UNISINOS

Programa de Pós-Graduação em

Computação Aplicada

Mestrado/Doutorado Acadêmico

Samuel Armbrust Freitas

DeepCADD: a Deep Neural Network for Automatic Detection of Coronary Artery Disease

São Leopoldo, 2022

Samuel Armbrust Freitas

DEEPCADD: A DEEP NEURAL NETWORK FOR AUTOMATIC DETECTION OF CORONARY ARTERY DISEASE

Master Thesis proposal presented as a partial requirement to obtain the Master's degree by the Applied Computing Graduate Program of the Universidade do Vale do Rio dos Sinos — UNISINOS

Orientador:
Prof. Dr. Gabriel de Oliveira Ramos

Coorientador:
Prof. Dr. Cristiano André da Costa

São Leopoldo
2021

F866d Freitas, Samuel Armbrust.
DeepCADD : a Depp neural network for automatic detection of coronary artery disease / Samuel Armbrust Freitas. – 2021.
108 f. : il. ; 30 cm.

Dissertação (mestrado) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Computação Aplicada, 2021.
"Orientador: Prof. Dr. Gabriel de Oliveira Ramos
Coorientador: Prof. Dr. Cristiano André da costa."

1. Coronary artery. 2. Coronary artery disease. 3. Deep learning. 4. Object detection. I. Título.

CDU 004.7

Dados Internacionais de Catalogação na Publicação (CIP)
(Bibliotecária: Silvana Dornelles Studzinski – CRB 10/2524)

ATA DE BANCA EXAMINADORA DE DISSERTAÇÃO DE MESTRADO Nº 1/2022

Aluno: Samuel Armbrust Freitas

Título da Dissertação: "DEEPCADD: A DEEP NEURAL NETWORK FOR AUTOMATIC DETECTION OF CORONARY ARTERY DISEASE"

Banca: Prof. Dr. Gabriel de Oliveira Ramos – Orientador
Prof. Dr. Cristiano André da Costa - Coorientador
Prof. Dra. Mariana Recamonde Mendoza - UFRGS
Prof. Dr. Rafael Kunst - UNISINOS
Prof. Dr. Sandro José Rigo - UNISINOS

Aos dezoito dias do mês de fevereiro do ano de 2022, às 10h, a Comissão Examinadora de Defesa de Dissertação composta pelos professores: Prof. Dr. Gabriel de Oliveira Ramos, Orientador – UNISINOS (participação por webconferência); Prof. Dr. Cristiano André da Costa – Coorientador – UNISINOS (participação por webconferência); Prof. Dra. Mariana Recamonde Mendoza, Avaliadora – Universidade Federal do Rio Grande do Sul, UFRGS (participação por webconferência); Prof. Dr. Rafael Kunst, Membro da Banca – UNISINOS (participação por webconferência),; e Prof. Dr. Sandro José Rigo, Membro da Banca – UNISINOS (participação por webconferência), para analisar e avaliar a Dissertação apresentada pelo Samuel Armbrust Freitas (participação por webconferência).

Considerações da Banca:
Após deliberação, os membros da banca avaliadora consideraram que o trabalho atende aos critérios esperados para uma dissertação de mestrado, além de apresentar originalidade científica na área de pesquisa. A banca sugeriu diversas melhorias para o trabalho no que se refere aos seguintes itens: organização/correção do texto, fundamentação das escolhas de projeto, definição da metodologia de avaliação, e análise e discussão dos resultados.

Ocorreu alteração do título? (X) Não ( ) Sim
Indicar o novo título: .....
.....
A Banca Examinadora, em cumprimento ao requisito exigido para a obtenção do Título de Mestre em Computação Aplicada, julga esta dissertação:
(X) APROVADA ( ) REPROVADA

Conforme Artigo 67 do Regimento do Programa o texto definitivo, com aprovação do Orientador, deverá ser entregue no prazo máximo de sessenta (60) dias após a defesa. O resultado da banca é de consenso entre os avaliadores. A emissão do Diploma está condicionada a entrega da versão final da

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

To my parents.

It's not a silly question if you can't answer it.
— JOSTEIN GAARDER, SOPHIE'S WORLD

ACKNOWLEDGEMENTS

This work was only possible due to the collaboration of several people in different aspects of the research. I would also like to thank my advisor Prof. Gabriel de Oliveira Ramos who accepted me under his supervision as a master's student, sharing his research excellence, time, and knowledge conducting this work. Gabriel trusted on me from our very first interaction and I will not forget his effort to make this study happen. I would also like to thank my co-advisor Prof. Cristiano André da Costa helping me in this research always bringing amazing feedback. I would like to thank Gabriel and Cristiano for several reviews, meetings, and for being references of excellence in research for me. I would also like to thank Felipe Zeiser for the amazing support, both in the idealization and sharing his brilliant technical skills with neural networks.

Special thank you to my family, for the emotional support throughout my life. To my father, Aldo, and my mother, Sônia, for being a synonym of respect, love, and power of will. Your determination inspired me in each and every decision I took. To my sisters, Rebeca and Brenda, for trusting in me, making every moment a moment of fun, and accepting my mistakes as part of my development.

I would like to thank my wife, Dra. Débora Nienow for sharing her amazing talent in medicine and generating the sparkle I needed to decide my final paper study in 2019, which catalyzed this master's degree proposal. Your partnership, enthusiasm, and excellence inspired me from the very first moment of this master's program so I was able to deliver the best of me.

I would like to thank Dr. Paulo Renato Mercio Machado and Sinara Leotte Garcia for opening the doors of Cardiosinos-NH to me and the partnership which resulted in the great dataset I consumed in this research. I would like to thank MsC. Graciela Racolte for her attention to my study and connecting me with Cardiosinos, creating the amazing opportunity to validate my study.

I would like to thank my colleagues from SAP for their support and trust in me during this period being able to deal with the master's degree and the daily job with excellence. Your motivation, attendance in the master's degree evaluation steps, and vibration with each small victory empowered me. A special appreciation to Guoyang Weng, which helped me in one of the most challenging steps of my literature review.

I would like to thank the excellent professors of Unisinos PPGCA (Postgraduate Program of Applied Computing) and the secretary staff for their patience during this challenging period. I am very grateful for this program's conduction, uncountable suggestions of improvement, and incentives for publishing papers. When I look back, I am impressed how every master's degree step is connected and leads to the amazing research I am delivering.

Finally, this research would not be possible without the financial support of CNPq and CAPES for my master's scholarship.

ABSTRACT

CONTEXT: Cardiovascular diseases represent the number one cause of death globally, which include the most common disorders in the heart's health, namely coronary artery disease (CAD). CAD is mainly caused by fat accumulated in the arteries' internal walls, creating an atherosclerotic plaque that impacts the functional behavior of the blood flow. Anatomical plaque characteristics are essential for a complete functional assessment of CAD. In fact, there is no unique method to assess all the coronary artery segments with high accuracy. **OBJECTIVE:** Such a panorama evidences the need for new techniques applied to image exams to improve the functional assessment of cardiovascular diseases by replacing manual activities with an automated segment selection. **METHODOLOGY:** This study presents a deep object detection neural network architecture, called DeepCADD to determine the lesion location in right coronary arteries (RCA) angiography exams. Using a Mask Region-Based Convolutional Neural Network (R-CNN), we expect to reach precision comparable to the gold standard, automating one step of the current protocol. We replace the Mask R-CNN's backbone with a ResNet-50 trained with coronary artery segments to improve the small features detection. We also train the whole DeepCADD architecture with angiographies collected in a local institution. **RESULTS:** DeepCADD outperformed similar networks in terms of sensitivity and presented a significant correlation with specialists during the validation, which suggests that DeepCADD can be used in the current angiography protocol. **CONCLUSION:** DeepCADD increases the correlation between the specialists and provides visual CAD suggestions, specially in multi-vessel lesions, which differentiates DeepCADD from the current literature. DeepCADD detects a high number of true positive candidates for lesion quantification, which we expect to extend for further arteries and dynamic evaluation in future research.

Keywords: Coronary artery. Deep learning. Coronary artery disease. Object detection.

RESUMO

CONTEXTO: Doenças cardiovasculares representam a causa número um de óbitos no mundo, e inclui a doença mais comum na saúde cardíaca, chamada de doença arterial coronariana (DAC). DAC é causada principalmente pelo acúmulo de gordura no interior das paredes arteriais, criando uma placa aterosclerótica que impacta o comportamento funcional no fluxo sanguíneo. As características anatômicas das placas são essenciais para a correta avaliação funcional das DACs. De fato, não há método único para avaliar todos os segmentos da artéria coronária com alta acurácia. **OBJETIVO:** O panorama apresentado, evidencia a necessidade de novas técnicas aplicadas em exames de imagem para melhorar a avaliação funcional de doenças arteriais coronárianas, substituindo etapas manuais com detecção automática de lesões. **METODOLOGIA:** Esse estudo apresenta uma arquitetura de rede neural para detecção de objetos, chamada DeepCADD para determinar a posição da lesão em exames de angiografias em artérias coronárias esquerdas. Usando uma rede neural convolucional baseada em regiões (Mask R-CNN), nós buscamos atingir precisão comparável ao padrão-ouro, automatizndo uma etapa manual no protocolo atual. Nós substituímos o backbone da Mask R-CNN com uma rede ResNet-50 trainada com segmentos de artérias coronárias para melhorar a detecção de pequenos objetos em imagens de angiografia. Nós também trainamos o DeepCADD com angiografias coletadas em uma instituição de saúde local. **RESULTADOS:** DeepCADD apresentou melhores resultados de sensibilidade em comparação com os estudos relacionados e correlação significante com os especialistas durante a validação, o que sugere seu uso no protocol atual da angiografia. **CONCLUSÃO:** DeepCADD aumentou a correlação entre os especilistas e proveu sugestões de DAC, especialmente em lesões com vários segmentos afetados, diferenciando a arquitetura proposta da atual literatura. DeepCADD detecta um grande número de candidatos verdadeiros positivos para a posterior quantificação das lesões. Com isso, esperamos expandir o uso do DeepCADD para as demais artérias e para a avaliação dinâmica de lesões em estudos posteriores.

Palavras-chave: Artérias coronárias. Aprendizado profundo. Doença arterial coronária. Detecção de objetos.

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ACS	Acute Coronary Syndrome.
AHA	American Health Association.
AI	Artificial Intelligence.
AMI	Acute Myocardial Infarction.
ANN	Artificial Neural Network.
CAD	Coronary Artery Disease.
CCA	Coronary Cine-angiography.
CCTA	Coronary Computed Tomography Angiography.
CFD	Computational Fluid Dynamics.
CMR	Cardiovascular Magnetic Resonance.
CNN	Convolutional Neural Network.
CVD	Cardiovascular Diseases.
DIP	Digital Image Processing.
DL	Deep Learning.
DS	Diameter Stenosis.
ESC	European Society of Cardiology.
FFR	Fractional Flow Reserve.
FMM	Fractional Myocardial Mass.
IFR	Instantaneous Wave-free Ratio.
IoU	Intersection Over Union.
IVUS	Intravascular Ultrasound.
LCA	Left Coronary Artery.
LL	Lesion Length.
LMS	Left Main Stem.
MAP	Mean Average Precision.
MAR	Mean Average Recall.
MCC	Matthews Correlation Coefficient.
ML	Machine Learning.
MLA	Minimal Lumen Area.
MLD	Minimal Lumen Diameter.
MLP	Multilayer Perceptron.
OCT	Optical Coherence Tomography.
QCA	Quantitative Coronary Angiography.
QFR	Quantitative Flow Ratio.
R-CNN	Region-Based Convolutional Neural Network.
RCA	Right Coronary Artery.

ResNet	Residual Network.
RGB	Red, Green And Blue.
ROIs	Regions Of Interest.
RPN	Region Proposal Network.
RVD	Reference Vessel Diameter.
SLR	Systematic Literature Review.
SVM	Support Vector Machine.
VOC	Visual Object Challenge.

CONTENTS

1 INTRODUCTION

Cardiovascular diseases (CVD) represents the number one cause of death in the world (OVALLE-MAGALLANES et al., 2020). More people die from CVD annually than from any other cause (World Health Organization, 2017). CVD is a group of disorders that impacts the heart's health and the circulatory system, which includes coronary disorders, also called coronary artery disease (CAD). CAD is mainly caused by the narrowing of coronary arteries, which is usually caused by fat and cholesterol that builds up in the internal walls of coronary arteries. This process is called atherosclerosis. From the anatomical point of view, atherosclerosis causes stenosis, which is the artery geometrical narrowing (GUYTON; HALL, 2006). From the hemodynamic point of view, atherosclerosis causes ischemia, which is the artery blood flow reduction (GUYTON; HALL, 2006).

The WHO (World Health Organization, 2017) estimates that around 23.6 million people will die from CVD up to 2030, mainly from CAD and stroke. According to the stenosis severity and localization, it can cause acute myocardial infarction (AMI) or stroke. AMI can be related to coronary artery narrowing. It reduces the blood flow limiting the oxygen transport to the myocardial tissue (HIDEO-KAJITA et al., 2019). The coronary artery system is composed of two main arteries: the left coronary artery (LCA), and the right coronary artery (RCA). The LCA and RCA consume around 75% and 25% of the aorta blood flow designated to the myocardial irrigation, respectively (GUYTON; HALL, 2006).

The coronary angiography (also known as coronary cine-angiography (CCA)) together with the FFR are described in the European Society of Cardiology (ESC) and American Health Association (AHA) as the gold standard for the CAD evaluation (HIDEO-KAJITA et al., 2019). Coronary angiography constitutes the conventional method for assessing the physiology and the severity of coronary artery disease by visual assessment (RAMASAMY et al., 2020). FFR, on the other hand, has been widely used to determine stenosis functional impact on blood flow during the maximal hyperemia (necessary to simulate the maximal blood flow), measuring the pressure decay caused by the artery narrowing (HIDEO-KAJITA et al., 2019). However, FFR is a high cost, invasive approach limited by the impacts of the vasodilators in some patients with asthma and chronic obstructive pulmonary diseases (FUKUOKA et al., 2020). Thus, FFR is being less employed than it is recommended. The functional assessment of the blood flow reduction is essential since the geometrical evaluation of the plaque is not sufficient for the correct lesion diagnosis. In such panorama, image-based techniques also provide valuable information to correlate with the renal blood flow reduction presenting the advantage of being less invasive. Nevertheless, their use is currently under discussion and usually limited to small validation (WESTRA et al., 2018; KO et al., 2015; MIN et al., 2015; PANG et al., 2016).

As seen, there is still no clear consensus on the best approach to perform a reliable functional assessment of coronary stenosis and several techniques are under development. The divergence of treatment options together with the number of emerging studies proposing new techniques to

offer more shreds of evidence of lesion severity create an opportunity to act in the patient triage phase. Based on this background, the main goal of this Master Thesis is to provide a second opinion in the atherosclerosis assessment by providing useful information for cardiologists in an understandable way, so it speeds up the lesion localization. To achieve this goal, we propose a model based on a Convolutional Neural Network (CNN) connected to a Region-Based Convolutional Neural Network (R-CNN) that work together to provide a diagnostic suggestion. The model obtains the classification and the location of the coronary narrowing in angiography exams with the highest sensitivity value in comparison with the recent literature, performing an important role of time-saving in the lesion analysis.

1.1 Motivation

CNNs have succeeded as a method to learn and discriminate characteristics (a.k.a. features) from raw data to classify lesions (OVALLE-MAGALLANES et al., 2020). Machine Learning (ML) is widely used in the healthcare area due to its ability to learn and provide faster insights to physicians (DEY et al., 2019). ML can learn rules and identify patterns progressively from large datasets, without being explicitly programmed or having any a priori assumptions (DEY et al., 2019). These advances have enabled the development and improvement of digital image processing techniques. Using these techniques, some Computer-aided Design systems applied to medical or biomedical imaging can achieve results close to human capacity (OVALLE-MAGALLANES et al., 2020). Especially in recent years, the use of ML for diagnosis, prognosis, and lesion classification has become a popular line of research (PYXARAS et al., 2018).

Recent literature states that the use of non-invasive techniques reaching high performance is promising to replace the invasive FFR (BENENATI et al., 2018). For this reason, the proposal of parallel processing in the coronary artery assessment can pave the way to interpretative data and the diagnosis speed-up in earlier lesion stages. Recent studies discussed alternatives to reduce the lesion evaluation invasiveness (MOON et al., 2021; CONG et al., 2021; PANG et al., 2021; RODRIGUES et al., 2021; WUA et al., 2020), as the angiography itself already provides important lesion physiological details (RODRIGUES et al., 2021).

These studies propose screening tools for patient triage, avoiding the need for additional exams such as FFR. The screening tools proposed are limited mainly by the multi-vessel identification, lacking specialist validations (RODRIGUES et al., 2021; WUA et al., 2020) and pure image classification (MOON et al., 2021; CONG et al., 2021; PANG et al., 2021), which was largely applied in the CAD context without effective success in real-life scenarios (FREITAS et al., 2021a). Current techniques and recent studies have focused on single lesion identification, which does not fit the most common scenarios. While pieces of evidence question the long-term benefit of decisions taken by them, recent studies lack specialist validations, evaluate models purely by numerical results, and work with deep learning architectures as black boxes.

1.2 Research Question

This work seeks to answer the following research question:

> *How to devise a model to detect coronary artery disease from angiography images based on a deep object detection neural network architecture with a performance comparable to the gold standard?*

The present work intends to identify the necessary architecture characteristics for an object detection system able to provide relevant information for cardiologists in angiography exams. To address the research question, we conducted a systematic literature review (SLR) to identify the state of the art, metrics, baselines, and opportunities available for coronary artery assessment.

In this Master Thesis, we are interested in analyzing how CNN learns coronary artery characteristics in angiography exams together with how object detection mechanisms can select relevant Regions of Interest (ROIs) in medical X-ray exams. Together with CNN-extracted features, we explore the use of preprocessing techniques and transfer learning to generate meaningful data for the diagnosis.

Finally, the model intends to present the significant lesion area, with bounding boxes aiming to apply measurement features to determine the exactly geometrical lesion information. Based on the CNN feature maps, an object detection network will be used to generate these bounding boxes.

1.3 Goals

The general goal of this Master Thesis is to investigate a Deep Learning (DL) model to detect CAD in coronary angiographies. With this model, we expect to improve the current angiography protocol, automating the lesion selection in the RCA segments with a frame-specific analysis to provide a second opinion in the atherosclerosis assessment. We expect these suggestions to create useful information for cardiologists in an understandable way. To do so, we have defined some specific goals:

- Understand the coronary artery assessment process;
- Identify the state of the art for coronary assessment, identifying challenges, techniques, useful metrics, and baselines through a Systematic Literature Review (SLR) ;
- Implement a classification model as a base for the hybrid final proposal;
- Propose an intelligent and hybrid architecture to classify and detect coronary lesion;
- Develop the proposed architecture;

- Validate the architecture based on the validation metrics and baselines identified in the SLR.

1.4 Scientific Contributions

In this Master Thesis, we investigate a Deep Learning (DL) model to detect CAD in coronary angiographies. With this model, we expect to improve the current angiography protocol, automating the lesion selection in the RCA segments with a frame-specific analysis. In this way, the main contributions of this work are detailed below.

- **Mapping of the state of the art metrics for CAD analysis**. With a literature review conducted over the recent literature of both computing and healthcare area. We used a SLR methodology to conduct this study analysis. In this context, we present a new taxonomy for CAD analysis and a common quantitative and qualitative analysis over the recent technologies applied to this end, resulting in the most suitable metrics that correlate with the CAD significance.

- **New convolutional architecture to automatically detect CAD in angiographies**. Based on the current gaps identified in the SLR, we proposed new architecture. In this way, the main scientific contribution of this dissertation is the proposal of a novel Deep Learning Architecture for Automatic Detection of Coronary Artery Disease, called DeepCADD, which outperforms the current literature for CAD evaluation focused on instance segmentation.

- **Explainability in the CAD detection.** The proposed model is based on the current angiography protocol, focused in the CAD detection in RCA segments. The pixel-level classification allows the local and global evaluation of lesions, also allowing the evaluation of border cases. Finally, we provide the heat maps of DeepCADD detection, which can help in the decision-making.

1.5 Text Organization

This Master Thesis is organized into 7 chapters. Chapter 2 presents the concepts related to coronary artery assessment, image processing, and machine learning, introducing the relevant technologies implemented in this work. In Chapter 3, the related works are discussed to present the state-of-the-art for coronary artery assessment. Chapter 4 presents the proposed model and design decisions. Chapter 5 presents the validation methodology and metrics for the model followed by the preliminary results in Chapter 6. Finally, Chapter 7 presents the conclusions obtained so far and future directions.

2 BACKGROUND

This chapter presents the main concepts used in this Master Thesis, which guided the development of intelligent architecture to assist in the classification and detection of coronary artery diseases in angiographies. Preliminary concepts regarding coronary arteries and their assessment are detailed in Section 2.1. In Section 2.2 we explain some important Digital Image Processing (DIP) features essential to the machine learning process. In Section 2.3 we present important ML concepts necessary to this work technical implementation. In Section 2.4 we present details about the architecture concepts in CNN necessary for the classification and for object detection on this work. Finally, in Section 2.5 we present the neural network architecture techniques to avoid over- and underfitting in models with limited datasets.

2.1 Coronary arteries

Coronary arteries lie on the surface of the heart and penetrate into the cardiac muscle and are responsible for almost all nutritive blood supply of the heart (GUYTON; HALL, 2006). The coronary circulation is more complex than the whole body circulation, as during the systole step the blood flow is reduced by contraction myocardial muscle, and then, in the diastole, the blood flow achieves its maximum (GUYTON; HALL, 2006). Figure 1 presents the main coronary arteries together with the RCA main segments. The RCA was defined as the study object as it has less morphology variation in comparison with the LCA. The RCA is composed mainly of three parts: proximal, medial, and distal segments. The bigger lumen diameter is on the proximal segment and reduces down to the distal segment (GUYTON; HALL, 2006).

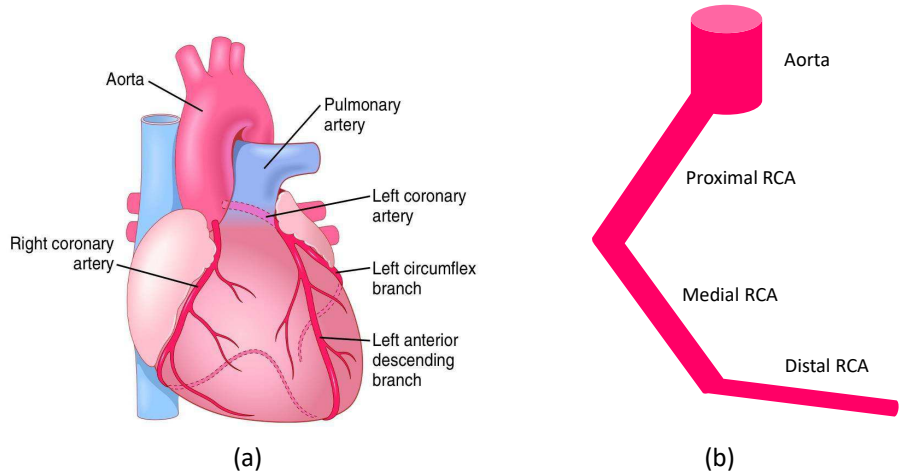
Figure 1 provides a simple model to demonstrate the main contour of the RCA. Simplifying the contour of the coronary artery we expect to reduce the problem complexity by focusing on the main artery segment. Also, this model represents the seed we have introduced to the CNN, focusing in the proximal, medial and distal RCA segments. For further contextualization, we will introduce the blood flow behavior up to its reduction impact in the right coronary system.

The blood flow is governed by Ohm's law of fluid flow (POLLOCK et al., 2020) in all bodies. Physiologically, it means that blood flow is equivalent to the pressure divided by the systemic resistance (POLLOCK et al., 2020). Basic mechanic concepts and clinical aspects of physiologic assessment present pieces of evidence of the FFR effectiveness in the correct assessment of coronary stenosis (HWANG; LEE; KOO, 2016).

2.1.1 Coronary Artery Disease

CAD is the most common disorder in heart health and especially responsible for cardiovascular diseases be the number one cause of death globally (GUYTON; HALL, 2006). CAD is mainly caused by fat accumulated in the arteries' internal walls, creating an atherosclerotic

Figure 1 – Coronary artery system (a) (GUYTON; HALL, 2006) and RCA representation (b).



Source: Adapted from (GUYTON; HALL, 2006)

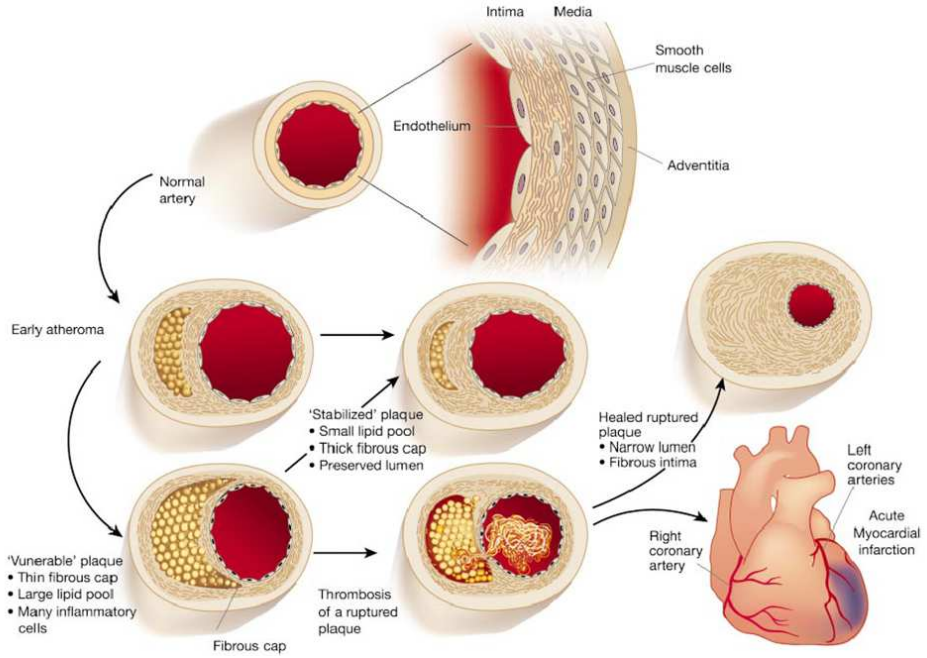
plaque that impacts the blood flow functional behavior. This process is called atherosclerosis. From the anatomical point of view, atherosclerosis causes stenosis, which is the artery geometrical narrowing (GUYTON; HALL, 2006).

The plaque caused by atherosclerosis presents some characteristics, which allow its classification as stable or unstable. The stable plaque expands gradually increasing the lipid (e.g., fat) according to the normal reproduction of the cells (GUYTON; HALL, 2006). On the other hand, the unstable plaque expands rapidly due to more frequent lipid deposition (GUYTON; HALL, 2006). From the clinical point of view, the unstable lesion differs from the stable one as it is prone to rupture, which increases the risk of acute thrombosis represented in Figure 2.

From the clinical point of view, the disease caused by plaque can be classified as stable or unstable angina. According to the ESC guideline, stable angina is represented by the presence of atherosclerosis but the symptoms may occur during physical activity and last for about 5 minutes, relieving with rest (GUYTON; HALL, 2006; MONTALESCOT et al., 2013). For unstable angina, symptoms occur even at rest, thus becoming more frequent, severe, or prolonged (GUYTON; HALL, 2006; MONTALESCOT et al., 2013).

Unstable angina significantly increases the risk of acute coronary syndrome (ACS). ACS represents the sudden, complete blockage of a coronary artery caused by an acute atherothrombotic event caused by the plaque rupture or erosion (MONTALESCOT et al., 2013). Unstable angina can have long stable periods, becoming unstable at any time (LIBBY, 2002). Atherosclerotic lesions can also be named as culprit or non-culprit by the cardiac ischemic event, designated AMI. The significant narrowing is designated as the reason for AMI and is named as culprit lesion. Any other narrowing not directly connected to AMI is named as non-culprit lesion (JOHNSON et al., 2016).

Figure 2 – Plaque build-up process (extracted from (LIBBY, 2002)). The constant increase of fat in the artery internal walls can lead to a stabilized or vulnerable plaques. The second can evolve to a artery blockage and reduce the blood flow drastically.



Source: Adapted from Libby (2002).

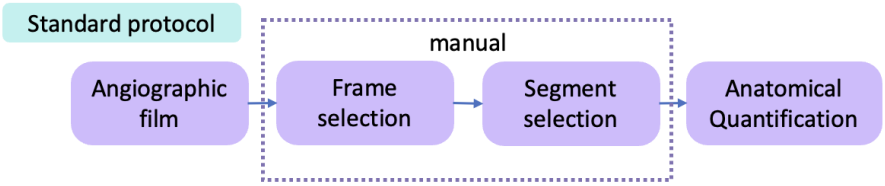
2.1.2 Coronary Artery Assessment

The CCA together with the FFR are described in the Brazilian Coronary Artery Guideline as the gold standard for the CAD evaluation (HIDEO-KAJITA et al., 2019). Coronary angiography is a less invasive technique, which focuses on the geometric characteristics of lesions. On the other hand, the FFR is a more invasive approach, which evaluates the pressure decay in the narrowed arteries responsible for the current lesion functional repercussion assessment. Nevertheless, these studies still suggest that the visual assessment or local pressure drop measurement alone do not represent the real coronary disease functional assessment (PYXARAS et al., 2018).

Coronary angiography constitutes the conventional method for assessing the physiology and the severity of coronary artery disease by visual assessment (RAMASAMY et al., 2020). Coronary angiography analysis is based on the quantitative coronary angiography (QCA) feature, a visual metric that measures the physiologic silhouette deformation caused by the fat plaque in each exam frame (MOERTL et al., 1995). QCA calculation is mainly composed by the artery minimal lumen diameter (MLD) metric. MLD is based on a reference vessel diameter (RVD), which represents the healthy segment without fat plaque. QCA main result is the quantitative percentage of diameter stenosis (DS).

Allied to the geometrical assessment, the FFR is an invasive approach that determines the likelihood that stenosis will prevent tissues from receiving oxygen (HIDEO-KAJITA et al.,

Figure 3 – Current protocol for functional assessment based on the angiography.



Source: Elaborated by the author

2019). The FFR measurement occurs at the maximum hyperemia induced by an intravenous infusion of adenosine (LEE et al., 2020). An FFR of 0.80 indicates that the stenotic coronary artery supplies 80% of the normal maximal flow and is considered functionally significant when the FFR is 0.80 or less (LEE et al., 2020). Its use in ACS is less common due to the maximal hyperemia required and the disease urgency (MARTINS et al., 2018).

Thus, FFR is being less executed than it is recommended. In fact, the functional assessment of the blood flow reduction is essential, since the geometrical evaluation of the plaque is not sufficient for the correct lesion diagnosis. In such panorama, image-based techniques also provide valuable information to correlate geometrical lesion data with the real blood flow reduction, whereas presenting the advantage of being less invasive. Their use is currently under discussion once the use of angiography alone can reach as good results as FFR, reinforced by benefits of the less invasive techniques use in patient's treatment (WESTRA et al., 2018; KO et al., 2015; MIN et al., 2015; PANG et al., 2016).

Current daily protocol for CAD assessment is represented in Figure 3, where the operator selects the frame where the contrast fluid has maximum adherence in the coronary artery, followed by the manual segment selection and then quantification based on QCA. QCA methodology receives the lumen border centerline manually traced by the operator in 2D angiographic projections (ATHANASIOU et al., 2017), constructs the lumen contour using as reference the distance between the lumen border detected before, and the plaque quantification is provided to the operator (ATHANASIOU et al., 2017).

Recently, a computational method has emerged, called quantitative flow ratio (QFR) index (WESTRA et al., 2018), considered the state of the art for functional assessment. QFR aims to be a less invasive method and has been validated with optimal determination of the narrowing significance considering the total flow from aorta (WESTRA et al., 2018). The impact of the narrowing in comparison with the aorta was validated with FFR (WESTRA et al., 2018). Based on the angiography, QFR evaluation is based on anatomical metrics such as DS, minimal lumen area (MLA) and MLD (PYXARAS et al., 2018).

Multi-center registries which evaluated the discordance between anatomical and physiologic assessment of coronary arteries using the fractional myocardial mass (FMM) index suggest that FMM improves the correlation between anatomical and physiologic assessments (KIM et al., 2016). FMM and MLD discovered vessels with FFR < 0.80 and with FFR > 0.80 better

than visual assessment (KIM et al., 2016). Nonetheless, FMM presents some limitations as it assumes that the artery segment is a cylindrical tube without volume compliance and flows turbulence, which may vary from the real case since blood is considered a non-Newtonian fluid (KIM et al., 2016; ADJEDJ et al., 2015). The coronary functional assessment is challenging and to the best of our knowledge, no single approach is the best option for assessing all possible lesions characteristics. Coronary assessment image-based emerging technologies can pave the way for further lesion and hemodynamic details, thus providing additional pieces of evidence for diagnosis decisions and treatment. Due to the nature of these methods, some important image processing concepts will be introduced in the next section.

2.2 Digital Image Processing

DIP can be defined as a set of computational techniques to acquire, represent, manipulate, and extract information from digital images (ANNADURAI, 2007). It means that DIP starts from image grabbing up to the recognition and interpretation. A DIP system have four major building blocks: acquisition, storage, processing, and communication interface. These blocks are important to guarantee the the communication between image processing systems (ANNADURAI, 2007). The use of DIP concepts is necessary on this work for image preparation module and correct filter selection to achieve the expected study objective.

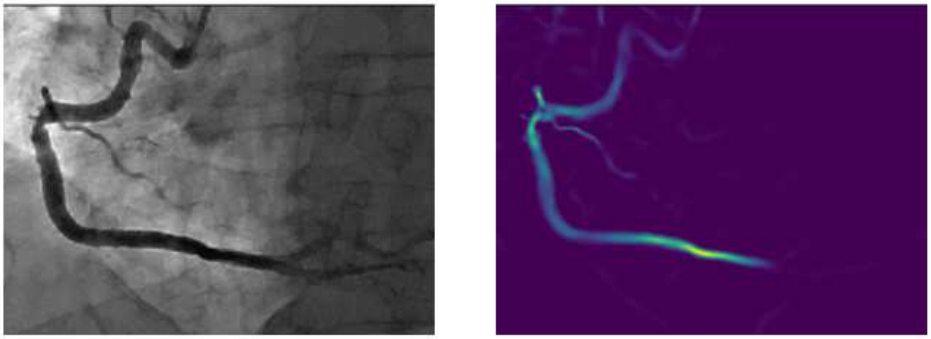
2.2.1 Digital Image

Any monochrome image can be represented by means of a two-dimensional (2D) light intensity function $f(x; y)$, where x and y denote spatial coordinates, and f corresponds to the intensity or level of gray in the image (ANNADURAI, 2007; GONZALEZ; WOODS, 2018). For a color digital image, in general, a three-dimensional (3D) representation is required (BURGER; BURGE, 2009). The most relevant 3D model for this work is the Red, Green and Blue (RGB) format, where each channel represents the intensity of each primary color. The intensities sum determine the final color (GONZALEZ; WOODS, 2018). In this work, RGB images are always converted to grayscale, converting a three-dimensional into a two-dimensional array due to the relevant data contained in the CCA.

2.2.2 Image Segmentation

Image segmentation intends to divide an image into regions of similar attributes (PRATT, 2013), using mainly two attributes: in monochrome images, the luminance; and in color images, the color components (PRATT, 2013). The division is carried out at the pixel level, so those pixels with the same label are connected concerning some visual or semantic property (GHOSH et al., 2019). In this section, we discuss appropriate methods for targeting CCA.

Figure 4 – Original angiography (a) and Frangi filter applied (b).



Source: Elaborated by the author

2.2.2.1 Frangi filter

Frangi filter works for vessel enhancement and is widely accepted in recent past research studies (KULATHILAKE et al., 2010). The Frangi filter is typically used to detect vessel-like or tube-like structures and fibers in volumetric image data (FRANGI et al., 1998). Nowadays, the Frangi filter is an excellent alternative for vessels identification and takes an important role in this work preprocessing module presented in the 3.1 section. The Frangi behavior is represented in Figure 4.

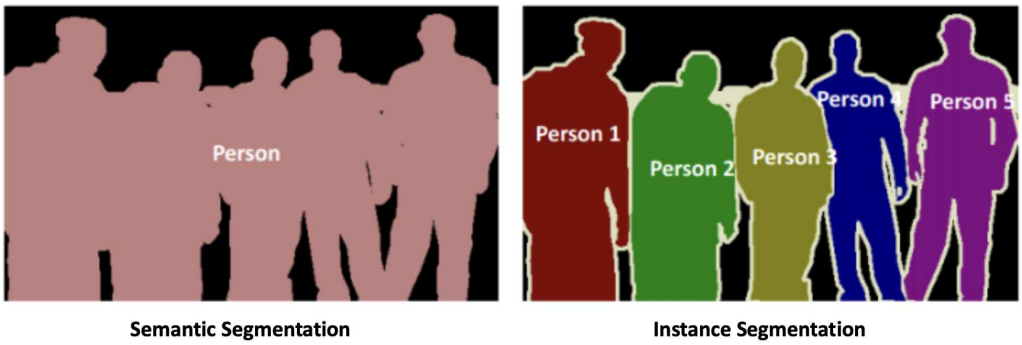
As Figure 4 presents, the Frangi filter provides a very useful technique that reduces the effort to select both vessels and arteries. Nevertheless, depending on the image contrast quality and the amount of overlaid vessels in some coronary angiography projections this filter will also enhance not relevant image features. In general, this filter is very useful and with the Frangi filter the CCA preprocessing tends to become easier, once it enhances the artery edges, providing useful information for image binarization.

Other binarization techniques are necessary for the best use of the Frangi filter. Among them, we highlight thresholding, Otsu's method, and adaptative thresholding. The thresholding technique works to classify at pixel level what is a background and what is an object in an image based on a predefined pixel intensity value (BURGER; BURGE, 2009). This technique is particularly useful to perform image binarization and to remove noise (GONZALEZ; WOODS, 2018), once pixels brighter than the threshold are mapped to unity display and all the remaining pixels below the threshold are mapped to zero level of display luminance (PRATT, 2013).

Otsu's method, proposed in 1979, provided a different approach to image binarization. Instead of choosing an arbitrary value for thresholding, Otsu's method determines it automatically (GONZALEZ; WOODS, 2018). The proposed algorithm iterates through all the image pixels looking for all peaks in the image histogram separating the image in background and object (GONZALEZ; WOODS, 2018).

Traditional segmentation using an arbitrary thresholding value works well only in images

Figure 5 – Difference between instance and semantic segmentation



Source: Adapted from (VARATHARASAN et al., 2019)

where there is a distinction between the colors of the background and the objects (GONZALEZ; WOODS, 2018). In this Master thesis' case, the CCA pixel intensity can vary for many reasons, so an arbitrary value is not enough. For this reason, the adaptative thresholding helps, once it determines the pixel-based threshold in small regions around it (GONZALEZ; WOODS, 2018). In other words, adaptive thresholding uses variable limits to perform segmentation, considering local characteristics of the image (BURGER; BURGE, 2009).

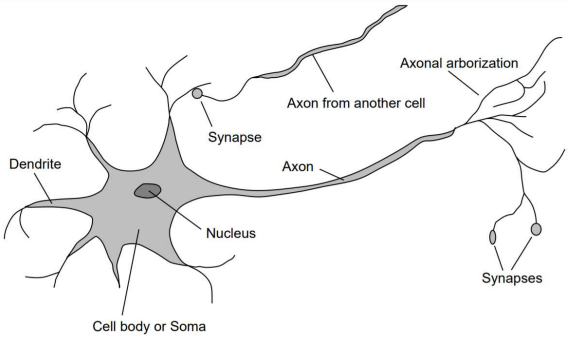
Nevertheless, even with a powerful segmentation technique, the identification of narrowed arteries is not guaranteed by them and the use of ML techniques become necessary and will be explored in the next section.

2.2.2.2 Instance segmentation

Different from the classification from CNN, object detection or localization provides both the detected classes in the image and the location of the image where the objects have been classified. The use of Mask R-CNN architecture includes, not only the bounding boxes and the classification but also the instance segmentation for the object detected (HE et al., 2018). Figure 5 presents the difference between semantic segmentation and instance segmentation. While the semantic segmentation provides the predicting labels' inference for every pixel included in the predicted class, the instance segmentation is the task of finding both the semantic segmentation and the object detection (HAFIZ; BHAT, 2020).

Instance segmentation provides different labels for each instance of the object detected belonging to the specific class (HAFIZ; BHAT, 2020). Instance segmentation is especially challenging for small objects. Small CNN presents the benefit of the narrow receptive field with richer detail, but the resolution is higher and the semantic sensitivity is lower (HAFIZ; BHAT, 2020). It remains an issue due to the CNN limitations, as the number of layers in the network increase. Small objects details can disappear in the higher layers, making the object detection quite challenging (HAFIZ; BHAT, 2020).

Figure 6 – Biological neuron main elements



Source: Adapted from Russell et al. (2010)

2.3 Machine Learning

Machine learning (ML) is widely used in healthcare due to its ability to learn and provide faster insights to doctors (DEY et al., 2019). ML can learn rules and identify patterns progressively from large datasets, without being explicitly programmed or having any a priori assumptions (DEY et al., 2019). Unfortunately, the image techniques to which ML has been applied to suffer from several limitations, such as data quality and pending image acquisition protocols (DEY et al., 2019; SHU et al., 2020). Therefore, back-propagation networks are known due to their ability to learn complex problems with both large and small collections of images. This capability allows the learning of high-dimensional and non-linear patterns in these collections (LECUN et al., 2010).

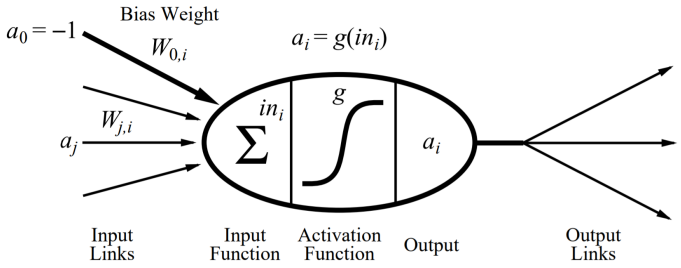
The main ML component discussed in this study is the Artificial Neural Network (ANN). The ANN is the artificial intelligence component created to simulate the human brain functioning using the biological neurons and their interconnections as inspiration (RUSSELL et al., 2010). In the next sections, we discuss some important concepts of ANN that are essential to the correct CCA evaluation.

2.3.1 Artificial Neuron

The artificial neuron is a representation of a natural neuron, that has four main elements: dendrites, axons, cell body, and synaptic terminals (RUSSELL et al., 2010), presented in Figure 6. The first artificial neuron model proposed in 1943 was composed of several binary inputs and binary output. Even very primitive, it allows the improvements which created, in 1958, the *Perceptron* concept. The *Perceptron* model has several numerical input and the output calculated by a threshold function (RUSSELL et al., 2010).

Currently, the artificial neuron model, which is presented in Figure 7 is composed by several inputs $a_{\text{i-j}}$, each one multiplied by a specific $w_{\text{i-j}}$ weight, which corresponds to the connection strength. Each neuron also have a constant value called $bias$ (RUSSELL et al., 2010). The

Figure 7 – Artificial neuron representation



Source: Adapted from Russell et al. (2010)

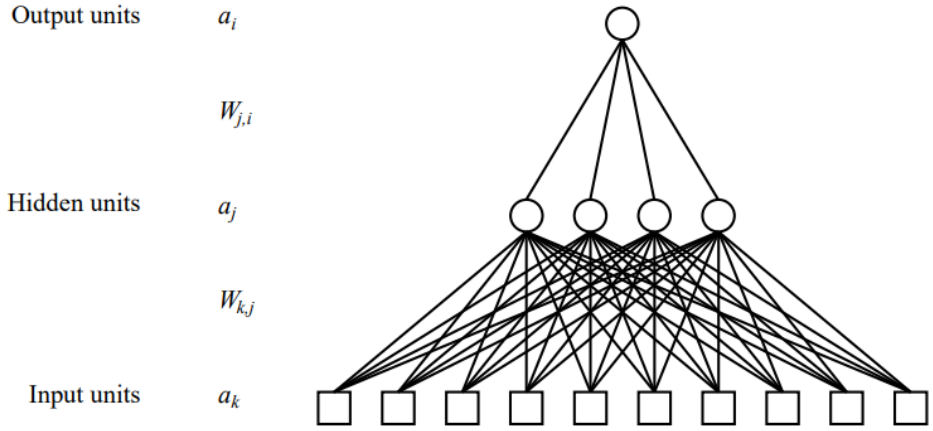
inputs are added together and then applied to a function g (a.k.a activation function), to produce the neuron output, which is then propagated to the next interconnections (RUSSELL et al., 2010).

2.3.2 Multilayer Perceptron

When the neurons are organized in layers, it is called Multilayer Perceptron (MLP) (RUSSELL et al., 2010). A typical MLP architecture is composed of one input layer, one or more hidden layers, and an output layer (RUSSELL et al., 2010) represented in Figure 8. The biggest computational process occurs in the hidden layers, given the neuron's interconnections and its activation functions. For binary problems, the output layer has only one neuron, while in multiclass problems the output layer has as many neurons as the number of classes. The weight update is an iterative process, while the x input signals are propagated from the input layer to the hidden layers producing an output $y0$, the back-propagation works in the opposite direction (RUSSELL et al., 2010; HEATON et al., 2018). The back-propagation process in the network produces what is called learning, which occurs by computing an error in the output layer. This error is measured by a loss function. This error is then back-propagated to the previous layers and the weights of the connections are updated with new values (HEATON et al., 2018). From the output $y0$ and its label y, the difference in the forecast is calculated with a loss function. The loss gradient is then computed (HEATON et al., 2018).

The gradient obtained from the loss function is propagated from the output layer to the hidden cameras to calculate the error in each neuron. The hidden layer error is obtained by calculating the partial derivative of the error concerning each weight employing the chain rule (HEATON et al., 2018). Finally, the neuron weights are then updated by subtracting a proportion of the gradient from the current value of the weights. The calculation of this proportion is defined by multiplying the gradient by the hyperparameter called learning rate (COPPIN, 2010). However, MLP are not the best idea to use for image processing. One of the main problems is that spatial information is lost when the image is flattened (matrix to vector) into an MLP. MLP using one perceptron for each input (e.g. pixel in an image), the amount of weights rapidly becomes unmanageable for large images.

Figure 8 – Multilayer perceptron representation



Source: Adapted from Russell et al. (2010)

2.3.3 Activation function

The objective of an activation function is to introduce non-linearity to the neuron (HEATON et al., 2018). This function defines the neuron output and must be capable of mapping entries between $\{-\infty, +\infty\}$ to an specific interval, normally between $\{0, 1\}$ or $\{-1, +1\}$ (HEATON et al., 2018).

There are a lot of activation functions available, such as sigmoid, tanh, maxout, SWISH, ReLU, and variants of ReLU, such as leaky ReLU, ELU, and PReLU. The most common activation function for image processing is the ReLU and its variations as they help in overcoming the vanishing gradient problem (KHAN et al., 2020). The most common activation functions implemented in the community are:

- **Step or threshold**: Map the input to $\{0, 1\}$, following Equation 2.1.

$$\alpha(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases} \tag{2.1}$$

- **Sigmoid**: Map the input to the interval $\{0, 1\}$, following Equation 2.2.

$$\alpha(v) = \frac{1}{1 + e^{-v}} \tag{2.2}$$

- **Tanh**: Map the input between $\{-1, 1\}$, following Equation 2.3.

$$\alpha(v) = tanh(v) \tag{2.3}$$

- **ReLU**: Map the input between $\{0, 1\}$, following Equation 2.4.

$$\alpha(v) = max(0, v) \tag{2.4}$$

- **Softmax**: Map the input between $\{0, 1\}$, following Equation 2.5.

$$\alpha(v) = \frac{exp(v)}{\sum exp(v)} \tag{2.5}$$

There is a consensus in the community about the best activation function in the last layer of ANN. For multi-class classification problems, the $softmax$ function has presented better results. For binary classification problems, $sigmoid$ function has presented a better result.

2.3.4 Loss functions

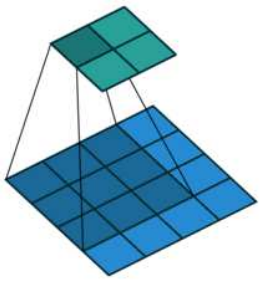
The gradient loss obtained during the back-propagation process is essential to the weights adjustment (RUSSELL et al., 2010). It intends to measure the difference between the estimated and the current values for a given training and it needs to be defined carefully (RUSSELL et al., 2010). The most common loss functions for classification are the Binary Crossentropy and the Categorical Crossentropy, and the most common for regression is the Mean Absolute Error and the Mean Squared Error.

Binary Cross-Entropy measures the classification model performance with a probability between 0 and 1. The closer the output is to 0, the better is the classification. A common sense from the community explains that Binary Crossentropy can only be used for binary classification problems. For multi-class classification problems, it is necessary the use Categorical Crossentropy, which calculates a separate loss for each observed class and makes a sum of the result to infer the error of the model (HEATON et al., 2018). For regression problems, the Mean Absolute Error represents the difference between the target and predicted values and the Mean Squared Error is the mean of the difference squared between the forecasts and the actual observations (HEATON et al., 2018).

2.4 Convolutional Neural Networks

The CNN (a.k.a ConvNet) learns from back-propagation strategy (KHAN et al., 2020) and for this reason CNN becomes popular as they combine three important architectural ideas: local receptive fields, shared weights and temporal sampling, which makes these networks excellent for pattern recognition (LECUN et al., 2010). CNN works well with data that has a spatial relationship (KHAN et al., 2020). For this reason, CNN has been widely applied in application areas including Image Classification and Segmentation, Object Detection, Video Processing, Natural Language Processing, and Speech Recognition (KHAN et al., 2020). CNN convolu-

Figure 9 – Convolutional layer representation



Source: Dumoulin e Visin (2016)

tional kernel also acts as a learning framework, reducing memory usage and computing faster (KHAN et al., 2020).

CNN is architectured by multiple stages and the input and output of each stage, called feature maps, are sets of arrays (LECUN et al., 2010). CNN are typically composed of multi 3-stage layers, which include a convolutional layer in the first stage. In the second stage, an activation function is applied to the convolutional result. Finally, in the last stage, a max-pooling layer is responsible for dimension reduction (LECUN et al., 2010).

The number of CNN necessary layers depends on the image complexity to capture low-level details. However, the computational cost also increases (KHAN et al., 2020). At the end of these connected 3-stage layers, a flatten layer is usually adopted. Considering these connected layers, over a series of epochs, the model becomes able to detect dominating low-level image features and classify the elements using the Softmax Classification technique (KHAN et al., 2020). Nevertheless, multiple connections that learn non-linear relations can also cause overfitting, which is important to avoid as they generate lots of false-positive classifications. These elements represent a common architecture of an image classification convolutional neural network and based on these elements, lots of architectures raised in the last decades, such as the AlexNet, ImageNet, Inception-ResNet, Wide ResNet, ResNeXt, VGG, etc (KHAN et al., 2020).

2.4.1 Convolutional Layers

The first stage of a typical CNN layer is the convolutional layer. This layer is responsible for the hover in the image with the kernel dimensions to extract sets of information from these image regions with a defined kernel, which represents a portion of the image limited by kernel dimension according to Figure 9. When the kernel stride by the image, changing the original image (Dumoulin; Visin, 2016).

Due to its weight sharing ability of the first stage, different sets of features within an image are extracted by sliding this kernel with the same dimension on the image and thus makes CNN parameter efficient as compared to the fully connected networks (KHAN et al., 2020). The generated kernels go through an activation function in the second stage which works as a gate

for given input data calculating neurons weight sum. It sums up similar information in the neighborhood pixels of the kernel and outputs the dominant response within this local region (LEE; GALLAGHER; TU, 2016).

A convolutional layer has some hyperparameters that need to be controlled, such as the filter size, which defines the size of the convolution filter; Number of filters, which defines the number of current layer filters; Padding, which defines edge value added to the original image; and Stride, which defines the step size the filters perform on convolution.

The last stage of a typical CNN layer is responsible for the dimensional reduction and can be composed by either Max Pooling or Average Pooling. This allows to reduce "unnecessary" information, and summarize the knowledge about a region (KHAN et al., 2020). While Max Pooling returns the maximum value from the portion of the image hovered by the Kernel and working as a noise suppressant, The Average Pooling returns the average of the values from the same covered image portion defined by the Kernel (KHAN et al., 2020). In the CNN the last layer, usually for classification problems, we want each final neuron represents the final class. The neurons behavior is connected to all previous layer activation, as a MLP (HEATON et al., 2018; KHAN et al., 2020).

2.4.2 Convolutional Neural Networks for Medical Image Classification

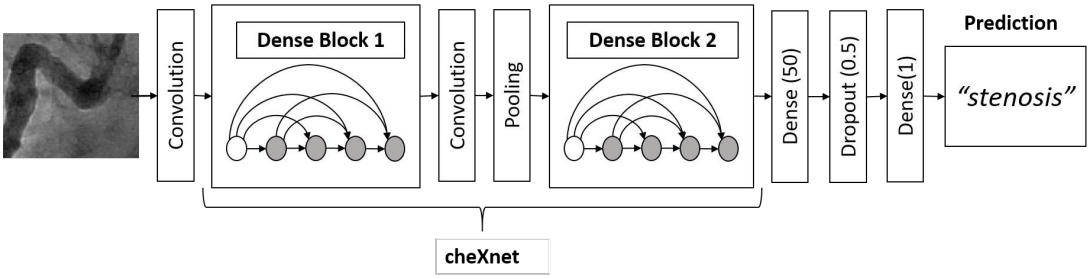
As an essential part of this Master Thesis, we propose the use of a pre-trained model to classify coronary artery segments from a synthetic dataset (ANTCZAK; LIBERADZKI, 2018). Recent studies suggest the better performance of some architectures in the classification of medical image data, such as the ResNet-50 (RAHIMZADEH; ATTAR, 2020), Xception (RAHIMZADEH; ATTAR, 2020; KHAN; SHAH; BHAT, 2020) and the recent CheXNet (RAJPURKAR et al., 2017), which is a DenseNet optimized to chest X-ray exams and have an image nature visually similar to angiographies. In the following subsections we will discuss three different architectures, the ChexNet, the Xception, and the ResNet:

2.4.2.1 CheXNet

The CheXNet architecture was proposed in 2016 focused on medical data classification problems (RAJPURKAR et al., 2017). CheXNet is a Dense Convolutional Network (DenseNet) with 121 layers trained with the ChestX-ray 14 dataset (RAJPURKAR et al., 2017). The use of DenseNet is particularly good to optimize the flow of information and gradients through the network, followed by a fully-connected layer with sigmoid nonlinearity activation (RAJPURKAR et al., 2017). Figure 10 presents the CheXnet architecture together with the proposed classification layer.

The CheXNet network weights are initialized with weights from ImageNet and trained end-to-end using Adam with standard parameters ($\beta 1 = 0.9$ and $\beta 2 = 0.999$) (RAJPURKAR et al.,

Figure 10 – CheXnet simplified architecture



Source: Adapted from (RAJPURKAR et al., 2017).

2017). CheXNet used a learning rate of 0.001 that is decayed by a factor of 10 each time the validation loss plateaus after an epoch (RAJPURKAR et al., 2017).

2.4.2.2 Xception

The Xception architecture was proposed in 2017 for classification problems and is a linear stack of depthwise separable convolution layers with residual connections (CHOLLET, 2017). The Xception has 36 convolutional layers which compose the network feature extraction base (CHOLLET, 2017).

In this network the data has an entry flow, which is followed by a middle flow repeated eight times and an exit flow, where all the Convolution layers are followed by Batch normalization (CHOLLET, 2017). Xception architecture has overperformed other architecture, such as VGG-16, ResNet and Inception V3 in classical classification challenges (CHOLLET, 2017).

2.4.2.3 Residual Network

Deep CNN has been widely used for image classification and is great at identifying low, mid, and high-level image features. Usually, stacking more layers in the network reaches better accuracy (HE et al., 2016). Nevertheless, the notorious problem of vanishing/exploding gradients impacts the performance of these networks. The result of vanishing gradient is the general inability to learn on a given dataset, which prematurely converges to poor solutions (HE et al., 2016).

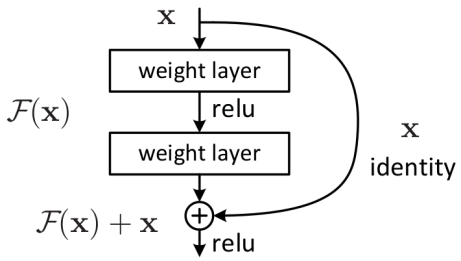
In detail, when deeper networks start converging, and the network depth increases, accuracy gets saturated and degrades rapidly (HE et al., 2016) not caused by overfitting. The addition of more layers, in this case, leads to higher training error (HE et al., 2016). In this context, the ResNet came up as modularized architectures that stack building blocks of the same connecting shape with the main advantage of skip connections. Figure 11 presents the ResNet architecture for different number of layers adapted from the original paper. ResNet that are over 100-layer deep have shown state-of-the-art accuracy for several challenging recognition tasks on ImageNet and MS COCO competitions (HE et al., 2016).

Figure 11 – ResNet architecture with variable number of layers.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Source: Adapted from He et al. (2016).

Figure 12 – Example of skip connection.



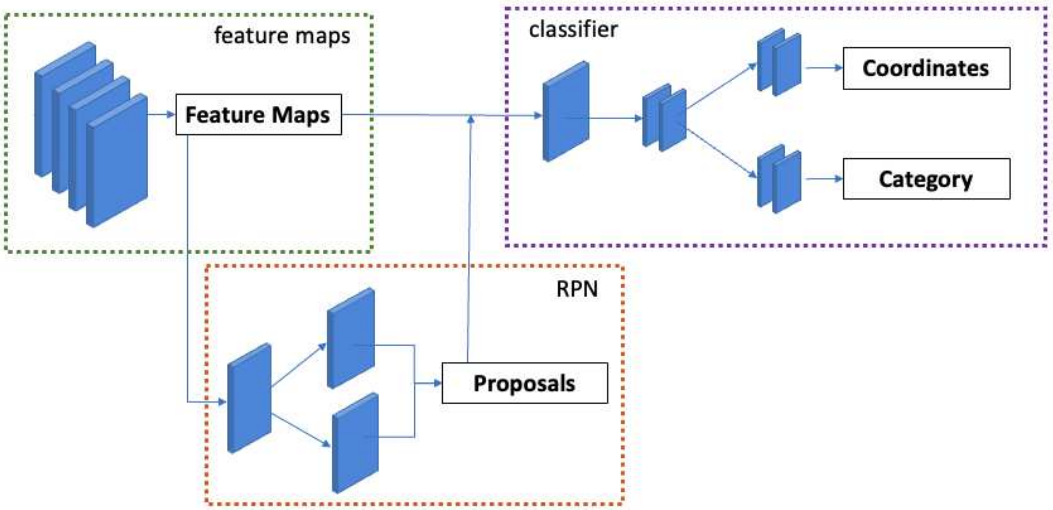
Source: Adapted from He et al. (2016).

Skip connections allow the propagation of small features learned by the initial convolutional layers to higher stacked layers. Figure 12 presents an example of skip connection also called as shortcut connection. The shortcut connections simply perform identity mapping, and their outputs are added to the outputs of the stacked layers (HE et al., 2016).

Skip connections are applied before the ReLU activation as shown in Figure 12. Researchers have found that this structure presents the best results (HE et al., 2016). Skip connections also allow the model to learn an identity function which ensures that the higher layer will perform at least as good as the lower layer, and not worse (HE et al., 2016).

Beyond the ResNet has won the 1st place in tracks in ILSVRC and COCO 2015 competitions, ResNet-50 is the default R-CNN backbone and have been widely used for medical data classification. ResNet has proven to be the best accuracy-to-speed ratio in coronary artery studies (DANILOV et al., 2021).

Figure 13 – Original Faster R-CNN architecture



Source: Elaborated by the author.

2.4.3 Convolutional Neural Networks for Object Detection

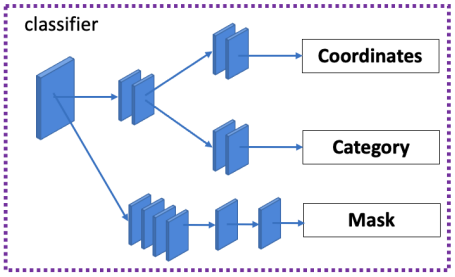
In this work, we propose an architecture to detect coronary artery diseases in angiographies and for this reason, the object detection task has an important role in the model. Object detection is the identification of an object in an image together with its localization and classification (CHAHAL; DEY, 2018). Software systems that can perform these tasks are called object detectors (CHAHAL; DEY, 2018).

Typically, there are three main steps in an object detection framework: the regions of interest or region proposals; the visual features are extracted for each of the bounding boxes; and in the final post-processing step, overlapping boxes are combined into a single bounding box. Object detection models are usually evaluated by Mean Average Precision (mAP) and are especially a good approach when the classification features are too small (CHAHAL; DEY, 2018). To do so, we decided to experiment with the R-CNN architectures.

R-CNN architectures are for object detection and contain a backbone, which is a CNN. The first R-CNN architecture proposed in the literature used selective search algorithm to determine ROIs using Support Vector Machine (SVM) to classify the different regions (CHAHAL; DEY, 2018). Recently, Region Proposal Network (RPN) came up to improve the capability of predicting regions of multiple scales. It replaced the high cost selective search algorithm that was considered as a bottleneck (CHAHAL; DEY, 2018). The use of parallel processing presented in Figure 13 and the inclusion of a RPN improved the network performance in the Faster R-CNN architecture (CHAHAL; DEY, 2018).

Recently, another enhancement from the Faster R-CNN showed up. The Mask R-CNN architecture extends Faster R-CNN by adding a branch for predicting an object mask in parallel with the existing branch for bounding box recognition (HE et al., 2018) presented in Figure 14.

Figure 14 – Mask R-CNN classifier module with the Mask parallel process



Source: Elaborated by the author.

Mask R-CNN is based on an *instance-first* strategy and the instance segmentation defines the object mask (HE et al., 2018) improved as a parallel prediction with the bounding-box proposal, being more flexible and simpler (HE et al., 2017).

2.5 Overfitting and Underfitting

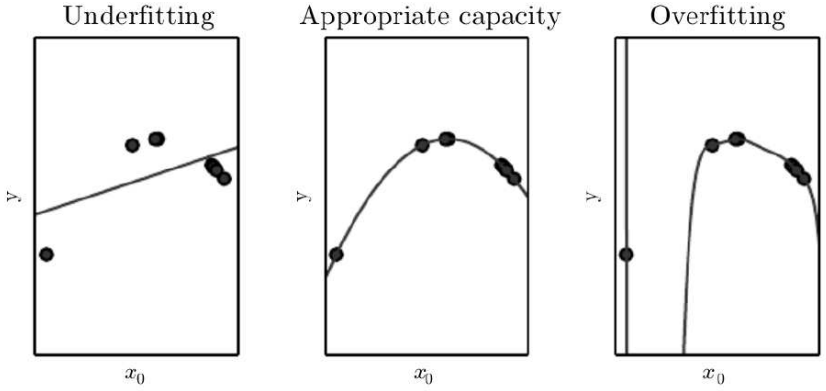
The training process requires the dataset division, at least into training and testing data. This separation is necessary, once an efficient model needs to perform well in both any data even, mainly after the training step. This capability is called generalization (HEATON et al., 2018). Usually, the dataset is divided into three parts (GÉRON, 2017): the training phase, which is responsible for the weight optimization process; the validation phase, which is responsible to monitor the model's behavior; and the test phase, which is responsible to measure the model performance while solving the problem.

There are two main challenges during training a model: underfitting and overfitting, represented in Figure 15. Underfitting occurs when the model does not obtain a low error for a given problem. The overfitting occurs when the gap between training error and validation is very large (HEATON et al., 2018). In the following two sections, we discuss some techniques to support these challenges and enhance the model generalization. We introduce the available techniques to enhance the current model without additional tools and the use of transfer learning.

2.5.1 Model and data preparation

In CNN architecture, it is popular the use of Dropout layers, which are responsible for a random dropping of some connections or units which produce abnormalities in the network architecture, also contributing to the network learning quality (KHAN et al., 2020). This layer "drops out" a random set of activations in that layer by setting them to zero. This makes the network more robust (kind of like you eating dirt builds up your immunity system, the network is more immune to small changes) and reduces overfitting. This is only used when training the network and is the Early Stopping, the Batch Normalization, and the Data Augmentation

Figure 15 – (Left) The model cannot capture the curvature that is present in the data, presenting underfitting. (Center) A quadratic function fit to the data generalizes well to unseen points. (Right) The model achieves a 9-degree polynomial function that fit exactly the training data suffering from overfitting.



Source: Heaton et al. (2018)

(HEATON et al., 2018).

The Early Stopping technique interrupts the model training based on loss function stop criteria in the training and validation sets. It guarantees the loss function does not be stuck in a local minimum. It avoids the waste of time and computational resources with excessive model adjustments.

The Batch Normalization raises as a strategy to make the network converge more quickly, once it is recommended to be added between after Dense layers. Given a set of features H of the layer to be normalized, Batch normalization is described as Equation 2.6:

$$H' = \frac{H - \mu}{\sigma} \tag{2.6}$$

Where μ is the average and σ is the layer feature maps standard deviation normalized. H' are normalized feature maps that are processed in the same way as if the network were processing H without normalization.

Finally, the Data Augmentation process consists of increasing the dataset via synthetic data based on the original dataset. The most common operations are rotation, translation, horizontal and vertical flip, subsampling, and distortions. These operations intend to generalize the classification avoiding memorizing image regions containing the class. However, some attention is needed to avoid generating noise in the training dataset.

2.5.2 Transfer Learning

The ability to learn from a large number of experiences, and export 'knowledge' into new environments is exactly what transfer learning is about (HEATON et al., 2018). The transfer learning process is another technique to avoid overfitting and underfitting, which consists of using trained weights for a given problem A in the initialization of another network to solve a

problem B. This technique is especially important in the healthcare area due to the difficulty to access labeled and public datasets (WONG; FORTINO; ABBOTT, 2020; SHU et al., 2020; ZREIK et al., 2018, 2019). In the training phase, we can adjust the weights of all layers or even freeze some of them leaving the original weights (HEATON et al., 2018).

In this context, the featuriser concept comes up. Essentially, in a featuriser we use the first layers of the network to determine the useful feature, but we do not use the output of the network, as it is too task-specific. There are two main ways to transfer learning: initialize a model with weights and retrain an existing model.

- **Initialize a model with weights**: Usually recommended when the number of images is lower than 1000 per class. During the training, we keep some layers fixed optimizing parameters in higher-level layers, once the initial layers have similar features regardless of the problem domain.

- **Retrain an existing model**: Usually recommended when the number of images is bigger than 1000 per class. An approach that most likely leads to overfitting, may need more parameter optimization than the number of images available. Given the similarity with the dataset originally used for training, then a large pre-trained network, this model can be used as a featuriser.

In the healthcare area, several techniques exist to improve performance, given the small number of public and labeled datasets. Typically require more work and fine-tuning (OVALLE-MAGALLANES et al., 2020). Typically, it suggests pre-train the weights of the classification network, by iteratively training each layer to reconstruct the images using convolutional or deconvolutional layers (OVALLE-MAGALLANES et al., 2020).

2.6 Final Remarks

Image-based techniques are helpful for burden plaque characterization, measure, and lumen attenuation gradient (DEY et al., 2019). Thus, the angiography nature studied highlighted the potential of image processing techniques to extract as much information as possible from exams. The understanding of exam characteristics allowed the best choice for filtering and segmentation tools. It is also important to highlight the potential of Artificial Intelligence (AI) to improve speed and quality of image acquisition (DEY et al., 2019) and classification, once the currently available protocol on the coronary artery assessment has lots of manual interactions that can be improved by the use of simple computational techniques. The helpful insights from the neural network suggest its wide use in the coronary lesion scenario, mainly by the use of classification and object detection models, that match this study proposal.

This chapter presented the main concepts involved in the proposal for an automatic coronary artery disease system. For this dissertation, we needed a deep understanding of our study

context, with the current angiography protocol and the process of atherosclerosis. This contextualization is necessary to determine how the angiography characteristics impact the model evaluation, as well as the communication with physicians. Then, we explored a few concepts about image processing, which lead to image segmentation techniques, especially instance segmentation. The instance segmentation concept created an opportunity to detect CAD, which also required the understanding of current techniques to apply it.

We discussed the use ML in the healthcare context, which has proven high effectiveness and we could differentiate from the current literature through the correct use of CNN features. In this context, we explored the neural network building concepts, from the image classification up to the image detection architectures. Due to the fact of having small datasets to support this study's goals, we also discussed the use of augmentation techniques to improve the model performance and avoid the most common problems caused by these limitations.

Therefore, we also needed to discuss a common language between the healthcare and the computing area. This becomes more evident in the difference between techniques evaluation in each study area. To this end, in the next chapter, we present a literature review to understand the state of the art to CAD assessment, together with their challenges and research opportunities that guide this Master Thesis.

3 RELATED WORK

This chapter presents the studies related to the coronary artery assessment and base for our proposed model. The literature review is divided into three main parts, methodology, results and discussion, and future directions. The main goal of this literature review is to generate a cross-area unique knowledge base from past few years of research to guide future CAD functional assessment (STAPIĆ et al., 2012). The methodology describes the procedures adopted for the selection of studies. In the discussion, we present an analysis of the selected studies, proposing a taxonomy for CAD in three main aspects. In the last part, we identified gaps and opportunities in the literature, indicating future research directions.

3.1 Methodology

In this section, we describe the protocol adopted for performing the SLR, presenting the research questions, the search and selection strategy, and how we obtain the information to answer our questions.

3.1.1 Study design

The focus of this work is to perform a systematic literature review of recent studies about CAD functional assessment techniques. Previous review studies were analyzed in the context of coronary stenosis functional assessment in terms of applicability of the current recommended techniques in today's context (LOEWE, 2019; JOHNSON et al., 2016; HWANG; LEE; KOO, 2016; KIM et al., 2016; BENTON et al., 2018). This work determines how the gold standards, newest promising technologies and approaches perform in the CAD functional assessment. The scope of the SLR is defined by the steps below:

- Research questions: introduce the research questions;
- Search strategy: present the strategy and the literature explored;
- Data extraction: present the strategy to extract data from studies;
- Article selection: explore the studies selection criteria;
- Techniques discussion: match the research questions with the literature;
- Evaluation of the discussion: explore if techniques fit the research objective;

3.1.2 Research questions

We defined general and specific questions to guide the research (STAPIĆ et al., 2012). The general questions intend to build a holistic view of the research field. We also explored details

Table 1 – Research questions.

Identifier	Issue
	General Questions (GQ)
GQ1	Which are the most common techniques in the coronary functional assessment?
GQ2	What are the most promising techniques in the coronary functional assessment?
GQ3	What are the challenges related to coronary functional assessment?
GQ4	What is the taxonomy for the functional assessment in coronary arteries?
	Specific Questions (SQ)
SQ1	Which are the benefits of each technique applied for coronary functional assessment?
SQ2	What are the comparison baselines applied for these methods?
SQ3	What are the current coronary functional assessment benchmarks?
SQ4	Which are the common ML techniques in a coronary assessment?

Source: Elaborated by the author.

of each exam and technique together with their limitations, defining the specific questions. Both general questions (GQ) and specific questions (SQ) are presented in Table 1.

With these questions, we provide a better understanding of current available techniques for coronary lesion significance assessment. GQ1 concerns the current guidelines and available exams in the healthcare system. GQ2 concerns the techniques not included in the patient treatment mainly due to pending bigger validations. GQ3 concerns the gap of the area comprising all the known challenges each exam face. GQ4 presents a proposed taxonomy to organize the functional assessment characteristics. SQ1 concerns the data evaluation to address gaps. SQ2 analyzes the most common baselines from the medical point of view. SQ3 presents the current evidences for benchmarks in the functional assessment. At last, SQ4 presents the most common ML techniques applied to the coronary artery assessment.

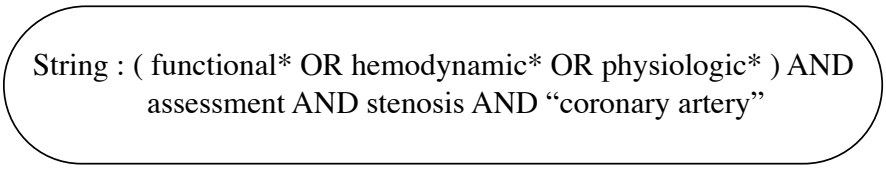
3.1.3 Search strategy

In order to find all relevant functional assessment techniques currently under study, we defined the search string presented in Figure 16. Our search string covers the relevant terms that associate hemodynamic and functional aspects from recent studies, which evaluated stenosis severity while also covering both outcome and correlation with gold standard exams.

Based on recently published SLR studies available in the literature, the search string was applied in known academic and scientific databases. The search query was created with the most relevant terms regarding coronary significance from medical trials, reviews and case reports, but also reaching promising basic science technical approaches based on mathematical models.

The article population was limited in the past 5 years including previous reviews, which inherit a larger period of time. To reach both medical-specific trials and promising basic science technical approaches, this research was performed in 8 different databases: ACM, IEEE, Elsevier, Springer, PubMed, Cochrane, Science Direct, and Google Scholar.

Figure 16 – Search string used in this work. The string covers all the relevant terms and combinations to reach the functional assessment relevant data for this meta-analysis.



Source: Elaborated by the author.

Quotation marks for coronary artery indicate that an exact match is sought to avoid search engines to include variations and non-coronary studies. Another important aspect of the search string is the use of asterisks to reach any word with the same stem.

The works selected from all databases were then handled using Mendeley[1] reference manager.

3.1.4 Article selection

This work includes the most recent studies which present analysis from the current gold standard and clinical available methods as well the most recent emerging techniques applied to the coronary assessment to increase the value of the cross-area knowledge. Inclusion criteria were: (i) articles written in English, (ii) articles addressing coronary artery functional assessment, (iii) articles evaluating methods performance, based on the full-text examination, and (iv) tests on big population for medical studies up to synthetically small databases for basic science studies.

This initial selection was composed of three main exclusion criteria to determine research relevance (STAPIĆ et al., 2012):

- Exclusion criteria 1: studies that do not address “coronary artery” and related acronyms (population criterion I).
- Exclusion criteria 2: studies that do not address “significance”, “functional”, and related acronyms regarding coronary narrowing severity (population criterion II).
- Exclusion criteria 3: studies which address secondary factors that impact the diagnosis of coronary artery narrowing disease such as “renal”, “pulmonary”, “stent”, “cancer”, and “bifurcation” (population criterion III).

The articles were filtered by title, which addresses retrospective methods comparison, method proposal, and validation/trials. Afterward, we performed body reading focusing on extracting

[1] https://www.mendeley.com

the relevant classification categories previously mentioned, thus determining the final list of studies.

3.1.5 Data extraction

After screening, the relevant articles were selected and read in full to confirm eligibility as recommended in the literature (PETTICREW; ROBERTS, 2008). This reading step separated the studies according to their relevance, focusing on four main categories: (i) comparison of methodologies performance according to outcomes; (ii) proposal of new techniques and the motivation for them; (iii) validation of existing methods via clinical trial; and (iv) literature reviews. Then, the selected studies were read and classified according to several characteristics to answer the study questions (PETTICREW; ROBERTS, 2008).

3.2 Results and discussion

In this section, we discuss the results based on the search strategy from the previous section.

3.2.1 Article selection

The search for articles using our string was finished on 31st October 2020. Papers published later than that were left aside from this survey. Our search process is detailed in Figure 17, which presents the whole initial study population as well as the exclusions and filtering steps. In total, our search string was able to find 1993 articles from the 8 selected databases. The selected articles were then filtered based on the three exclusion criteria, resulting in 303 works.

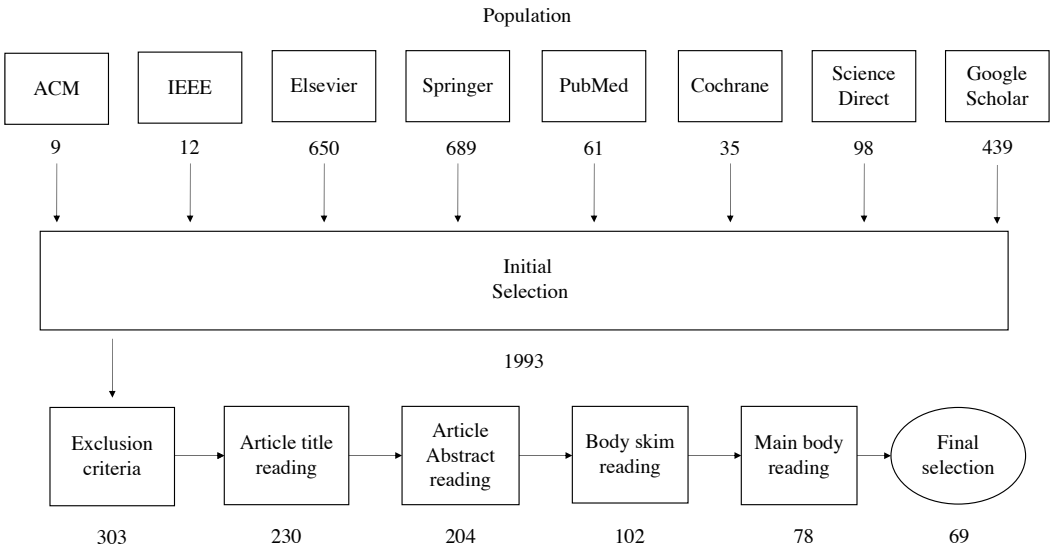
The exclusion step intended to remove secondary factors that imply in the exam diagnosis, such as the vessel bifurcation analysis and other associated diseases such as cancer and pulmonary limitations. Then, we filtered articles by title and by abstract reading. Afterward, we performed a body skim reading to reach the classification items previously described, and, at last, we performed full-text reading for those where skim-reading was not enough. In the end, this process resulted in 69 articles as the final population.

3.2.2 Data extraction

In this section, we answer the general and specific questions presented in Section 3.1.2, as well as discuss the state-of-the-art approaches for coronary artery functional assessment.

To enhance presentation, we cluster the works based on the classification criteria. Precisely 27 (twenty-seven) classification items were defined. From the most important clinical relevance, this work highlights the invasiveness, in which the studies were assessed according to the evaluated technique invasiveness degree. The baseline for comparison was identified together with its

Figure 17 – Research population selection process. From a diverse group of databases, this work connects science basic studies with well-known trials ans exam validation.



Source: Elaborated by the author.

correlation value. Regardless of the main use of FFR, some of the recent studies were compared to the patient outcome, which is the metric to evaluate the efficiency of an assessment method due to its long-time benefit (RAMASAMY et al., 2020). Trial studies usually evaluate stable CADs due to their complexity and occasionally artery at ACS are also functionally investigated. The measurement classification is in most cases at rest, due to the efforts of non-invasive techniques with lower artery stress. Nevertheless, lots of them still examine the artery at hyperemia, depending on patient conditions (RAMASAMY et al., 2020).

Metrics evaluated are important to match the difference between clinical and scientific points of view together with their performance. Population size was also classified together with the use of AI, Computational fluid dynamics (CFD) and image-based techniques, etc.

GQ1 – Which are the most common techniques in the coronary functional assessment?

The minority of patients who undergo a coronary angiography undergo a clinically invasive stress test, such as FFR or instantaneous wave-free ratio (iFR) (ADJEDJ et al., 2015), evidencing CCA as the main source for CAD diagnosis and most common exam applied. The use of imaging techniques to assess coronary diseases has been widely used in recent years (46% of studies). Nevertheless, several improvements are expected to arise in the next years, with more randomized and broad studies (TANIGAKI et al., 2019; PYXARAS et al., 2018). More than achievable indexes of coronary impact, the functional assessment via anatomical characteristics has been largely discussed (ADJEDJ et al., 2015), mainly due to the reference of angiography, which provides poor correlation with functional assessment (ADJEDJ et al., 2015; BENTON et al., 2018; PYXARAS et al., 2018; RAMASAMY et al., 2020). Especially in moderate stenosis, the luminal narrowing correlates poorly with resultant ischemia (BENTON et al., 2018;

PYXARAS et al., 2018; RAMASAMY et al., 2020). In general, this discordance between visual and functional assessment remains uncertain (KIM et al., 2016; PYXARAS et al., 2018; BAIBHAV et al., 2018) as currently there is no simple relationship between the visual parameters, the stenosis hemodynamic significance, and the plaque vulnerability (AGUJETAS et al., 2018). For this reason, there are doubts if angiography will remain indispensable in patients with known or suspected CAD (ADJEDJ et al., 2015).

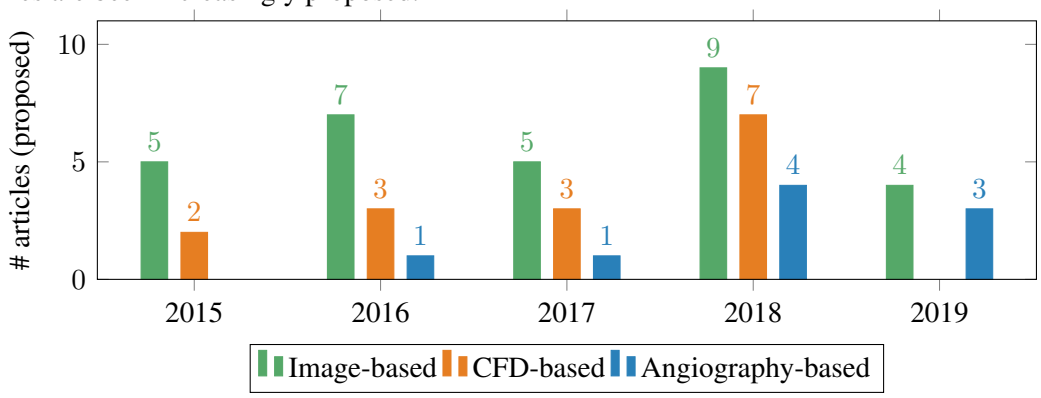
The gold standard is considered essential to classify a narrowing impact and to determine a CAD diagnosis, reinforced by the AHA (PARK; LERMAN; HERRMANN, 2017) and ESC guidelines (TEBALDI et al., 2018). The AHA recommends the use of FFR to assess the narrowing significance, which was first angiographically determined as intermediate (50–70%). This process guides to revascularization in patients with stable ischemic heart disease when FFR ≤ 0.80 (PARK; LERMAN; HERRMANN, 2017). ESC recommends FFR for the coronary disease assessment showing at least one coronary lesion with 50% to 90% DS angiographically assessed (TEBALDI et al., 2018).

Recently, the British Cardiovascular Intervention Society audited the use of FFR and detected that it was used in only 8% of the PCI performed in the UK (RAMASAMY et al., 2020). Another analysis estimates that FFR is used in fewer than 10% of the cases in many parts of the world (De Waard et al., 2018) due to several restrictions. In terms of patient conditions, we can highlight the chest discomfort caused by the procedure and that FFR is not recommended in patients with severe asthma, hypotension, or atrioventricular conduction disturbances (YAZAKI et al., 2017; De Waard et al., 2018). From the applicability point of view, FFR has a high cost for both training and execution (BHATT, 2018; RAMASAMY et al., 2020).

In fact, every technique must prove to be good to guide towards the right diagnosis. Even some studies have shown poor correlation of $r^2 = 0.58$ between FFR and the outcome (MIN et al., 2015). Other works identified limitations in the initial test with small population (BENTON et al., 2018; COENEN et al., 2017; KIM et al., 2016; LEE et al., 2020; Di Serafino et al., 2016), mainly for left main stem (LMS) lesion (RAMASAMY et al., 2020; LEE et al., 2020). In general, all the methods based uniquely on visual assessment suffer from low-quality images, missing protocols for data acquisition, and the lack of labeled datasets. Also important, these methods usually create simplifications to reduce the simulation complexity, such as CFD (AGUJETAS et al., 2018), which tends to deteriorate the simulation of natural blood flow behavior. CT MPI compared to FFR, for example, correlates $r = 0.61$ (COENEN et al., 2017) and the coronary computed tomography angiography (CCTA) compared with FFR with $r = 0.61$ (KIM et al., 2016).

Intravascular ultrasound (IVUS)-based algorithms did not present good results in side diffuse and tandem branches lesions (LEE et al., 2020). The ADDED method also correlates poorly with FFR and with iFR, with $r = 0.59$ and $r = 0.61$, respectively (Di Serafino et al., 2016). IVUS quantifies LMS lesion severity, rather than qualitative information from angiography. A MLA for LMS $> 6mm^2$ has been shown to be a cut-off value for deferring revascular-

Figure 18 – Common approaches either proposed in recent studies. As seen, image-based and CFD-based approaches are been increasingly proposed.



Source: Elaborated by the author.

ization (BAYDOUN et al., 2019).

Optical Coherence Tomography (OCT) performs better to identify non-left main coronary artery functional significance (RAMASAMY et al., 2020). These evidences suggest that OCT is superior to IVUS for non-left main stem diseases to assess lesion morphology (AGUJETAS et al., 2018), but still poor for functional assessment (RAMASAMY et al., 2020). Both IVUS and OCT provide useful information about coronary artery morphology, burden plaque, and lumen dimension (RAMASAMY et al., 2020). Unfortunately, some of the emerging techniques are not compared against FFR (LI et al., 2020; KWAN et al., 2020; AL'AREF et al., 2020; BAUMANN et al., 2020).

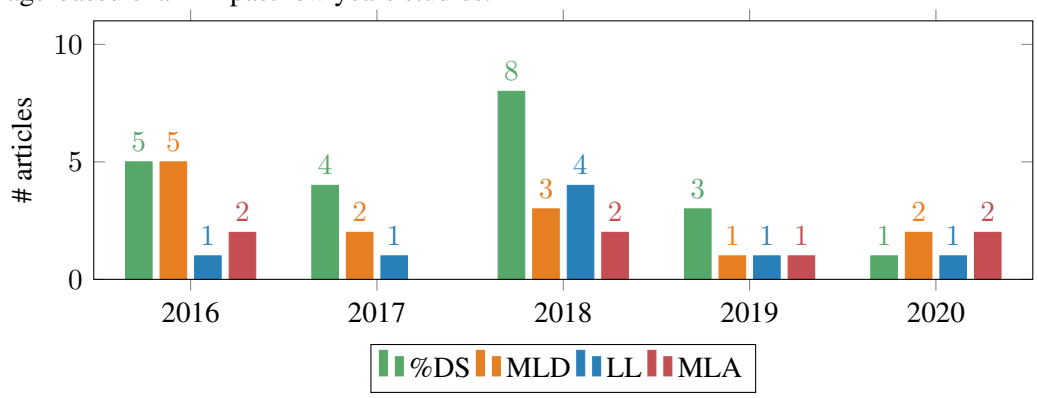
FFR_{CT} works better in low-to-intermediate lesions using CFD simulation. FFR_{CT} reached good correlation with FFR, but it produces lots of false positives (BENTON et al., 2018). DILEMMA results have a moderate correlation with FFR, which could defer pressure wire assessment in about 50% of the cases. For DILEMMA scores ≤ 2 and ≥ 9, a good performance is attainable (MICHAIL et al., 2019). Considering a cut-off value of 0.80 for the FFR, the ADDED index showed a higher area under the curve (AUC; of 0.94 [0.90–0.99], with $p < 0.001$) as compared to MLD alone (of 0.76 [95% CI: 0.70 to 0.86], with $p < 0.001$) (Di Serafino et al., 2016).

GQ2 – What are the most promising techniques in the coronary functional assessment?

To determine the most promising techniques, we evaluated each promising technique grouping them according to the basic concepts, such as the use of image, blood flow simulations and the need of an angiographic previous analysis. Aside the commonly adopted techniques for CAD assessment, several other approaches have been proposed in the literature. These techniques can be basically grouped in image-based, CFD-based and angiography-based. Figure 18 presents the frequency with which each of technique principles were investigated in the past few years.

Recently, the use of image-based data extraction, CFD techniques for blood flow simulation,

Figure 19 – Use of anatomical metrics. This figure presents the most common anatomical data extracted from image-based exam in past few years studies.



Source: Elaborated by the author.

and angiographic geometrical information are commonly studied. These studies suggest that MLD (MICHAIL et al., 2019; Di Serafino et al., 2016; PYXARAS et al., 2018), lesion length (LL) (MICHAIL et al., 2019; Di Serafino et al., 2016; PYXARAS et al., 2018), and percentage of diameter stenosis (%DS) (PYXARAS et al., 2018; HU et al., 2020) are essential to determine CAD significance. In these studies, the anatomical data presented good performance to predict a severe stenosis, reducing the need for pressure wire usage. The frequency of these studies over the last five years is presented in Figure 19.

From 2018 onwards, some other techniques exploring the use of AI emerged, and we also evaluate their promising results. Other image-based exams such as IVUS and OCT have also been used for daily clinical analysis, although they present worse performance on classifying significant stenosis. IVUS provides important lesion characteristics, such as vessel size, lesion length (LL), and the extent of calcification (LEE et al., 2020), which can be helpful for the holistic evaluation of CAD. Due to these important characteristics, Lee et al. (LEE et al., 2020) proposed six IVUS-based supervised machine learning algorithm to identify lesions with an FFR ≤ 0.80. These algorithms are trained with six clinical variables, namely age, gender, body surface area, vessel type, involved segment, and involvement of the proximal left anterior descending artery. Since 2018, several other ML techniques also emerged to improve the diagnosis performance (LEE et al., 2020; LOEWE, 2019; TRAN; NGUYEN; LEE, 2018; DEY et al., 2018). These techniques usually focus on the multivessel CAD evaluation (LEE et al., 2020; LOEWE, 2019; TRAN; NGUYEN; LEE, 2018) via metrics such as LL (LEE et al., 2020; TRAN; NGUYEN; LEE, 2018; DEY et al., 2018) and %DS (TRAN; NGUYEN; LEE, 2018; DEY et al., 2018), which showed good correlation with FFR (TRAN; NGUYEN; LEE, 2018).

The major part of the ML-based techniques proposed in recent studies follows the image-based approach (PANG et al., 2021; JAFARIAN et al., 2020; CRUZ-ACEVES; HERNANDEZ-AGUIRRE; VALDEZ, 2016; HU et al., 2020; LONCARIC et al., 2020; EBERHARD et al., 2020; Kaplan Berkaya; Ak Sivrikoz; GUNAL, 2020; LI et al., 2020) using (deep) neural networks, which typically require large databases to be properly trained. This need for a huge

amount of data associated with the lack of openly available datasets has also led to synthetically generated databases (EBERHARD et al., 2020). The plaque characteristics are mainly used in ML-based techniques and are challenging to reproduce (HU et al., 2020; TRAN; NGUYEN; LEE, 2018; AL'AREF et al., 2020). For this reason, we noticed the increasing use of transfer learning approaches to leverage the performance of the network (LONCARIC et al., 2020; Kaplan Berkaya; Ak Sivrikoz; GUNAL, 2020) or creating a complete sequence of steps to stenosis detection (PANG et al., 2021). Transfer learning is the process of improving the target predictive function by using the related information from an already trained function, improving a small part of the network parameters, and reusing the initial layers, which represent the most generic features (WEISS; KHOSHGOFTAAR; WANG, 2006).

In the image-based context, region of interest (ROI) (LI et al., 2020), features segmentation (LI et al., 2016; CRUZ-ACEVES; HERNANDEZ-AGUIRRE; VALDEZ, 2016; Kaplan Berkaya; Ak Sivrikoz; GUNAL, 2020) and data reduction (BENHAR; IDRI; L Fernández-Alemán, 2020) are the most regarded elements. These studies suggest that an image is not enough to determine the lesion significance yet (LI et al., 2016; Kaplan Berkaya; Ak Sivrikoz; GUNAL, 2020; CRUZ-ACEVES; HERNANDEZ-AGUIRRE; VALDEZ, 2016; BENHAR; IDRI; L Fernández-Alemán, 2020; LI et al., 2020). The importance of non-image-based characteristics to determine the lesion severity, such as patient conditions, can provide a better view of the patient, not limited to the visual plaque characteristics. Current evidences suggest the need for improvements in these methods to reach better correlation with FFR in terms of outcome (LONCARIC et al., 2020; Kaplan Berkaya; Ak Sivrikoz; GUNAL, 2020; BENHAR; IDRI; L Fernández-Alemán, 2020; LI et al., 2020), but showing good performance even in low contrast regions (LI et al., 2016).

In a different direction, but also aiming to provide better insights in the functional assessment, blood flow model approaches based on the analogy to hydraulic and electrical systems, also known as 0D models, emerged to simulate blood behavior reconstructed from CCTA (UUS et al., 2015). 0D models consider the resistance as viscous dissipation, the capacitors as volume compliance, and the inductors as the blood inertia (UUS et al., 2015). Recent studies present a valuable analysis of 0D and 3D models for coronary significance assessment (GREEN et al., 2018), showing the hydraulic impact of a stenosis in the coronary segment (TRAN; NGUYEN; LEE, 2018; UUS et al., 2015). These studies employ 0D models for simulation and achieved similar results to FFR. For 0D models, hyperemia is simulated by fluid parameters. In such panorama, ML-based algorithm obtained results close to an expert assessment (Kaplan Berkaya; Ak Sivrikoz; GUNAL, 2020; TRAN; NGUYEN; LEE, 2018; HU et al., 2020; GREEN et al., 2018) and some better results than CCTA alone (HU et al., 2020).

GQ3 – What are the challenges related to coronary functional assessment?

In order to answer our third general question, we start by summarizing the advantages, disadvantages, and gaps of the state of the art approaches discussed in the previous answer. These results are present in Table 2. As already discussed, the visual assessment alone has

proven insufficient to determine the hemodynamic significant epicardial stenosis (CONTE et al., 2020; KO et al., 2015). The %DS metric is commonly associated with a good correlation with FFR diagnosis (32% of studies). Nevertheless, there are evidences that some anatomical indices such as %AS, %DS and MLD have weaker correlations ($\rho = -0.26, P = 0.001; \rho = -0.37, P < 0.0001; \rho = 0.52, P < 0.0001$) with FFR (YAZAKI et al., 2017). These evidences do not mean directly that anatomical information are not correlated with outcome, but that metrics are essential to plaque characterization. There are also evidences that FFR is not able to determine lesion anatomy and plaque, which represents a reason of discordance (RAMASAMY et al., 2020). In fact, they are complementary, requiring a holistic lesion analysis combining approaches to assess both anatomical and functional impacts (ADJEDJ et al., 2015).

The QFR is currently considered the state of the art for coronary functional assessment due to its less invasive approach, improvements in the treatment of blood flow reduction, and high agreement with FFR (PYXARAS et al., 2018). QFR inherits the FFR assumptions of proportional linear relationship between coronary perfusion pressure, flow, and coronary microvascular resistance (PARK; LERMAN; HERRMANN, 2017).

FFR and QFR require the angiography as the reference exam for evaluation, since CAD evaluation always starts with an angiographic assessment. Although FFR has been widely used, some studies that evaluate the long-term results of FFR-assessed severe stenosis have indicated the importance of including complete coronary tree hemodynamic assessment. This inclusion is due to the fact that FFR-only do not reach the correct diagnosis in patients with stable angina (XAPLANTERIS et al., 2018). In-procedure QFR is clinically feasible and is superior to angiographic assessment alone (WESTRA et al., 2018). QFR bears the potential to expand the adoption of physiological lesion assessment (WESTRA et al., 2018). QFR also has good correlation between with FFR and FFR_{CT}. Considering the ≤ 0.8 FFR cut-off for severe lesions, QFR provides higher diagnosis accuracy (TAKAHASHI et al., 2019; TANIGAKI et al., 2019).

FFR use in LMS disease is less universally accepted due to a known method limitation. Recently, the application of IVUS together with FFR seems to have worked well for LMS lesion classification (SHAO; HOU; CHIU, 2014; PARK; LERMAN; HERRMANN, 2017), suggesting a promising future applicability of FFR on this segment. FFR_{CT}, is then limited by image quality (i.e., artifacts from calcification), patient-specific factors (such as anemia, which causes reduced viscosity that may impact the calculation of FFR), and processing time for computation (PARK; LERMAN; HERRMANN, 2017).

FFR_{CT} is the FFR measurement from computer tomography (CT) exams reconstructed with CFD techniques. FFR_{CT} is also based on volumetric finite element mesh, which simulates the blood flow incorporating the fluid properties of blood (CONTE et al., 2020). FFR_{CT} has improved upon stenosis evaluation in terms of prediction of functional significance compared to anatomical evaluation alone (CONTE et al., 2020).

From the image processing perspective, QFR spent time on the calculation procedure did not include the time for angiographic acquisition despite producing good results (WESTRA et al.,

Table 2 – Exams and techniques comparison evaluated in terms of advantages, disadvantages and gaps evaluation I.

Exam	Advantages	Disadvantages	Gaps
FFR	good outcome correlation, pressure drop	contraindications, induced hyperemia, high cost, invasive	poor results on LMS lesions
iFR	good FFR correlation, no induced hyperemia, pressure drop	cutoff values vary, several limitations, invasive	no impact on outcome, poor on LMS lesions
IVUS	image-based, good on LMS lesions, non-invasive	limited to plaque characterization	FFR improve results, require bigger validation
FFR_{CT}	low-to-intermediate lesions, non-invasive, image-based	low-quality images, lots of false positives	patient-specific factors, require bigger validation
QFR	non-invasive, image-based, excellent FFR correlation	time for processing	linear pressure, local segment assessment, under validation
CMR	exclude severe lesion, non-invasive, image-based	limited availability, contraindications	require bigger validation
0D	non-invasive, possible benchmark, close system, cover coronary tree	simplified blood flow	require bigger validation
DILEMMA	good predictive FFR, non-invasive, image-based, score 0-10	few variables	gray zone (3-8)
ADDED	good FFR correlation, non-invasive, image-based, score 0-10	good for FFR > 0.8	require bigger validation
OCT	cover coronary tree, good for non-LMS, non-invasive, image-based	poor FFR correlation	require bigger validation, angiography-based

Source: Elaborated by the author.

2018). QFR results are not tested with FFR for all the vessel segments, which limits the holistic artery quantitative evaluation (WESTRA et al., 2018). Considering the whole QFR solution, the proposed method does not consider the data transfer time from angiographic equipment to the QFR workstation (WESTRA et al., 2018), being open for enhancements and optimization.

iFR presents good results in comparison with FFR but has several limitations since the cutoff values vary too much among the studies (PARK; LERMAN; HERRMANN, 2017). Further studies are required before iFR can be adopted to determine revascularization in patients with intermediate LMS lesions (BAYDOUN et al., 2019). iFR has not demonstrated yet its impact on clinical outcomes (ADJEDJ et al., 2015).

Cardiovascular magnetic resonance (CMR) imaging perfusion imaging has also several limitations that restrict its utility in the assessment of myocardial ischemia. CMR limitations include the low availability of MRI scanners, the patient claustrophobia, poor gating, motion artifacts, and contraindications, such as pacemakers or defibrillators (KO et al., 2015).

From the holistic CAD evaluation, the DILEMMA score varies from 0 to 10, where 0 represents the lowest significance of the stenosis and 10 represents the highest level of significance (MICHAIL et al., 2019). According to the results, values of ≤ 2 demonstrated a negative predictive value to identify lesions with FFR > 0.8 and values of ≥ 9 demonstrated excellent positive predictive value for lesions with FFR ≤ 0.8 (MICHAIL et al., 2019), suggesting this approach as a good example of score based on visual lesion characteristics together with patient clinical data (MICHAIL et al., 2019).

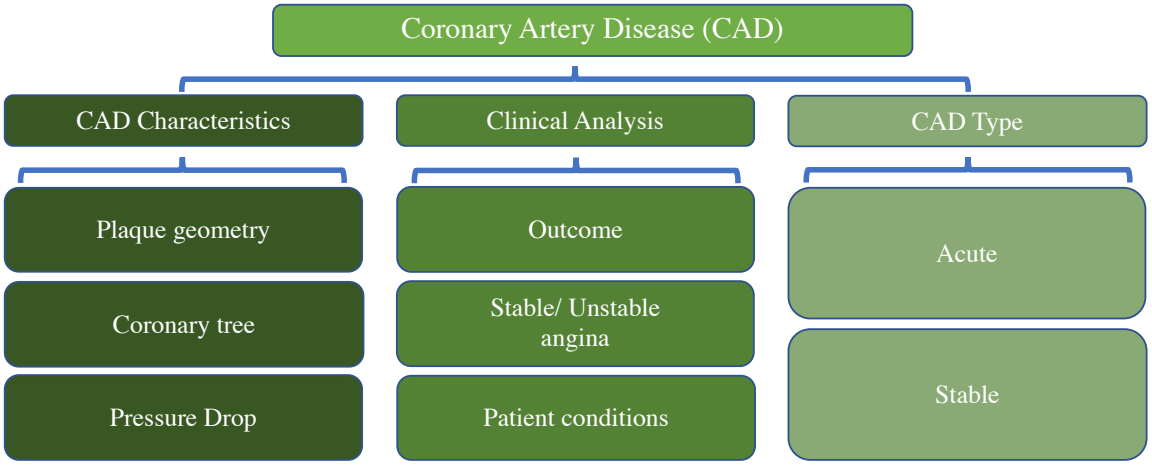
In general, IVUS and OCT also present good results on predicting FFR. Some studies compared OCT prediction with FFR obtaining a values of 73% with 61 stenosis, 85% with 62 stenosis and 80% with 55 stenosis (PYXARAS et al., 2018). These findings suggest a weak correlation between OCT and FFR (PYXARAS et al., 2018). In the same review, they have presented a study with 42 patients using IVUS to predict severe stenosis assessed by FFR achieving a accuracy of 88% (PYXARAS et al., 2018). Even though IVUS seems promising to predict FFR, it still requires larger trials to validate its applicability.

GQ4 – What is the taxonomy for the functional assessment in coronary arteries?

To obtain a better comprehension of the functional assessment of coronary arteries and the lack of an existing taxonomy that connects computing with healthcare, this work proposes a new taxonomy. The proposed taxonomy organizes the concepts and the connections between the exams and techniques to assess the CAD severity. As a result, we present three distinct taxonomies, which consider the following perspectives: CAD (Figure 20), technique (Figure 21), and interaction characteristics (Figure 22).

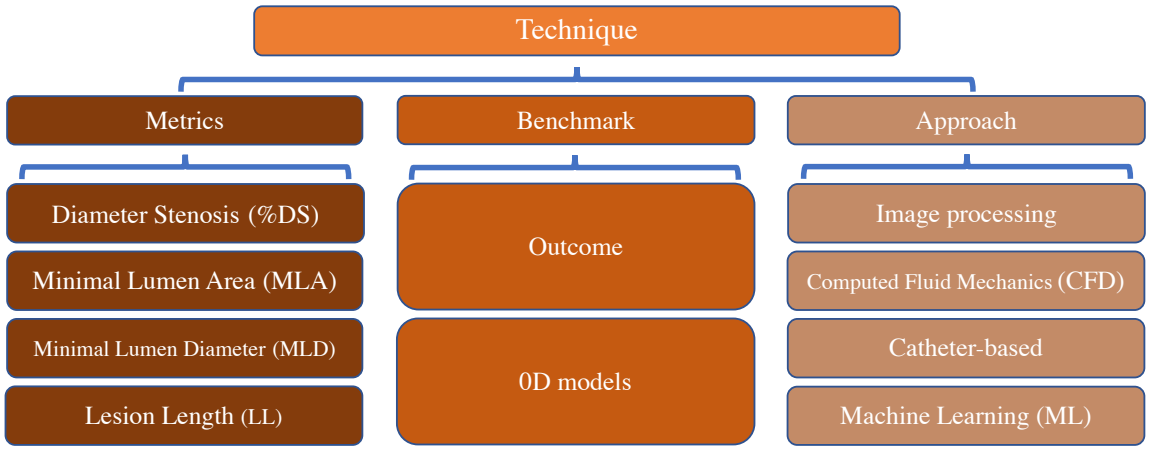
The first taxonomy presented in Figure 20, CAD, have three nodes, characteristics, clinical analysis, and lesion type. The first node, CAD characteristics present the measurable lesion aspects relevant in the currently available exams. The second node, clinical analysis, presents the clinical lesion important evidence that determines the diagnosis. The last one, type, presents the CAD condition, which can be acute or stable.

Figure 20 – Blood flow functional assessment taxonomy from the disease perspective. It summarizes disease main characteristics.



Source: Elaborated by the author.

Figure 21 – Blood flow functional assessment taxonomy from the performance evaluation perspective. It presents the metrics for performance evaluation.

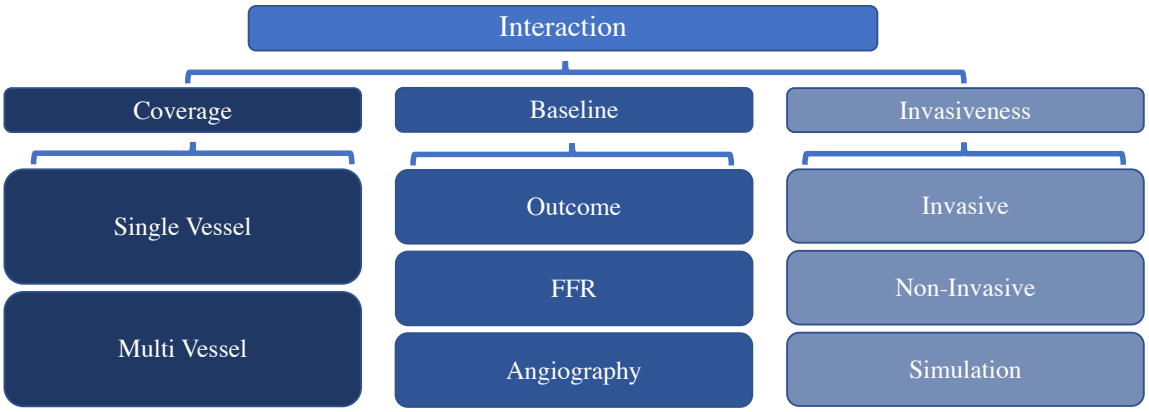


Source: Elaborated by the author.

The second taxonomy presented in Figure 21, Technique, describes the types of assessments as well as their characteristics. This taxonomy intends to show the technical details of promising techniques and current exams in terms of numerical evaluation. This taxonomy is composed of three nodes, metrics, benchmark, and approach. The first node, Metrics, presents the most relevant metrics studied in the current literature. The second node presents the available benchmarks for technique and exams determination, evidencing a promising future common benchmark on 0D models. The third node, Approach presents the technique principle more frequently in the literature.

The third and last taxonomy presented in Figure 22, Interaction, categorizes the exams and techniques characteristics in terms of coverage, baseline, and invasiveness degree. The first

Figure 22 – Blood flow functional assessment taxonomy from the operational perspective. It summarizes how the techniques interact with the patient.



Source: Elaborated by the author.

node, the coverage presents the two possible coronary tree analyses. Some methods evaluate the culprit lesion alone to determine the significance. On the other hand, other methods focus on multi-vessel assessment due to the importance of complete coronary assessment. The second node presents the available baselines for CAD assessment, which can be the outcome, FFR or the angiography. The third node, invasiveness presents that exams and techniques can be invasive, non-invasive, or even based on blood flow simulation by real objects simplification.

The three taxonomies intend to evaluate the CAD assessment from a different perspective. The group elements from both exams and emerging techniques provide a unique base for comparison. The taxonomies cover all the relevant aspects of CAD in terms of application, evaluation, and guide to diagnosis.

SQ1 – Which are the benefits of each technique applied for the coronary functional assessment?

CAD assessment benefits can be measured by two main aspects, the correlation with the gold standard and the improvement in kinds of lesions for which the gold standard is not recommended. The correlation with the gold standard is based on a numerical evaluation, while the improvements are measured in terms of outcome evaluation.

The FAME trial demonstrated that routine use of FFR during PCI in patients with multivessel CAD has proven long-lived and cost-saving (AGUJETAS et al., 2018; PARK; LERMAN; HERRMANN, 2017; PYXARAS et al., 2018). FAME II also suggests that the CAD diseases cannot be fully explained by the stenosis alone (HWANG; LEE; KOO, 2016). Despite the use of FFR and the already mentioned limitations of FFR use, we evaluated the correlation factor of further techniques, which are usually compared with FFR.

Table 3 presents both quantitative and qualitative evaluations of all techniques and exams discussed before in terms of correlation with FFR. Items are disposed of according to the higher general correlation down to high correlations with very specific cases and the moderate correla-

Table 3 – Exams with higher correlation factors with FFR.

Method	Quantitative FFR correlation	Qualitative FFR correlation
iFR	$0.80 - 0.94$	Strong
IVUS-based	$0.70 - 0.92$	Strong
FFR_{ML}	$0.73 - 0.8$	Strong
QFR	$0.69 - 0.83$	Moderate/Strong
FFR_{CT}	$0.61 - 0.82$	Moderate/Strong
ADDED	0.96	Strong (FFR > 0.8)
OCT	$0.73 - 0.80$	Strong (FFR > 0.8)
CMR	-	Strong (FFR > 0.75)
DILEMMA	-	Strong (FFR < 0.8)
0D	-	Moderate

Source: Elaborated by the author.

tions. An important remark is that all the reported data came from the literature and the studies were evaluated with different samples.

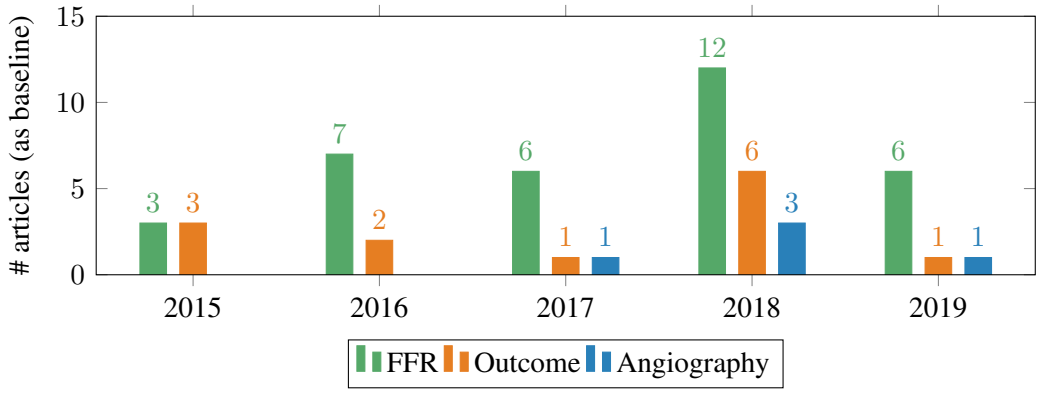
iFR predicts the FFR value with the same accuracy as the mere resting Pd/Pa ratio (ADJEDJ et al., 2015; PYXARAS et al., 2018). iFR is a good option in countries where Adenosine is expensive and/or other vasodilators are not available (ADJEDJ et al., 2015; De Waard et al., 2018). iFR presents a good correlation between 0.8 and 0.94 for severe or low-grade stenosis (PYXARAS et al., 2018; BENTON et al., 2018; FUKUOKA et al., 2020; HWANG; LEE; KOO, 2016; LEONE et al., 2018), but a poor correlation in moderate cases (FUKUOKA et al., 2020; PYXARAS et al., 2018; BENTON et al., 2018).

From the image-based techniques point of view, CCTA, FFR_{CT}, and QFR presented the best results. FFR_{CT} using CFD from CCTA, also known as Virtual Functional Assessment Index (vFAI), has a correlation with FFR between 0.61 and 0.82 (SIOGKAS et al., 2017; TAKAHASHI et al., 2019; COENEN et al., 2017; TANIGAKI et al., 2019). FFR_{ML} method presented a correlation value of 0.73 with FFR (TRAN; NGUYEN; LEE, 2018).

QFR results suggest that it exceeds the limits of FFR (CESARO et al., 2018). QFR has better results than iFR, presenting excellent correlation with FFR (CESARO et al., 2018; TANIGAKI et al., 2019) with good correlation with FFR_{CT} (TANIGAKI et al., 2019). From several studies, the correlation between QFR and FFR varies between 0.69 and 0.83 (CESARO et al., 2018; WESTRA et al., 2018; YAZAKI et al., 2017; TIES et al., 2018; TANIGAKI et al., 2019). Considering the level of analysis, QFR had a correlation with FFR of 0.83 per-vessel and 0.8 per-patient (WESTRA et al., 2018).

IVUS-based techniques, such as 3D-based IVUS (fFFR), present a correlation value of 0.92 with FFR (WANG et al., 2018). IVUS-based approach showed a correlation factor between 0.8 and 0.92 with FFR. Another approach using supervised machine learning to predict FFR ≤ 0.80 based on the IVUS images presented a correlation between 0.75 and 0.8 (LEE et al., 2020). This technique presented results between 82% and 86% (using L2-penalized logistic regression), 80% and 85% (using artificial neural network), 83% and 87% (using random forest), 83% and

Figure 23 – Baselines commonly used in recent studies. As seen, FFR has been steadily used as the baseline approach.



Source: Elaborated by the author.

87% (using AdaBoost), 81% and 85% (using CatBoost), and 81% and 85% (using support vector machines) (LEE et al., 2020). The ML-based techniques achieved good correlation with FFR (LONCARIC et al., 2020; KWAN et al., 2020; AL'AREF et al., 2020; Kaplan Berkaya; Ak Sivrikoz; GUNAL, 2020) and are considered helpful for severe lesion exclusion (KAWASAKI et al., 2020; BAUMANN et al., 2020). CCTA is commonly associated with the use of logistic regression (HU et al., 2020; AL'AREF et al., 2020). Nevertheless, CCTA was also associated to predictive method (AL'AREF et al., 2020) and random forest (KAWASAKI et al., 2020) in terms of severe lesions identification.

CMR is good for detecting CAD by FFR ≤ 0.8 (KOTECHA et al., 2019). Myocardial perfusion are highly correlated with FFR using a cutoff value of 0.75 to exclude severe lesion (KO et al., 2015) as well as the combination of CCTA and CTP (G. et al., 2018). On the other hand, the FMM/MLD discovered vessels with FFR < 0.80 and vessels with FFR > 0.80 better than anatomical stenosis alone (KIM et al., 2016). FMM and MLD together improved the %DS accuracy and promise to be the missing link between the anatomical and physiological discordance (KIM et al., 2016).

SQ2 – What are the comparison baselines applied for these methods?

The major part of the considered studies used FFR as the baseline. This reason is due to its good results improving people outcome (57% of studies). Besides the use of FFR, several other studies compared the methods results with the patients outcome (19% of studies), which shows the tendency to evaluate the clinical outcome over the easier technique proof of concept (PYXARAS et al., 2018). Finally, the comparison against visual assessment, mainly via angiography (6% of studies), is often present in the daily clinical job (PYXARAS et al., 2018). Even though not typically recommended, the visual assessment is still a routine for CAD functional assessment due its presence in the worldwide adopted guidelines and ease of use (PYXARAS et al., 2018).

Figure 23 presents the frequency with which each of the above techniques have been inves-

tigated in the selected articles, limited for the range of 2015 and 2019 due to availability of data. As seen, the FFR is the most common baseline for techniques comparison due its relevance in the patients outcome improvement. Nevertheless, it is important to evidence the presence of outcome as the second major baseline comparison in the past few years. It suggests the need of medical follow-up besides the numerical comparison with FFR cut-offs. In the chart is also possible to see the recent use of angiography for comparison, which can be associated with the restrictions to medical records access.

SQ3 – What are the current coronary functional assessment benchmarks?

Since there are some discordance about the stable angina treatment already explored, the current benchmark for functional assessment are the long-term outcome. The long-term outcome evaluation range from mortality to morbidity, quality of life, and resource use (CARSON et al., 2019; NUDI et al., 2020).

The recent techniques which measure and simulate the blood flow behavior aim to provide a quantitative evaluation to replace the unique outcome evaluation. To provide an initial perspective of benchmark, 4 models were evaluated including 0D and 3D modeling and provided valuable insights about the use of 0D models correlating it with the FFR results (CARSON et al., 2019). Results suggest that prediction of FFR via numerical models still need further investigation (CARSON et al., 2019), but will probably be the future benchmark for CAD functional assessment.

SQ4 – Which are the common ML techniques in a coronary assessment?

In the past few years, several studies emerged implementing ML techniques for coronary artery assessment, such as deep learning (WONG; FORTINO; ABBOTT, 2020; HAN et al., 2020; SEIFERT et al., 2020; GUPTA et al., 2019) and supervised learning (BRAUN et al., 2020; HAE et al., 2018; KWAN et al., 2020; SHU et al., 2020) including regression algorithms (AAKKARA; AAISUEB; AEELANUPAB, 2020; KWAN et al., 2020) and specifically neural networks (HE et al., 2020; WANG, 2020; TESCHE; GRAY, 2020; OVALLE-MAGALLANES et al., 2020; PODGORSAK et al., 2020; ZREIK et al., 2018, 2019).

The major part of the studies were guided with small dataset, with less than 500 vessels (HAN et al., 2020; KWAN et al., 2020; AAKKARA; AAISUEB; AEELANUPAB, 2020; OVALLE-MAGALLANES et al., 2020; PODGORSAK et al., 2020; SHU et al., 2020; ZREIK et al., 2018, 2019). Nevertheless, some other studies achieved their results with more than 500 vessels (BRAUN et al., 2020; HE et al., 2020; GUPTA et al., 2019; HAE et al., 2018). The use of ML techniques is highly associated to images (WONG; FORTINO; ABBOTT, 2020; HAN et al., 2020; HE et al., 2020; TESCHE; GRAY, 2020; PODGORSAK et al., 2020; ZREIK et al., 2018, 2019; GUPTA et al., 2019; HAE et al., 2018), while some of them focused only in patient numerical lesion data (SEIFERT et al., 2020; KWAN et al., 2020; BRAUN et al., 2020; AAKKARA; AAISUEB; AEELANUPAB, 2020; SHU et al., 2020).

The most common ML technique is supervised learning, used in many approaches for both

images and numerical data. The neural network approach has focused on CCTA (HAN et al., 2020; ZREIK et al., 2018, 2019; HAE et al., 2018), patient characteristics (BRAUN et al., 2020; PODGORSAK et al., 2020), OCT exams (HE et al., 2020), plaque characteristics (WANG, 2020) and lesion numerical characteristics from FFR CT (TESCHE; GRAY, 2020; MANNIL et al., 2020; SHU et al., 2020) and CCA classification (OVALLE-MAGALLANES et al., 2020; GUPTA et al., 2019).

Several regression algorithms in patient data (AAKKARA; AAISUEB; AEELANUPAB, 2020) and multivariate logistic regression (SHU et al., 2020) were presented, but apparently they do not provide excellent results mainly by the small validation dataset and reduced population. Over 70% of treatment decisions still rely on visual estimation of angiographic stenosis, which has limited accuracy (about 60%–65%) for the prediction of FFR ≤ 0.80 (HAE et al., 2018), mainly due to the time and expense associated with FFR-guided decision-making (HAE et al., 2018). Invasive FFR has been a standard tool to detect ischemia-producing lesions with FFR ≤ 0.80 (HAE et al., 2018).

3.3 Challenges and future directions

FFR is based on a proportional linear relationship between coronary blood flow and perfusion pressure during maximal vasodilation (PARK; LERMAN; HERRMANN, 2017) and has been widely validated (IHDAYHID et al., 2018). However, its cost, availability, vasodilator need (BHATT, 2018; RAMASAMY et al., 2020), and poor hemodynamic correlation (STEGEHUIS et al., 2018; NOUS et al., 2020) makes its use for physiological assessment meet the current guideline indicators in only 50% of the cases (DEY et al., 2018). FFR importance is questioned (STEGEHUIS et al., 2018) as well as hyperemia is suggested not to be essential to the correct diagnosis (BENENATI et al., 2018). Recent evidence affirms that the use of non-invasive techniques reaching similar results is promising to replace the invasive FFR (BENENATI et al., 2018).

Both hemodynamic and anatomical assessment are necessary for correct functional assessment (BAIBHAV et al., 2018; DEY et al., 2018; KIM et al., 2016). Emerging techniques based on CCTA have been widely used for its availability for basic science research (CONTE et al., 2020; BENTON et al., 2018; G. et al., 2018; DANAD et al., 2017; KO et al., 2015). Once deaths caused by coronary heart disease or nonfatal myocardial infarction was 3 times more frequent in patients with adverse plaque (WILLIAMS et al., 2019), obtain plaque characteristics is essential to the correct CAD evaluation (BHATT, 2018; WANG et al., 2018). For this mean, the OCT, CCTA (AGUJETAS et al., 2018), IVUS (QURESHI; AGRAWAL, 2017; SIOGKAS et al., 2016), and angiography (DEY et al., 2018) are suggested as good sources of information. Unfortunately, current guidelines lack protocols to obtain high quality exams and make them available for larger studies and validation (WILLIAMS et al., 2019).

QFR and FFR_{CT} (MIN et al., 2015) have good correlation with FFR (TAKAHASHI et al.,

2019; TANIGAKI et al., 2019; YAZAKI et al., 2017; PANG et al., 2016; CESARO et al., 2018) and superior to angiographic assessment alone (WESTRA et al., 2018). While some methods achieved good results excluding severe lesion (BENTON et al., 2018; G. et al., 2018; TIES et al., 2018; Di Serafino et al., 2016), some emerging techniques and iFR (BENENATI et al., 2018) are promising to identify FFR < 0.8 (TANIGAKI et al., 2019; MICHAIL et al., 2019; COENEN et al., 2017) or FFR > 0.8 only (KO et al., 2015; SIOGKAS et al., 2017; RAMASAMY et al., 2020; PYXARAS et al., 2018). In general, these studies require further validations with bigger population (TSOMPOU et al., 2018; LEONE et al., 2018; CARUSO et al., 2016; PANG et al., 2016) to be applicable in larger population.

These findings evidence the applicability of the *no free lunch* theorem to coronary artery assessment (WOLPERT; MACREADY, 1997), meaning that no existing approach can achieve the best results in every type of lesion. Image-based exams, such as CCTA and QFR has shown good results and promise to be the future standard non-invasive characterization of plaque morphology. Most important than plaque characteristics, image-based exams' ability to differentiate low- and high-risk lesions has good preliminary results for FFR correlation, but still have space for improvements to reach good results in different lesion types.

Not only limited by image-based techniques, FMM and MLD together have shown improvement in the accuracy of CAD evaluation and provide a good relationship between anatomical and physiological assessment (KIM et al., 2016). FMM is suggested as the link between anatomical-physiological discordance (KIM et al., 2016) also requiring further validation. Nevertheless, FMM is turning real the fact of clinical data together with plaque characteristics provide better patient condition evaluation. This evidence suggests the increasing of image-based technique use for quick and reliable risk stratification of each patient mainly for severe lesions exclusion.

While angiography is the gold standard, studies suggest that is necessary for the whole patient view to provide the correct diagnosis, observing more than plaque, but also patient conditions and patient scores. To automate part of the current protocol and allow the dedication on the selection of patient data for a better lesion evaluation, this work proposes the removal of manual interaction. This approach may lead to better evidence and address the existing gaps in the gold standard.

Once automated and generating valuable data, the use of the 0D model can improve to become the benchmark for hemodynamic analysis. Even with improvements in the protocol, bigger labeled datasets are required. The current use of augmentation techniques and small open datasets lead to unclear technique evaluation, once evaluate models with synthetic data eventually do not reflect real patient circumstances.

3.4 Discussion

The literature review demonstrated that currently available exams represent several limitations, mainly caused by the invasiveness and the single-lesion evaluation. The current evidence suggests the need for combined techniques to reach the real CAD significance, so the use of the gold standard FFR is even more questionable. The use of simulation techniques together with the anatomical metrics raises an opportunity for improvement, once they generate valid insights for the patient lesion classification. The currently accepted techniques are mainly based on angiography, which is an invasive method and all further procedure depends on angiography employment.

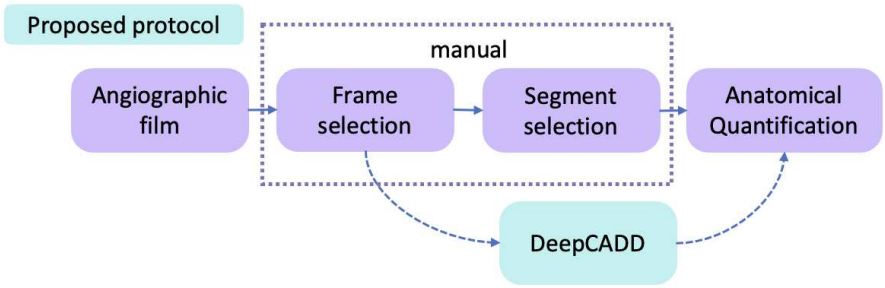
In this context, with angiography being the gold standard together with FFR for lesion severity and the lack of automated tools for multi-vessel identification, this study looks to the use of anatomical evaluation as the approach to detect CAD in RCA. The current architecture proposal is our contribution to the literature. We combined metrics that are relevant for deep learning model, together with the evaluation by physicians to provide an intelligent model in the current angiography acquisition protocol. This study make the best use of current invasiveness level, aiming to reduce additional invasive techniques and reaching comparable performance with the gold standard.

4 DEEPCADD MODEL

Coronary assessment to detect CAD depends on the correct lesion identification and quantification. The process is usually done manually via frame selection and stenosed segment region mark, followed by the quantification of the stenosis percentage as performed by the acquisition software. During this protocol, there are lots of manual interactions that can be enhanced with the use of intelligent software. To automate part of the protocol, this Master Thesis proposes a parallel step after the image acquisition and manual determination to help in the identification of the lesion area. Figure 24 presents the standard protocol (purple squared boxes) and the proposed protocol, with the DeepCADD architecture between the exam and the final lesion quantification.

In the standard protocol, the angiography exam generates films, usually in 15 frames per second, which highlight the coronary artery segment. Then, the frame that better highlights the artery with the fluorescent dye is manually selected (*Frame selection* on Figure 24). With the best frame selected, the visual analysis determine the image location with narrowed segments (*Segment selection* on Figure 24). Once the segment is selected, the coordinates of the lesions in the image are transferred to the QCA, which determines the percentage of the lesion.

Figure 24 – Proposed protocol for functional artery assessment based on angiography with DeepCADD

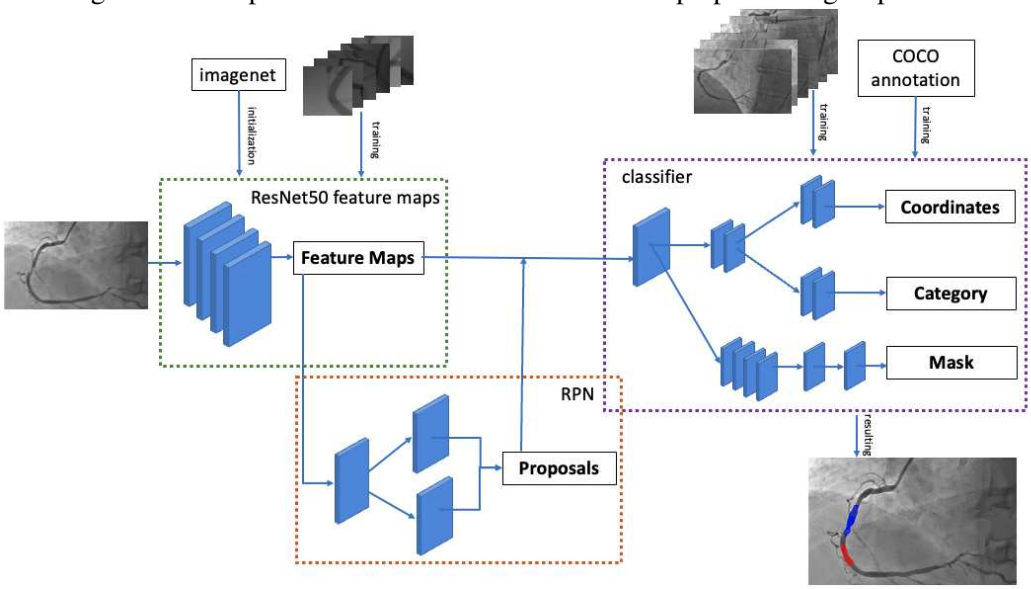


Source: Elaborated by the author.

Coronary angiography datasets availability is rare and the presence of personal data in the original exam requires anonymization. For this work, we use two datasets, presented in the next section which are in part synthetically created and in the other part, obtained manually in a local institution's medical records.

Due to the dataset limitation and the challenging task of classifying angiography images, we have evaluated previous pipelines to automatically identify lesions using selective search (FREITAS et al., 2021b). We identified the potential of selective search, but we also noticed that the RPN module already outperformed selective search. In the same study we evaluated the use of CNN for CAD classification (FREITAS et al., 2021b). Accordingly, the fact that DeepCADD is composed of an R-CNN-based architecture is a matter of technology improvement as we decided to use the newest technologies which performed better in the recent literature. DeepCADD presents several steps to improve the original Mask R-CNN performance and this

Figure 25 – Proposed architecture with datasets and preprocessing steps



Source: Elaborated by the author.

chapter details all the parts of the model, starting with an architecture overview in Section 4.1 and then presenting each module.

4.1 Architecture

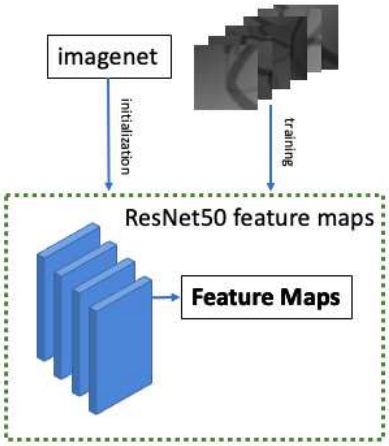
In this section, we present the architecture proposed to detect CAD in angiography images of RCA based on a deep object detection neural network. We expect to achieve a model that can detect CAD with performance comparable with the gold standard. Based on incremental development of classification model up to an object detection model, we decide to implement a R-CNN-based architecture, with improvements to specialize it for CAD context.

To this end, the proposed architecture is composed of the Mask R-CNN base. The original backbone module is composed of a CNN. Here we replaced it with a new ResNet-50 feature extractor. The architecture presented in Figure 25 shows the use of ResNet-50 feature extractor as the architecture backbone, together with the RPN and classifier layer. As the output, we expect to have a class label, bounding-box offset, and a third branch that outputs the object mask. The choice for a Mask R-CNN as a base for DeepCADD is a design decision, in which we expect to provide more interpretability for DeepCADD output.

From the architecture point of view, Figure 25 presents the model receiving the angiographies. As a design decision, we resized the angiographies to 224x224 pixels, followed by the ResNet-50 convolutional layers trained with the segments dataset. Once the architecture for image classification is trained with the segments datasets, we expect to connect the object detection layer from Mask R-CNN.

Figure 25 also details the datasets for each one of the modules. The backbone module is

Figure 26 – Backbone details on DeepCADD architecture



Source: Elaborated by the author.

trained with angiography segments dataset to learn the angiography complex image features. Then, the classifier will be trained with the angiography dataset. Further details about DeepCADD architecture are detailed in the next sections.

4.1.1 Backbone

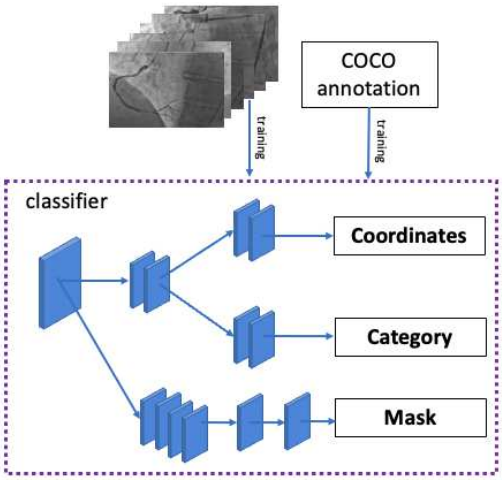
The backbone module is responsible for the feature maps extraction used in the object detection architecture. Figure 26 presents the backbone, which learns all the image features, from the most simple up to the most complexes. For the DeepCADD model, we are training the backbone with a ResNet-50 architecture, initialized with ImageNet weights. We compared different architectures, such as VGG16, MobilenetV2, ChexNet, Xception, and ResNet-50 which were the most promising CNN architectures in the literature applied to coronary artery assessment. For a better understanding of the experiments using the CNN architectures, we detail them in Chapter 5.2.1.

DeepCADD's backbone is planned to perform binary classification. The resulting feature maps from this binary classification model will map the stenosed and non-stenosed into positive and negative values. Thus, the important outcome from this step is the complex features extraction powered by the existing skipping connections from ResNet-50, which improve the propagation of small image features for deeper layers (HE et al., 2017).

The backbone replacement is an important step in the DeepCADD architecture, which leads to the feature extraction and provides the feature maps necessary to the classification layer interpret the angiography and the small objects in the whole frame. In the next section, we explain the classifier layer details.

4.1.2 Classifier

Figure 27 – Classifier details on DeepCADD architecture



Source: Elaborated by the author.

We re-use the original classification top layer from Mask R-CNN. The classifier module contains the coordinates together with the class determination and the additional instance segmentation that generates masks. These features will help in the automatic determination and classification of the lesion. This module will require a training step with a reduced number of images. We planned to use 100 angiography images labeled with bounding boxes identifying the CAD.

Image annotation was planned to be in the COCO's polygon format, which is the expected format to the Mask R-CNN architecture input. This format provides the bounding box class and the four points which represent the 2D position in the image (x_{min}, x_{max}, y_{min} and y_{max}) together with the class.

We also planned this module to perform binary object detection. This implementation is simpler than the multi-class, but in fact, it is enough to reach the CAD identification defining CAD candidates for QCA execution. In the current architecture definition, there is no need for preprocessing in the angiographies. Mask R-CNN just requires the annotation and the data to be resized to the correct input data, which must be 224x224 pixels. In this context, the next section explains the image segmentation pre-processing created by this study. This image segmentation is not necessary at this moment but can be used in future studies, such as in the automation of more steps of the standard protocol.

5 EXPERIMENTS AND RESULTS

In this chapter we will detail the materials and methods employed in this research. We proposed a structure that includes the methodology and the results. In the following section we will detail every step of DeepCADD implementation and preprocessing.

5.1 Methodology

The methodology starts with the datasets explanation and their applicability in the DeepCADD structure in Section 5.1.1. Section 5.1.2 explains the annotation process executed together with two physicians. Then, we present the implementation details in Section 5.1.3. We discuss the evaluation metrics for classification and for object detection in Section 5.1.4. Section 5.1.5 details every step designed to the validation with physicians.

5.1.1 Datasets

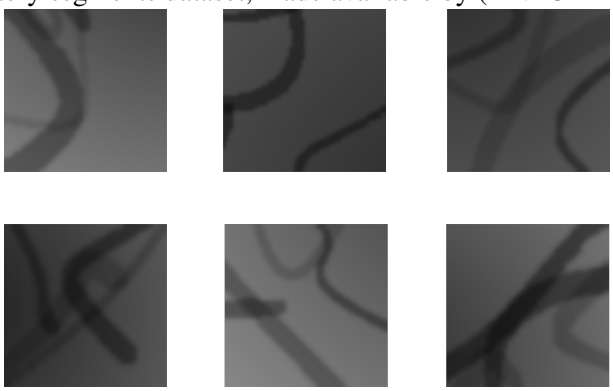
Coronary artery angiographies are rare to find in open databases. The most common angiography images are samples from published studies which do not provide access to the original dataset. For this reason, we have defined two strategies to achieve data for the proposed model. The first one is the use of coronary artery segments dataset (ANTCZAK; LIBERADZKI, 2018). The second one is the acquisition of frames from coronary angiographies executed before and after angioplasty procedures in partnership with the Cardiosinos Cardiologia Intervencionista private company in Novo Hamburgo, RS, Brazil with the consent of the National Research Ethics Committee. Angioplasty is the invasive procedure to include a stent in the interior of the coronary arteries to reestablish the blood flow for a healthy level.

5.1.1.1 Synthetic dataset

The coronary artery segments dataset, is composed by normalized synthetic images with size of 32x32 and 64x64 pixels (ANTCZAK; LIBERADZKI, 2018). The synthetic dataset contains two different classes, the narrowed arteries and the normal arteries. This dataset was synthetically created, so they represent arteries as perfect tubes (ANTCZAK; LIBERADZKI, 2018). The artery segments were generated using the Bézier curves technique, commonly used in cerebral and coronary artery simulations (ANTCZAK; LIBERADZKI, 2018). It means that it does not have the natural deformations of coronary arteries. This approach is valid and enables the feature extraction even with limited real angiography datasets available. Figure 28 presents a few samples of this dataset.

This dataset has a group of segment images with 32x32 pixels with 2,500 images of each of the classes: negative and positive cases, which represent, respectively, presence and absence of

Figure 28 – Synthetic artery segments dataset, made available by (ANTCZAK; LIBERADZKI, 2018)



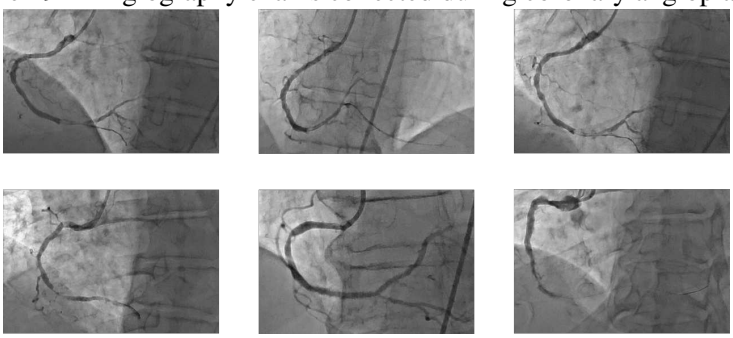
Source: Elaborated by the author.

narrowed arteries. Also, it has another group of segment images with 64x64 pixels with 1,000 images of each class. All the data is available in Pickle files already normalized, so they are not visible directly and only after the image be multiplied by a value of 255. We decided to use the group of data already in 64x64 pixel due to the quality of the dataset.

5.1.1.2 Angiography dataset

Artery segments only are not enough to reach the study objective. So we obtained access to an angioplasty database with the consent of the National Research Ethics Committee. Each angiography contains films from different incidences before and after the angioplasty. It means that we received raw films with RCA with reduced blood flow and the same arteries with blood flow reestablished for a healthy volume. We extracted the frames from angiographic films and after data anonymization, we collected approximately 500 frames. Some examples of the angiography dataset are presented in Figure 29.

Figure 29 – Angiography exams collected during coronary angioplasties



Source: Elaborated by the author.

With 50 exams from different patients, and considering the artery position variation during the film frames, we obtained 500 images, divided into two classes:

- **Severe stenosis**: this class represents the right coronary artery angiogram with a visible

narrowing in any of the main segments: the proximal RCA, the medial RCA and the distal RCA before an angioplasty.

- **Reestablished blood flow**: this class represents the images with a right coronary artery with a stent after an angioplasty. It means that coronary blood flow has been reestablished for a healthy volume and no significant narrowing is present.

Each one of the frames was saved using screen capture tools removing the personal data of the patients. In this way, all the image contains only the coronary artery image. The reduced number of images was identified as a challenge. Fortunately, R-CNN architecture generalizes well and does not require too much data to reach good results on object detection. From the 500 frames, we selected 132 frames with distinct arteries, visible narrowing, and containing the full RCA highlighted with fluorescent contrast. From the final angiography dataset, 118 frames were used for training and validation and 14 frames for testing.

5.1.2 Annotations

Object detection architectures, such as Mask R-CNN require labeled images for the training phase. These architectures support VGG JSON and COCO's polygon format. As a design decision, we decided to use the Make Sense alpha[1] tool, which provides an easy user interface to label images in the COCO's polygon format. The Make Sense tool also provides an easy way to add new images to the current labeled dataset, which is very helpful considering the incremental acquisition of new angiographic frames. Figure 30 presents an example of the annotation created for this study.

Figure 30 – Angiography exam annotated with Make Sense AI tool.

Source: Elaborated by the author.

In this study, we needed to label the angiography dataset. Each angiographic exam was labeled by a physician with the COCO's polygon format using the Make Sense tool. The an-

[1] https://www.makesense.ai/

notation process was executed in the presence of this Master's thesis author and the physician. The selection criteria for the lesions in the angiographic frame was defined as all the visible narrowing lesions in the RCA, apart from if the lesion can lead or not to angioplasty.

5.1.3 Implementation details

DeepCADD construction started with the CNN implementation, testing the plannedVGG16, MobilenetV2, ChexNet, Xception and ResNet-50 VGG16, MobilenetV2, ChexNet, Xception and ResNet-50. We identified that R-CNN's backbone only accepts sequential architectures. We noticed that VGG16 architecture is not able to be implemented directly. We studied the Mask R-CNN input and we reached the common 224x224 size.

We implemented the CNN architecture using ImageDataGenerator from Keras, which allows the automatic data augmentation, which avoids the need for manual augmentation. We divided both synthetic and angiography datasets into the training and validation in the proportions 0.7 and 0.3. We instantiate the CNN with the top layer, which is usually a Dense layer with 1000 neurons. This last layer is then, replaced by a Dense layer with 2 neurons and *softmax* activation, which performed better than a Dense with 1 neuron and *sigmoid* activation.

CNN was then trained between 5 and 100 epochs, which presented to be the highest boundaries we could reach with the dataset without overfitting. After that, we evaluated the confusion matrix to reach the model performance looking for an accuracy higher than 90%. This parameter of 90% was initially established as a target due to the top 75% performance achieved before the use of ResNet-50. We decided the challenge ourselves to reach at least this percentage that would mean a significant performance with small loss values. With the model defined, we save the model weights with and without the top layer. The top layer removal in the model weights is necessary as the Mask R-CNN implementation allows both initialization.

Once the CNN with the best performance is reached, we started the R-CNN implementation. We implemented the original Mask R-CNN architecture first to understand the benefits and drawbacks, together with the technical difficulties, such as Python, Keras, and environmental versions. Aside from these are merely technical details, usually version conflicts require too much time to enable simple network testing. At this point, Google Colaboratory was, once again, the best alternative.

After the initial implementation and testing of Mask R-CNN, we went through the backbone replacement. We replaced the chunk of code initializing the original ResNet-50 by the new CNN pre-trained with the synthetic CAD dataset. In fact, we just finish the DeepCADD setup. Once this replacement is correctly concluded, we will be able to re-train the DeepCADD with the angiography dataset.

We decide to follow the head training approach, which trains only the R-CNN top layer, with all the other layers frozen. For the DeepCADD head training, we require the COCO's polygon dataset annotation. To do so, we annotate all the 132 frames together with a physician,

generating a JSON file with the COCO's polygon annotation. Then, we trained the DeepCADD with a few epochs from 5 up to 20 also identified as the boundaries without critical overfitting. Finally, we evaluate the DeepCADD with the object detection metrics already described and prepare the elements for the specialist validation, which is composed of two steps and is detailed in the following section.

5.1.4 Metrics

Considering the architecture described, the two main stages of the model proposed in this Master Thesis require specific evaluation metrics. From the classification point of view, the use of confusion matrix will derive several metrics which allow the model evaluation to reach accuracy higher than 90%. On the object detection module, the use of Intersection Over Union (IoU) and mAP will produce the necessary model adjustments necessary to the correct bounding boxes determination. The next subsections describe these metrics.

5.1.4.1 Metrics Based on the Confusion Matrix

In the binary classification problem from backbone module, instances are mapped into two classes. Then, for any binary classification, there are four possible results (FAWCETT, 2006) shown in Table 4.

Table 4 – Confusion matrix comparing the real classes with predictions

		Positive	*Negative*
Predicted	*Positive*	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

Source: Adapted from (FAWCETT, 2006).

From the confusion matrix, it is possible to calculate the classifier performance metrics (RUUSKA et al., 2018). The current and planned evaluations are:

- **Sensitivity**: Represents the model's ability to detect the presence of the disease, defined by Equation 5.1:

$$sen = \frac{TP}{TP + FN} \tag{5.1}$$

- **Specificity**: Represents the models' ability to detect the absence of the disease, defined by Equation 5.2:

$$spec = \frac{TN}{TN + FP} \tag{5.2}$$

- **Precision**: Represents the fraction of positive instances classified correctly, defined by Equation 5.3:

$$prec = \frac{TP}{TP + FP} \tag{5.3}$$

- **Accuracy**: Represents the model's abaility to correctly classify each case, defined by Equation 5.4:

$$acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.4}$$

- **F1-Score**: Represents the harmonic measure between precision and sensitivity, defined by Equation 5.5:

$$F1score = \frac{2 * TP}{2 * TP + FN + FP} \tag{5.5}$$

These metrics presented intend to be enough for the classification model evaluation. Nevertheless, based on the enhancement required in the results, some other metrics can be added to this work.

5.1.4.2 Metrics for Object Detection

As a challenging task, Object detection has some common metrics as well as other specifics. The most important metrics are Precision and Recall, which are the same as described in the 5.1.4.1 section, the IoU optimization, and the Average Precision (AP) metrics. In the object detection task, the model targets the list of class probabilities and the offset coordinates for each reason proposed by the Selective Search algorithm. The IoU is the area of intersection between the two boxes, divided by the union of the area occupied by the original box and the predicted box (CHAHAL; DEY, 2018). It helps to adjust the original bounding box shape to match the exact object (CHAHAL; DEY, 2018). The mAP is one of the most popular metric to evaluate object detection tasks and has been widely used in recent papers (CHAHAL; DEY, 2018). This is basically the mean of Average Precision (AP) represented in Equation 5.6 [2]:

$$\text{AP} = \sum_{n} (R_n - R_{n-1}) P_n \tag{5.6}$$

The YOLO paper as well as the PASCAL Visual Object Challenge (VOC) (LIN et al., 2015) dataset details mAP to be the same quantity as AP. The COCO's polygon dataset (EVERINGHAM et al., 2010) however, uses a modification to this metric called the Mixed Average Metric (CHAHAL; DEY, 2018).

[2] https://github.com/blmoistawinde/ml_equations_latex

5.1.4.3 Additional metrics

Aside the current object detection and image classification metrics, we noticed the need of additional metrics not based on confusion matrix or IoU. For this reason, we decided to explore two other metrics, the Matthews Correlation Coefficient (MCC) and the Cohen's Kappa coefficient.

MCC was introduced by Brian W. Matthews to evaluate the predicted structure of an enzyme, in a biochemical study in 1975 (CHICCO; WARRENS; JURMAN, 2021). Since 2000, MCC has been used as a standard metric in several scientific competitions, such as the Kaggle competition to detect power line fault detection (CHICCO; WARRENS; JURMAN, 2021). The Equation 5.7 presents how to achieve MCC.

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{5.7}$$

Cohen's Kappa coefficient is a statistical measure used for qualitative items. This metrics is considered more robust than a simple percent agreement calculation (VIEIRA; KAYMAK; SOUSA, 2010). Aside some criticism about the Kappa measure, it is statistically robust (BEN-DAVID, 2008).

The Cohen's Kappa coefficient has been widely implied in social sciences, biology and in medical sciences for a couple of decades (BEN-DAVID, 2008). However, within machine learning and data mining communities, Cohen's Kappa has not received much attention as a metric for accuracy (VIEIRA; KAYMAK; SOUSA, 2010). Kappa has two different equations, considering binary and multi-class problems. The Equation 5.8 presents how to achieve non-weighted Cohen's Kappa.

$$\kappa = \frac{2 \times (TP \times TN - FN \times FP)}{(TP + FP) \times (FP + TN) + (TP + FN) \times (FN + TN)} \tag{5.8}$$

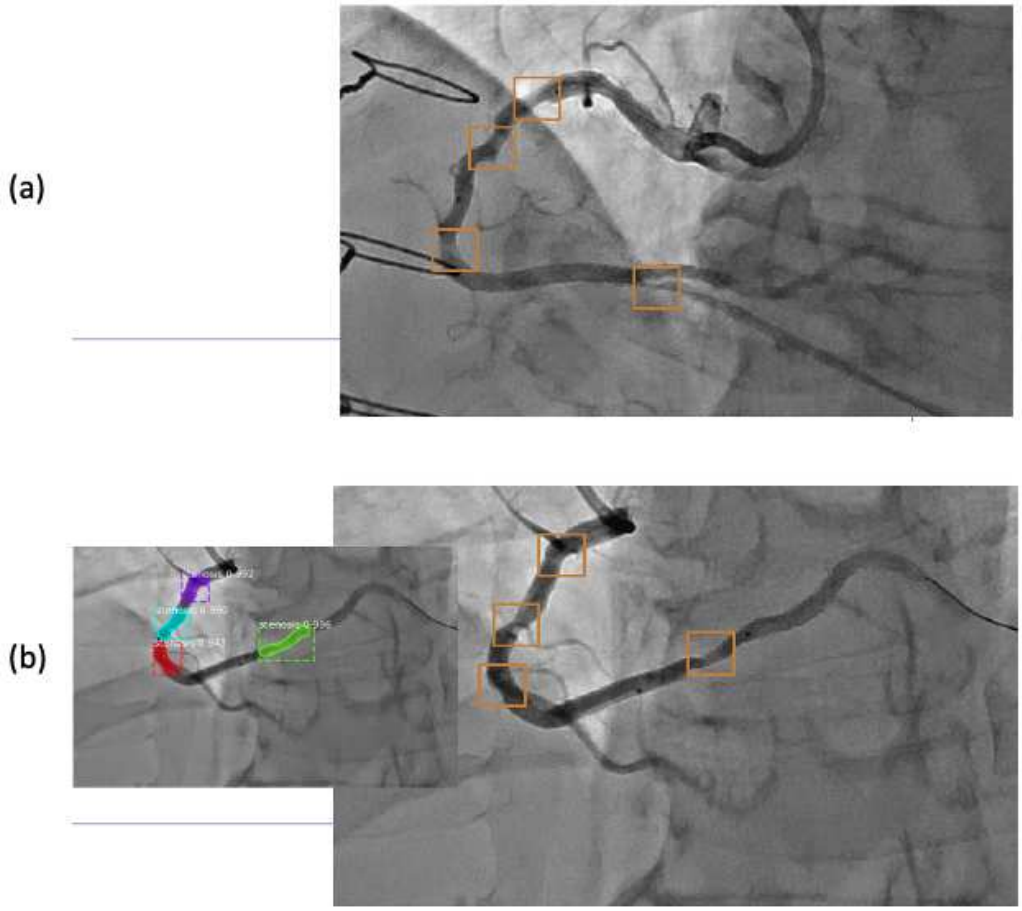
In this study we will only apply the equation for binary classification problem as we are only working with two classes.

5.1.5 Validation with specialist

Besides the defined object detection metrics, initial DeepCADD training with angiography dataset evaluated by metrics based on IoU suggested the need for additional validation and alternative metrics reach the correct model performance. For this reason, we perform validation with two physicians, one specialist in the cardiac area, and a clinician. In Figure 31 we present the validation slides prepared.

The validation was composed of 20 angiographic frames, due to the specialist and dataset availability. We prepared a slide presentation with 10 angiographies without any annotation and we asked the physicians to identify CAD in the angiographic frames, considering all the artery

Figure 31 – Slides of the validation with the doctors.



Source: Elaborated by the author.

narrowing as presented in Figure 31 item (a). In item (a) we present the original angiography with the lesions marked with squares.

The last 10 angiographies were presented to physicians together with DeepCADD lesion suggestions. We asked the physicians for validation to identify if DeepCADD prediction reaches the same specialist understanding, presented in Figure 31 item (b). In item (b) we present the DeepCADD lesion suggestions on the left side and the specialist validation on the right side.

After this validation, we counted the number of lesions detected in the 20 images in terms of true positives, false negatives, and false positives. These data allowed us to calculate the model sensitivity and precision. Due to the validation modeling and the object detection binary classification, we were not able to count the true negatives, once everything that is not detected as a lesion in the image could be considered as true negative. Therefore, we noticed the importance of sensitivity and precision as metrics for patient triage in screening tools (RODRIGUES et al., 2021).

5.2 Results

With an object detection architecture, we wanted to validate the use of DeepCADD to detect CAD in angiographies. The presence and absence of lesion in angiography exams is a complex task for the human eye, once the difference between presence and absence of lesion is very subtle and it may be only a small narrow in the main artery segment, as well as a unique important narrowing a completely unhealthy artery. As pointed out in the materials and methods chapter, we started the experiments with the CNN optimization. The Xception, InceptionV3 and ResNet CNN architectures are complex and may lead to overfitting, so we needed to run several experiments to optimize the hyperparameters. For more details refer to Appendix B.

We implemented transfer learning reusing a CNN architecture pre-trained on ImageNet and then, we trained with the specific datasets. We present the classification model results in Section 5.2.1. Section 5.2.2 presents the object detection results. Finaly, Section 5.2.3 details the validation with specialists.

The experiments in this Master Thesis have been implemented using the Google Colaboratory PRO [3] with GPU P100 and 7.3 gigabytes of available RAM to reduce the time spent in the training phase, as well as avoid the version compatibility issues when configuring a specific machine or using Conda environments. The use of pre-installed lib in a common environment saved a lot of time.

5.2.1 Classification model

In the classification model, we expected to achieve an excellent performance classifying CAD presence and absence in coronary artery segments. We executed several experiments using different CNN architectures. We tested different classification layers from only one Dense layer with 1 neuron up to a sequence of 4 Dense layers of 100x100x100x1 neurons. Table 5 presents the best performance of each architecture in the angiography classification test using the same dataset. In the same table, we detail the final dense layer and the hyperparameters of each one.

Table 5 – Classification model architecture configurations

Base	Learning rate	Final dense layer (neurons)	Accuracy	Loss	Trainable par.
Xception	$1e-7$	100x100x1	0.75	0.54	8M
CheXNet	$1e-4$	100x100x100x1	0.76	0.63	all
InceptionV3	$1e-4$	256x2	0.96	0.65	all
ResNet-50	$1e-5$	2	0.90	0.39	23M

Source: Elaborated by the author.

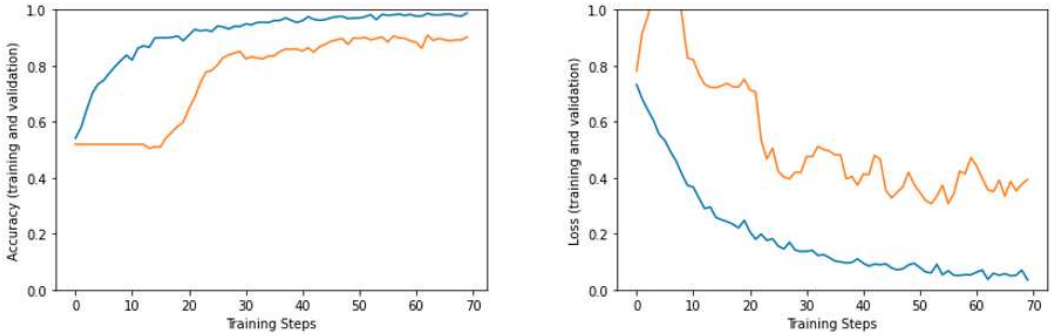
In the experiments, we used the synthetic dataset (ANTCZAK; LIBERADZKI, 2018) with 2,000 CCA exams of presence and absence of CAD. We decided to use binary classification for

[3] https://colab.research.google.com/

feature maps extraction with a final dense layer with 2 neurons and *softmax* activation, as well as 1 neuron with *sigmoid* activation. For all the tested CNN architectures we resized the input up to 150x150 pixels to match their input parameter.

We noticed that a binary classification with the ResNet-50 CNN using 2 neurons with *softmax* outperformed the other architectures with high accuracy and the lowest loss values from all the others. For this reason, we decided to continue the DeepCADD implementation with ResNet-50 architecture as the backbone. Using a batch size of 32 images and the Adam optimizer, we could reach a good performance in only 70 training epochs, as presented in Figure 32.

Figure 32 – ResNet-50 validation and training graphs.

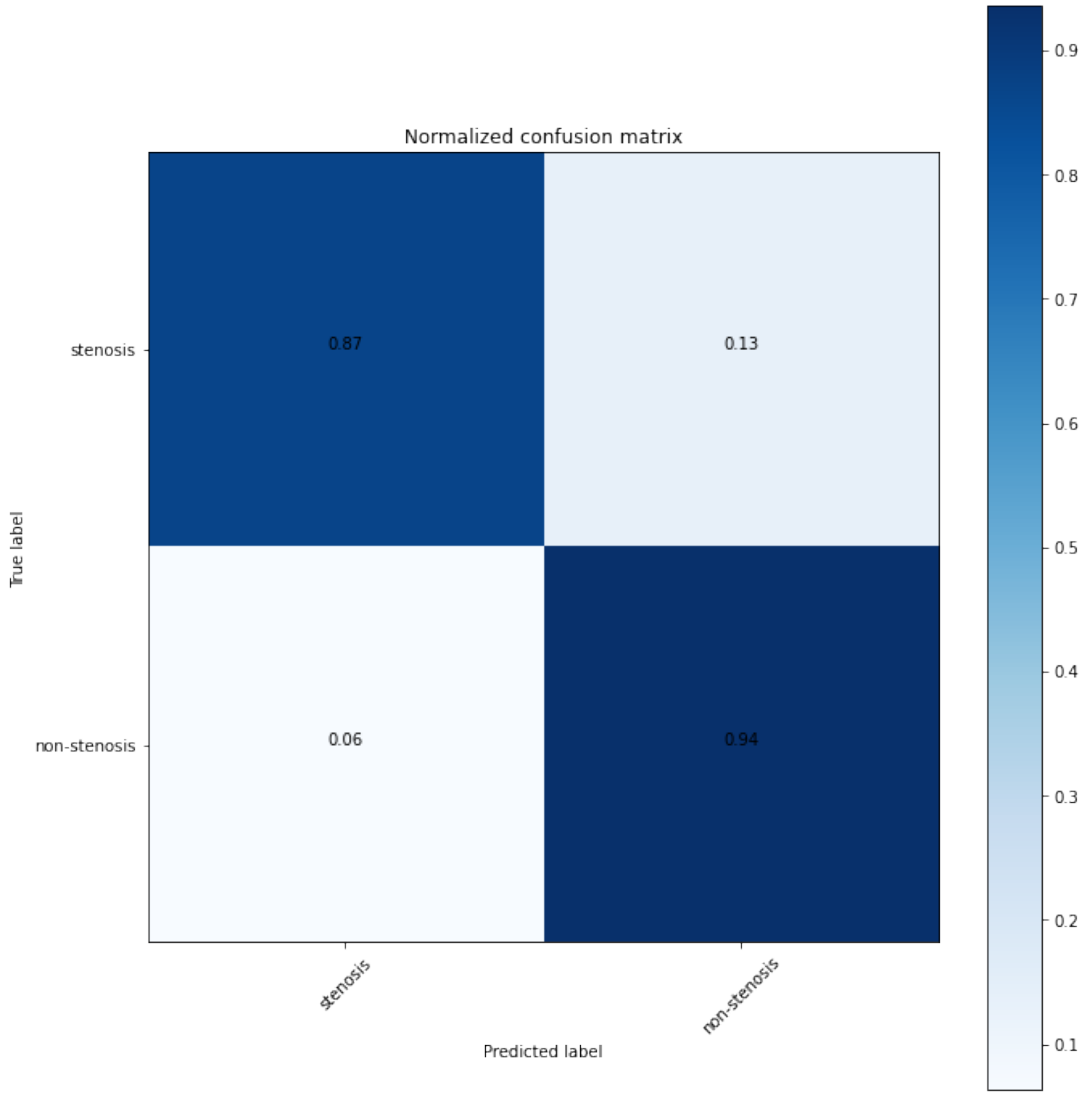


Source: Elaborated by the author.

The training parameter is not too stable as an ideal model, but we can see in Figure 32 the learning curve increasing properly. ResNet-50 learned most of the features between the 20th and 30th epoch but also learned up to the 70th epoch as well as the loss decreasing, which reached a value of 0.39. Figure 33 presents very interesting results of the ResNet-50 confusion matrix.

From the confusion matrix, we identified a sensitivity of 0.87, with a specificity of 0.94. We also reached a precision of 0.90, with a high recall and f1-score of 0.9. Fortunately, we could reach a very powerful CAD classification model with ResNet-50 architecture, which is helpful for the R-CNN object detection model implementation. Then, we expect that, with this high performance, we could improve the Mask R-CNN original performance in the CAD detection problem. We understood this improvement in the performance using ResNet-50 due to the fact of skip connections in the architecture, once it enhances the performance of higher layers understand small details in the image classification task. CAD can be interpreted as small image features in the whole angiography and as the grayscale frame, it can suffer from noise in the image acquisition, presence of other components in the image (e.g. stent), or calcification. With these results, we decided to implement the R-CNN and we detail it in the next section.

Figure 33 – ResNet-50 backbone confusion matrix.



Source: Elaborated by the author.

5.2.2 Object detection model

In the object detection model, we wanted to be able to detect CAD in angiographic frames with precision comparable to the gold standard. The gold standard, in this case, is the specialist visual assessment, as we focused on the segment selection. It is important to point out that, at this moment, the standard protocol for the angiography exam already happened, so the frame was selected and we are evaluating static frames. To do so, we first annotated the images with a specialist, reaching 132 frames labeled with the polygon COCO's polygon format. Due to the complexity of the existing object detection architecture and the promising results in the literature, we decided to experiments an R-CNN-based architecture. To do so, we implemented a Mask R-CNN going through the backbone replacement.

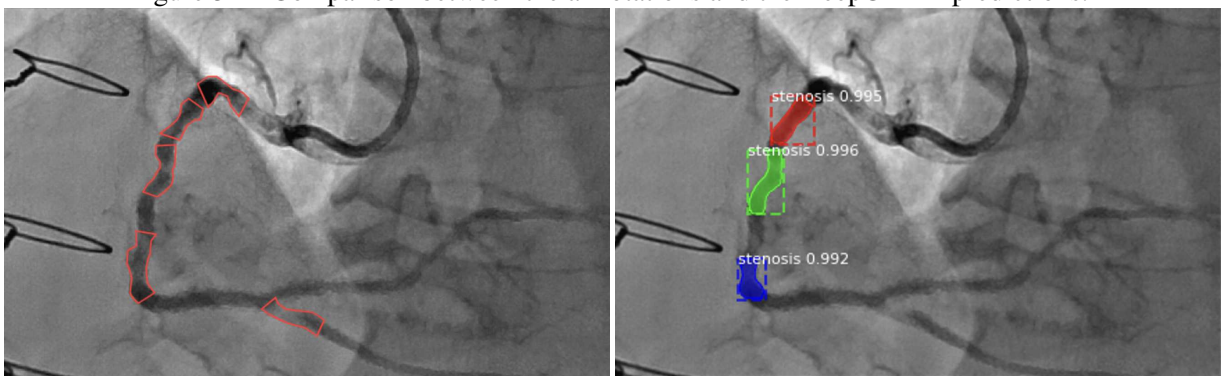
We developed the importing of model weights from the classification model, preferably

without the top layer. We initialized DeepCADD weights from our classification ResNet-50 model directly in the backbone. In the training phase, we imported the whole 132-image angiography dataset and the COCO's polygon annotation file. We tested different epochs but the best result happened with 5 epochs training only the network head with a learning rate of $1e-3$. We used the expected training and validation separation of 70% and 30%, respectively.

Evaluating the standard metrics, we reached mAP of 0.248, Mean Average Recall (mAR) of 0.5 and f1-score of 0.33. These numerical results are not so promising as they evaluate the IoU, which we identified during the tests and validation as not correctly correlating with the study objective. IoU measures the matching between the area predicted and the area annotated, but we are not interested in the lesion extension, but the lesion presence.

We needed additional metrics to evaluate the model performance. One of them is the visual assessment, between the annotations and the prediction. Figure 34 presents some examples comparing the annotation created by one specialist in the area and the model predictions.

Figure 34 – Comparison between the annotations and the DeepCADD predictions.



Source: Elaborated by the author.

Figure 34 shows the annotations created to the angiography (left image), and the DeepCADD detection (right image). DeepCADD identified three of five lesions identified by a specialist. This example is especially important because the two lesions not identified by the model at this time are considered as noncalcified lesions, which present a *fogging* aspect, making it harder to identify and were consequently, less frequent in the dataset.

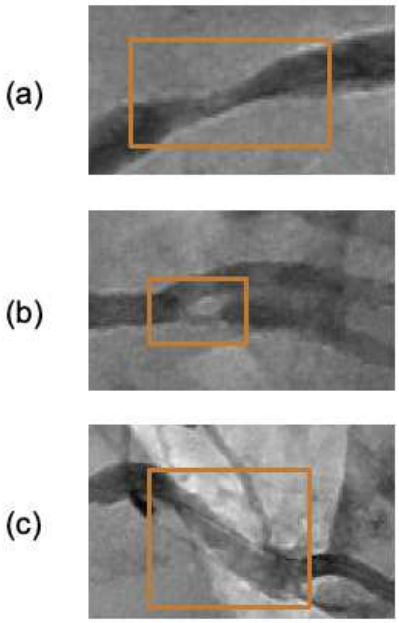
We selected 20 images for the visual assessment validation, resulting in 104 lesions. DeepCADD was able to recognize CAD lesions in 70% of the cases. There are some aspects of the exam impacting the model detection. These aspects are presented in Section 5.2.2.1, where we describe the different types of lesions.

5.2.2.1 Lesion characteristics

During the model validation with the specialists, we noticed three different types of the lesion for visual assessment. This bias is completely handled by the specialist, which highlights the importance of physician evaluation. In Figure 35 shows three different types of lesions.

Figure 35(a) we represent the most common lesion identified by the angiography visual assessment with high contrast between the artery and the background. Figure 35(b) shows the lesion close to a bifurcation, which creates noise in the visual identity as the lesion diameter is confused by the other branch. In Figure 35(c) depicts the plaques with lower density compared to contrast-enhanced lumen, which represent the noncalcified lesions (NILSSON et al., 2014).

Figure 35 – The three different types of lesion identified in the visual assessment.



Source: Elaborated by the author.

The kind of lesion represented in Figure 35(a) is characterized by the presence of a severe narrowing and the high contrast with the background, which makes the CAD visible. We executed an additional test separating the dataset into these three types of the lesion and we noticed that DeepCADD easily identify it, also representing the highest mAP performance achieved. For this type of lesion, DeepCADD is very similar to the Mask R-CNN alone initialized with ImageNet.

Figure 35 item (b) is characterized by the presence of bifurcation close to the lesion. This lesion is very challenging due to the visual noise generated in the real lesion diameter. Additionally, during the validation with the specialist, we detected that QCA tool, under development for the past 40 years also faces issues with bifurcation, so it was identified as a possible limitation, but not critical as the similarity with the current protocol equipment. DeepCADD presents some false negatives caused by this type of lesion.

Figure 35 item (c) presents the most challenging lesion identification. Noncalcified lesions present this aspect of *fogging*, caused by unstable plaques in the arteries. Noncalcified lesions are critical and during the specialist validation, we noticed that these lesions are also challenging

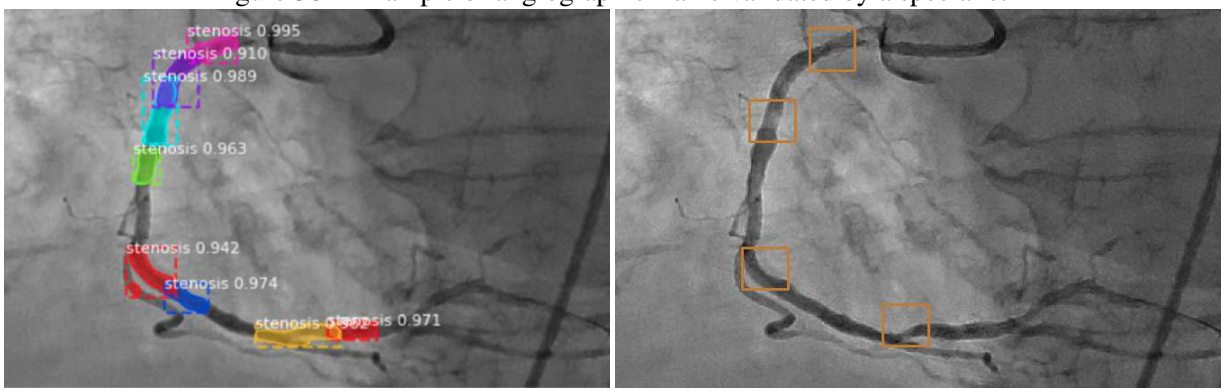
for QCA. DeepCADD presents also a good performance on these lesions mainly due to the comparison with the Mask R-CNN alone initialized with ImageNet. For this type of lesion, we noticed a very good improvement in the model caused by the backbone replacement.

Beyond the mentioned lesions, we also noticed interesting results in the lesions $< 30\%$, called wall-irregularities (ROOS et al., 2013). This kind of lesion is also considered a CAD, but with a lower severity due to its small artery narrowing (ROOS et al., 2013). We noticed the existence of this CAD mainly in the RCA stem. Therefore, as these lesions are not too significant, they are usually not selected by the specialists, as it does not represent risks. It can be considered as a gap in the understanding before the specialist's validation. We can also consider the wall irregularities as identified lesions, due to their anatomy, but in this current study, the impact on the final results and model performance is not significant. Wall irregularities can lead a future studies. In the next section, we present the results from the evaluation with specialists.

5.2.3 Model validations

The objective of this validation was to understand the real DeepCADD performance. We prepared a validation with 20 images, composed of 10 frames in which the specialist informed the presence of lesions including colored squares on them. In this phase, the specialist informed the lesion presence using the regions of RCA anatomy, such as proximal, medial, and distal segments. An example of lesions in each one of the segments is presented in Figure 36, wherein on the left image we have the suggestions from DeepCADD, and in the right image, we present the angiography marked by the specialist with one lesion in the proximal segment, two lesions in the medial segment and one lesion in the distal segment.

Figure 36 – Example of angiographic frame validated by a specialist



Source: Elaborated by the author.

In the second phase, we provided the DeepCADD output as angiography with the lesion marked with a mask and bounding boxes, followed by the original angiography in a bigger size to make the evaluation easier by the specialist. This step result is the same as Figure 36 right image, but this phase was used to confirm if the suggested lesions from DeepCADD are CAD as well adding the CAD identified by specialists and missed by DeepCADD.

The example of Figure 36 is very interesting to discuss both the performance and limitations of DeepCADD. During the validation, we received very interesting feedback from the specialists. One example is the visual presentation of DeepCADD, in which we currently present the mask, the bounding boxes, and the class with the percentage of confidence in the output. The feedback from specialists was to reduce the number of information in the DeepCADD output, leaving only the mask, without the bounding boxes and the class name, without visual impact in the RCA segment.

As the common object detection metrics are not enough, we also analyzed the model sensitivity and precision. We compared DeepCADD with both specialists and the performance over the 104 lesions. A significant attention point in the evaluation of the metric is the limitation caused by the objective of this proposed model. DeepCADD is trained to identify the presence and absence of CAD, but it detects only the presence of a lesion, so we reached a problem to determine true negative values. It impacts mainly the metrics which depend on correct true negative values. Therefore, we evaluated the correct detection of 104 lesions compared with each one of the specialists with DeepCADD detection. In Table 6 we present that DeepCADD achieved high sensitivity and relatively good precision values compared with each specialist in the CAD detection.

Table 6 – DeepCADD performance evaluation with the specialists

Metric	DeepCADD vs Specialist I	DeepCADD vs Specialist II
sensitivity	0.8	0.89
precision	0.83	0.57

Source: Elaborated by the author.

Table 7 presents important DeepCADD results, in which it outperforms other architectures. A direct quantitative comparison with related works is difficult, mainly due to datasets availability, different methodologies, data selection, and different classification or detection objectives. In this way, we organized in Table 7 a quantitative comparison with similar works, presenting the best results achieved by recent studies in terms of sensitivity and precision achieved. This comparison considers similar methodology and architectures, but with different datasets. While DeepCADD performance was achieved using the angiography dataset with 104 images, further studies used their datasets. Even with different datasets, but taking into account the similarities, DeepCADD outperforms similar investigations of sensitivity, which is particularly important due to good evaluation in multi-vessel CAD.

Besides the good sensitivity numbers, Table 7 also presents the limitation in the precision, mainly caused by the limitation of CAD absence detection as already stated. These results suggest a good performance of DeepCADD and consider it as a candidate to be used as a screening tool. The angiography protocol automation through stenosis detection is possible due to the high sensitivity model value, allowing the suggestion of artery segments to go through the QCA lesion quantification.

Table 7 – Comparison of DeepCADD model with the related works

Method	Prec. (%)	Sens. (%)	Reference
Libra R-CNN	58.44	79.51	(PANG et al., 2019)
Cascade R-CNN	68.41	67.9	(CAI; VASCONCELOS, 2018)
Faster R-CNN	66.3	74.81	(REN et al., 2015)
Mask R-CNN	70.91	76.47	(HE et al., 2018)
Stenosis-DetNet	94.87	82.2	(RODRIGUES et al., 2021)
Automatic detection	79.5	87.2	(WUA et al., 2020)
DeepCADD	**83.0**	**89.13**	**Proposed method**

Table 8 – DeepCADD MCC evaluation with the specialists

Comparison	MCC
DeepCADD vs Specialist I	-0.19
DeepCADD vs Specialist II	-0.22

Source: Elaborated by the author.

Even suggested as better than Cohen's Kappa, we noticed that MCC requires the number of true negatives and as already stated, which lead to poor values. Then, we analyzed the performance through the Cohen's Kappa value, as presented in Table 9, where we noticed that the absence of true negatives also impact negatively. We can also notice that the performance between the specialist individually are very similar, unfortunately presenting a minimal correlation.

Table 9 – DeepCADD Cohen's Kappa evaluation compared with specialists

Comparison	Cohen's Kappa
DeepCADD x Specialist I	0.22
DeepCADD x Specialist II	0.14
DeepCADD x Both specialists	0.61

Source: Elaborated by the author.

We also explored the analysis between DeepCADD and both specialists, combining the specialists lesion identification. It means that we have extracted part of the original confusion matrix where both specialist agree with the presence and the absence of stenosis. It means that, frames are considered with lesion, if at least one specialist indicate the same. So, Table 9 presents a substantial correlation between DeepCADD. The joint specialist analysis, which is very interesting, can suggest the use of DeepCADD for CAD analysis standardization.

As the validation was composed by two steps, we also decided to investigate the impact of each one of the validation steps in this Cohen's Kappa correlation value. We identified that correlation value differ significantly between the validation steps. The first validation step, composed by the CAD selection by specialists without any DeepCADD suggestion; and the second one with the specialist validating the suggestions from DeepCADD. Results from this partial evaluation are presented in Figure 10.

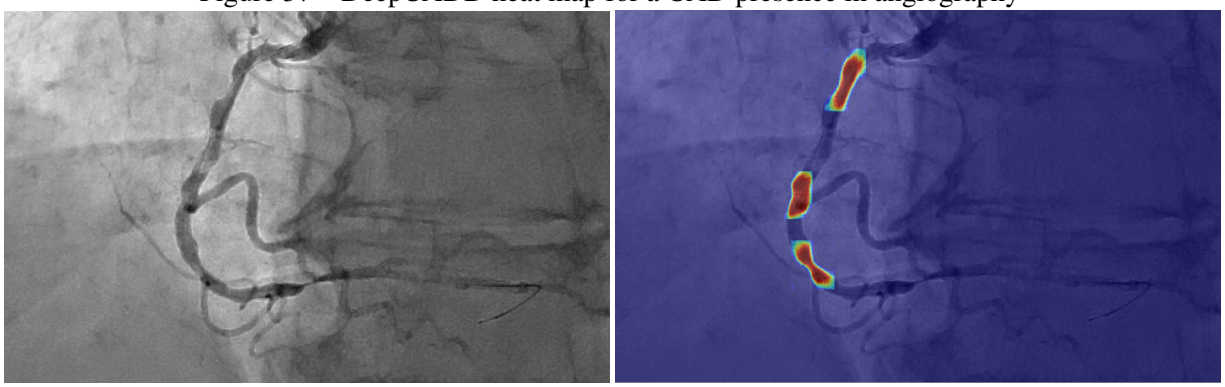
Table 10 – DeepCADD Cohen's Kappa evaluation compared with specialists and the validation steps

Comparison	Validation Step	Cohen's Kappa
DeepCADD x Both specialists	specialist CAD selection	0.47
DeepCADD x Both specialists	validate DeepCADD CAD	0.74

Source: Elaborated by the author.

Using the same angiography dataset, we identified that the agreement between the specialists and DeepCADD was higher in the second part of the validation when the specialists validated the lesions in the angiography with DeepCADD CAD suggestions. In fact, we came to an important discussion point, which we were not able to understand yet if DeepCADD is able to standardize the CAD detection system or created a new bias in the CAD identification. This is a very important topic to be addressed in a future study. Currently, the use of DeepCADD helps in the lesion identification, resulted in a significant correlation, and is a good suggestion for patient triage. To help in the DeepCADD explainability, we went one step further to analyze the DeepCADD heat map.

Figure 37 – DeepCADD heat map for a CAD presence in angiography



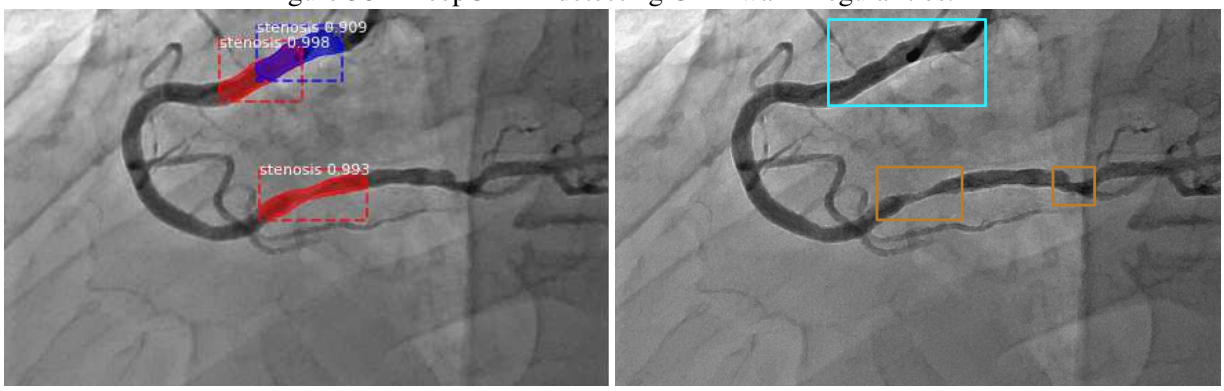
Source: Elaborated by the author.

The behavior shown in Figure 37 shows another important study contribution, where we see that DeepCADD was able to generate an activation for the suspected region, being able to also support in border cases, where the CAD is not completely identified. The heat map presented in Figure 37 brings pixel-level details in the DeepCADD prediction, informing the exactly RCA area where the lesion was identified.

Besides the agreement in the lesion identification, we also noticed that DeepCADD was able to identify the CAD $< 30\%$, already stated as the wall-irregularities. Even not representing a huge impact on the coronary blood flow, these irregularities are caused by CAD. We noticed during the validation that physicians do not select them due to this small impact. One example of this detection is presented in Figure 38.

In the left image of Figure 38 we present the DeepCADD CAD suggestions. In the right side image, we present the physician annotations. The wall-irregularity is present in the blue square in the right image and we noticed that DeepCADD was able to detect two CAD instances

Figure 38 – DeepCADD detecting CAD wall-irregularities.



Source: Elaborated by the author.

in the wall-irregularity region.

Finally, beyond the numerical analysis, we also received some visual improvement suggestions from specialists, such as the change of the mask colors. As DeepCADD detection of presence and absence is a binary problem, we left only one color for all the lesions. We provide an example of the suggested model improvement, reducing the information in the output in Figure 39. We also received a suggestion for, in a future study, providing three different colors depending on the lesion significance, which could make it even better for the specialist analysis, but will require changes in the architecture to support multi-class detection.

Figure 39 presents four different angiographic frames from different lesion with DeepCADD CAD detection. These suggestions implemented in Figure 39 are especially important, as the use of only one color also makes the detection easier, as we are not interested in the lesion extension, but the presence of it. If DeepCADD detects more lesions in the segment between two lesions, such as the 4 first top detection in Figure 36 would not be a problem as the lesion is compromised by CAD lesions and acceptable from the specialist's point of view.

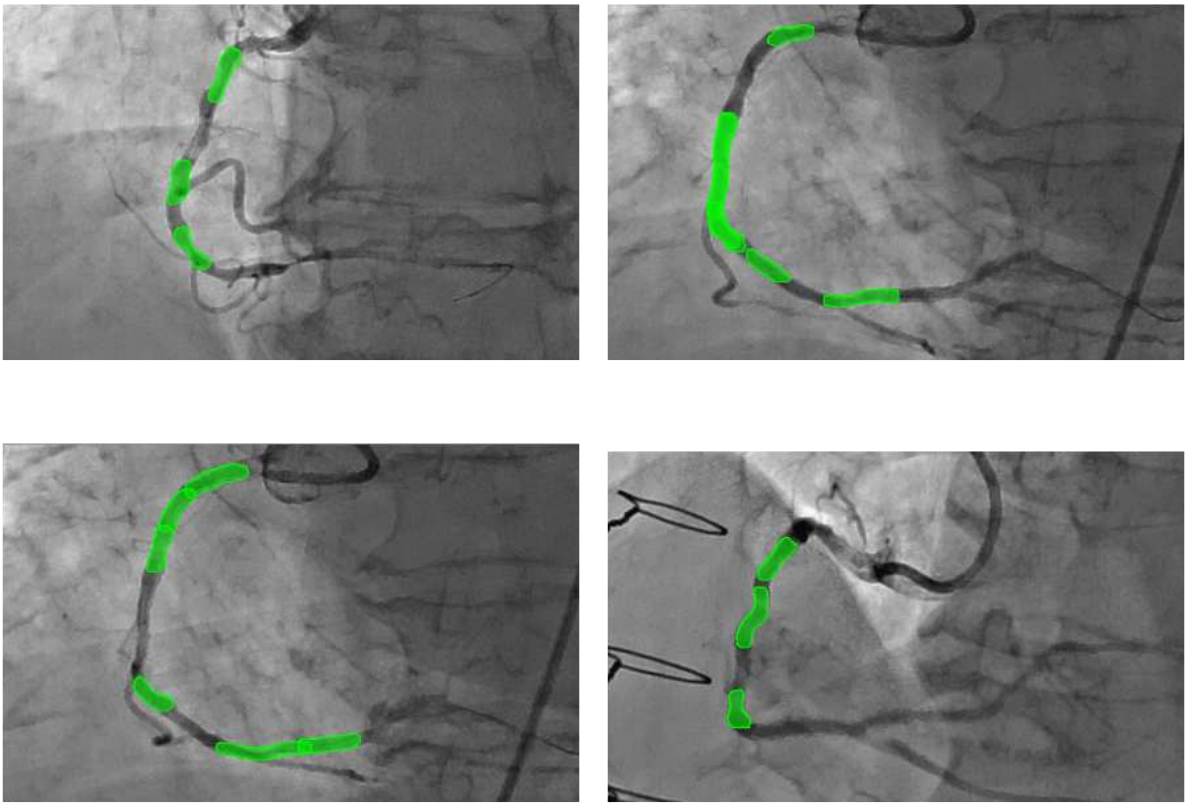
In summary, DeepCADD presented good numerical results in the validation and was able to identify most of the lesions in RCA angiographies. DeepCADD suffered improvements since its first version and we were able to propose a useful tool with high sensitivity and correlation with the daily visual assessment. DeepCADD can be enhanced and become even better for more types of lesions, which is natural for recent development, and the characteristics of the complex lesions, which are challenging the current QCA technique for a couple of decades.

5.3 Discussion and limitations

DeepCADD is an object detection architecture enhanced for CAD detection in RCA segments. This architecture was able to identify lesions with high performance but lacked the true negative evaluation, which lead to difficulty to measure additional metrics.

Even the absence of lesion visually evaluated by the absence of masks in the DeepCADD output is not enough to calculate the model numerical performance with excellence. In fact, in

Figure 39 – Four different angiographic frames provided by DeepCADD after specialist suggestions



Source: Elaborated by the author.

future research, we can also propose a new numerical way to measure the model performance in this specific context.

MCC was suggested in the literature as superior for binary problems. Nevertheless, the insignificant MCC values were outperformed by the Cohen's Kappa numerical results, which were able to identify that DeepCADD suggestions increased the correlation between the specialists.

Also, the identification of different perspectives between the specialist's evaluation and the common must ask for dynamic CAD analysis. This aspect is especially important and must be further explored, once the frame-specific analysis is helpful but must be associated with the dynamic analysis. Given the current DeepCADD structure, we expect that dynamic evaluation can be easily implemented in the future. Refer to Appendix A.

DeepCADD presented promising results, especially in the classical artery narrowing, on images with high contrast and a high presence of fluorescent dye. DeepCADD improvements can also be reached with a bigger labeled dataset, in which we can explore the local hospital partnership. The current partnership will improve the performance mainly in lesions close to bifurcations and the noncalcified which caused the most of false negatives.

6 CONCLUSION

In this work, we considered the angiography protocol and one of the manual steps necessary to obtain the CAD assessment. We detailed the CAD context and the current literature advances and gaps. The SLR has shown that image-based approaches are essential to the correct lesion assessment, once it produces important anatomical metrics for the functional lesion assessment. During the experiments, the use of neural network architecture has shown to be powerful to learn complex angiography image features. Current results also show that a CNN for classification model is helpful and require further improvements to produce a better model performance. Once the model reaches better results, object detection will be evaluated.

We then introduced the DeepCADD architecture to automatically determine multi-vessel CAD using frame-based object detection. We provided theoretical results, showing that DeepCADD outperforms the current literature in the sensitivity index, which makes DeepCADD a good candidate for the daily patient triage as a screening tool. Our theoretical findings are supported by a series of experimental results based on specialists validation and statistical analysis.

The DeepCADD model seeks to address the gaps identified in the related works identified in the SLR. In this sense, the development of the DeepCADD will result in scientific contributions that represent additions to state-of-the-art and use in the daily diagnosis. The main expected contributions of this work are highlighted below:

The first contribution is the DeepCADD architecture for CAD classification and automatic identification in angiographies. DeepCADD is based on small datasets and transfers learning, reaching performance comparable with the gold standard specialist evaluation. Therefore, DeepCADD still requires improvements in the detection of complex lesion types, such as the presence of bifurcation and noncalcified lesions, so the use of bigger labeled datasets can improve the architecture performance and provide even more insights to the daily angiography protocol. We were also able to raise the question about standardization, which can be better explored and confirmed as helpful or not in the daily tool use.

The second contribution, as an output from SLR, we provided the most suitable performance metrics compared to the state of the art for CAD detection in angiographies as well as the use of transfer learning and artery segments dataset for DeepCADD feature extractors, which added more significant variability of information for whole coronary angiography analysis.

The third contribution, DeepCADD was able to outperform current literature in the multi-vessel CAD identification. We remark that DeepCADD, using mask-based object detection, increases the convergence between specialists' assessments, which can be important to improve the current angiography visual assessment protocol. The use of a mask-based system differs from current CAD detection literature and produces more visual insights to the operator. We explored pixel-level object detection and we identified a powerful tool to detect difficult regions or to support border cases.

Besides the contributions, DeepCADD is a frame-based detection tool limited for RCA seg-

ments and with calcified plaque lesions. The use of bigger datasets and labeled angiographies will enhance the architecture's performance. Hence, DeepCADD is also limited for lesions that require more frames for dynamic analysis. Therefore, DeepCADD can be used in future studies for multi-frame analysis.

6.1 Publications

As partial contributions throughout the research, some articles were produced for publication in journals and events. The published and submitted articles are:

- Published articles:

  1. FREITAS, Samuel; RAMOS, Gabriel; SCHMITH, Jean; COSTA , Cristiano. Nodal Analysis for Coronary Artery Ischemia Diagnosis. XX Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS), 2020. Porto Alegre: Sociedade Brasileira de Computação, p. 262-272. Available at: https://doi.org/10.5753/sbcas.2020.11519.

  2. FREITAS, Samuel A.; COSTA, Cristiano A. da; RAMOS, Gabriel de O.. Coronary Artery Disease Automatic Classification. ESCOLA REGIONAL DE COMPUTAÇÃO APLICADA À SAÚDE (ERCAS), 8. , 2021, São Paulo. Anais. Porto Alegre: Sociedade Brasileira de Computação, 2021 . p. 26-29. DOI: https://doi.org/10.5753/ercas.2021.17431

  3. FREITAS, Samuel; NIENOW, Débora; DA COSTA, Cristiano A.; RAMOS, Gabriel de O. Functional Coronary Artery Assessment: a Systematic Literature Review. Wiener klinische Wochenschrift. 2021. DOI: https://doi.org/10.1007/s00508-021-01970-4.

- Published abstracts:

  1. NIENOW, Débora; FREITAS, Samuel; SCHMITH, Jean. Software para auxílio no diagnóstico de estenoses em artérias coronárias. Arq. Bras. Cardiol: 113 (2 sup l. 1) p. 254. 2019. Available at: http://publicacoes.cardiol.br/portal/abc/portugues/2019/v11303/pdf/sbc-2019-programa-trabalhos.pdf.

- Submitted articles:

  1. FREITAS, Samuel; ZEISER, Felipe; COSTA, Cristiano A. da; RAMOS, Gabriel de O. DeepCADD: a Novel Deep Learning Architecture for Coronary Artery Disease Automatic Detection. IJCNN 2022.

6.2 Future Work

The DeepCADD architecture presents many opportunities for improvement. As future work, we would like to further investigate the use of different CNN architectures as DeepCADD backbone, mainly the branch models which are not currently supported.

The lesion types limitation caused by the small number of labeled is also a important DeepCADD challenge we expect to mitigate with a bigger dataset, the same architecture and more specialists validations. The current heat map analysis would be deeper explored in order to reach an explainability DeepCADD module as result in better CAD suggestions.

An important aspect to consider in further studies refers to the dynamic CAD detection, which can be achieved using the frame-based DeepCADD tool in the angiography films without the need of additional modules. The current angiography protocol still have further manual interactions, which can be replaced by automatic further DeepCADD steps.

Finally, we look forward to extend our DeepCADD architecture to support LCA as well as the RCA, making it a more widespread tool for whole coronary artery system assessment.

REFERENCES

AAKKARA, H.; AAISUEB, A.; AEELANUPAB, A. Comparing Classifiers for the Prediction of the Stenosis of Coronary Artery. **International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)**, [S.l.], p. 747–750, 2020.

ADJEDJ, J. et al. Invasive Measures of Myocardial Perfusion and Ischemia. **Progress in Cardiovascular Diseases**, [S.l.], v. 57, n. 6, p. 555–565, 2015.

AGUJETAS, R. et al. Numerical analysis of the pressure drop across highly-eccentric coronary stenoses: application to the calculation of the fractional flow reserve. **BioMedical Engineering Online**, [S.l.], v. 17, n. 1, p. 1–23, 2018.

AL'AREF, S. J. et al. A Boosted Ensemble Algorithm for Determination of Plaque Stability in High-Risk Patients on Coronary CTA. **JACC: Cardiovascular Imaging**, [S.l.], v. 13, n. 10, p. 2162–2173, 2020.

ANNADURAI, S. **Fundamentals of digital image processing**. [S.l.]: Pearson, 2007. (Always learning).

ANTCZAK, K.; LIBERADZKI, Ł. Stenosis Detection with Deep Convolutional Neural Networks. **MATEC Web of Conferences**, [S.l.], v. 210, p. 1–6, 2018.

ATHANASIOU, L. S. et al. 3 - quantitative coronary angiography methods. In: **Atherosclerotic plaque characterization methods based on coronary imaging**. Oxford: Academic Press, 2017. p. 49–69.

BAIBHAV, B. et al. Role of invasive functional assessment in surgical revascularization of coronary artery disease. **Circulation**, [S.l.], v. 137, n. 16, p. 1731–1739, 2018.

BAUMANN, S. et al. Correlation of machine learning computed tomography-based fractional flow reserve with instantaneous wave free ratio to detect hemodynamically significant coronary stenosis. **Clinical Research in Cardiology**, [S.l.], v. 109, n. 6, p. 735–745, 2020.

BAYDOUN, H. et al. Revascularization of Left Main Coronary Artery. **Cardiovascular Revascularization Medicine**, [S.l.], v. 20, n. 11, p. 1014–1019, 2019.

BEN-DAVID, A. Comparison of classification accuracy using cohen's weighted kappa. **Expert Systems with Applications**, [S.l.], v. 34, n. 2, p. 825–832, 2008.

BENENATI, S. et al. Invasive "in the cath-lab" assessment of myocardial ischemia in patients with coronary artery disease: when does the gold standard not apply? **Cardiovascular Revascularization Medicine**, [S.l.], v. 19, n. 3, p. 362–372, 2018.

BENHAR, H.; IDRI, A.; L Fernández-Alemán, J. Data preprocessing for heart disease classification: a systematic literature review. **Computer Methods and Programs in Biomedicine**, [S.l.], v. 195, 2020.

BENTON, S. M. et al. Noninvasive derivation of fractional flow reserve from coronary computed tomographic angiography: a review. **Journal of Thoracic Imaging**, [S.l.], v. 33, n. 2, p. 88–96, 2018.

BHATT, D. L. Fractional flow reserve measurement for the physiological assessment of coronary artery stenosis severity. **JAMA - Journal of the American Medical Association**, [S.l.], v. 320, n. 12, p. 1275–1276, 2018.

BRAUN, T. et al. Detection of myocardial ischemia due to clinically asymptomatic coronary artery stenosis at rest using supervised artificial intelligence-enabled vectorcardiography – A five-fold cross validation of accuracy. **Journal of Electrocardiology**, [S.l.], v. 59, p. 100–105, 2020.

BURGER, W.; BURGE, M. J. **Principles of digital image processing**: advanced methods. 1. ed. [S.l.]: Springer, 2009.

CAI, Z.; VASCONCELOS, N. Cascade r-cnn: delving into high quality object detection. In: IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2018., 2018. **Anais...** [S.l.: s.n.], 2018. p. 6154–6162.

CARSON, J. M. et al. Non-invasive coronary CT angiography-derived fractional flow reserve: a benchmark study comparing the diagnostic performance of four different computational methodologies. **International Journal for Numerical Methods in Biomedical Engineering**, [S.l.], v. 35, n. 10, p. 1–22, 2019.

CARUSO, D. et al. Dynamic CT myocardial perfusion imaging. **European Journal of Radiology**, [S.l.], v. 85, n. 10, p. 1893–1899, 2016.

CESARO, A. et al. **Functional assessment of coronary stenosis: an overview of available techniques. is quantitative flow ratio a step to the future?** [S.l.]: Taylor & Francis, 2018. 951–962 p. v. 16, n. 12.

CHAHAL, K. S.; DEY, K. **A survey of modern object detection literature using deep learning**. 2018.

CHICCO, D.; WARRENS, M. J.; JURMAN, G. The matthews correlation coefficient (mcc) is more informative than cohen's kappa and brier score in binary classification assessment. **IEEE Access**, [S.l.], v. 9, p. 78368–78381, 2021.

CHOLLET, F. **Xception**: deep learning with depthwise separable convolutions. 2017.

COENEN, A. et al. Integrating CT Myocardial Perfusion and CT-FFR in the Work-Up of Coronary Artery Disease. **JACC: Cardiovascular Imaging**, [S.l.], v. 10, n. 7, p. 760–770, 2017.

CONG, C. et al. Deep learning-based end-to-end automated stenosis classification and localization on catheter coronary angiography. **medRxiv**, [S.l.], 2021.

CONTE, E. et al. FFRCT and CT perfusion: a review on the evaluation of functional impact of coronary artery stenosis by cardiac ct. **International Journal of Cardiology**, [S.l.], v. 300, p. 289–296., 2020.

COPPIN, B. **Inteligência artificial**. 1. ed. Rio de Janeiro: LTC, 2010. 636 p.

CRUZ-ACEVES, I.; HERNANDEZ-AGUIRRE, A.; VALDEZ, S. I. On the performance of nature inspired algorithms for the automatic segmentation of coronary arteries using gaussian matched filters. **Applied Soft Computing**, [S.l.], v. 46, p. 665–676, 2016.

DANAD, I. et al. Diagnostic performance of cardiac imaging methods to diagnose ischaemia-causing coronary artery disease when directly compared with fractional flow reserve as a reference standard: a meta-analysis. **European Heart Journal**, [S.l.], v. 38, n. 13, p. 991–998, 2017.

DANILOV, V. V. et al. Real-time coronary artery stenosis detection based on modern neural networks. **Nature scientific report**, [S.l.], 2021.

De Waard, G. A. et al. Coronary autoregulation and assessment of stenosis severity without pharmacological vasodilation. **European Heart Journal**, [S.l.], v. 39, n. 46, p. 4062–4071, 2018.

DEY, D. et al. Integrated prediction of lesion-specific ischaemia from quantitative coronary CT angiography using machine learning: a multicentre study. **European Radiology**, [S.l.], v. 28, n. 6, p. 2655–2664, 2018.

DEY, D. et al. Artificial Intelligence in Cardiovascular Imaging: jacc state-of-the-art review. **Journal of the American College of Cardiology**, [S.l.], v. 73, n. 11, p. 1317–1335, 2019.

Di Serafino, L. et al. FFR prediction model based on conventional quantitative coronary angiography and the amount of myocardium subtended by an intermediate coronary artery stenosis. **International Journal of Cardiology**, [S.l.], v. 223, p. 340–344, 2016.

Dumoulin, V.; Visin, F. A guide to convolution arithmetic for deep learning. **ArXiv e-prints**, [S.l.], mar 2016.

EBERHARD, M. et al. Machine learning-based CT fractional flow reserve assessment in acute chest pain: first experience. **Cardiovascular diagnosis and therapy**, [S.l.], v. 10, n. 4, p. 820–830, 2020.

EVERINGHAM, M. et al. The pascal visual object classes (voc) challenge. **Int. J. Comput. Vision**, USA, v. 88, n. 2, p. 303–338, June 2010.

FAWCETT, T. An introduction to roc analysis. **Pattern Recognition Letters**, [S.l.], v. 27, p. 861–874, 2006.

FRANGI, A. F. et al. "Multiscale Vessel Enhancement Filtering", Medical Image Computing and Computer-Assisted Interventation. , [S.l.], 1998.

FREITAS, S. A. et al. Functional coronary artery assessment: a systematic literature review. **Wiener Klinische Wochenschrift**, [S.l.], 2021.

FREITAS, S. et al. Coronary artery disease automatic classification. **Escola Regional de Computação Aplicada à Saúde (ERCAS)**, Porto Alegre, p. 26–29, 2021.

FUKUOKA, S. et al. Clinical usefulness of instantaneous wave-free ratio for the evaluation of coronary artery lesion with prior myocardial infarction: a multi-center study. **IJC Heart and Vasculature.**, [S.l.], v. 26, 2020.

G., P. et al. Diagnostic accuracy of simultaneous evaluation of coronary arteries and myocardial perfusion with single stress cardiac computed tomography acquisition compared to invasive coronary angiography plus invasive fractional flow reserve. **International Journal of Cardiology**, [S.l.], v. 273, p. 263–268, 2018.

GHOSH, S. et al. Understanding deep learning techniques for image segmentation. **ACM Comput. Surv.**, New York, NY, USA, v. 52, n. 4, Aug. 2019.

GONZALEZ, R. C.; WOODS, R. E. **Digital image processing**. 4. ed. [S.l.]: Pearson, 2018.

GREEN, M. et al. 3-d neural denoising for low-dose coronary ct angiography (ccta). **Computerized Medical Imaging and Graphics**, [S.l.], v. 70, p. 185–191, 2018.

GUPTA, V. et al. Performance of a Deep Neural Network Algorithm Based on a Small Medical Image Dataset: incremental impact of 3d-to-2d reformation combined with novel data augmentation, photometric conversion, or transfer learning. **Journal of Digital Imaging**, [S.l.], p. 431–438, 2019.

GUYTON, A.; HALL, J. **Textbook of medical physiology**. [S.l.]: Elsevier Saunders, 2006.

GÉRON, A. Hands-on machine learning with scikit-learn and tensorflow: concepts. , [S.l.], 2017.

HAE, H. et al. Machine learning assessment of myocardial ischemia using angiography: development and retrospective validation. **PLoS Medicine**, [S.l.], v. 15, n. 11, p. 1–19, 2018.

HAFIZ, A. M.; BHAT, G. M. A survey on instance segmentation: state of the art. **International Journal of Multimedia Information Retrieval**, [S.l.], v. 9, p. 171–189, 9 2020.

HAN, D. et al. Deep learning analysis in coronary computed tomographic angiography imaging for the assessment of patients with coronary artery stenosis. **Computer Methods and Programs in Biomedicine**, [S.l.], v. 196, 2020.

HE, C. et al. Automated classification of coronary plaque calcification in OCT pullbacks with 3D deep neural networks. **Journal of Biomedical Optics**, [S.l.], v. 25, n. 09, p. 1–13, 2020.

HE, K. et al. Deep residual learning for image recognition. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), 2016. **Proceedings...** [S.l.: s.n.], 2016.

HE, K. et al. Mask r-cnn. In: IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV), 2017., 2017. **Anais...** [S.l.: s.n.], 2017. p. 2980–2988.

HE, K. et al. **Mask r-cnn**. 2018.

HEATON, J. et al. Deep learning. , [S.l.], p. 305–307, 2018.

HIDEO-KAJITA, A. et al. Atualização sobre tecnologias fisiológicas baseadas em angiografia coronariana - update on coronary angiography-based physiology technologies. , [S.l.], 2019.

HU, W. et al. Novel radiomics features from CCTA images for the functional evaluation of significant ischaemic lesions based on the coronary fractional flow reserve score. **International Journal of Cardiovascular Imaging**, [S.l.], v. 36, n. 10, p. 2039–2050, 2020.

HWANG, D.; LEE, J. M.; KOO, B. K. Physiologic assessment of coronary artery disease: focus on fractional flow reserve. **Korean Journal of Radiology**, [S.l.], v. 17, n. 3, p. 307–320, 2016.

IHDAYHID, A. R. et al. A Practical Guide for Fractional Flow Reserve Guided Revascularisation. **Heart Lung and Circulation**, [S.l.], v. 27, n. 4, p. 406–419, 2018.

JAFARIAN, K. et al. Automating detection and localization of myocardial infarction using shallow and end-to-end deep neural networks. **Applied Soft Computing**, [S.l.], v. 93, p. 106383, 2020.

JOHNSON, N. P. et al. Invasive FFR and Noninvasive CFR in the Evaluation of Ischemia: what is the future? **Journal of the American College of Cardiology**, [S.l.], v. 67, n. 23, p. 2772–2788, 2016.

Kaplan Berkaya, S.; Ak Sivrikoz, I.; GUNAL, S. Classification models for SPECT myocardial perfusion imaging. **Computers in Biology and Medicine**, [S.l.], v. 123, n. February, p. 103893, 2020.

KAWASAKI, T. et al. Evaluation of Significant Coronary Artery Disease Based on CT Fractional Flow Reserve and Plaque Characteristics Using Random Forest Analysis in Machine Learning. **Academic Radiology**, [S.l.], p. 1–9, 2020.

KHAN, A. et al. A survey of the recent architectures of deep convolutional neural networks. **Artificial Intelligence Review**, [S.l.], p. 1–70, 2020.

KHAN, A. I.; SHAH, J. L.; BHAT, M. M. Coronet: a deep neural network for detection and diagnosis of covid-19 from chest x-ray images. **Computer Methods and Programs in Biomedicine**, [S.l.], v. 196, p. 105581, 2020.

KIM, H. Y. et al. Physiological Severity of Coronary Artery Stenosis Depends on the Amount of Myocardial Mass Subtended by the Coronary Artery. **JACC: Cardiovascular Interventions**, [S.l.], v. 9, n. 15, p. 1548–1560, 2016.

KO, S. M. et al. Multi-modality imaging for the assessment of myocardial perfusion with emphasis on stress perfusion CT and MR imaging. **International Journal of Cardiovascular Imaging**, [S.l.], v. 31, p. 1–21, 2015.

KOTECHA, T. et al. Automated Pixel-Wise Quantitative Myocardial Perfusion Mapping by CMR to Detect Obstructive Coronary Artery Disease and Coronary Microvascular Dysfunction: validation against invasive coronary physiology. **JACC: Cardiovascular Imaging**, [S.l.], v. 12, n. 10, p. 1958–1969, 2019.

KULATHILAKE, K. A. S. H. et al. A segmentation method for extraction of main arteries from coronary cine-angiograms. , [S.l.], p. 10–11, 2010.

KWAN, A. C. et al. Prediction of revascularization by coronary CT angiography using a machine learning ischemia risk score. **European Radiology**, [S.l.], 2020.

LECUN, Y. et al. Convolutional networks and applications in vision. **ISCAS 2010 - 2010 IEEE International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems**, [S.l.], p. 253–256, 2010.

LEE, C.-Y.; GALLAGHER, P.; TU, Z. Generalizing pooling functions in convolutional neural networks: mixed, gated, and tree. **In: Artificial Intelligence and Statistics**, [S.l.], 2016.

LEE, J. G. et al. Intravascular ultrasound-based machine learning for predicting fractional flow reserve in intermediate coronary artery lesions. **Atherosclerosis**, [S.l.], v. 292, n. September 2019, p. 171–177, 2020.

LEONE, A. M. et al. Contrast Fractional Flow Reserve (cFFR): a pragmatic response to the call for simplification of invasive functional assessment. **International Journal of Cardiology**, [S.l.], v. 268, p. 45–50, 2018.

LI, S. et al. The impact of iterative reconstruction algorithms on machine learning-based coronary CT angiography-derived fractional flow reserve (CT-FFRML) values. **International Journal of Cardiovascular Imaging**, [S.l.], v. 36, n. 6, p. 1177–1185, 2020.

LI, Z. et al. Automatic coronary artery segmentation based on multi-domains remapping and quantile regression in angiographies. **Computerized Medical Imaging and Graphics**, [S.l.], v. 54, p. 55–66, 2016.

LIBBY, P. Inflammation in atherosclerosis. **Nature**, [S.l.], v. 420, p. 868–874, 2002.

LIN, T.-Y. et al. **Microsoft coco**: common objects in context. 2015.

LOEWE, C. Hemodynamically significant coronary stenosis: detection with ct myocardial perfusion imaging versus machine learning coronary ct fractional flow reserve. **Radiology**, [S.l.], v. 293, n. 2, p. 315–316, 2019.

LONCARIC, F. et al. Integration of artificial intelligence into clinical patient management: focus on cardiac imaging. **Revista Española de Cardiología (English Edition)**, [S.l.], n. x, 2020.

MANNIL, M. et al. Artificial Intelligence and Texture Analysis in Cardiac Imaging. **Current Cardiology Reports**, [S.l.], v. 22, n. 11, p. 1–9, 2020.

MARTINS, J. L. et al. Fractional flow reserve-guided strategy in acute coronary syndrome. A systematic review and meta-analysis. **Arquivos Brasileiros de Cardiologia**, [S.l.], v. 111, n. 4, p. 542–550, 2018.

MICHAIL, M. et al. Application of the DILEMMA score to improve lesion selection for invasive physiological assessment. **Catheterization and Cardiovascular Interventions**, [S.l.], v. 94, n. 3, p. E96–E103, 2019.

MIN, J. K. et al. Noninvasive fractional flow reserve derived from coronary CT angiography clinical data and scientific principles. **JACC: Cardiovascular Imaging**, [S.l.], v. 8, n. 10, p. 1209–1222, 2015.

MOERTL, D. et al. Comparison of quantitative coronary angiography with visual. , [S.l.], p. 145–148, 1995.

MONTALESCOT, G. et al. Guidelines on the management of stable coronary artery disease. **British Journal of Cardiac Nursing**, [S.l.], v. 8, n. 11, p. 519–520, 2013.

MOON, J. H. et al. Automatic stenosis recognition from coronary angiography using convolutional neural networks. **Computer Methods and Programs in Biomedicine**, [S.l.], v. 198, p. 105819, 2021.

NILSSON, L. et al. Neutrophil/lymphocyte ratio is associated with non-calcified plaque burden in patients with coronary artery disease. **PLOS ONE**, [S.l.], v. 9, n. 9, p. 1–8, 09 2014.

NOUS, F. M. et al. Impact of machine-learning CT-derived fractional flow reserve for the diagnosis and management of coronary artery disease in the randomized CRESCENT trials. **European Radiology**, [S.l.], v. 30, n. 7, p. 3692–3701, 2020.

NUDI, A. et al. Promises and pitfalls of relying on angiography-derived indexes to identify myocardial ischemia: a tale of romulus and remus. **Journal of Nuclear Cardiology**, [S.l.], 2020.

OVALLE-MAGALLANES, E. et al. Transfer learning for stenosis detection in X-ray Coronary Angiography. **Mathematics**, [S.l.], v. 8, n. 9, p. 1–20, 2020.

PANG, C. L. et al. Determining the haemodynamic significance of arterial stenosis: the relationship between ct angiography, computational fluid dynamics, and non-invasive fractional flow reserve. **Clinical Radiology**, [S.l.], v. 71, n. 8, p. 750–757, 2016.

PANG, J. et al. Libra r-cnn: towards balanced learning for object detection. In: IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), 2019., 2019. **Anais...** [S.l.: s.n.], 2019. p. 821–830.

PANG, K. et al. Stenosis-detnet: sequence consistency-based stenosis detection for x-ray coronary angiography. **Computerized Medical Imaging and Graphics**, [S.l.], v. 89, p. 101900, 2021.

PARK, J. Y.; LERMAN, A.; HERRMANN, J. Use of fractional flow reserve in patients with coronary artery disease: the right choice for the right outcome. **Trends in Cardiovascular Medicine**, [S.l.], v. 27, n. 2, p. 106–120, 2017.

PETTICREW, M.; ROBERTS, H. **Systematic reviews in the social sciences**: a practical guide. [S.l.]: John Wiley & Sons, 2008.

PODGORSAK, A. R. et al. Initial evaluation of a convolutional neural network used for noninvasive assessment of coronary artery disease severity from coronary computed tomography angiography data. **Medical Physics**, [S.l.], v. 47, n. 9, p. 3996–4004, 2020.

POLLOCK, J. D. et al. **Physiology, cardiovascular hemodynamics**. [S.l.: s.n.], 2020.

PRATT, W. **Introduction to digital image processing**. [S.l.]: Taylor & Francis, 2013.

PYXARAS, S. A. et al. Invasive assessment of coronary artery disease. **Journal of Nuclear Cardiology**, [S.l.], v. 25, n. 3, p. 860–871, 2018.

QURESHI, A. M.; AGRAWAL, H. Catheter-based anatomic and functional assessment of coronary arteries in anomalous aortic origin of a coronary artery, myocardial bridges and Kawasaki disease. **Congenital Heart Disease**, [S.l.], v. 12, n. 5, p. 615–618, 2017.

RAHIMZADEH, M.; ATTAR, A. A modified deep convolutional neural network for detecting covid-19 and pneumonia from chest x-ray images based on the concatenation of xception and resnet50v2. **Informatics in Medicine Unlocked**, [S.l.], v. 19, p. 100360, 2020.

RAJPURKAR, P. et al. **Chexnet**: radiologist-level pneumonia detection on chest x-rays with deep learning. 2017.

RAMASAMY, A. et al. Optical coherence tomography enables more accurate detection of functionally significant intermediate non-left main coronary artery stenoses than intravascular ultrasound: a meta-analysis of 6919 patients and 7537 lesions. **International Journal of Cardiology**, [S.l.], v. 301, p. 226–234, 2020.

REN, S. et al. Faster r-cnn: towards real-time object detection with region proposal networks. **Advances in neural information processing systems**, [S.l.], v. 28, p. 91–99, 2015.

RODRIGUES, D. L. et al. Automated detection of coronary artery stenosis in x-ray angiography using deep neural networks. , [S.l.], 3 2021.

ROOS, C. J. et al. **Int J Cardiovasc Imaging**, [S.l.], 2013.

RUSSELL, S. et al. **Artificial intelligence**: a modern approach. [S.l.]: Prentice Hall, 2010. (Prentice Hall series in artificial intelligence).

RUUSKA, S. et al. Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. **Behavioural Processes**, [S.l.], v. 148, p. 56–62, 2018.

SEIFERT, R. et al. AI and Machine Learning in Nuclear Medicine: future perspectives. **Seminars in Nuclear Medicine**, [S.l.], 2020.

SHAO, Y. E.; HOU, C.-D.; CHIU, C.-C. Hybrid intelligent modeling schemes for heart disease classification. **Applied Soft Computing**, [S.l.], v. 14, p. 47–52, 2014. Special issue on hybrid intelligent methods for health technologies.

SHU, Z. Y. et al. Predicting Chronic Myocardial Ischemia Using CCTA-Based Radiomics Machine Learning Nomogram. **Journal of Nuclear Cardiology**, [S.l.], n. 158, 2020.

SIOGKAS, P. K. et al. Computational estimation of the severity of coronary lesions with intravascular ultrasound images: a pilot study. **Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS**, [S.l.], v. 2016-Octob, p. 2664–2667, 2016.

SIOGKAS, P. K. et al. Computational estimation of the hemodynamic significance of coronary stenoses in arterial branches deriving from CCTA: a proof-of-concept study. **Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS**, [S.l.], p. 1348–1351, 2017.

STAPIĆ, Z. et al. Performing systematic literature review in software engineering. In: CECIIS 2012-23RD INTERNATIONAL CONFERENCE, 2012. **Anais...** [S.l.: s.n.], 2012.

STEGEHUIS, V. E. et al. Fractional Flow Reserve or Coronary Flow Reserve for the Assessment of Myocardial Perfusion: implications of ffr as an imperfect reference standard for myocardial ischemia. **Current Cardiology Reports**, [S.l.], v. 20, n. 9, 2018.

TAKAHASHI, K. et al. TCT-326 Diagnostic Performance of Angiography-Based Quantitative Flow Ratio With Respect to Fractional Flow Reserve Derived From Computed Tomography Angiography: insight from the syntax iii trial. **Journal of the American College of Cardiology**, [S.l.], v. 74, n. 13, p. B324, 2019.

TANIGAKI, T. et al. QFR Versus FFR Derived From Computed Tomography for Functional Assessment of Coronary Artery Stenosis. **JACC: Cardiovascular Interventions**, [S.l.], v. 12, n. 20, p. 2050–2059, 2019.

TEBALDI, M. et al. Evolving Routine Standards in Invasive Hemodynamic Assessment of Coronary Stenosis: the nationwide italian sici-gise cross-sectional eris study. **JACC: Cardiovascular Interventions**, [S.l.], v. 11, n. 15, p. 1482–1491, 2018.

TESCHE, C.; GRAY, H. N. Machine Learning and Deep Neural Networks Applications in Coronary Flow Assessment: the case of computed tomography fractional flow reserve. **Journal of thoracic imaging**, [S.l.], v. 35, n. May, p. S66–S71, 2020.

TIES, D. et al. Computational quantitative flow ratio to assess functional severity of coronary artery stenosis. **International Journal of Cardiology**, [S.l.], v. 271, p. 36–41, 2018.

TRAN, D. M.; NGUYEN, M. T.; LEE, S. W. Machine learning based evaluation of functional index for coronary lesion severity. **ACM International Conference Proceeding Series**, [S.l.], p. 1–4, 2018.

TSOMPOU, P. I. et al. Comparison of 3D reconstruction methods based on different cardiovascular imaging: a study of multimodality reconstruction method. **Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS**, [S.l.], v. 2018-July, p. 899–902, 2018.

UUS, A. et al. Assessment of stenosis introduced flow resistance in CCTA-reconstructed coronary arteries. **2015 22nd International Conference on Systems, Signals and Image Processing - Proceedings of IWSSIP 2015**, [S.l.], p. 313–320, 2015.

VARATHARASAN, V. et al. Improving learning effectiveness for object detection and classification in cluttered backgrounds. In: WORKSHOP ON RESEARCH, EDUCATION AND DEVELOPMENT OF UNMANNED AERIAL SYSTEMS (RED UAS), 2019., 2019. **Anais...** [S.l.: s.n.], 2019. p. 78–85.

VIEIRA, S. M.; KAYMAK, U.; SOUSA, J. M. C. Cohen's kappa coefficient as a performance measure for feature selection. In: INTERNATIONAL CONFERENCE ON FUZZY SYSTEMS, 2010. **Anais...** [S.l.: s.n.], 2010. p. 1–8.

WANG, J. OCT Image Recognition of Cardiovascular Vulnerable Plaque Based on CNN. **IEEE Access**, [S.l.], v. 8, p. 140767–140776, 2020.

WANG, X. et al. Functional assessment of stenotic coronary artery in 3d geometric reconstruction from fusion of intravascular ultrasound and X-ray angiography. **IEEE Access**, [S.l.], v. 6, p. 53330–53341, 2018.

WEISS, K.; KHOSHGOFTAAR, T.; WANG, D. A survey of transfer learning. **Big Data**, [S.l.], v. 3, n. 9, 2006.

WESTRA, J. et al. Diagnostic performance of in-procedure angiography-derived quantitative flow reserve compared to pressure-derived fractional flow reserve: the favor ii europe-japan study. , [S.l.], 2018.

WILLIAMS, M. C. et al. Coronary Artery Plaque Characteristics Associated With Adverse Outcomes in the SCOT-HEART Study. **Journal of the American College of Cardiology**, [S.l.], v. 73, n. 3, p. 291–301, 2019.

WOLPERT, D. H.; MACREADY, W. G. No free lunch theorems for optimization. **Trans. Evol. Comp**, [S.l.], v. 1, n. 1, p. 67–82, Apr. 1997.

WONG, K. K.; FORTINO, G.; ABBOTT, D. Deep learning-based cardiovascular image diagnosis: a promising challenge. **Future Generation Computer Systems**, [S.l.], v. 110, p. 802–811, 2020.

World Health Organization. **Cardiovascular diseases (cvds)**. 2017.

WUA, W. et al. Automatic detection of coronary artery stenosis by convolutional neural network with temporal constraint. **Computers in Biology and Medicine**, [S.l.], v. 118, 2020.

XAPLANTERIS, P. et al. Five-year outcomes with pci guided by fractional flow reserve. **New England Journal of Medicine**, [S.l.], v. 379, n. 3, p. 250–259, 2018. PMID: 29785878.

YAZAKI, K. et al. Applicability of 3-dimensional quantitative coronary angiography-derived computed fractional flow reserve for intermediate coronary stenosis. **Circulation Journal**, [S.l.], v. 81, n. 7, p. 988–992, 2017.

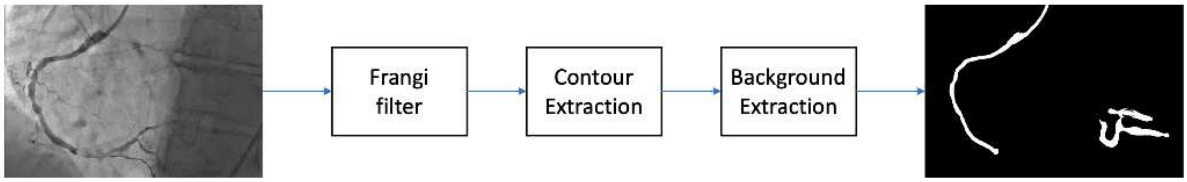
ZREIK, M. et al. Automatic Detection and Characterization of Coronary Artery Plaque and Stenosis using a Recurrent Convolutional Neural Network in Coronary CT Angiography. **arXiv**, [S.l.], n. 2, p. 1–13, 2018.

ZREIK, M. et al. A Recurrent CNN for Automatic Detection and Classification of Coronary Artery Plaque and Stenosis in Coronary CT Angiography. **IEEE Transactions on Medical Imaging**, [S.l.], v. 38, n. 7, p. 1588–1598, 2019.

APPENDIX A – ANGIOGRAPHY SEGMENTATION FOR DYNAMIC EVALUATION

Angiography exams have variable threshold as the X-ray grayscale may vary from acquisition equipment and the acquisition procedure. For this reason, we have created a segmentation pipeline to extract the main coronary artery segment. This strategy works better for RCA due to the less complex artery structure, which does not have additional relevant bifurcations. For this reason, the proposed pipeline has three main steps presented in Figure 40, where the main steps are the Frangi filter and the background extraction.

Figure 40 – Preprocessing steps from original coronary angiography up to main segment extraction



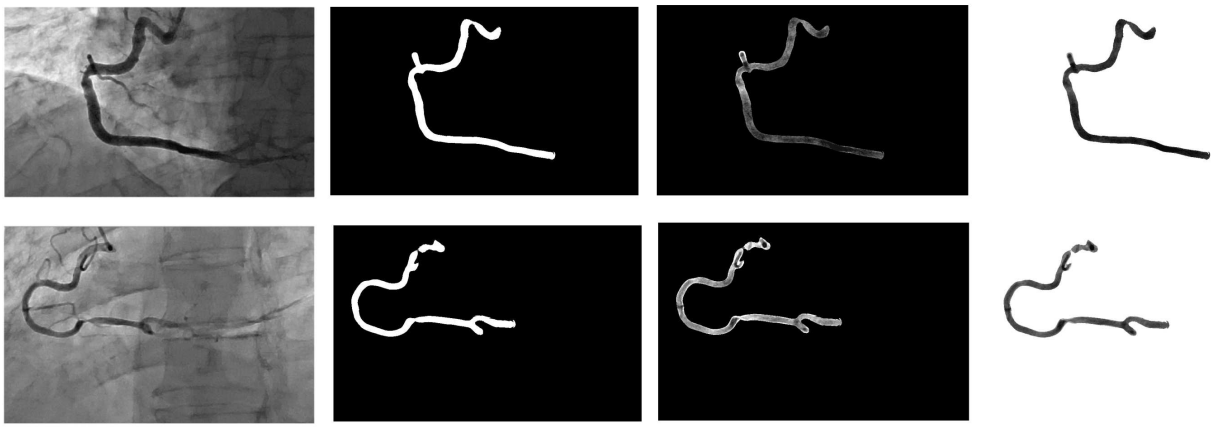
Source: Elaborated by the author.

In this pipeline, the Frangi filter is the main responsible for identifying the coronary segments in the whole exam. Frangi filter has been widely recommended in the literature due to its potential to enhance vessels and arteries (FRANGI et al., 1998). Frangi filter produces a mask with the contour of every possible artery of vessel in the image, and we employed it as a design decision. This step requires also a threshold evaluation for the amount of area present in the mask. For this reason it is also possible to check in Figure 40 that some non-relevant areas can also be detected by the preprocessing module, such as in the right bottom side of output image, which does not present an artery, but a wrong segmentation. Nevertheless, the preprocessing module is able to detect and extract the RCA main segment and reduce the image complexity to be used in the R-CNN classification module.

The preprocessing pipeline is not useful for the proposed model at this time, once we focused in the Segment selection and we defined an architecture which does not require the image segmentation. We plan to use this tool to automate the frame selection in Figure 24 and increase the automation in the angiography protocol, once it will allow the dynamic angiography evaluation.

DeepCADD model is an object detection architecture based on Mask R-CNN which provides classification, bounding boxes and mask to determine CAD location, which matches this Master's Thesis goal to automatically detect CAD. DeepCADD can be used for multi-vessel detection, as it provides instance segmentation. DeepCADD detection is based in anatomical RCA data, such as the MLD and %DS, and we aim to reduce the need of additional invasive tools for CAD significance determination. For the preprocessing step we wanted to test the use of image processing features to extract the main coronary artery segment in RCA angiographies. Figure 41 presents some results on segmenting the RCA in angiographies.

Figure 41 – Angiography segmentation results, with two different RCA from different incidences with the main artery segment enhanced using mask and either black and white background



Source: Elaborated by the author.

We reached these results using mainly the Frangi filter, adaptative thresholding, and contour finding. From Figure 41, it is possible to identify the segmentation sequence, which includes the original image in the first image of each line (left side), followed by the segmentation mask extracted from the original image, followed by a bitwise operation between the original angiography and mask with a black background, and finally the original angiography with a white background (right side). We intended to use this segmentation to train the DeepCADD object detection module, but we noticed that it was not necessary.

The coronary artery segmentation is helpful to identify the main coronary artery segment of RCA. We choose the RCA due to its smaller anatomical variability, which was a design decision and limitation. This segmentation depends on the clear artery highlight caused by the injection of contrast in the interior of coronary artery walls. This technique identifies the coronary segment and can be used in further studies.

In a prospective analysis, we decided to leave these results to support future studies in the automation of the frame selection. As the frame selection in the standard protocol looks for the frame with the highest presence of contrast, this image segmentation technique can be helpful in the CAD detection, so we can use it to choose the best frame in the angiographic film.

APPENDIX B – ARCHITECTURES COMPARISON FOR IMAGE CLASSIFICATION

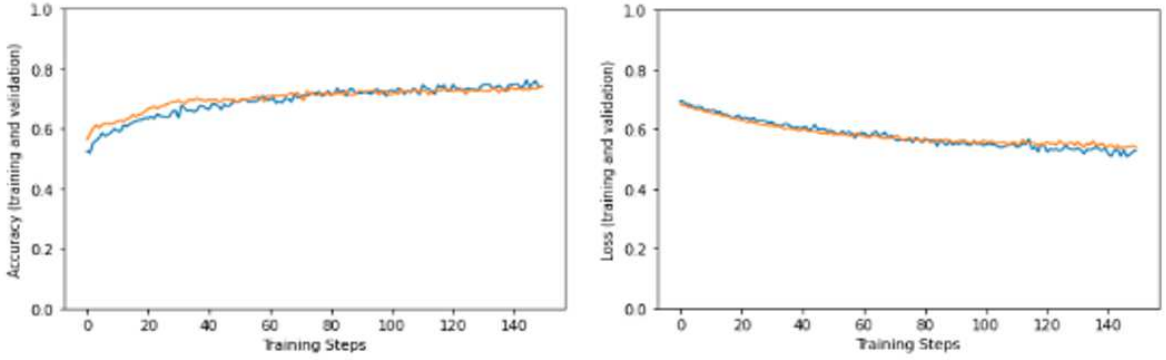
During this study we have experimented several architectures for image classification using the synthetic dataset. Among them, we highlighted the use of Xception, CheXNet, InceptionV3, and ResNet-50 architectures. To pave the way for future studies in the architecture decision for image classification of coronary artery assessment we will present the results obtained during this research considering the learning rate, number of training epochs (epochs), trainable parameters (trainable param.), final dense layers composition (dense layers) and the results as validation loss (loss), and validation accuracy (accuracy).

Table 11 – Hyperparameters optimization for Xception architecture

Base	learning rate	epochs	trainable param.	dense layers	loss	accuracy
Xception	$1e-6$	50	8M	50x1	0.56	0.70
Xception	$1e-6$	50	8M	100x1	0.55	0.73
Xception	$1e-6$	45	all	50x1	0.55	0.75
Xception	$1e-6$	45	8M	50x1	0.63	0.68
Xception	$1e-7$	800	12M	100x100x100x1	0.54	0.73
Xception	$1e-6$	100	8M	25x1	0.63	0.74
Xception	$1e-6$	300	6M	25x1	0.63	0.74
Xception	$1e-6$	300	7M	256x25x1	0.66	0.72
Xception	$1e-6$	1000	530k	256x25x2	0.66	0.69
Xception	$1e-5$	200	530k	256x25x1	0.67	0.69
Xception	$1e-5$	200	530k	256x25x2	0.66	0.69
Xception	$1e-5$	200	530k	256x25x1	0.66	0.69
Xception	$1e-5$	200	530k	256x25x2	0.60	0.72
Xception	$1e-6$	200	530k	256x25x1	0.59	0.71
Xception	$1e-5$	200	4M	256x25x2	0.60	0.70
Xception	$1e-6$	600	4M	256x25x1	1.20	0.70
Xception	$1e-8$	20000	4M	256x25x1	0.62	0.63
Xception	$1e-7$	150	8M	100x100x1	0.54	0.75

Source: Elaborated by the author.

Figure 42 – Best Xception performance on CAD with accuracy and loss graphs



Source: Elaborated by the author.

Results from Table 11 are not too promising to indicate a good performance in the coronary artery disease classification with a maximum accuracy of 0.75 presented in the table line. This best result for Xception architecture is also represented in the Figure 42. In this figure can see that training and validation performance are very close, so it represents a consistent learning curve, but it stop in the maximum of 0.75 accuracy and does not improve the 0.54 loss.

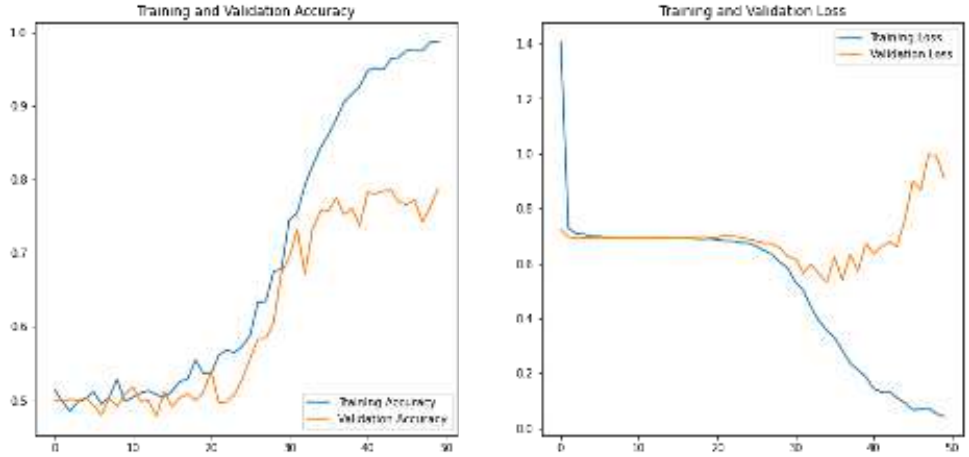
So, Table 12 presents the experiments' results for CheXNet architecture, which also does not present a good performance on our classification task. Even CheXNet presenting good performance in the literature for chest x-ray exams, this architecture does not achieved the expected results. At the time we decided to use this architecture, we expected that it would outperform the other architectures due to the nature of the coronary artery disease exam, which is also a x-ray exam in grayscale. Nevertheless, CheXNet proved not enough to classify coronary artery lesions.

Table 12 – Hyperparameters optimization for CheXNet architecture

Base	learning rate	epochs	trainable param.	dense layer	loss	accuracy
CheXNet	$1e-4$	30	all	100x100x50x1	0.70	0.68
CheXNet	$1e-4$	30	all	100x100x100x1	0.70	0.76
CheXNet	$1e-4$	50	all	50x1	1.00	0.78
CheXNet	$1e-4$	50	all	256x1	1.30	0.65
CheXNet	$1e-4$	100	all	100x100x50x1	2.00	0.73
CheXNet	$1e-5$	50	all	256x2	0.70	0.54
CheXNet	$1e-5$	150	all	100x100x50x1	0.70	0.60
CheXNet	$1e-8$	20000	3M	100x100x50x2	0.69	0.54
CheXNet	$1e-4$	100	all	100x100x100x1	0.63	0.76
CheXNet	$1e-4$	50	all	100x1	1.00	0.80

Source: Elaborated by the author.

Figure 43 – Best CheXNet performance on CAD with accuracy and loss graphs



Source: Elaborated by the author.

Figure 43 presents the best result of CheXNet, where we can notice that ChexNet start with a interesting learning curve stopping close to an accuracy of 0.8 and with the loss diverge

after 30 epochs. According to the literature, we also decided to experiments the InceptionV3 architecture. Table 13 presents the hyperparameter optimization executed using InceptionV3 architecture. Besides it perform a little bit better than Xception and ChexNet, these results still present high loss values, impacting in the overall model performance.

Table 13 – Hyperparameters optimization for InceptionV3 architecture

Base	learning rate	epochs	trainable param.	dense layer	loss	accuracy
InceptionV3	$1e-4$	30	all	100x100x50x1	0.72	0.68
InceptionV3	$1e-4$	30	all	100x100x100x1	0.66	0.76
InceptionV3	$1e-4$	50	all	100x1	0.70	0.80
InceptionV3	$1e-4$	50	all	256x1	0.67	0.65
InceptionV3	$1e-4$	100	all	100x100x50x1	1.00	0.73
InceptionV3	$1e-5$	50	all	256x2	0.56	0.54
InceptionV3	$1e-5$	150	all	100x100x50x1	0.75	0.60
InceptionV3	$1e-4$	50	all	50x1	0.65	0.78

Source: Elaborated by the author.

Finally, we experimented the use of ResNet-50 architecture in the coronary artery lesions classification. Table 14 presents the results using ResNet-50. In this context, as well as the other architectures, we tested many different final dense layers in the model. We noticed that the use of 2 neurons in the final dense layer presented the best results from all the other compositions.

Table 14 – Hyperparameters optimization for ResNet-50 architecture

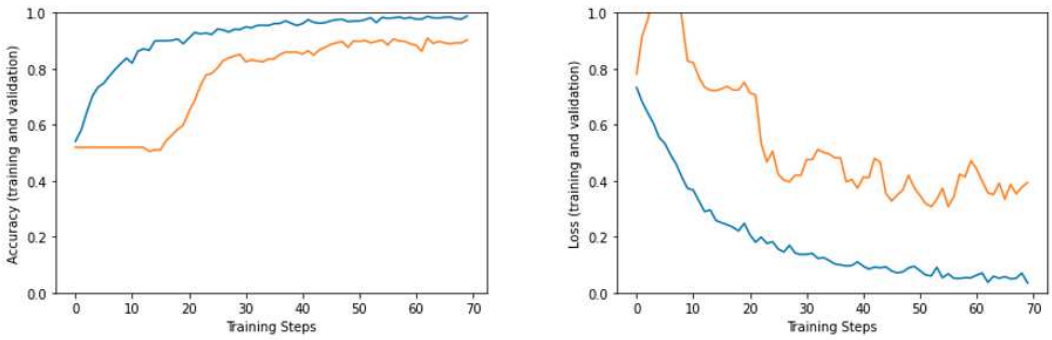
Base	learning rate	epochs	trainable param.	dense layer	loss	accuracy
ResNet-50	$1e-4$	30	all	100x100x50x1	0.64	0.68
ResNet-50	$1e-4$	30	all	100x100x100x1	0.54	0.76
ResNet-50	$1e-4$	50	all	50x2	0.45	0.78
ResNet-50	$1e-4$	50	all	100x1	0.65	0.80
ResNet-50	$1e-4$	50	all	256x1	0.47	0.65
ResNet-50	$1e-4$	100	all	100x100x50x1	1.00	0.73
ResNet-50	$1e-5$	50	all	256x1	0.56	0.54
ResNet-50	$1e-5$	150	all	100x100x50x1	0.43	0.60
ResNet-50	$1e-5$	70	all	2	0.39	0.90

Source: Elaborated by the author.

From the very first ResNet-50 architecture built, we noticed that the achieved accuracy was superior from the others. We experimented several other final dense layers compositions, but the use of 2 neurons showed far better accuracy. Figure 44 presented the learning behavior during the first 70th epochs. Using the ResNet-50 we reported all the experiments using the all 23M trainable parameters, as the use of less trainable parameters did not present significant difference. As we initialized the network with ImageNet weights and the network is trainable in a suitable time, we did not need to reduce the trainable parameters.

The accuracy presented a peak of learning between the 20th and the 30th epochs but continued learning after it, reaching a percentage of 0.90. This accuracy result, together with the

Figure 44 – ResNet-50 validation and training graphs.



Source: Elaborated by the author.

continuous decreasing of loss in both training and validation phases agree and indicates that ResNet-50 learnt the coronary artery lesions context.