that wider vessels have a larger influence on the end result than smaller vessels since the performance is evaluated based on the accuracy of segmentation compared to the gold standard segmentation. Also this method has high computational burden. The gold standard segmentation is the manual segmentation carried out by the first observer out of two observers.

5.2.2 Filter based Methods

Filter-based methods apply filters to images which remove unwanted information in the image. Steerable filters [80] are used to speed up the search for orientation information since a pair of orthogonal filter responses are combined proportionally to produce the response of the filter at various orientations.

Perez-Rovira et al. [81] apply 2D steerable filters and thresholding to detect vessels in retinal images. The Hessian filter is the most popular steerable filter which uses three second derivatives of the Gaussian. One of the major drawbacks is the difficulty in choosing the right scale suitable for vessels of all widths. The small filters will perform poorly on wide vessels and the large filters will perform poorly on the narrow vessels. Strong responses are normally generated around non-vessel structures such as the optic disc and field of view. This consequently reduces accuracy hence post processing techniques are required for their removal [82]. Below we describe a few common filter approaches used for vessel segmentation.

5.2.2.1 Matched filtering

A Matched filter detects known signals that have been contaminated by noise. This is achieved by cross-correlating the signal with the filter. The filter maximizes the signal to noise ratio (SNR) of the signal being detected with respect to noise.

Figure 43 shows a model of a matched filter:

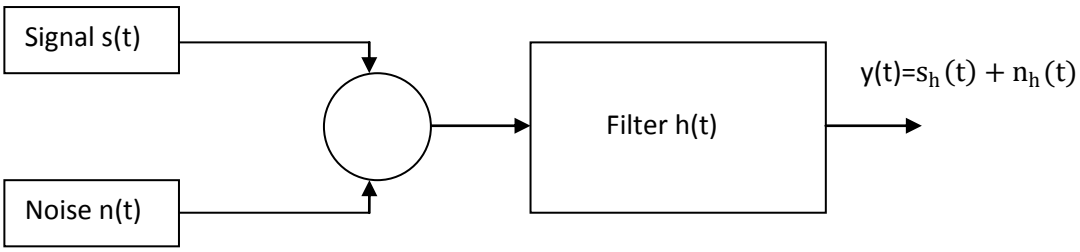


Figure 46 . A matched filter

s(t) is the input signal, n(t) is noise, then the maximum SNR at the output will occur when the filter has an impulse response that is the time-reverse of the input signal. The filter h(t) maximizes the SNR of the output y(t) [83]. When the input signal is a wavelet w(t), of length T then the matched filter is:

$$h(t)=w(T-t) \tag{19}$$

The drawback of applying matched filters is that many junction points and fine vessels are missed out during detection. False detections also arise as strong responses occur for non vessel edges such as the edges of red lesions and bright blobs in the retinogram. Also, since vessel properties such as thickness, shape and orientation vary across the image, designing filters that can model every possibility becomes difficult.

Al-Rawi et al [84] apply an optimization method to parameters of a matched filter in order to improve its performance. The matched filter kernel is rotated to all possible vessel orientations and the maximum response is obtained. The optimization procedure carries out a search for the best vessel segment length the kernel adapts to, and the best spread of the intensity profile. This method improves on the classical matched filter and outperforms it by utilizing better filter parameters. It is applied to images of the DRIVE database. The optimization procedure is applied to the 20 training images. The best filter parameters are chosen and applied.

5.2.2.2 Wavelet based Methods

Wavelets are used to decompose an image into multi-resolution versions [85]. A full description is given in section 3.3. Gabor filters are closely related to wavelets and are capable of tuning the specific frequencies and orientations, they allow for simultaneous vessel enhancement and noise filtering, but do not guarantee perfect reconstruction.

Bankhead et al. [85] threshold wavelet coefficients and apply spline fitting to extract vessel orientations, followed by a search for zero crossings of the second derivative in a direction perpendicular to the vessels. This method provides a high level of accuracy, low measurement error and a much shorter processing time than graph based methods [86]. It has a straightforward implementation. The thresholding of the wavelet coefficients causes the method to be much faster than unsupervised methods and achieves comparable accuracy. Although a low true positive rate is achieved because the narrow vessels cannot be detected, the accuracy is not compromised because of the low false negative rate. Due to the existence of noise, the accurate quantification of changes in the diameter of vessels that are between 1-2 pixels is not achieved.

5.2.3 Hessian based Methods

The Hessian matrix consists of the second order gradients of an image and is used to measure the curvature of a function. The vector of partial derivatives (gradient) is used to summarise slope information and when there are slopes in so many directions, the change in these slopes can be determined by the hessian matrix. Considering $f(x, y)$ as a real-valued function and providing all real valued derivatives exist, the gradient of the above function is given by:

$$\nabla f(x, y) = f_x \hat{\imath} + f_y \hat{\jmath} = [f_x f_y] \tag{20}$$

and the hessian matrix is

$$H_f(x, y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \tag{21}$$

Salem et al. [87] detect vessel centrelines based on the larger eigenvalue of the hessian matrix. The smaller value of the matrix is used to obtain the vessel orientations. Since

vessels have different diameters, different scales are used to calculate the eigen values, the maximum response is then chosen for each pixel at different scales. The variation of the direction of the eigenvectors in a pixel over six different scales is less for blood vessel pixels than for non-blood vessel pixels. One of the advantages of this method is that it only one parameter needs to be set for this algorithm which is the scale *s* which makes it simple and easy to implement. Supervised methods produce better results than the above method but they have the expense of a need for manual segmentation which is time consuming and the necessity of good feature vectors for classifiers used.

5.2.4 Mathematical Morphology

Morphological operators are used to analyse the topology of images. Dilation and erosion are the main operators. With dilation, objects grow and nearby structures are connected together which will fill holes and improve connectivity in image structures and the filling up of small holes. Erosion causes objects to shrink and erodes away the boundaries of regions of foreground pixels.

The structuring element controls the size, shape and direction of the operators hence affecting the objects in the image. Opening is erosion followed by dilation whereas closing is a dilation followed by erosion.

Abdurrazaq et al. [88] apply mathematical morphology to segment images. Opening and closing operations are used to produce tophat transformations. This results in an enhancement of bright or dark regions within the image. Morphological closing operation with an octagonal structuring element is carried out followed by dilation in order to reduce noise and to obtain the edge features of the image. The difference between the result and the original image is then calculated. This method is able to outline all changing pixels by subtracting the original image from the erosion or dilation result image.

The drawback of this method is that small vessels that are not connected to the rest of the vascular tree are not detected and disappear into the noise. They are not detected due to the structuring element being too large compared to the resolution of the image. On the larger vessels a bright strip appears on the centre of the vessel which causes misinterpretation of them being two small vessels running in parallel. Another disadvantage is that pairs of edges are misinterpreted as vessels.

5.2.5 Adaptive Local Thresholding

Thresholding is applied to images in order to binarize them into two classes, background and foreground. There are two types of thresholding, local and global thresholding. Global thresholding is used when there is a uniform contrast distribution of background and foreground whereas local thresholding is used where variation in intensity exists and the image is very noisy.

Jiang et al. [89] propose a vessel detection algorithm using adaptive local thresholding based on a verification-based multi-threshold probing scheme. The algorithm consists of a classification scheme which identifies regions within a binary image that can be accepted as objects. Several thresholds are used to probe the image. The binary results are combined to form the final segmentation. If an image contains dark objects and a bright background, thresholding at a particular threshold will result in a binary image in which pixels with intensity lower than or equal to the threshold will be objects and other pixels will be background. If circular objects are the desired objects to be found, a verification procedure can be designed which can identify whether a region is a true circular one. A component labelling operation is performed on each binary image and the verification test is carried out on all connected regions. A new binary image is obtained by retaining regions that passed the test hence the binary image will contain circular regions only.

As an example, the algorithm is applied to a binary image of uniform background intensity of 255 containing a circular object and a square object of intensity 20 and 200 respectively. The thresholds that are below 20 will lead to an empty binary image. The circular object is marked as object pixels for any threshold $20 \leq T < 200$. Also, any threshold $200 \leq T < 255$ will mark both the circular and square object as object pixels. If the search is for circular objects and the verification procedure P tells us whether an image region is a true circular one, a component labelling operation is performed on each binary image and P is applied to all connected regions. A new binary image results from the regions that have passed the test from P.

The object features used for classification in the verification procedure are shape, colour, intensity, and contrast. Properties such as scale and orientation invariance can also be used depending on the application. The probe thresholds can also be adjusted according to the

application. The simplest form takes values ranging from the minimum intensity to the maximum intensity of an image at some constant step.

Hence the efficiency of the algorithm is influenced by the probing thresholds and the verification stage. The algorithm then consists of three steps:

1. Apply Euclidean distance transform on binary image. The distance between the candidate vessel point to its nearest background pixel and the position of the background pixel are defined within the resulting distance map.

2. using the distance map, prune the vessel candidates to retain curvilinear bands which are defined by their respective centre line pixels. Pruning is done with respect to the properties width, size, contrast and angle.

3. From the centre line pixels reconstruct the curvilinear bands.

The method has been used on 20 retinal images where 10 of them are normal and the other 10 are abnormal. The advantage of this method is its simplicity and the fact that it fully utilizes all relevant information about the objects of interest. Vessel boundaries contain misclassified pixels. Therefore a post processing step is added where non-maximal suppression is applied. On the resulting image, closing is carried out. This step has a large effect on the boundaries and improves results. The disadvantage is that the filters show less accuracy when applied to abnormal retinograms.

5.2.6 Edge Detection

Edge detection methods identify edges in images. The first order discontinuities in an image are called step edges [90]. There are two main techniques used for edge detection. The first one is enhancement/thresholding where discontinuities in the image gray tone are enhanced by neighbourhood operators. The presence of noise degrades the performance of these algorithms as it decreases the sharpness of the edges. The second one employs an edge fitting method. The drawback of these techniques is that the large-scale properties of vessels are ignored in order to retain computational feasibility [91].

Akram et al. [92] apply a 2D Gabor wavelet to enhance the vascular pattern. Prior to extracting vessels from the enhanced image, sharpening filters are used to sharpen the

vessels. Vessel segmentation mask is created by extracting vessel boundaries using the Canny edge detector. Morphological dilation is then applied to produce the final refined vessel segmentation mask.
The proposed method produces an accuracy of 0.944±0.025 in comparison to the best method that produces an accuracy of 0.952±0.033.

5.2.7 Combination of Low levels

Some methods combine two or three of the above low level methods to extract blood vessels in medical images:

Jelinek et al. [93] applied the Gabor wavelet transform in order to obtain features to be fed to a Bayesian classifier. The continuous wavelet transform is applied to fluorescein-labelled retinal blood vessels and the maximum response over all angles for all scales is taken for each pixel along with its intensity value for the features of the classifier. The wavelet parameters such as scale, frequency, elongation were tuned so that the best matching between vessels and wavelets could be reached. The classifier features were normalized for training before the tuned classifier was applied to the images. A forward-backward feature-selection process using linear discriminant analysis was performed to select the best parameters. The above method was capable of classifying images that contain neovascularisation within the retina.

Cornforth et al. [103] apply the combination of wavelets, adaptive thresholding and supervised classification to segment retinal blood vessels. They use three different techniques on 20 images from the STARE database. The images are pre-processed using a mean filter. Method 1 uses continuous wavelet transform and adaptive thresholding, Method 2 uses wavelet transform plus supervised classifiers and Method 3 uses wavelet transform plus pixel probabilities combined with adaptive thresholding. A Bayesian classifier consisting of a mixture of Gaussians and was used for classification. Morphology was used for post-processing in order to fill the vessels and improve their connectivity. Methods 2 and 3 using the supervised classifier give the best results where more of the smaller vessels are detected. The method allows noise filtering and vessel enhancement in a single step through the wavelet transforms.

Sekhar et al. [95] detect blood vessels in images by examining their geometric and spatial distributions. Scale-space analysis is used to segment them. A hessian matrix is applied to obtain the orientations and centrelines hence the average thickness of these vessels. A 2D Gaussian kernel is convolved with the image to separate out information in the image at different scales by obtaining multi-scale derivatives. The edges and ridges in the image are then obtained via first and higher-order derivatives of the image. To obtain the edge points, the gradient of the image is taken. The local maximum points in the direction where the gradient image undergoes the largest change are the ridge points. To obtain narrow vessels, small scales are convenient whereas for wide vessels large scales are appropriate. From the eigenvectors of the Hessian matrix, the vessel orientation is calculated. The angle of the eigenvectors on average gives the blood vessel orientation. To obtain the width, a coordinate system is defined for each vessel since each vessel has a different orientation. The number of vessel pixels are counted beginning from the centre-line of the vessel in the positive and negative directions. Then the thin blood vessels are removed via a morphological top-hat operation. Then a parabolic Hough transform (PHT) is applied which is fast and robust on thick vessels [95]. The blood vessel binary image is transformed to a set of three parameters, representing the accumulator, focal length, and vertex. The accumulator will have a set of edge points where each edge point contributes a parabola with vertex (x_c, y_c) and focal length a in the accumulator. The peak of the accumulator corresponds to a parabola that shows the main blood vessel arcade with (a, x_c, y_c) as the parameters. The vertex is used to identify the optical disc location. The technique has been applied to the DRIVE and STARE databases.

Wu et al. [96] apply a multi-scale matched filter approach. The enhancement filter consists of a combination of Hessian-based filters and matched filters which operates with edge constraints on multiple scales and is easy to implement. A few pre-processing steps are carried out where the green channel of the retinograms are extracted and the resulting grey-scale images show a large contrast between vessels and background. The images are then inverted so that vessels have higher intensity than the background. Curvatures of the Hessian-based filters, which are second order derivatives, determine whether the analysed region is a vessel or non-tubular structure. Vessels have large curvatures whilst non-tubular structures such as planes have zero curvatures. The Eigenvectors of the matrix provide the orientation of vessels and whilst enhancing their size, estimate their directions. The ratio between the smallest and largest eigenvalues determines the filter response. But hessian-filters cannot distinguish between vessels and step edges effectively, therefore a matched

filter per scale per pixel is used which makes the filter more efficient as multiple kernels are used at each scale instead of requiring multiple filters per pixel. A multi-scale matched filter operates slightly differently, it convolves the image with second order derivatives of the Gaussian function at multiple scales and then extracts the maximum response of the output and thus enhances the vessels whatever their size whilst obtaining their direction. The matched filter has the benefit of distinguishing between a tube and a step edge, hence the proposed method combines the two filter types whilst taking in consideration edge information and treating vessels as two parallel edges. The responses of spurious boundary edges are suppressed via the edge constraints. They ensure that two edges exist at the boundary of vessels. The maximal response at multiple scales is the final output. The new filter is analysed by comparing it with a single multi-scale matched filter and the Frangi filter (Hessian-based vessel enhancement filter) using the DRIVE database. The filters are applied to the 40 images of the database. The proposed filter shows to be insensitive to the central reflex which is removed by a large Gaussian kernel. The matched filter produces more noisy responses than the hessian filter but shows more details of the image. So, the hybrid filter is much more capable as it is more sensitive to the central reflex and produces less noisy responses. Hence the hybrid filter produces the highest area under the ROC compared to the matched and hessian filter. This method is highly effective in eliminating false detections that are caused by lesions and abnormalities next to vessels or bifurcation points. It requires very little tuning and most of the small and large vessels are detected. Although the false detections are caused because of incorrectly identified initial tracing points.

5.3 Intermediate level analysis

Intermediate- level methods improve the connectivity of the vessels in images. The following section describes the most common Intermediate-level methods used to detect vessels in medical images. Some of the methods below combine Intermediate level methods with low-level methods to detect blood vessels.

5.3.1 Hough Transform

The Hough transform is used for detecting lines, circles and various shapes within an image. It works by detecting the global features via local measurements. The equation of a line is given below:

$$x \cos\theta + y \sin\theta = r \tag{22}$$

where r is the normal distance from the origin to this line and θ is the angle of r with respect to the x-axis.

If (x_i, y_i) are coordinates of edge points, they are the constants of the line equation above and r and θ are the unknowns. To map this point along with other edge points of shapes from the Cartesian grid to curves in the polar Hough parameter space, a Hough transformation is carried out. In order to detect a line in an image, the Hough parameter space is divided into accumulator cells. Each time a point is transformed into a curve, the accumulator cells which lie on the curve are incremented and the peaks of the accumulator array will show that a straight line has been detected. The same process can be applied to other shapes in the image.

Leander et al [97] develop a blood vessels segmentation method based on spatial high pass filtering. A median filter was also applied to reduce noise. The vessels were extracted with some bifurcations and intersections missing. Then the Hough transform was used to reintegrate vessel segments and transections at intersections and bifurcations. The Hough transform parameters were 10 line pixels (min number of pixels required for a line), 0-180 degrees for the range of angles along whose sides pixels constructing a line were to be connected, 6 pixels for each line segment. The algorithm was applied to the DRIVE images and the SNR, RMS error and PFM were computed to find the amount of match between the automatically-segmented and manually segmented images. The segmented images with and without the Hough transform were compared with the observer-segmented images. After applying the Hough transform the SNR and RMS error improved by 10%. The PFM was decreased by 6%. Mismatches between the two image sets were found due to missing smaller vessels.

5.3.2 Active Contour

Active contours or snakes are used in image analysis for the detection of object boundaries. They are influenced by external constraint forces that minimize the energy and internal forces that drag the contour towards lines and edges of structures.

Espona et al. [98] detect the vessel trees of retinal images by constructing a snake model which deforms and fits to the vessels and is guided by a vessel crease (ridge). The crease is a continuous area of points which have highest or lowest levels in their environment. The blood vessels are identified as tubular and extreme regions in their neighbourhood. The crease is used to determine the initial snake and hence is used to guide the contour expansion. Once the creases are identified, they are expanded in order to eliminate any discontinuities. When the snake is placed on the image, the global energy which is the sum of the internal and external energies is minimised in order to fit the snake to the vessels. The internal energy controls the elasticity of the snake whilst the external energy drives it towards the edges of the vessels. The snake is formed by setting its nodes to three states: normal, crease and edge. The crease nodes are placed on the vessel creases and form the vessel centrelines whilst the edge nodes are located on the vessel edges. The normal nodes are located on the remaining locations which cause the snake to expand in different directions at an intermediate level. A set of weighting factors and energies are defined and imposed onto the snake in order to control its deformation. The accuracy, sensitivity and specificity are determined when the model is applied to retinal angiograms. The sensitivity of the system is low due to thin unsegmented vessels. By dynamically tuning the vessel width parameter, this error can be reduced. The method takes low computation time. The drawback is that a lot of the thin vessels are unsegmented hence it produces low pixel sensitivity.

Al-Diri et al. [99] form a retinal vascular graph from segmented retinal vessels using a junction resolution algorithm. The graph is capable of determining the branching patterns of arteries and veins. The limitation of previous methods in detecting branches and crossings is that they fail at complex junctions whereas the proposed method overcomes this limitation and can recover from various segmentation failures, such as broken up vessels and overlaps. The method applies a search algorithm with a neural cost function. The image set consists of 21 images where 12 are normal. 435 junctions are marked and the resulting information is used to determine which features to include in the system. The

initial step is to initialize the algorithm with typically disconnected vessel segments. Each segment consists of a number of profiles that each have a centre point and width. Each segment consists of two segment ends (SEs), which we define as either:

1. Bridge- links a segment end to another one to form a continuous vessel.
2. Junction- Links an SE to two other ones.
3. Leaf- If the SE does not connect to anything.

An SE has a specific location, angle, width and intensity. Therefore bridges and junctions are identified and the disconnected SE neighbours which are caused by segmentation errors and undiscovered vessel segments are joined up based on cost functions that consider the junction geometry. A segmentation problem that arises is that SEs can be some distance from their proper joining point. Therefore each segment is extended by first assigning a flag to each SE initialized to “unresolved” and then extending the centrelines a step at a time until it is resolved. The vessel ridge is tracked via a twin contour which converges onto the ridge from either side. Overlapping segments are processed differently and they are replaced with two bridges. Each pair of overlapping SEs is connected via a bridge and if the cost is lower than a threshold for both bridges, then the bridge segments are created and the overlap segments are discarded. The results of the overlapping segment detection algorithm are assessed using the DRIVE database. The database contains 153 overlapping segments. Errors are caused by undetected overlapping segments and segmentation errors. The main junction resolution algorithm is assessed using the same database. 2803 JSE (Junction segment end) sets are identified, of which 21.8% are not solved at all, 72.1 % are solved correctly. The errors are caused by poor segmentation where 2.0% are false positive segments, 5.7% are missing segments and 0.7% are segmentation algorithm errors. The method is able to determine the connectivity of most junctions in the retinal graph.

5.3.3 Graph Based

In order to achieve accurate labelling when the set of labelled examples is small during classification, a graph on the whole training set is built using edge weights. This graph will show the similarity between examples and partitioning it will best explain the known labelled examples. In a training set where there are U unlabelled examples and L labelled examples (containing positive and negative examples), in order to find a labelling of the

examples in U that minimizes the leave-one-out cross validation error of a nearest neighbour algorithm, the graph cut method is used. It connects similar examples through edges with large weights and dissimilar examples through edges with small weights. A partitioning of the positive and negative examples is then carried out using the minimum cost cut.

Freiman et al. [100] present a nearly automatic graph min-cut segmentation method for the carotid arteries and aorta. For the vessel segmentations, a new graph-cut with tubular shape prior formulation is used and applied to Computed Tomography Angiography (CTA) scans and then applied to remove vein segments and to fill in vessel discontinuities that have been caused by large intensity variations. An edge weights function is also used for the graph min-cut method which couples the geometric vesselness shape prior, the intensity prior and the voxel intensity. The method is robust, accurate and easy to implement and is not time consuming and laborious compared to semi-automatic methods.

Salazar-Gonzalez et al. [101] segment retinal vascular network using the graph-cut technique. They enhance blood vessels then gather prior information with spatial pixel connection to construct the graph. The algorithm is tested on the DRIVE and STARE images. The advantage is that it does not require ad hoc tuning of parameters on different datasets.

5.3.4 Region Growing

Region growing methods work by comparing a starting pixel or seed pixel with its surrounding pixels and obtaining the similar ones and merging them with the initial pixel in order to form a region. The process is continued so the region no longer can grow. When the region stops growing, another seed pixel is chosen and other regions are formed. The drawback of this method is that the regions created are too dependent on the seed points, therefore there is a bias towards the regions that are initially segmented before testing other seed points. Also, when a region is growing, it will dominate the growth and the segmentation result.

Mendonca et al. [102] propose an algorithm which extracts vessel centrelines using a set of four directional differential filters and connect the resulting candidate points using a region growing process. The centreline segments are then validated based on the characteristics of

the line segments which is done in each one of the four directions and combined to obtain the vessel centreline map. During the vessel segmentation phase, vessels with different widths are enhanced using a top hat transform with different size structuring elements. A binary morphological reconstruction method is then applied followed by a region growing process to fill the vessels in the resulting binary maps. The method is applied to the DRIVE and STARE databases. The advantage of this method is that it is adaptable to image intensity since most algorithms are based on threshold values computed from local and global image information. Also the use of the vascular skeleton causes the exclusion of noise and better detection of thin vessels. Under segmentation of vessel regions occur as a consequence of the variability of intensity in the vessel regions. Also the variation of intensity and contrast in the vessels can sometimes make it difficult to distinguish between vessel and background noise resulting in some vessels not being reconstructed properly.

Staal et al [103] develop a method to identify vessel centrelines within an image containing a vessel network by extracting image ridges. Ridges are defined as points where the image has an extremum in the direction of the largest surface curvature. They are extracted to form primitives that are grouped into sets that approximate straight line elements. These sets divide the image into patches by assigning every pixel of the image to its nearest primitive. A local coordinate frame is defined by the line element in each patch in which local features can be extracted for every pixel. The features are used to classify the pixels in the patch into vessel and non-vessel. This is explained in more detail below:

To extract the image ridges, the green channel of the colour fundus images is used since it gives the highest contrast between vessel and background. Derivatives are taken at a scale using the Gaussian scale-space technique and they are convolved with the image. A search is carried out in the image for points where the first derivative of the intensity in the direction of the largest surface curvature changes sign. In order to do this, the hessian matrix of the image is calculated which is the eigenvector of the matrix of second order derivatives which corresponds to the largest absolute eigenvalue. The sign of the eigenvalue determines whether a local minimum or a local maximum is found. From this, a scalar field is defined over the image which takes the values -1 for ridges of local minima, 1 for ridges of local maxima and 0 elsewhere.

The next step is to group the ridge pixels that belong to the same ridge in order to form the primitives for the vessels which would represent line elements. A region growing algorithm is used for this where a grouped ridge pixel is compared with ungrouped ridge

pixels in a neighbourhood with a specified radius. To compare the two pixels together two conditions are considered:

1. The ridge pixels should have similar eigenvector directions.
2. If condition 1 is satisfied, the pixels should be on the same ridge (not on parallel ridges).

To check the first condition, if the scalar product of the eigenvectors at the location of the pixels is close to 1, it means that the pixels have similar orientation. The second condition is checked by computing the unit-length normalized vector between the locations of the two pixels under consideration and taking the vector product between the normalized vector and the eigenvectors of the grouped pixel. The vector product will be 1 if the pixels are on the same segment. The grouped ridge pixels are referred to as affine convex sets. They are convex since they approximate straight line elements and affine because instead of straight line (Euclidean) convexity they have a geodesic convexity. Depending on parameter settings, highly curved convex sets or straighter convex sets are formed. Hence based on these convex sets, each pixel is allocated to the most similar set and the image is divided into patches called convex set regions (CSR).

In order to select the classifier features, sequential forward selection is used where the best feature that satisfies some criterion function is chosen in the feature set at each step. Then the set that gives the best performance is chosen with the area under the ROC being the criterion function.

From the local appearance profiles perpendicular to the convex sets, features are extracted in order to distinguish the vessel sets from the non-vessel ones. From every point in the convex set, a profile is sampled from the image's green plane and is averaged. This mean profile is used to extract features. The blue channel contains a lot of noise and is often empty therefore it is not used to extract features. The features that are extracted are:

1. The height of the profile.
2. The width of the profile which is the length between the right and left edge of the profile.
3. The height over the width (h/w).
4. The edge strength.
5. The edge strength over the width.
6. The edge height.

7. The difference between the height and the edge height.
8. The height over the edge height.

Other features are also extracted which are not extracted from the profiles but can give useful information such as:

9. The distance between the first and last point of a convex set.
10. The length of a convex set.
11. The convex set curvature.
12. The mean of a rectangular image patch sampled from the green plane around a convex set.
13. The standard deviation for the green patch.
14. The mean value of the green plane at the locations of the convex set divided by the mean value of the red plane.
15. The ridge strength which is obtained by computing the mean value at the locations of the convex set.

Another set of features are extracted by fitting a straight line through a convex set and defining a local coordinate system for every convex set region (CSR). The centre of mass of the convex set is chosen for the origin of the coordinate system. One of its axes is defined along the direction of the fitted line while the other is perpendicular to the first and its direction is chosen so that a right hand oriented coordinate system is established. The CSR features are divided into three sets of features:

(1) Features that take into account convex set information:

a. The chance that the corresponding convex set belongs to a vessel.
b. The distance between the pixel and the closest point on the convex set.
c. The ratio of the red values of the pixel and the closest point on the convex set.
d. The difference in the red values of the pixel and the closest point of the convex set.

(2) Features that take no information of the convex set:

a. The ratio of the green values and red values of the pixel.

b. The value of the green plane of the image at the pixel location.

c. The value of the red plane of the image at the pixel location.

(3) Features that use the local coordinate system information:

a. The coordinate of the pixel with respect to the first axis.

b. The coordinate of the pixel with respect to the second axis.

c. Derivatives of the Gaussian at different scales.

The results of Staals method are compared to the results of Hoover et al. [104]. The ridges are extracted from the green channel. The pixel intersection of the convex sets with the vessel pixels of the ground truth images are used to construct training data for the convex sets. If more than 50% of the convex set pixels intersect, they are labelled as vessel else as non-vessel. In order to enhance local contrast, the pixels of every colour channel are locally normalized to zero mean and unit variance. 18 features are extracted per convex set. For extracting features for the CSR, the selection is done on a leave-one-out fashion to obtain probabilities for the convex sets. Then the convex sets of one image are classified based on the training data obtained from convex sets of other images. Derivatives of the local coordinate system are taken at different scales, resulting in 27 features. The area under the ROC (Az) is used as the function criterion which is averaged for all images to decide whether a feature should be included or not. Each feature is normalized to zero mean and unit variance since a KNN classifier is used which is sensitive to scaling between different features.

In the result images, under segmentation and over segmentation is caused which is important when determining vessel width for some applications. The vessel branches are not completely detected since mostly the small vessels are missed out. Also a high probability of vessel pixels is found around the boundaries of the FOV and the optical disc. Some of this is due to the blurring that is induced to extract ridges. Also the training data might be insufficient since the pathologies in the images are mostly classified as vessel since enough pixels with pathology might not have been included in the training set. Also the convex sets in the training sets could have been incorrectly labelled or no ridge might have been detected at a vessel location which is likely when detecting small vessels.

Staal et al. [103] established a benchmark for vessel detection and introduced the DRIVE database. Staals method was better than the vessel detection methods that were introduced

before it. By producing the DRIVE database and making it publically available they allowed subsequent methods to be compared with theirs.

Martinez-Perez et al. [105] segment the vasculature of retinograms by using the combination of region growing and scale-space analysis. Scale-space analysis gives information on the width, orientation and size of structures in images. The topological information is obtained from the first and second order derivatives of the intensity image where the two geometrical features, edges and maximum principal curvature are extracted. This will overcome the variations in contrast in the images. The first and second order-derivatives result from the difference between the result of filtering images with a Gaussian function at different scales. . The derivatives are translation, scale and rotation invariant. To obtain the edges, the magnitude of the gradient of the image is extracted over scales and weighted. The magnitude of the gradient image is high at edges and low at uniform regions. To obtain the maximum principal curvature, ridge pixels are identified from the second derivative information. The second derivative information is obtained in the form of a hessian matrix and the eigenvalues are used to measure the convexity and concavity to get ridge points. The maximum principal curvature corresponds to the maximum eigenvalue. The ridge pixels identified indicate a vessel region. Both these features are normalized by intensity scaling whilst retaining their local maxima. The intensity scaling is carried out because for large vessels, the local maxima response tends to be larger due to the fact that there is more absorption of non-red light in the red-free images hence increased fluorescein in fluorescein images in larger vessels. So the intensity scaling factor which is related to the width of the vessels is used to normalise each feature.

A multiple pass region growing procedure is then applied which uses the histograms of the above two features (ridge strength and edge strength) combined with spatial information from the 8-neighbouring pixels. This is done in order to label the pixels as vessel or non-vessel pixels. The first stage of the growing is fast since it is constrained to regions of low gradient magnitude, therefore regions outside boundaries will have rapid growth and vessels will grow only when the values of the maximum principal curvature lie within a wide interval. This is done by alternating the vessel and background region growing and lowering the feature thresholds after each iteration until no new pixel is added to the classes. For one feature, only one class is used (low gradient). For the other feature, the histogram is divided into two classes and their means and standard deviations are calculated.

At the final stage of the growing, borders between classes are identified by removing the gradient restriction and simultaneously growing vessel and background classes. This algorithm is applied to red-free and fluorescein retinal images. The fluorescein retinal images are segmented better than the red-free images. The variation of intensity in the images degrades the performance of this algorithm and the optical disc of the retina causes over segmentation. The region growing algorithm is fast since the growth is restricted to regions with low gradients which allow vessels to grow where the values of the maximum principal curvature lie within a wide interval. Therefore rapid growth of regions is achieved outside the boundaries.

Hoover et al. [104] apply a segmentation algorithm which uses the combination of global and local vessel features. The image is convolved with a matched filter and an area of the response image is probed, iteratively decreasing the threshold. Each pixel in the probe queue is used as a starting point for the probing and the initial threshold is the matched filter response image value at the starting pixel. The attributes of the considered region are computed at each iteration in order to determine the threshold and to identify whether the region belongs to a vessel and if further probing is required. Pixels that are not classified as vessel need to undergo probing again. For each probe, a queue of points is initialized. If the region belongs to a vessel, the end points of the region will be added to the queue after completion of probing. Therefore different thresholds (probes) can be applied throughout the image. In order to initialize a queue of pixels as starting points for probes, the following steps are carried out:

1. The image is convolved with a matched filter.
2. The histogram of the result image is obtained and the image is thresholded above a certain threshold.
3. The thresholded image is thinned.
4. The branch points within the thinned image are erased. Branch points are defined as any pixel for which a traverse yields more than two transitions. This will break up the foreground into segments that contain two endpoints each. An end point is defined as any pixel for which a traverse of the eight bordering pixels in clockwise order yields only one foreground-to-background transition.
5. Segments shorter than ten pixels are discarded.
6. Remaining endpoints are placed in the probe queue.

These starting points can be placed anywhere within the vessel network therefore branches and pathology do not cause parts of the network to be missed. Then each pixel is taken from the queue and used as a starting point for threshold probing. Using a conditional paint-fill technique, at each iteration of the probing, a region is grown from the start pixel and the paint fill spreads over all connecting pixels that are above the current threshold that need labelling. When the paint fill is complete the attributes of the region are tested. If the region passes the test, then the threshold is decreased by one and a new iteration begins. Once the probe is complete, the region is labelled as vessel. Then the endpoints of the vessel region are added to the queue and the next point is selected for probing and the algorithm continues until the queue is empty. This algorithm was applied to all 20 images and ROC curves were constructed. The false positive rate obtained is reduced by 15% over the basic thresholding of a matched filter response. An 80% true positive rate is achieved. A false positive rate comparable to a human observer is achieved since ¾ of the vessels are segmented. The method reduces false positive rate by a factor of 15 times compared to the basic thresholding of matched filter. The disadvantage is that the basic thresholding produces small groups of isolated pixels.

5.3.5 Mathematical Morphology

This method can be used on an intermediate level basis where it is applied to the segmented result in order to improve the connectivity of vessels.

You et al. [106] use radial projection and a semi-supervised approach to extract retinal vessels. Vessel centrelines are identified through radial projection as described below. A steerable complex wavelet is used to enhance the retinogram. Radial projection is performed based on piecewise linear property (vessels are piecewise linear). The first step of pre-processing is to grow the region of interest (ROI) which reduces the false detection around the border. For each pixel in the ROI, radial projection is then applied. The pixel is taken as the centre of radial projection to judge whether it belongs to a vessel or notand then projection is performed along different radial directions. A projected curve is hence produced for each pixel containing a series of value of radial projection (VRP) at a particular orientation. This will distinguish vessel pixels from background points. If the pixel belongs to a vessel branching, the projected curve will have at least three prominent peaks.

To filter out vessel pixels that have falsely been detected, a pruning operation is carried out by examining their intensity and length of segment. If the intensity value is lower than the median value, the pixel belongs to a vessel.

For the classification, a steerable complex wavelet transform is used where the line strength at different scales and the average intensity and largest value at an orientation are used for the feature selection. The feature vector is then fed to an SVM classifier. To connect discontinuity between vessels, the morphological bridge operation is applied to the image.

The misclassified background pixels that surround vessels are identified via the four neighbourhood connectivity. The algorithm is applied to the DRIVE and STARE database images that contain ground truth. The accuracy, sensitivity and specificity are used for evaluation measures. The results when compared to previous methods show that the common problem of inaccuracy for extracted narrow low contrast vessels is overcome. There is false detection due to noise around the optic disk border and pathological regions. Also, some narrow vessels are overestimated because of the vessel-like noise.

5.3.6 Tracking

Tracking-based methods track vessel edges by finding the best path by being guided by local information [107]. The tracking starts off with seed points that are either manually or automatically selected. Manually selecting seeds is time consuming, labour intensive and dependant on the expertise of the user [108]. Most tracking methods estimate the width and orientation of the current point in a region and track the next point based on the orientation of the current point and the process is repeated until the full vessel is tracked.
A drawback of these methods is their tendency for termination at branch points and crossings and being too dependent on the starting points. Also low contrast vessels cause discontinuities and the termination of the tracking path.

Can et al [109] apply a vessel tracking method which starts by searching the image along a grid of one –pixel-wide lines and estimating the intensity levels in order to find the local intensity minima. The minima points obtained are used for seed points and the search starts at each of these seed points. A set of two-dimensional correlation kernels are used

perpendicular to the direction of the vessels. The method is efficient since direct processing is done on intensity data without any pre-processing and avoiding low-level image operations such as edge detection and morphological processing. The algorithm is applied on normal and abnormal retinograms and performs well regardless of significant changes in illumination and is able to trace branches and crossovers. (24

5.4 Comparison

In this project, we aim to detect blood vessels as accurately as possible so we are interested in a method that is adaptable to detecting vessel networks of varying shapes comprising of coarse and fine vessels arranged uniquely within each network.

From the low level methods, the classification methods can be used to distinguish between vessel and non-vessel pixels and to analyse the vessel network at a pixel level. We have chosen SVMs and random forests based on the reasons explained in section 4.13.

From the filter based methods we have chosen wavelet based methods and particularly the dual-tree complex wavelet transform [3] which is the only method that is shift-invariant. We have also used mathematical morphology which is suitable for this application and improving results after classification (Chapter 7).

5.5 Classifier Evaluation

The performance of a classifier can be evaluated using various techniques, although the intended application and the characteristics of the data are important factors that affect the evaluation results.

Common measures for classifier performance are false positives, false negatives, true positives and true negatives. False positives are data that have been incorrectly classified as having positive labels compared to the labelled data. False negatives are data that have been incorrectly classified as having negative labels compared to the labelled data. True positives are data that have been correctly classified as positive labels compared to the labelled data and true negatives are data that have been correctly classified as negative labels compared to the labelled data. The receiver operating characteristic (ROC) curve is a

common method for summarising the trade-off that generally exists between true and false positives [110].

Classification evaluation is often based on prediction accuracy (the percentage of correct prediction over the total predictions). In order to avoid bias in evaluating a classifier, the data on which the classifier is tested is different from the data used to build it so separate sets for training and testing are used.

The most common technique used to determine the accuracy is cross validation where a proportion of the data is used for training and the rest of it is used for testing.

The most common cross validation technique is the leave-one-out cross validation where each sample is taken as the evaluation set. A single sample of the full data set is initially omitted, the model is then built, the sample is then tested on and finally the above steps are repeated until each sample has been at least omitted and tested once. Leave-one-out testing is very slow and not feasible for large datasets, it is computationally expensive and a large number of training runs are required [111].

5.5 Performance of existing methods on the DRIVE database

Table 1 shows results obtained from the DRIVE website which is a comparison of the earlier methods applied to the DRIVE images. The automated methods are compared to the human observer method (Refer to 2.6.1).

The performance is evaluated based on the maximum average accuracy (MAA) of the segmentation in comparison with the first-observer based gold standard. MAA is achieved by initially thresholding the result images at an optimal threshold determined on the training set. Using these thresholded images the accuracy per image is determined by taking the total number of correctly classified vessel and non-vessel pixels for each image and then dividing this sum by the total number of pixels in the FOV. The MAA is then determined by taking the average value of the accuracy for all images [2]. For methods where the result is a probability image rather than binary, the MAA is obtained by thresholding the probability image at the threshold of the probability image that gives the best MAA.

The kappa value is also used to assess the measure of agreement between the segmented results and the ground truth images. Kappa is a measure of agreement between two observers to see how often they agree in their interpretations. (See 8.2.2.2 for a detailed explanation). The ROC curve is also used which displays the true positive rate (sensitivity) over the false positive rate (1- specificity) whilst the threshold is varied. The area under the ROC (Az) and accuracy are used as the main measures for comparison between methods [112].

Table 1. Comparison of retinogram classification methods provided on the DRIVE website. The authors provide standard errors (in brackets in table) on the MAA.

Method	Maximum Average Accuracy	Kappa	A_z
Human Observer	0.947 (0.0048)	0.76	n/a
Staal	0.944 (0.0065)	0.73	0.952
Niemeijer	0.942 (0.0065)	0.71	0.929
Zana	0.938 (0.0077)	0.70	0.898
Al-Diri	0.926 (0.0126)	0.67	n/a
Jiang	0.921 (0.0076)	0.64	0.911
Martinez-Perez	0.918 (0.0240)	0.64	n/a
Chaudhuri	0.877 (0.0232)	0.33	0.788

We have constructed Table 2 which is an update of Table 1 and includes more methods that have been applied to the DRIVE images for comparison purposes. The ordering is done by initially ordering them by the area under the ROC ($\boldsymbol{A_z}$). For methods that have not provided $\boldsymbol{A_z}$, the accuracy has been used for ordering, and for methods that have not included $\boldsymbol{A_z}$ or accuracy, they have been ordered alphabetically.

Table 2. Comparison of retinogram classification methods on the DRIVE images.

The ordering is done by initially ordering them by the area under the ROC(A_z). For methods that have not provided A_z, the accuracy has been used for ordering, and for methods that have not included A_z or accuracy, they have been ordered alphabetically. , Sens. =sensitivity, Spec. =specificity, TPF=True positive rate, FPF=False positive rate.

Authors	Accuracy	Standard Error	Kappa	A_z	Sens.	Spec.	TPR	FPR
Wu				0.972	0.9517	0.9023		
Akram	0.947	0.0053		0.963				
Wu				0.963				
Carnimeo				0.962				
Lam	0.947			0.961				
Soares	0.947			0.961				
Rangayyan				0.960				
Oloumi				0.960				
Ricci	0.963			0.965				
Moghimirad	0.966			0.965				
Staal	0.944	0.0065	0.73	0.952				
Kande	0.891			0.952				
Rangayyan				0.950				
Al-Rawi	0.953			0.943				
Adjeroh	0.964			0.943	0.7346	0.9568		
Cinsdikici	0.929			0.941				
Amin	0.920			0.943				
Niemeijer	0.942	0.0065	0.71	0.933				
Garg	0.936			0.933				
Jiang	0.921	0.0076	0.64	0.911				

Zana	0.938	0.0077	0.70	0.898				
Cummings	0.924			0.864				
Chaudhuri	0.877	0.0232	0.33	0.788				
Amiri	0.985							
Osareh	0.968					0.9650	0.9710	
Saleh	0.965						0.8431	0.0283
Saleh	0.963						0.8423	0.0342
Zhang	0.961						0.754	0.0228
Lin	0.950							
Salazar-Gonzalez	0.948						0.6782	0.0271
Chui	0.947	0.0048					0.7760	0.0275
Human Observer	0.947	0.0048	0.76					
Mendonca	0.946	0.0065					0.7315	0.0219
Rezatofighi	0.946						0.7358	0.0233
Miri	0.946						0.7352	0.0205
Marin	0.945				0.707	0.9801		
Fraz	0.944				0.731	0.9680		
You	0.943				0.741	0.9751		

Pourmorteza	0.940						0.7492	0.0384
Zhang	0.938						0.7120	0.0276
Bankhead	0.937						0.7027	0.0283
Martinezperez	0.934						0.7246	0.0345
Espona	0.932				0.731	0.9600		
Espona	0.932				0.663	0.9682		
Vlachos	0.928	0.0088			0.747	0.9551		
Al-Diri	0.926	0.0126	0.67					
Palomera-perez	0.925	0.011					0.644	0.033
Amin	0.920	0.94						
Martinez-Perez	0.918	0.0240	0.64					
Alonso-monte	0.918							
All Background	0.873	0.0123						
Siddalingswy	0.842							
Abdurrazaq							0.8214	0.0941
Adel							0.8	0.094
Bhuiyan					0.844	0.9961		
Hoover					0.682	0.9806		
Muramatsu							0.793	0.410

Oost					0.845	0.9559		
Patasius					0.401	0.9893		
Reza							0.7704	0.0693
Wang					0.841	0.966		
Yin							0.73	0.039
Zhang							0.5766	0.0028

Table 2 shows the results of all methods applied to the DRIVE database. The evaluation has been carried out using the measures given in the table and those used by the authors of the DRIVE website.

A comparison of the early methods on the website has been done by Niemeijer et al. [2]. In [2] the DRIVE database is constructed by the authors and comparison is carried out between the methods by creating Table 1.

Niemeijers comparison identified Staal's classification method [103] as giving the best results in terms of detecting both coarse and fine vessels. This is concluded from the fact that in methods such as **Chardhuri et al.** [113] and **Martinez-Perez et al.** [105], over-segmentation results from the optical disc and in methods such as **Niemeijer et al.** [2] under-segmentation is caused from the area around the optical disc. Also, Staal's method provides an area under the ROC curve of 0.952 which is higher in comparison to the ROC curves of the other DRIVE methods. Although, a disadvantage in this case is that when taking the MAA as the criteria of accuracy, the wider vessels have a larger influence on the end result than the smaller vessels. Also, methods should be compared based on the application in which they are being used. If small vessel detection is more important in an application than large vessel detection, then if a method was to only detect small vessels would be marked as the best method.

5.6 Other work on analysing retinal images

5.6.1 Quantitative Analysis

The quantification of vessel features such as diameter, width, shape and tortuosity are used for the prediction of diseases such as diabetes and hypertension. Tortuosity is the measure of how tortuous or bent the blood vessels are. Below gives a description of a few methods that have been applied for quantitative analysis of features:

Zhou et al. [114] apply a vessel tracking algorithm which works by initially identifying the vessel centrelines and choosing the next location based on a number of vessel properties, under the constraint of a Gaussian modelled vessel profile. Each vessel contains unlimited segments with corresponding directional vectors. Their method also identifies the curvature index, the width of the vessel lumen and the edge location as well as the vessel centreline.

Wu et al. [115] obtain the radius at each pixel of vessel centrelines by observing the relation between vessel radius and the derivatives of Gaussian kernels. Bhuiyan et al. [116] develop a new method for detecting retinal vascular tortuosity of vessels with different widths and directions. The vessel edge, centreline and width information are considered. The centrelines are fragmented at branch points to create vessel segments. The tortuosity of the segment is measured on its edge and sample pixels are computed for angle measurement. The absolute angle difference is obtained from each pair of pixels. The average angle is then obtained which gives the tortuosity of the vessels.

Hatanaka et al. [117] track vessels and extract centrelines by using a thinning algorithm. Points that had two neighbouring pixels were defined as one of the elements of a straight line and points with three neighbouring pixels were defined as bifurcations and points with four neighbouring pixels were defined as furcations. The direction of a vector was then determined by the angle made by a bifurcation point and a furcation. The straight lines achieved from above were approximated on the centrelines and then a perpendicular line was determined using the approximated line of the centreline. The next step was to determine vessel walls using the zero-crossing method [118]. The vessel diameters were then measured based on the distance between the vessel walls on the approximated line.

5.6.2 Vessel crossings

Vessel crossings are the points on the vascular tree where vessels cross each other and coincide. Bifurcations are points where vessels split into smaller vessels. The connectivity of the network is highly dependent on the bifurcation and crossing points.

Calvo et al. [119] classify detected vascular points into bifurcations and crossovers by initially skeletonising the retinal vessel tree and using a median filter to remove noise. Then a Tophat filter is then applied which enhances the vessels which are big and dark. Two steps are then carried out in order to extract the vascular. The first step is a segmentation step where hysteresis thresholding is used followed by the removal of isolated pixels to get rid of spurious structure that does not belong to the vascular structure. The second step is that mathematical morphology is used to fill the gaps inside the detected vessels. An intersection number is calculated for each feature point based on local information and analysis of the neighbours of each point where the crossovers and bifurcations have a high intersection number.

Ortega et al. [120] detect landmark points which are crossovers, bifurcations and endpoints in vessels. Dilation is initially carried out to fill in the gaps in the vessels using median filtering. A skeletonization of the result is then done by a thinning process which erases the pixels from the borders towards the vessel centre without affecting the connectivity. The detected points are divided into two sets: vessel end points and bifurcation or crossover points using the intersection number obtained using local information from surrounding neighbours of each point. Then a vessel tracking algorithm is applied using the above information to label all segments in the image and to find the segments containing branches. A pruning task is then applied which erases all branches that are shorter than a chosen threshold.

Bhuiyan et al. [121] detect bifurcations and crossovers by initially applying a morphological thinning operation to find the vessel centrelines. Then a filter is applied on the centreline image and the geometrical and topological properties of the vessels passing through the identified landmark points are considered and the number of these blood vessels is used for identifying whether the point is a crossover point or landmark point.

5.6.3 Detecting the optic disc

The optic disc is an important anatomical landmark in the retina and one of the problems during detection is caused by the existence of bright lesions in the image which have a higher contrast than the optic disc. [18]

Lu et al. [122] detect the optic disc using a line operator. Line operators detect line segments of specific length at multiple orientations centred on each pixel. The image intensities along all oriented line segments are stored in a matrix where each row represents the intensity at a specific line segment. The image variation along the oriented line segments is computed. The orientation of the line segment with the maximum and minimum variations is used to locate the optic disc.

5.6.4 Abnormality Detection

Once the position of the major anatomy including the optic disc is detected, it is easier to detect the abnormal structures within the image, especially abnormalities that are analysed based on the position of a lesion in relation to the major anatomy [15].

One of the signs of diabetic retinopathy as described in chapter two is the appearance of red lesions (i.e. microaneurysms and haemorrhages). Bright lesions that appear due to diabetic retinopathy are cottonwool spots and exudates [15].

Niemeijer et al. [123] detect red lesions in retinal images based on a pixel classification approach. They initially apply a tophat transformation in order to discriminate between circular red lesions and the vasculature. The tophat transformation applies mathematical opening to the images with structuring elements of different orientations and a length that is bigger than the largest red lesion in the set. 12 structuring elements were used which resulted in 12 opened images. The maximum value at each pixel from the images is extracted to develop a map of the vasculature. The map is then subtracted from the original pre-processed image in order to obtain an image containing the lesions. A matched filter was then used to enhance the contrast between the lesions and the background. Then the filtered image was thresholded in order to produce a binary image. A region growing procedure is then applied to grow the pathologies. A KNN classifier is used which allows

for soft classification. The resulting pixel classification result is then thresholded to obtain the final binary image containing the red lesions.

5.7 Summary

In this chapter, a literature review on current state of the art vessel segmentation methods was presented. In particular, results comparing different methods on the DRIVE database. Table 2 presented the results of the methods applied to the DRIVE database including the earlier methods presented on the DRIVE website. All methods have common limitations which is difficulty in detecting the small, low contrast vessels. The filter-based and classifier-based methods are the most widely used methods due to their high ability in detecting local structure.

Our aim is to achieve better results than the previously reported results and to achieve this we propose to use state of the art features and classification methods, 1. Random forests (chapter 4) and the dual-tree complex wavelet transform (chapter 3) which provide a rich and complete but computationally efficient description of local structure and 2. Support vector machines and the dual-tree complex wavelet transform. The DT-CWT is shift invariant and we believe that it will improve the classification of vessel and non-vessel pixels.

In the next chapter we present the above two state of the art methods and apply the best to the DRIVE images. For comparison, we present results using the same evaluation measures as those in Table 2.

CHAPTER 6: Learning to Detect Vessels

6.1 Introduction

This chapter describes in more detail the proposed approach to analysing retinograms using the dual-tree complex wavelet transform and random forests or SVM classifiers together with the results of evaluation using both synthetic and real data. The experiments using synthetic data are carried out to assess the approach and obtain centreline classification results as a proof of concept experiment. Experiments on real data are carried out with different feature sets for the classifiers in order to optimise the results. The vessel pixel classification of the real retinograms has been carried out with two approaches: the first classifies all vessel pixels from the background and the second approach classifies vessel centrelines from the background. Given these baseline low-level results we present Intermediate-level processing methods designed to improve performance.

6.2 Overview

Classification using features is used as the machine learning technique chosen for this project, as identified in chapter 4 it is the best approach for this project since it is more suitable for this application of identifying vessel and non-vessel pixels in retinograms.

The dual-tree complex wavelet transform is used to obtain features for the classifiers since it is rotation, scale and shift invariant and provides rich information on each pixel of an image which is suitable for training the classifiers.

In section 6.4.1 we will explore the different ways of obtaining features to construct the feature vectors for the classifier, based on the real and imaginary dual-tree complex wavelet transform coefficients at different scales and orientations obtained at each pixel. Training data is gathered from the example of vessel pixels and non-vessel pixels to train classifiers and then applied to the test data.

For SVM classification the LIBSVM software [124] has been used. The package includes a parameter search code which allows one to find the best parameters for a given kernel. We have used the parameter search code to find the kernel ideal for the experiments described

below. The kernel used for the pixel classification tasks below is the Gaussian kernel. For random forest classification, in-house software by M Berks [125] was used.

6.3 Working with Synthetic Data

6.3.1 Design of Experiments

In the synthetic data approach, images of Gaussian bars of different widths and angles are created. The reason why Gaussian bars are chosen is because vessels resemble Gaussian bars. The bars are placed in the centre of each image. The dual-tree complex wavelet transform is then applied to these images at 6 scales. Then interpolation [61] between samples is carried out as described in chapter 3 (section 3.4.3). All responses in terms of magnitude across all 6 sub-bands (for each subband) and at each scale are extracted. The phase that corresponds to the magnitudes is also extracted. The magnitude and phase vectors hence are features that are fed to both a support vector machine and a random forest classifier. The classifier is trained on bars of angles 0, 60 and 130. Both SVM and Random forest (RF) classifiers are tested on bars of angles 15, 30, 45, 75, 90, 105, 120, 135, 150, 165 and 180. The centreline pixels are classified from the remaining regions. The centrelines of the Gaussian bars are known, therefore when observing the result of the classifier, a clear evaluation can be carried out and the pixels misclassified can be clearly observed. The bars widths range from 3 to 13 pixels width in steps of 2. The narrowest vessels in the retinograms are 3 pixels wide and the widest ones are typically 9 pixels wide.

Figure 44 shows example Gaussian bars of different angles and widths. As image pixel values are in the range [0 1], the Gaussian bars have values in the same range as shown with the colorbar in Figure 44.

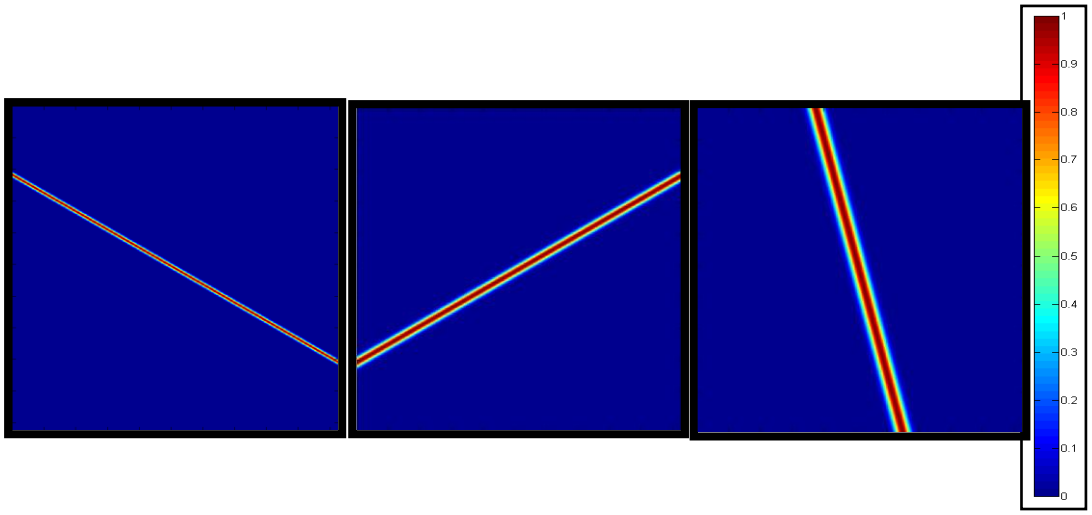


Figure 47 . Gaussian bars.

(Left) Bar of width = 3 pixels, angle =30. (Middle) Bar of width = 7 pixels, angle =150, (Right) Bar of width = 11, angle =75.The colours are pixel values.

We carry out the above on 2 sets of Gaussian bar images. The first set of images are noise free and the second set contains images in which we add Gaussian white noise and we add the variance: 0.002, 0.004, 0.006,0.008, 0.0016, 0.0032, 0.064, 0.128 and 0.256, doubling it each time.

Figure 45 shows the same bars as Figure 44 with increasing levels of noise.

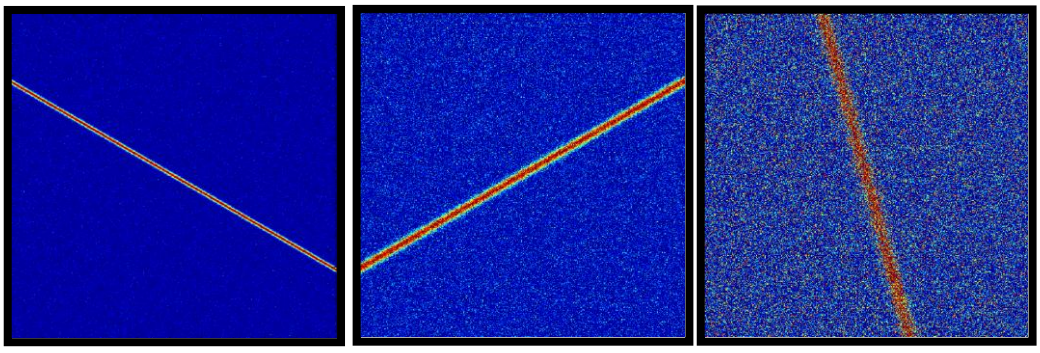


Figure 48 . Gaussian bars.

(Left) Bar of width = 3 pixels, angle =30, noise variance = 0.004. (Middle) Bar of width = 7 pixels, angle =150, noise variance = 0.032, (Right) Bar of width = 11, angle =75, noise variance = 0.256.

6.3.2 Results

Figure 46 shows the classification results of applying the SVM classifier to the Gaussian bars in Figure 44.

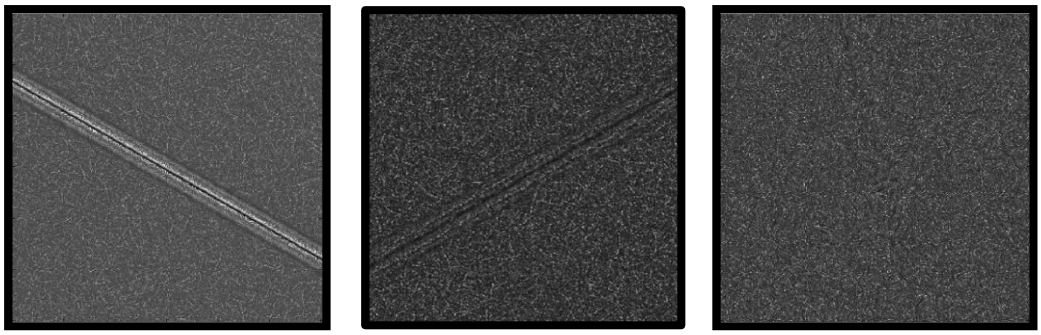


Figure 49 . Classification results of Gaussian bars in Figure 44 using an SVM classifier.

(Left) SVM Classification result of bar of width = 3 pixels, angle = 30, noise variance = 0.002. (Middle) Bar of width = 7 pixels, angle = 150, noise variance = 0.032. (Right) Bar of width = 11, angle = 75, noise variance = 0.256.

Figure 47 shows the classification results of applying the RF classifier to the Gaussian bars in Figure 45.

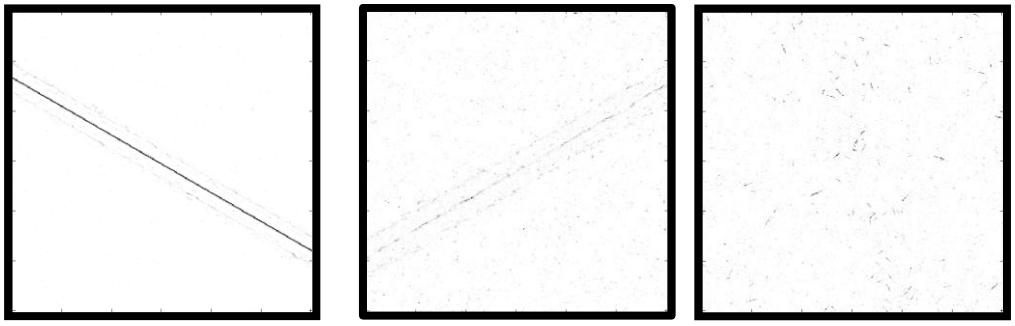


Figure 50 . Classification results of Gaussian bars in Figure 45 using a random forest classifier.

(Left) RF Classification result of bar of width = 3 pixels, angle = 30, noise variance = 0.002. (Middle) Bar of width = 7 pixels, angle =150, noise variance = 0.032. (Right) Bar of width = 11, angle =75, noise variance = 0.256.

Table 3 shows the results of applying the random forest classifier to Gaussian bars of all test angles highlighted in 6.3.1. The average mean over all angles is shown for each bar width as well as the standard errors. Table 4 shows this for SVM classification.

Table 3. Random forest classification results on Gaussian bars.

The mean and standard errors of area under the ROCs are presented.

Random Forest						
Width	**3**	**5**	**7**	**9**	**11**	**13**
Mean	0.999	0.977	0.832	0.620	0.558	0.569
SE	0.0003	0.0062	0.0173	0.0124	0.0054	0.0062

Table 4. SVM classification results on Gaussian bars.

The mean and standard errors of area under the ROCS are presented.

SVM						
Width	**3**	**5**	**7**	**9**	**11**	**13**
Mean	0.982	0.940	0.774	0.566	0.525	0.516
SE	0.0020	0.0060	0.0143	0.0065	0.0036	0.0034

Table 5 shows the classification results of applying the random forest classifier to Gaussian bars of all test angles averaged over all widths. The average mean over all angles for the noise levels are shown in the table. Table 6 shows this for SVM classification.

Table 5. Classification results of the Random forest classifier over all noise levels.

The mean and standard errors of area under the ROCS are presented.

Random Forest		
Noise level	**Mean**	**SE**
0	0.844	0.0212
0.002	0.831	0.0214
0.004	0.804	0.0233
0.006	0.778	0.0259
0.008	0.749	0.0279
0.016	0.734	0.0275
0.032	0.720	0.0263
0.064	0.698	0.0257
0.128	0.677	0.0240
0.256	0.643	0.0261

Table 6. Classification results of the SVM classifier over all noise levels.

The mean and standard errors of area under the ROCS are presented.

SVM		
Noise level	**Mean**	**SE**
0	0.744	0.0265
0.002	0.735	0.0270
0.004	0.729	0.0269
0.006	0.723	0.0270
0.008	0.721	0.0269
0.016	0.717	0.0260
0.032	0.711	0.0243
0.064	0.701	0.0225
0.128	0.673	0.0200
0.256	0.640	0.0268

Full results are given in appendix 6 for random forest classification and appendix 7 for SVM classification.

As seen in Figure 46, the SVM classifier detects the centrelines of the Gaussian bars but causes edge effects where false positives are detected around the edge of the bars. In the random forest classification (Figure 47) edge effects are very less and a significant reduction in the false positives can be seen.

6.4 Working with Retinograms

To classify a raw retinogram, it is converted from a colour image (Figure 49) to a grey-scale image using the Matlab default function (rgb2gray) which converts an RGB image to a gray scale intensity image by eliminating the saturation and hue information retaining the intensity, as shown in Figure 50. Labels are extracted from the corresponding ground truth image (Figure 48).

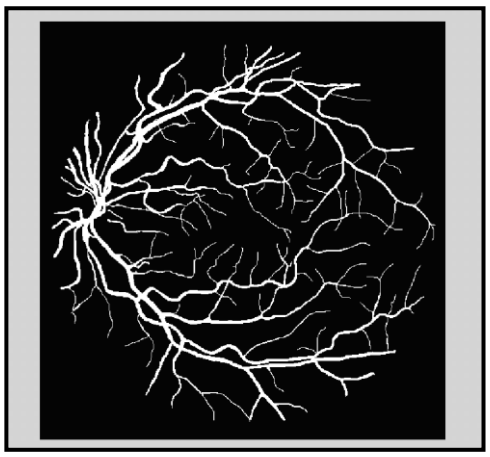


Figure 51 . The binary version (Ground truth) of the retinogram in Figure 49.

Figure 52 . A raw retinogram.

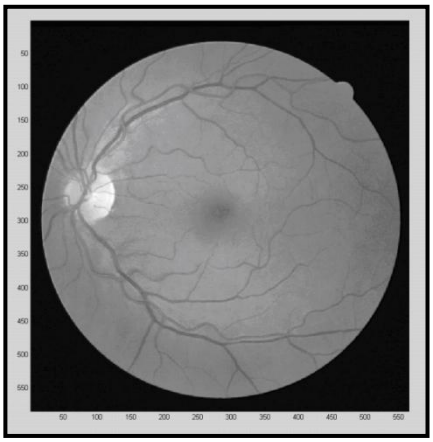


Figure 53 . The grey-level image of Figure 56.

The pixel classification process begins with applying the dual-tree complex wavelet transform at six levels to the original raw retinograms. Then interpolation [61] between samples is carried out as described in chapter 3 (section 3.4.3), giving 6 (levels) × 6 (sub-bands) × 2 (real + imaginary) = 72 coefficients. These are used in various combinations, as described in section 6.4.1.

To detect centrelines, the classifier is trained by initially applying skeletonisation [126] using an iterative thinning algorithm [127] to the training image which involves preserving the original structure of the object and discarding the foreground surrounding pixels.

The skeletonised image provides the centreline pixels that are fed to the classifier during the training phase. Therefore a set of centreline ground truth images are created by taking the ground truth training images and applying skeletonisation. Similarly, a set of centreline ground truth images are created from the test images.

Figure 51 shows a skeletonised version of the ground truth image. This image is used to identify the vessel centrelines for training.

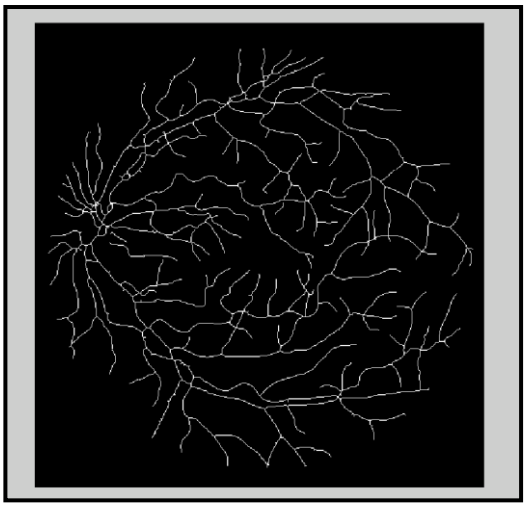


Figure 54 . A skeletonised binary retinogram.

The process involves removing the foreground pixels and preserving the original structure which in this case are the vessel centrelines.

6.4.1 Classifying using different feature sets

In this section, two classifiers are applied to the retinograms and a series of different feature vectors are tested in order to obtain the best classifier approach and the best results.

The dual-tree complex wavelet transform is applied at 6 scales to the 20 training images of the DRIVE database. The 6 sub-bands at each scale and their responses in both complex and magnitude and phase form have been obtained and sample pixels are taken from each image. The samples are either weighted or un-weighted. In the weighted form, an equal number of vessel and non-vessel pixels are used as input data. These samples are fed to a classifier which is then applied to the 20 test retinograms and the vessel and non-vessel pixels of each retinogram are predicted. The receiver operating characteristic (true positives vs false positives – ROC) is plotted for all the images in the test set, by thresholding at a series of levels and comparing the result to the ground truth. The ROC data is summarised by measuring A_z the area under the curve (an area of 1 indicates perfect classification). We are looking for a single measure for the performance and there will be a trade-off between false positives and false negatives and the area under the curve provides a good measure of performance of the classifier under all conditions without having to use a particular trade-off between false positives and false negatives.

As shown in Table 7, four different approaches to the selection of features used to train the random forest classifier have been considered:

Table 7. Different approaches used to train the random forest classifier.

Name	Method	Dimensions
Full feature vector	6 scales×6 sub-bands × magnitude/phase	72
Maximum sub-band	Only the response with the largest magnitude across sub-bands at each scale	12
Re-ordered sub-bands	Full feature vector, but cyclically reordered so that the maximum response across scales is always first	72
3×3 neighbourhood	Concatenation of all the feature vectors in a 3×3 neighbourhood (pixel values) around the pixel	648

A random forest classifier of 100 trees is initially used to predict the class of each pixel. For the weighted samples, around 3000 vessel pixels and 3000 background pixels have been randomly chosen from each image in the training set – set of 120000 pixels

The magnitude and phase of each pixel has been extracted using the DT-CWT and fed to each classifier to observe the results. The complex value, magnitude and phase separately were used as a feature set and the results have been recorded. Table 8 shows the results for both classifiers and all feature sets.

6.4.2 Results

Table 8 shows the results of applying both the SVM and Random forest classifiers to a single retinogram. The area under the ROCs are recorded.

Table 8. Area under the ROC. Standard errors are in brackets beneath the area under ROCs. Different pixel information used to train the random forest classifier and support vector machine. The standard error is shown in brackets.

	Magnitude & Phase	Complex	Magnitude only	Phase only
SVM	**0.895** **(0.0051)**	0.780 (0.0062)	0.886 (0.0061)	0.546 (0.0061)
Random Forest	**0.957** **(0.0021)**	0.907 (0.0032)	0.874 (0.0058)	0.854 (0.0057)

The SVM classifier does not classify well compared to the random forest classifier. The best results are for the random forest classifier using magnitude and phase representation which is the combination we have used for training the classifiers from this point onwards.

Table 9 compares the area under the ROC curve for each method using a random forest classifier. The best results are obtained using the weighted full feature vector, which gives an average A_z of 0.944.

Table 9. Area under the ROC.

Weighted (equal number of pixels for background and vessel pixels fed to the classifier for training) and un-weighted (random selection of vessel and background pixels fed to the classifier for training).The standard error is shown in brackets.

	Weighted	Un-Weighted
All Sub-bands	**0.944** **(0.0041)**	0.912 (0.0030)
Maximum Sub-band	0.929 (0.0027)	0.902 (0.0020)
Re-ordered Sub-band	0.929 (0.0028)	0.903 (0.0024)
3×3 Neighbourhood	0.936 (0.0031)	0.924 (0.0021)

The best of the above approaches which is the weighted all sub-bands approach has been evaluated using random forest classifiers using a variable numbers of trees. Table 10 shows the results which indicate that as more trees are added, the area under the curve increases.

Table 10. Comparison of results for random forest classifiers with different numbers of trees.

Trees	100	200	300	400	500	600	800	900	**1000**
A_z	0.940 (0.0021)	0.941 (0.0021)	0.942 (0.0023)	0.943 (0.0023)	0.944 (0.0023)	0.945 (0.0023)	0.945 (0.0023)	0.946 (0.0022)	**0.946 (0.0022)**

From Table 10 , it can be seen that a random forest classifier with 900/1000 trees provides for the highest area under the ROC , therefore it has been applied to the centreline probability images and the average area under the ROC is found to be 0.909(±0.0022).

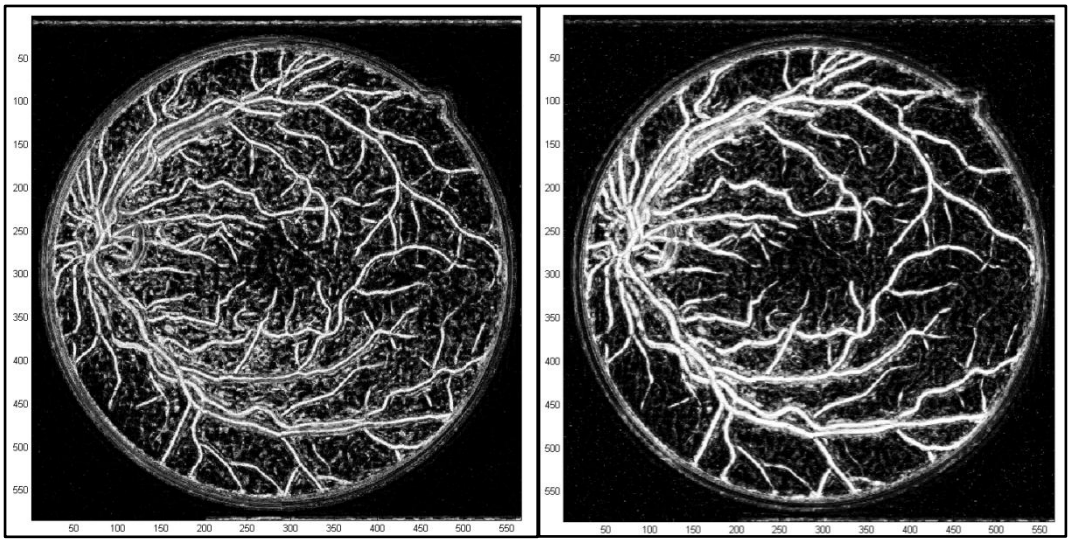


Figure 55 . Probability results of Figure 50.

Left: The centreline pixels of the retinogram are predicted. Right: The vessel pixels of the retinogram are predicted.

Figure 52 shows the probability results of Figure 50 for centreline and All-vessels detection.

In retinogram classification, the false positive rate is high due to the fact that in the retinograms, edge effects cause misclassification of the background as vessels. This can also be seen in the binary images.

6.5 Improving the Results

One of the observations from the results of segmenting the retinograms is that a high proportion of false positives are created around the boundary of the field of view (Figure 52), so methods were sought to eliminate them.

6.5.1 In painting

In order to reduce the edge effects and the difference in contrast between the pixels surrounding the field of view of the retinograms, several in-painting techniques developed by a colleague have been applied to the retinograms prior to vessel detection:

Method 1: In this method intensity dilation is applied using a disk structuring element which makes the pixel values outside the mask similar to those immediately inside the mask. .

Method 2: In this method the image is filled using nearest neighbour interpolation where each pixel outside the mask is assigned the value of the pixel it is closest to within the mask.

Method 3: The image is filled sector by sector with the average of an internal sector. For each sector the outside pixels is filled with the mean of the background pixels obtained from the corresponding sector inside the mask.

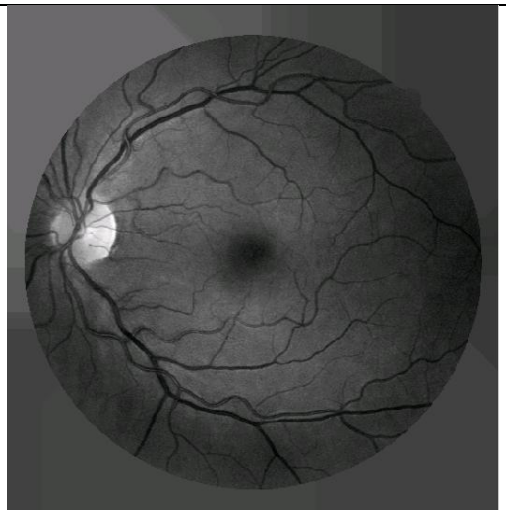
Method 4: Same as method 3 but instead of replacing the pixels outside the mask with a single mean value, a value is randomly selected from the pixels inside the mask near the edge.

Method 5: Fill all the pixels outside the mask with the mean of all background pixels inside the mask.

Figure 53 shows the original image before applying in-painting techniques to it.

Figure 56 . The original raw retinogram.

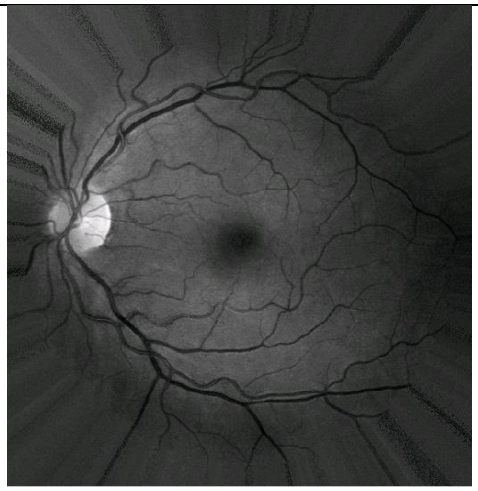
Without In-painting



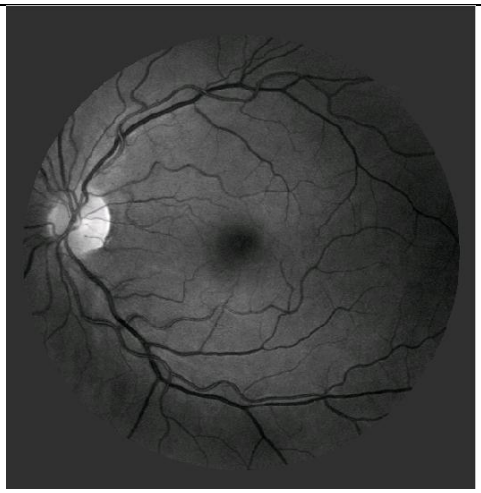
Method1

Method2

Method3



Method4



Method5

Figure 57 . Result images of applying in-painting methods to the retinogram in Figure 53.

Table 12 shows the classification results (Figure 54) achieved when applying the various in-painting methods to the raw retinograms prior to vessel detection and obtaining the area under the ROC. Classification is carried out for both all-vessels detection and centreline detection.

Table 11. Area under the ROC. Comparison of the different in-painting methods.

Method 2 provides for better results as it gives a higher area under the ROC.

Method	**Az – All-vessels**	**Az-Centrelines**
Without In-painting	0.946 (0.0036)	0.909 (0.0022)
Method 1	0.944 (0.0041)	0.907 (0.0026)
Method 2	**0.951 (0.0021)**	**0.915 (0.0025)**
Method 3	0.949 (0.0030)	0.912 (0.0055)
Method 4	0.945 (0.0042)	0.908 (0.0027)
Method 5	0.917 (0.0038)	0.880 (0.0043)

As can be seen, method 2 provides for a higher area under the ROC curve with a low standard error for both all-vessels and centreline detection.

Figure 55 shows an example of a centreline probability image and an all-vessels probability image which are the outputs of the random forest classifier.

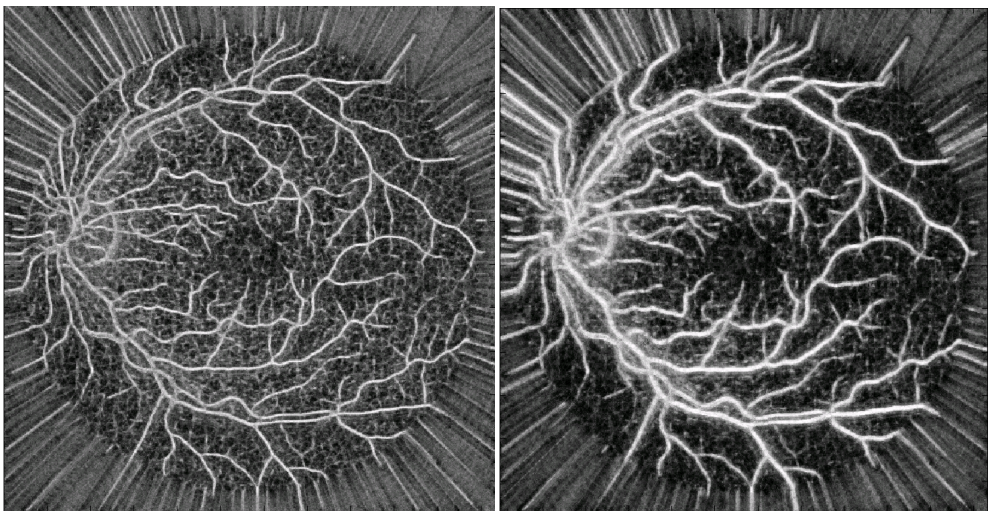


Figure 58 . Probability results of Figure 53.

Left: The centreline pixels of the retinogram are predicted. Right: The vessel pixels of the retinogram are predicted.

6.5.2 Shrinking

To discard the extra edge pixels surrounding the field of view (FOV), we have shrunk this region in the result images and compared it to the ground truth in order to observe how well the classifier is performing. A disk structuring element is used and applied with the erosion operation [128] with a radius of 3 pixels. We have considered two strategies for this:

1. Shrinking Ground Truth - In this approach we have shrunk the ground truth field of view by a few pixels and simultaneously shrunk the result image by the same amount. This way the evaluation would only be based on the pixels inside the FOV.
2. Not Shrinking Ground Truth - In this approach we have carried out the above strategy but with the difference that the ground truth FOV is not shrunk. The reason for this is so that a direct comparison can be made between our method and the other segmentation methods that have used the same images since they have not shrunk the ground truth images. This approach will automatically introduce false negatives but will still eliminate the majority of false positives produced around the edges.

Tables 13 and 14 give a summary of the results for both All-vessels and centreline detection.

Table 12. Area under the ROC. Shrinking methods using classifier of 1000 trees applied to detect All-vessels.

No. of pixels shrunk Radius of disk structuring element	Shrinking of GT mask Average area under ROC	Shrink result and NOT GT mask Average area under ROC
0	0.951 (0.0021)	0.951 (0.0021)
1	0.951 (0.0022)	0.950 (0.0021)
2	0.951 (0.0024)	0.949 (0.0020)
4	0.952 (0.0020)	0.947 (0.0036)
8	0.953 (0.0026)	0.934 (0.0021)
16	0.953 (0.0030)	0.900 (0.0023)
32	0.953 (0.0026)	0.827 (0.0055)

Table 13. Area under the ROC. Shrinking methods using classifier of 1000 trees applied to detect centrelines.

No. of pixels shrunk Radius of disk structuring element	Shrinking of GT mask Average area under ROC	Shrink result and not GT mask Average area under ROC
0	0.915 (0.0045)	0.915 (0.0045)
1	0.915 (0.0021)	0.915 (0.0044)
2	0.915 (0.0023)	0.914 (0.0035)
4	0.916 (0.0022)	0.911 (0.0026)
8	0.917 (0.0024)	0.899 (0.0082)
16	0.916 (0.0025)	0.867 (0.0074)
32	0.915 (0.0024)	0.803 (0.0245)

As the results shows, when shrinking both the result image and the ground truth, the area under the ROC improves.

6.6 Non maximal Suppression

To improve the results for centreline detection we applied non-maximal suppression.

In non-maximal suppression (NMS), a local maximum search is carried out within the image. Each neighbourhood region is considered to find the local maximum. In a 1D a $(2n+1)$ pixel neighbourhood is a set of $(2n+1)$ pixels centred at $(N+1)$. In 2D a $(2n+1)$ x $(2n+1)$ neighbourhood is a region of $(2n+1)$ x $(2n+1)$ pixels centred at (n+1, n+1). -When the search is done, the pixels that are not part of the local maxima are set to zero and suppressed [129].

In Figure 56, the vessel probability of the centre pixel is considered, if it is lower than that of a neighbour in a direction across the vessel, then its magnitude is set to zero. If it has a higher value than the surrounding pixels, it is left unchanged. To improve the results for centreline detection we applied non-maximal suppression. Since NMS requires vessel orientation, we applied random forest regression [130], implemented by a colleague to get an orientation for each pixel in the probability images, using the same feature vector as for classification. Table 9 shows results of applying NMS to centreline probability images.

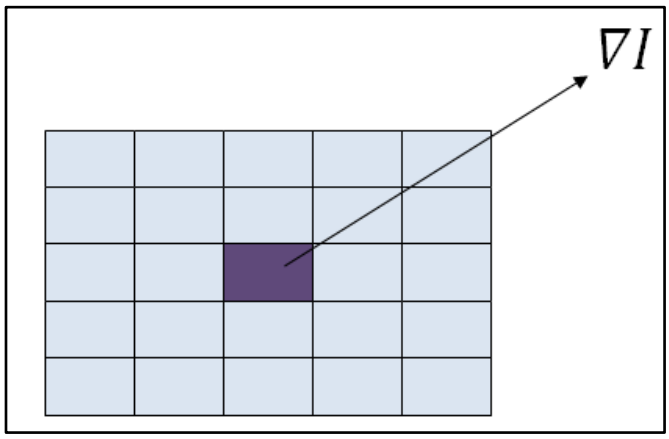


Figure 59 . Illustration of a pixel in its neighbouring region and the direction of gradient.

We applied random forest regression trees [130] implemented by a colleague to get an orientation for each pixel in the probability images using the same feature vector as for classification. Table 15 shows results of applying NMS to centreline probability images.

Figure 57 shows an example of applying Regression and NMS to a retinogram. The orientations are displayed on the same image. Vessels with the same orientation are displayed in the same color.

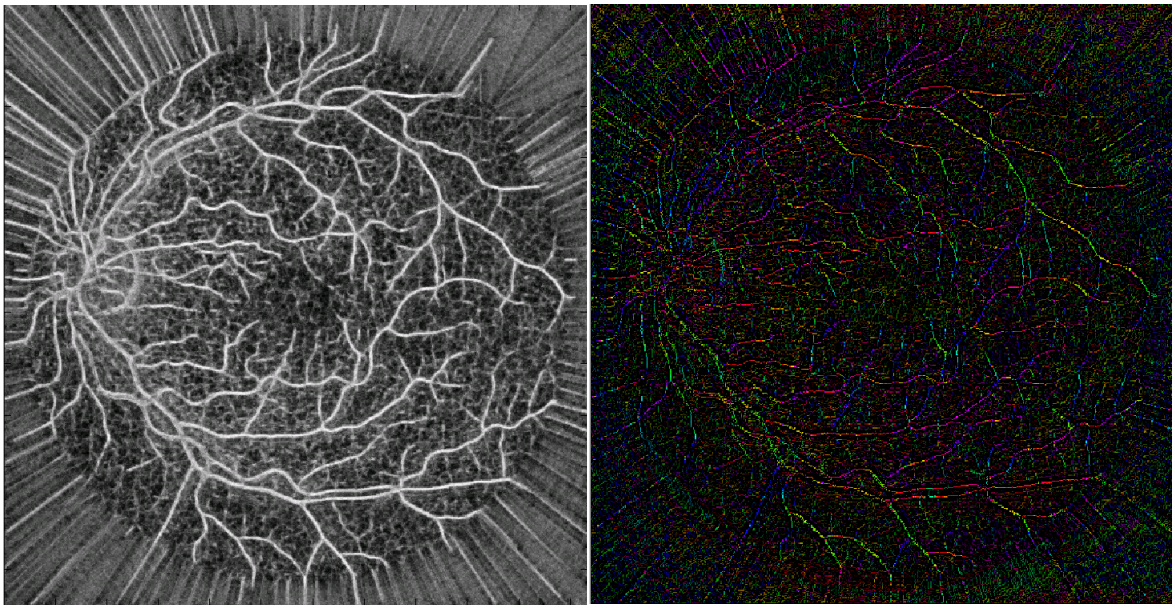


Figure 60 . Result of applying NMS.

NMS is applied to the centreline probability image on the left. The colours in the NMS result (right) show the orientations of the vessel pixels. Each colour shows a different orientation.

Table 14. Results of applying NMS to centreline probability images. Area under the ROC. Standard errors are shown in brackets.

Weighted centerline (Average area under the ROC)	0.915 (0.0025)
After NMS (Average area under the ROC)	**0.926** **(0.0022)**

6.7 Discussion

From the synthetic bar experiments as expected, the performance of both classifiers becomes poor with modest levels of noise. Referring to the mean values, random forest produces higher results than the SVM for all bar angles and widths. As Tables 3 and 4 showed, the classifier performance for narrow bars is good but degrades as the bar

becomes wider. Although the area under the ROC is not close to 1, the results show that both methods degrade significantly with the addition of noise.
For bar widths that are higher than 9 pixels, a dramatic drop can be seen in the area under the ROC which should not affect retinogram classification since the widest vessels in the retinograms are 7 pixels wide and the narrowest ones are 3 pixels wide.
In retinogram classification, from Table 3 we found that using magnitude and phase information provides for better results than the Magnitude Only or Phase only or complex versions. Hence we used Magnitude and phase for the rest of the experiments and tried all feature vectors and found that the weighted all sub-bands approach provided for better results. We found that 1000 trees provided better results than 100 trees so we applied a random forest classifier of 1000 trees to extract centrelines from the retinograms. We observed edge effects and to reduce them we applied several in-painting methods and found Method 2 to give a higher area under the ROC and to also remove edge affects around the FOV .We also mask the results using mask images of each retinogram provided in the database. This removes everything around the mask and out of the FOV and evaluates structure within the FOV.
We also applied NMS and found it to improve centreline detection which we incorporated in our state of the art method in chapter 8.

6.8 Summary

From comparing the results, the random forest classifier takes less time to train and provides high accuracy with less training data. The SVM is adaptable to larger training data but will consequently take more time to train and produces less accurate results.

In-painting and shrinking were pre-processing methods applied in order to improve results. The best in-painting method was chosen and applied to the retinograms. Also, NMS was applied to narrow the vessels and provide for more accurate centreline detection. The results are improved after applying the NMS. The next chapter describes some post-processing methods that can be used to improve results, applied to the best method found in this chapter. The best settings are found to be weighted-full feature vector-inpainting using method 2, random forest, 1000 trees, detecting centrelines using regression and NMS.

CHAPTER 7: Post Processing

7.1 Introduction

The previous chapter described methods for obtaining vessel probability independently at each pixel. This chapter looks at whether the results can be improved by taking contextual information into account. The techniques discussed are mathematical morphology and hysteresis thresholding which are standard methods from literature and a novel tracking method. These techniques are applied to the probability images and results are compared.

7.2 Mathematical Morphology

Mathematical morphology is a tool for extracting shapes and objects from images [128]. It deals with the shape of objects within an image and image features can be altered for processing.

When a structuring element is applied to an image, a morphological transformation takes place where the relation between the structuring element and image pixels extracts image features. A structuring element, a set of point coordinates set to be a particular shape (e.g. 3×3 square) is applied to every point in the input image and its intersection with the image structure is examined [128].

Fundamental binary operators applied to images are Dilation, Erosion, Opening and Closing [131]. The latter three are used in this project and described below.

7.2.1 Erosion

Erosion is a basic morphological operator that is applied to binary images in order to eliminate boundaries of regions of foreground pixels [131].

For example, Figure 58 shows an original object and the result after erosion using a structuring element of disk with radius 15.

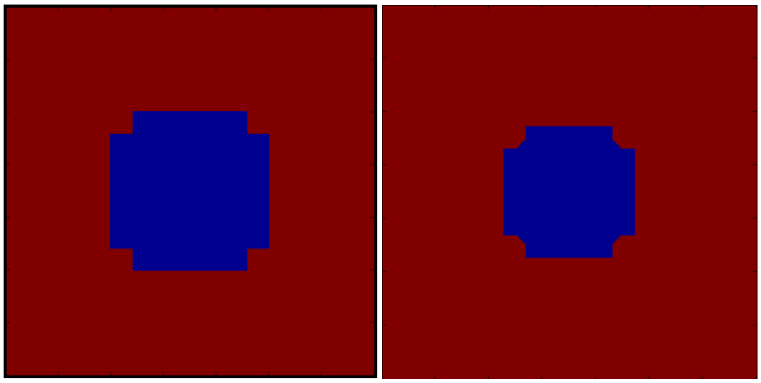


Figure 61 . (a) Original Image. (b) Result after erosion

Figure 62 . Disk structuring element.

7.2.2 Dilation

Dilation is a basic morphological operator that is applied to the foreground pixels in images in order to expand boundaries of regions and borders which belong to objects in the image. Therefore foreground pixels will increase and any holes within the region will shrink. Gaps and areas that are separated become connected [131].

For example, Figure 60 shows an original object (left) and the result after dilation (right) with a disk structuring element with radius 15.

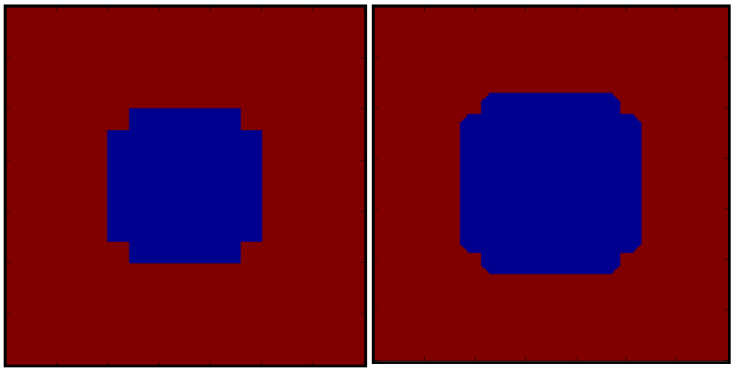


Figure 63 . (Left) Original Image. (Right) Result after dilation.

7.2.3 Closing

Closing is defined as dilation followed by erosion using the same structuring element. Closing improves connectivity of objects in the same neighbourhood and their smoothness as well as filling up small holes [128]. Figure 61 shows closing applied to an image of circles.

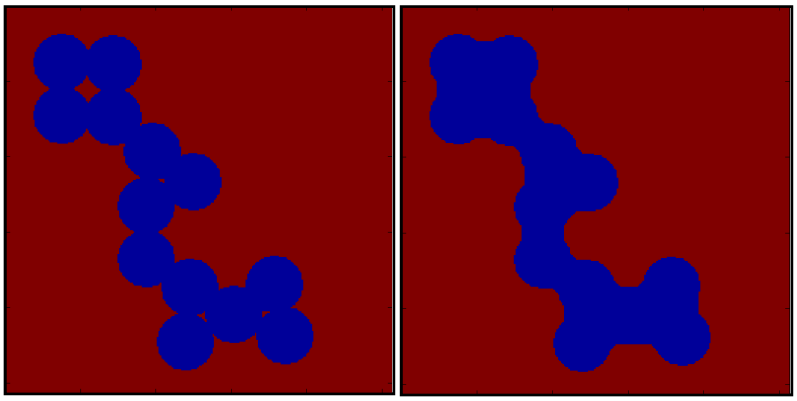


Figure 64 . Closing applied to the image of circles in the left of this Figure.

7.2.4 Applying Morphology to Retinograms

Binary closing was applied to the resulting binary images after thresholding at each of the levels used to construct the ROC. The structuring element that we used was a disc with radius of 3. Using a line structuring element would improve connectivity for only vessels that had the same angle and the line structuring element but would not affect the majority of the vessels that had other angles.

Closing with a disc structuring element was applied to the probability images in order to bridge gaps between disconnected vessels particularly fine vessels and to also eliminate the false positives in the background.

Table 16 shows the average area under the ROC as a result of applying the mathematical operations closing to the 20 centreline and all-vessels probability images obtained as described in the previous chapter.

Table 15. Area under the ROC with and without closing applied to detect All-vessels and centrelines.

	Centrelines	All-vessels
Without Closing	0.915 (0.0021)	0.951 (0.0022)
With Closing	**0.943** **(0.0024)**	**0.951** **(0.0022)**

From the results it can be seen that the closing operation has slightly improved the area under the ROC for the centreline images. The standard error shows that there is statistically significant difference.

An example of applying closing to a binary image of a vessel in a retinogram is shown in Figure 62. It can be seen that more vessel pixels are connected together by applying closing.

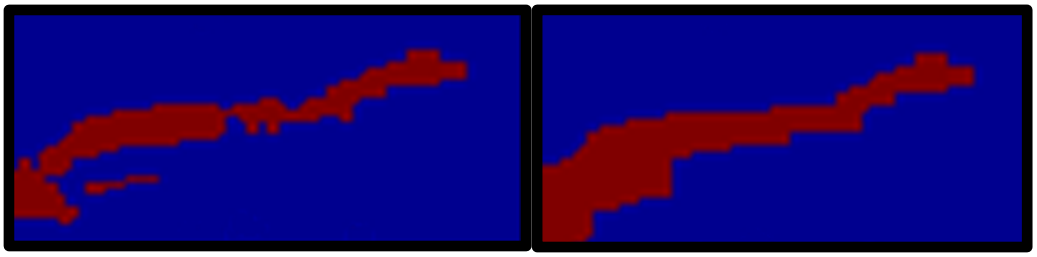


Figure 65 . (a) Original vessel. (b) Result after applying closing.

7.3 Tracking

Tracking applies a local search to find the location of features in images. We described the tracks of vessels in the probability images in chapter 6 obtained from the classifier output.

The starting point for the tracking process is selected from pixels above a selected threshold in the probability image. This is done so that the starting points would be likely to be on vessels. A subset is then randomly chosen from these starting points. These starting points are used on the original image and tracking is carried out on the original image. The probability image here is only used to define the starting points.

The next step is applying the dual-tree complex wavelet transform to the original image at 6 levels and then interpolating. For each starting point, the 3 by 3 neighbourhood region is obtained and the magnitudes of the 6 sub-bands at that point are considered.

We obtain new probability images and define a probability for each of the pixels surrounding the starting point so that the probability of that pixel being a vessel is computed, the following algorithm is used:

Resulting magnitudes from the DT-CWT are divided by the sum of all magnitudes belonging to the sub-bands.

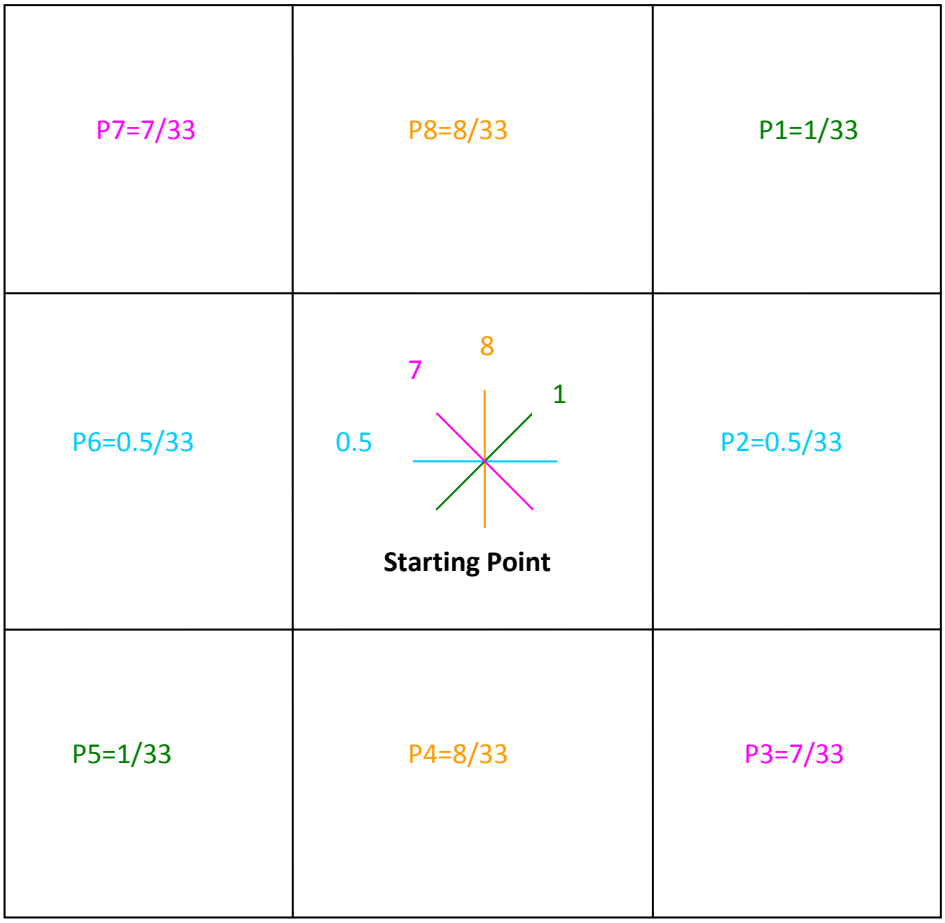


Figure 66 . Probability assigning process.

Each pixel surrounding the starting points is assigned a probability based on the magnitudes of each sub-band at the starting point.

Figure 63 shows the probability assigning process. When the probability of each pixel in the neighbourhood region is obtained, a random number is generated by throwing a die. Depending on which probability interval the number falls within, the operator moves to the

next pixel in the neighbourhood region (either pixels in the probability interval). This process continues until the vessel pixels in the retinogram are found.

The neighbourhood pixels are found based on 3 different methods. In the **random scale** method a random scale is chosen for each neighbouring pixel then the magnitudes of all sub-bands are found based on choosing the largest response, these are then allocated to the neighbouring pixels, then the probabilities are obtained based on Figure 63 and then depending on which probability interval the number falls within the next pixel is chosen. Then this process continues by moving onto the next pixel which would be the next tracked point and then randomly choosing another scale.

In the **maximum scale per track** method, the maximum magnitude (sub-band) across all sub-bands and scales are obtained and then the magnitudes are allocated to the neighbouring pixels according to Figure 63 and then depending on which probability interval the number falls within the next pixel is chosen. So we add up all responses (sub-bands) at each scale and then get the largest response.

In the **maximum image** method, the maximum magnitude across scales for each sub-band is found and assigned to each neighbouring pixel and then the probabilities are obtained (based on Figure 63), then a die is thrown and the next point is determined by the interval the die falls within. So we choose the scale at which we get the largest response.

7.3.1 Results of Tracking

Table 17 draws a comparison between the areas under the ROC curve obtained for one test retinogram for the three different tracking methods. Different thresholds are compared in order to find the best one.

Table 16. Results (Az) of different tracking methods applied to a retinogram at different thresholds.

Threshold	Random Scale	Maximum Image	Max scale per track
0.1	0.925	0.923	0.921
0.2	0.924	0.923	0.921
0.3	0.925	0.923	0.921
0.4	0.924	0.923	0.921
0.5	0.925	0.924	0.921
0.6	0.925	0.924	0.921
0.7	0.925	0.923	0.921
0.8	0.924	0.923	0.922
0.9	**0.926**	0.923	**0.922**

Table 18 compares the results of applying the tracking methods to all 20 retinograms at the best threshold (0.9). The random scale procedure provides for better results.

Table 17. Area under the ROC after applying tracking methods to all probability images to detect all vessels. Standard errors are shown in brackets.

	Random Scale	Maximum Image	Max scale per track
Average area under ROC	0.912 (0.0026)	0.910 (0.0026)	0.907 (0.0027)

Figure 64 shows the tracking process on a retinogram. The red dots show the starting points and the yellow pixels show the tracked vessel pixels.

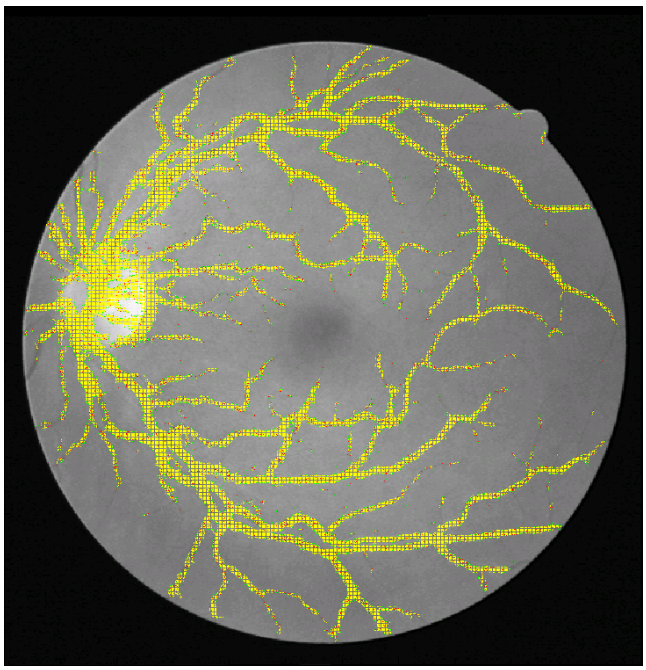


Figure 67 . Tracking the vessel pixels of a retinogram.

The red dots show the starting points. The yellow dots show the tracked pixels.

7.4 Hysteresis Thresholding

7.4.1 Hysteresis Thresholding using Neighbours

The probability images contain a lot of background noise. In order to reduce the amount of false positives, hysteresis thresholding [131] can be applied where two thresholds are required. One threshold contains a high value and the other contains a low value. Hysteresis thresholding is performed by the following: starting points are computed as the pixels above the high threshold on the initial probability image (chapter 6). Then the 3 by 3 neighbourhood regions of the starting points are obtained by dilating the resulting high thresholded image and then multiplying this with the low thresholded binary image. Thus any pixel in the low threshold image, adjacent to one in the high threshold image, is set to one.

Any pixel value below the low threshold is set to zero. This effectively eliminates any sparse white pixels on a black background therefore the false positives in the result image are significantly reduced. 100 low thresholds were used from 1 to 100 in steps in order to

generate an ROC. For the high thresholds we used 0.5 to 1 in steps of 0.1. This process is repeated until convergence is achieved.

Table 19 shows the area under the ROC curves for 4 high threshold values and 100 low thresholds.

Table 18. Area under the ROCs using a neighbouring strategy for hysteresis thresholding.

Hi_Threshold	**0.6**	**0.7**	**0.8**	**0.9**
A_z	0.938 (0.0031)	0.950 (0.0032)	**0.956** **(0.0031)**	**0.956** **(0.0031)**

7.4.2 Hysteresis Thresholding using Morphology

We used an alternative implementation that allowed us to look at neighbourhoods of different size using morphology. The same technique (hysteresis thresholding) has been applied using morphological processes. A high threshold is chosen, and then structuring elements are applied to the low thresholded binary images. In order to connect linear structures of unknown orientations, we applied disks of different radius. As shown in Table 20, 4 high thresholds are chosen (0.6, 0.7, 0.8 and 0.9), then disks of radius between 1-16 are applied to the low thresholded image each time it is thresholded. The area under the ROC is obtained for each high threshold.

Table 19. Area under the ROCs using morphological processes for hysteresis thresholding. Standad errors are presented in brackets.

Hi_Threshold	0.6	0.7	0.8	0.9
Disk –radius=1	0.938 (0.0026)	0.950 (0.0026)	0.955 (0.0027)	0.955 (0.0028)
Disk –radius=2	0.938 (0.0026	0.950 (0.0026)	0.956 (0.0027)	0.957 (0.0028)

	)			
Disk–radius= 3	0.938 (0.0026)	0.950 (0.0026)	0.956 (0.0027)	0.957 (0.0028)
Disk –radius=4	0.938 (0.0026)	0.951 (0.0025)	0.957 (0.0025)	0.958 (0.0028)
Disk–radius= 8	0.938 (0.0026)	0.951 (0.0025)	0.957 (0.0025)	**0.960 (0.0027)**
Disk –radius=16	0.939 (0.0024)	0.952 (0.0028)	0.958 (0.0027)	0.959 (0.0023)

7.5 Summary

From the above experiments, it is concluded that closing improves results for both All-vessels and centreline detection. When tracking is applied to the retinograms, it provides for good results although they are not as good as the results obtained after applying hysteresis thresholding and closing. The presence of noise affects the tracking performance and produces lower results. The best result is achieved after applying hysteresis thresholding with a disk structuring element of 8 pixels. The area under the ROC of 0.960(±0.0027) is achieved.

CHAPTER 8: Evaluation Strategies

8.1 Introduction

Niemeijer et al. [2] describe an evaluation methodology for retinal vessel segmentation and compare several approaches to retinal vessel segmentation experimentally, using the publicly available DRIVE database. Their results show Staal's method [103] to be the best of those tested, using an approach to evaluation that we will refer to as the standard method. In this approach the detected vessel map is compared with a ground-truth vessel mask. For detection methods that give a continuous response related to vessel probability at each pixel or voxel, results are often presented as receiver operator characteristic (ROC) curves, obtained by applying a series of thresholds to the response image(s), (alternatively the less informative total error rate or kappa statistic are used). We argue that the standard pixel-wise approach of TPs and FPs produce an ROC that is uninformative, because the results tend to be dominated by the detection of large vessels and fail to capture adequately whether or not finer, lower contrast vessels have been detected successfully [4]. We are looking for evaluation methods that provide more information.

We describe four strategies for evaluating vessel detection, three of which are new, and use them to compare the performance of our new algorithm for detecting blood vessels in retinograms to existing methods. The four evaluation strategies are:

1. DRIVE method
2. Centrelines only method.
3. Modified centrelines method
4. Standardised tolerance centreline method.

In method 2, detected and ground-truth centrelines are compared. 3 & 4 are hybrid strategies that define true positives similarly to the centreline method and false positives similarly to the standard method.

8.2 Evaluation

Here we describe existing evaluation strategies that are based on area overlap or boundary localisation.

8.2.1 Evaluation using area overlap

There are numerous methods for assessing the accuracy of a segmentation with respect to a ground truth segmentation. The simplest involve a measure of the area overlap between the test segmentation and the ground truth, normalised in some way. Different choices of normalisation result in the Tanimoto [132], Dice [133], and Jaccard [134] indices, or Cohen's Kappa [135]. For a detection problem (such as vessel detection) it is also common to record separately the true positive rate (proportion of ground truth vessel pixels overlapping detected pixels) and false positive rate (proportion of ground truth non-vessel pixels overlapping detected pixels). Alternatively, where error rates are low, it can be more helpful to record FNs rather than TPs. Where the detection method gives a continuous output (e.g. probability of vessel), it is common to apply a series of thresholds and plot true positives against false positives, giving a receiver operating characteristic (ROC) curve. This is the approach recommended for evaluating vessel detection methods using the DRIVE retinogram database [2].When a single performance measure is required, the area under the ROC is often used [4].

8.2.1.1 Tanimoto

The Tanimoto coefficient [132] measures the overlap between two sets. It is calculated by obtaining the size of the intersection of the set and dividing it by the size of their union.

If the two sets are C and D, the Tanimoto coefficient is calculated by:

$$J(C, D) = | C \cap D | / | C \cup D | \quad (23)$$

If C and D have no common members, the index is zero. If they are identical, the index is 1. Therefore when the index is close to 1, it indicates a better agreement.

For example, if the two blood vessel networks in Figure 67 are being compared, and if N_1 represents the number of vessels in network1 and N_2 represents the number of vessels in network 2, and N_{12} is the number of vessels that are common in both networks, then the Tanimoto coefficient would be:

$$T = N_{12} / (N_1 + N_2 - N_{12}) \quad (24)$$

Figure 65 shows networks 1 and 2:

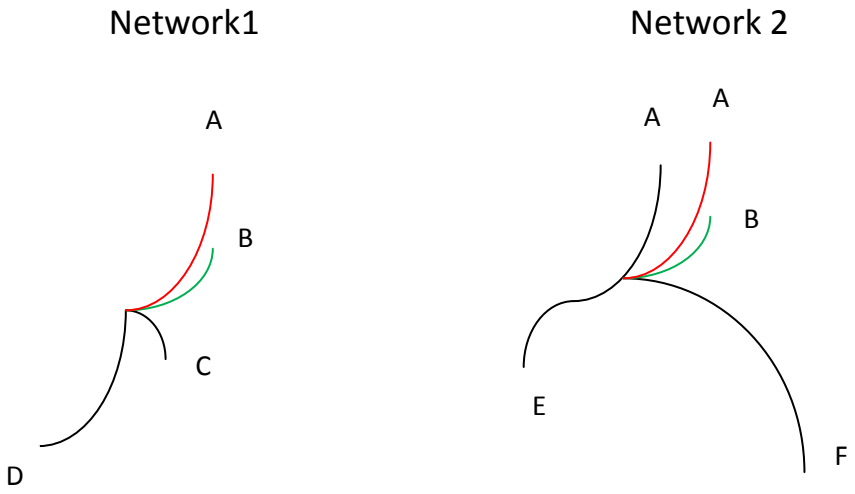


Figure 68 . Blood vessel network examples.

Table 21 compares the two networks in order to find their Tanimoto measurement:

Table 20. Networks 1 and 2 are compared to obtain their Tanimoto overlap measurement [141].

Fragments	A	B	C	D	E	F
Network1	1	1	1	1	0	0
Network2	2	1	0	0	1	1

$$T = \frac{N_{12}}{N_1 + N_2 - N_{12}} = \frac{2}{4+5-2} = 0.3 \quad (25)$$

8.2.1.2 Dice

The Dice coefficient [133] is obtained by dividing the size of the overlap of the two sets by the total size of the two sets:

D (A, B) = 2 |A ∩ B | / (|A|+|B|)

A Dice coefficient of 1 indicates perfect agreements and 0 indicates no overlap.

For the above example it would be:

$$D = \frac{N_{12}}{N_1+N_2} = \frac{4}{5+6} = 0.4 \quad (26)$$

8.2.2 Evaluation using boundary localisation

A second class of evaluation methods use a measure of boundary localisation – for example the distribution of distances between each point on the ground truth boundary and the nearest boundary point on the test segmentation [136, 137, and 138] (or vice versa – typically summarised by the mean distance). Although we could find no specific reference to it, there is an obvious adaptation of this method to vessel detection, where the distances between ground truth and test segmentation centrelines are measured, and we are sure this must have been described. The disadvantage of this approach is that it does not satisfactorily capture the trade-off between true and false positive detections, which is an important practical consideration.

8.2.2.1 Kappa

Kappa is a measure of agreement between two observers to see how often they agree in their interpretations. The equation for kappa is

Kappa = (Observed Agreement-Expected Agreement)/(1-Expected Agreement) (27)

To calculate the observed agreement, the agreement is calculated:

$$\text{Observed Agreement} = (a+d)/N \quad (28)$$

where (a) is the agreement between the two observers and (d) is the disagreement between them.

To calculate the expected agreement, the expected agreement is calculated:

$$\text{Expected (random) Agreement} = (\text{Expected (a)} + \text{Expected (d)})/N \quad (29)$$

For example, suppose two doctors were specifying whether patients had diabetes or not. Table 22 shows the results:

Table 21. Data showing the classification of patients into diabetic and non-diabetic by two doctors [143].

		Doctor 2	Doctor 2
		Diabetics	Non-Diabetics
Doctor 1	Diabetics	20	5
Doctor 1	Non-Diabetics	10	15

They both agreed, 20 times and both disagreed 15 times therefore

$$\text{Observed Agreement} = (a+d)/N = (20+15)/50 = 0.7 \quad (30)$$

To calculate the random agreement, doctor 1 specifies that 25 of the patients are diabetic (Table 22, row3) and 25 are non-diabetic (Table 18, row4) therefore $25/50 \times 100 = 50\%$ are specified as diabetic patients. Doctor 2 specifies that 30 are diabetic (Table 22, column 3) and 20 are non-diabetic (Table 22, column 4). Therefore $30/50 \times 100 = 60\%$ are specified as diabetic patients.

So both agree on diabetic patients by $0.50(50\%) \times 0.60(60\%) = 0.30$ and on non-diabetic patients by $0.50(100\%-50\%) \times 0.40(100\%-60\%) = 0.20$. Therefore, the total expected agreement will be:

$$\text{Total random (expected) agreement} = (0.2+0.3) = 0.5 \quad (31)$$

Therefore

Kappa: (Observed Agreement-Expected Agreement)/(1-Expected Agreement)

Kappa= (0.7-0.5)/(1-0.5)=0.4 (32)

8.2.2.2 Maximum average accuracy

The maximum average accuracy (MAA) is calculated by establishing the point on the ROC where FN + FP is minimised, and taking that value averaged over the test set. (Refer to section 5.4).

8.3 Evaluation methods on retinograms

Given these observations, we adopt an overlap-based ROC approach. However, we will argue below that the standard overlap approach recommended for evaluating performance on the DRIVE database has unfortunate characteristics.

We have obtained from the authors probability images generated by the state-of-the-art methods of Staal et al. [103] and Niemeijer et al. [2], the first and second authors of the DRIVE database. In order to assess the performance of the proposed evaluation methods, they have been applied to probability images obtained from our method and those provided by Staal et al. [103] and Niemeijer et al. [2]. For centreline detection, NMS has initially been applied to the probability images.

In the following sections different methods for evaluating vessel detection are explained. To explain our evaluation strategies we use the following terms: GT = ground-truth mask; GTS = skeleton of ground-truth mask; DV = thresholded vessel detection map; DC = thresholded vessel centreline detection map; DE = result of non-maximally suppressing DC thresholding and then dilating; TP = true positive map; FP = false positive map; / is the logical ‘not’ operation.

8.3.1 DRIVE method

The DRIVE method is the primary approach used to compare methods using the DRIVE database [2]. At a given threshold, the true positive and false positive fraction are computed as

$$TP = AND(DV, GT) \quad FP = AND(DV, \overline{GT}) \tag{33}$$

where DV is the detection result. In Figure 66, two vessel detection results (A and B) are shown for a region of a retinogram which contains both thick and thin vessels. Result A shows that the thick vessel has been completely detected but the thin one hasn't been detected. Result B shows that the thick vessel has not been completely detected and the fine vessel is partly detected. The colour-coded images on the right correspond to the colours in equation 53. The true positives are red, the false negatives which correspond to the undetected vessels are light blue and the false positives are shown in green. Result A is rated better than Result B because fewer vessel pixels have not been detected and the false positives are equivalent in both results. But result B should be preferred since a proportion of the thin vessel has been detected whereas in A, none of it has been detected. This illustrates that the DRIVE method gives a performance measure that is biased towards the detection of thick vessels [4].

A good vessel detection evaluation method should separate out the issue of whether vessels have been detected from that of whether their widths have been estimated accurately. The evaluation methods described below aim to achieve this.

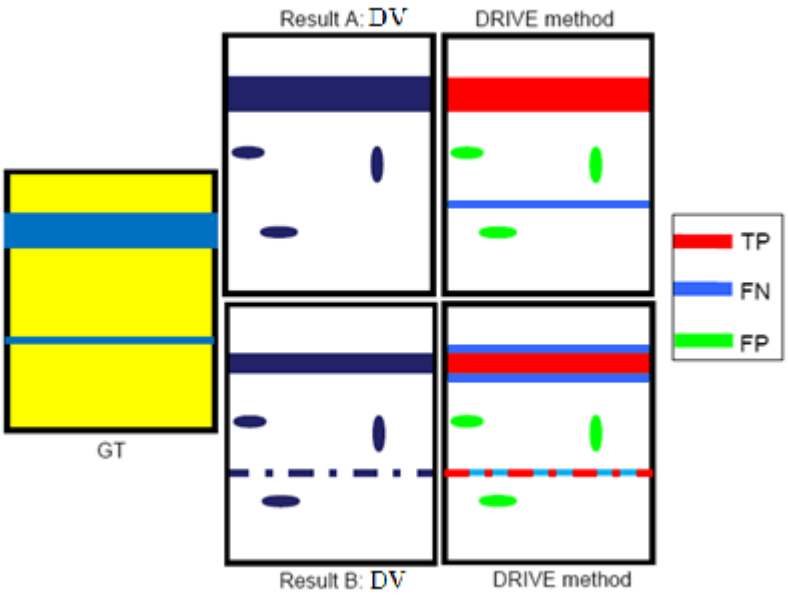


Figure 69 . DRIVE method [4].

Left) Ground truth (Middle) Results of vessel detection (classifier outputs). (Right) Red shows the true positives, green shows the false positives and blue shows the false negatives (pixels that have not been detected). The top right Figure shows the colour coded version of result A and the bottom right shows the colour coded versions of result B. Result A is rated better than result B [4].

8.3.2 Centrelines Only

The centrelines only method detects only centreline pixels and compares the detection result with a skeletonised version of the original ground truth image.

Thus true positives and false positives are computed as

$$TP = AND(DC, GTS), FP = AND(DC, /GTS) \quad (34)$$

where GTS is the skeletonised ground truth and DC is the centreline detection result. To obtain the centreline probability images, NMS was applied to the All-vessels probability images, to allow a fair comparison of our method with those of Staal's and Neimeijer's, who do not offer a centreline variant.

As shown in Figure 67, this method removes the bias due to vessel width but tends to underestimate performance, since any broadening of the detected centrelines results in an

increase in false positives, whilst small location errors produce an increase in both false positives and false negatives. In both results it can be seen that the centreline detection at the top has been broadened which results in an increase of false positives. This method would rate result A as the best result but result B is preferred since it has detected more vessel centrelines than A, albeit with a small (e.g. 1 pixel) location error which produces false positives [4].

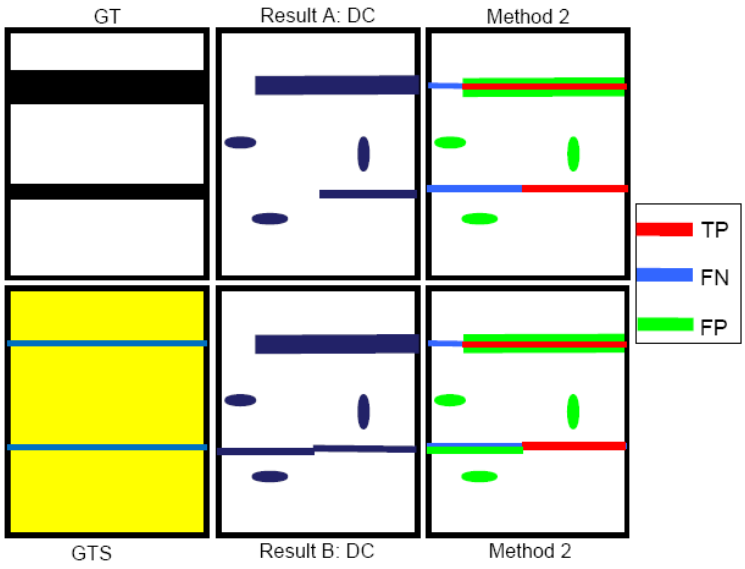


Figure 70 . Method 2 of vessel detection.

(Left) Ground truth (GT) and skeletonised ground truth (GTS). (Middle) DC is the centrelines detection result for result A and B. (Right)Colour coded versions of results. True positives are shown in red, false positives are shown in green and false negatives are shown in blue. This method rates result A as better than B [4].

8.3.3 Modified Centrelines

In this method false positives are only registered for detected centrelines that lie outside the original ground truth. This will not underestimate performance and penalise small location errors as method 2 does. Therefore the bias due to vessel thickness is removed and small location errors are tolerated (because centreline detection typically gives a 'broadened' result), whilst avoiding penalising the broadening (as occurs in the Centreline Method).

Figure 68 shows the results after applying method 3. As can be seen, the false positives around the centrelines have been removed. However, Result B still does not receive true positives for the added detection, and in this case method 3 would rate Result A and B equivalently.

$$TP = AND(DC, GTS), \quad FP = AND(DC, /GT) \tag{35}$$

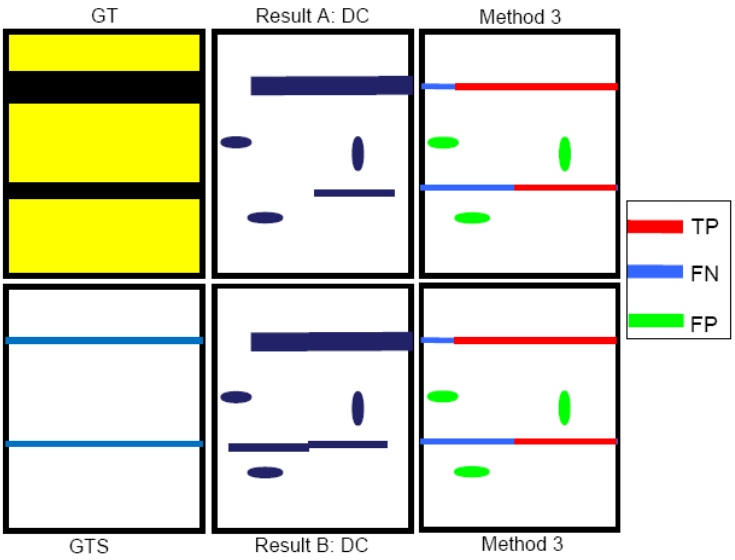


Figure 71 . Method 3 of vessel detection [4].

(Left) Ground truth and skeletonised ground truth. (Middle) Results A and B for centreline detection. (Right) Colour coded versions of results. In this method the vessel pixels that lie within the ground truth vessels that have been misclassified as centreline pixels are not counted as false positives. Therefore, a small displacement is allowed. Result A and Result B would be rated equivalently [4].

8.3.4 Standardised tolerance centreline

In this method, a tolerance is specified and applied instead of relying on broadening of the centreline detection. Non-maximal suppression is used in order to thin the detected vessel centrelines. The pixels are then dilated by a fixed amount (in our experiments we have used a 3 pixel diameter disk). This provides for controlled tolerance to location errors. The true positives and false positives are then obtained from the dilated NMS result. The degree of latitude allowed in the centreline location can be controlled by increasing or decreasing the dilation which is applied to the thresholded NMS responses. This method is shown in Figure 69. By allowing a small tolerance in location, result B now records true positives along the full length of the vessel near the bottom of the image, and as a result, method 4 would rate result B as better than result A as desired.

$$TP = AND(DE, GTS), \quad FP = AND(DE, /GT) \tag{36}$$

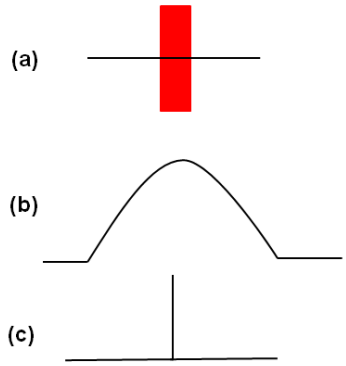


Figure 72 . Non-maximal suppression.

In (a), a vessel is shown in red. In (b), the cross-sectional response of (a) is shown. In (c), the result after NMS shows the exact centreline of the vessel which has been obtained by taking the maximum response at the peak in (b) and setting the surrounding pixel responses to zero [4].

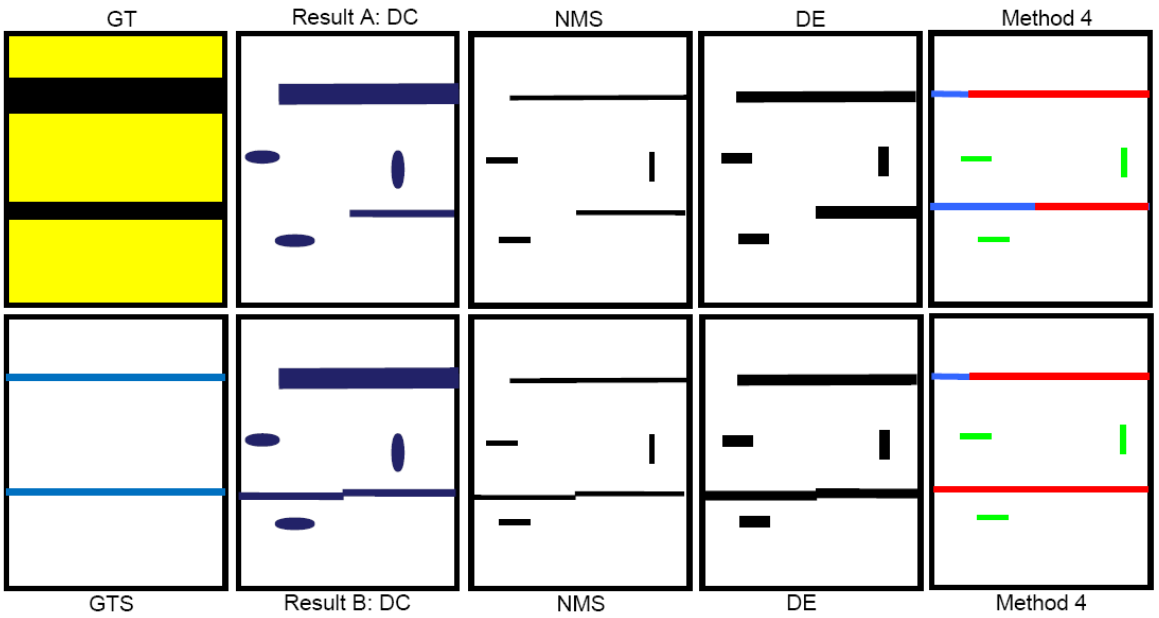


Figure 73 . Method 4 of vessel detection.

NMS is applied to the detection results before they are thresholded and dilated (middle). The centreline detections that have minor displacements will be counted as true positives and thus result B is rated better than result A [4]. DE shows the results after dilation.

8.3.5 Comparison of Methods

Table 23 draws a comparison between the standard method applied to the DRIVE images with earlier methods that have been applied to the same images. Our method is second to Staal's method (highlighted in table). The general evaluation methods used are kappa and maximum average accuracy.

Table 22. Comparison of area under the ROC between methods applied to the DRIVE database. The standard errors (SE) are shown in brackets. Authors provide this on the maximum average accuracy.

Method	A_z	MAA (SE)	Kappa
Human Observer	n/a	0.947 (0.0048)	0.759
RF/DT-CWT	**0.960 (0.0021)**	**0.937 (0.0094)**	**0.699 (0.0022)**
Staal	0.952	0.944 (0.0065)	0.734
Niemeijer	0.929	0.942 (0.0065)	0.714
Zana	0.898	0.938 (0.0077)	0.697
Al-Diri	n/a	0.926 (0.0126)	0.672
Jiang	0.911	0.921 (0.0076)	0.640
Martinez-Perez	n/a	0.918	0.639

		(0.0240)	
Chaudhuri	0.788	0.877 (0.0232)	0.336
All Background	n/a	0.873 (0.0123)	0

Table 24 draws a comparison between the results of applying the proposed evaluation methods to our probability images and those of Staal's and Niemeijer's:

Table 23. Comparison of area under the ROC between our methods applied to the DRIVE database. Our methods and the standard method have all been applied to Staal and Neimeijer's probability images and the results are compared. The standard errors are shown in brackets.

	DT-CWT/RF	STAAL	NEIMEIJERS
Standard method	0.960 (0.0012)	0.952 (0.0011)	0.929 (0.0015)
Centrelines Only(with NMS)	0.608 (0.0025)	0.622 (0.0037)	0.633 (0.0042)
Modified Centrelines(with NMS)	0.612 (0.0026)	0.625 (0.0038)	0.636 (0.0041)
Standard Tolerance(with NMS)	0.962 (0.0014)	0.967 (0.0017)	0.966 (0.0015)
Standard Tolerance on classifier trained on centrelines(with NMS)	**0.977** **(0.0013)**	**0.967** **(0.0017)**	**0.966** **(0.0015)**

Table 24 shows that our method, the standard tolerance on classifier trained on centrelines outperforms Staal's and Niemeijer's methods.

Figure 71 shows sections of the same retinogram at a particular threshold (0.7), which are colour coded to show where the true positives, false positives and false negatives are for each variation of the evaluation methods. Figure 72 shows the average ROCs for all the evaluation methods.

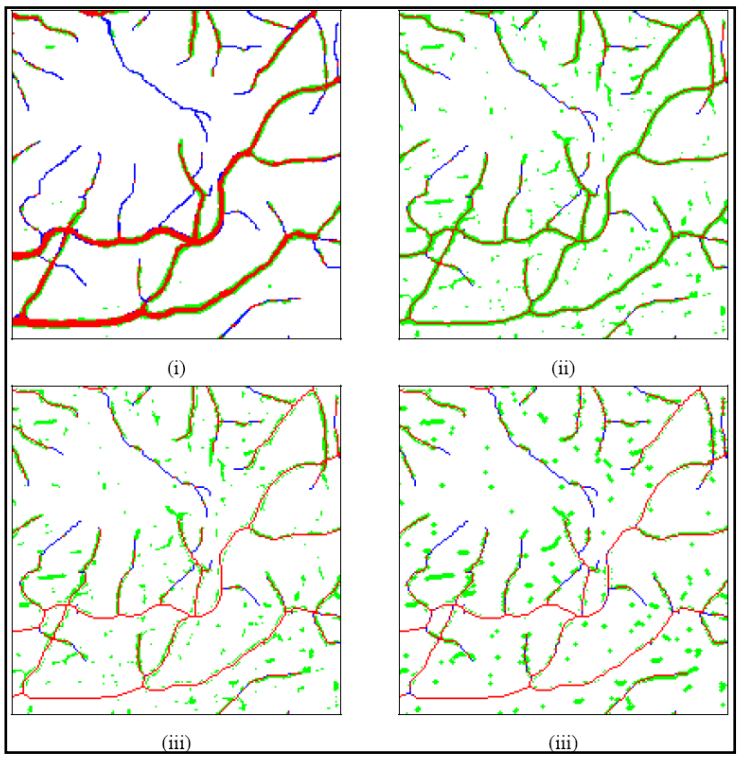


Figure 74 . Images showing the vessel detection methods.

True positives are shown in red, false positives are shown in green and false negatives are shown in blue for a region of a retinogram. (i)DRIVE method (ii) centreline method (iii) modified centreline method (iv) standardized tolerance method. For each evaluation method a threshold of 0.7 was used.

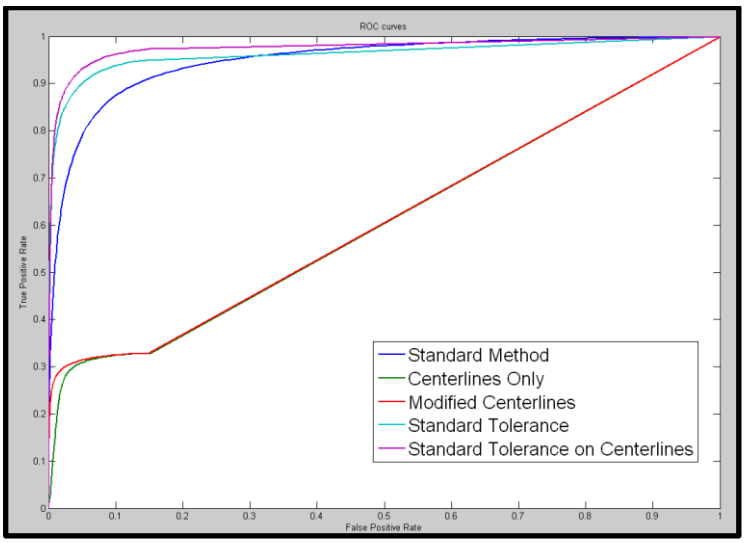


Figure 75 . ROC curves of the evaluation methods applied to the DRIVE database images.

As anticipated, the standard tolerance on centrelines method gives the highest area under the ROC curve. The standard tolerance method is second best. The DRIVE method gives the third largest area under the curve, because it tends to overestimate performance for the reasons explained in 8.3.1. Similarly, method 2 (Centrelines Only) gives the smallest area under the curve, because it tends to underestimate performance for the reasons given in 8.3.2. The modified centrelines method also underestimates performance and lies very close to the centrelines only method.

8.4 Summary

In this chapter the importance of having a good evaluation method has been emphasized. The current standard method also called the DRIVE method tends to be biased towards the detection of large vessels. An evaluation method that is capable of separating the detection of vessel widths from the detection of vessel lengths has been proposed. The evaluation methods have been applied to vessel probability images obtained from the DRIVE database , using our vessel detection method and two existing state of the art methods applied to the same database. Methods 3 & 4 provide performance estimates that depend on the length of vessels detected, irrespective of width, whilst the DRIVE method over-estimates and the centrelines only method under-estimates performance [4].

The results show that the performance of a classification method is dependent on the evaluation method and in the case of retinograms an evaluation method that evaluates the detection of both coarse and fine vessels is desirable which is not the case for the DRIVE method. The standard tolerance on classifier trained on centrelines outperforms Staal's and Niemeijer's methods.

CHAPTER 9: Conclusions and Further Work

9.1 Introduction

In this chapter we summarise the work presented in this thesis and discuss possible areas for future work.

The goal of the work described in this thesis was to develop a method that can detect blood vessels in the back of the eye capturing both the coarse and fine vessels, regardless of their position and orientation in the image. The common drawback of retinal vessel detection techniques is their inability to detect the fine vessels with low contrast. Also, to develop an appropriate evaluation method that evaluates the detection of both fine and coarse vessels since previous evaluation methods are biased towards the detection of thick vessels.

A novel retinal vessel segmentation technique has been developed in this thesis which combines the dual-tree complex wavelet transform with a random forest classifier. The method is applied to retinograms of the DRIVE database. The DT-CWT is used to extract magnitude and phase information from every pixel in the image which is then fed to a random forest classifier. The classifier is used to classify the retinogram pixels into those that belong to vessels and those that belong to background.

The method performs two types of vessel extraction: An extraction which is based on extracting the vessels as a whole throughout the retinogram and a method that extracts the centrelines of the vessels. The All-vessels extraction method provides results that are comparable to the results of previously published methods applied to the same database. The centreline method outperforms the methods of Staal et al. [103] and Niemejier et al. [2].

When applied for segmenting the whole vessel network, it achieves an area under the ROC curve (A_Z) of 0.960(±0.0021) compared to the best result of 0.972 obtained by Wu [96].

When applied for vessel centreline detection, it achieves an A_Z of 0.977(±0.0013) which is higher and statistically significant than the previous best method (Staal et al. [103]) with an A_Z of 0.967(±0.0017). These results are obtained using an evaluation method proposed in this project (chapter 8) that eliminates the drawback of current evaluation methods hence removes the bias towards the detection of thick vessels.

9.2 Further development of our work

The general goal is to develop a generative model which has the capability of modelling both the local (stochastic) and global structures within retinograms and other medical images that contain similar structures as blood vessel networks.

In order to accomplish that, the initial step would be to effectively model the local structure and to understand how it is characterized. The thesis described the state-of-the-art which accomplished the initial step. As for the global structure, a discriminative model should be implemented to be able to distinguish between legal and illegal examples of retinograms. Experiments should initially be carried out on synthetic images to determine if the classifier works efficiently before testing it on retinograms.

9.3 Extracting global structure using phase fields

An extended phase field model [139] has been partly implemented and applied to extract the global structure within retinograms. The model aims to extract vessel-like structures such as roads from images and has two main parts:

1. Prior term
2. Image term

The prior model is independent of an image and has two higher order active contour (HOAC) terms:

1. Non-linear Nonlocal (HOAC) Term
2. Linear Nonlocal HOAC Term

Both parts consist of a standard model which is derived based on phase field energy terms (E0 + ES). A phase field is a function which defines a region via a threshold. By differentiating this function, the boundaries of that region can be identified since the field has two distinct values (+1 and -1) in each of its phases and a smooth change between both

values in the zone around the interface which corresponds to the boundary area in the region.

The standard model consists of the two terms E0 and ES. E0 derives an energy that is based on the gradient of the phase field and a smoothing function. ES is based on the interaction between different points in the region, hence when it is added to E0, it drives the evolution of the model.

The standard HOAC model provides a poor model of networks with straight narrow branches or highly curved, wide branches. This is because the curve of a branch is dependent on its width.

The Prior model consists of parameters that need to be tuned. The extension to this standard model separates the width and curve analysis and it contains the nonlinear nonlocal HOAC term and the linear Nonlocal HOAC term. They are each added to the standard model.

Once the prior model is developed, it is then applied to an image which in this case would be the vessel probability image and by minimizing it the model gradually evolves to fit to the boundaries within the image. So the lower the overall energy, the better the fit.

The future work would be to complete the phase field model and combine it with the local vessel detection method described in this thesis and to then apply it to the DRIVE images.

Appendix 1: Table of DT-CWT responses

Table 24. DT-CWT results using the Interpolation method. The bar size (sigma) associated with each level can be observed along with the orientations associated to each sub-band of the dual-tree complex wavelet transform.

BAND1				BAND2			
Level	Sigma	Angle	Response	Level	Sigma	Angle	Response
1	1	162	0.3234	1	1	135	0.1426
2	1	159	0.7937	2	1	135	0.6533
3	2	159	0.7934	3	2	135	0.6707
4	4	159	0.8109	4	3	135	0.6677
5	9	159	0.8096	5	7	135	0.6723
6	18	159	0.8104	6	14	135	0.6717
BAND3				**BAND4**			
Level	Sigma	Angle	Response	Level	Sigma	Angle	Response
1	1	108	0.3234	1	1	72	0.3174
2	1	111	0.7937	2	1	69	0.7949
3	2	111	0.7934	3	2	69	0.7937
4	4	111	0.8109	4	4	69	0.811
5	9	111	0.8096	5	9	69	0.8096
6	18	111	0.8104	6	18	69	0.8104
BAND5				**BAND6**			
Level	Sigma	Angle	Response	Level	Sigma	Angle	Response
1	1	45	0.1487	1	1	18	0.3174
2	1	45	0.6569	2	1	21	0.7949
3	2	45	0.6716	3	2	21	0.7937
4	3	45	0.6679	4	4	21	0.811
5	7	45	0.6723	5	9	21	0.8096
6	14	45	0.6717	6	18	21	0.8104

Appendix 2: Retinal Vessel Detection Methods

Table 25. Retinal vessel detection methods.

HL=Higher-level, A=Author, Ct=Citation, Yr=Year, CB=Classification-based, FB=Filter-based, HB=Hessian-based, WB=Wavelet-base, MM=Mathematical morphology, ALT=Adaptive Local Thresholding, ED=Edge Detection, FC=Fuzzy Clustering, AC=Active contour, NN=Neural network, CL=clustering,DCM=Deformable contour model, GC=Graph Cut, N=None, RG=Region Growing, T=Tracking

HL	A	Ct	Yr	CB	FB	HB	WB	MM	ALT	ED
FC	Kande	0	2010		x					
AC	Zhu	4	2010				x			
AC	Alonso-Montes	7	2008							x
AC	Pang	0	2009					x		
AC+NN	Al-Diri	2	2010							x
CL	Saleh	0	2011		x					
DCM	Espona2	6	2007							
DCM	Espona	0	2008							
FC	Kande1	0	2008		x					
GC	Salazar-Gonzalez	0	2010							
MM	Miri	0	2011		x			x		x
MM	You	1	2011	x			x			
MM	Reza	4	2008		x					x
MM	Akram2	0	2009				x			x
MM	Pourmorteza	0	2008	x						
MM	Akram3	0	2009		x		x			x
MM	Odstrcilik	1	2009		x					
N	Wu2	0	2007		x	x				
N	Rangayyan	3	2007				x			
N	Zana	2	1997		x			x		
N	Niemeijer	**123**	2004	x						
N	Jiang	**136**	2003						x	
N	Chaudhuri	**317**	1989		x					
N	Sekhar	1	2011		x	x		x	x	
N	Amin	0	2011	x			x			
N	Lam	6	2010			x				
N	Oloumi	2	2007	x			x			
N	Bhuiyan	0	2007	x			x			
N	Ricci	**36**	2007	x						
N	Lian	0	2007		x					
N	Adjeroh	2	2007	x	x					
N	Jelinek	**10**	2007	x			x			
N	Yang	4	2008	x				x		
N	Abdurrazaq	0	2008					x		
N	Chui	0	2008						x	
N	Marin	6	2011	x	x			x		
N	Fraz	1	2011		x			x		

N	Soares	127	2006	x			x			
N	Muramatsu	1	2010	x	x					
N	Al-Rawi	26	2007		x					
N	Zhang3	0	2010	x	x					
N	Zhang2	3	2010		x					
N	Cinsdikici	6	2009		x					
N	Kose	0	2011							
N	Salem	1	2008			x				
N	Moghimirad	0	2010			x			x	
N	Azemin	0	2010	x			x			
N	Rothaus	6	2009	x						x
N	Rothaus 2	0	2011	x						
N	Osareh	1	2009	x			x			
N	Carnimeo	0	2009	x	x	x	x			
N	Akram	1	2009				x		x	
N	Chang	0	2009	x						
N	Lam2	14	2008		x					x
N	Chen	3	2008				x			
N	Bankhead	0	2012				x			
N	Balkys	0	2012						x	
N	Zhu2	0	2011		x					
N	Saleh2	0	2012		x			x		
N	Dua	0	2005		x					x
N	Perez-Rovira	0	2011		x				x	
N	Cummings	0	2010							
N	Franchi	3	2009		x					
N	Goatman	1	2011	x						
N	Qin	1	2006				x		x	x
N	Joshi	0	2012		x			x		
N	Kondermann	0	2007	x						
N	Rangayyan2	5	2008				x			
N	Wu3	0	2007		x	x				
N	Zhang	4	2009					x	x	
N	Rezatofighi	0	2009	x						
N	Sinthanayothin	154	1999	x						
N	Siddalingaswamy	0	2009				x			
N	Dittrich	0	2009			x				
N	Cornforth	22	2004	x			x		x	
N	Wang	12	2007							
N	Xu	1	2009	x	x					
N	Amin2	0	2009				x			
N	Harangozo	0	2009	x						
N	Kumar	0	2009		x					
N	Amiri	1	2009					x		
RG	Staal	315	2004	x						
RG	Hoover	290	2000	x	x				x	

RG	Palomera-Perez	0	2010			x				
RG	Mendonca	76	2006	x	x			x		x
RG	Garg	2	2007							
RG	Martinez-Perez	36	1999	x		x				
RG	Martinez-Perez2	50	2007			x				
RG+AC	Al-Diri2	18	2009		x					x
T	Hatanaka	8	2004							
T	Adel	2	2010							x
T	Wu	26	2006				x			
T	Adel2	1	2009							x
T	Yin	1	2010							x
T	Tolias	116	1998	x						
T	Can	148	1999							
T	Li	0	2006				x			
T	Sofka	64	2006		x					
T	Cree	11	2005							
T	Zhou	82	1994							x
T	Patasius	0	2009							
T	Lin	0	2009		x					
T+MM	Vlachos	2	2010	x	x					

Appendix 3: Vessel Detection Methods

Table 26. Vessel detection methods.

A=Author, Ct=Citation, Yr=Year, HL=Higher level, AC=Active contour, LS=Level set, T=Tracking, DM=Deformable model, GB=Graph-based, HT=Hough Transform, MRF=Markov Random Field, MM=Mathematical Morphology, PM=Pattern Matching, RG=Region Growing, T=Tracking, DM=Deformable Model, N=None.

HL	A	Ct	Yr	CB	FB	HB	WB	MM	ALT	ED
AC	Shang	1	2011			x				
AC	Shang1	0	2010	x						
AC	Nain	29	2004		x					
AC	Chung	0	2006							
AC+LS	Gooya1	15	2008							
AC+LS	Sun	0	2012					x		
AC+LS	Tek	7	2001		x					
AC+T	Tang	33	2004							
DM	Valverde	5	2004							x
DM	Shen	2	2011							
DM+LS	Law2	19	2007							
DM+LS	Lorigo	132	2001							
GB	Freiman	5	2009	x						
HT	Yang	33	1997							
LS	Pizaine	0	2011							
LS	Manniesing	47	2006							
MRF	Kubassova	3	2007							
MM	Zana	158	2001		x			x		x
MM	Toumoulin	13	1990					x		
PM	Koese	0	2008		x					
RG	Lemaitre	1	2011		x					x
RG	Kozaitis	0	2011							
RG	Yim	76	2000							x
RG	Cai	0	2006			x				
T	Delibasis	1	2010							
T	Friman	7	2008							
T	Benmansour	2	2011							
T	Lo	0	2010		x					
T	Jelinek	15	2005	x	x					
T	Avants	12	2000							
T	Schaap	7	2007							
T+DM	Toledo	5	2000							
T+DM	Toledo1	0	2000							
T+AC	Zhao	0	2011							
T+RG	Lo2	1	2009							
N	Lathen	1	2010		x					x
N	Agam	49	2005		x					
N	Law	19	2007							x

N	Villalobos-Castaldi	0	2010							x
N	Cooper	0	2010			x				
N	Gooya	3	2007							x
N	Zhao2	0	2011							
N	Foruzan	0	2012			x				
N	Chen	0	2011		x	x				
N	Korfiatis	1	2011	x	x	x				
N	Yu	0	2011	x		x				
N	Qian	16	2009		x					
N	Truc	9	2009		x					
N	Hao	0	2008							
N	Liu	0	2011							
N	Yuan	1	2008		x	x				
N	Aylward	248	2002							
N	Sun2	1	2011		x			x		
N	Nekovei	47	1995	x						
N	Leandro	2	2003	x	x		x			
N	Zhang	1	2005		x					
N	Zhang1	0	2007							
N	Lorenz	18	1997		x	x				
N	Lorenz1	88	1997		x	x				

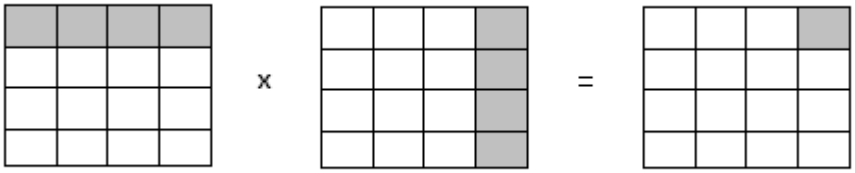
Appendix 4: 2D Oriented Dual-Tree CWT

The dual-tree CWT produces oriented wavelets [6] by:

If $\psi(x)$ is a complex wavelet given by $\psi(x) = \psi_h(x) + j\psi_g(x)$, the 2D wavelet $\psi(x,y) = \psi(x)\psi(y)$ then becomes:

$$\psi(x,y) = [\psi_h(x) + j\psi_g(x)][\psi_h(y) + j\psi_g(y)] = \psi_h(x)\psi_h(y) - \psi_g(x)\psi_g(y) + j[\psi_g(x)\psi_h(y) + \psi_h(x)\psi_g(y)] \quad (37)$$

The Fourier spectrum of this complex wavelet is illustrated below:

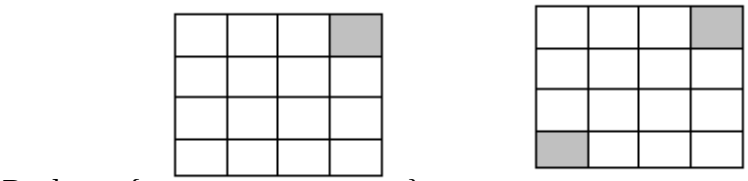


Since the spectrum of the 2D wavelet $\psi(x,y)$ is supported in only one quadrant of the 2D frequency plane, the complex 2D wavelet is oriented.

The real part of this complex wavelet is obtained by the sum of two separable wavelets

$$\text{Real Part } \{\psi(x,y)\} = \psi_h(x)\psi_h(y) - \psi_g(x)\psi_g(y) \quad (38)$$

The spectrum of this real wavelet is supported in 2 quadrants of the 2D frequency plane as illustrated below:



This wavelet is oriented at -45°. This is due to $\psi_g(t)$ being approximately the Hilbert transform of $\psi_h(t)$, $[\psi_g(t) \approx \mathcal{H}\{\psi_h(t)\}]$ and the complex wavelet $\psi(x) = \psi_h(x) + j\psi_g(x)$ being approximately analytic.

The first term in expression 44, $(\psi_h(x)\psi_h(y)$ is the HH wavelet of a 2D real wavelet transform which has been implemented using the filters $\{h_0(n), h_1(n)\}$.

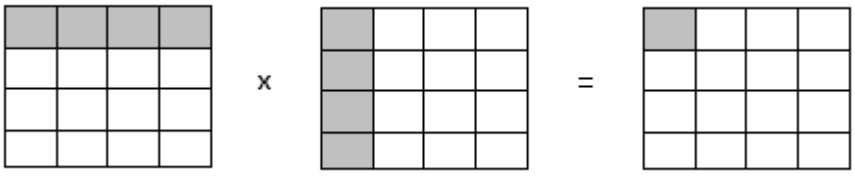
Similarly, the second term $\psi_g(x)\psi_g(y)$ is the HH wavelet of a 2D real wavelet transform which has been implemented using the filters $\{g_0(n), g_1(n)\}$.

To obtain a 2D real wavelet oriented at +45°, considering the complex 2D wavelet $\psi_2(x,y) = \psi(x)\psi(y)$ ($\psi(y)$ is the complex conjugate of $\psi(y)$, $\psi_2(x,y)$ becomes:

$$\psi_2(x,y) = [\psi_h(x) + \mathrm{j}\psi_g(x)]\overline{[\psi_h(y) + \mathrm{j}\psi_g(y)]} =$$

$$[\psi_h(x) + \mathrm{j}\psi_g(x)][\psi_h(y) - \mathrm{j}\psi_g(y)] = \psi_h(x)\psi_h(y) + \psi_g(x)\psi_g(y) + j[\psi_g(x)\psi_h(y) - \psi_h(x)\psi_g(y)]. \quad (39)$$

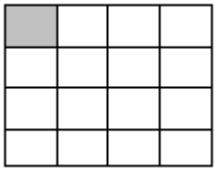
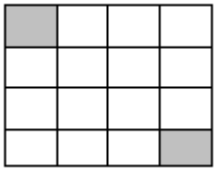
The support spectrum of this wavelet in the 2D frequency plane is shown below:



The Real part of this wavelet $\psi_2(x,y)$ becomes:

$$\text{Real Part } \{\psi_2(x,y)\} = \psi_h(x)\psi_h(y) - \psi_g(x)\psi_g(y) \quad (40)$$

Again, the spectrum is supported in two quadrants of the 2D frequency plane:

Real part {  } = 

To obtain another 4 oriented real 2D wavelets, the above steps are repeated for the below complex 2D wavelets:

$$\emptyset(x)\psi(y), \psi(x)\emptyset(y), \emptyset(x)\psi(y), \psi(x)\emptyset(y)$$

where $\psi(x) = \psi_h(x) + j\psi_g(x)$ and $\emptyset(x) = \emptyset_h(x) + j\emptyset_g(x)$

When we take the real part of the above four complex wavelets, we obtain four real oriented 2D wavelets.

Hence, the total six real oriented wavelets are:

$$\psi_{1,1}(x,y) = \emptyset_h(x)\psi_h(y), \tag{41}$$

$$\psi_{1,2}(x,y) = \psi_h(x)\emptyset_h(y), \tag{42}$$

$$\psi_{1,3}(x,y) = \psi_h(x)\psi_h(y), \tag{43}$$

$$\psi_{2,1}(x,y) = \emptyset_g(x)\psi_g(y), \tag{44}$$

$$\psi_{2,2}(x,y) = \psi_g(x)\emptyset_g(y), \tag{45}$$

$$\psi_{2,3}(x,y) = \psi_g(x)\psi_g(y). \tag{46}$$

To develop a 2-D wavelet transform that is both oriented and complex, the imaginary part of (43):

$$\text{ImagPart}\ \{\psi(x,y)\} = \psi_g(x)\psi_h(y) + \psi_h(x)\psi_g(y) \tag{47}$$

The spectrum of the imaginary part is the same as its real part. Therefore considering the imaginary parts of $(x)\psi(y)$, $\emptyset(x)\ \psi(y)$, $\psi(x)\emptyset(y)$, $\emptyset(x)\psi(y)$, and $\psi(x)\emptyset(y)$

where $\psi(x) = \psi_h(x) + j\psi_g(x)$ and $\emptyset(x) = \emptyset_h(x) + j\emptyset_g(x)$

Six oriented wavelets given by:

$$\psi_i(x,y) = \frac{1}{\sqrt{2}}(\psi_{3,i}(x,y) + \psi_{4,i}(x,y)) \tag{48}$$

$$\psi_{i+3}(x,y) = \frac{1}{\sqrt{2}}(\psi_{3,i}(x,y) - \psi_{4,i}(x,y)) \tag{49}$$

Normalization $\frac{1}{\sqrt{2}}$ has been used so the sum/difference operation creates an orthonormal operation.

For i = 1, 2, and 3, where the two separable 2-D wavelet bases are defined as:

$$\psi_{3,1}(x,y) = \emptyset_g(x)\psi_h(y), \quad (50)$$

$$\psi_{3,2}(x,y) = \psi_g(x)\emptyset_h(y), \quad (51)$$

$$\psi_{3,3}(x,y) = \psi_g(x)\psi_h(y), \quad (52)$$

$$\psi_{4,1}(x,y) = \emptyset_h(x)\psi_g(y), \quad (53)$$

$$\psi_{4,2}(x,y) = \psi_h(x)\emptyset_g(y), \quad (54)$$

$$\psi_{4,3}(x,y) = \psi_h(x)\psi_g(y). \quad (55)$$

Appendix 5: Inter-Coefficient Product (ICP)

Figure 73 shows the sub-band offset of a step edge. As shown it is defined in the direction normal to the sub-band:

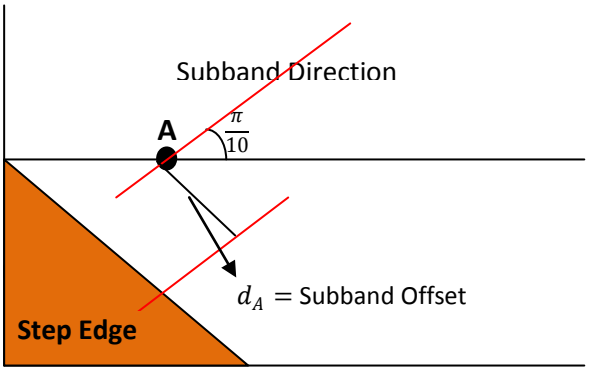


Figure 76 . The sub-band offset d_A between the step edge and the coefficient location A [58].

In [58] it is shown that the relationship between coefficient phase and feature offset is $\frac{-4.49}{D_W cos\frac{\pi}{10}}$ radians per unit length for sub-bands 1, 3, 4 and 6 and $\frac{-4.49}{D_W cos\frac{\pi}{4}}$ radians per unit length for subbands 2 and 5. D_W is the distance between adjacent coefficients.

Figure 74 shows a step edge in an image that spans the support regions of two coefficients that are horizontally adjacent at locations A=(x,y) and B=(x+1,y) at some arbitrary level [58]. The vertical and horizontal lines correspond to x, y coordinate system. The trigonometric relationship between the angle of the feature and its sub-band offsets in relation to the coefficients is demonstrated. ICP will determine the orientation of this step edge:

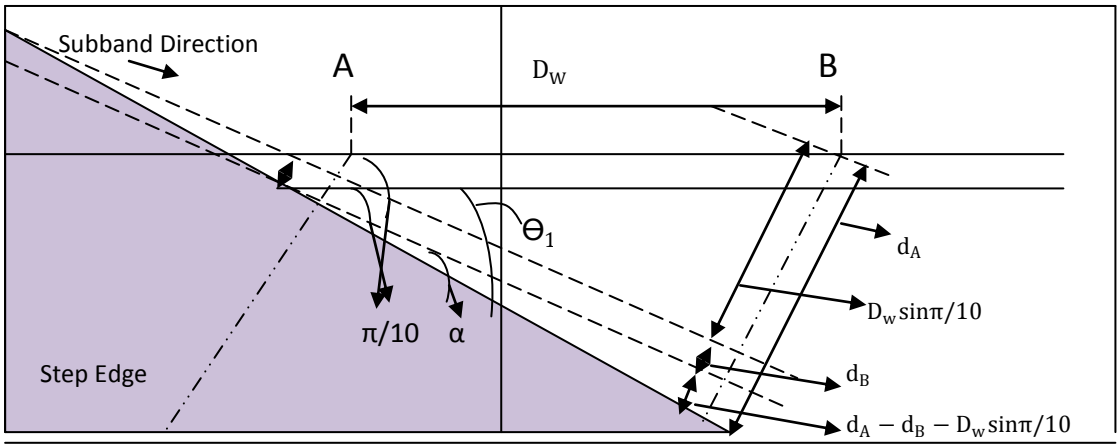


Figure 77 . Orientation of the step edge approximated using the ICP method [62].

The angle of the step edge is θ_1. The sub-band offsets are d_A and d_B. The trigonometric relationship between the angle and the offsets to the two horizontally adjacent $\frac{\pi}{10}$ sub-band coefficients located at A and B is demonstrated.

The initial step is to choose a sub-band which has the closest orientation to that edge so as to obtain a strong and linear response from the coefficients (In Figure 76, this is the $\frac{\pi}{10}$ sub-band) and then to identify the two adjacent coefficients where the edge lies in between A and B (coefficients) and to obtain the distance between them (D_W). In this case, the two coefficients are $W_l(x, y, 1)$ and $W_l(x+1, y, 1)$ which will have large magnitudes in the $\frac{\pi}{10}$ sub-band.

The next step is to measure the sub-band offsets d_A and d_B with respect to the coefficients. These two offsets have lengths of $\frac{D_W}{4.72} < W_l(x, y, 1)$ and $\frac{D_W}{4.72} < W_l(x+1, y, 1)$

where W_l represents the complex coefficient at position (x,y).

Referring to Figure 76, the angle θ_1 can be calculated by

$$\theta_1 = \frac{\pi}{10} + \alpha = \frac{\pi}{10} + \tan^{-1} \frac{d_A - d_B - D_W \sin\frac{\pi}{10}}{D_W \cos\frac{\pi}{10}} \tag{56}$$

Through simplifications it can be substituted with

$$\theta_1 = \frac{1}{4.49}[\angle W_1(x, y, 1) - \angle W_1(x+1, y, 1)] \tag{57}$$

The same approach can be used for other sub-bands taking into account the change of orientation and whether the angle needs to be measured with respect to the horizontal or vertical axis.

The phase differences of the coefficients can be obtained through conjugate products:

$$W_{l\Delta}(x, y, 1) = W_l(x, y, 1) \times W_l(x+1, y, 1)^* \tag{58}$$

The ICP is then represented by ψ_1 which is a constant-phase complex value where its magnitude indicates the strength of a feature and its phase gives the orientation [58].

$$\psi_l(x, y, 1) = \sqrt{\left|W_{l\Delta}(x, y, 1)\right|} \times e^{i\left(\frac{1}{4.49}\angle W_{l\Delta}(x, y, 1)\right)} \tag{59}$$

Appendix 6: Random forest classification results on Gaussian bars

Noise/Width		3	5	7	9	11	13
	0	0.99996412	0.9999172	0.998849	0.889884	0.532888	0.6305
Angle = 15	**0.002**	0.9999629	0.99989158	0.998731	0.841837	0.553137	0.662875
	0.004	0.99996569	0.99988202	0.998181	0.754656	0.566928	0.624368
	0.008	0.99993977	0.99987175	0.996881	0.606624	0.53343	0.589395
	0.016	0.99989275	0.99986302	0.992027	0.576358	0.536484	0.518991
	0.032	0.99990235	0.99979735	0.931275	0.534332	0.541494	0.520066
	0.064	0.9998534	0.99956433	0.836694	0.524587	0.524671	0.53305
	0.128	0.99955421	0.99676849	0.703224	0.541584	0.523034	0.523691
	0.256	0.99815342	0.95419813	0.572337	0.540573	0.559675	0.530932

Noise/Width		3	5	7	9	11	13
	0	0.99997327	0.99992946	0.991296	0.771019	0.653811	0.685032
Angle = 30	**0.002**	0.99995825	0.99992065	0.989676	0.695309	0.679608	0.675259
	0.004	0.99996849	0.99990734	0.957061	0.613901	0.634069	0.657125
	0.008	0.99994101	0.99981015	0.947667	0.527011	0.557034	0.583994
	0.016	0.99983715	0.99963058	0.885166	0.5106	0.540722	0.544654
	0.032	0.99979156	0.99944961	0.796636	0.523249	0.515513	0.536961
	0.064	0.99981783	0.99749911	0.685413	0.52862	0.515907	0.534683
	0.128	0.99941645	0.99151322	0.594212	0.538625	0.529915	0.536897
	0.256	0.99707489	0.86136965	0.590353	0.543881	0.531888	0.51801

Noise/Width		3	5	7	9	11	13
	0	1	0.99374253	0.904558	0.656199	0.709401	0.68466
Angle = 45	**0.002**	0.99999264	0.99512888	0.80613	0.641132	0.684359	0.678748
	0.004	0.99997758	0.99637412	0.661222	0.598887	0.650693	0.657274
	0.008	0.99996382	0.99386006	0.665479	0.584333	0.599527	0.589615
	0.016	0.99989857	0.98594638	0.551803	0.528796	0.55133	0.537534
	0.032	0.99997239	0.95253774	0.549283	0.527124	0.530881	0.543723
	0.064	0.99996746	0.88575402	0.569292	0.522796	0.527989	0.542736
	0.128	0.9998532	0.73474807	0.554162	0.520983	0.542637	0.515517
	0.256	0.98154687	0.65260818	0.538767	0.533939	0.537567	0.514566

Noise/Width		3	5	7	9	11	13
	0	0.99996373	0.99991608	0.99855	0.88766	0.530608	0.636528
Angle = 75	**0.002**	0.99995777	0.99992509	0.998104	0.835425	0.588944	0.654856
	0.004	0.99994883	0.99988597	0.997624	0.704538	0.535841	0.608913
	0.008	0.99994651	0.99986722	0.997444	0.598187	0.55148	0.578945
	0.016	0.99993132	0.99979906	0.981549	0.522716	0.495562	0.550204
	0.032	0.99984814	0.99975975	0.878742	0.519708	0.521338	0.515356
	0.064	0.99971453	0.9993595	0.830416	0.519468	0.538706	0.532753
	0.128	0.99950858	0.98924227	0.716716	0.533767	0.519069	0.503713
	0.256	0.99894542	0.96083378	0.629575	0.522599	0.540921	0.515668

Noise/Width		3	5	7	9	11	13
	0	1	1	0.999998	0.881353	0.635824	0.528403
Angle = 90	**0.002**	1	1	0.999745	0.878354	0.558849	0.542591
	0.004	1	1	0.999354	0.869486	0.563857	0.543599
	0.008	1	1	0.999219	0.790207	0.535881	0.543873
	0.016	1	1	0.996797	0.620386	0.514293	0.535992
	0.032	1	1	0.991066	0.589215	0.529222	0.531732
	0.064	1	0.99989621	0.90394	0.555967	0.530095	0.509393
	0.128	0.99999733	0.99941579	0.734869	0.532951	0.526009	0.528091
	0.256	0.99978871	0.95646161	0.64399	0.542677	0.521002	0.528938

Noise/Width		3	5	7	9	11	13
	0	0.99996246	0.99989114	0.998381	0.92137	0.550311	0.620415
Angle = 105	**0.002**	0.99995782	0.99988862	0.99644	0.844856	0.558244	0.648487
	0.004	0.99997051	0.99989156	0.99735	0.730792	0.563197	0.596951
	0.008	0.9999373	0.99986201	0.997235	0.581843	0.553258	0.606462
	0.016	0.99990524	0.99982235	0.975435	0.531448	0.526068	0.522616
	0.032	0.99988852	0.99956412	0.939462	0.515701	0.511945	0.526179
	0.064	0.99974149	0.99913994	0.778582	0.534292	0.538217	0.511224
	0.128	0.99960963	0.98225237	0.676019	0.544302	0.522784	0.524188
	0.256	0.99816629	0.95068028	0.630744	0.536839	0.539666	0.528232

Noise/Width		3	5	7	9	11	13
	0	0.99997051	0.99992971	0.986486	0.809918	0.685266	0.729185
Angle = 120	**0.002**	0.99996208	0.99994135	0.980417	0.655765	0.629753	0.702696
	0.004	0.99995439	0.99990209	0.971093	0.595187	0.614646	0.667782
	0.008	0.99992595	0.99986171	0.958952	0.543393	0.571511	0.572888
	0.016	0.99990774	0.99964926	0.906467	0.524953	0.524649	0.554066
	0.032	0.99980538	0.99945508	0.79035	0.527247	0.507186	0.538159
	0.064	0.99949781	0.99554479	0.694544	0.51436	0.522301	0.524999
	0.128	0.99935701	0.98047541	0.625845	0.544717	0.52125	0.512473

	0.256	0.99807567	0.95137255	0.58556	0.524391	0.506358	0.509271

Noise/Width		**3**	**5**	**7**	**9**	**11**	**13**
	0	1	0.99541323	0.889111	0.663269	0.759735	0.700129
Angle = 135	**0.002**	0.99999998	0.9956783	0.779072	0.652825	0.737586	0.688649
	0.004	0.99998663	0.99597298	0.659591	0.602614	0.648047	0.6133
	0.008	0.99995818	0.99123856	0.588733	0.552524	0.571546	0.591706
	0.016	0.99993872	0.98243048	0.544575	0.517952	0.55426	0.558987
	0.032	0.99993925	0.93702092	0.555186	0.529675	0.526028	0.524109
	0.064	0.99998498	0.89818602	0.555548	0.525006	0.53317	0.525565
	0.128	0.99988427	0.78203556	0.554819	0.527011	0.529279	0.527623
	0.256	0.97200505	0.70671321	0.538301	0.52859	0.54068	0.528542

Noise/Width		**3**	**5**	**7**	**9**	**11**	**13**
	0	0.99996935	0.99990462	0.991669	0.761777	0.645715	0.723286
Angle = 150	**0.002**	0.99996785	0.99988671	0.979797	0.6293	0.643062	0.695151
	0.004	0.99995736	0.9998679	0.963304	0.614983	0.629288	0.638511
	0.008	0.9999341	0.99984825	0.957491	0.536412	0.56488	0.582477
	0.016	0.99989281	0.99952593	0.868593	0.539672	0.541785	0.540685
	0.032	0.99978043	0.9975807	0.763801	0.513358	0.522058	0.516869
	0.064	0.9997894	0.99546858	0.691788	0.532044	0.52539	0.520792
	0.128	0.99937182	0.9827446	0.574983	0.522524	0.521951	0.514657
	0.256	0.99258979	0.84553196	0.575036	0.543892	0.525346	0.517177

Noise/Width		**3**	**5**	**7**	**9**	**11**	**13**
	0	0.99995347	0.99989938	0.998578	0.903599	0.533973	0.635045
Angle = 165	**0.002**	0.9999532	0.99990099	0.997801	0.802238	0.524002	0.688233
	0.004	0.99995409	0.99983648	0.997352	0.697647	0.536857	0.622634
	0.008	0.99993547	0.99986791	0.993723	0.620267	0.565545	0.564471
	0.016	0.99992438	0.99978383	0.986629	0.552217	0.500745	0.528097
	0.032	0.99992141	0.99971111	0.942011	0.526469	0.517075	0.514817
	0.064	0.9997753	0.99958569	0.847966	0.535589	0.535612	0.53065
	0.128	0.99953723	0.99737634	0.693314	0.53893	0.509503	0.508761
	0.256	0.99816599	0.89117483	0.578059	0.554195	0.534806	0.545928

Noise/Width		**3**	**5**	**7**	**9**	**11**	**13**
	0	1	1	0.999865	0.872165	0.613582	0.512832
Angle = 180	**0.002**	1	1	0.999167	0.887177	0.618839	0.528582
	0.004	1	1	0.999595	0.845546	0.55292	0.551902
	0.008	1	1	0.999313	0.811879	0.519449	0.559955
	0.016	1	1	0.998741	0.641008	0.505492	0.518175
	0.032	1	0.9999995	0.993678	0.573554	0.538379	0.511322
	0.064	1	0.99970716	0.877524	0.537587	0.530187	0.51994

	0.128	0.99999341	0.99684299	0.815823	0.57594	0.521788	0.514883
	0.256	0.99977596	0.97558235	0.60088	0.527668	0.548697	0.553813

Appendix 7: SVM classification results on Gaussian bars

Noise/Width		3	5	7	9	11	13
	0	0.99552	0.983682	0.926777	0.669589	0.517691	0.481312
Angle = 15	**0.002**	0.994874	0.971038	0.92321	0.642976	0.493153	0.489926
	0.004	0.995337	0.956877	0.917175	0.602689	0.495189	0.532616
	0.008	0.990981	0.960589	0.91938	0.570814	0.503072	0.468192
	0.016	0.994416	0.968256	0.888524	0.540023	0.538947	0.488548
	0.032	0.988252	0.961887	0.898486	0.549146	0.556554	0.5053
	0.064	0.96452	0.970223	0.797851	0.571233	0.545331	0.532578
	0.128	0.963923	0.962779	0.737369	0.54273	0.574965	0.538995
	0.256	0.937457	0.919935	0.642271	0.551908	0.587928	0.54731

Noise/Width		3	5	7	9	11	13
	0	0.997958	0.963562	0.876307	0.516537	0.491682	0.455501
Angle = 30	**0.002**	0.997554	0.969426	0.852295	0.527117	0.478839	0.488866
	0.004	0.996063	0.970074	0.851831	0.508604	0.501292	0.506131
	0.008	0.996684	0.961179	0.806543	0.491058	0.481026	0.486096
	0.016	0.987908	0.961011	0.807489	0.533845	0.520444	0.480024
	0.032	0.972595	0.971899	0.75714	0.484885	0.499742	0.502314
	0.064	0.976449	0.972673	0.592998	0.520958	0.527609	0.501249
	0.128	0.953455	0.938107	0.649659	0.54177	0.539597	0.529586
	0.256	0.95921	0.869216	0.565318	0.550547	0.546376	0.532298

Noise/Width		3	5	7	9	11	13
	0	0.999756	0.914299	0.628474	0.518177	0.529491	0.499533
Angle = 45	**0.002**	0.997717	0.91715	0.610407	0.477003	0.49935	0.49914
	0.004	0.998627	0.92176	0.564414	0.506045	0.5067	0.483734
	0.008	0.993028	0.896037	0.543083	0.515605	0.481757	0.525116
	0.016	0.980859	0.90047	0.511808	0.50925	0.505111	0.480288
	0.032	0.978297	0.900628	0.6185	0.492978	0.516619	0.527688
	0.064	0.985754	0.849232	0.559873	0.544799	0.514459	0.541671
	0.128	0.974414	0.796242	0.532371	0.56149	0.565487	0.530903
	0.256	0.936187	0.629497	0.556309	0.553303	0.53752	0.555588

Noise/Width		3	5	7	9	11	13
	0	0.993326	0.961944	0.92881	0.665455	0.501652	0.485071
Angle = 75	**0.002**	0.990105	0.966326	0.941376	0.622122	0.502641	0.49281
	0.004	0.994249	0.955921	0.912601	0.563519	0.493259	0.468054
	0.008	0.992975	0.955792	0.916369	0.543514	0.504408	0.49641
	0.016	0.986818	0.970569	0.919557	0.55899	0.479525	0.442237
	0.032	0.974283	0.960162	0.846743	0.563649	0.508034	0.507953
	0.064	0.985285	0.979566	0.784158	0.573866	0.528316	0.516788
	0.128	0.949316	0.961982	0.709157	0.566263	0.560958	0.577303
	0.256	0.943982	0.845923	0.623505	0.542521	0.560555	0.542287

Noise/Width		3	5	7	9	11	13
	0	0.999966	0.999537	0.908337	0.754983	0.637295	0.600689
Angle = 90	**0.002**	0.999975	0.987906	0.93209	0.728373	0.602747	0.56229
	0.004	0.99998	0.971371	0.91943	0.729315	0.578741	0.529879
	0.008	0.999951	0.978484	0.932366	0.683626	0.557557	0.559978
	0.016	0.999897	0.976128	0.914495	0.617358	0.554996	0.504658
	0.032	0.996944	0.978352	0.941164	0.595147	0.542469	0.539154
	0.064	0.995004	0.968245	0.8844	0.565983	0.552689	0.601125
	0.128	0.981878	0.957285	0.763533	0.579313	0.597639	0.529205
	0.256	0.925253	0.886211	0.66669	0.563253	0.569188	0.562653

Noise/Width		3	5	7	9	11	13
	0	0.993805	0.976877	0.927279	0.685	0.492016	0.455891
Angle = 105	**0.002**	0.994606	0.960433	0.909841	0.630461	0.501649	0.461563
	0.004	0.991491	0.953507	0.92131	0.635177	0.488666	0.483244
	0.008	0.995591	0.954768	0.901699	0.555053	0.496212	0.464594
	0.016	0.985525	0.951326	0.900499	0.553519	0.523436	0.490057
	0.032	0.990945	0.963219	0.812824	0.509938	0.540949	0.495646
	0.064	0.976066	0.974996	0.807363	0.578576	0.516827	0.482372
	0.128	0.953619	0.959277	0.732792	0.559572	0.556821	0.533798
	0.256	0.931119	0.918772	0.673232	0.554341	0.543905	0.555544

Noise/Width		3	5	7	9	11	13
	0	0.99662	0.969718	0.875699	0.53714	0.510646	0.485754
Angle = 120	**0.002**	0.994958	0.9601	0.88252	0.548114	0.483819	0.49694
	0.004	0.997312	0.955027	0.838934	0.529569	0.509809	0.481866
	0.008	0.995203	0.967605	0.829862	0.50604	0.499534	0.504458
	0.016	0.987911	0.971902	0.816622	0.505833	0.540697	0.528794
	0.032	0.979078	0.971058	0.744077	0.51729	0.506021	0.533797
	0.064	0.961504	0.961889	0.714312	0.532413	0.583671	0.552905
	0.128	0.948008	0.935909	0.623209	0.53449	0.542382	0.548356
	0.256	0.942166	0.834703	0.573261	0.506702	0.547393	0.534443

Noise/Width		3	5	7	9	11	13
	0	0.999701	0.895779	0.625382	0.502018	0.502402	0.475728
Angle = 135	**0.002**	0.999666	0.930187	0.582524	0.498955	0.46028	0.498963
	0.004	0.999699	0.914102	0.599587	0.493958	0.496682	0.497878
	0.008	0.995859	0.936187	0.557204	0.493467	0.511091	0.519173
	0.016	0.995441	0.919105	0.540007	0.504647	0.535166	0.506324
	0.032	0.986945	0.905164	0.529266	0.49574	0.505232	0.505995
	0.064	0.978577	0.844257	0.53434	0.54927	0.530923	0.548744
	0.128	0.96647	0.760135	0.542388	0.532308	0.53334	0.560559
	0.256	0.945354	0.663017	0.534847	0.533272	0.565556	0.554885

Noise/Width		3	5	7	9	11	13
	0	0.998144	0.96232	0.874234	0.545261	0.48331	0.493809
Angle = 150	**0.002**	0.997246	0.973601	0.845346	0.49579	0.456791	0.481669
	0.004	0.995106	0.970477	0.830898	0.493713	0.481997	0.50537
	0.008	0.995195	0.972789	0.829465	0.496557	0.512523	0.4593
	0.016	0.986945	0.975582	0.779709	0.549382	0.464788	0.540858
	0.032	0.983971	0.966093	0.732125	0.532599	0.504039	0.53822
	0.064	0.948551	0.963406	0.658774	0.491539	0.50787	0.517533
	0.128	0.954132	0.957067	0.576243	0.573539	0.559263	0.535321
	0.256	0.9414	0.868345	0.587502	0.559044	0.528761	0.554128

Noise/Width		3	5	7	9	11	13
	0	0.993632	0.983142	0.923573	0.687486	0.479498	0.465405
Angle = 165	**0.002**	0.992583	0.973537	0.916679	0.662819	0.491822	0.474557
	0.004	0.987513	0.962946	0.920313	0.627198	0.479002	0.478291
	0.008	0.988336	0.96221	0.899554	0.552477	0.510632	0.471137
	0.016	0.992322	0.970944	0.932622	0.539613	0.497028	0.504135
	0.032	0.985118	0.965966	0.889788	0.545723	0.47517	0.51797
	0.064	0.958599	0.960438	0.78562	0.557081	0.543555	0.522246
	0.128	0.95619	0.972498	0.712519	0.547901	0.531131	0.546581
	0.256	0.932169	0.889123	0.617827	0.572805	0.581019	0.544016

Noise/Width		3	5	7	9	11	13
	0	0.999949	0.999423	0.924609	0.762167	0.628242	0.589679
Angle = 180	**0.002**	0.99994	0.987166	0.921896	0.742808	0.61622	0.549497
	0.004	0.999912	0.981615	0.931484	0.692246	0.518537	0.542306
	0.008	0.999864	0.966869	0.941014	0.671649	0.564066	0.573226
	0.016	0.999813	0.972157	0.946806	0.596242	0.531366	0.526552
	0.032	0.999697	0.976283	0.928207	0.596837	0.523558	0.543838
	0.064	0.997429	0.98159	0.839208	0.578039	0.546717	0.540626
	0.128	0.970893	0.972793	0.738527	0.57333	0.539401	0.570597
	0.256	0.93297	0.887598	0.682493	0.564768	0.556409	0.55726

Bibliography

[1] J. B. Saaddine, A. A. Honeycutt, K. M. V. Narayan, X. Zhang, R. Klein, and J. P. Boyle (2008) "Projection of diabetic retinopathy and other major eye diseases among people with diabetes mellitus: United States, 2005–2050," *Archives of Ophthalmology*, **126**(12):1740–1747.

[2] M. Niemeijer, J.J. Staal, B. van Ginneken, M. Loog, and M.D. Abramoff, (2004) "Comparative study of retinal vessel segmentation methods on a new publicly available database",*SPIE Medical Imaging*, 648–656.http://www.isi.uu.nl/Research/Databases/DRIVE/.Date of last access: 09/02/2014.

[3] R Sadeghzadeh, M Berks, S M Astley, C J Taylor, (2010)"Detection of retinal blood vessels using complex wavelet transforms and random forest classification". *Medical Image Understanding and Analysis*; Warwick. BMVA, 127-132.

[4] R Sadeghzadeh, M Berks, S M Astley, C J Taylor,(2011)"Evaluation of blood vessel detection methods"*Medical Imaging 2011: Image Processing. Proceedings of the SPIE,* **7962**:79623U-79623U-12.

[5]L Breiman, (2001),"Random Forests", *Machine Learning*, *university of California***45** (1):5–32.

[6] I. W. Selesnick, R.G. Baraniuk, and N. G. Kingsbury,(2005)"The dual-tree complex wavelet transform", *IEEE Signal Processing Magazine*, **22**(6),123-151.

[7] M. D. Abramoff, M. K. Garvin, and M. Sonka,(2010) "Retinal imaging and image analysis," *IEEE Reviews in Biomedical Engineering*, **3**:169–208.

[8] M.I Iqbal, N.S Gubbal, A.M Aibinu, A. Khan (2006), "Automatic diagnosis of diabetic retinopathy using fundus images", *Blenkinge Institute of Technology*.

[9] R.M Rangayyan, F. Oloumi (2007), "Detection of blood vessels in the retina using Gabor filters", *Canadian Conference on Electrical and Computer Engineering*, **1-3:** 717-720.

[10] H. A.Harry. Quigley; A.E. Brown; J.D. Morrison, MD; S M. Drance (1990), "The Size and Shape of the Optic Disc in Normal Human Eyes" *Arch Ophthalmol* : 108(1):51-57.

[11] S. Ravishankar, A. Jain, A. Mittal (2009), “Automated feature extraction for early detection of diabetic retinopathy in fundus images”, *IEEE Conference on Computer Vision and Pattern Recognition.*

[12] H. Kolb, E Fernandez and R. Nelson (2002), Webvision: The Organization of the Retina and Visual System. http://www.webvision.med.utah.edu. Date of last access: 09/02/2014.

[13] A. Osareh, B. Shadgar (2009), "Automatic blood vessel segmentation in colour images of retina” *Iranian Journal of Science and Technology Transaction B-Engineering*, **33**(B2): 191-206.

[14] K S, Coyne , MK Margolis, T K Martin , TM Baker, R Klein , M D Paul (2004), “The impact of diabetic retinopathy: perspectives from patient focus groups”. *FamPract***21**:447–453.

[15] C .Agurto, E. S. Barriga,V. Murray, S. Nemeth, R. Crammer,W. Bauman, G. Zamora, S. M Pattichis, and P Soliz (2011), “Automatic detection of diabetic retinopathy and age-related macular degeneration in digital fundus images”. *Investigative Ophthalmology and Visual Sciences (IOVS).***52**(8):5862-71.

[16] M. J. Cree and H. F. Jelinek, “Image analysis of retinal images”, chapter 11:263.

[17] H.P Hammes, M Porta (2010), “Experimental approaches to diabetic retinopathy”. *Front Diabetes. Basel, Karger*, **20**:1–19.

[18] M. Niemeijer (2006), “Automatic detection of diabetic retinopathy in digital fundus photographs”, PhD thesis, Image Sciences Institute, Utrecht University.

[19] M. García, C. Sanchez, M. López, J. Poza, R. Hornero. (2009), “Detection of hard exudates in retinal images using a radial basis function classifier“, *Annals of biomedical engineering*, **37**(7):1448–1463.

[20] T. Kauppi, V. Kalesnykiene, J.K. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kälviäinen, J. Pietilä (2006), “DIARETDB0: Evaluation Database and Methodology for Diabetic Retinopathy Algorithms*”*, http://www.it.lut.fi/project/imageret/diaretdb0/index.html.Date of last access: 09/02/2014.

[21] J.V.B Soares, J. J. G. Leandro (2006), "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification." *IEEE Transactions on Medical Imaging,* **25**(9): 1214-1222.

[22] J. G. Fujimoto,W. Drexler (2008), "Introduction to optical coherence tomography"*optical coherence tomography Institute of Technology Cambridge. Biological and Medical Physics, Biomedical Engineering*, 1-45.

[23] M. Ali and R. Parlapalli (2010),"Signal processing overview of optical coherence tomography systems for medical imaging", *Texas Instruments Incorporated*.

[24] Topcon connecting visions. http://www.topcon-medical.co.uk. Date of last access: 09/02/2014.

[25] R. Scarf M.D. https://rmscharf.ipower.com/symptomsdiseases.html.Date of last access: 09/02/2014.

[26] Optical Coherence Tomography,http://www.oct-ophtalmo.fr.Date of last access: 09/02/2014.

[27] D.Y. Lin, M.S. Blumenkranz, R. Brothers (1999) ,"The Role of Digital Fundus Photography in Diabetic Retinopathy Screening", *Diabetes technology and therapeutics,* **1**(4):477-487.

[28] C.C. Hull (2009), "Digital Imaging and Screening for Diabetic Retinopathy". *Optometry Today*, 28-35.

[29] A. Hoover, V. Kouznetsova(2000), "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response" *IEEE Transactions on Medical Imaging,* **19**(3): 203-210.

[30] J. Zhang, T. Tan (2002) , "Brief review of invariant texture analysis methods", *Pattern Recognition,* **35** :735–747.

[31] H.H. Arsenault, Y. Sheng (1986), "Properties of the circular harmonic expansion for rotation-invariant pattern recognition", Applied Optics.**25** (18) 3225–3229.

[32]A. Khotanzad, Y.H. Hong (1990), Invariant image recognition by Zernike moments, *IEEE Trans. PAMI,* **12:** 489–497.

[33] R.T. Chin (1988), "On image analysis by the methods of moments*"*, *IEEE Trans. PAMI,* **10:** 496–513.

[34] M. Pietikainen, T. Ojala, Z. Xu (2000), “Rotation-invariant texture classification using feature distributions*”, Pattern Recognition,* **33:** 43–52.
[35] D.G. Lowe (1999), "Object recognition from local scale-invariant features". *Proceedings of the International Conference on Computer Vision,* **2**:1150–1157.
[36] J. Davernoy (1984), “Optical digital processing of directional terrain textures invariant under translation, rotation, and change of scale”, *Applied Optics,* **23** (6) : 828–837.
[37] J. Mao, A.K. Jain (1992), “Texture classification and segmentation using multi-resolution simultaneous autoregressive models*”, Pattern Recognition,* **25** (2) :173–188.
[38] A. Kashyap, A. Khotanzad (1990), *“*A model based method for rotation invariant texture classification*”, IEEE Trans. PAMI,* **12**: 489–497.
[39] F. Cohen (1992), “Classification of rotated and scaled texture images using Gaussian Markov random field models*”, IEEE Trans. PAMI,* **13** (2): 192–202.
[40] J.L Chen, A. Kundu (1994), “Rotation and gray scale transform invariant texture recognition using hidden Markov model”, *IEEE Trans. Pattern Analysis and Machine Intelligence,* (16):208-214.
[41] D. Gabor (1946), "Theory of Communication", *J. Inst. Elec Eng,* **93**: 429 – 457.
[42] A.K. Jain, F. Farrokhnia (1991), “Unsupervised texture segmentation using Gabor filters*”, Pattern Recognition,* **24** (12):1167–1186.
[43] T.P. Weldon , W.E Higgins , D.F. Dunn (1996), “Efficient Gabor filter design for texture segmentation*”, Pattern Recognition,* **29** (12): 2005-2015.
[44] W Freeman, E. Adelson (1991), "The Design and use of Steerable Filters", *IEEE Trans. pattern anal. Machine intel.* **13**:89 1-906.
[45] E.P. Simoncelli, W.T. Freeman (1995), "The Steerable Pyramid: A Flexible Architecture for Multi-Scale Derivative Computation", *IEEE Second International Conference on Image Processing, Washington DC*, **3**: 444-447.
[46] J. Portilla, E.P. Simoncelli (2000), “A parametric texture model based on joint statistics of complex wavelet coefficients”, *International journal of computer vision,* **40** (1): 49-71.

[47] R. Goyal, W.L. Goh, D.P. Mital, K.L.Chan (1995) “Scale and rotation invariant texture analysis based on structural property”. *Proceedings of IECON,* **1-2**: 1290–1294.
[48] G. Eichmann, T. Kasparis (1998), “Topologically invariant texture descriptors”, *Comput. Vision Image Process*, **41**:267–281.
[49] R.K. Goyal, W.L. Goh, D.P. Mital, K.L. Chan (1994), “A translation rotation and scale invariant texture analysis technique based on structural properties”, *Proceedings of the Third International Conference on Automation Technology*.
[50] D.P. Mital, W.L. Goh, K.L. Chan, R.K. Goyal (1994), “A translation rotation and scale invariant texture analysis technique based on image granularity”, *Proceedings of the Fifth International Symposium on Robotics and Manufacturing*.
[51] R. Goyal, W.L. Goh, D.P. Mital, K.L. Chan (1994), “Scale and Rotation invariant texture analysis based on structural property”, *Proceedings of 7th Al Conference*, 522–529.
[52] R.K. Goyal, W.L. Goh, D.P. Mital, K.L. Chan (1994), “Invariant element compactness for texture classification”, *Proceedings of the Third International Conference on Automation, Robotics and Computer Vision, Singapore,* 1902–1096.
[53] W.K. Lam, C.K. Li (1997), “Rotated texture classification by improved iterative morphological decomposition”, *IEE ProcVision Image Signal Process,* **144** (3):171–179.
[54] I. Daubechies (1992), “Ten Lectures on Wavelets*”, Society for Industrial and Applied Mathematics,* Philadelphia.
[55] S.K. Jashanbir, S. Reecha (2012), “Comparative Performance Analysis of Haar, Symlets and Bior wavelets on Image compression using Discrete Wavelet Transform”, *International Journal of Computer and Distributed System,* **1**(2):11.
[56] S. Arivazhagan , L. Ganesan (2003), “Texture classification using wavelet transform”, *Pattern Recognition Letters*, **24**(9-10):1513-1521.
[57] S. Arivazhagan , L. Ganesan (2003), “Texture segmentation using wavelet transform”, *Pattern Recognition Letters*, **24**(16) : 3197-3203.
[58] R. Anderson, N. Kingsbury, J. Fauqueur (2005), “Determining multi-scale image feature angles from complex wavelet phases”, Image analysis and recognition, **3656**: 490-498**.**
[59] N. Kingsbury (2006), "Rotation-invariant local feature matching with complex wavelets", *Proc. European Conference on Signal Processing, Florence*.
[60] E. Alpaydin (2004), “Introduction to Machine Learning”, *The MIT Press,* ISBN 0-262-01211-1.
[61] S. B. Kotsiantis (2007), “Supervised Machine Learning: A Review of Classification Techniques”, *proceedings of Emerging Artificial Intelligence,* 3-24.

[62] D. Michie, D.J. Spiegelhalter, C.C. Taylor (1994), "Machine Learning, Neural and Statistical Classification*". EllisHorwood Limited, Upper Saddle River, NJ, USA*, ISBN 0-13-106360-X.
[63] K. Gurney (1997), "An Introduction to Neural Networks*", UCL Press.(Taylor & Francis group).*
[64] A. Khajanchi (2003), "Artificial neural networks: the next intelligence", *university of South Carolina*. Available at www.globalriskguard.com/resources/market/NN.pdf.Date of last access: 09/02/2014.
[65] Wu, R Chung-Fern (1994), "Integrating Neural Computing and Auditing Expertise", *Managerial Auditing Journal*, **9**(3): 20-26.
[66] P. Cunningham, S. J. Delany (2007), "K-Nearest Neighbour Classifiers", *Technical Report, UCD-CSI, Dublin: Artificial Intelligence Group.*
[67] Press, H. W. Teukolsky, A. Saul, W.T. Vetterling, B.P. Flannery (2007), "Numerical Recipes: The Art of Scientific Computing (3rd ed.)", Section 16.5.Support Vector Machines, *New York: Cambridge University Press*. ISBN 978-0-521-88068-8.
[68] V.N.Vapnik (2000), "The Nature of Statistical Learning Theory*". Springer, second edition.*
[69] V.N.Vapnik ,A.Y.Chervonenkis (1974),"Theory of pattern recognition*", Nauka, Moscow.*
[70] V. Podgorelec, P. Kokol, B. Stiglic, I, Rozman (2002), "Decision trees: an overview and their use in medicine", *Journal of Medical Systems*, **26**:445–463.
[71] P.O. Gislason (2004), "Random forest classification of multisource remote sensing and geographic data". *Geoscience and remote sensing symposium*, **2**:1049-1052.
[72] P.O. Gislason, J.A. Benediktsson, J. R. Sveinsson (2006), "Random Forests for land cover classification", **27** (4): 294-300.
[73] R. Diaz-Uriarte, S. Alvarez De Andres (2006), "Gene selection and classification of microarray data using random forest", *BMC bioinformatics*, **7**(1):3.

[74] D. R. Cutler**,** T. C. Edwards **,** K H. Beard**,** A Cutler**,** KT. Hess**,** J Gibson, and J. J. Lawler (2007), **"**Random forests for classification in ecology". *Ecology,* **88**(11): 2783–2792.

[75] P. Xu, F. Jelinek, "Random Forests in Language Modelling", *Centre for Language and Speech Processing the Johns Hopkins University*, *EMNLP*, 325-332.

[76] A. D Flaxman**,** A.Vahdatpour**,** S Green**,** S. LJames**,** C.J.LMurray **(**2011), **"**Random forests for verbal autopsy analysis: multisite validation study using clinical diagnostic gold standards". *Population Health Metrics Research Consortium*, **9**(29).

[77] E. Ricci, R. Perfetti (2007), "Retinal blood vessel segmentation using line operators and support vector classification." *IEEE Transactions on Medical Imaging* **26**(10): 1357-1365.

[78] Dixon, R. N. and C. J. Taylor (1979). Automated asbestos fibre counting. Machine Aided Image Analysis,. W. E. Gardner. London, Institute of Physics. 44: 178-185.

[79] C. Sinthanayothin, J. F. Boyce (1999), "Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images. "*British Journal of Ophthalmology* **83**(8): 902-910.

[80] W. Freeman, E. Adelson (1991), "The Design and use of Steerable Filters", *IEEE Trans. pattern anal. Machine intel,* **13**:89 1-906.

[81] A. Perez-Rovira, K. Zutis (2011), "Improving vessel segmentation in ultra-wide field-of-view retinal fluorescein angiograms.", *Conference proceeding : Annual International Conference of the IEEE Engineering in Medicine and Biology Society:*2614-7.

[82] N. Lian , V. Zagorodnov (2007),"Retinal vessel detection using self-matched filtering", *IEEE International Conference on Image Processing*, **1-7:** 2829-2832.

[83] J. C. Bancroft (2002),"Introduction to matched filters", http://www.crewes.org/ForOurSponsors/ResearchReports/2002/2002-46.pdf.Date of last access: 09/02/2014.

[84] M. Al-Rawi, M. Qutaishat (2007), "An improved matched filter for blood vessel detection of digital retinal images.", *Computers in Biology and Medicine,* **37**(2): 262-267.

[85] P. Bankhead, C. N. Scholfield, J. G. McGeown,T. M. Curtis (2012) "Fast retinal vessel detection and measurement using wavelets and edge location refinement," *Plos one* **7**(3):e32435.

[86] Xu X, Niemeijer M, Song Q, Sonka M, Garvin MK, et al. (2011) Vessel boundary delineation on fundus images using graph-based approach. IEEE Trans Med Imaging 30: 1184–91.

[87] N.M Salem, A. K. Nandi (2008), “Unsupervised Segmentation of Retinal Blood Vessels Using a Single Parameter Vesselness Measure”, *Computer Vision, Graphics & Image Processing*,528:534.

[88] I. Abdurrazaq, S. Hati (2008), “Morphology approach for features extraction in retinal images for diabetic retionopathy diagnosis”, *International Conference on Computer and Communication Engineering*, **1-3:**1373-1377.

[89] X.Y. Jiang, D. Mojon (2003), "Adaptive local thresholding by verification-based multi-threshold probing with application to vessel detection in retinal images.",*IEEE Transactions on Pattern Analysis and Machine Intelligence*,**25**(1): 131-137.

[90] S. Dua, N. Kandiraju (2005), “Design and implementation of a unique blood-vessel detection algorithm towards early diagnosis of diabetic retinopathy”, *International Conference on Information Technology: Coding and Computing*,**1:**26-31.

[91] M. Sofka, C. V. Stewart (2006). "Retinal vessel centreline extraction using multiscale matched filters, confidence and edge measures." *IEEE Transactions on Medical Imaging,* **25**(12): 1531-1546.

[92] M.U. Akram, A. Tariq (2009), “Retinal image blood vessel segmentation”, *International Conference on Information and Communication Technologies*. *New York,* 140-144.

[93] H.F. Jelinek, M. J. Cree (2007), "Automated segmentation of retinal blood vessels and identification of proliferative diabetic retinopathy."*Journal of the Optical Society of America a-Optics Image Science and Vision,* **24**(5): 1448-1456.

[94] D. J. Cornforth , H. F. Jelinek , J. J. G. Leandro , J. V. B. Soares , R. M. Cesar-Jr , M. J. Cree , P. Mitchell and T. Bossamaier (2004) "Development of retinal blood vessel segmentation methodology using wavelet transforms for assessment of diabetic retinopathy", *Proc. 8th Asia Pacific Symp. Intell.Evolution. Syst* :50-60.

[95] S. Sekhar, F. E. Abd El-Samie (2011), "Automated localization of retinal features.", *Applied Optics*, **50**(19): 3064-3075.

[96] D. Wu, M. Zhang (2006), "On the adaptive detection of blood vessels in retinal images."*IEEE Transactions on Biomedical Engineering*, **53**(2): 341-343.
[97] R. Leander , P.I. Bidari , T.A. Mohammed , M. Das , E. S. Umbaugh (2010), "New algorithm for detecting smaller retinal blood vessels in fundus images", *Proc. SPIE 7624, Medical Imaging 2010*: *Computer-Aided Diagnosis,* 76243:10.1117/12.844996.
[98] L. Espona, M. J. Carreira (2007),"A snake for retinal vessel segmentation", *Pattern Recognition and Image Analysis Proceedings*, **4478:** 178-185.
[99] B. Al-Diri, A. Hunter (2010), "Automated analysis of retinal vascular network connectivity.", *Computerized Medical Imaging and Graphics*, **34**(6): 462-470.
[100] M. Freiman, N. Broide (2009),"Vessels-Cut: A Graph Based Approach to Patient-Specific Carotid Arteries Modelling", *Modelling the Physiological Huma, Berlin, Springer-Verlag Berlin*, **5903:** 1-12.
[101] A.G Salazar-Gonzalez, L. Yongmin (2010),"Retinal blood vessel segmentation via graph cut". *11th International Conference on Control Automation Robotics & Vision,* **1**:225-230.
[102] A.M. Mendonca, A. Campilho (2006),"Segmentation of retinal blood vessels by combining the detection of centrelines and morphological reconstruction.", *IEEE Transactions on Medical Imaging*, **25**(9): 1200-1213.
[103] J. Staal, M. D. Abramoff (2004), "Ridge-based vessel segmentation in colour images of the retina."*IEEE Transactions on Medical Imaging*, **23**(4): 501-509.
[104] A. Hoover, V. Kouznetsova(2000), "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response" *IEEE Transactions on Medical Imaging*, **19**(3): 203-210.
[105] M.E. Martinez-Perez, A. D. Hughes (1999),"Retinal blood vessel segmentation by means of scale-space analysis and region growing", *Medical Image Computing and Computer-Assisted Intervention, Miccai99, Proceedings*, **1679:** 90-97.
[106] X.You, Q. Peng (2011),"Segmentation of retinal blood vessels using the radial projection and semi-supervised approach.", *Pattern Recognition*, **44**(10-11):2314-2324.
[107] M. Adel, M. Rasigni (2009), "Statistical-based linear vessel structure detection in medical images, *16th IEEE International Conference on Image Processing, New York,* **1-6**:649-652.

[108] C. Muramatsu, Y. Hatanaka (2010), “Automated detection and classification of major retinal vessels for determination of diameter ratio of arteries and veins”, *Medical Imaging 2010: Computer - Aided Diagnosis. Bellingham, Spie-IntSoc Optical Engineering*, **7624**.

[109] A. Can, H. Shen (1999), "Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms." *IEEE transactions on information technology in biomedicine: a publication of the IEEE Engineering in Medicine and Biology Society*, **3**(2): 125-38.

[110] T. Fawcett (2003), ”ROC Graphs: Notes and Practical Considerations for Data Mining Researchers”. *Technical Report HPL-2003–4, Palo Alto, CA: HP Laboratories.*

[111] Kohavi, Ron (1995), "A study of cross-validation and bootstrap for accuracy estimation and model selection". *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence* **2** (12): 1137–1143.

[112] C.E. Metz (1978), “Basic principles of ROC analysis”. *Seminars in Nuclear Medicine*, **8**:283-298.

[113] S. Chaudhuri, S. Chatterjee (1989). "Detection of blood-vessels in retinal images using two-dimensional matched filters”, *IEEE transactions on medical imaging*, **8**(3): 263-269.

[114] Z. Liang, M. Rzeszotarski, L. Singerman, J. Chokreff (1994), “The detection and quantification of retinopathy using digital angiograms”, *IEEE Medical Imaging*, **13**(4):619:626.

[115] W. Changhua, J. Jennifer, K. Derwent, P. Stanchev (2009), “Retinal vessel radius estimation and a vessel centre line segmentation method based on ridge descriptors”, *Signal Processing Systems*, **55**(1-3):91-102.

[116] A. Bhuiyan, B. Nath (2012), "Automated analysis of retinal vascular tortuosity on colour retinal images." *Journal of Medical Systems*, **36**(2): 689-697.

[117] Y. Hatanaka, T. Hara (2004), “Automated analysis of the distributions and geometries of blood vessels on retinal fundus images”, *Medical Imaging 2004: Image Processing, Bellingham, Spie-IntSoc Optical Engineering*, **5370:** 1621-1628.

[118] K.R. Hoffmann, K. Doi, H.P. Chan, L. Fencil, H. Fujita and A. Muraki (1986), “Automated tracking of the vascular tree in DSA images using a double-square-box region-of search algorithm”, *Proc. of SPIE-The International Society for Optical Engineering, Medicine*, **626**:326-333.

[119] D. Calvo, M. Ortega (2011), "Automatic detection and characterisation of retinal vessel tree bifurcations and crossovers in eye fundus images", *Computer Methods and Programs in Biomedicine,* **103**(1):28-38.

[120] M. Ortega, J. Rouco, J. Novo, M.G.Penedo (2009), "Vascular landmark detection in retinal images", *Computer aided systems theory,* **5717**:211-217.

[121] A. Bhuiyan, B. Nath, J. Chua , K Ramamohanarao (2007)*, "*Automatic detection of vascular bifurcations and crossovers from colour retinal fundus images", *IEEE signal image technologies and internet-based systems,* 711-718.

[122] Shijian Lu (2011)*, "*Automatic optic disc detection from retinal images by a line operator, *Computer vision and image understanding,* **58**(1) 88-94.

[123] M. Niemeijer, B. van Ginneken (2010). "Retinopathy online challenge: Automatic detection of microaneurysms in digital colour fundus photographs." *IEEE Transactions on Medical Imaging,* **29**(1): 185-195.

[124] A Library for support vector machines (LIBSVM) http://www.csie.ntu.edu.tw/~cjlin/libsvm/.Date of last access: 09/02/2014.

[125] Berks, M., Chen, Z., Astley, S., Taylor (2011), C.: Detecting and Classifying Linear Structures in Mammograms Using Random Forests. In: Székely, G., Hahn, H.K. (eds.) IPMI 2011. LNCS, vol. 6801, pp. 510–524. Springer, Heidelberg

[126] Skeletonization http://www.inf.u-szeged.hu/~palagyi/skel/skel.html.Date of last access: 09/02/2014.

[127] R.M. Haralick, L.G. Shapiro (1992),"Computer and Robot Vision",**1**:170-171.ISBN:0201108771.

[128] J. Serra (1982), "Image Analysis and Mathematical Morphology". Academic Press, San Diego.

[129] R. C Gonzalez, R. E Woods, "Digital Image Processing", ISBN 8177581686.

[130] A.O Sykes (1986), "An Introduction to Regression Analysis".http://www.law.uchicago.edu/files/files/20.Sykes_.Regression.pdf.Date of last access: 09/02/2014.

[131] M. Sonka, V. Hlavac, R. Boyle (1998), "Image Processing, Analysis, and Machine Vision*",* second edition, 256:260.

[132] T.T. Tanimoto (1957), "IBM Internal Report".

[133] L R Dice (1945), "Measures of the amount of ecologic association between species," Journal of Ecology, **26**(3) 297-302.

[134] P. Jaccard (1912), "The distribution of flora in alpine zone ", New Phytosociology, **11**: 37-87.

[135] J. Cohen, " Weighted Kappa: Nominal scale agreement with provision for scaled disagreement or partial credit", *Psycological Bulletin*.**70**:213-220.
[136] J. Freixenet, X. Munoz, D. Raba, J. Marti, and X. Cuff (2002), "Yet another survey on image segmentation:region and boundary information integration," *Proc.EuropeanConfComputer Vision*: 408-422.
[137] Q. Huang, B. Dom (1995), "Quantitative methods of evaluating image segmentation," *Proc.IEEE Int'l Conf. Image Processing*, 53-56.
[138] Dogra, K Shaillay (2007), "Script for computing Tanimoto coefficient",*QSAR World*.
[139] T. Peng, I. H. Jermyn, V. Prinet, J. Zerubia (2010), "Extended phase field higher-order active contour models for networks", *International journal of computer vision*, **88**(1):111-128.