FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Cardiovascular Risks Detection Using Fundus Image

David Joel Nogueira Azevedo

Mestrado Integrado em Engenharia Informática e Computação

Supervisor: Prof. Luís F. Teixeira, FEUP

Second Supervisor: Dr. Inês Sousa, Fraunhofer

July 19, 2019

Cardiovascular Risks Detection Using Fundus Image

David Joel Nogueira Azevedo

Mestrado Integrado em Engenharia Informática e Computação

Approved in oral examination by the committee:

Chair: Doctor Henrique Lopes Cardoso
External Examiner: Doctor Catarina Carvalho
Supervisor: Doctor Luís Filipe Teixeira
July 19, 2019

Abstract

According to the World Health Organization, cardiovascular diseases have been the leading causes of death between the years of 2000 and 2016. Additionally, diseases of the circulatory system take up a high percentage of government budgets on health care systems, which could also be reduced with early detection. Recent studies have shown that age, blood pressure and the size of the optical disc vessels are correlated. Other studies claim that there is a relationship between retinal vessels and cardiovascular diseases.

Fraunhofer Portugal AICOS has been working on research and development of innovative solutions for handheld retinal fundus image acquisition and computer-aided detection of Diabetic Retinopathy. Fraunhofer's solution uses the EyeFundusScope prototype, a portable retinograph accompanied by an Android application, capable of processing the images captured by the device. This thesis aims to expand the potential use of this prototype, in the scope of preventive medicine, for screening Cardiovascular Diseases Risk through the automatic analysis of retinal fundus images.

During this dissertation, a fundus photographs dataset, annotated using the SCORE cardiovascular risk calculator, was produced. We proposed a retinal blood vessel enhancement pipeline for improving the contrast between the retinal background and these structures. The resulting images were used to train a MobileNet architecture, pretained on ImageNet, achieving a three-class classification of 58.82% for low, moderate and high cardiovascular risk. The final model was able to identify the presence of risk (moderate or high) with a 79.41% accuracy.

The final results are encouraging and lead to the conclusion that there is a direct relationship between retinal blood vessels and cardiovascular risk. The solution shows it is possible to obtain cardiovascular risk stratification directly from fundus images in a mobile setting.

Resumo

Segundo a Organização Mundial da Saúde, as doenças cardiovasculares têm sido as principais causas de morte entre os anos de 2000 e 2016. Adicionalmente, as doenças do sistema circulatório representam uma percentagem elevada nos orçamentos governamentais dos sistemas de saúde, que também poderiam ser reduzidos com uma detecção precoce. Estudos recentes demonstraram que a idade, a pressão arterial e o tamanho dos vasos do disco óptico estão correlacionados. Outros estudos afirmam que existe uma relação entre os vasos da retina e doenças cardiovasculares.

A Fraunhofer Portugal AICOS tem trabalhado na investigação e desenvolvimento de soluções portáteis e inovadoras para a aquisição de imagens de fundo da retina e para o diagnóstico assistido por computador de Retinopatia Diabética. A solução da Fraunhofer usa o protótipo EyeFundusScope, um retinógrafo portátil acompanhado por uma aplicação Android, capaz de processar as imagens capturadas pelo dispositivo. Esta tese tem como objetivo ampliar o potencial de uso deste protótipo, no âmbito da medicina preventiva, para triagem do Risco de Doenças Cardiovasculares, através da análise automática das imagens do fundo do olho.

Durante esta dissertação, uma base de dados de fotografias do fundus, anotado utilizando a calculadora de risco cardiovascular SCORE, foi produzido. Propomos uma *pipeline* para realçar os vasos sanguíneos da retina para melhorar o contraste entre o fundo da retina e estas estruturas. As images resultantes foram usadas para treinar uma arquitetura *MobileNet*, pré treinada na *ImageNet*, obtendo uma classificação de três classes de $58,82\%$ para risco cardiovascular baixo, moderado e alto. O modelo final foi capaz de identificar a presença de risco (moderado ou alto) com uma precisão de $79,41\%$.

Os resultados finais são motivadores e levam à conclusão de que existe uma relação direta entre os vasos sanguíneos da retina e o risco cardiovascular. A solução mostra que é possível obter estratificação de risco cardiovascular directamente a partir de imagens do fundo do olho num ambiente móvel.

Acknowledgements

I want to start by expressing my sincerest appreciation to Professor Luís Filipe Teixeira, from FEUP, and Dr Inês Sousa, from Fraunhofer AICOS Portugal, for the mentoring and supervision during my dissertation. Their input was valuable towards the completion of this work and their expertise proved to be essential and pragmatic.

To Fraunhofer AICOS Portugal for providing a balanced work environment with powerful and modern tools and a welcoming workplace setting. To FEUP for the effort put in by all members of staff which thrive for excellence when forming engineers.

A special thanks to my university friends for five years filled with laughter, joy and, most importantly, educational environment that helped me grow as a student, as a professional and as a person. To José Martins, my Erasmus buddy and close friend, for the support and advice given during these last months. To my friends and girlfriend for the emotional and motivational support provided.

Finally, an extraordinary thank you to my family for the outstanding support always given to me during all 23 years of my life so far. Their hard work helped become the person I am today and achieve my goal of graduating with a Masters Degree.

David Nogueira Azevedo

"Machine learning will not single-handedly determine the future, any more than any other technology; it's what we decide to do with it that counts, and now you have the tools to decide."

Pedro Domingos

Contents

List of Figures

LIST OF FIGURES

List of Tables

Abbreviations

AMD	Age-related Macular Degeneration
ANN	Artificial Neural Network
AREDS	Age-Related Eye Disease Study
AUC	Area Under Curve
AV	Arteriovenous
BMI	Body Mass Index
BP	Blood Pressure
CAD	Computer-Aided Diagnosis
CASE	Computer-Aided Software Engineering
CLAHE	Constrast Limited Adaptative Histogram Equalization
CNN	Convolutional Neural Network
CVD	Cardiovascular Disease
DBP	Distolic Blood Pressure
DNN	Deep Neural Network
DRIVE	Digital Retinal Images for Vessel Extraction
EFS	EyeFundusScope
FRS	Framingham Risk Score
HbA1c	Glycated hemoglobin
HRF	High-Resolution Fundus
kNN	k-nearest neighbours
MACE	Major Adverse Cardiac Events
NEI	National Eye Institute
REVIEW	Retinal Vessel Image set for Estimation of Widths
SBP	Systolic Blood Pressure
SCAS	Semana das Ciências Aplicadas à Saúde
SCORE	Systematic Coronary Risk Evaluation
SUACE	Speeded Up Adaptive Contrast Enhancement
SVM	Support Vector Machine

Chapter 1

Introduction

This chapter focuses on contextualising and defining the problem this study aims to solve. The first section 1.1 describes the context of this work. Section 1.2 justifies why this work is useful and whom it helps. It also sets the final goals and methodologies used to achieve them. To close this chapter, section 1.3 gives a brief overview of the remaining chapter's structure and describes their purpose.

1.1 Context

According to the World Health Organisation[1], cardiovascular diseases (CVD) have been the leading cause of death between the years of 2000 and 2016[Wor18]. Figure 1.1 shows the number of deaths, for the ten most common reasons, relative to the year of 2016. Improved risk prediction and detection will help in reducing these values. Additionally, diseases of the circulatory system take up a high percentage of government budgets on health care systems, according to Eurostat [Eur]. These numbers can be reduced with better risk stratification and preventive treatment. The need for the early detection of cardiovascular risks is clear.

In recent years, there has been a growing trend regarding the application of Machine Learning techniques. Machine Learning has seen a rise in the past years due to recent advances in computing power and storage capabilities. Medical applications are a big research topic nowadays not limited to but also given the lack of availability of medics and specialists in underdeveloped populations. Litjens 2017 *et al*. [LKB+17] provides an extensive list (out of 308 analysed papers) of Machine Learning applications in the context of medical image analysis. The application of such techniques to fundus images is not novel. However, the combination of these to predict CVD is somewhat a new concept.

[1] https://www.who.int/

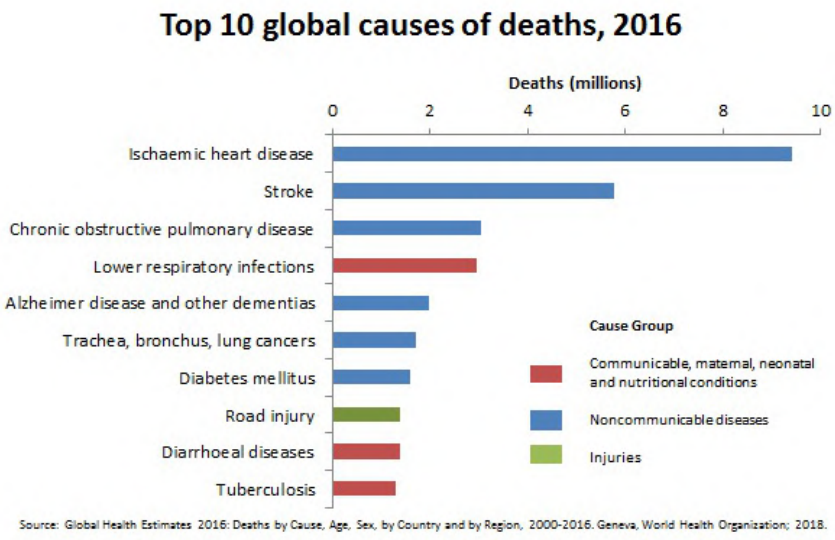

Figure 1.1: Top 10 global causes of deaths 2016 [Wor18].

Fraunhofer's [2] researchers have been developing a mobile application which integrates the EyeFundusScope (EFS) - a mobile retinograph 1.2 - capable of taking fundus images with a smartphone camera. The application then processes these images in order to detect diabetic retinopathy, using Machine Learning and Computer Vision techniques. The application shows great potential for future expansion of features and diagnostics. The Research Center's *Colaborar* [3] network is a group of public and private entities - hospitals, clinics, senior homes - that connects researchers to users, which is very useful for data collection.

1.2 Motivation and Goals

Medical studies have been trying to explain the relationship between age, blood pressure and optical blood vessel size [WKK+03] [KKMW07]. Other studies hint at a relationship between these blood vessels and cardiovascular problems [WKK+01] [MLM+09] [IdJV+04]. Witt 2006 *et al.* [WWH+06] studies the association between retinal microcirculation abnormalities and hypertension and concludes that the abnormalities visible in retinal vessels are an indication of death from stroke and Coronary artery disease, both diseases of the circulatory system. It wraps up by stating:

> *"A detailed assessment of the retinal microvascular network from digitized photographs may be useful in the noninvasive assessment of target organ damage and cardiovascular risk." [WWH+06]*

[2] https://www.fraunhofer.pt/en/home.html
[3] http://colaborar.fraunhofer.pt/en/

Figure 1.2: EyeFundusScope Version 0.8

We believe that this study can take advantage of the current computing power and leverage it with modern medicine in order to help both doctors and patients.

Fraunhofer's research with EyeFundusScope has been mainly focused on detecting diabetic retinopathy. However, there is a possibility of expanding this system to detect other problems, namely the risk of fatal CVD.

During an initial approach, the aim is to study the relationship between retinal blood vessels and cardiovascular risks. For this, a private dataset was created. The data acquisition was performed in a public environment using the EFS prototype. To estimate cardiovascular risk, the SCORE risk calculator was practised. As the central goal, this dissertation sets out to produce a model capable of detecting cardiovascular problems through fundus images as reliable as currently available techniques used by cardiologists. As a secondary objective, this model should be able to be integrated with a mobile application, to be used with the EFS prototype.

1.3 Thesis Structure

Besides this introductory chapter, this thesis contains five additional chapters. Chapter 2, explains key concepts, relative to this thesis, that a person with a general knowledge of computer science might not be familiar with. In chapter 3, the recent related work is shown as well as how this area has been developing over the years, including public databases and current technologies. This chapter describes the current problems and further motivates the tasks performed. In chapter 4, a more in-depth and complete problem statement is given and all the steps executed are described. Chapter 5 shows the final results obtained and gives a first look into the closing conclusions. Lastly, chapter 6 reviews the final conclusions, lists the limitations of the project and outlines possible approaches for future works.

Chapter 2

Background

This chapter focuses on explaining key concepts relevant to this work that a person with a general knowledge of informatics might not be familiar with. Section 2.1 explains what are fundus images, how they are obtained and why they are important. Section 2.2 introduces cardiovascular diseases. Section 2.3 gives a brief overview of deep learning and its purpose. Finally, section 2.4 introduces the most commonly used classification metrics.

2.1 Fundus Photography

A fundus image is a photograph of the back of the eye taken through the pupil. In it one can visualise the retina, the optic disk and the surrounding optical blood vessels [ST02]. It is used by ophthalmologists in order to diagnose diseases of the eye. To produce these photographs a retinograph is used, which is a medical tool designed for this sole purpose.

Fundus images are the only non-invasive way of observing blood vessels in the human body [PAM+06]. These structures present similar characteristics to those found in the human brain. Figure 2.1 shows an example of a fundus image. In the context of this dissertation, this type of images will be analysed and processed using machine learning and computer vision techniques in order to predict cardiovascular risks.

2.2 Cardiovascular Diseases

According to the World Health Organization [1], diseases involving the heart and blood vessels are grouped into Cardiovascular diseases. Some examples of CVDs are coronary heart disease, cerebrovascular disease, peripheral arterial disease, rheumatic heart disease, congenital heart disease,

[1] https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

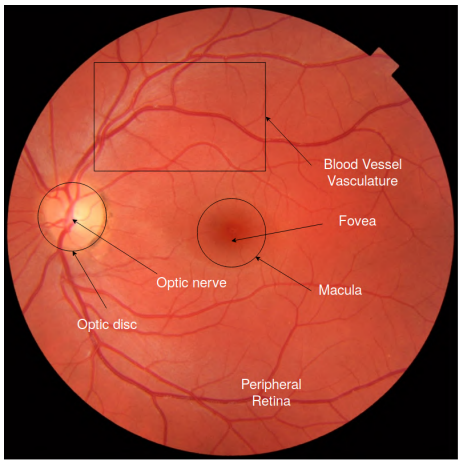

Figure 2.1: Example of a fundus photograph.

deep vein thrombosis and pulmonary embolism [1]. They are the number one cause of death worldwide [WH10]. Most CVDs can be avoided by simply changing routine habits. Early detection and preventive treatment is the best way to address these issues.

CVDs are diagnosed by taking into account several risk factors. The most common factors include smoking, hypertension, diabetes, unhealthy diet and obesity, physical inactivity. These risks contribute to a higher probability of major adverse cardiovascular events (MACE), such as heart attack and stroke, as well as other previously mentioned diseases. Heart attack (or myocardial infarction) and stroke are acute events caused by a blockage of the blood vessels resulting in less blood flow to the heart or brain. Both events result in permanent damage to the human body and in the worst case can be fatal.

To stratify the risk, cardiologists often use cardiovascular risk calculators. An example of such a calculator is the Systematic Coronary Risk Evaluation (SCORE) [Con03]. SCORE takes as input blood pressure, age, gender, smoking status and cholesterol levels and outputs the likelihood of fatal CVD within a ten-year period. After the assessment of risk, common preventive measures include: better dieting, regular exercise, stop smoking, blood-pressure-lowering medication [WH10]. It is important to note that these measures are merely preventive of future cardiovascular events and diseases.

2.3 Deep Learning

Deep learning [Sch15] is a subclass of algorithms of Machine Learning. It comprises artificial neural networks (ANN) with multiple layers, also known as deep neural networks (DNN). They have proven to excel at classification and pattern recognition tasks [Sch15], with results comparable to those of humans [CMS12]. An example of a simplistic ANN is presented in figure 2.2. As the image depicts, basic ANNs are composed of neuron layers, the first is called the input layer and the last is usually named the output layer while all layers in between these two are considered

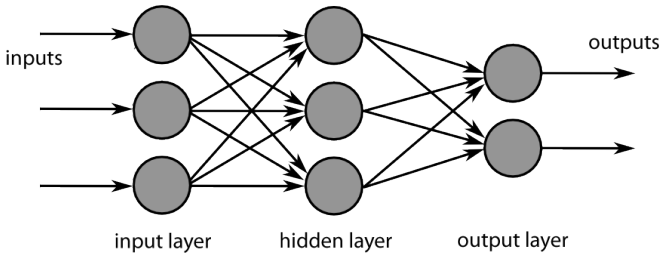

Figure 2.2: Example of a neural network.

hidden layers. For context, in an image classification problem, one would have a single input, the image, with the number of outputs being equal to the number of classes. It is recognised as deep learning once multiple hidden layers are present. Training is done by successively updating the values of the connections between nodes in consecutive layers. These values are called weights or parameters. The weights are initialised randomly and are updated to converge to a minimum of the desired error function [CMS12].

In traditional Machine Learning image classification problems, image preprocessing is a crucial part of the data preparation phase, which requires domain knowledge. Deep Learning gives us the freedom to have less domain knowledge by extracting features from the input images. In order to do this, convolutional neural networks (CNN) are typically used, where most hidden layers are convolutional layers [Dap] to the exception of the last few layers, which are mostly consisted by fully-connected layers, such as figure 2.2 depicts.

2.3.1 Convolutional Neural Networks

Although not a new concept, current advances in processing units have made CNN's the state-of-the-art for image classification problems [TSG+16]. A typical CNN architecture involves two stages, as it is shown in figure 2.3.

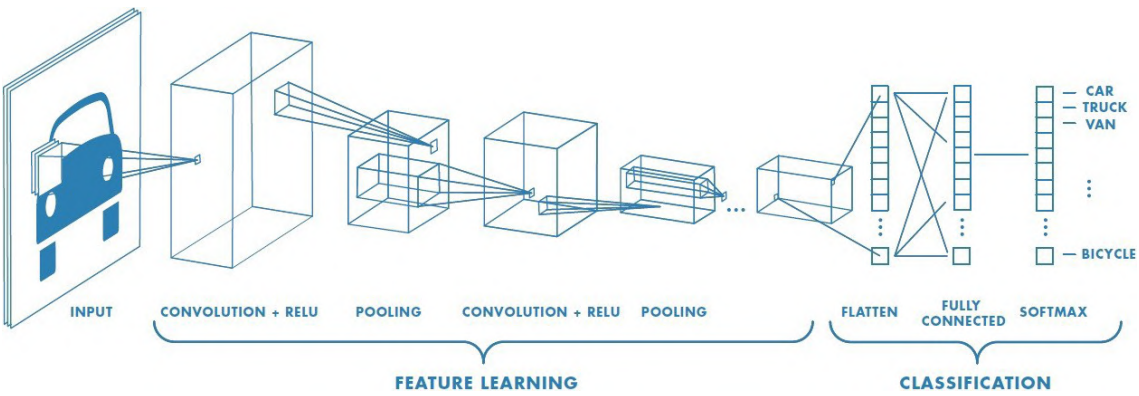

Figure 2.3: Example of a convolutional neural network architecture [TSG+16].

The feature extraction phase, labelled in the image as "Feature Learning", is where the network focuses on extracting features from the images in order to better classify them. This is done with convolution layers that apply convolution operations, shown in figure 2.4. In a convolution operation, a kernel, or filter, is shifted through the input image matrix. The sum of the element-wise product of the kernel with the corresponding portion of the input results in a single value, as depicted. In this case, the values of the filter matrix represent the weights.

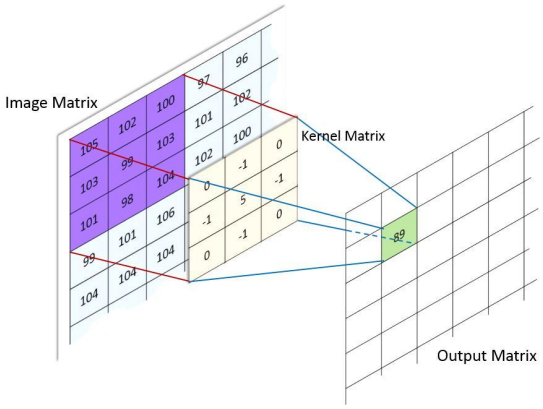

Figure 2.4: Low-level diagram of a convolution [Bri].

In the classification phase, the features are firstly flattened into a one-dimensional array and then pass through one or more fully-connected layers before passing through a classification function. The classification function divides the resulting values into the desired classes. Given the high number of parameters that need to be fine-tuned, CNN's demand high amounts of data [TSG+16].

2.3.2 Transfer Learning

In the context of Deep Learning, transfer learning is when the weights are initialised with values from previous training on different datasets and for different classification problems. Transfer Learning is relevant when the available data is scarce [LPX+17], as is the case in most medical imaging classification problems. High-level image features, such as colour, texture, corners and edges, are common between different types of images. This leads to less training to be needed for the model to learn how to classify the new input images.

2.4 Classification Metrics

To evaluate any system, metrics are necessary. Classification problems are not different. This section introduces some metrics which were used to assess the final results.

To better understand how these next equation work, core concepts are needed. Namely, the concepts of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN),

when considering binary classification. These four base concepts are used to calculate more complex values. True Positive is a case where a positive example is correctly predicted. True Negative, in turn, is when a negative example is correctly predicted. A False Positive example is when a model wrongfully predicts a negative example as positive. Likewise, a False Negative is the incorrect prediction of a positive example as being negative.

The accuracy metric, equation 2.1, represents the percentage of correct predictions made. In multi-class problems, accuracy is a good metric when there is a balance in the number of examples of each class.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.1}$$

Precision, given by equation 2.2, is the fraction of correctly predicted positive examples over the total number of predicted positive cases. This metric is relevant when diagnosing diseases, where it is necessary to know the likelihood of a person being ill when the model predicts so.

$$Precision = \frac{TP}{TP + FP} \tag{2.2}$$

Recall, equation 2.3, also known as sensitivity, is the fraction of correctly predicted positive examples over the total number of positive cases.

$$Recall = \frac{TP}{TP + FN} \tag{2.3}$$

F1-Score, equation 2.4, is the harmonic mean between recall and precision, encapsulating both values in one. The harmonic mean is close to the average when both values are close. When the values differ greatly, the harmonic mean is closer to the smaller value.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{2.4}$$

Chapter 3

State of the Art

In this chapter, the state of the art is described along with current open problems in the area. Section 3.1 shows what has been studied with Machine Learning applied to Ophthalmology. Section 3.2 describes different ways of how cardiovascular risks are detected, both in medicine and using computer-aided diagnosis (CAD). In section 3.3, an overview of fundus images datasets is provided. Section 3.4 states the most notable technologies relevant to this work. Finally, section 3.5 summarises the relevant aspects of this chapter and justifies the problem at hand.

3.1 Deep Learning in Medical Image Analysis

In the last decades, mostly due to the increase in computing power, machine learning has been a growing research area. Researchers and students have been trying to apply these techniques in many different areas. Indeed, Litjens 2017 *et al.* [LKB+17] presents more than 300 papers where Deep Learning was employed to medical image analysis. Better screening technologies have eased and improved the way doctors diagnose patients. Within this domain, the focus is mainly on segmentation, detection and classification. In recent years, the best results have been achieved using CNN's, particularly Transfer Learning due to the small sizes of available images and costly procedures to obtain them. Some results have even been able to outperform medical experts [GPC+16]. As emphasised by these contributions, the application of Deep Learning in the medical image analysis domain is very active and still has many open challenges. In such cases, the most prominent difficulties relate to the size of the datasets and the lack of data annotated by experts.

3.1.1 Opthalmology

Ophthalmology is one of the central areas where medical image analysis is conducted. Given that most diagnoses are performed visually by experts. The focus in this area revolves around

retinal image analysis, as described in Patton 2006 *et al*. [PAM+06]. Although these images denote primarily ocular diseases, they can also be used to analyse the retinal vascular morphology. This subsection presents studies where Machine Learning techniques were used over fundus images. Rahimy 2018 *et al*. [Rah18] comprises a review of numerous deep learning applications in the field of ophthalmology. Typical applications include abnormalities detection, segmentation of anatomical structures, quality assessment and diagnosis of eye diseases [LKB+17]. In Varadarajan 2018 *et al*. [VPB+18], researchers propose how to extract the refractive error from fundus photographs. This was the first proposal where the refractive error was automatically extracted from fundus images with high accuracy. Demonstrating that there is potential to make novel predictions using fundus images and Deep Learning. Choi 2017 *et al*. [CYS+17] applies Transfer Learning to the STARE database, mentioned in section 3.3, reporting a 72.8% accuracy over multi-categorical disease classification. Concluding that retinal disease detection from fundus images is a novel approach.

3.1.1.1 Diabetic Retinopathy

The most popular application is the detection of diabetic retinopathy [LPX+17], a diabetes-related disease which affects the human eye. Gulshan *et al*. [GPC+16] applies deep learning algorithms to detect diabetic macular oedema and diabetic retinopathy automatically. More accurately, CNN's were used, given their prowess with image analysis. Their solution involves using Transfer Learning with Google's Inception-V3 architecture with weights pretrained on the ImageNet dataset. This work still lacks improvement but establishes that medical imaging and deep learning are very much linked, and the possibilities are endless. A very similar study was done by Hemanth *et al*. [HDK19]. Although, in this case, the results achieved were better. Deep learning was preceded by the application of image processing and computer vision methodologies. This step ensured better contrast in the input pictures by highlighting the retinal structures. In this work, the dataset used was much smaller, showing that preprocessing the images resulted in better results, although it is possible that some degree of overfitting has occurred. A comparison was done between different classification techniques where Deep Learning showed to be the superior approach.

Garcia *et al*. [GSL+09] applies neural networks, on a total of 117 images, in order to detect hard exudates, an effect of diabetic retinopathy. The proposed solution is an ensemble of three neural network models each with different training for different types of analysis.

Orlando 2018 *et al*. [OPdFB18] proposes an approach to detect red lesion, a side effect to diabetic retinopathy. Here feature extraction was being done by hand by experts and Deep Learning managed to achieve results very close to those of experts. The important aspect of this work was the use of both CNN's and domain knowledge for feature extraction. With this, the authors managed to achieve fairly good results with a small dataset.

3.1.1.2 Retinal Vessel Analysis

The analysis of blood vessels in retinal images is quite popular given that blood vessels are the dominant structure in these photographs. The focus of research is largely in the segmentation of vessels. Some researchers study the vessels and relate them with cardiovascular risk factors [PVB+18], while others use this detection to remove the vessels from the image so that other structures become more noticeable. The segmentation of blood vessels located around the optical disk is a common application of machine learning to fundus images [FRH+12]. Various techniques have been applied to achieve it. Proper vessel segmentation is important in this context because this is the only relevant ocular structure when considering cardiovascular diseases [WWH+06].

Following a chronological order, Akita 1982 *et al*. [AK82] is the first proposal for a computer method for scanning fundus images. In Staal 2004 *et al*. [SAN+04], an approach using k-nearest neighbours (kNN) is used with a resulting AUC metric of 95%, which is excellent. Fast forward to 2015 where Maji 2015 *et al*. [MSG+15] describes the use of an ensemble - combination - of deep neural networks and random forests to detect retinal blood vessels. Lahiri 2016 *et al*. [LRSB16] had a very similar approach to that of Maji 2015 *et al*.and managed to surpass the previous work by 2% accuracy (from 93% to 95%). Leopold 2017 *et al*. [LOZL17] is the publication with the best results among current literature. They managed to achieve these results whilst only leveraging convolutional neural networks. The three last cited works are comparable given that all three worked with the DRIVE database.

An important note concerning this section is that many of the literature regarding optical blood vessel segmentation mentions the importance of this step towards classification and prediction of cardiovascular diseases and risks. A key difference between these works and this research is that none of them made use of low-quality fundus images. Publicly available datasets were all captured using standard retinographs.

The classification of blood vessels as arteries and veins is very relevant to this work since the arterio-venous ratio (AV ratio) is strongly related to cardiovascular problems [LADC18]. The AV ratio is typically used to identify hypertension [WKK+01], one of the most prevalent cardiovascular diseases. Although, for this to be performed, preceding precise segmentation is required. As Girard 2007 *et al*. [GC17] states:

> *"One clinical measure used to assess the severity of cardiovascular risk is the retinal arterio-venous ratio (AVR), which significantly depends on the accuracy of vessel classification into arteries or veins"*

In this study, they use a CNN and likelihood score propagation to distinguish previously segmented retinal blood vessels between arteries and veins. Their reported result of 93.3% accuracy was better than the state of the art at that time.

3.1.1.3 Other applications

According to Rahimy 2018 *et al*. [Rah18], other applications comprise glaucoma detection and age-related macular degeneration (AMD) detection. Many of these works focus on common deep learning architectures and show encouraging results. Such is the case of Grassmann 2018 *et al*. [GMB+18], which focuses on AMD prediction. In this work, an ensemble of 6 different neural network architectures bettered human graders, but the authors still consider their work to be unready for clinic practice. Likewise, Gomez-Valverde 2019 *et al*. [GVAF+19] studied how data set size, architecture, transfer learning and custom architecture affect the performance of a CNN for automatic glaucoma classification. Employing Transfer Learning with VGG19 architecture achieved the best results. These results approximate those of experts, with an AUC performance of 0.94.

3.2 Detecting Cardiovascular Risks

Predicting the risk of fatal cardiovascular disease in any given individual is crucial in order to administrate early treatment and preventive measures. These diseases include hypertension, stroke and heart attack. This section lists common ways used by cardiologists to evaluate future cardiovascular complications in patients and elaborates on recent attempts to achieve the same results with computer-aided diagnosis (CAD).

3.2.1 Using Traditional Medicine

Cardiovascular risks calculators are used to evaluate the risk of cardiovascular diseases in patients.

The Systematic COronary Risk Evaluation (SCORE) calculator uses smoking, age, total cholesterol, systolic blood pressure and gender to predict fatal cardiovascular occurrences over a ten-year period. The method was developed with data from 205,178 people and a total of 7,934 deaths related to cardiovascular problems [Con03]. SCORE is the most relevant calculator for European countries, including more than 2.5 million person-years of observations. The study resulted in two charts, as shown in figures 3.1 and 3.2. Populations were divided into two classes, low and high risk of CVD. Portugal's population is considered to have a low risk of CVD. Before this proposal only methods that targeted specific diseases were available. A drawback of this method is that it emphasises in fatal events. Another limitation is that the cardiovascular risk factors it uses are very limited and the addition of more factors would increase its accuracy. Nevertheless, the authors mention that the impact of new factors on the models' accuracy is uncertain given that too many factors do not guarantee a gain in predictive ability. An online interactive tool exists and it is recommended by the European Society of Cardiology [1].

The Framingham risk score is another calculator which is very similar to the SCORE calculator. However, it was performed much earlier than SCORE with a North-American population and fewer subjects [WDL+98]. The factors used in this study were equivalent to those of SCORE. A

[1] https://www.escardio.org/Education/Practice-Tools/CVD-prevention-toolbox/SCORE-Risk-Charts

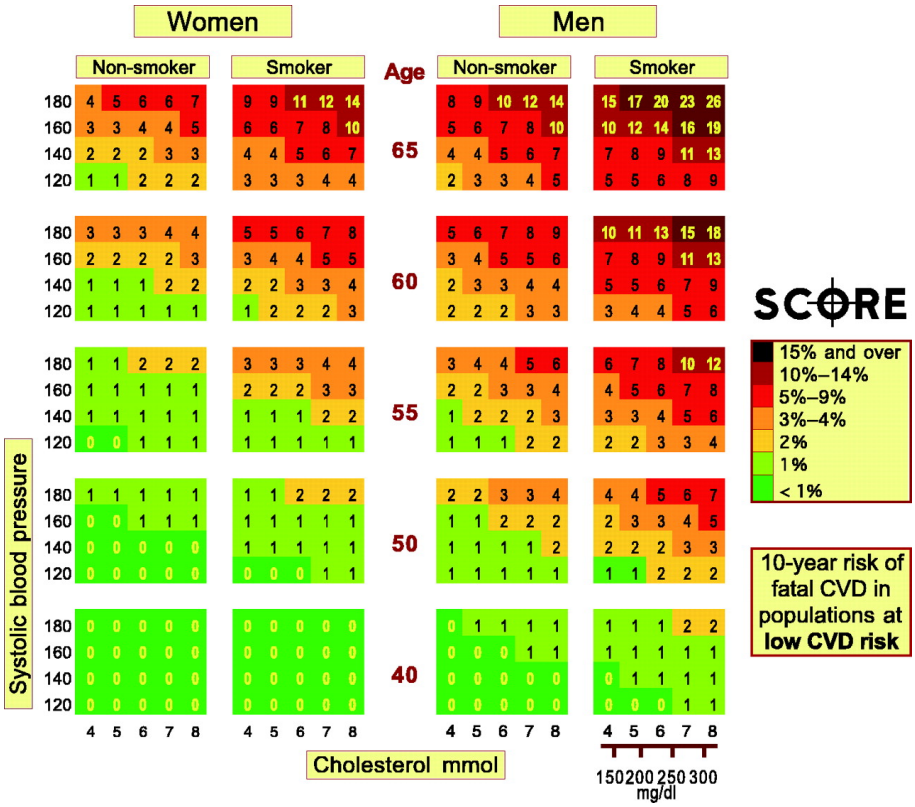

Figure 3.1: SCORE low CVD risk chart [Con03].

comparative study shows both these approaches to be close in terms of results [SKH+14]. Both Framingham and SCORE were able to generalise and stratify risk in an Asian population. A mutual limitation to both calculators is that many of the factors required are not readily available for every person, which makes the whole process more time consuming for everyone involved.

3.2.2 Using Computer Aided Diagnosis

Some studies can be found leveraging Machine Learning and fundus images for the prediction of cardiovascular risks. However, all of them seem to be limited to stroke prognosis.

Poplin 2018 *et al.* [PVB+18] starts by using the fundus image to predict cardiovascular risks factors (age, smoking status, gender, systolic blood pressure), which SCORE also uses. This forecast, achieved using CNN's, turns out to be relatively accurate. In the following step, these factors are used in combination with retinal fundus images to predict the risk of stroke in the patient within a 5-year timeline. The results show similar performance to that of SCORE being around 72%. The dataset used from the UK Biobank had data from over 250,000 patients. The authors approach focused on the basis that standard cardiovascular risk factors influence the optical blood vessels. Their work shows a clear correlation between risk factors and anatomical structures present in retinal fundus photography, such as the optical disk and blood vessels. The authors used the Inception-V3 architecture with a weight initialisation trained on the ImageNet dataset.

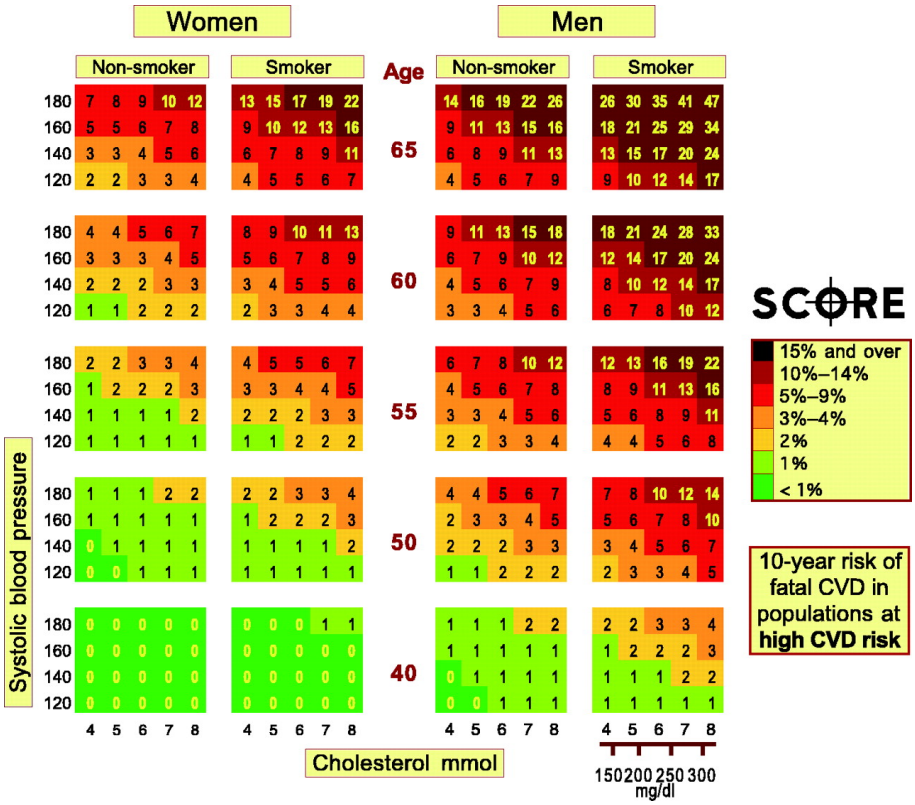

Figure 3.2: SCORE high CVD risk chart [Con03].

In RS 2017 *et al.* [RK17], fundus images were used to extract essential factors that indicate stroke risks. The prognostication of stroke is achieved through the prediction of retinal ischaemia. The use of artificial neural networks for classification is discussed. Results report an 89% accuracy for the prediction of retinal ischaemia. These results reinforce the relationship found between retinal images and stroke.

Jeena 2019 *et al.* [JSM19] studies the variations in the retinal vasculature to predict stroke risk. Factors such as branching coefficients and angle, AV ratio, fractal dimension and asymmetry factor were extracted from the images and used in an SVM classifier to achieve the results.

Finally, the work done by Fathalla 2016 *et al.*[FESG16] focuses on approximating a Machine Learning model to the Framingham Risk Score (FRS). The main characteristic is that the data used involved many different measures such as artery and vein responses over 350 seconds and retinal vessel diameters. The data collection comprises complicated and costly processes. After labelling the data using the FRS method, the authors define three groups, low, medium and high risk. The classes are highly imbalanced which required oversampling. A comparative analysis is performed using five established classification algorithms where Random Forest obtained the leading performance. This work combines the risk factors with manually computed metrics (Maximum diameter, Percentage Dilation, Dilation Amplitude and others) and uses them as input features, which makes this method unable to be performed automatically.

3.3 Fundus Image Datasets

Regarding fundus images, there are many available datasets resulting from previous studies. These studies mainly focus on the detection of diabetic retinopathy and retinal blood vessel segmentation. The list of available datasets found so far, which are relevant for vessel segmentation is as follows, in no particular order.

- The **High-Resolution Fundus** [2] (HRF) dataset was created for the purpose of supporting the study of segmentation algorithms on retinal fundus photographs. It is freely available for research studies. In total there are 45 high-resolution images divided into 3 equally sized groups - healthy patients, patients with diabetic retinopathy and glaucomatous patients. The annotations available for this data which is useful for our application scenario is the retinal blood vessel segmentation that was performed by a group of experts [KBK+13].

- The **Age-Related Eye Disease Study** [3] (AREDS) is a study conducted by the National Eye Institute (NEI). It contains over 135,000 fundus images with high-quality of more than 4,600 participants. An important and deciding factor of this dataset is that the participants were clinically followed during 5 or more years. The medical annotations related to diseases these subjects experienced during the study is of great importance for our work. In total, 124 publications have used this dataset. Varadarajan 2018 *et al*. [VPB+18] has used this dataset to predict refractive error with deep learning applied to fundus images.

- The **Digital Retinal Images for Vessel Extraction** [4] (DRIVE) is a dataset made for retinal blood vessel segmentation [SAN+04]. It is also freely available for research and educational purposes. This dataset was not used due to its small size and lack of cardiovascular risk classification, nevertheless, it is worth mentioning since many publications used it and reference it [LADC18, LOZL17, MSG+15, LRSB16].

- The **Retinal Vessel Image set for Estimation of Widths** [ADHS+08] (REVIEW) dataset comprises 16 images with 193 vessel segments and a total of 5,066 manually marked profiles. The importance of this dataset is that the blood vessels have been manually measured by experts. This gives us a valid dataset to evaluate the performance of segmentation.

- The **UK Biobank** [5] dataset is the biggest and most relevant dataset for this study found to date. It tracked more than 500,000 volunteer participant over the course of 5 years. Fundus images are annotated with several cardiovascular risks factors as well as clinical conditions of the subjects. Poplin *et al*. [PVB+18] and Varadarajan *et al*. [VPB+18] used this dataset to predict the risk of stroke using deep neural networks with fundus images and the refractive error respectively. The major drawback is the slow, and costly, application procedure which can take up to 3 months to be approved.

[2] https://www5.cs.fau.de/research/data/fundus-images/
[3] AREDS Study webpage
[4] https://www.isi.uu.nl/Research/Databases/DRIVE/
[5] http://www.ukbiobank.ac.uk/about-biobank-uk/

Other datasets found which were not considered are: STARE [6] which has used by Choi 2017 *et al*. [CYS+17] to detect retinal diseases with the aplication of deep convolutional neural networks, DIARETDB1 [7], DRIONS [8], Project Baseline [9].

All of these datasets are relevant for researchers who wish to study fundus images. Given that almost every dataset found does not contain annotations related to cardiovascular risk factors and diseases, the more feasible option is to build a private dataset using the EFS prototype and the SCORE risk calculator.

3.4 Technologies

There are many robust machine learning libraries in production today. Here we will address only the ones chosen for solving the problem at hand.

For this study, tensorflow [10] is the best solution. Tensorflow is an open source Machine Learning library developed by the Google Brain team and used by many industry leaders and researchers alike. It has strong support for deep learning which is the main reason for this choice. This library was used to build, train and validate deep neural networks. Tensorflow will be used via the keras library [11], which provides a friendlier API. Keras has built-in methods, architectures and weights for the ImageNet dataset. The aforementioned out-of-the-box methods allow for straightforward implementation of Transfer Learning and data augmentation.

OpenCV [12] is an open source library focused on image processing and computer vision. It was particularly useful to segment and enhance the retinal blood vessels in fundus photographs.

3.5 Summary

First and foremost, this chapter establishes that the use of Deep Learning in Medical Image Analysis has become a prevailing topic of research. Particularly in the domain of ophthalmology with fundus images.

From this chapter, we can conclude that although different ways of assessing a patients' risk of cardiovascular events exist in the medical area, these have many limitations. Many deep learning approaches were proposed but so far no one has been able to solve the problem with better results than those of methods such as SCORE or Framingham. There is room for improvement in the computer-aided diagnosis of cardiovascular risks. Furthermore, CNN's have shown great results on segmenting retinal blood vessels and classifying multiple diseases in fundus images. Figure 3.3 shows the relationships that were studied and those that still lack further investigation. From the literature overview provided in this chapter, it is clear that there is a bi-directional relationship

[6] http://cecas.clemson.edu/ ahoover/stare/
[7] http://www.it.lut.fi/project/imageret/diaretdb1/
[8] http://www.ia.uned.es/ ejcarmona/DRIONS-DB.html
[9] https://www.projectbaseline.com/
[10] https://www.tensorflow.org/
[11] https://keras.io/
[12] https://opencv.org/

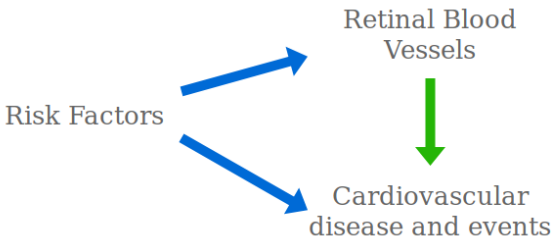

Figure 3.3: Relationship between risk factors, retinal blood vessels and cardiovascular disease and events.

between cardiovascular risk factors and retinal blood vessels, as shown in subsection 3.2.2. In the image, blue arrows represent the dominant focus of study from the research community, while the green arrow represents the relation which lacks focus. In this thesis, the goal will be to achieve cardiovascular risk stratification directly from fundus images. For this purpose, there is a need to create a dataset where fundus images will be acquired in parallel with cardiovascular risk stratification. Given the low amounts of data collected, Transfer Learning was the most viable option. This fact is backed by the current literature.

Chapter 4

Methodology

The central idea of this chapter is to outline the methods used to produce the results. Section 4.1 starts by defining the problem which this work is trying to solve. Section 4.2 illustrates the data acquisition protocol and how the dataset used was acquired. It wraps up by presenting an overview of the available data. In section 4.3, the image preprocessing and data cleansing steps are defined. It focuses on how the dataset was filtered by excluding low-quality examples. Section 4.4 begins by listing the deep learning architectures employed, how parameter optimisation was accomplished and how model selection was performed, finishing by presenting the obtained validation metrics. Finally, in section 4.5 a summary for this chapter is done and some preliminary conclusions are taken.

4.1 Problem Statement

Cardiovascular risks are currently calculated using medical formulas which take into account many factors. Very often, these factors are not readily available for every patient and require additional tests. Another problem is that this process requires patients to travel to a hospital. However, for some populations, this is nearly impossible or expensive. As such, a portable solution, and one that requires less invasive measures is ought to be of great value for both patients and medical staff. In light of these limitations, this thesis aims to further expand the utility of fundus images in order to predict diseases of the circulatory system.

Fundus images have been mainly used for ophthalmology diagnosis, nevertheless, studies indicate that these images might contain valuable insights in other areas, such as neurovascular and cardiovascular diseases. Medical researchers have shown the existence of a strong relationship between the optic blood vessels and cardiovascular problems. Indeed, some studies have used Support Vector Machines (SVMs) to distinguished healthy fundus images from those of victims of stroke.

This thesis intends to further study the link between fundus photographs and cardiovascular diseases. By combining risk stratification and risk factors with these photographs we want to be able to predict problems - like hypertension and stroke - in the next ten years using deep neural networks, more precisely CNN's. Convolutional Neural Networks work best with big amounts of data. The principle is that the model will only be as good as the data we provide for training.

Another possible problem regarding this work is related to the quality of the fundus photographs obtained with the EyeFundusScope prototype in comparison to those of regular retinographs. This represents a problem due to the small dimensions and distinguishability of the retinal blood vessels. Given that CNN's work with lower resolutions, the vessel dimension problem is minimised, but not solved.

The expected result is a model capable of predicting cardiovascular risks with a performance at least as good as current calculators designed for this purpose. At the same time, complexity should be kept at a minimum in order for the resulting model to be used in a mobile setting.

4.2 Data Collection

Due to the seemingly nonexistence of public datasets which relate cardiovascular risk stratification with fundus images, a data collection specific to this dissertation was performed. The acquisition made use of the EFS prototype, as shown in figure 1.2. During this process, the focus was taking fundus photographs of consenting patients from the general population while at the same time stratifying their cardiovascular risk using the SCORE table for low-risk countries, present in figure 3.1. From the 20th to the 23rd of March 2019, we participated in the *Week of Applied Sciences in Health* (SCAS, in portuguese) event at Coimbra's Alma Shopping. It is coordinated by the student association of the *Escola Superior de Tecnologia da Saúde de Coimbra* and unifies students and patients to offer free medical evaluations to the general public.

A total of 100 people were surveyed. The examination consisted of taking multiple fundus photographs of the persons left and right eyes using the EFS prototype, measuring the blood pressure and pulse wave, and SCORE risk stratification. In order to stratify cardiovascular risk, a small questionnaire was employed, this questionnaire can be found in appendix B. For SCORE, only the age, gender, smoking status, systolic blood pressure and total cholesterol level are required but other parameters were also recorded so that this data may be used in other studies. Priority was given to people who had already measured their cholesterol levels in a different booth. The pulse wave measurement is a cardiovascular risk factor in itself [AAJMC07], despite not being used in this work. So that the data could be used in this study, an informed consent was needed from the participants. The document used can be found in appendix A.

4.2.1 Acquisition Protocol

A protocol was defined and established so that other studies may also use the data collected with an emphasis on consistency and scientific rigour. The most important aspect of this process was to keep the same conditions throughout data collection. Ambient and prototype settings were

the most impactful variables. Due to confidentiality reasons, specific aspects of the EFS prototype and its settings cannot be disclosed. The parameters for the accompanying EFS mobile application were set before the first examination and remained unchanged for the entirety of the process.

When performing fundus image acquisitions, a dark environment is highly recommended. In a setting where the lighting is minimal, pupil dilation occurs, meaning that the 'window' to the person fundus is maximised. For this purpose, a dark room with low-intensity lighting was used, which can be seen in figure 4.1. A chair was placed at 2.50 meters from the wall and a fixed spot (marked with a red circle in the middle image) at 2.00 meters high was identified. The patients were asked to look at this spot while the picture was being taken. With this, it was ensured that the resulting photos remained consistent throughout the acquisition process, since it allowed the optic disc to be aligned more easily by the experts handling the prototype.

Figure 4.1: Environment used during data acquisition.

Since it takes some time for the human pupil to dilate and get used to a darker surrounding, this time was spent informing the subject about the scope of the project and the signing of the informed consent. A medical researcher was responsible for measuring the persons' blood pressure and pulse wave with an appropriate device while filling out the provided survey. Ultimately, the persons' risk was classified using the measured factors according to the SCORE risk table.

4.2.2 Data Overview

As mentioned previously, one hundred people participated in the data collection. The factors required to calculate the SCORE risk were surveyed. SCORE outputs a percentage of the likelihood that the patient has of having a fatal cardiovascular disease over a ten-year period. These percentages are divided into ranges from $< 1\%$ to $\geq 15\%$, as depicted in figure 3.1. These values are often aggregated into ranges. Given that CNN's work very well in image classification, and that the dataset size comprises very few examples, a decision was made to group SCORE outputs in three classes. The classes defined were '*Low Risk*' for values smaller than 1%, '*Moderate Risk*' for values ranging from 1% to 15% and '*High Risk*' for values higher than 15%. This decision was

based on the fact that for preventive treatment to be administrated, the actual percentage resulted from SCORE is irrelevant, knowing that a person as a moderate or high risk is enough.

Characteristics	Value
Number of people	100
Gender: % male	38%
Ethnicity	5% Black, 95% White
Current Smoker: %	14%

Table 4.1: Demographics on data collected.

As table 4.1 shows, the population included in this study is predominantly Caucasian. These values are adequate to keep in mind because cardiovascular diseases affect different cultures and races in different ways. In figure 4.2, the distribution of age groups and risk classes are presented. In terms of classes, this dataset has a reasonably satisfactory balance, contrary to most medical image analysis datasets, as stated in chapter 3. The age groups represented in the chart were not chosen at random. In SCORE, age is one of the most deciding factors. If a person is younger than 40 years and does not have any previous cardiovascular disease or event, their risk is immediately classified as low (lower than 1%). In the same way, people who are over 65 years of age, even non-smokers with healthy cholesterol and systolic blood pressure levels, are considered to be at high risk (equal to or higher than 15%). This threshold is extremely steep and it is important to be aware of. The presence of previous cardiovascular disease or event automatically attributes the classification of high risk to an individual. Overall, the dataset obtained is rather small in size, especially in the context of Deep Learning.

It is necessary to note that risk labels are given per person, while images are considered independent per pair of person-eye (left or right), meaning a total of 200 entries were collected. That occurs because eyes have differentiating patterns, much like fingerprints from different hands [RCS+16].

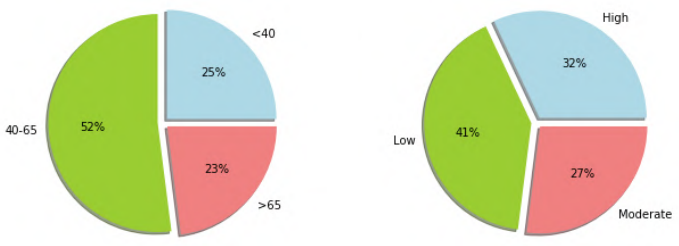

Figure 4.2: 1. Age distribution chart. 2. Risk classes distribution chart.

4.3 Data Preprocessing

After collecting the dataset, which is very relevant for this study, the next logical step is to handle the data and preprocess the images. This step is necessary to ensure the data, that will be used to train the chosen Deep Learning models, will have the potential to solve the problem at hand and contain consistent examples. Preprocessing is even more relevant with this dataset because the quality of the obtained fundus images is subpar when compared to medical retinographs, since the competitive advantage of the EFS lies on mobility. The prototype and the accompanying Android application are not yet ready for public release and still lack further improvements.

The main goal of this process is to ensure the removal of bad examples within the data. By observing the dataset, a common occurrence is the existence of unwanted reflections in the image in people who have been operated to correct cataracts. In this procedure, the optic lens is broken and replaced with an artificial one made of a special type of glass. The glass then results in unwanted reflections from the EFS flash at the moment of capture. This effect can be observed in the second photograph of figure 4.3.

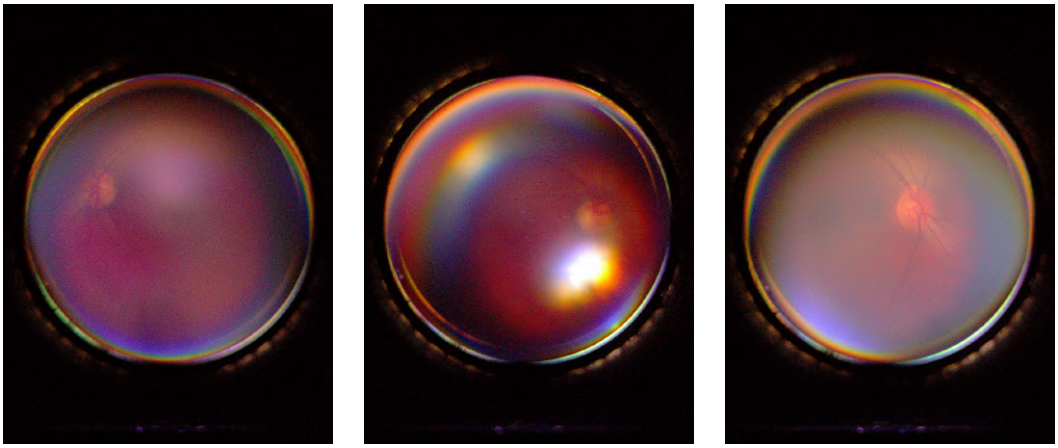

Figure 4.3: Examples from the acquired SCAS dataset.

Figure 4.4 displays some examples from the DRIVE database, mentioned in section 3.3. As it is visible, the quality of both datasets is clearly different. DRIVE images are more consistent in terms of sharpness and tone. This database is used in most studies related to retinal vessel segmentation. Another clear distinction is the number and overall visibility of the vessels present. In the SCAS dataset, vessels are most distinguishable in and around the optic disc.

4.3.1 Manual Cropping

Given the low quality and poor contrast in areas further away from the optic disc, a decision was made to crop the images around the disc. This structure is where retinal blood vessels converge, which means they are thicker and more prominent in and around it. Ideally, an automatic pipeline should be created which is able to achieve risk prediction from the initial photographs. To make the images more uniform in terms of shape and tone, we want to locate the optic disc and crop

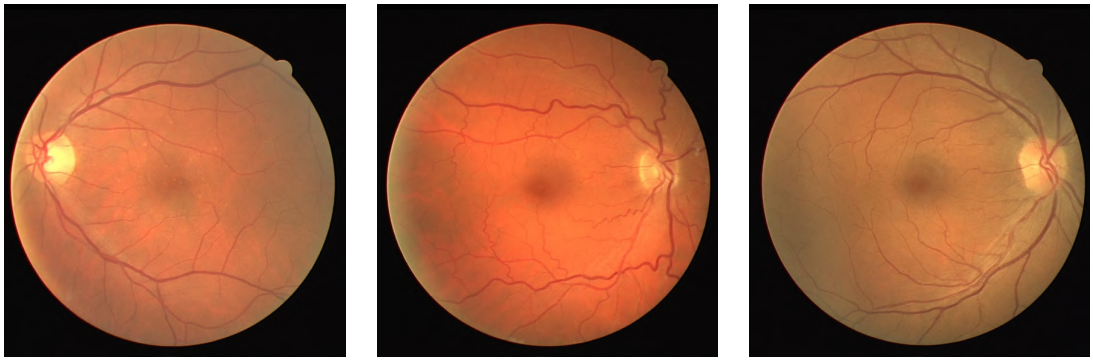

Figure 4.4: Examples from the DRIVE database [SAN+04].

the surrounding area. The result should be a more apparent vessel distinction from the noisy background. In the SCAS dataset, this step is especially important since the peripheral regions of the fundus picture have many reflections and blurs which do not hold any viable knowledge about the persons' health.

Some initial attempts to automatically detect the optic disc for further cropping were made but, given the difference between SCAS and other fundus datasets, results were discouraging. One such attempt consisted on the application of MNet_DeepCDR[1] [FCX+18b] [FCX+18a], the state-of-the-art solution for Glaucoma screening. One of the first and most critical steps of this project is the segmentation of the disc, which is the structure affected by the Glaucoma disease. With this approach, it was impossible to detect and segment the disc with good consistency. For these reasons, a tool to manually crop these images was created.

The tool requires an evaluator to indicate the centre of the optic disc, by clicking on the image. The cropper then proceeds to use the chosen point as the centroid of a square with a size of 840 pixels. The size of 840 pixels was obtained by manually measuring the optic disc in three images, averaging and duplicating the value to capture the vessels around the optic disc. Examples of the resulting images from this process can be viewed in figure 4.5. The resulting pictures are now square, with an 840x840 dimensions, centred in the disc and uniform in terms of shape and the location of the eye structures, differently to figures 4.3 and 4.4.

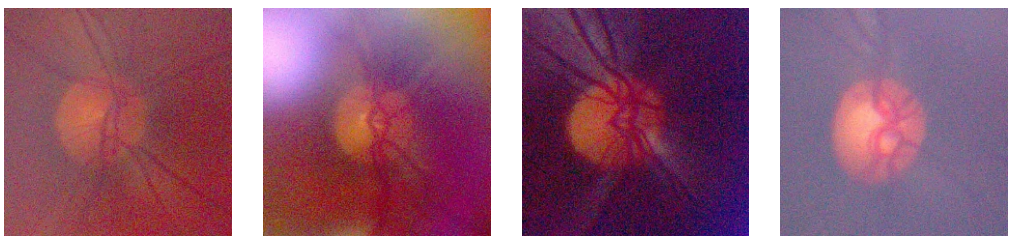

Figure 4.5: Examples of resulting images after cropping.

[1] http://hzfu.github.io/proj_glaucoma_fundus.html

4.3.2 Vessel Enhancement

Like previously mentioned, and according to current literature, blood vessels seem to be the ocular structure most influenced by cardiovascular risk factors. Considering that the images are noisy, and vessel contrast varies a lot from different subjects, in order to focus the learning in the retinal vessels, a vessel enhancement pipeline was established. This pipeline produced a new preprocessed dataset for later comparison. The purpose of the dataset is to examine if models will perform better with information solely based on the vessels and without bias from image quality, noise, reflections and tones.

The proposed pipeline was inspired by the work of Bandara 2017 *et al.*[BG17]. This work focuses on comparing different contrast enhancement techniques. Different techniques were compared based on true positive rate, false positive rate and accuracy, the top performer overall was the Speeded Up Adaptive Contrast Enhancement (SUACE) algorithm, which the authors developed themselves. The Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm came in as a close second showing a performance very similar to that of SUACE. The final decision was to use CLAHE as it comes implemented with OpenCV, given that the difference in performance to SUACE was minimal.

The established pipeline consists of four steps, which were implemented using the OpenCV library. Intermediate and final results are shown in figure 4.6.

- The first step was to **resize the image**. Resizing an image is a type of denoising from the way the pixels are grouped and averaged in order to reduce the size. The original cropped images were 840 by 840 pixels. Since most Deep Learning architectures expect input of 224 by 224, as will be detailed in section 4.4, this was the chosen target for the resize operation.

- Secondly, **gray scaling** was applied. By gray scaling the images, the intention is to standardise the dataset and reduce the influence of different shades, which resulted from imperfections from the EFS prototypes' parameters.

- The third step was the application of the **CLAHE** algorithm. The focal point of CLAHE, as the name suggests, is improving the contrast of a picture. Histogram equalisation works by stretching the pixel values, resulting in higher value differences, further separating the light and dark pixels in terms of intensity. Simple histogram equalisation regards solely the global contrast of a picture. CLAHE applies the same principal to smaller patches of the image at a time, called tiles. This works particularly well inside of the optic disc. The optic disc is a bright structure and by applying CLAHE instead of normal histogram equalisation the contrast is enhanced much better inside the disc. The contrast limiting aspect of CLAHE further helps reduce noise by avoiding to amplify it in the equalisation.

- **Gaussian Blur** is used to smooth edges by blurring the image using a Gaussian filter. This method is extremely powerful at reducing Gaussian noise.

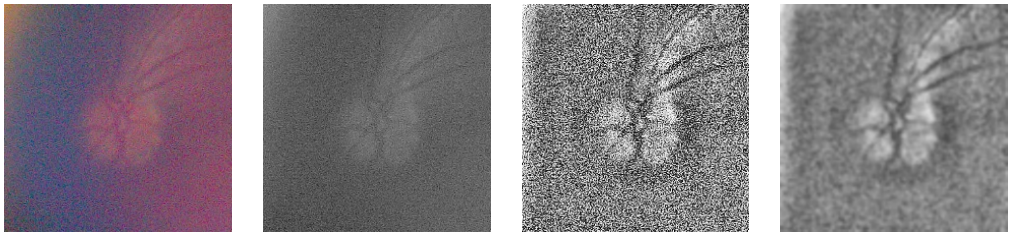

Figure 4.6: Intermediate steps of vessel enhancement. From left to right:
a) Resize (224x224). b) Gray Scale. c) CLAHE. d) Gaussian Blur.

By observing the progress of the pipeline from start to finish, in figure 4.6, there is a clear improvement in image contrast, principally concerning blood vessels. With this method, a second dataset, to be assessed, arose. From this point on, two datasets are to be considered, the coloured dataset and the vessel enhanced one. The new dataset focuses on studying the direct relationship between retinal blood vessels and cardiovascular risks.

4.3.3 Outlier Removal

Upon observing the resulting images, two types of outliers were detected. Outliers are observations that significantly diverge from regular records. Outlier removal ensures that these cases do not influence negatively the results given that they do not yield valuable knowledge, although they can help reduce overfitting.

The first type identified was images that had minimal contrast, as shown in figure 4.7. To remove these examples, the standard deviation was calculated on the images outputted by the method detailed in subsection 4.3.2. Low standard deviation values translate to images which do not have enough contrast for the retinal vessels and optic disc to be discernible.

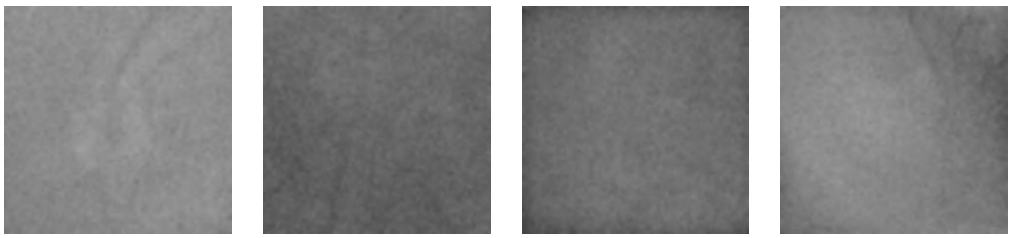

Figure 4.7: Examples of outliers with low standard deviation.

The second identified type consisted of images with white reflections. As mentioned previously, this occurrence was common in patients who have had cataract surgery. In most instances, these reflections were present on the outermost regions of the photograph and were effectively cropped out with the process defined in subsection 4.3.1. To further exclude these cases, a threshold on white pixels was applied. If the image included a substantial number of completely white pixels they were discarded. Some examples can be observed in figure 4.8. As can be seen, the presence of large reflection spots highly reduced the distinctness of the ocular structures.

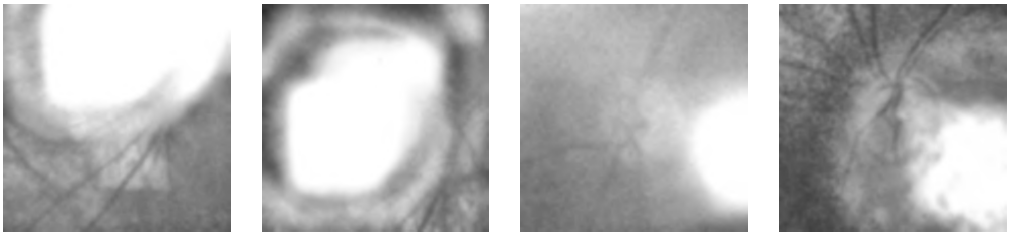

Figure 4.8: Examples of outliers with white reflections.

At the end of these processes, images excluded were also removed from the original coloured dataset. A total of 21 photographs were filtered out by having a low standard deviation and 6 for big white spots caused by reflections. The final image count was 432.

4.4 Model Training

This section focuses on explaining how the training of Deep Learning models occurred. Only pre-established models were considered for this work. These architectures have the advantage of being already publicly available within the keras framework and with weights trained on the ImageNet [2] dataset. The ImageNet dataset is a well-known performance benchmark for Deep Learning architectures. The database comprises over 14 million images hand labelled in over 21 thousand different categories. The models have already been subject to various investigations and applied to different contexts using Transfer Learning. A custom architecture was not defined in this project. Multiple studies demonstrate that pretrained models achieve similar or greater performance comparative to problem specific architectures [TSG+16] [LKB+17] [LPX+17], as detailed in chapter 3.

Furthermore, the chosen models were trained with two axes of comparison in mind. The first axis is the vessel enhancement preprocessing, denoting two datasets (**grayscale** and **coloured**). The second axis is the weight initialisation scheme (**ImageNet** and **random**) which tests whether the models are able to transfer the learnt feature extraction from the ImageNet dataset to the fundus photographs dataset.

4.4.1 Pretrained Networks

Based on the fact that one of the secondary objectives of this work is to achieve a model capable of being integrated into a mobile application for cardiovascular risk prediction, all mobile-friendly architectures preloaded with keras were subject of training and comparison. Mobile networks trade-off slightly lower performance for fewer requirements in computing power and memory space, allowing them to be used in mobile applications. Additionally, the InceptionV3 architecture was also considered since it is state-of-the-art for transfer learning in medical image analysis. InceptionV3 was the model used by several studies, such as Poplin 2018 *et al.*[PVB+18], giving

[2] http://www.image-net.org/

this study a term of comparison with current literature and a comparative baseline to the theoretically less powerful mobile architectures.

The model architectures that will be taken into account are **Mobile Net**, **Mobile Net V2**, **NASNet Mobile** and **InceptionV3**. A comprehensive table comparing these four models in terms of the number of parameters and accuracy on the ImageNet dataset can be observed in table 4.2. As the table shows, the performance comparison is very little compared to the reduction in the total number of parameters. A smaller number of parameters results in faster training and faster prediction once the model is optimised. All three of the mobile networks expect as input 224 by 224 images (in pixels) while the InceptionV3 network accepts 299 by 299 images, although it also accepts the previous format. An important note is that all of the architectures used expect an input with 3 channels for RGB. The grayscale images were saved with these three channels to comply with the expected input.

	Parameters (in millions)	**Accuracy on ImageNet**
Mobile Net [HZC+17]	4.2	70.6%
Mobile Net V2 [SHZ+18]	3.4	72%
NASNet Mobile [ZVSL17]	5.3	74%
Inception V3 [SVI+15]	23.8	76.6%

Table 4.2: Comparison between models in terms of number of parameters and accuracy on ImageNet dataset.

4.4.2 Train-Validation-Test Split

The next logical step is to divide the datasets (grayscale and coloured) into 3, the train, validation and test sets. Each of these sets has a specific goal.

- The **train set** is used to iteratively train the models' parameters which are fit to optimise prediction on these images.
- The **validation set** is to check the model performance on unseen data after each iteration. This allows checking whether the model is able to generalise the information learned from training to unseen data.
- The **test set** remains untouched during training and is used at the moment to report final results on model performance in unseen data.

There are two commonly used splitting values, 70-15-15 and 80-10-10. Since the amount of available data is limited, a decision was made to use the 70-15-15 approach. At first, the method was applied over the entire images, which, recalling the explanation given in section 4.2, contain up to four slightly different examples for the same persons' eye. This first try resulted in a precise split over the three established classes in which similar data was present both in the test and

validation sets. To avoid possible data leakage [3] a new split was done. In this new split, instead of considering images as single items, they were grouped by different eyes (all different images for patient 001s' left eye were grouped) and afterwards the splitting was applied. The resulting split is detailed in table 4.3. The exact same split was applied to both datasets.

	Low	**Moderate**	**High**	**Total**
Train	123	82	88	293
Validation	37	15	19	71
Test	16	23	29	68
Total	176	120	136	432

Table 4.3: Train-Validation-Test Split overview.

4.4.3 Training

All 16 distinct models, permutations between four different architectures, two datasets and two separate weight initialisation schemes, were trained under the same circumstances. These parameters were fine-tuned first using the Mobile Net architecture and fixed throughout all individual permutations. The fixed parameters include the classification layers, as shown in figure 2.3, the data augmentation scheme, the optimiser and its learning rate and the model selection scheme.

Data augmentation is a procedure used to artificially enlarge the size of a dataset with modified versions of the pictures it contains. Other benefits include a reduction in overfitting and better generalisation ability of the model. When implementing data augmentation, different modifications can be applied, however, it is essential to maintain the resulting images as realistic as possible. For this reason, only alterations which do not change the vessels' signature were used. These include random rotations of up to 45 degrees, up to 5% width and height shift and both horizontal and vertical flipping.

The chosen optimiser was Adam [KB14], which is an adaptive learning rate optimisation algorithm. This selection was made according to state-of-the-art literature [RKK18]. The learning rate chosen for Adam was 1.0×10^{-4}. This value controls the speed at which the neural network learns information by updating its weights. The optimiser typically does not vary the performance widely and as such this decision was trivial.

In order to classify the input images into the desired classes, a fully connected architecture was implemented on top of the original backbone of the chosen neural network. These last layers are responsible for converting the features extracted by the initial layers into prediction probabilities for each class. The first added layer is a Global Average Pooling layer, a common practice, which is responsible to downsample the extracted features into a one-dimensional array by computing the mean value for each feature map. In this work, the output of this first layer is then fed into successive fully connected (or dense) layers. The scheme used can be observed in figure 4.9. A

[3] Data leakage happens when knowledge not present in the training dataset influences the models' training, leading to a biased model. This phenomenon produces invalid and overly promising predictive models. A common way to detect data leakage is when model performance seems too good to be true.

total of four dense layers were added, the first three of these are normal *'relu'* activated layers which focus on reducing the feature space into 512 neurons and then 256. The last dense layer, also called the *softmax* layer, is where class probability is calculated. The *softmax* layer outputs three values between the values of 0 and 1, which have a direct correspondence with the likelihood of each class (low, moderate, high).

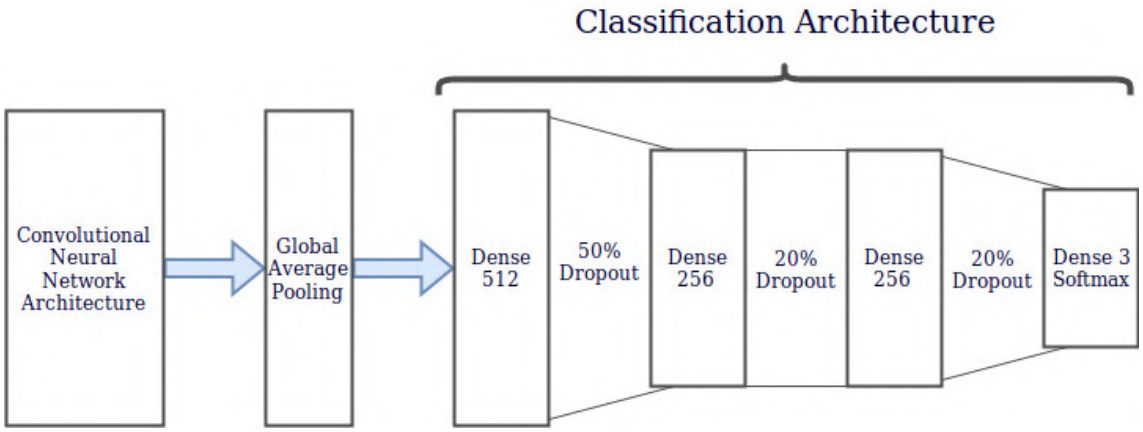

Figure 4.9: Implemented classification architecture.

As depicted in figure 4.9, dropout was applied between consecutive dense layers. Dropout is a regularisation measure, and it means that a percentage of weights will be randomly set to 0 each pass. In doing so the goal is to avoid overfitting the model to the training data. Since the amount of available data is considerably limited, overfitting is very likely and should be dealt with. The downside of using dropout is that it makes the model harder to converge. A value of 50% dropout was used between the first two fully connected layers and a 20% value used for the remaining separations. As previously stated, this arrangement showed to be the most beneficial using the Mobile Net architecture and was kept in place throughout.

A common debate within the scientific community is whether to retrain all the layers of a network or just the top ones in Transfer Learning [TSG+16] [LPX+17]. The argument lies in the fact that the feature extraction portion of the pretrained neural network already contains the knowledge to extract distinguishing features. This argument is more accurate for datasets with similar characteristics to those present in the ImageNet database, which is not the case in this work. Following this logic, it was decided to perform full training of all network architectures to ensure the best possible results.

The last aspect of training contemplated was model selection. This is done at each passing by observing performance on validation data. Given that the problem at hand is a multi-categorical classification problem, the loss function chosen was categorical cross entropy, detailed in equation 4.1. Which calculates the error or difference between the correct probabilities and the ones predicted by the *softmax* function. The training goal is to minimise this error on the training data while model selection chooses the model which resulted in the minimum value for the validation data. Early stopping was also implemented and training would stop when the validation loss

metric did not improve, by at least 0.01, after 50 iterations.

$$CategoricalCrossEntropyLoss = -\sum_{c}^{C} y_{o,c} log(p_{o,c}) \tag{4.1}$$

4.5 Summary and Conclusions

This chapter describes states the problem and the steps performed during the work. Initially, and due to non-availability of public datasets, a private dataset, using the EFS prototype, was generated, denominated the SCAS dataset. The results contain SCORE cardiovascular risk stratification for 100 people with accompanying fundus images of both their left and right eyes. Based on the output of SCORE, each person was labelled with low, moderate or high risk, turning the problem into a classification task. By observing the images it was concluded that the overall quality of the photographs was reduced. This lead to the definition of a preprocessing pipeline after manually cropping the pictures centred on the optic disc. The pipelines' purpose is to enhance the distinction between the retinal blood vessels and other eye structures and image background. This was done to validate the relationship between the vessels and CVDs. After the application of preprocessing, two distinct datasets emerged, the coloured and the grayscale datasets. Before using these datasets, commonly occurring outliers were removed. A total of 432 images were divided into training, validation and testing sets to be used in a deep learning pipeline. Mobile neural network architectures were the preferred choice given that the dataset comes from a mobile device and the secondary objective of this work was to achieve a mobile pipeline for cardiovascular risks detection from fundus images. The chosen models were trained under similar conditions and the goal was to minimise the categorical cross entropy error.

Chapter 5

Results

This chapter focuses on showing the results obtained for the prediction of cardiovascular risk using the methodology defined in the previous chapter. Section 5.1 shows the values for the loss and accuracy obtained during training, validation and testing. In section 5.2, the confusion matrices for the best performing models are shown. Afterwards, section 5.3 displays the attention maps which indicate the regions of the images from where the model is observing. Finally, section 5.4 summarises the chapter with the most important aspects.

5.1 Model Comparison

5.1.1 Training and Validation

Table 5.1 shows the metrics calculated during training for the 16 models. The naming scheme chosen for the models is Model-Weights-Dataset. Where Weights can be 'I' or 'R' for ImageNet trained weights or randomly initialised weights respectively and Dataset can be 'G' for the grayscale dataset with vessel enhancement or 'C' for the coloured cropped dataset. From now on this naming convention will be used to refer to the models. The values shown comprise the results for the training accuracy and loss and validation accuracy and loss for the epoch where validation loss was minimal during training. The epoch number is also present in the table. The model which reached higher performance for the training dataset was NASNetMobile-I-C. The InceptionV3-I-C model was able to get an 84.62% accuracy in the validation set while the MobileNet-I-C model managed to reach the lowest validation loss with 0.5437. These values hint that the combination I-C may be the top performing combination but it is necessary to keep in mind that these results can be different during the evaluation on the test set.

Figures 5.1 and 5.2 display the training and validation plots for the InceptionV3-I-C and MobileNet-I-C models respectively. The orange line shows the values for the validation set while the blue line covers the training set. In these charts, both model convergence and overfit to the training data is clear. The model is overfitting to the training data in the right-hand side of the

Model-Weights-Dataset	**Train Acc**	**Train Loss**	**Val Acc**	**Val Loss**	**Epoch**
MobileNet-I-G	0.7542	0.6006	0.6471	0.8916	38
MobileNet-R-G	0.5903	0.8852	0.6410	0.8061	51
MobileNet-I-C	0.8697	0.3423	0.7435	**0.5437**	63
MobileNet-R-C	0.5670	0.9127	0.6667	0.8576	56
MobileNetV2-I-G	0.5977	0.8769	0.6154	0.8288	20
MobileNetV2-R-G	0.4598	1.0563	0.6154	0.8716	15
MobileNetV2-I-C	0.5440	0.9819	0.6923	0.8340	9
MobileNetV2-R-C	0.6130	0.8755	0.6154	0.8472	63
NASNetMobile-I-G	0.7701	0.5735	0.5897	0.8743	32
NASNetMobile-R-G	0.5211	0.9550	0.6923	0.8437	57
NASNetMobile-I-C	**0.9349**	**0.1681**	0.8205	0.5947	65
NASNetMobile-R-C	0.6245	0.8053	0.6923	0.7582	36
InceptionV3-I-G	0.7361	0.6818	0.5897	0.8672	21
InceptionV3-R-G	0.5057	1.0168	0.6923	0.7507	**87**
InceptionV3-I-C	0.8046	0.4780	**0.8462**	0.6308	32
InceptionV3-R-C	0.6897	0.7601	0.6923	0.6923	68

Table 5.1: Performance metrics during training and validation of the models.
I - ImageNet weights; R - Random weights; G - Grayscale; C - Coloured.

plots where the distance between both values is greater. Convergence occurs when the training curves (blue line) reach a plateau. The high variance can be explained by the small size of the validation dataset (71 images in total).

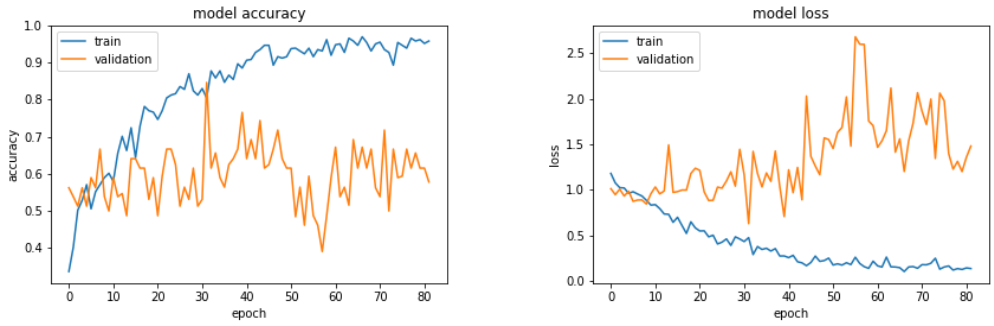

Figure 5.1: Training and Validation plots for model InceptionV3-I-C.

Based on these results, some early conclusions can be drawn.

- Models are able to generalise knowledge from the training to the validation datasets. This is indicated by the validation scores which are higher than random guessing. This fact indicates that there is a higher chance that by using a bigger and higher quality dataset the problem is solvable. Establishing a direct relationship between retinal blood vessels and cardiovascular risks.

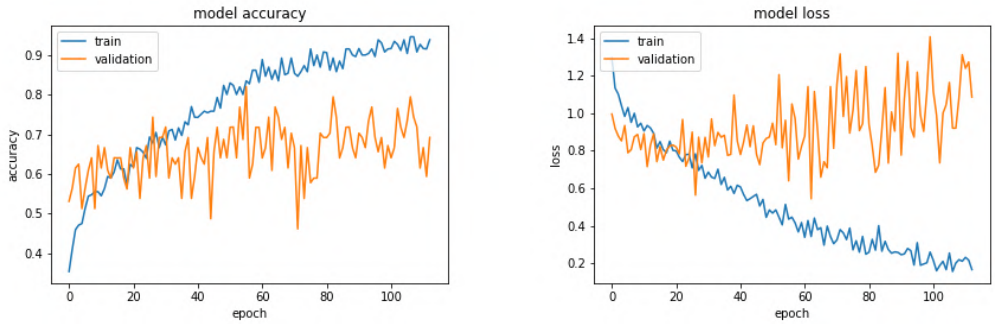

Figure 5.2: Training and Validation plots for model MobileNet-I-C.

- As expected, models starting with pretrained weights converge faster and, in general, are able to achieve better performance. From this fact, we reach the conclusion that, as indicated by the state of the art review, Transfer Learning based on the ImageNet database is possible with fundus photographs.
- A trend can be seen where models achieve better results on validation data with the coloured dataset. This makes sense given the way CNNs are implemented to extract information from the input of three RGB channels.

The conclusions discussed were made solely based on the training and validation performance. For a more realistic view on real-world performance a test set, which has never been observed by the model, is required.

5.1.2 Test

The test dataset has the objective of providing an unbiased evaluation of a model. It is the most relevant metric for assessment since the training dataset directly influences the model. The validation set also indirectly affects the final model, given that it is used to stop training. The values present in table 5.2 were computed by applying the predictive capabilities of each model to the images reserved for testing.

From table 5.2 it is possible to note that the best performing model combination was the MobileNet architecture pretrained on the ImageNet dataset and trained to classify grayscale examples with vessel enhancement. The model reached an accuracy of 58.82%.

Overall, the values obtained on the test set deviate considerably with respect to the validation dataset. As shown in the training plots, in figures 5.2 and 5.1, the variance of the validation metrics is quite high. The fact that training set accuracy increases, without improvement over the validation dataset, leads us to the conclusion that most of the models were indeed overfitted to these examples and unable to learn information for unseen images. The high values obtained for the cross-entropy loss metric indicate that the models' confidence is weak. Although the model is able to predict correctly some examples, the probability outputted for the correct class is relatively low.

Model-Weights-Dataset	**Test Loss**	**Test Acc**
MobileNet-I-G	**0.8238**	**0.5882**
MobileNetV2-I-G	1.2264	0.4559
NASNetMobile-I-G	1.5007	0.3971
InceptionV3-I-G	1.5084	0.5294
MobileNet-I-C	3.1649	0.4117
MobileNetV2-I-C	1.3071	0.2941
NASNetMobile-I-C	1.6051	0.3235
InceptionV3-I-C	1.7776	0.4118
MobileNet-R-G	2.0765	0.2647
MobileNetV2-R-G	1.2854	0.2353
NASNetMobile-R-G	1.5306	0.3824
InceptionV3-R-G	2.1677	0.2941
MobileNet-R-C	1.6007	0.2941
MobileNetV2-R-C	1.8122	0.3529
NASNetMobile-R-C	1.1252	0.5
InceptionV3-R-C	1.1997	0.5294

Table 5.2: Performance metrics for the models on the test dataset.

Regarding the use of pretrained weights, the results show that models pretrained on the ImageNet dataset were able to outperform those which were randomly initialised. The randomly initialised models, in general, were unable to produce performance higher than random guessing, which would be at around 33.33% accuracy since we are dealing with three classes. For the four neural network architectures evaluated, a trend emerged where pretrained models worked best for the grayscale dataset while models which had random initial weights performed better with the coloured dataset. Effectively, the second and third best combinations in terms of loss were the NASNetMobile-R-C and the InceptionV3-R-C respectively. Out of the 16 different combinations tested, a total of 10 managed to reach an accuracy performance higher than random guessing. Indicating a plausible relationship between fundus images and cardiovascular risk.

The main difference between the performance on the validation and test sets can be attributed due to the data split performed. The way the images are grouped per person-eye pair resulted in a correct balance between the total sizes of the three different sets but in a poor balance between the number of examples per class. Both the training and validation sets have a higher number of low risk examples, while the test set has a higher number of both moderate and high risk examples in comparison to low risk images.

5.2 Confusion Matrices

The accuracy metric of a model alone does not give a lot of useful information regarding the performance of a model across all classes. Accuracy is simply the percentage of correct predictions from the total predictions made. A confusion matrix allows for a more visual interpretation of the performance between the predicted classes. It shows which types of examples the model is able

to distinguish and which ones result in confusion. This is particularly important for imbalanced datasets, which is not the case. These matrices show the number of examples from the test set with the true label crossed with the models' predicted label. A normalised confusion matrix shows the percentage of predicted labels over the total number of examples from each class. Both types of matrices are presented for the top three performing model combinations in figures 5.3, 5.4 and 5.5.

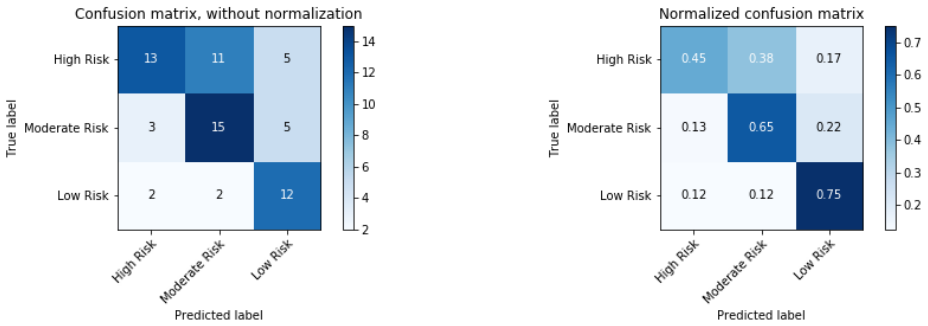

Figure 5.3: Confusion Matrix for model MobileNet-I-C.
a) non-normalised. b) normalised.

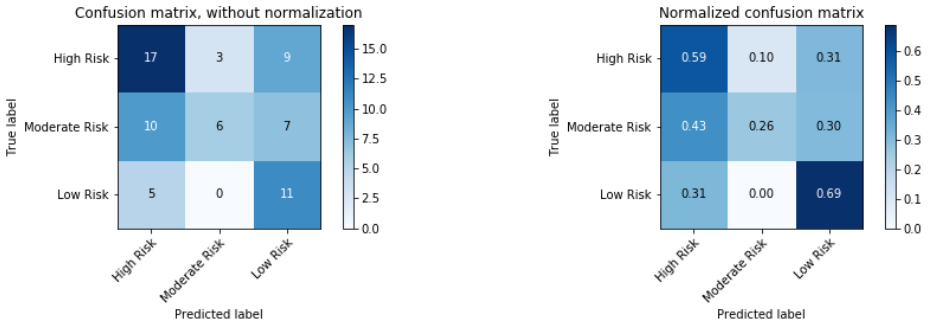

Figure 5.4: Confusion Matrix for model NASNetMobile-R-C.
a) non-normalised. b) normalised.

Figure 5.3 shows the confusion matrix obtained for the top model, MobileNet-I-C. As can be observed by the colouring scheme, this model is able to distinguish the classes low and moderate risk fairly well. Its main problem lies in the classification between the high risk and moderate risk classes where 11 high risk examples were predicted as being of moderate risk. Since knowing that a patient has some type of risk is sufficient for preventive treatment to be considered, it is interesting to see that in a binary classification of low or moderate/high risk this model was able to score a 79.41% accuracy.

Figures 5.4 and 5.5, present the matrices for the NASNetMobile-R-C and InceptionV3-R-C combinations respectively. Again these models are very good at classifying for low or moderate/high risk, with accuracy values of 69.12% and 76.47%. Contrary to MobileNet-I-C, these models tend to classify moderate risk examples with a high risk. Although the InceptionV3-R-C

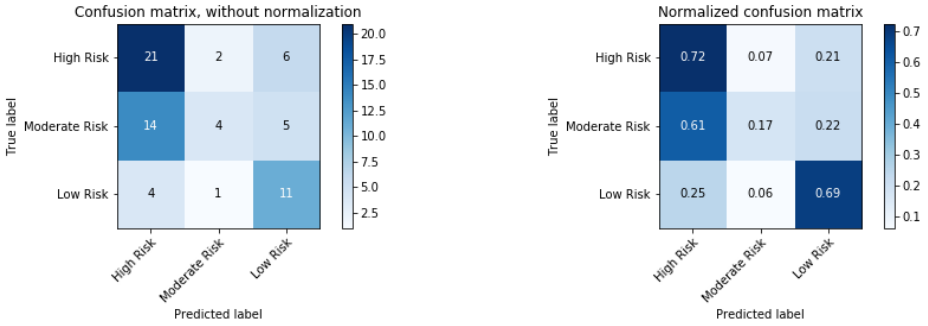

Figure 5.5: Confusion Matrix for model InceptionV3-R-C.
a) non-normalised. b) normalised.

model obtained a higher cross-entropy loss (less is better), its ability to distinguish between lower and higher risks seem to be better as shown by the binary classification accuracy of 76.47%, in comparison to the NASNetMobile-R-C model.

5.2.1 Classification Metrics

From the confusion matrix, it is possible to compute the values of precision, recall and f1-score. These metrics were designed for binary classification problems. Since in this case there are three classes, these values are calculated for each class by considering the negative class as the conjunction of the remaining two classes. Tables 5.3, 5.4 and 5.5, show the classification metrics for the three top-performing models. The tables show the precision, recall, f1-score and the total number of examples divided per class. The last line of each table presents the weighted average for each value.

	precision	**recall**	**f1-score**	**examples**
high	0.72	0.45	0.55	29
low	0.55	0.75	0.63	16
moderate	0.54	0.65	0.59	23
weighted avg	0.62	0.59	0.58	68

Table 5.3: Classification report for model MobileNet-I-C.

	precision	**recall**	**f1-score**	**examples**
high	0.53	0.59	0.56	29
low	0.41	0.69	0.51	16
moderate	0.67	0.26	0.38	23
weighted avg	0.55	0.50	0.48	68

Table 5.4: Classification report for model NASNetMobile-R-C.

	precision	recall	f1-score	examples
high	0.54	0.72	0.62	29
low	0.50	0.69	0.58	16
moderate	0.57	0.17	0.27	23
weighted avg	0.54	0.53	0.49	68

Table 5.5: Classification report for model InceptionV3-R-C.

5.3 Activation Maps

Class activation maps are a technique which helps to visualise the attention of the CNN over the inputted image. CNNs are typically a black-box and have very low interpretability, with an activation map it is possible to visualise the regions of the input which had a bigger influence over the models' prediction. To visualise attention over input in this work, the framework keras-vis [Kc17] was used. This framework uses the Grad-CAM technique proposed by Selvaraju 2016 *et al.*[SCD+16]. Grad-CAM makes use of the spatial information present in the penultimate convolutional layers output. According to the authors, this information gets completely lost in the final dense layers of a model (the classification layers) and their technique provides a more faithful visual explanation for model decisions, as stated by the authors. Figures 5.6, 5.7 and 5.8 show six, two of each class, correctly predicted, examples for the top three models.

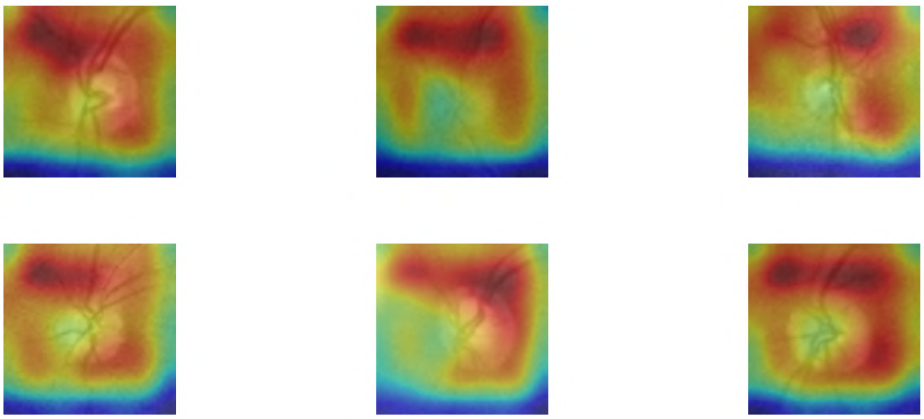

Figure 5.6: Activation maps example for MobileNet-I-G.
From left to right, top to bottom: 2 examples of each class (low, moderate and high)

The attention maps for the grayscale images, used for the MobileNet-I-G model, present in figure 5.6, show a larger focus in the areas surrounding the optic disc, where retinal blood vessels converge. The attention is fairly spread out to where blood vessels are visible. This means that the vessel enhancement preprocessing pipeline has ensured that vessels are the most prominent eye structure. Ultimately translating into better results, which establish a relationship between retinal vessels and cardiovascular risk. From these results, it is difficult to pinpoint the exact features of

the vessels that influence prognostication the most, be it tortuosity, ramification or vessel diameter.

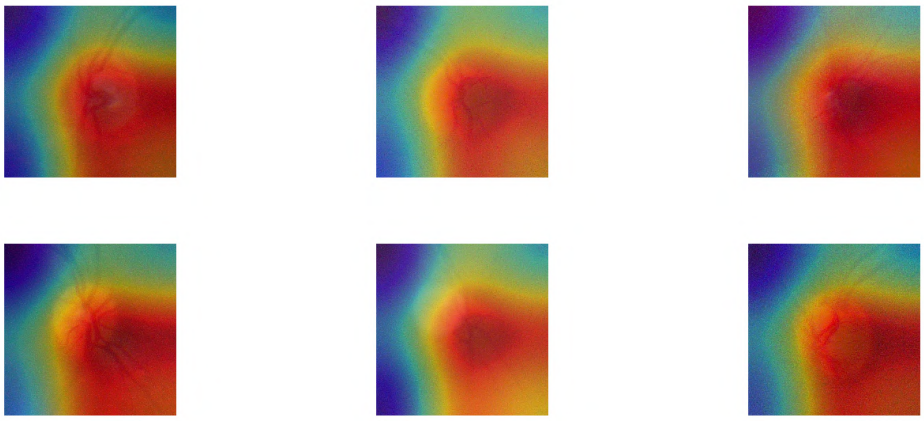

Figure 5.7: Activation maps example for InceptionV3-R-C.
From left to right, top to bottom: 2 examples of each class (low, moderate and high)

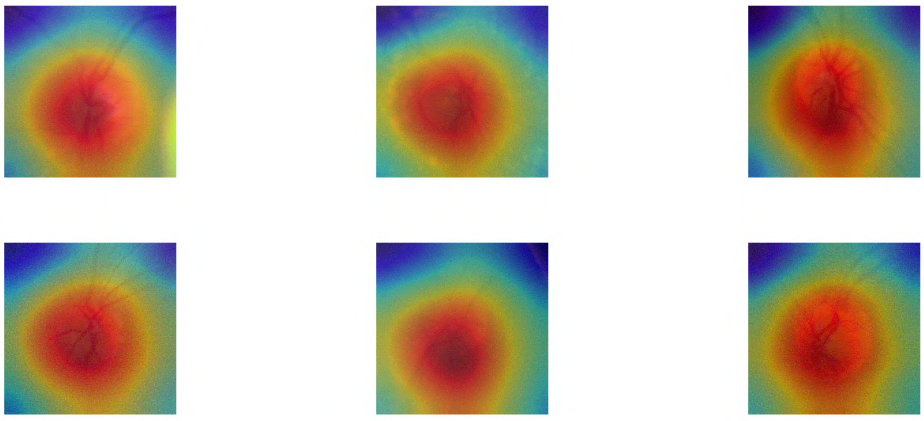

Figure 5.8: Activation maps example for NASNetMobile-R-C.
From left to right, top to bottom: 2 examples of each class (low, moderate and high)

For the coloured dataset, shown in figures 5.8 and 5.7, it is possible to observe a clear difference from the grayscale examples. Models trained on the coloured dataset focus mainly on the optic disc, with little to no attention paid to the more peripheral retinal blood vessels present in the image. In the particular case of InceptionV3-R-C, the model tends to pay more attention to the lower right-hand side of the image. The similarity across the six images for each model indicates that these model do not focus on any particular structure, but rather a general area around the centre of the picture.

5.4 Summary

This chapter exposed the intermediate and final results for the models that were described in the preceding chapter. The best model was the MobileNet-I-G, the MobileNet architecture pretrained on the ImageNet dataset with the vessel enhanced grayscale dataset. The model managed to predict accurately 58.82% of the examples in the portion of available data reserved for testing with a 0.8238 cross-entropy loss. Model performance, in general, deviated from the validation to the test sets, leading to the conclusion of overfitting to the training and validation examples. The usage of pretrained weights on the ImageNet dataset resulted in faster convergence and better results. For the preprocessed dataset, the ImageNet weights initialisation proved more beneficial while for the coloured dataset random initialisation was capable of achieving higher performance. The second and third best models were NASNetMobile-R-C and InceptionV3-R-C respectively.

The confusion matrices revealed that the top three models were able to distinguish the low-risk class quite properly from the remaining two. All three of the top performing models had difficulties in differentiating the moderate and high-risk classes. This might be due to the slight class imbalance originated from the train-validation-test split performed. Nonetheless, the models were able to detect the presence of risk (moderate or high) with up to 79.41% accuracy.

The activation maps displayed provided insight into the inner workings of the models. These images revealed that for the grayscale dataset, which obtained the highest results, the models focus on the blood vessels morphology surrounding the optic disc. For the coloured dataset, model attention was centred in the optic disc.

When taking into consideration the reduced size and quality of the available dataset, the results look very encouraging. The relationship between retinal blood vessels and cardiovascular risk is apparent.

Chapter 6

Conclusions

This chapter summarises the work performed. Section 6.1 opens the chapter with the most relevant conclusions achieved. Section 6.2 gives an overview of the limiting factors that influenced the final outcome. In section 6.3 the main contributions to the scientific community are presented. Finally, section 6.4 describes possible enhancements and what can still be done regarding the relationship between cardiovascular risks and fundus photography.

6.1 Final Conclusions

In this work, we have shown how Machine Learning can be used to analyse images of the retinal vascular morphology. How cardiovascular risks are currently calculated in modern medicine and their corresponding limitations. Most importantly, how we can apply deep learning to fundus images in order to predict the risk of fatal cardiovascular disease. Lastly, the problem of cardiovascular risks detection was defined and a possible solution was investigated.

A literature review uncovered new insight regarding commonly used methodologies and obtained results. Leveraging the power of deep learning is of great significance for the early detection of cardiovascular risks which in turn can lead to early treatment and prevention.

During the implementation of the proposed solution, a private dataset was built. In this dataset, fundus photographs were taken, using the EyeFundusScope prototype, alongside the stratification of cardiovascular risk, employing the SCORE risk calculator. The dataset produced a multi-categorical image classification problem with three distinct classes, low, moderate and high risk.

Following the application of a preprocessing pipeline, consisting of retinal blood vessel enhancement and manual cropping centred on the optic disc, deep neural network architectures were trained to classify the resulting 432 images. Training considered the use of pretrained and random weights and the application of the aforementioned preprocessing scheme.

A comparison was done between mobile friendly deep learning architectures so that a pipeline able to run on an Android application could be created. The MobileNet architecture managed

to produce the best results, with a classification accuracy of 58.82%, using a Transfer Learning approach with values pretrained on the ImageNet dataset and the vessel enhanced images. The model was able to distinguish the presence of risk (moderate or high risk) with a 79.41% accuracy, which is relevant for the employment of preventive treatment.

The study of the confusion matrices showed that the top three performing models were able to differentiate well low-risk images from the remaining examples but had a harder time classifying the moderate and high-risk photographs. From the attention maps, it was possible to conclude that the final outcome was mainly influenced by the retinal blood vessels surrounding the optic disc. These observations demonstrate an existing relationship between cardiovascular risk and fundus photographs, in particular, retinal vessels, which was one of the principal aims of this work.

6.2 Limitations

The obtained results were conditioned by the reduced amount of available examples, which limited the architectures' learning capabilities. The difference in image quality from retinographs and the EFS prototype, which is still in development at Fraunhofer, made it more difficult to segment retinal structures. This meant that manual cropping had to be done in order to reduce image errors, such as reflections and blurs. The manual cropping step impeded the creation of an automatic pipeline for the acquisition of fundus images and real-time prediction of cardiovascular risks.

In general, the fundus images obtained from a younger demographic were better in terms of definition and, most importantly, vessel contrast. Given that younger people tend to have healthier cardiovascular risk factors, and in turn, reducing the risk of cardiovascular disease, it is possible that the final outcomes may have been influenced by image contrast. To confirm this, a bigger dataset and further study are required.

A clear limitation comes from the fact that the labels used were obtained with the SCORE calculator. This limits the results of only being able to be as good as this calculator, which is reported to have a 0.72 AUC [PVB+18].

6.3 Contributions

This work contributes to the scientific study of retinal fundus photographs, in the area of medical image analysis. The dataset obtained, although private, is not only useful for this work, but also for studies related to the EyeFundusScope project since eye diseases were also annotated. The dataset created comprises information for multiple risk factors, which were not considered in this project, such as pulse wave measures and body mass index. Overall, these annotations may be of great value for future studies.

The application of ImageNet based Transfer Learning proved to be beneficial for the classification of retinal fundus images. The training of Deep Learning architectures requires substantial amounts of properly annotated data, a common problem within the field of medical image analysis. This problem is minimised by the employment of Transfer Learning.

Enhancing the contrast between the retinal background and the blood vessels produced the best results. Making it possible to denote a direct relationship between these structures and cardiovascular risk.

Ultimately, the work performed opens the possibility of defining a mobile cardiovascular risk detection tool based on fundus images since the best model was of low complexity and is able to be integrated into an Android application. This tool is able to bypass the needed measurement for systolic blood pressure and cholesterol levels in traditional cardiology techniques.

6.4 Future Work

This study has the possibility of further improvement in different ways. First and foremost, a bigger dataset would help improve significantly the obtained results. Based on the secondary objective of mobile detection, solely mobile architectures should be considered. The application of vessel enhancement to coloured images should also be considered, given the affinity of Convolutional Neural Networks to coloured input. A binary classifier which simply predicts whether a person has a high risk or not is also possible, and according to the results should work better with the defined dataset.

A possible improvement would be to add readily available factors, like age, gender and smoking status, to the classification pipeline. Considering the influence of these factors in SCORE, better accuracy is ought to be achieved. Since they are readily available it maintains the prospect of defining a mobile pipeline capable of bringing cardiovascular risk stratification to underdeveloped populations.

Ideally, a dataset based on yearly follow up of patients should be considered. This dataset would comprise several follow-up fundus photographs and the presence of major adverse cardiovascular events, making the prediction more reliable and accurate.

References

[AAJMC07] S.R. Alty, N. Angarita-Jaimes, S.C. Millasseau, and P.J. Chowienczyk. Predicting Arterial Stiffness From the Digital Volume Pulse Waveform. *IEEE Transactions on Biomedical Engineering*, 54(12):2268–2275, dec 2007.

[ADHS+08] Bashir Al-Diri, Andrew Hunter, David Steel, Maged Habib, Taghread Hudaib, and Simon Berry. REVIEW - A reference data set for retinal vessel profiles. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 2008, pages 2262–2265. IEEE, aug 2008.

[AK82] Koichiro Akita and Hideki Kuga. A computer method of understanding ocular fundus images. *Pattern Recognition*, 15(6):431–443, jan 1982.

[BG17] A. M. R. R. Bandara and P. W. G. R. M. P. B. Giragama. A retinal image enhancement technique for blood vessel segmentation algorithm. In *2017 IEEE International Conference on Industrial and Information Systems (ICIIS)*, pages 1–5. IEEE, dec 2017.

[Bri] Brilliant.org. Convolutional neural network.

[CMS12] D. Ciresan, U. Meier, and J. Schmidhuber. Multi-column deep neural networks for image classification. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3642–3649. IEEE, jun 2012.

[Con03] R Conroy. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *European Heart Journal*, 24(11):987–1003, jun 2003.

[CYS+17] Joon Yul Choi, Tae Keun Yoo, Jeong Gi Seo, Jiyong Kwak, Terry Taewoong Um, and Tyler Hyungtaek Rim. Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database. *PLOS ONE*, 12(11):e0187336, nov 2017.

[Dap] Daphne Cornelisse. An intuitive guide to Convolutional Neural Networks.

[Eur] Eurostat. Cardiovascular diseases statistics - Statistics Explained.

[FCX+18a] Huazhu Fu, Jun Cheng, Yanwu Xu, Damon Wing Kee Wong, Jiang Liu, and Xiaochun Cao. Joint Optic Disc and Cup Segmentation Based on Multi-label Deep Network and Polar Transformation. *CoRR*, jan 2018.

[FCX+18b] Huazhu Fu, Jun Cheng, Yanwu Xu, Changqing Zhang, Damon Wing Kee Wong, Jiang Liu, and Xiaochun Cao. Disc-aware Ensemble Network for Glaucoma Screening from Fundus Image. *CoRR*, may 2018.

[FESG16] Karma M. Fathalla, Aniko Ekart, Swathi Seshadri, and Doina Gherghel. Cardiovascular risk prediction based on Retinal Vessel Analysis using machine learning. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 000880–000885. IEEE, oct 2016.

[FRH+12] M.M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, and S.A. Barman. Blood vessel segmentation methodologies in retinal images – A survey. *Computer Methods and Programs in Biomedicine*, 108(1):407–433, oct 2012.

[GC17] Fantin Girard and Farida Cheriet. Artery/vein classification in fundus images using CNN and likelihood score propagation. In *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 720–724. IEEE, nov 2017.

[GMB+18] Felix Grassmann, Judith Mengelkamp, Caroline Brandl, Sebastian Harsch, Martina E. Zimmermann, Birgit Linkohr, Annette Peters, Iris M. Heid, Christoph Palm, and Bernhard H.F. Weber. A Deep Learning Algorithm for Prediction of Age-Related Eye Disease Study Severity Scale for Age-Related Macular Degeneration from Color Fundus Photography. *Ophthalmology*, 125(9):1410–1420, sep 2018.

[GPC+16] Varun Gulshan, Lily Peng, Marc Coram, Martin C. Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan, Kasumi Widner, Tom Madams, Jorge Cuadros, Ramasamy Kim, Rajiv Raman, Philip C. Nelson, Jessica L. Mega, and Dale R. Webster. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*, 316(22):2402, dec 2016.

[GSL+09] María García, Clara I. Sánchez, María I. López, Daniel Abásolo, and Roberto Hornero. Neural network based detection of hard exudates in retinal images. *Computer Methods and Programs in Biomedicine*, 93(1):9–19, jan 2009.

[GVAF+19] Juan J. Gómez-Valverde, Alfonso Antón, Gianluca Fatti, Bart Liefers, Alejandra Herranz, Andrés Santos, Clara I. Sánchez, and María J. Ledesma-Carbayo. Automatic glaucoma classification using color fundus images based on convolutional neural networks and transfer learning. *Biomedical Optics Express*, 10(2):892, feb 2019.

[HDK19] D. Jude Hemanth, Omer Deperlioglu, and Utku Kose. An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network. *Neural Computing and Applications*, jan 2019.

[HZC+17] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *CoRR*, apr 2017.

[IdJV+04] M. Kamran Ikram, Frank Jan de Jong, Johannes R. Vingerling, Jacqueline C. M. Witteman, Albert Hofman, Monique M. B. Breteler, and Paulus T. V. M. de Jong. Are Retinal Arteriolar or Venular Diameters Associated with Markers for Cardiovascular Disorders? The Rotterdam Study. *Investigative Opthalmology & Visual Science*, 45(7):2129, jul 2004.

REFERENCES

[JSM19] R. S. Jeena, A. Sukeshkumar, and K. Mahadevan. Retina as a Biomarker of Stroke. In *Computer Aided Intervention and Diagnostics in Clinical and Medical Images*, pages 219–226. Springer, Cham, 2019.

[KB14] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations*, dec 2014.

[KBK+13] Thomas Kohler, Attila Budai, Martin F. Kraus, Jan Odstrcilik, Georg Michelson, and Joachim Hornegger. Automatic no-reference quality assessment for retinal fundus images using vessel segmentation. In *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, pages 95–100. IEEE, jun 2013.

[Kc17] Raghavendra Kotikalapudi and contributors. keras-vis. https://github.com/raghakot/keras-vis, 2017.

[KKMW07] Shweta Kaushik, Annette Kifley, Paul Mitchell, and Jie Jin Wang. Age, Blood Pressure, and Retinal Vessel Diameter: Separate Effects and Interaction of Blood Pressure and Age. *Investigative Opthalmology & Visual Science*, 48(2):557, feb 2007.

[LADC18] Gabriel Lepetit-Aimon, Renaud Duval, and Farida Cheriet. Large Receptive Field Fully Convolutional Network for Semantic Segmentation of Retinal Vasculature in Fundus Images. In *Computational Pathology and Ophthalmic Medical Image Analysis*, pages 201–209. Springer International Publishing, 2018.

[LKB+17] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, dec 2017.

[LOZL17] H. Leopold, J. Orchard, J. Zelek, and V. Lakshminarayanan. Segmentation and feature extraction of retinal vascular morphology. *Medical Imaging 2017: Image Processing*, 10133:UNSP 101330V, 2017. PT: S; CT: Conference on Medical Imaging - Image Processing; CY: FEB 12-14, 2017; CL: Orlando, FL; SP: SPIE, Alpin Med Syst; UT: WOS:000405564600028.

[LPX+17] Xiaogang Li, Tiantian Pang, Biao Xiong, Weixiang Liu, Ping Liang, and Tianfu Wang. Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification. In *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pages 1–11. IEEE, oct 2017.

[LRSB16] A. Lahiri, Abhijit Guha Roy, Debdoot Sheet, and Prabir Kumar Biswas. Deep neural ensemble for retinal vessel segmentation in fundus images towards achieving label-free angiography. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, volume 2016, pages 1340–1343. IEEE, aug 2016.

[MLM+09] K. McGeechan, G. Liew, P. Macaskill, L. Irwig, R. Klein, B. E. K. Klein, J. J. Wang, P. Mitchell, J. R. Vingerling, P. T. V. M. de Jong, J. C. M. Witteman, M. M. B. Breteler, J. Shaw, P. Zimmet, and T. Y. Wong. Prediction of Incident Stroke Events Based on Retinal Vessel Caliber: A Systematic Review and Individual-Participant Meta-Analysis. *American Journal of Epidemiology*, 170(11):1323–1332, dec 2009.

[MSG+15] Debapriya Maji, Anirban Santara, Sambuddha Ghosh, Debdoot Sheet, and Pabitra Mitra. Deep neural network and random forest hybrid architecture for learning to detect retinal vessels in fundus images. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3029–3032. IEEE, aug 2015.

[OPdFB18] José Ignacio Orlando, Elena Prokofyeva, Mariana del Fresno, and Matthew B. Blaschko. An ensemble deep learning based approach for red lesion detection in fundus images. *Computer Methods and Programs in Biomedicine*, 153:115–127, jan 2018.

[PAM+06] Niall Patton, Tariq M. Aslam, Thomas MacGillivray, Ian J. Deary, Baljean Dhillon, Robert H. Eikelboom, Kanagasingam Yogesan, and Ian J. Constable. Retinal image analysis: Concepts, applications and potential. *Progress in Retinal and Eye Research*, 25(1):99–127, jan 2006.

[PVB+18] Ryan Poplin, Avinash V. Varadarajan, Katy Blumer, Yun Liu, Michael V. McConnell, Greg S. Corrado, Lily Peng, and Dale R. Webster. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nature Biomedical Engineering*, 2(3):158–164, mar 2018.

[Rah18] Ehsan Rahimy. Deep learning applications in ophthalmology. *Current Opinion in Ophthalmology*, 29(3):254–260, may 2018.

[RCS+16] P. K. Roy, R. Chakravorty, S. Sedai, D. Mahapatra, and R. Garnavi. Automatic eye type detection in retinal fundus image using fusion of transfer learning and anatomical features. In *2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pages 1–7, Nov 2016.

[RK17] Jeena R S and Sukesh Kumar A. Stroke Prognosis through Retinal Image Analysis. *Advances in Image and Video Processing*, 5(2):13, apr 2017.

[RKK18] Sashank J. Reddi, Satyen Kale, and Sanjiv Kumar. On the convergence of adam and beyond. In *International Conference on Learning Representations*, 2018.

[SAN+04] J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, and B. van Ginneken. Ridge-Based Vessel Segmentation in Color Images of the Retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, apr 2004.

[SCD+16] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *CoRR*, oct 2016.

[Sch15] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, jan 2015.

[SHZ+18] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. MobileNetV2: Inverted Residuals and Linear Bottlenecks. *CoRR*, jan 2018.

[SKH+14] Sharmini Selvarajah, Gurpreet Kaur, Jamaiyah Haniff, Kee Chee Cheong, Tee Guat Hiong, Yolanda van der Graaf, and Michiel L. Bots. Comparison of the Framingham Risk Score, SCORE and WHO/ISH cardiovascular risk prediction models in an Asian population. *International Journal of Cardiology*, 176(1):211–218, sep 2014.

REFERENCES

[ST02] Patrick J. Saine and Marshall E. Tyler. *Ophthalmic photography : retinal photography, angiography, and electronic imaging*. Butterworth-Heinemann, 2002.

[SVI+15] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the Inception Architecture for Computer Vision. *CoRR*, dec 2015.

[TSG+16] Nima Tajbakhsh, Jae Y. Shin, Suryakanth R. Gurudu, R. Todd Hurst, Christopher B. Kendall, Michael B. Gotway, and Jianming Liang. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Transactions on Medical Imaging*, 35(5):1299–1312, may 2016.

[VPB+18] Avinash V. Varadarajan, Ryan Poplin, Katy Blumer, Christof Angermueller, Joe Ledsam, Reena Chopra, Pearse A. Keane, Greg S. Corrado, Lily Peng, and Dale R. Webster. Deep Learning for Predicting Refractive Error From Retinal Fundus Images. *Investigative Opthalmology & Visual Science*, 59(7):2861, jun 2018.

[WDL+98] Peter W. F. Wilson, Ralph B. D'Agostino, Daniel Levy, Albert M. Belanger, Halit Silbershatz, and William B. Kannel. Prediction of Coronary Heart Disease Using Risk Factor Categories. *Circulation*, 97(18):1837–1847, may 1998.

[WH10] Ruth Webster and Emma Heeley. Perceptions of risk: understanding cardiovascular disease. *Risk management and healthcare policy*, 3:49–60, 2010.

[WKK+01] Tien Yin Wong, Ronald Klein, Barbara E.K Klein, James M Tielsch, Larry Hubbard, and F.Javier Nieto. Retinal Microvascular Abnormalities and their Relationship with Hypertension, Cardiovascular Disease, and Mortality. *Survey of Ophthalmology*, 46(1):59–80, jul 2001.

[WKK+03] Tien Yin Wong, Ronald Klein, Barbara E. K. Klein, Stacy M. Meuer, and Larry D. Hubbard. Retinal Vessel Diameters and Their Associations with Age and Blood Pressure. *Investigative Opthalmology & Visual Science*, 44(11):4644, nov 2003.

[Wor18] World Health Organization. The top 10 causes of death, 2018.

[WWH+06] Nicholas Witt, Tien Y. Wong, Alun D. Hughes, Nish Chaturvedi, Barbara E. Klein, Richard Evans, Mary McNamara, Simon A. McG Thom, and Ronald Klein. Abnormalities of Retinal Microvascular Structure and Risk of Mortality From Ischemic Heart Disease and Stroke. *Hypertension*, 47(5):975–981, may 2006.

[ZVSL17] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V. Le. Learning Transferable Architectures for Scalable Image Recognition. *CoRR*, jul 2017.

REFERENCES

Appendix A

Data Collection Informed Consent

CONSENTIMENTO INFORMADO, LIVRE E ESCLARECIDO PARA PARTICIPAÇÃO EM INVESTIGAÇÃO

de acordo com a Declaração de Helsínquia[1] e a Convenção de Oviedo[2]

Por favor, leia com atenção a seguinte informação. Se achar que algo está incorreto ou que não está claro, não hesite em solicitar mais informações. Se concorda com a proposta que lhe foi feita, queira assinar este documento.

Título do estudo: Avaliação de um sistema móvel de aquisição de imagens da retina e obtenção de dados para suporte a estudos de investigação exploratórios acerca da previsão do risco cardiovascular ou doenças oculares.

Enquadramento: Este estudo de investigação está a ser promovido pela Associação Fraunhofer Portugal Research – Fraunhofer Portugal Research Center for Assistive Information and Communication Solutions (Fraunhofer Portugal AICOS) no âmbito do projeto EyeFundusScope que tem como coordenador científico e investigador principal o Doutor Filipe Soares. Este estudo contribuirá também para a recolha de dados a serem utilizados em dissertações de Mestrado em Engenharia Informática da Faculdade de Engenharia da Universidade do Porto associadas à Fraunhofer Portugal AICOS.

Explicação do estudo: O estudo de investigação acima mencionado tem como finalidade avaliar um sistema móvel para aquisição de imagens da retina em termos de qualidade de imagem e adequabilidade ao fim de diagnóstico de doenças oculares. Um outro objetivo do estudo é obter dados que suportem estudos exploratórios, que irão investigar a relação entre as imagens da retina e o risco cardiovascular. A sua participação neste estudo consiste na recolha de imagens da retina usando um protótipo desenvolvido pelo centro de investigação Fraunhofer Portugal AICOS. A recolha das imagens da retina tem uma duração de cerca de 5 minutos para os dois olhos, é indolor e não representa qualquer risco ou desconforto. As imagens serão gravadas, de forma anonimizada, para posterior análise; o sistema não fornece no imediato qualquer resultado de análise das imagens e estes resultados não serão lhe comunicados. Pedimos também que responda a um questionário/entrevista, que terá uma duração prevista de 5 minutos, com o objetivo de recolher dados demográficos, clínicos e laboratoriais relacionados com o risco cardiovascular.

Condições e financiamento: Não serão necessárias deslocações adicionais no âmbito deste estudo. Não haverá lugar a qualquer pagamento pela sua participação. Este estudo é financiado pelos programas Norte 2020, Portugal 2020 e pela União Europeia. A sua participação neste estudo é voluntária, podendo em qualquer altura cessá-la sem qualquer tipo de consequência.

Confidencialidade e anonimato: As informações recolhidas são confidenciais e a sua identificação não será revelada. Os dados recolhidos serão utilizados exclusivamente no âmbito dos objetivos do presente estudo. Os dados serão recolhidos e analisados pelos investigadores identificados abaixo ou por outro elemento com autorização delegada pelo investigador principal e serão apresentados em publicações científicas e conferências ou outro tipo de evento científico ou de divulgação do projeto, assim como nos diversos meios de comunicação social. A sua identidade não será revelada em qualquer situação. Os dados pessoais, assim como as respostas aos questionários/entrevistas, serão mantidos numa base de dados, de forma anonimizada.

Gostaríamos de contar com a sua participação. Agradecemos muito o seu contributo, fundamental para a nossa investigação!

Filipe Soares

Investigador Sénior na Fraunhofer Portugal AICOS

Sandra Ferreira

Presidente da Associação de estudantes da Escola Superior de Saúde de Coimbra

---

1 http://portal.arsnorte.min-saude.pt/portal/page/portal/ARSNorte/Comiss%C3%A3o%20de%20%C3%89tica/Ficheiros/Declaracao_Helsinquia_2008.pdf

2 http://dre.pt/pdf1sdip/2001/01/002A00/00140036.pdf

Coordenador científico e Investigador Principal:

Nome: Filipe Soares

Instituição: Associação Fraunhofer Portugal Research – Fraunhofer Portugal Research Center for Assistive Information and Communication Solutions

E-mail: filipe.soares@fraunhofer.pt

Assinatura: Filipe Cruz Gomes Soares

-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-o-

Declaro ter lido e compreendido este documento, bem como as informações verbais que me foram fornecidas pela/s pessoa/s que acima assina/m. Foi-me garantida a possibilidade de, em qualquer altura, recusar participar neste estudo sem qualquer tipo de consequências. Desta forma, aceito participar neste estudo e permito a utilização dos dados que de forma voluntária forneço, confiando em que apenas serão utilizados para esta investigação e nas garantias de confidencialidade e anonimato que me são dadas pelo/a investigador/a.

Nome: _____

Assinatura: _____

Data: ____ /_____ /_____

SE NÃO FOR O PRÓPRIO A ASSINAR POR IDADE OU INCAPACIDADE

(se o menor tiver discernimento deve também assinar em cima, se consentir)

NOME: _____

BI/CD Nº: _____ DATA OU VALIDADE _____ /_____ /_____

GRAU DE PARENTESCO OU TIPO DE REPRESENTAÇÃO: _____

ASSINATURA: _____

ESTE DOCUMENTO É COMPOSTO DE 2 PÁGINAS E FEITO EM DUPLICADO:

UMA VIA PARA O/A INVESTIGADOR/A, OUTRA PARA A PESSOA QUE CONSENTE

Appendix B

Data Collection Questionnaire

Questionário | Data: __/__/__

Identificação: |SCAS __ __ __|

Género: Masculino ☐ Feminino ☐

Idade:_____ anos

Etnicidade: _____

Peso:____Kg Altura:____m

Massa Magra: _____ % Massa Gorda: _____ %

Perímetro da cintura:____cm Perímetro da anca:____cm

Colesterol Total: _____ mg/dL

Colesterol HDL: _____ mg/dL

Fumador? _____ (Cigarros/dia: ____)

Consome bebidas alcoólicas?____ frequência?_____

HTA (S/N): _____ Medicação (S/N)? _____

Diabetes (S/N): _____Medicação (S/N)? _____

Dislipidemia (S/N): _____ Medicação (S/N)? _____

Antecedentes Cardiovasculares (S/N): ____Qual?_____

Antecedentes cardiovasculares na família (S/N):_____

Problemas de visão?_____ Qual? _____

Heart SCORE : ____ Classificação do risco: _____