ANGIOGRAPHY AND BLOOD PARAMETERS BASED PREDICTION OF FRACTIONAL FLOW RESERVE IN CORONARY ARTERIES USING MACHINE LEARNING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MEHMET NAZİR KAÇAR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
ELECTRICAL AND ELECTRONIC ENGINEERING

JUNE 2026

Approval of the thesis:

ANGIOGRAPHY AND BLOOD PARAMETERS BASED PREDICTION OF FRACTIONAL FLOW RESERVE IN CORONARY ARTERIES USING MACHINE LEARNING

submitted by **Mehmet Nazir Kaçar** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** i**n Electrical and Electronic Engineering, Middle East Technical University** by,

Prof. Dr. Naci Emre Altun
Dean, **Graduate School of Natural and Applied Sciences** _____

Prof. Dr. Nevzat Güneri Gençer
Head of the Department, **Electrical and Electronic Engineering** _____

Prof. Dr. İlkay Ulusoy
Supervisor, **Electrical and Electronic Engineering, METU** _____

Examining Committee Members:

Prof. Dr. Çağrı Yayla
Cardiology, University of Health Sciences _____

Prof. Dr. İlkay Ulusoy
Electrical and Electronic Engineering, METU _____

Assoc. Prof. Dr. Yeşim S. Doğrusöz
Electrical and Electronic Engineering, METU _____

Assoc. Prof. Dr. Erdem Akagündüz
Informatics Institute, METU _____

Prof. Dr. Gökhan Çiçek
Cardiology, University of Health Sciences _____

Date: 26.06.2026

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name Lastname : Mehmet Nazir Kaçar

Signature :

ABSTRACT

ANGIOGRAPHY AND BLOOD PARAMETERS BASED PREDICTION OF FRACTIONAL FLOW RESERVE IN CORONARY ARTERIES USING MACHINE LEARNING

Kaçar, Mehmet Nazir
Doctor of Philosophy, Electrical and Electronic Engineering
Supervisor: Prof. Dr. İlkay Ulusoy

June 2026, 115 pages

Coronary artery disease (CAD) is the most common cardiovascular disease with high morbidity, disability, and societal burden. The CAD occurs due to atherosclerotic occlusions of the coronary arteries. Coronary angiography is the gold standard of clinical diagnosis of coronary heart disease. Measurement of fractional flow reserve (FFR) during routine coronary angiography by using a pressure wire to calculate the ratio between coronary pressure distal to a coronary artery stenosis (after the blockage in the vessel) and aortic pressure aids clinicians to make decisions in certain circumstances. However, FFR is an invasive, complicated and expensive process and has to be performed during angiography. Several methods of FFR estimation by angiography without requiring the instrumentation of the coronary artery have been developed. These methods often employ various approaches to fluid dynamics computation (CFD). Complexity of CFD process has led to various assumptions regarding boundary conditions and blood physical properties. Various machine learning (ML) models are being applied to predict FFR using angiography images. The main limitation of angiography-based predictions is its failure to consider the

physical properties of the fluid being measured: blood. Because FFR measures blood flow impairment, it is inherently affected by this rheological behaviour. We propose the use of blood parameters with X-ray coronary angiography images to improve prediction, and a novel model that combines blood parameters with angiography videos. We retrospectively collected patient data and experimentally showed that adding blood parameters improves prediction.

Keywords: Angiography, Machine Learning, Coronary Artery Disease, Fractional Flow Reserve, Blood Parameters.

ÖZ

MAKİNE ÖĞRENMESİ İLE KORONER ARTERLERDE KESİRLİ AKIŞ REZERVİNİN ANJİYOGRAFİ VE KAN PARAMETRELERİ TABANLI TAHMİNİ

Kaçar, Mehmet Nazir
Doktora, Elektrik ve Elektronik Mühendisliği
Tez Yöneticisi: Prof. Dr. İlkay Ulusoy

Haziran 2026, 115 sayfa

Koroner arter hastalığı (KAH), yüksek ölüm oranı, sakatlık ve toplumsal yükü olan çok yaygın bir kardiyovasküler hastalıktır. KAH, koroner arterlerin aterosklerotik tıkanıklıklarından kaynaklanır. Koroner anjiyografi, koroner kalp hastalığının klinik tanısında altın standarttır. Rutin koroner anjiyografi sırasında bir koroner arter stenozu distalindeki koroner basınç (damardaki tıkanmadan sonra) ile aort basıncı arasındaki oranı hesaplamak için basınçlı tel kullanılarak (fraksiyonel akış rezervi (FFR) ölçümü) klinisyenlerin belirli durumlarda karar vermesine yardımcı olur. Bununla birlikte, FFR invaziv, karmaşık ve pahalı bir işlemdir ve anjiyografi sırasında gerçekleştirilmesi gerekir. Koroner arter enstrümantasyonu gerektirmeden anjiyografi verisi kullanılarak FFR tahmini için çeşitli yöntemler geliştirilmiştir. Bu yöntemler genellikle akışkanlar dinamiği hesaplamasına (CFD) yönelik çeşitli yaklaşımlar kullanır. CFD sürecinin karmaşıklığı, sınır koşulları ve kanın fiziksel özellikleri ile ilgili çeşitli varsayımlara yol açmıştır. Anjiyografi görüntüleri kullanılarak FFR'yi tahmin etmek için çeşitli makine öğrenmesi (ML) modelleri uygulanmaktadır. Anjiyografi tabanlı tahminlerin temel sınırlaması, ölçülen sıvının

(kanın) fiziksel özelliklerini dikkate almamasıdır. FFR, kan akışı bozukluğunu ölçtüğü için, bu reolojik davranıştan doğal olarak etkilenir. Tahmini iyileştirmek için X-ışını koroner anjiyografi görüntüleriyle birlikte kan parametrelerinin kullanılmasını ve kan parametrelerini anjiyografi videolarıyla birleştiren yeni bir model öneriyoruz. Hasta verilerini geriye dönük olarak topladık ve kan parametrelerinin eklenmesinin tahmin yeteneğini geliştirdiğini deneysel olarak gösterdik.

Anahtar Kelimeler: Anjiyografi, Makine Öğrenmesi, Koroner Arter Hastalığı, Fraksiyonel Akış Rezervi, Kan Parametreleri.

To my wife Dilek and my children Zeynep Ece and Selim Giray...

ACKNOWLEDGMENTS

I would like to thank my supervisor, Prof. Dr. İlkay Ulusoy, for her endless patience and for always trusting my work throughout this process.

I am also grateful to Prof. Dr. Çağrı Yayla for sharing his valuable clinical data and for his helpful comments that kept this work grounded in reality. My thanks go to Assoc. Prof. Dr. Yeşim S. Doğrusöz for her guidance and for the critical feedback that helped me improve my methods.

Finally, I owe a great debt of gratitude to Asst. Prof. Dr. Dilek Kaçar. Her hematology expertise was essential to this study; she helped me develop the hemorheological indices from scratch and guided me through the literature. Her support with the statistical analysis and data validation was also critical to the reliability of our results. I truly appreciate all the time and effort she put into this work alongside me.

TABLE OF CONTENTS

ABSTRACT v

ÖZ vii

ACKNOWLEDGMENTS x

TABLE OF CONTENTS xi

LIST OF TABLES xv

LIST OF FIGURES xvii

LIST OF ABBREVIATIONS xviii

CHAPTERS

1 INTRODUCTION 1

1.1 Motivation 1

1.2 Problem Statement 2

1.3 Contribution of the Thesis 3

1.4 Outline of the Thesis 4

2 MEDICAL BACKGROUND 5

2.1 Coronary Arteries and Coronary Circulation 5

2.2 Coronary Artery Disease 7

2.2.2 Invasive Coronary Angiography (ICA) 8

2.3 Fractional Flow Reserve (FFR) 10

2.4 The Composition and Mechanical Properties of Blood 11

2.4.1 Red Blood Cells (Erythrocytes) 11

2.4.2 Platelet Indices 12

2.4.3 White Blood Cell (WBC-Leukocyte) Differential 12

2.4.4 Total Serum Protein 13

2.5 Fluid Mechanics of Luminal Stenosis and Cellular Rheology 13
2.6 AI in the Management of Coronary Artery Disease (CAD) 15
2.6.1 Advanced Imaging and Automated Interpretation 16
2.6.2 Functional Assessment and AI-FFR 16
2.6.3 Predictive Analytics and Risk Stratification 17
3 MACHINE LEARNING BACKGROUND 19
3.1 Artificial Neural Networks and Optimization Mechanics 19
3.2 Convolutional Neural Networks and Spatio-Temporal Extraction 21
3.3 Vision Transformers and Shifting Attention Mechanisms 23
3.4 Gradient Boosted Regression Trees 23
3.5 Multimodal Data Fusion and Representation Conditioning 24
3.5.1 Taxonomy of Multimodal Fusion: Early, Hybrid, and Late Fusion 25
3.5.2 Feature-wise Linear Modulation (FiLM) 26
4 LITERATURE REVIEW 29
4.1 Commercial software predicting FFR without AI 29
4.2 Commercial software using AI 29
4.3 AI applied to patient anatomical data 30
4.4 AI Applied to CCTA Images 31
4.5 AI applied to ICA key frames 31
4.6 AI Applied to ICA Videos 32
4.7 AI applied to synthetic and other kinds of data 34
4.8 Challenges and Limitations in FFR Prediction 35
5 DATASET 39
5.1 Institutional Ethics and Data Acquisition 39

5.2 Cohort Selection and Exclusion Criteria 40
5.3 Angiographic Image and Temporal Signal Processing 41
5.3.1 Temporal Standardization and Padding 41
5.3.1 Multi-Scale Spatial Datasets (Downscaling and Cropping) 42
5.3.2 Spatio-Temporal Velocity Fields (Optical Flow) 43
5.4 Hematological Dataset and Feature Engineering 45
5.4.1 Baseline Statistical Analysis & Predictive Potential 49
6 EXPERIMENTS 53
6.1 Performance comparison of common models with video-only training 54
6.1.1 Comparison with angiography-based studies 56
6.2 Hemorheological Feature Analysis and Predictive Benchmarking 56
6.3 Designed Model Architectures 58
6.3.1 Tabular-Only model 58
6.3.2 Video -Only model 60
6.3.3 Multimodal Late Fusion Model 62
6.3.4 Multimodal FiLM Model 64
6.3.5 Output Heads and Ratio Loss 65
6.4 Training Methodology 68
6.4.1 Patient level data splitting 68
6.4.2 Class Imbalance Handling 69
6.4.3 Training Protocol 69
6.4.4 Reproducibility 70
6.4.5 Visualization 70
6.5 Results 70

6.5.1 Cross Validation benchmarking 71
6.5.2 Test Set Generalization 71
6.5.3 Predictive Precision and Clinical Reliability 72
6.5.4 GradCAM Visualization 81
6.6 Discussion 81
6.6.1 Video Features Alone Are Insufficient 81
6.6.2 Blood Parameters Carry Independent FFR Signal 84
6.6.3 FiLM Outperforms Late Fusion 86
6.6.4 Cross-Validation vs. Test-Set Metrics 86
6.6.5 Composite Output Layer 86
6.6.6 Comparison with similar work 87
6.7 Limitations and Future Work 88
7 CONCLUSION 91
REFERENCES 97
A. Minimum and Maximum values of Modality Performances 111
CURRICULUM VITAE 115

LIST OF TABLES

TABLES

Table 5.1 Dataset Statistics 41
Table 5.2 Dataset Statistics (continued) 42
Table 5.3 Prepared Datasets 45
Table 5.4 Input Feature Matrix Specifications for the Multimodal Blood Fluidity Model 48
Table 5.5 Hemorheological Predictors of FFR-defined Ischemia in the Logistic Regression Model 50
Table 5.6 Baseline demographic, Clinical, and Comprehensive Hemorheological Profiles of the Study Population, stratified by gender, Functional Ischemia Status, and Continuous FFR Correlations 51
Table 6.1 Performance comparison of standard models (MAE) 55
Table 6.2 Comparative Performance Metrics: Raw Hematological Indices vs. Engineered Hemorheological Vector 59
Table 6.3 Output Heads 65
Table 6.4 r and p values of Spearman's correlation for modality predictions 72
Table 6.5 Single output head test set results for FiLM model for single seed 73
Table 6.6 Comparison of Cross-Validation Modality Performance with all blood parameters 79
Table 6.7 Comparison of Cross-Validation Modality Performance with selected blood parameters 79
Table 6.8 Comparison of Test Set Modality Performance with all blood parameters 80
Table 6.9 Comparison of Test Set Modality Performance with selected blood parameters 80
Table 6.10 Comparison with FFR predicting studies 89

Table A.1 Minimum values of Cross-Validation Modality Performance with all blood parameters 111
Table A.2 Minimum values of Cross-Validation Modality Performance with selected blood parameters..... 111
Table A.3 Maximum values of Cross-Validation Modality Performance with all blood parameters 112
Table A.4 Maximum values of Cross-Validation Modality Performance with selected blood parameters..... 112
Table A.5 Minimum values of Test Set Modality Performance with all blood parameters..... 113
Table A.6 Minimum values Test Set Modality Performance with selected blood parameters..... 113
Table A.7 Maximum values of Test Set Modality Performance with all blood parameters..... 114
Table A.8 Maximum values of Test Set Modality Performance with selected blood parameters..... 114

LIST OF FIGURES

FIGURES

Figure 2.1. Coronary Arteries Source: Wikipedia 7
Figure 2.2. Coronary Angiographic Views 10
Figure 5.1. FFR Value occurrences in dataset 42
Figure 5.2. Cropping around stenosis coordinates 43
Figure 5.3. Original image and RAFT Optical Flow Model output 44
Figure 5.4. Original image and Farneback Optical Flow output 44
Figure 6.1. Tabular-Only model structure 60
Figure 6.2. Video-only model structure 61
Figure 6.3. Late Fusion model structure 63
Figure 6.4. A FiLM Layer for a CNN 65
Figure 6.5. Proposed multimodal FiLM model. Two FiLM layers are stacked after 3D-CNN layers. 66
Figure 6.6. Test Set ROC with selected blood parameters for best seed 75
Figure 6.7. Bland-Altman plot for multimodal FiLM model (top) and multimodal late fusion model (bottom) for all seeds 76
Figure 6.8. Scatter plots of: tabular-only model (top), video-only model (bottom) 77
Figure 6.9. Scatter plots of: multimodal late fusion model (top) and multimodal FiLM model (bottom) for all seeds 78
Figure 6.10. GradCAM comparison of a positive class video 82
Figure 6.11. GradCAM comparison of a negative class video 83

LIST OF ABBREVIATIONS

ABBREVIATIONS

AI: Artificial Intelligence

ANN: Artificial Neural Network

AUC: Area Under Curve

BPNN: Backpropagation Neural Network

CA: Coronary Artery

CAD: Coronary Artery Disease

CAU: Caudal

CBC: Complete Blood Count

CCTA: Cardiac Computed Tomography Angiography

CFD: Computational Fluid Dynamics

CFR: Coronary Flow Reserve

CNN: Convolutional Neural Network

CRA: Cranial

CTP: Computed Tomography Perfusion

CV: Cross-validation

CVD: Cardiovascular Disease

DL: Deep Learning

FFNN: Feed Forward Neural Network

FFR: Fractional Flow Reserve

FiLM: Feature-wise Linear Modulation

FN: False Negative

FP: False Positive

GradCAM: Gradient-weighted class activation mapping

HCT: Hematocrit

HGB: Hemoglobin

ICA: Invasive Coronary Angiography

LAD: Left Anterior Descending

LAO: Left Anterior Oblique

LCx: Left Circumflex

LMCA: Left Main Coronary Artery

LSTM: Long Short-Term Memory

MAD: Mean Absolute Deviation

MAE: Mean Absolute Error

MCH: Mean corpuscular hemoglobin

MCHC: Mean corpuscular hemoglobin concentration

MCV: Mean corpuscular volume

MPR: Multivariate Polynomial Regression

MPV: Mean platelet volume

ML: Machine learning

MSE: Mean Squared Error

N/A: Not Applicable

NPZ: Numpy array Zipped

OCT: Optical Coherence Tomography

PCT: Plateletcrit

PDW: Platelet distribution width

PINN: Physics-Informed Neural Network

PLT: Platelet count

SD: Standard Deviation

QCA: Quantitative Coronary Angiography

QFR: Quantitative Flow Rate

R3D: Resnet 3D

RAO: Right Anterior Oblique

RBC: Red Blood Cells

RCA: Right Coronary Artery

RDW: Red blood cell distribution width

ROC: Receiver Operating Characteristic

S3D: Separable 3D CNN

TN: True Negative

TP: True Positive

ViViT: Video Vision Transformer

WBC: White Blood Cells

WBV: Whole Blood Viscosity

CHAPTER 1

INTRODUCTION

1.1 Motivation

Cardiovascular diseases (CVDs) remain the primary cause of global mortality, accounting for nearly 20 million deaths annually as of 2024 [1]. Among these, Coronary Artery Disease (CAD) represents the most significant clinical burden, characterized by the progressive accumulation of atherosclerotic plaques in the coronary arteries, arteries of the heart that compromise its blood supply. In Turkey, CAD accounts for approximately 42.9% of all circulatory system-related deaths [2], highlighting an urgent need for more precise and cost-effective diagnostic workflows.

There are both invasive and non-invasive tools for assessing coronary blood circulation and diagnosing CAD. Non-invasive tools include electrocardiography, echocardiography, stress testing, computed tomography, and cardiac magnetic resonance imaging (MRI), which are easier to perform and have fewer complications. Invasive tools include intravascular imaging modalities and invasive coronary angiography (ICA), both of which require inserting instruments into the body [3]. The choice of diagnostic procedure depends on the severity of the patient's symptoms.

The current gold standard for anatomical assessment of CAD is ICA. While ICA provides high-resolution 2D visualizations of the coronary tree, visual interpretation is inherently subjective, particularly in the diagnosis of intermediate stenoses (50–70% luminal narrowing). Studies indicate significant inter-observer variability in assessing the functional significance of such lesions, which can lead to either unnecessary revascularization or overlooked ischemia. To overcome these

limitations, FFR has been introduced as the clinical gold standard for functional assessment. By calculating the pressure ratio across a stenosis, FFR provides an objective metric of blood flow impairment. However, the widespread adoption of FFR is hindered by its invasive nature, the requirement for specialized pressure-sensitive wires, increased procedural time, and higher costs.

Beyond anatomical narrowing, the hemodynamic significance of a lesion is profoundly influenced by blood rheology. Blood is a non-Newtonian, multi-phase suspension whose viscosity varies dynamically with shear rate and composition. Specifically, hemogram parameters - such as hematocrit levels and total plasma protein concentrations - are critical determinants of calculated Whole Blood Viscosity (WBV). Clinical evidence suggests that elevated WBV can trigger hemodynamically significant pressure drops (FFR $\leq$ 0.80) even in lesions with moderate anatomical narrowing [4] [5]. Despite its importance, this physiological synergy between vessel geometry and blood rheology is often neglected in standard angiographic evaluations.

1.2 Problem Statement

The emergence of Artificial Intelligence (AI) and Machine Learning (ML) offers a transformative opportunity to bridge the gap between anatomical imaging and functional physiology. Recent advancements in deep learning (DL) have demonstrated the potential to automate plaque characterization and reconstruct 3D geometries from 2D angiographic frames. However, most existing AI-FFR models rely heavily on imaging data alone, often ignoring the patient-specific microfluidic properties and cellular suspension profiles that govern the absolute fluid dynamics within the living vessel [6].

To address this critical gap, we hypothesized that integrating specific erythrocyte and platelet indices derived from routine complete blood count (CBC) tests would provide the bio-rheological context to calibrate these computational models.

However, review of the current literature revealed a lack of publicly available datasets that concurrently pair raw ICA video sequences with corresponding hematological profiles.

To overcome this limitation, we created a unique multidimensional dataset. This study was conducted with the institutional collaboration of Ankara Bilkent City Hospital, which provided the angiography video sequences with synchronized blood parameters, along with the necessary institutional ethical approval.

1.3 Contribution of the Thesis

This thesis proposes a novel AI-based framework for predicting FFR by integrating two distinct data modalities: 2D ICA sequences and patient-specific blood parameters. A fundamental contribution of this research is the compilation, standardization, and creation of a unique and multimodal dataset that aligns real-world angiographic video sequences with synchronized, comprehensive, and patient-specific hematological profiles—routinely obtained before the procedure—thereby addressing a critical data gap in computational cardiology.

Shifting from geometry-centric fluid simulations, this research introduces an expert hematological perspective to model the coronary lumen as a non-Newtonian cellular suspension. To operationalize this biophysical hypothesis, we engineered specific hemorheological indices derived from routine complete blood counts to capture intravascular fluid dynamics, microhemolysis, and spatial cell-to-cell interactions: the Erythrocyte Fragility Index (EFI), the Cell-Free Hemoglobin Index (CFHI), and the Erythrocyte Platelet Volume Balance (EPVB). These newly developed indices, alongside Mean Platelet Volume (MPV) as an isolated raw predictor and gender as an essential physiological calibration factor, were integrated to condition spatial-temporal convolutional neural network (CNN) feature maps.

By leveraging this multimodal architecture, this research aims to develop a non-invasive, real-time functional assessment tool that simplifies clinical workflows, reduces the need for invasive pressure wires, and provides a precision-medicine-driven diagnostic pathway for patients with coronary artery disease.

The systematic validation of this framework is executed across three distinct experimental steps:

First, models trained solely on raw angiographic video sequences were tested to establish a visual baseline performance.

Second, the non-linear standalone associations between the engineered indices, raw hematological variables, and absolute FFR values were investigated.

Finally, four models were trained together and compared: a blood parameter only model, a video only model, a late fusion model combining blood parameters with video features using baseline cross-modal fusion techniques, and a specific Feature-wise Linear Modulation (FiLM) model. A training scheme that provided the best optimization and improvement was applied.

1.4 Outline of the Thesis

First chapter summarizes the problem and motivation. A brief medical background is given in Chapter 2. In Chapter 3, ML methods used in this thesis are explained briefly. A literature review on FFR estimation is provided in Chapter 4. In Chapter 5, the preparation and properties of the dataset are given. In Chapter 6, experiments performed are listed. Discussion of the results and a conclusion is presented in Chapter 7.

CHAPTER 2

MEDICAL BACKGROUND

CVDs are the leading cause of death worldwide. According to 2024 reports, CVDs cause nearly 20 million deaths globally each year [1]. This represents an increase of over 20% over the past decade, despite therapeutic advances that have reduced the age-standardized death rate by almost 15%. The rising prevalence is due to an aging population, which will continue to grow in the coming years [7]. In European Union countries, 11% of the entire healthcare budget is spent on CVDs, which remains the leading cause of death on the continent. In middle-income countries, mortality rates were at least 2.5 times higher than those in high-income countries [8]. Diseases of the circulatory system are also the leading cause of death in Turkey, accounting for 36% of all deaths, with CAD being the most common at 42.9% in 2024 [2].

Given the increasing frequency and associated costs, diagnostic tests for this disease should be streamlined, with complex tests reserved for patients who require them. AI/ML may play a role in CAD management through signal processing, advanced medical imaging, feature extraction, and predictive risk stratification [9]. There is an urgent need to develop AI/ML tools to create practical, cost-effective workflows for precision medicine that address and simplify core clinical needs.

2.1 Coronary Arteries and Coronary Circulation

The heart is a highly metabolically active organ in the human body, with the highest oxygen consumption per unit mass. To meet this demand, the coronary circulation, which receives approximately 5% of cardiac output, functions as a hierarchical hydraulic network optimized for dynamic perfusion. The coronary arteries, which supply blood to the heart muscle (myocardium), arise from the aortic root. The system follows a specific bifurcating topology:

- Right Coronary Artery (RCA): Arising from the anterior aortic sinus, it supplies the right atrium, right ventricle, and the conduction system (SA and AV nodes).
- Left Main Coronary Artery (LMCA): Arising from the left posterior aortic sinus, it quickly bifurcates into the Left Anterior Descending (LAD) and the Left Circumflex (LCX) arteries, which primarily perfuse the high-pressure left ventricle.

From a fluid-mechanics perspective, the system transitions from low-resistance epicardial conductance vessels (large-diameter, surface-level vessels) to high-resistance intramyocardial microvascular arterioles and dense capillary beds. While epicardial vessels act as primary distributors, the distal microvasculature provides the necessary hydraulic resistance to modulate flow via autoregulatory mechanisms. A defining mechanical feature of the coronary network is its phasic flow behavior, which diverges from systemic arterial patterns. In most peripheral tissues, flow peaks during ventricular systole. However, coronary perfusion is governed by intramyocardial pump effects:

- Systolic Compression: During ventricular contraction (systole), high intramyocardial pressure compresses the penetrating vessels, significantly increasing flow resistance and nearly halting coronary inflow.
- Diastolic Perfusion: Flow occurs almost exclusively during ventricular relaxation (diastole), when extravascular pressure drops, allowing the pressure gradient between the aorta and the myocardium to drive perfusion.

And a sophisticated autoregulatory feedback loop maintains the stability of this hydraulic network. This system responds to myogenic pressure changes, metabolic signals, and fluid shear stress, thereby increasing the ratio of maximum coronary blood flow to resting coronary blood flow up to fivefold—a measure known as Coronary Flow Reserve (CFR) [10] [11].

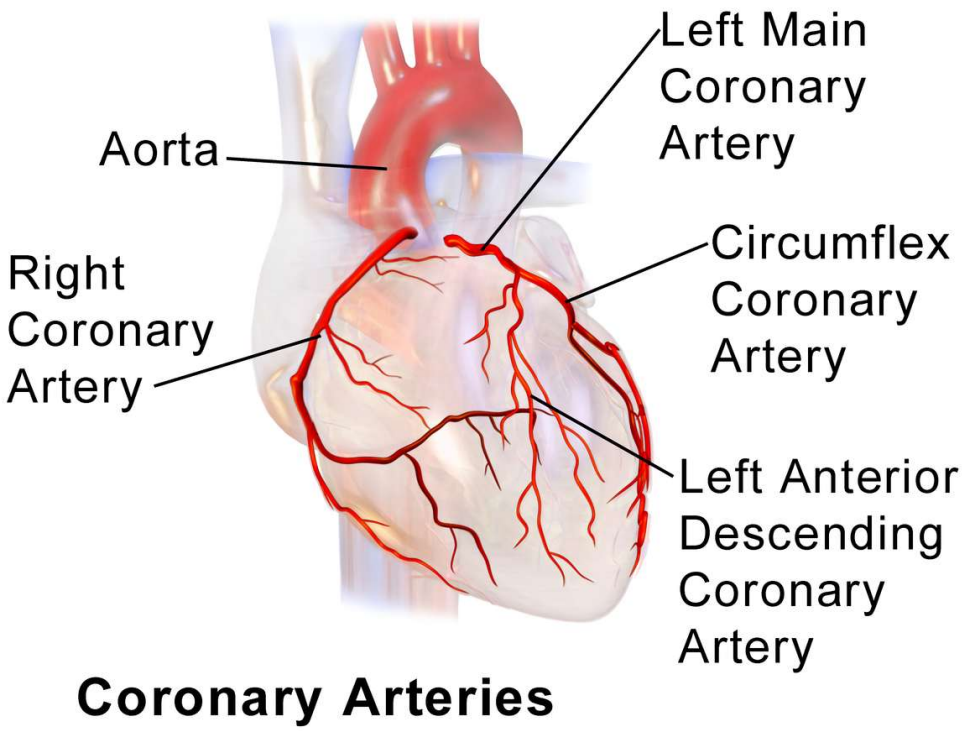

Figure 2.1. Coronary Arteries Source: Wikipedia

2.2 Coronary Artery Disease

CAD is a complex, progressive pathophysiological process driven by the accumulation of atherosclerotic plaques within the CAs, leading to progressive narrowing, occlusion, and impaired myocardial blood supply. CAD is multifactorial, with etiological factors including non-modifiable factors such as gender, age, family history, and genetics, and modifiable factors such as hypertension, smoking, obesity, lipid levels, and psychosocial variables. Men are more predisposed to CAD than women. And inflammation is also a substantial risk factor for CAD.

In Western countries, a faster lifestyle has increased the consumption of fast, unhealthy meals, leading to a higher prevalence of CAD. On the other hand, improved health care among middle and higher-socioeconomic status groups in

developed nations has shifted the incidence of CAD to later in life. Although CAD rates are expected to decline in these countries, factors such as immigration and an aging population may temporarily increase the incidence of coronary events. Meanwhile, developing nations exhibit substantial variation in CAD incidence, with Western dietary patterns and sedentary lifestyles likely contributing to increased cases.

From a hemodynamic perspective, CAD represents progressive structural degradation of the coronary hydraulic network, with the primary consequence being an increase in localized flow resistance. The clinical and functional severity of this impairment is strictly governed by the degree of luminal diameter reduction. In non-significant stenosis (<50% reduction), the system's innate autoregulatory mechanisms—primarily vasodilation of the distal microvasculature—compensate for the localized pressure drop, thereby maintaining adequate perfusion at rest. However, as the lesion progresses to a hemodynamically significant stenosis (≥50%), the hydraulic resistance reaches a threshold where the oxygen supply-demand balance is disrupted during peak metabolic activity, manifesting as angina pectoris (chest pain). Finally, in cases of critical stenosis (>90%), the obstruction becomes so severe that the pressure gradient is insufficient to drive adequate flow even under baseline conditions, leading to chronic myocardial ischemia [7] [12].

2.2.2 Invasive Coronary Angiography (ICA)

The most common procedure employed to diagnose CAD is ICA. An interventional cardiologist guides a catheter through the major arteries of the body until it reaches the aorta, positioning its tip within one of the coronary arteries (CAs). Subsequently, a radio-opaque contrast agent is injected via the catheter, and a sequence of X-ray images is captured. These images render the coronary arteries and capillaries visible for a limited duration before the contrast agent is dissipated. The data collected from coronary angiography consist of sequences of monochrome images, which are often visualized as videos because multiple frames are captured per second. Within each

angiographic video, a key frame exists wherein the contrast agent has adequately dispersed, revealing the entire coronary vasculature.

The angiographic images are two-dimensional, whereas the coronary arteries are three-dimensional structures. The visibility of different segments of the coronary arteries varies depending on the camera's position. The camera can be oriented toward the right side of the patient (Right Anterior Oblique – RAO) or the left side (Left Anterior Oblique – LAO) along the horizontal axis (ranging from -90 degrees to +90 degrees). Additionally, the camera can be directed toward the patient's head (cranial view, CRA) or toward the feet (caudal view, CAU) along the vertical axis. These perspectives are referred to as coronary angiographic views. To obtain more explicit images and prevent overlapping of branches, the viewing angles are adjusted for proximal, mid, and distal segments of the coronary arteries. Some of these views are illustrated in Figure 2.2. During angiography, the cardiologist adjusts the camera position to obtain comprehensive visualization of the coronary arteries and records one or multiple videos accordingly.

Each angiography video possesses a distinct angiographic view. Heart size and coronary artery diameters vary among patients. Furthermore, the distance from the x-ray source to the patient and to the detector affects the overall contrast and spatial resolution [13]. During angiography, the patient continues to breathe, and the heart continues to beat, introducing complex and unique movements into the angiographic footage. Due to the presence of the patient’s body and organs, background objects such as the spine, ribs, diaphragm, muscles, implants, dye catheters, probes, and stitches are always visible. Interventional cardiologists analyze the angiography videos to evaluate lesions, plaques, bifurcations, and stenosis within the coronary arteries. Immediate diagnostic angiography is recommended for patients presenting with acute coronary syndrome. Depending on the significance of ICA findings, invasive management may be conducted within the same session [3]. Through visual inspection of the angiography, mild and severe stenoses, along with coronary artery occlusions below 50% and exceeding 70%, are more readily diagnosed. Conversely, intermediate stenoses (50% to 70%) present greater diagnostic challenges [14].

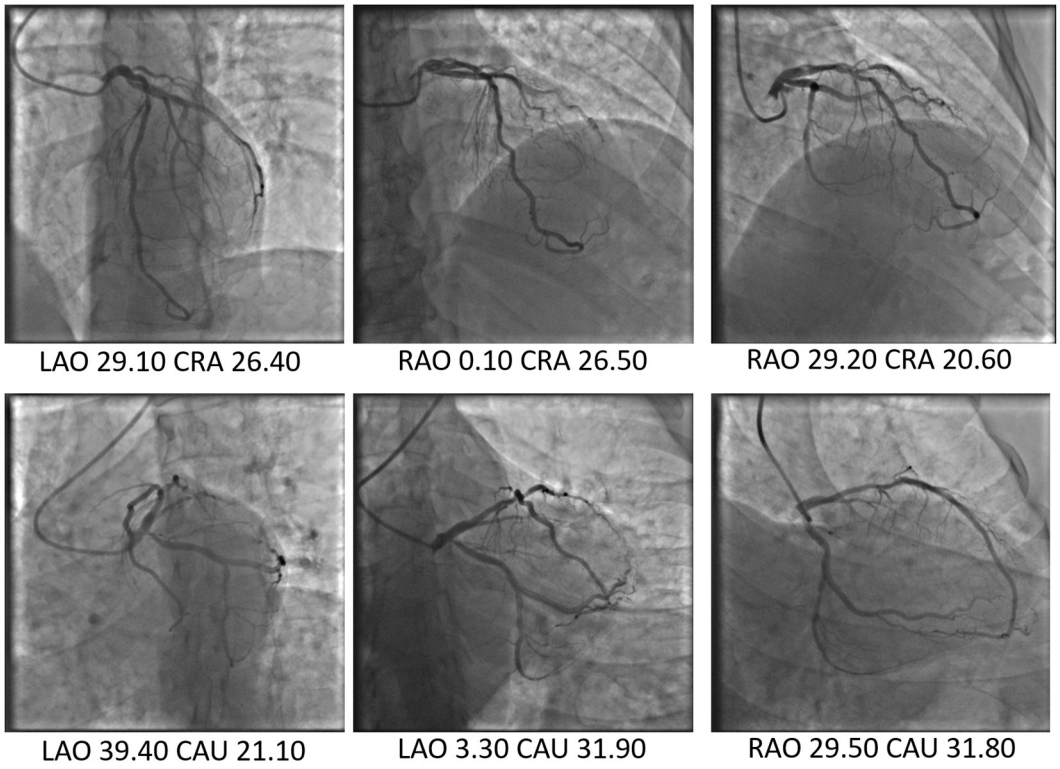

Figure 2.2. Coronary Angiographic Views. Abbreviations: CAU, caudal; CRA, cranial; LAO, left anterior oblique; RAO, right anterior oblique.

2.3 Fractional Flow Reserve (FFR)

Based solely on visual inspection of angiography, occlusions of the CAs below 50% and above 70% are more straightforward to diagnose, whereas stenoses between 50% and 70% pose greater diagnostic challenges. Expert evaluations rely heavily on experience and opinion, leading to potential disagreements among specialists [14]. To address this issue, in selected cases, a more advanced invasive procedure known as FFR measurement is employed. FFR is defined as the ratio of the pressure distal to a stenosis to the pressure proximal to the stenosis [15]. This measurement is obtained via a pressure-sensitive wire introduced through the angiography catheter, which records pressures before and after the stenosis. The FFR sensor is calibrated to a value of 1 at the aorta. An FFR exceeding 0.75 is associated with a lower risk of cardiac death or myocardial infarction related to the stenosis, with an annual risk

below 1%. Further research suggested that 0.80 serves as a more appropriate cutoff value [16], with FFR values between 0.75 and 0.80 designated as the gray zone [17].

2.4 The Composition and Mechanical Properties of Blood

Blood is a specialized non-Newtonian fluid that serves as the primary transport system for oxygen, nutrients, and immune cells. It consists of a liquid matrix called plasma and three main cellular categories: red blood cells (RBCs, erythrocytes), white blood cells (WBCs), and platelets (thrombocytes). The cellular phase is dominated by RBCs, which account for 40–45% of the total blood volume (hematocrit). WBCs and platelets together constitute only about 1% of the volume. The remaining 55% of the volume, plasma, is approximately 93% water and contains organic molecules, electrolytes, and critical proteins such as albumin, globulins, and fibrinogen. Information on blood cell counts, percentages, and many structural details of cells can be obtained from CBC or hemogram tests [18].

2.4.1 Red Blood Cells (Erythrocytes)

The most abundant cells (4.2–6.2 million cells/μL) are primarily responsible for oxygen transport via hemoglobin. Their biconcave shape (diameter ~6 μm) and high deformability enable them to navigate microvessels and significantly influence blood viscosity [19]. The RBC indices in CBC tests are;

- RBC: Number of RBCs.
- Hemoglobin (HGB): Hemoglobin is a protein contained in RBCs, and makes 96% of a RBC's dry weight.
- Hematocrit (HCT): Hematocrit is the volume percentage of RBCs in blood, as a calculation of MCV and RBC.
- Mean Corpuscular Volume (MCV): Represents the average size of RBCs.
- Mean Corpuscular Hemoglobin (MCH): Quantifies the average amount of hemoglobin per RBC.

- Mean Corpuscular Hemoglobin Concentration (MCHC): Indicates the hemoglobin density in a given volume of RBCs. The internal viscosity of an RBC is primarily determined by its hemoglobin concentration; higher MCHC values can reduce cell deformability.
- Red Cell Distribution Width (RDW): Measures the variation in RBC size.

2.4.2 Platelet Indices

Platelets are the primary cells of hemostasis. Beyond the absolute count, specific indices define their functional state:

- Platelet count (PLT): Platelet count is the number of platelets in blood.
- Mean Platelet Volume (MPV): Indicates the average size of platelets. Larger platelets contain denser granules and are enzymatically more active.
- Plateletcrit (PCT): Plateletcrit is the calculation of volume occupied by platelets in the blood, as a calculation of PLT and MPV.
- Platelet Distribution Width (PDW): Reflects the variation in platelet size. An increased PDW is a hallmark of platelet activation and heterogeneity.

2.4.3 White Blood Cell (WBC-Leukocyte) Differential

A standard CBC also includes a leukocyte differential, which quantifies the five major leukocyte subtypes. Each subtype has a distinct role in the inflammatory processes.

- Neutrophils: The primary responders to acute inflammation and tissue damage.
- Lymphocytes: Key mediators of the adaptive immune response.
- Monocytes: They are a critical part of the innate immune system, providing vital elements like macrophages and dendritic cells.

- Eosinophils & Basophils: Involved in allergic responses and specialized inflammatory pathways.

2.4.4 Total Serum Protein

Total Serum Protein represents the concentration of protein in serum. Serum contains many proteins, including serum albumin, a variety of globulins, and many others.

2.5 Fluid Mechanics of Luminal Stenosis and Cellular Rheology

The estimation of fluid dynamics and pressure drops across an arterial narrowing has historically focused on macro-rheological properties, primarily quantified through Whole Blood Viscosity (WBV). A widely recognized method for this is the **De Simone formula**, which calculates viscosity based on HCT and Total Protein (TP) levels [20]:

$$For\ low\ shear\ rates: WBV\ = (0.12\ \times HCT) + (0.17\ \times Protein) - 2.07$$

$$For\ high\ shear\ rates: WBV\ = (3.53\ \times HCT) + (6.36\ \times Protein) - 119.47$$

Clinical investigations utilizing this framework have demonstrated that a higher estimated WBV significantly correlates with reduced pressure-wire measurements, implying that bulk fluid thickness independently intensifies pressure gradients across a stenosis [4] [5]. Furthermore, multi-dimensional mathematical simulations based on the Navier-Stokes equations confirm that local dynamic viscosity directly modulates the velocity profiles and subsequent pressure drops surrounding a stenotic segment. These mathematical models emphasize that viscosity is a critical boundary variable that must be accounted for in functional calculations to prevent diagnostic errors arising from purely anatomical assessments [21]. However, while traditional formulations provide useful generalized macroscopic estimates, they treat blood as a uniform, homogenous fluid. This continuum simplification inherently fails to capture the intricate, non-linear micro-mechanical interactions taking place among

discrete cellular components within the traveling suspension. Intravascular hemodynamics are fundamentally governed by these non-Newtonian cellular dynamics, which are in turn dictated by the structural, volumetric, and rheological alterations of circulating cells under stress. While intravascular hemodynamics are undeniably influenced by a multitude of complex, overlapping systemic factors, we hypothesize that the cumulative impact of these systemic conditions, alongside patient-specific genetic predispositions, is directly mirrored by the structural, volumetric, and rheological alterations of circulating blood cells. Mature erythrocytes and platelets are anucleated elements with distinct lifespans—averaging 120 days for RBCs and 7 to 10 days for platelets—and they continuously traverse the entire macro- and microcirculatory beds during their life cycles. Consequently, they effectively function as highly personalized, circulating biological sensors of systemic vascular stress. Crucially, this relationship operates as a dynamic, bidirectional feedback loop; while these circulating cells are continuously remodeled by the surrounding vascular environment, they concurrently act as the primary determinants of local fluid dynamics, particularly through the dominant role of erythrocytes in governing blood viscosity and shear-induced behavior. Utilizing these cellular parameters offers a profoundly practical clinical advantage, as these precise rheological signatures can be extracted directly from a routine CBC—a baseline laboratory assessment universally performed on every patient prior to coronary angiography.

Within a biophysical framework, the spatial topography of circulating blood cells under physiological shear stress governs fluidic efficiency. In healthy microvascular environments, blood exhibits a distinct multicellular organization where large, highly deformable RBCs occupy the high-velocity central axial stream. This movement hydrodynamically drives smaller platelets toward the vessel periphery via shear-induced lateral displacement, establishing a pronounced concentration peak near the endothelial boundary [22] [23]. Structurally robust platelets with high MPV are efficiently excluded by this central RBC core. Because these larger elements cannot fit into the microscopic interstitial gaps between traveling erythrocytes, they

are forcefully driven toward the extreme periphery [23]. Crucially, by remaining marginated, these large platelets leave the central RBC core uninhibited, preserving ideal fluidic harmony and maximizing downstream perfusion. Conversely, when platelet volume is depleted (low MPV), smaller platelets more readily escape this geometric exclusion and infiltrate the microgaps within the erythrocyte core. This intra-erythrocytic infiltration directly disrupts the structural and fluidic alignment of traveling RBCs, increasing the frequency of chaotic cell-to-cell collisions and amplifying internal viscosity [23]. This spatial disharmony is further compounded by structural vulnerabilities within the erythrocyte compartment itself, characterized by increased erythrocyte size heterogeneity (high RDW) and a concomitant loss of membrane deformability [24] [25]. When these rigid RBCs encounter a stenotic segment, their altered microfluidic behavior forces them to deviate from their axial trajectory, leading to aberrant margination and direct mechanical friction against the vascular endothelium, thereby promoting downstream microvascular damage [26] [27]. Under the extreme turbulent and mechanical shear forces operating within a stenosis, these fragile erythrocytes undergo subclinical micro-fragmentation, releasing cell-free hemoglobin into the plasma. This cell-free plasma hemoglobin acts as an exceptionally aggressive, high-affinity scavenger of endothelial nitric oxide (NO), instantly neutralizing local NO bioavailability and dismantling essential vasodilatory protective pathways. This localized depletion of NO triggers acute smooth muscle contraction, inducing severe microvascular vasospasms and further restricting downstream absolute blood flow independently of macrovascular luminal narrowing [28] [29].

2.6 AI in the Management of Coronary Artery Disease (CAD)

AI and ML are transforming the paradigm of CAD management by enhancing diagnostic precision and optimizing risk stratification across the patient care continuum. According to recent scientific statements [9], the application of AI in CAD can be categorized into three primary domains:

2.6.1 Advanced Imaging and Automated Interpretation

AI algorithms, particularly DL models, are increasingly utilized to automate the interpretation of coronary computed tomography angiography (CCTA). These tools enable:

- Automated Plaque Characterization: Identifying and quantifying atherosclerotic plaque volume and composition (e.g., calcified vs. non-calcified) with high reproducibility [30] [31].
- Geometric Modeling: Reconstructing 3D coronary architectures from 2D imaging data, which serves as a prerequisite for non-invasive functional assessments [32].

2.6.2 Functional Assessment and AI-FFR

One of the most significant breakthroughs is the integration of AI with fluid dynamics. Taylor et al. [33] used CCTA image slices to model patient's heart anatomy in 3D. The luminal areas of major vessels and their branches are identified using image segmentation algorithms. The topology of the coronary tree is extracted, and every vessel, branch, and plaque is analyzed. A mesh is generated from the geometric model, and the resulting system is solved for velocity and pressure using the Navier-Stokes equations. Although blood is a complex non-Newtonian fluid, the study suggests that it can be approximated as a Newtonian fluid with constant viscosity in large arteries. Traditional FFR requires invasive pressure wires; however, AI-driven models can now estimate FFR from CCTA (AI-FFR) or invasive angiography by utilizing:

- Reduced-order Modeling: Accelerating computational fluid dynamics (CFD) simulations through ML to predict pressure drops across stenoses [34].

- Feature-based Prediction: Leveraging large datasets to correlate anatomical features directly with functional significance without the need for full CFD simulations [35].

2.6.3 Predictive Analytics and Risk Stratification

Beyond imaging, AI integrates electronic health records (EHR), genomic data, and biomarkers to predict major adverse cardiovascular events (MACE). These multimodal models outperform traditional scoring systems (e.g., the Framingham Risk Score) by capturing complex, nonlinear interactions among risk factors, thereby enabling more personalized therapeutic interventions [36] [37].

In summary, AI serves as a force multiplier in CAD diagnostics, transitioning the field from purely anatomical assessment to a more nuanced, functional, and predictive framework.

CHAPTER 3

MACHINE LEARNING BACKGROUND

ML encompasses a diverse suite of computational methodologies designed to systematically extract underlying statistical patterns from high-dimensional datasets, enabling robust generalization to previously unseen evaluation domains. Drawing heavily upon computational statistics, mathematical optimization, and information theory, ML systems construct predictive mapping functions directly from complex representations without relying on deterministic, hand-coded heuristics.

Within this paradigm, a ML model represents a specific parameterized function optimization framework tailored to resolve a concrete analytical task. These frameworks are broadly classified into supervised learning architectures, which calibrate internal parameters against rigorous ground-truth targets to perform discrete classification or continuous regression, and unsupervised learning architectures, which isolate inherent structural topology, latent sub-spaces, or clusters within the feature distribution without objective target labels.

3.1 Artificial Neural Networks and Optimization Mechanics

The elemental computational unit of a DL architecture is the artificial neuron. Mathematically, a neuron operationalizes a multi-input mapping function that executes a linear affine transformation on an input vector $x \in \mathbb{R}^d$, parameterized by a learnable weight vector $w \in \mathbb{R}^d$ and a scalar bias $b \in \mathbb{R}$, which is subsequently mapped through a non-linear activation function $\sigma(\cdot)$ to yield a scalar output y:

$$y = \sigma\left(\sum_{i=1}^{d} w_i\, x_i + b\right) = \sigma(\boldsymbol{w}^T \boldsymbol{x} + b)$$

The global model architecture is established by organizing these analytical units into stacked, hierarchical topologies termed layers. To prevent the collapse of deep

architectures into trivial linear operations, specific non-linear activation functions are deployed based on layer-specific objectives:

- Linear Activation: Primarily utilized in the terminal output layer of regression architectures to map latent variables directly to an unconstrained continuous real-world numeric space ($y \in \mathbb{R}$).
- Sigmoid Activation: Deployed in binary classification tasks or specific modulation layers to map continuous real-valued inputs into a strictly bounded interval of (0,1), effectively transforming arbitrary scores into independent probability metrics or scaling factors.

$$Sigmoid(x) = \sigma(x) = \frac{1}{1 + e^{-x}}$$

- Softmax Activation: Deployed exclusively in multi-class classification nodes to normalize an unconstrained logit vector **z** into a valid probability distribution over K discrete classes, satisfying the axiomatic constraint:

$$Softmax(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \; where \sum_{i=1}^{K} Softmax(z_i) = 1$$

Model optimization is achieved through supervised iterative training loops over multiple epochs. During each forward pass, the model projects predictions that are evaluated against the true clinical targets via a differentiable objective loss function, $\mathcal{L}(\theta)$. To update the parameterized network weights (θ), a backpropagation algorithm systematically computes downstream partial derivatives via the chain rule of calculus, propagating the calculated error gradient backward through the network topology.

In this framework, the Adaptive Moment Estimation (Adam) optimizer is widely utilized to dynamically modulate individual parameter learning rates. Adam achieves stable convergence by calculating exponentially decaying moving averages of both the past gradients (first raw moment, m_t) and the squared uncentered gradients (second variance moment, v_t), combining the principles of classical momentum and Root-Mean-Square Propagation (RMSProp).

To measure empirical error in continuous regression tasks, standard objective formulations include Mean Squared Error (MSE), which penalizes larger outlier errors quadratically, and Mean Absolute Error (MAE), which scales linearly:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \, , \, MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$$

3.2 Convolutional Neural Networks and Spatio-Temporal Extraction

CNNs introduce structural weight-sharing and localized receptive fields to drastically reduce total parameter cardinality compared to fully-connected topologies, ensuring computational tractability over structured arrays. CNNs leverage discrete translation-invariant kernel convolutions to autonomously extract hierarchical features directly from raw data representations across diverse dimensional domains:

- 1D-CNNs: Utilize localized sliding kernels along a singular temporal or spatial axis, optimal for processing sequential signals or localized blood parameter vector spaces.
- 2D-CNNs: Slide kernels across a two-dimensional spatial lattice to capture spatial dependencies, serving as the standard paradigm for parsing static visual inputs.
- 3D-CNNs: Concurrently extend kernel convolutions across both 2D spatial dimensions and a 1D temporal axis, enabling the simultaneous capture of static anatomical structures and dynamic motion trajectories across consecutive video frames.

Spatial and temporal dimensionality reductions are achieved via Max Pooling layers, which selectively propagate only the dominant maximum activation response within a localized kernel window, effectively filtering low-amplitude noise. Conversely, Dense (Fully-Connected) layers map these extracted visual features into global representations prior to final prediction. To mitigate structural overfitting and

regularize complex deep representations, Dropout layers randomly deactivate a predefined fraction of forward activations during training, preventing co-dependency among hidden nodes.

To model the complex spatial and temporal dynamics of invasive angiographic sequences, this thesis leverages three distinct state-of-the-art spatio-temporal architectures:

- ResNet3D (R3D): Extends the classic residual skip-connection framework into a three-dimensional convolutional space, optimizing video action recognition by preserving gradient flow across deep spatial and temporal layers simultaneously [38]. Pretrained on the Kinetics-400 dataset [39]—a massive benchmark comprising 400 distinct human action classes—this framework provides highly generalized low-level edge and motion feature detectors.
- Separable 3D-CNN (S3D): Replaces heavy, computationally expensive $3 \times 3 \times 3$ volumetric convolutions with a decoupled sequence consisting of a $1 \times 3 \times 3$ 2D spatial convolution followed immediately by a $3 \times 1 \times 1$ 1D temporal convolution [40]. By structurally isolating spatial geometry from temporal velocity fields, S3D yields highly efficient feature extraction with a reduced parameter footprint, serving as a critical diagnostic baseline to evaluate whether angiography sequences are best processed as discrete independent images or unified spatio-temporal video arrays.
- Convolutional Long Short-Term Memory (ConvLSTM): Combines the recurrent sequence-modeling capacity of Long Short-Term Memory (LSTM) cells with the localized feature extraction of CNNs by embedding convolutional structures directly within the recurrent state transitions [41] [42]. Originally developed for complex meteorological forecasting, a 2D ConvLSTM layer extracts spatial features while maintaining a persistent internal memory cell to model frame-to-frame contrast transit dynamics,

mimicking the visual interpretation process executed by a clinical cardiologist.

3.3 Vision Transformers and Shifting Attention Mechanisms

Shifting away from localized convolutional inductive biases, Transformers rely on global multi-head self-attention mechanisms to explicitly compute pairwise relational dependencies across all tokens within an input sequence [43]. In natural language processing, text inputs are serialized into sequential discrete tokens; similarly, the Vision Transformer (ViT) adapts this logic to computer vision by segmenting a static image into non-overlapping spatial patches, projecting them into a latent embedding space, and appending positional encodings to preserve spatial topology [44]. Video transformers extend this paradigm by extracting spatio-temporal 3D tubelets from raw video arrays, transforming continuous multi-frame clips into high-dimensional tokens for dense relationship mapping, a technique pioneered by the Video Vision Transformer (ViViT) [45].

To optimize the feature extraction performance across varying vessel scales, the Swin Transformer model is utilized [46]. Unlike standard Vision Transformers that compute attention globally—resulting in a quadratic computational complexity relative to image size—the Swin architecture restricts self-attention computation to non-overlapping localized windows while introducing a shifted windowing strategy across consecutive layers. This mechanism achieves hierarchical representation learning by gradually merging adjacent patches at deeper layers, enabling multi-scale spatial feature aggregation that is uniquely suited for detecting variable-sized coronary artery stenoses.

3.4 Gradient Boosted Regression Trees

In contrast to the deep feature representation learning of neural networks, Decision and Regression Trees execute non-linear feature space partitioning by learning

sequential, orthogonal decision boundaries based on deterministic optimization criteria. A single regression tree recursively splits tabular data to predict a continuous numeric outcome by minimizing internal variance within the generated leaf nodes. However, individual trees are highly prone to variance and overfitting, acting as weak learners.

Gradient Boosting constructs highly robust, predictive ensembles by sequentially combining these weak regression tree learners in an iterative boosting framework [47]. The optimization sequence begins with an initial baseline prediction (typically the empirical mean of the target distribution). In each subsequent iteration t, a new regression tree is trained specifically to fit the negative gradient of the loss function—representing the pseudo-residuals or predictive errors—left unresolved by the cumulative ensemble from iteration $t-1$. The prediction of the new tree is then scaled by a constrained regularization shrinkage parameter (learning rate, η) and added to the cumulative model prediction:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta f_t(\boldsymbol{x}_i)$$

This loop repeats over hundreds of iterations, with each sequential tree systematically minimizing the remaining residual error space of its predecessor. Within this domain, XGBoost (Extreme Gradient Boosting) stands as an advanced, highly scalable implementation of gradient boosting, incorporating a formalized regularization term directly into the objective function to penalize model complexity and prevent overfitting [48]. In this thesis, XGBoost is leveraged as the core analytical framework to evaluate the direct statistical and non-linear relationships between clinical FFR values and engineered peripheral blood fluidity vectors.

3.5 Multimodal Data Fusion and Representation Conditioning

In modern ML applications, integrating heterogeneous data sources has emerged as a critical element for enhancing model generalization capability and predictive performance. Especially within clinical decision support systems, synchronous

processing of high-dimensional spatiotemporal features inherent in medical imaging or video streams (e.g., invasive coronary angiography videos) alongside lower-dimensional tabular clinical parameters—such as patient demographics and laboratory findings—is paramount. Merging these distinct modules, which exhibit fundamentally disparate mathematical structures and statistical distributions, necessitates advanced multimodal data fusion and representation conditioning methodologies.

3.5.1 Taxonomy of Multimodal Fusion: Early, Hybrid, and Late Fusion

In multimodal learning literature, information integration frameworks are categorized based on the architectural stage at which the modalities merge, primarily divided into early fusion, hybrid fusion, and late fusion. Early fusion approaches aggregate raw data streams or input-level feature vectors at the earliest processing phase before feeding the unified representations into the model. However, this approach often yields unstable optimization dynamics due to the extreme dimensionality imbalance between massive data spaces (such as video frames) and low-dimensional tabular attributes, frequently resulting in the suppression of modality-specific non-linear characteristics. To circumvent these architectural limitations, late fusion architectures govern information integration by independently processing each modality through specialized, decoupled sub-networks. Within this paradigm, a spatiotemporal DL pipeline (e.g., a 3D-CNN or Vision Transformer branch) captures contextual cues from video data, while a parallel network topology (e.g., a Gradient Boosted Regression Tree or Multi-Layer Perceptron) encodes the tabular patient characteristics. Immediately prior to the final decision layer or regression/classification head, the high-level latent representations (h_v and h_t) generated by these independent branches are systematically synthesized via a mathematical operator. The mathematically prevalent mechanism employed in

late fusion is the sequential concatenation of feature vectors. Denoting the combined multimodal representation vector as z, the formulation is expressed as follows:

$$z = h_v \parallel h_t$$

Here, the ‖ operator signifies vector concatenation. Alternatively, under constraints where feature dimensions are symmetric, element-wise summation or weighted averaging operations may be utilized:

$$z = \omega_v \cdot h_v + \omega_t \cdot h_t$$

Where ω_v and ω_t denote learnable weight coefficients optimized to adaptively calibrate the relative contribution of each modality to the final decision boundary. The primary theoretical asset of late fusion lies in its capacity to preserve the integrity of distinct feature spaces up to the deepest abstraction layer, while providing structural resilience against missing data artifacts across individual modalities.

3.5.2 Feature-wise Linear Modulation (FiLM)

While late fusion configurations effectively merge highly abstracted representations at the output layer, they lack the capacity to capture cross-modal interactions within the intermediate latent layers of deep architectures. To alleviate this constraint and dynamically govern deep spatiotemporal feature maps with secondary contextual data, FiLM [49] was introduced as an adaptive conditioning mechanism. FiLM conditions the primary network (e.g., a 3D-CNN or Vision Transformer) by applying an affine transformation over its intermediate feature maps based on the representations produced by an auxiliary conditioning network (e.g., mapping tabular clinical features). Let x represent the input vector containing tabular clinical variables. This conditioning source is projected via distinct linear or non-linear mapping functions (f and g) to compute adaptive scaling (γ) and shifting (β) parameters:

$$\gamma = f(x), \qquad \beta = g(x)$$

The resulting coefficients γ and β are then applied feature-wise (channel-wise) to the intermediate activation tensor F_c within the corresponding convolutional layer. Given a channel index c, the mathematical formulation of the FiLM operation is defined as:

$$FiLM(F_c) = \gamma_c \cdot F_c + \beta_c$$

In this framework, the scalar multiplier γ_c scales the activations of the target feature map channel to selectively amplify or attenuate spatiotemporal cues, whereas the bias translation β_c introduces an adaptive threshold shift to alter the neural activation pattern based on the biological or clinical context. Consequently, the convolutional layers cease to operate purely on visual input isolated from context; instead, they undergo dynamic restructuring to extract features conditioned precisely on the specific patient history and risk profiles.

CHAPTER 4

LITERATURE REVIEW

As it is a valuable but challenging procedure due to its invasive and complicated nature, many methods were employed to predict FFR.

4.1 Commercial software predicting FFR without AI

There are commercial non-invasive techniques that employ CFD on coronary angiograms that predict FFR. However, these techniques have specific requirements for angiograms, such as multiple imaging angles, certain image quality, high costs, and have long calculation times, like hundreds of seconds [50] [51] [52]. A recent study argues that FFR pressure wire through a cardiac catheter can cause an adverse impact on FFR and this impact is not considered in existing CFD-based noninvasive FFR (FFR_{CFD}) technologies [53]. Major companies offer FFR_{CFD} prediction software. Overall accuracy of FFR_{CFD} was found to be low in the gray zone (0.75 to 0.80) and often followed by invasive FFR in clinical situations [52] [54]. Fearon et al. [55] used a commercial software that tries to reconstruct a 3D representation of the coronary tree using multiple-angle angiograms and estimate flow using calculated vessel volume. Zhang et al. [56] used commercial software to construct a 3D reconstruction of CAs utilizing cardiac computed tomography angiography (CCTA) images.

4.2 Commercial software using AI

Several approaches exist in the estimation of FFR using AI. Commercial software using AI is already available. Roguin et al. [57] used an ANN network to calculate

the FFR value by multiple-angle angiograms directly from the angiography machine in 37 seconds. Although they reported 90% accuracy level for predicting a FFR value lower or higher than 0.8, the dataset comprised 31 patients, which is not enough to make a decision. Griffin et al. [58] used CCTA images to train an AI model and compared the results with lab-interpreted QCA and measured FFR. The commercial software uses convolutional network models (VGG-19 and 3D U-Net). Lab-interpreted QCA was used for ground truth. The AI model performed similarly to QCA, but both were inferior to FFR when it comes to sensitivity.

4.3 AI applied to patient anatomical data

A common approach is to gather data about the patient's physiological characteristics. This approach is also utilized together with image or video processing.

Zimmerman et al. [59] used resting aortic and distal coronary pressure curve data trained by CNN and recurrent neural networks (RNN) to predict FFR values. They found the diagnostic accuracy 77-79%. They concluded that only resting pressures are not enough and hyperemia induction should remain the standard practice in FFR prediction.

Lin et al. [60] used CCTA-based quantitative plaque features and PET/CT (positron emission tomography) based myocardial blood flow data to train an XGBoost network and compared results with FFRCT derived from CCTA. An independent laboratory performed plaque quantification. FFRCT was derived using dedicated commercial software. Results were consistent with FFRCT but inferior to invasive FFR.

Lee et al. [54] used morphological features such as flow and geometrical features which were extracted from CCTA to train two neural networks and six ML algorithms and compared results with measured FFR value. CCTA images were processed and morphological features of coronary arteries were extracted. Synthetic

vessel models were generated based on patient data to amplify the data quantity. The multi-layer perceptron model showed the highest gray zone accuracy percentage of 98.9%, but the accuracy was 44.1%. Models were tuned only for the best performance in the gray zone. Best accuracy was 67.6%.

4.4 AI Applied to CCTA Images

Another approach is to process non-invasive CCTA images. Candemir et al. [61] used CCTA images to train a 3D-CNN network to detect stenoses, and compared results with expert review. First a centerline extraction method was applied to CCTA sequences to locate coronary arteries and their branches within the CCTA image datasets. A straightened longitudinal view of vessel was acquired. Clamping and length equalization was performed. Data was applied to a 3D-CNN architecture. Although average classification accuracy was 90.9%, positive predictive value was 58.8%. Prepared vessel views may also be used for flow rate detection, but it wasn't mentioned. Fossan et al. [62] used CCTA images to train feed forward neural network (FFNN) and compared results with measured FFR value. A hybrid approach was proposed. Vessel segmentation, 3D modeling and 3D flow calculation was done using commercial software. The dataset was very limited to 90 patients, but neural network performed better than CFD model. And purely data-driven FFNN performed better than combined physics-informed FFNN. Neural networks had a positive bias (predicts higher FFR values), whilst the CFD model had a negative bias.

4.5 AI applied to ICA key frames

A third approach is to process the angiography key frame or part of the key frame that shows the lesion and its surroundings. Using a single image limits outcomes since coronary vessels are 3D objects and stenoses can be eccentric, and sometimes more than one projection is needed to analyze physiology. Selecting one frame among many frames of ICA may result in a loss of information.

Zhou et al. [63] used ICA video key frames of right coronary artery obtained from right anterior oblique and left anterior oblique projections. They used ResNet-18 for key frame extraction and U-Net for vessel segmentation. Results were compared with expert reviews. Stenosis size is measured by counting pixels on the key frame, which is not different from what expert is doing.

Cho et al. [64] measured angiographic lumen diameters, computed 24 angiographic parameters based on the diameter plot, and added four clinical features to form a dataset. They trained the XGBoost algorithm and compared the results with the measured FFR value. Angiographic lumen segmentation on key frames was done using commercial edge detection software. Erroneous segments were manually corrected, increasing time spent and sacrificing automation. Angiography-based ML model predicted FFR≤0.80 with a sensitivity 84%, a specificity 80%, and an overall accuracy of 82%. Although not directly related to FFR estimation, the endpoints comply with the intent of the FFR procedure.

Mahendiran et al. [65] used ICA image patches to train ResNet18 algorithm. They compared the results with expert review, commercially quantitative flow rate (QFR), and stenosis parameters against the actual patient condition. They used two angiographic views and tried to estimate future culprit lesions. Feature engineering was required to prepare image patches. DL algorithm had 0.78 accuracy, whereas expert review, stenosis diameter, stenosis area and QFR accuracy were 0.58, 0.58, 0.60 and 0.70 respectively. These results indicate that DL outperforms other approaches in predicting future culprit lesions.

4.6 AI Applied to ICA Videos

The last approach is to process the angiography video directly or after preprocessing. Zhao et al. [66] used ICA videos to train a CNN for vessel segmentation and compared the results with manually calculated CFR (coronary flow reserve). Two loops of the same angiographic projection, with and without hyperemia, are used to

calculate the CFR. CFR is different from FFR as it is defined as hyperemia-to-rest flow ratio, and measures microvascular dysfunction. Patients with FFR>0.8 but low CFR have worse prognosis. FFR and CFR disagree in approximately 32% of the cases. Vessel length is calculated and length/time curve is derived using frame count. Flow velocity is achieved by fitting a straight line to length/time curve. CFR is calculated by dividing hyperemic flow velocity to rest flow velocity. Same method is applied by a laboratory to calculate CFR_{manual}. Automatically calculated CFR_{auto} showed moderate agreement with CFR_{manual}. Manual calculation of CFR is not feasible.

Langlais et al. [67] retrospectively collected 182.418 coronary angiography videos in 5 years and introduced an AI-driven pipeline that detects the coronary vessel type, finds stenosis, stabilizes stenosis location, applies segmentation to detect stenosis affected vessel and predicts stenosis blockage percentage. A number of different AI models were used including Xception, RetinaNet, Swin3D models. Final step of the pipeline, stenosis percentage calculation step had around 20% MAE when clinical reports used as ground truth and 7.75% MAE when QCA is used as ground truth. Despite the vast effort expended, the stenosis percentage used for training and testing is based on clinician interpretation from coronary angiography videos, thus predicting FFR without a solid reference and resulting in similar mean error with ground truth.

Mineo et al. [68] used a private dataset comprising 778 coronary angiographies collected retrospectively. They used a Resnet3D model pretrained on the Kinetics-400 dataset for feature extraction and then fed features into a factorized spatio-temporal transformer for classification and regression. Classification accuracy was 89.4% at its best. Regression predictions resulted in a MAE of 0.063, meaning 6% mean error, which is not enough to make good predictions.

4.7 AI applied to synthetic and other kinds of data

Xie et al. [69] used 6600 synthetic coronary trees to train a Physics-informed neural network (PINN). Synthetic coronary trees were simulated for blood pressure ground truth values. Verification was done with data of 143 X-ray and 40 CT angiography videos of real patients. Results showed accordance with CFD results of the patients. However, PINN model has poor performance without the morphology and boundary condition data, which are hard to come by in real patients.

Carson et al. [70] generated virtual single and multi-vessel networks, trained multivariate polynomial regression (MPR), FFNN, and long-short term memory on synthesized data, and compared results with 25 real patients' data. Diagnostic accuracy of 1-D single vessel model was 72% with real world data. Absence of real-world data is the main limitation of the study.

Sun et al. [71] generated 3028 idealized stenosis models to train a back propagation neural network (BPNN) and tried to achieve stenosis resistance. They tested the network with 30 patients' FFR data. Stenosis resistance is a key parameter to calculate FFR_{CT}. FFR_{CT} was calculated by 3D CFD simulation of artery models reconstructed from CCTA images of the patients. BPNN-based FFRCT and CFD-based FFR_{CT} are compared. FFR_{CT} based on the BPNN-predicted stenosis resistance had an excellent accuracy rate compared with the clinically measured FFR. Because of the CFD analysis, blood and vessels are assumed to be ideal.

Muscogiuri et al. [72] used computed tomography perfusion (CTP) images to train a CNN network and compared CCTA, CTP, and CCTA combined with CTP results with measured FFR values. Image preprocessing was done manually to decrease computational burden. It was found that CTP significantly improved the detection of hemodynamically significant CAD compared to CCTA alone. Best sensitivity and specificity compared to invasive FFR were 93% and 83%, respectively.

Xue et al. [73] created a 3D hybrid model using ICA key frame and 2D CCTA model, used 13000 synthetic and 180 real coronary trees to train a multilayer perceptron

network and a bidirectional RNN, and compared results with measured FFR values. The hybrid model is formed with the overall spatial structure from the CCTA centerline and ICA-derived vessel radii. The test dataset was small, comprised of 67 vessels from 40 patients. Sensitivity was 100% for all types of vessels and vessel segments, but specificity varied from 71.43% to 94.44%. Statistically, the dataset was not enough to come to a conclusion.

Farajtabar et al. [74] used geometrical features of the artery and velocity boundary conditions to train an FFNN. These features were obtained by carrying out CFD simulations on synthetically generated vessels. Given the features for each artery in the arterial network, FFNN predicts pressure and velocity values at a given point in the artery network. Since pressure values at any point can be predicted, FFR can be easily calculated. Results were tested on three real coronary artery geometries. Results were highly accurate. Still, gathering features for each artery is a time-consuming job, and only three real cases are not enough to come to a conclusion.

Cha et al. [75] used features obtained from optical coherence tomography (OCT) to train a Random Forest algorithm with added features and compared results with measured FFR values. The dataset comprised 356 individual lesions in 130 patients. Seven features with different weights were selected from the OCT image analysis. These features were percent area stenosis, minimal lumen area, target vessel, distal lumen area, lesion length, plaque area, and proximal lumen area. The sensitivity, specificity, and accuracy of OCT-based ML-FFR were 98.3%, 61.5%, and 91.7%, respectively. OCT is not a standard device in hospitals. Also, required feature engineering slows the process.

4.8 Challenges and Limitations in FFR Prediction

Currently, FFR estimation is attempted by processing non-invasive methods with and without AI. However, ICA maintains its superiority as it has the opportunity for treatment in the same session and provides more information than non-invasive

methods. The main limitations of methods other than ICA that estimate FFR are as follows: Physiological data-based studies extract features about patient physiology from imaging and diagnostic tools and patient data. Feature extraction requires expertise and slows the process. The most used non-invasive procedure CCTA, is recommended for low-to-medium risk patients. CCTA lacks the temporal information that ICA supplies, and patient heart rate is required to be low to acquire good quality images. Usage of CCTA may be limited in patients with tachycardia, established CAD previous stents, or extensive coronary calcification [3].

The number of studies estimating FFR with ICA is limited. Studies that focus on angiography key frame discard flow information which is nontrivial in expert decision. Since some parts of the image don't show any vessel or stenosis, image must be preprocessed automatically or manually to single out relevant areas. The limitation of view to a single angle is not compatible with the 3D nature of the CAs. On the other hand, processing the angiography output as a video needs excessive computing power. Newer angiography machines have down to 0.1mm imaging precision, and one video of one angle can be up to 100 Mbytes. Naturally, gathering angiography data is difficult. For prospective studies, it is difficult to get permission from ethical committee, and it is exhausting to work with physicians considering the daily operation of the hospitals. Patients should sign informed consents. Physicians should follow established workflow. For retrospective studies, it is difficult to gather data from hospital databases, mainly because there are different physicians taking medical notes, and angiography databases are not perfectly integrated to hospital databases. There are differences in angiography device output resolutions and image qualities. Physicians have different opinions about angiography angles, vessels and lesions and have their own style. Multicenter studies take years to finish. To our knowledge, there are no commonly available video angiography datasets with FFR values attached. Most of the studies on the subject are not based on large datasets. Multicenter large datasets are only seen as commercial software evaluation. FFR estimation problem could be seen as a classification problem if we look for FFR>0.8 or not. But if an FFR value is to be predicted, problem becomes a regression problem.

Angiography images are X-ray pictures thus tissues, organs, blood, bones in the body are seen in images as background. Some patients have stitches, stents, implants that are seen in angiography. Multiple lesions on the same vessel or on different vessels on the same image need to be taken into consideration. Every patient has different cardiac characteristics, i.e. vessels differ in size and shape. Some patients have congenital heart defects that mess up with the blood flow characteristics. These drawbacks spoil acquired data and render it unusable.

On the other hand, some of the invasive and non-invasive diagnostic tools like OCT, cardiac MRI, intravascular imaging, CTP and even FFR measurement systems are not available in every center. Because of this, the number of studies based on these methods are limited.

Although recent efforts to derive FFR through non-invasive or less invasive modalities represent a significant step forward, these approaches require further refinement to achieve the clinical precision necessary for routine practice. Given the frequent and urgent need for hemodynamic assessment during ICA, integrating FFR estimation directly into the ICA workflow offers substantial diagnostic advantages. However, as [76] emphasizes, current CFD models are inherently limited by the assumption of rigid vessel walls, failing to account for the dynamic interplay between plaque biology, vasodilation, and stenosis resistance. Furthermore, these models often struggle to characterize coronary microcirculatory resistance due to the omission of complex, patient-specific compensatory mechanisms. Therefore, there is a critical need to transition from purely geometric CFD models toward more biologically informed frameworks that integrate these pathophysiological nuances to enhance the diagnostic reliability of non-invasive FFR estimation.

CHAPTER 5

DATASET

To date, there is no universally standard or publicly available multi-modal dataset that synchronizes functional invasive pressure wire measurements with live angiographic sequences and patient-specific blood profiles. To address this gap in the literature, a comprehensive clinical repository was constructed retrospectively. This chapter details the multi-stage data acquisition protocols, institutional ethics framework, rigorous cohort selection criteria, visual and temporal signal processing operations applied to the angiographic video streams, and the engineering of novel, biology-driven hemorheological feature vectors.

5.1 Institutional Ethics and Data Acquisition

The execution of this retrospective, single-center study was formally reviewed and approved by the Institutional Review Board of Ankara Bilkent City Hospital under Approval Numbers E2-23-5672 and E2-26-13895. Due to the strictly retrospective design of the analysis and the utilization of fully de-identified legacy records, the requirement for active patient informed consent was officially waived by the board.

A comprehensive, cross-relational search was systematically executed across digital medical archives, including patient epicrisis reports, ICA operational logs, and localized cardiologist charts. Initially, digital archives identified a consecutive cohort of 1,255 patients who underwent invasive, pressure-wire-based FFR assessment for suspected or documented coronary artery disease between February 1, 2019, and October 1, 2023. This comprehensive search yielded a baseline video repository consisting of 16,920 raw angiographic sequences, totaling approximately 616 gigabytes of digital data.

5.2 Cohort Selection and Exclusion Criteria

To eliminate physiological confounding factors and guarantee a homogenous training distribution for the ML models, strict exclusion criteria were implemented sequentially:

- Data Pairing Truncation: First, 466 patients were excluded from the baseline repository because their objective invasive FFR values could not be definitively paired or synchronized with the corresponding digital angiographic video sequences.
- Anatomical Uniformity Standardization: Among the remaining 789 patients, clinical logs indicated that 751 patients (95.2%) underwent functional evaluation specifically for Left Anterior Descending (LAD) coronary artery stenoses. Consequently, patients presenting with isolated Right Coronary Artery (RCA) or Left Circumflex (LCx) artery interventions were completely excluded to ensure strict anatomic and geometric uniformity across the dataset.
- Visual Quality Filters: Stenosis coordinates were meticulously extracted from cardiologist notes, and target lesion windows were mapped using custom-developed frame annotation software. Following this, each video sequence underwent independent visual quality screening. Video arrays characterized by low contrast resolution, severe movement artifacts, geometric vessel overlaps obscuring the target lesion, the presence of previous coronary stents, or overlapping diagnostic catheters were meticulously eliminated.

Following these filtering steps, a highly curated dataset consisting of 256 high-fidelity, single-stenosis, grayscale angiographic video segments from 179 distinct patients was finalized, with each sequence directly mapped to its corresponding wire-derived functional FFR value.

5.3 Angiographic Image and Temporal Signal Processing

All baseline invasive coronary angiograms were acquired using standard clinical protocols and preserved in lossless Digital Imaging and Communications in Medicine (DICOM) format at a native spatial resolution of 1000x1000 pixels.

5.3.1 Temporal Standardization and Padding

To eliminate initial and terminal catheter-positioning noise and contrast artifacts, the first and last frames of each raw sequence were discarded. The native length of the clinical video clips ranged from 29 to 144 frames. Dynamic physiological analysis demonstrated that the contrast medium comprehensively opacified the entire epicardial tree and washed away completely within 64 frames.

Therefore, all dynamic sequences were strictly standardized to a uniform temporal dimension of 64 frames to fit DL inputs. For short sequences (starting from 29 frames), the final fully opacified frame was frozen and repeated up to the 64-frame threshold. Conversely, longer sequences were clipped exactly at the 64th frame, ensuring a homogeneous temporal dimension for multi-frame DL processing. The descriptive statistics of the final dataset are presented in Table 5.1 and Table 5.2, while the absolute FFR target distribution is visualized in Figure 5.1.

Table 5.1 Dataset Statistics

FFR Value	*< 0.75*	*0.75-0.8 (Gray Zone)*	*> 0.8*	*Total Number of Videos*
Male	23 (12,64%)	48 (26,37%)	111 (60,99%)	182 (71,1%)
Female	2 (2,7%)	14 (18,92%)	58 (78,38%)	74 (29,9%)
Total	25 (9,77%)	62 (24,21%)	169 (66,02%)	256 (100%)

Abbreviations: FFR, fractional flow reserve

Table 5.2 Dataset Statistics (continued)

	Avg. FFR	*Avg. Age*	*Patients*	*Frames*	*Avg. Frames*
Male	0,8267	55,15	124 (69,27%)	10682 (70,21 %)	59,34
Female	0,8555	64,16	55 (30,73%)	4532 (29,79%)	59,63
Total	0,835	57,76	179 (100%)	15214 (100%)	59,43

Abbreviations: Avg, average; FFR, fractional flow reserve

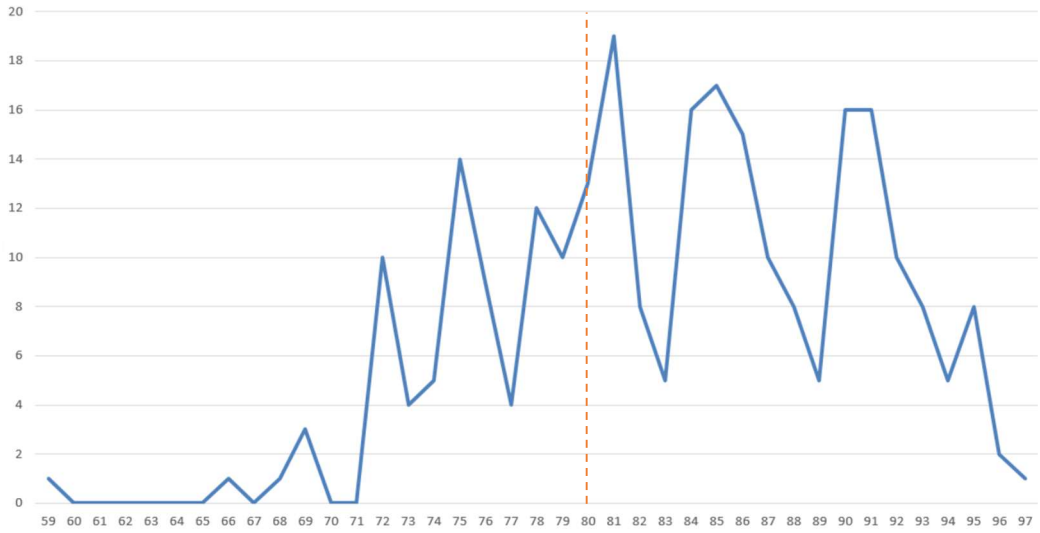

Figure 5.1. FFR Value occurrences in dataset.

5.3.1 Multi-Scale Spatial Datasets (Downscaling and Cropping)

To systematically evaluate the impact of spatial resolution and model parameter weights, the native 1000x1000 pixel sequences were downscaled using the OpenCV inter-area interpolation algorithm to form three secondary datasets:

- Downscaled Dataset 1: 500x500 pixels
- Downscaled Dataset 2: 250x250 pixels
- Downscaled Dataset 3: 125x125 pixels

In addition to downscaling the global view, unscaled regional crops focused directly on the manually logged stenosis coordinates were extracted to isolate local

narrowing dynamics. These cropped secondary sets were processed at 480x480 pixels and 256x256 pixels without spatial scaling (Figure 5.2).

Figure 5.2. Cropping around stenosis coordinates

5.3.2 Spatio-Temporal Velocity Fields (Optical Flow)

To explicitly capture multi-scale velocity fields and fluid motion vectors independently of raw pixel intensity, dense spatio-temporal dynamics were extracted across two distinct frameworks:

- RAFT Optical Flow: Utilizing the PyTorch Recurrent All-Pairs Field Transforms (RAFT) [77] pretrained DL architecture, motion fields were converted to RGB optical flow streams at resolutions of 480x480 and 240x240 pixels (Figure 5.3).
- Farneback Optical Flow: Utilizing the traditional OpenCV Farneback optical flow algorithm, a secondary differential motion dataset was processed at 240x240 pixels (Figure 5.4).

In total, a comprehensive suite of eight multi-modal imaging sub-datasets was generated for cross-computational validation, as outlined in Table 5.3.

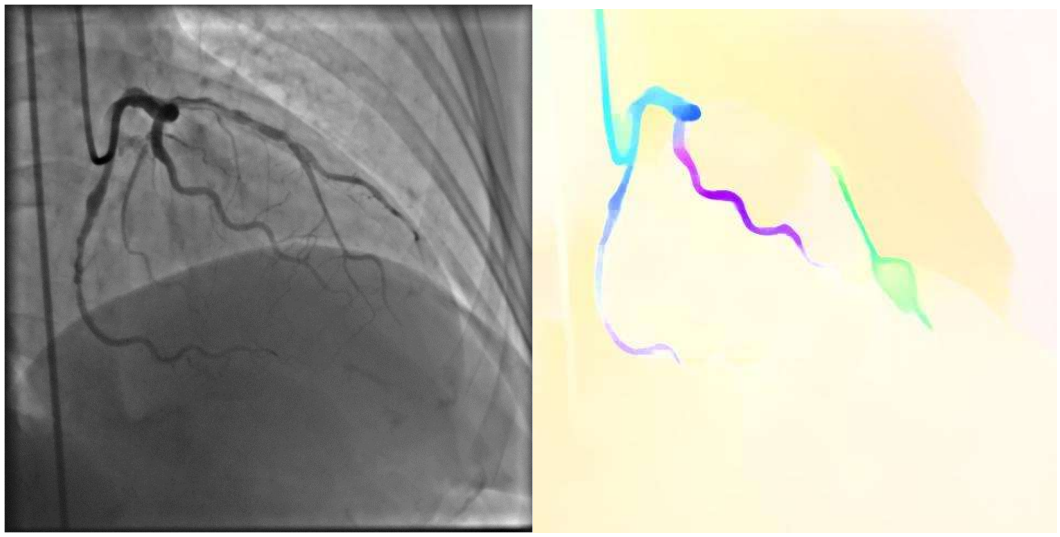

Figure 5.3. Original image and RAFT Optical Flow Model output

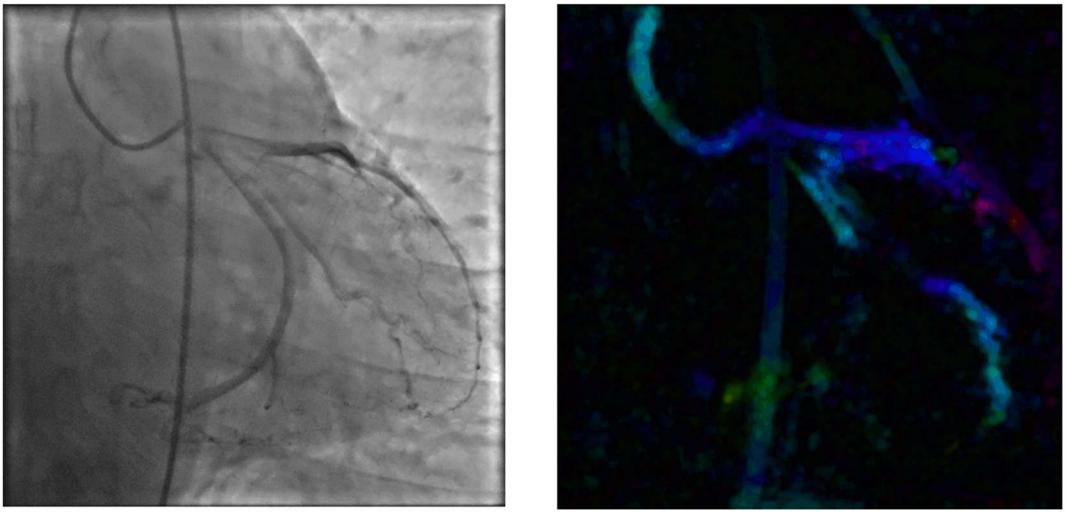

Figure 5.4. Original image and Farneback Optical Flow output

Table 5.3 Prepared Datasets

Dataset	*Resolution*	*Color*	*Method*	*File Size*
Original	1000x1000	Grayscale	None	135 MB
Downscaled 1	500x500	Grayscale	OpenCV InterArea	33,8 MB
Downscaled 2	250x250	Grayscale	OpenCV InterArea	8,46 MB
Downscaled 3	125x125	Grayscale	OpenCV InterArea	2,11 MB
Cropped 1	480x480	Grayscale	Window Crop	31,2 MB
Cropped 2	256x256	Grayscale	Window Crop	8,87 MB
Optical Flow 1	480x480	RGB	Pretrained RAFT	92,9 MB
Optical Flow 2	240x240	RGB	Pretrained RAFT	23,2 MB
Optical Flow 3	240x240	RGB	OpenCV Farneback	25,2 MB

Abbreviations: MB, megabytes; RAFT, Recurrent All-Pairs Field Transforms; RGB, red green blue

5.4 Hematological Dataset and Feature Engineering

To operationalize the biophysical hypothesis of blood as a non-Newtonian cellular suspension, patient-specific hematological profiles collected during the pre-procedural baseline window were integrated into the multimodal dataset. Routine blood assays performed universally prior to the index ICA procedure were retrieved to capture the patient's baseline rheological fingerprint. Whole blood samples were collected in standard ethylenediaminetetraacetic acid (EDTA) tubes and analyzed using automated hematology analyzer systems.

For each patient, raw hematological parameters were extracted, including RBC, HGB, HCT, MCV, MCHC, RDW, PLT, PCT, and MPV. Gender was formalized as a binary operational variable (assigned as 1 for female and 0 for male), and chronological age at the time of the FFR procedure was recorded to complete the baseline demographic matrix. These tabular parameters were structurally mapped and linked to their respective 2D ICA video sequences, using the angiographic filenames as unique relational keys, to serve as the baseline feature layers for subsequent multi-dimensional ML architectures.

While routine CBC panels provide comprehensive leukocyte metrics, all WBC indices were strategically excluded from the ML framework. From a fluid-mechanics and stochastic-modeling perspective, nucleated leukocytes are highly volatile and responsive to external stimuli, such as acute infections or systemic inflammation. Including WBC counts would introduce significant noise and potential predictive bias, as their fluctuations reflect immune responses rather than baseline hemodynamic resistance. Consequently, feature engineering focused exclusively on anucleated erythrocytes and platelets, which serve as stable indicators of intravascular fluid dynamics. Similarly, chronological age was excluded as an independent predictor in the final fluidic framework, based on the hypothesis that cumulative cellular and volumetric alterations inherently encapsulate biological vascular wear far more accurately than calendar years.

To translate the raw laboratory metrics into computable biophysical vectors capable of capturing flow disruptions under pathological shear stress, three novel hemorheological indices were engineered:

- Erythrocyte Fragility Index (EFI): Designed to capture the simultaneous impact of erythrocyte size heterogeneity and internal cellular density. It indexes the micro-rheological vulnerability and rigidity of the red cell mass, where elevated values serve as a proxy for a less deformable erythrocyte collective susceptible to trajectory deviations: EFI = RDW (%) x MCHC(g/dL).
- Cell-Free Hemoglobin Index (CFHI): Developed to mathematically flag the subclinical footprint of intravascular microhemolysis. While automated analyzers measure total hemoglobin across both intact and ruptured compartments at low levels of hemolysis, hematocrit uniquely reflects the true packed cellular volume. The ratio thus serves as an index of the hemoglobin fraction that has escaped the cellular capsule to drive endothelial NO scavenging: CFHI = [HGB (g/dL) / HCT (%)] × 100.
- Erythrocyte Platelet Volume Balance (EPVB): Introduced to model the spatial volumetric balance and structural carrying capacity between the

central driving erythrocyte stream and the marginated peripheral platelet barrier: EPVB= HCT (%) / PCT (%).

Ultimately, these three engineered biophysical indices (EFI, CFHI, EPVB), alongside raw MPV and gender dictating baseline macrovascular diameters and hematological setpoints, constitute the five core features of the blood fluidity vector. These features and their specifications are given in Table 5.4.

To ensure stability during neural network convergence and prevent feature dominance, a rigorous per-fold Z-score normalization was applied to all five rheological features using the StandardScaler architecture of the sklearn ML library, isolating the training folds to prevent data leakage. The StandardScaler was fitted on the training fold only and applied to the validation and test folds. Furthermore, while the absolute invasive FFR values within the curated clinical repository ranged between [0.59, 0.97], the target variable was linearly mapped from its initial [0.5, 1.0] boundary to a normalized [0, 1] dynamic range. Under this transformation matrix, the critical clinical ischemia threshold of 0.80 is mathematically mapped to exactly 0.60, optimizing the loss function's sensitivity around the clinical decision boundary.

Table 5.4 Input Feature Matrix Specifications for the Multimodal Blood Fluidity Model

Feature Name	Symbol / Formula	Data Type	Biophysical / Modeling Rationale
Gender	Gender (0,1)	Binary	Calibrates baseline macrovascular diameters and hematological setpoints.
Mean Platelet Volume	MPV (fL)	Continuous (Raw)	Identifies baseline platelet structural state.
Erythrocyte Fragility Index	EFI = RDW (%) x MCHC(g/dL)	Continuous (Engineered)	Proxies RBC rigidity, size heterogeneity, and propensity for trajectory deviation.
Cell-Free Hemoglobin Index	CFHI = [HGB (g/dL) / HCT (%)] × 100	Continuous (Engineered)	Flags subclinical intravascular microhemolysis and the risk of endothelial NO scavenging.
Erythrocyte Platelet Volume Balance	EPVB = HCT (%) / PCT (%)	Continuous (Engineered)	Models the spatial volumetric balance between central axial RBC flow and peripheral platelet margin.

Abbreviations: CFHI, cell-free hemoglobin index; dL, desiliter; EFI, erythrocyte fragility index; EPVB, erythrocyte platelet volume balance; fL, femtoliter; g, gram; HCT, hematocrit, HGB, hemoglobin; MCHC, mean corpuscular hemoglobin concentration, MPV, mean platelet volume; NO, nitrous oxide; PCT, plateletcrit; RBC, red blood cells; RDW, red cell distribution width

5.4.1 Baseline Statistical Analysis & Predictive Potential

The statistical evaluation of blood parameters and hemorheological features was performed in four distinct stages, utilizing IBM SPSS Statistics (v26.0). Normality was assessed using the Shapiro-Wilk test, with Student's t-tests applied to normally distributed variables (mean ± SD) and Mann-Whitney U tests to non-normally distributed variables (medians and ranges). Categorical data were compared using Chi-square tests. In this context, statistical significance was defined as a p-value of less than 0.05, representing a threshold at which observed effects are highly unlikely to be due to random chance in the data.

We first quantified physiological variance across gender to isolate potential confounding factors (Table 5.6). This analysis revealed significant gender-based dimorphism: the female cohort was significantly older (64.2 vs. 55.5 years, $p < 0.001$), while male patients had lower median FFR values ($p = 0.002$) and a higher prevalence of functional ischemia ($p = 0.027$). These findings necessitate a gender-conditioned approach in our predictive modeling.

We further examined the discriminatory capacity of our features relative to the functional ischemia threshold (FFR ≤ 0.80). Univariate comparisons indicated that individual markers lack robust diagnostic power in isolation. While baseline MPV (p=0.066) and EFI (p=0.074) exhibited descriptive trends, the lack of statistical significance suggests that ischemic potential is governed by a multi-dimensional interaction rather than a single hematological threshold.

Continuous linear correlation assessments confirmed these baseline limitations. As shown in Table 5.6 , only baseline MPV demonstrated a weak but statistically significant correlation with absolute FFR values ($r = 0.159$, $p = 0.034$). No other parameter showed continuous linear dependence, establishing a baseline limit for conventional univariate analytical frameworks.

Finally, to assess the latent predictive utility of our five-feature vector, we constructed a binary logistic regression model (Table 5.5). Independent variables were standardized via Z-score transformation. Model performance was validated

using the Omnibus test and the Nagelkerke statistic. Unlike the univariate stages, this multivariate integration captured a significant hemodynamic signature (p = 0.037). The individual contribution of EFI (p=0.055) and the overall AUC of 0.652 demonstrate that, while individual markers exhibit modest predictive power, their collective integration acts as a valuable surrogate for systemic ischemic potential.

This multi-layered validation confirms that our hemorheological indices operate within a non-linear feature space, justifying their integration into our DL pipeline.

Table 5.5 Hemorheological Predictors of FFR-defined Ischemia in the Logistic Regression Model

Predictor	B (Beta)	*p* value	Odds Ratio (Exp(B))	95% CI
Female Gender	-0.801	0.061	0.449	0.194–1.039
EFI	0.326	0.055	1.386	0.994–1.932
MPV	-0.194	0.278	0.824	0.581–1.169
CFHI	0.096	0.588	1.101	0.778–1.558
EPVB	0.063	0.721	1.065	0.753–1.508

Abbreviations: CFHI, cell-free hemoglobin index; CI, confidence interval; EFI, erythrocyte fragility index; EPVB, erythrocyte platelet volume balance; MPV, mean platelet volume

Table 5.6 Baseline demographic, Clinical, and Comprehensive Hemorheological Profiles of the Study Population, stratified by gender, Functional Ischemia Status, and Continuous FFR Correlations

	Gender			Ischemia			FFR Correlation	
	Male (n=124, 69.3%)	Female (n= 55, 30.7%)	*p* value	FFR≤0.80	FFR>0.80	*p* value	r_s	*p* value
Age (years, mean±SD)	55.5 ±9.2	64.2 ±9.4	**<0.001**	57.9 ±7.6	58.3 ±11.1	0.830	0.035	0.646
HGB (g/dL, mean±SD)	14.6 ± 1.6	12.7 ±1.3	**<0.001**	14.1±1.7	14 ± 1.8	0.581	-0.060	0.424
HCT (%, median/mean±SD)	44.5 (27.8-54.2)	39.3 (25.6-50)	**<0.001**	42.6±5	42.4±5	0.909	-0.020	0.789
RBC (x10^{12}/L, mean±SD/median)	5.1 ± 0.5	4.6±0.6	**<0.001**	4.9 (3.1-6)	4.9 (2.9-6.8)	0.773	-0.014	0.847
MCV (fL, median)	88.1 (64.7-98.8)	86.3 (60.3-110)	**0.016**	87.8 (60.3-110)	87.4 (63.1-98.8)	0.800	0.024	0.753
MCH (pg, median)	29.4 (20.9-33.2)	28.3 (19.8-40.2)	**<0.001**	29.2 (19.8-40.2)	29.2 (20.1-33.2)	0.713	0.063	0.403
MCHC (g/dL, median)	33.3 (28.7-36.6)	32.7 (29.8-35.1)	**0.001**	33.1 (30.4-36.6)	33.1 (28.7-36.1)	0.354	-0.100	0.183
RDW (%, median)	13.6 (12-16.7)	14.3 (11.8-19.3)	**0.001**	14 (12.6-16.7)	13.7 (11.8-19.3)	0.376	-0.011	0.883
PLT (x10^{9}/L, median)	247.5 (70 - 484)	285 (134 - 579)	**0.006**	252 (70-579)	265 (112-486)	0.580	0.042	0.580
MPV (fL, median)	8.2 (6.5-11.6)	8.5 (6.7 - 13.2)	**0.034**	8.1 (6.5-11.5)	8.3 (6.6-13.2)	0.066	0.159	**0.034**
PCT (%, median)	0.20 (0.06-0.38)	0.25 (0.14-0.57)	**<0.001**	0.20 (0.06-0.57)	0.22 (0.10-0.38)	0.085	0.137	0.067
PDW (%, mean±SD)	51.8± 8.6	51.1 ±9.1	0.601	52.6± 7.7	51.1±9.6	0.263	0.004	0.958
EFI (median)	454 (396-556)	467 (388-580)	**0.048**	467 (417-556)	453 (388-581)	0.074	-0.082	0.276
CFHI (median)	33.3 (28.6-36.7)	32.7 (29.8-35.1)	**0.001**	33.2±1.2	32.9±1.4	0.142	-0.107	0.152
EPVB (median)	218 (110-608)	158 (73-294)	**<0.001**	218 (73-608)	187 (90-413)	0.101	-0.137	0.068
FFR (median)	0.83 (0.59-0.96)	0.87 (0.66-0.97)	**0.002**	0.76 (0.59-0.80)	0.87 (0.81-0.97)	<0.001	-	-
Positive FFR (n, %)	48, 26.8%	12, 6.7%	**0.027**	-	-	-	-	-

Abbreviations: CFHI, cell-free hemoglobin index; dL, desiliter; EFI, erythrocyte fragility index; EPVB, erythrocyte platelet volume balance; FFR, fractional flow reserve, fL, femtoliter; g, gram; HCT, hematocrit, HGB, hemoglobin; L, liter; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration, MPV, mean platelet volume; PCT, plateletcrit; PDW, platelet distribution width; pg, picogram; RBC, red blood cells; RDW, red cell distribution width; r_s, Spearman's rank correlation coefficient; SD, standard deviation

CHAPTER 6

EXPERIMENTS

The experimental framework was structured in three sequential phases: dataset preparation, feature analysis, and architectural ablation. Given the high-dimensional nature of angiography video sequences (up to 135 MB per file), we developed a custom data generator to enable efficient batch processing. Initially, we conducted a benchmarking study to identify the most discriminative model and feature sets. While literature [4] [5] [21] suggests a physiological link between blood rheology and coronary perfusion, our experimental results indicate that these relationships are highly complex and non-linear, necessitating a ML approach to extract meaningful diagnostic signals. To evaluate the contribution of each modality, we designed four distinct architectures for ablation: a pure tabular baseline, a pure video model, and two multimodal fusion strategies (Late Fusion and FiLM). Recognizing that our dataset exhibits a class imbalance—with a distribution ratio of approximately 1:1.94 between positive (ischemic) and negative (non-ischemic) cases—we implemented balanced oversampling and positive class weight in the classification loss. The task of FFR prediction is formulated as a dual objective: continuous regression for FFR value estimation and binary classification based on the clinical threshold of 0.80 [13]. In this context, FFR values ≤ 0.80 are defined as positive (ischemic), while values >0.80 are defined as negative. We prioritize the classification head's integrity because, clinically, a regression error that results in a threshold crossing (e.g., predicting 0.81 for a true 0.79 case) carries significant diagnostic consequences. Thus, our evaluation framework tracks metrics specifically to ensure the model's clinical reliability around this 0.80 decision boundary.

6.1 Performance comparison of common models with video-only training

Most DL architectures for video analysis are natively designed for classification tasks, where the output is a probability distribution across discrete classes. For instance, models trained on the Kinetics-400 dataset are optimized to predict labels such as specific human actions via Softmax activation. To repurpose these architectures for FFR estimation—a continuous regression task—we modified the final dense output layer to a single-output unit utilizing linear activation. Parameters were optimized for best result.

Six established architectures were evaluated for initial experimentation. PyTorch was employed for Video ResNet (R3D) [38], its Kinetics-400 pretrained variant, Video S3D (S3D) [40] and Video Swin Transformer Tiny Model (Swin) [46]. Conversely, ViViT [45] and ConvLSTM2D [42] were implemented via the Keras library. Additionally, a custom 3D-CNN [78] architecture was designed, consisting of three sequential 3D-CNN and Max Pooling layers (extended to four layers for 480 and 500-pixel resolution inputs), followed by dense and dropout regularization layers.

The dataset was split into 80% training and 20% validation sets. Training was conducted for 300 epochs per model, utilizing Mean Squared Error (MSE) for loss optimization and Mean Absolute Error (MAE) as the primary evaluation metric. Inputs were set to a fixed duration of 64 frames, with varying spatial resolutions.

Table 6.1 presents the best validation MAE results observed during training. To contextualize these metrics, we calculated the Mean Absolute Deviation (MAD) of our dataset as 0.05757. Since a dummy regressor predicting the median would yield an MAE equal to the MAD, we established this value as our baseline; models with an MAE lower than 0.05757 were considered to possess predictive capability.

Table 6.1 Performance comparison of standard models (MAE)

Model	*125x125*	*250x250*	*500x500*	*256x256 cut*	*480x480 cut*	*240x240 RAFT flow*	*480x480 RAFT flow*	*250x250 Opt. flow*
ViViT	0.05575	0.05566	0.05561	0.04766	0.05596	0.05574	0.05665	0.05249
R3D	0.08904	0.08273	0.11656	0.06800	0.07801	0.06917	0.06453	0.08362
R3D (pretrained)	0.23638	0.39839	0.15438	0.45796	0.68900	0.16473	0.10409	0.09536
3D-CNN	0.05036	0.05145	0.05280	0.05299	0.04984	0.05521	0.05382	0.05048
ConvLSTM2D	0.06023	0.06192	N/A	0.05425	N/A	0.05654	N/A	0.05199
S3D	N/A	0.55901	N/A	0.71590	N/A	0.60593	N/A	0.35824
Swin	0.06088	0.06058	N/A	0.06092	N/A	0.06084	N/A	N/A

Abbreviations: 3D, three dimensional; CNN, convolutional neural network; ConvLSTM2D, convolutional two dimension long-short term memory; Opt., optical; R3D, Resnet 3D; RAFT, Recurrent All-Pairs Field Transforms; S3D, separable three dimensional convolutional neural network; Swin, swin transformer; ViViT, video vision transformer

The experimental results revealed consistent performance patterns across models. Notably, the Kinetics-pretrained R3D model performed worse than the version trained from scratch and essentially reached random-guessing performance, highlighting the limited efficacy of transfer learning in medical imaging, where contrast dispersion dynamics differ significantly from natural video features. S3D, which separates spatial and temporal processing, failed to achieve satisfactory performance. High-resolution inputs posed significant hardware constraints, while resource-intensive models like ViViT and Swin Transformer offered no meaningful improvement over simpler architectures. Furthermore, datasets augmented with RAFT and Farneback optical flow inputs showed negligible performance differences. Ultimately, the 256x256 coordinate-cut dataset yielded the best results, as it allowed models to better focus on the stenosis region. The 256x256 coordinate-cut dataset was selected for further experimentation. Although ViViT displayed lower MAE in some instances, it did not translate to superior classification performance. Consequently, the 3D-CNN architecture was selected for subsequent experiments due to its rapid training cycle, lower resource footprint, and the ease of optimizing regularizing parameters.

6.1.1 Comparison with angiography-based studies

The model proposed in Mineo et al. [68] was implemented and trained on our 256x256 coordinate cut dataset. The best MAE performance indicated in the study was 0.063. Best MAE with 256x256 coordinate cut dataset was 0.05502, similar to other models and our own results.

Other relevant studies presented challenges in direct implementation. Langlais et al. [67] introduced a 6-step AI-driven pipeline. The final step of the pipeline, the stenosis percentage calculation step, had around 20% MAE when clinical reports were used as ground truth and 7.75% MAE when commercial QCA was used as ground truth. No further metric was given. Despite the vast effort expended, the stenosis percentage used for training and testing is based on clinician interpretation from coronary angiography videos, thus predicting FFR without a solid reference and resulting in similar mean error with ground truth. Since no code was shared, it was not possible to train the proposed model on our dataset.

Zhao et al. [66] calculated flow velocity by counting the number of frames needed to fill the vessel with blood. They achieved 83,4% F1 score, 91,9% sensitivity, and 76,3% positive prediction value. Due to the unavailability of source code for these studies, direct training on our dataset was not feasible.

Finally, commercial solutions [50] [55] [57] consistently report high diagnostic accuracy (88%–94%), yet these systems are fundamentally different from our approach as they mandate multi-angle (bi-plane) angiography inputs, whereas our model maintains high diagnostic robustness using only single-view video data.

6.2 Hemorheological Feature Analysis and Predictive Benchmarking

To determine the maximum achievable predictive capacity using strictly hematological markers, we employed the XGBoost [48] ML framework. Our objective was two-fold: first, to establish a baseline for individual parameter

performance, and second, to evaluate the efficacy of our engineered biophysical indices. Functional ischemia was defined as an FFR≤0.80. All experiments were conducted using a 5-fold cross-validation strategy, tested with 16 seeds, with performance quantified via the Area Under the Receiver Operating Characteristic Curve (AUC), accuracy, sensitivity, and specificity. While strong associations between WBV and FFR have been documented in literature [4] [5] [21], this study represents one of the first systematic attempts to quantify these relationships through a machine-learning-based classification framework.

The results (Table 6.2) reveal a significant discrepancy between raw hematologic data and our engineered features. While individual raw indices (e.g., RBC, MCV, RDW) displayed limited discriminative power (AUC < 0.50), the Engineered Hemorheological Vector provides the most robust discriminative performance (AUC: 0.602 ± 0.028). A crucial observation is the synergistic contribution of the selected features: while individual indices such as EFI exhibit high sensitivity (0.780), they often suffer from lower specificity (0.542), leading to a higher rate of false positives. By integrating MPV, CFHI, EPVB, and Gender into the engineered vector, we successfully elevated the specificity to 0.678, significantly enhancing the model's clinical utility. Several critical insights emerge from this benchmarking:

The all raw parameters configuration (AUC: 0.475) performed inferiorly to our engineered vector. This indicates that raw hematological data contains substantial hematologic noise that obscures the hemodynamic signal. Our engineered features—specifically the EFI, CFHI, and EPVB—effectively act as a low-pass filter, isolating the biophysical variables that directly correlate with distal perfusion variance.

The predictive ceiling observed across all blood-based models (AUC peaking at 0.607) validates the limitations of systemic peripheral markers in predicting localized coronary ischemia. This result provides the primary justification for our multimodal fusion approach; since blood markers alone capture only a fraction of the FFR variance, they must be spatially anchored to the coronary anatomy.

The superior performance of the Engineered Hemorheological Vector (AUC: 0.602) compared to raw indices validates our biophysical hypothesis: the interplay between erythrocyte fragility and platelet volume is a more robust indicator of coronary vulnerability than standard cell counts.

6.3 Designed Model Architectures

Following the establishment of the hemorheological baseline, we formulated a systematic experimental framework to explore the integration of these markers with angiographic data. Four distinct architectures were designed for ablation study: a pure tabular baseline, a pure video model, and two multimodal fusion strategies. To ensure statistical comparability, all models were trained concurrently using identical fold data, allowing for a precise evaluation of how different fusion methodologies—and the inclusion of hemorheological conditioning—impact diagnostic performance.

6.3.1 Tabular-Only model

The Tabular-only architecture serves as a baseline for evaluating the predictive power of hemorheological data in isolation. The model consists of two dense layers, each followed by a dropout layer to prevent overfitting. Given that our input dimensionality is constrained—at most 14 parameters, with the Engineered Hemorheological Vector containing only 5—the model is intentionally kept lightweight, comprising 2,594 parameters. This compactness is strategic; it ensures that the model learns representative features from the tabular data without falling into memorization, which is particularly crucial given our limited sample size. The structure of this baseline model is illustrated in Figure 6.1. Single layer and 3-layer architectures performed worse. Layer dimensions were optimized to use with dropout.

Table 6.2 Comparative Performance Metrics: Raw Hematological Indices vs. Engineered Hemorheological Vector

Blood Parameter	*AUC*	*Accuracy*	*Sensitivity*	*Specificity*
RBC	0.514 ± 0.030	0.629 ± 0.033	0.518 ± 0.077	0.685 ± 0.083
HGB	0.566 ± 0.022	0.596 ± 0.024	0.688 ± 0.082	0.551 ± 0.072
HCT	0.559 ± 0.032	0.587 ± 0.035	0.747 ± 0.102	0.507 ± 0.097
MCV	0.455 ± 0.032	0.542 ± 0.045	0.660 ± 0.146	0.483 ± 0.135
MCH	0.423 ± 0.028	0.617 ± 0.034	0.357 ± 0.113	0.747 ± 0.105
MCHC	0.477 ± 0.036	0.603 ± 0.050	0.505 ± 0.122	0.652 ± 0.134
RDW	0.432 ± 0.026	0.549 ± 0.058	0.579 ± 0.167	0.533 ± 0.171
PLT	0.482 ± 0.027	0.582 ± 0.032	0.584 ± 0.111	0.581 ± 0.098
MPV	0.557 ± 0.035	0.642 ± 0.032	0.554 ± 0.080	0.686 ± 0.075
PCT	0.490 ± 0.040	0.496 ± 0.036	0.870 ± 0.081	0.308 ± 0.093
PDW	0.446 ± 0.041	0.573 ± 0.044	0.545 ± 0.129	0.587 ± 0.126
Age	0.563 ± 0.033	0.629 ± 0.037	0.576 ± 0.095	0.656 ± 0.101
Gender	0.581 ± 0.001	0.545 ± 0.026	0.705 ± 0.067	0.464 ± 0.072
All raw parameters	0.475 ± 0.027	0.594 ± 0.036	0.576 ± 0.111	0.603 ± 0.103
EFI	0.607 ± 0.022	0.622 ± 0.021	0.780 ± 0.056	0.542 ± 0.050
EPVB	0.528 ± 0.023	0.549 ± 0.031	0.761 ± 0.082	0.442 ± 0.084
CFHI	0.520 ± 0.030	0.625 ± 0.047	0.537 ± 0.124	0.669 ± 0.131
Engineered Hemorheological Vector (EFI, MPV, CFHI, EPVB, Gender)	0.602 ± 0.028	0.656 ± 0.029	0.611 ± 0.078	0.678 ± 0.069

Abbreviations: AUC, area under curve; CFHI, cell-free hemoglobin index; EFI, erythrocyte fragility index; EPVB, erythrocyte platelet volume balance; HCT, hematocrit, HGB, hemoglobin; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration, MPV, mean platelet volume; PCT, plateletcrit; PDW, platelet distribution width; RBC, red blood cells; RDW, red cell distribution width

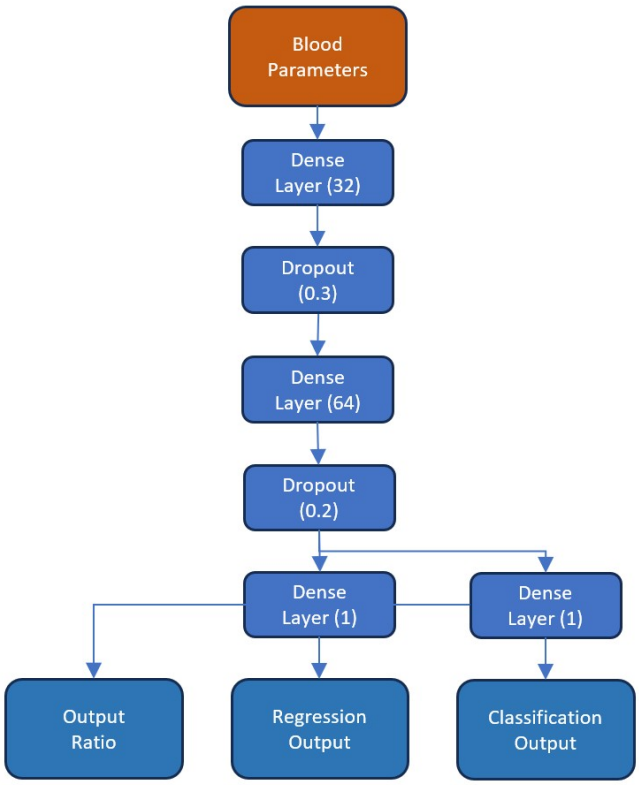

Figure 6.1. Tabular-Only model structure

6.3.2 Video -Only model

The Video-only architecture is designed to capture the complex, time-varying hemodynamic patterns inherent in coronary angiography, which static images cannot convey. The model utilizes two stages of 3D-CNN layers to process the video input, where the 3D convolution kernels effectively downsample the data across both temporal and spatial dimensions.

To refine the feature extraction process, these convolutional stages are followed by pooling and dropout layers, which mitigate overfitting and reduce computational redundancy. A global average pooling layer then collapses the multidimensional feature maps into a consolidated vector, which is fed into a final dense-dropout combination for classification and regression. Given the depth of the 3D-CNN, the model comprises 271,010 parameters.

Single layer and 3-layer models performed worse. 3D-CNN kernel sizes were optimized and (2,8,8) kernel size was selected. Smaller kernel sizes exhausted memory quickly and did not extract meaningful features. Larger kernel sizes could not separate features. 2 frame temporal windows were selected because we were looking for frame differences. No temporal window and more than 2 frame temporal windows performed worse.

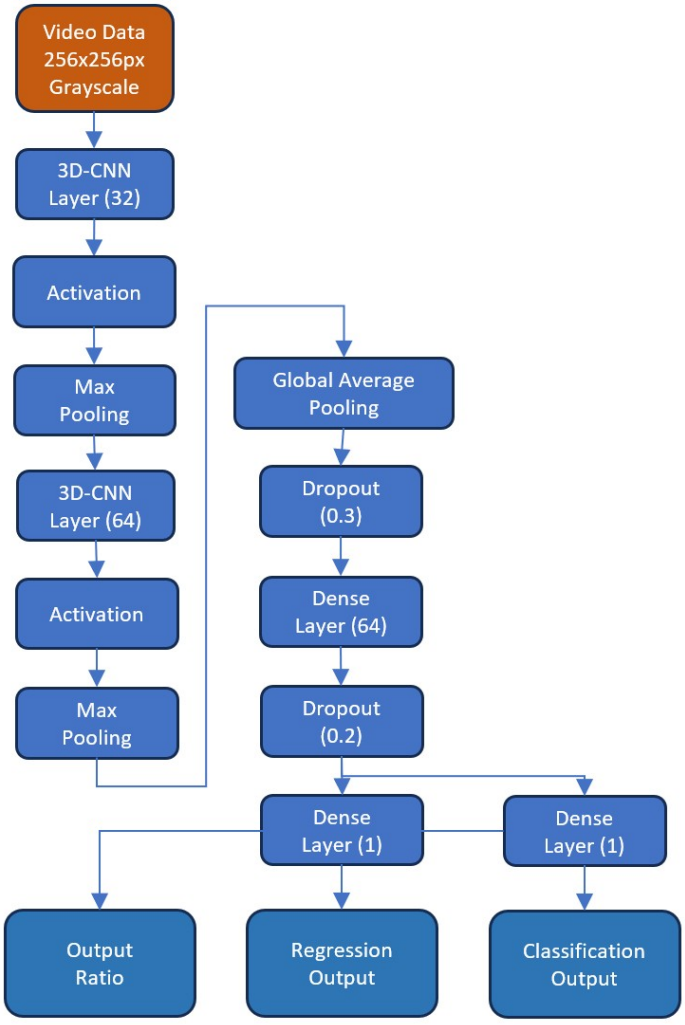

Figure 6.2. Video-only model structure

Crucially, we integrated a Gradient-weighted class activation mapping (GradCAM) layer immediately after the first 3D-CNN layer to extract saliency maps. This design choice enables us to visualize the spatiotemporal regions that the model identifies as most diagnostic, providing a critical interpretability check for the decision-making process. The architectural flow is detailed in Figure 6.2.

6.3.3 Multimodal Late Fusion Model

The Late Fusion architecture serves as a baseline for multimodal integration, representing a decision-level fusion strategy. In this model, the tabular and video data modalities are processed through independent, specialized pipelines before their feature representations are merged. The video branch maintains the same 3D-CNN architecture described in the Video-only model, while the tabular branch processes the blood parameters.

Upon reaching the fusion junction, the 64-dimensional feature vector extracted from the video stream is concatenated with the 128-dimensional feature vector derived from the blood parameters. This merged feature set is then fed into a final dense-dropout layer combination for joint classification and regression. This strategy is foundational for our ablation study, as it provides a benchmark to assess whether simply aggregating features is sufficient to capture the intricate synergy between blood rheology and vascular geometry. The architectural configuration for this model is illustrated in Figure 6.3. Layer dimensions and kernel sizes were tested for different values and best performing values were used. Kernel sizes were (2,8,8).

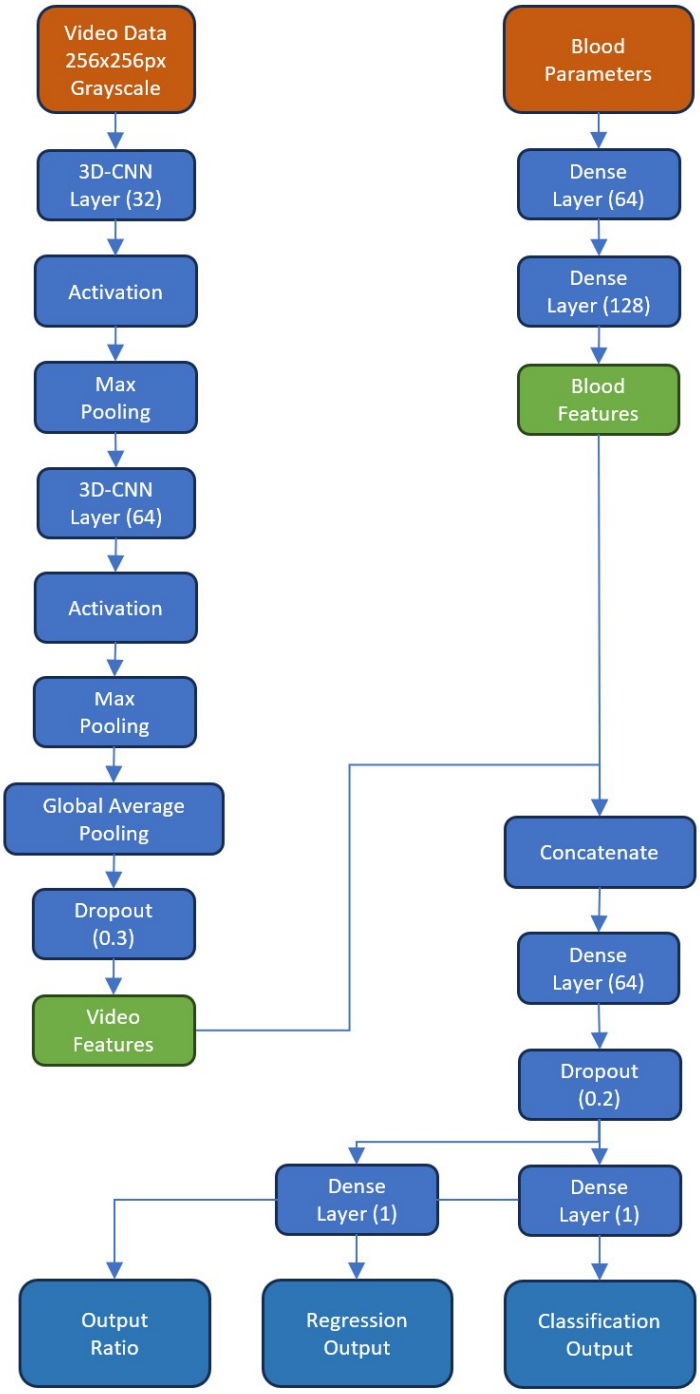

Figure 6.3. Late Fusion model structure

6.3.4 Multimodal FiLM Model

The Multimodal FiLM architecture represents our proposed solution for deep-level feature integration. Unlike the Late Fusion model, which aggregates information only at the final decision stage, the FiLM approach enables an interactive integration of hemorheological data throughout the feature extraction process of the 3D-CNN. Drawing from the principles of conditional computation [49], we utilize FiLM layers to modulate the video feature representations based on patient-specific blood parameters. The feature-wise transformation is defined by the following operation:

$$FiLM(F_{i,c}|\gamma_{i,c}, \beta_{i,c}) = \gamma_{i,c}F_{i,c} + \beta_{i,c}$$

Where $\gamma_{i,c}$ and $\beta_{i,c}$ represents the scale and shift parameters dynamically generated by the tabular branch, and $F_{i,c}$ represents the intermediate video features extracted by the 3D-CNN. This transformation allows the hemorheological profile to act as a gating mechanism. By applying FiLM before the ReLU activation, the tabular branch gains the ability to selectively suppress entire channels within the CNN—essentially instructing the network to ignore or prioritize specific spatiotemporal features based on the patient's clinical profile. A FiLM layer for a CNN is illustrated in Figure 6.4. T is the number of frames and W and H are frame width and height. C is number of features. A γ and β coefficient is kept for each feature coming from the CNN layer.

Our architecture implements this through two hierarchical conditioning stages; initial stage focuses on spatial vessel boundaries and catheter localization, and second stage targets abstract flow patterns and hemodynamic indicators. A FiLM block is used for each stage.

The integration process involves applying a 3D residual block architecture where blood parameters are mapped to γ and β coefficients via dense layers. Additionally, we maintain a residual path by directly concatenating the blood vector at the final fusion layer. This dual-path design—gated modulation through FiLM and direct concatenation—ensures that the model benefits from both the high-level anatomical context provided by the video and the direct physiological insights provided by the

blood indices. This comprehensive integration strategy is detailed in Figure 6.5. Layer dimensions and kernel sizes were tested for different values and best performing values were used. Kernel sizes were (2,8,8).

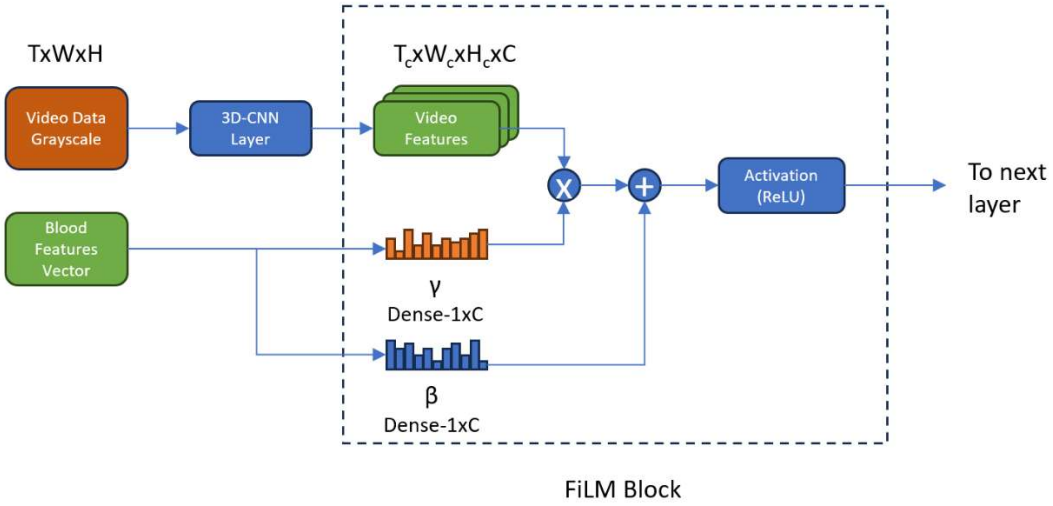

Figure 6.4. A FiLM Layer for a CNN

6.3.5 Output Heads and Ratio Loss

All four models share three output heads. These are given in Table 6.3 .

Table 6.3 Output Heads

Output	*Activation*	*Target*	*Loss*
Regression	Sigmoid	FFR Value	MSE
Classification	Sigmoid	Binary Label	Label-smoothed BCE
Ratio	None	Binary Label	Label-smoothed BCE

Abbreviations: BCE, binary cross entropy; FFR, fractional flow reserve; MSE, mean squared error

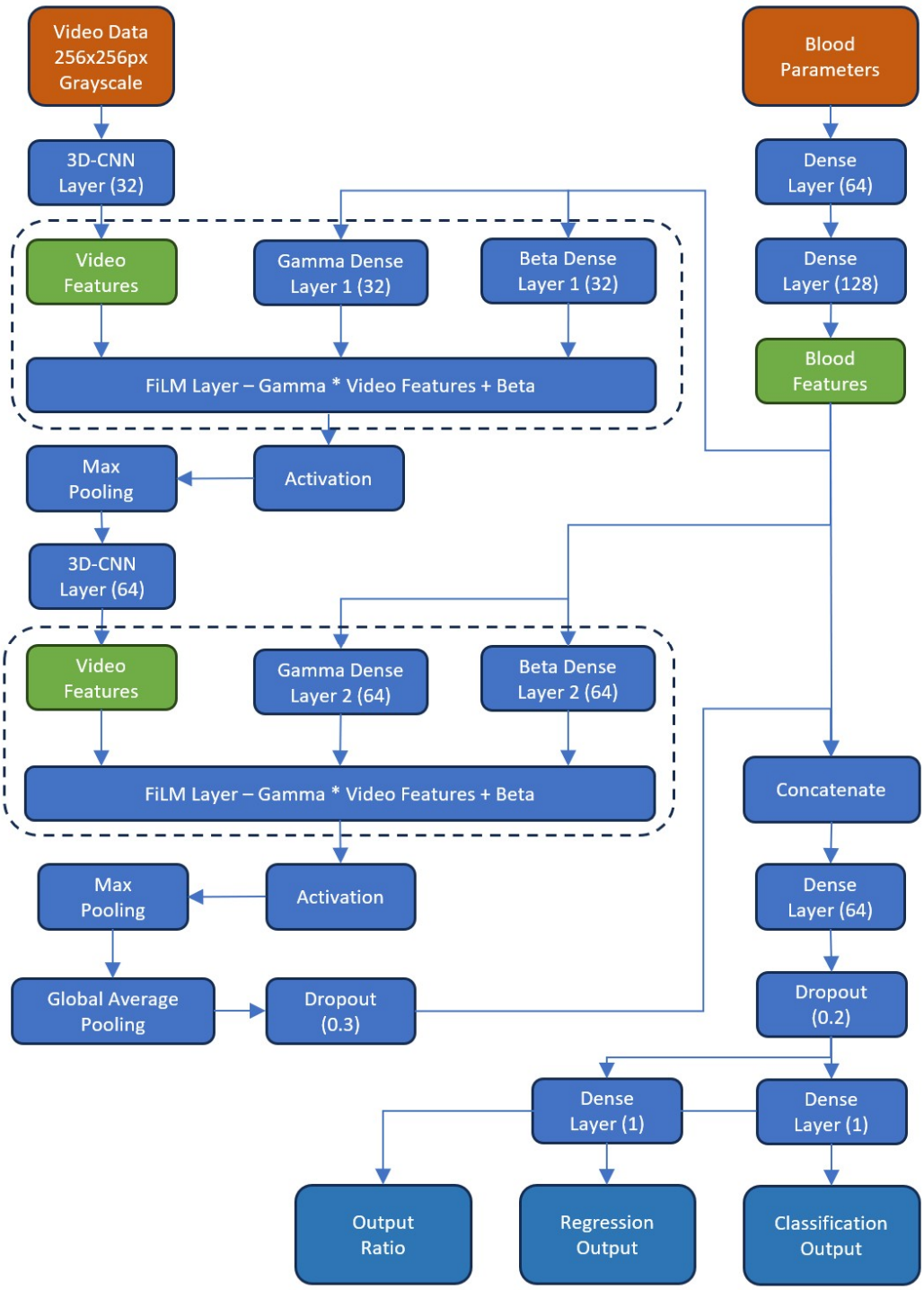

Figure 6.5. Proposed multimodal FiLM model. Two FiLM layers are stacked after 3D-CNN layers.

The ratio output is defined as:

$$r = \frac{p_{cls}}{p_{cls} + p_{reg} + \varepsilon}$$

where p_{cls} is the classification probability and p_{reg} is the regression output. This score is high for positive patients (high p_{cls}, low p_{reg}) and low for negative patients. Training it against the binary label creates cross-gradients that simultaneously push p_{cls} and p_{reg} in complementary directions, encoding the prior knowledge that the two heads should behave as a coherent pair. Statistical analysis (SPSS) confirmed that the ratio score p_{cls} / p_{reg} is correlated more strongly with true FFR values than either output alone (r = 0.033 and p = 0.615 for prediction – r = -0.249 and $p < 0.01$ for probability – $r = 0.261$ and $p < 0.01$ for ratio), motivating its inclusion as a supervised auxiliary signal. Formula was altered to p_{cls} / (p_{cls} + p_{reg}) to keep the result between 0 and 1. A very small epsilon value was added to avoid division by zero.

The boundary MSE loss adds an extra penalty when the regression prediction crosses the clinical threshold:

$$\mathcal{L}_{bMSE} = MSE(\tilde{y}, \hat{y}) + 0.5 \cdot \mathbb{E}\left[\mathbf{1}[sign(\tilde{y} - \tau) \neq sign(\hat{y} - \tau)] \cdot (\tilde{y} - \hat{y})^2\right]$$

where τ = 0.60 is the normalized threshold. This directly penalizes clinically consequential regression errors. The composite training objective is:

$$\mathcal{L} = 1.0 \cdot \mathcal{L}_{bMSE} + 0.3 \cdot \mathcal{L}_{cls} + 0.2 \cdot \mathcal{L}_{ratio}$$

Common practice for regression output layer activation function is to use linear activation. Linear output lets gradients flow without squashing, and makes no assumptions about the target range. It's the default for unbounded continuous targets. In our model, The FiLM layers generate per-channel scale factors γ from the tabular branch. With no constraint on γ, feature activations can grow very large. A linear regression head then outputs raw dot products of these amplified features, which produced predictions in the 10–20 range rather than [0, 1], our dataset mapping. FFR is physiologically bounded to roughly [0.5, 1.0]. After normalizing and mapping,

every ground-truth label is in [0, 1]. A sigmoid output is structurally matched to this range.

Sigmoid gradients saturate near 0 and 1. If the model were ever very confident and very wrong (prediction near 0, true value near 1), the gradient signal becomes small. In practice this is not a serious problem here because:

- The clinical threshold (0.80 → normalized 0.60) sits in the central, high-gradient region of sigmoid.
- The boundary MSE loss adds extra penalty near the threshold, keeping predictions away from the saturation zone.
- The classification head (separate sigmoid) provides a parallel gradient path.

6.4 Training Methodology

6.4.1 Patient level data splitting

To mitigate stochastic dependency between samples from the same patient, we adopted a patient-level splitting strategy. Standard random splitting would inadvertently introduce data leakage, as multiple videos from a single patient would contaminate both the training and validation sets. We employed the StratifiedGroupKFold strategy, ensuring that all videos belonging to the same patient remain within the same fold, regardless of the target label. Patients with mixed FFR labels across their videos (2 of 179) are handled naturally: all videos of the same patient belong to the same fold regardless of label.

Furthermore, a 20% held-out test set reserving approximately 36 patients (≈51 videos) was sequestered prior to any model development, serving as a 'blind' benchmark that was never utilized for hyperparameter tuning or early stopping criteria. The remaining 80% of patients formed the development set on which 5-fold cross-validation was performed.

The test set ensures that reported metrics reflect genuine generalization. Cross-validation metrics are optimistic because early stopping and model checkpointing select the best epoch based on the validation fold; a completely independent test set breaks this dependency chain. Patient-level grouping prevents blood parameter leakage that would otherwise inflate performance for the tabular and multimodal models.

6.4.2 Class Imbalance Handling

The dataset exhibits a class imbalance ratio of approximately 1:1.94 (positive vs. negative). To prevent model bias toward the majority class, we implemented Balanced Batch Sampling, utilizing an oversampling generator that provides an equal number of positive and negative samples per training epoch.

6.4.3 Training Protocol

The training protocol utilized the Adam optimizer (initial LR: 2×10^{-4}) with an adaptive ReduceLROnPlateau scheduler and early stopping with patience of 20 epochs. Balanced accuracy is chosen as the monitoring metric as it is the mean of sensitivity and specificity; robust to class imbalance; and sensitivity or specificity collapses give BA=0.5. We observed that increasing the training duration to 30 or 50 epochs triggered the scheduler excessively, causing the optimizer to prematurely converge to local minima with limited generalization capacity. A 20-epoch training window, however, maintained a more robust global trajectory.

Furthermore, a batch size of 1 was empirically determined as optimal; larger batch sizes (2, 4, 8, 16) exhibited degraded performance, likely due to a reduction in gradient update frequency and the sensitivity of internal covariate shift correction in our specific 3D-CNN architecture. Gradient clipping is employed to prevent gradient explosion with batch size 1. Video length was fixed to 32 frames, because all videos revealed the entire coronary vasculature before 32 frames and less memory is used.

6.4.4 Reproducibility

In TensorFlow, Keras, and Python, all random processes were made deterministic by defining seed values. At the start of each fold, fold number is added to the seed, ensuring generator shuffles and weight initialization are independent across folds and model types.

6.4.5 Visualization

Gradient-weighted Class Activation Mapping (GradCAM) [79] was applied to the first ReLU activation layer of video processing networks to produce spatiotemporal saliency maps. For each temporal feature position, per-channel weights are computed by global-average-pooling the gradients over the spatial dimensions, then linearly combined with the feature activations and passed through ReLU. The resulting 32 × 32 saliency maps are bilinearly upsampled to the original 256 × 256 resolution. Temporal interpolation uses linear blending between adjacent feature positions to avoid the frame-duplication artefact that arises from nearest-neighbor mapping because the temporal stride is 2.

6.5 Results

The performance of our proposed architectures was evaluated through a systematic ablation study, comparing unimodal baselines against multimodal fusion strategies. All experiments were conducted using a 5-fold cross-validation (CV) strategy across 16 random seeds to account for sampling variance, with final evaluations performed on a fully independent held-out test set ($n = 36$ patients, ≈ 51 videos).

6.5.1 Cross Validation benchmarking

Table 6.6 and Table 6.7 summarize the CV performance using all blood parameters versus our selected hemorheological feature vector. The Multimodal FiLM model consistently achieves the highest CV performance (AUC: 0.672). Moving from all parameters to the selected vector yielded a marked improvement, particularly for the FiLM architecture. This indicates that raw hematological data introduces substantial noise, whereas our engineered features (EFI, CFHI, EPVB, Gender) effectively act as a low-pass filter, isolating the biophysical markers most relevant to hemodynamic variance.

6.5.2 Test Set Generalization

While CV metrics provide an optimistic estimate, the independent test set (Table 6.8 and Table 6.9) offers an honest assessment of the model's generalization capacity. The Multimodal FiLM model consistently outperformed all baselines in test AUC (0.600) and Balanced Accuracy (0.559). FiLM architecture outperforms Late Fusion, confirming that early/intermediate conditioning (gating spatiotemporal features at the convolutional stage) is superior to late-stage concatenation. This suggests that the model benefits from tailoring the feature extraction process to the patient's specific physiological profile. The Video-only model barely exceeds chance (AUC ≈0.47), underscoring the hemodynamic-geometric gap: vascular geometry alone is insufficient to predict FFR at this dataset scale, justifying the necessity of our multimodal fusion approach.

The Video-Only model achieves sensitivity and specificity very close to 1.000 in cross-validation tests because its best epoch corresponds to a collapse where all samples are classified as positive or negative. The simultaneously high metrics arise from different epochs being selected for each metric independently, this is a limitation of reporting best-epoch statistics.

6.5.3 Predictive Precision and Clinical Reliability

To further validate our findings, we assessed the relationship between predicted and ground-truth values.

Our multimodal models demonstrated a statistically significant correlation between predicted and true FFR values. The correlation coefficient r and the statistical significance p values for modalities are given in Table 6.4. FiLM model has the best correlation and p value. A comparison of the scatter plots of modalities is given in Figure 6.8 and Figure 6.9.

Table 6.4 r and p values of Spearman's correlation for modality predictions

Modality	Correlation Coefficient r_s	Statistical Significance *p-value*
Tabular only	0.117	0.067
Video only	-0.002	0.975
Multimodal Late Fusion	0.103	0.106
Multimodal FiLM	0.236	**0.0002**

Abbreviations: FiLM, feature-wise linear modulation; r_s, Spearman's rank correlation coefficient

A Bland-Altman analysis (Figure 6.7) was utilized to assess agreement between the methods and invasive FFR. By plotting the difference between predicted and true values against their mean, we identified a Mean Bias of -0.017 for FiLM model and -0.029 for late fusion model. This minimal bias confirms that our models do not exhibit systematic over- or under-estimation across the diagnostic range.

Compared to using either output individually with statistical analyses, the ratio loss $p_{cls}/(p_{cls} + p_{reg})$ demonstrated a stronger correlation with true FFR values, an effect especially pronounced in the tabular and multimodal models. This loss function generates cross-gradients that force the two heads to become complementary rather than redundant. With FiLM model, the correlation is not strongly seen. Using only ratio output, sensitivity was zero. In contrast, removing

ratio output from outputs diminished result. Single output head test set results for FiLM model for single seed are given in Table 6.5.

Table 6.5 Single output head test set results for FiLM model for single seed

Output Head	*Regression*	*Classification*	*Ratio*	*All 0.3/1/0.2*	*All 1/0.3/0.2*
AUC	0.629	0.675	0.671	0.687	**0.714**
MAE	**0.136**	0.256	0.399	0.163	0.147
Sensitivity	0.514	0.400	0.000	0.386	**0.614**
Specificity	0.622	0.795	**0.978**	0.768	0.676
Balanced Accuracy	0.568	0.597	0.489	0.577	**0.645**

Abbreviations: AUC, area under curve; MAE, mean absolute error

Table 6.8 and Table 6.9 report results on the independent test set (36 patients, ≈51 videos) that was never used during training or model selection. These are honest single-pass evaluations of the saved checkpoint. The Multimodal FiLM model achieves the highest test AUC and balanced accuracy across all tests. Notably, the multimodal FiLM model outperforms the tabular-only and video-only baselines confirming that the video branch and its FiLM conditioning provide independent predictive information.

Minimum and maximum values of results for cross-validation and test sets are given in the Appendix.

The Video-Only model barely exceeds chance, and its cross-validation performance is inflated by a sensitivity collapse. This result, consistent across all epoch settings and dataset sizes tested, suggests that the spatiotemporal patterns in a single-vessel angiogram are insufficient for reliable FFR estimation at this dataset scale with this architecture. The 3D-CNN may require substantially more data or a larger architecture to learn the subtle flow dynamics that distinguish hemodynamically significant from non-significant stenoses.

The Tabular-Only model achieves a competitive test AUC, confirming that hematological parameters are genuinely informative for FFR estimation. This is physiologically plausible.

A key experimental question was whether the benefit of multimodal integration depends on the location of fusion — early/intermediate (FiLM) vs. late (concatenation). The Late Fusion model performs better than baseline models, but underperforms FiLM despite having access to the same information. This gap is consistent with the theoretical motivation: FiLM allows blood parameters to gate which spatiotemporal features the CNN attends to at each convolutional stage, effectively telling the network to focus on these vessel characteristics using patient's hematological profile. Late fusion can only weight the final representations, not reshape the feature extraction process itself. AUC curve of top AUC value is given in Figure 6.6. Clearly, multimodal FiLM model surpasses late fusion model.

A systematic gap exists between CV and test metrics. This reflects the difficulty of selecting the best validation fold under the StratifiedGroupKFold constraint: with ≈143 development patients divided into 5 folds, each validation set contains ≈29 patients, and a single misclassification shifts fold AUC by approximately 0.035. Because of this, multiple (16) seed runs were performed to minimize this sampling variance.

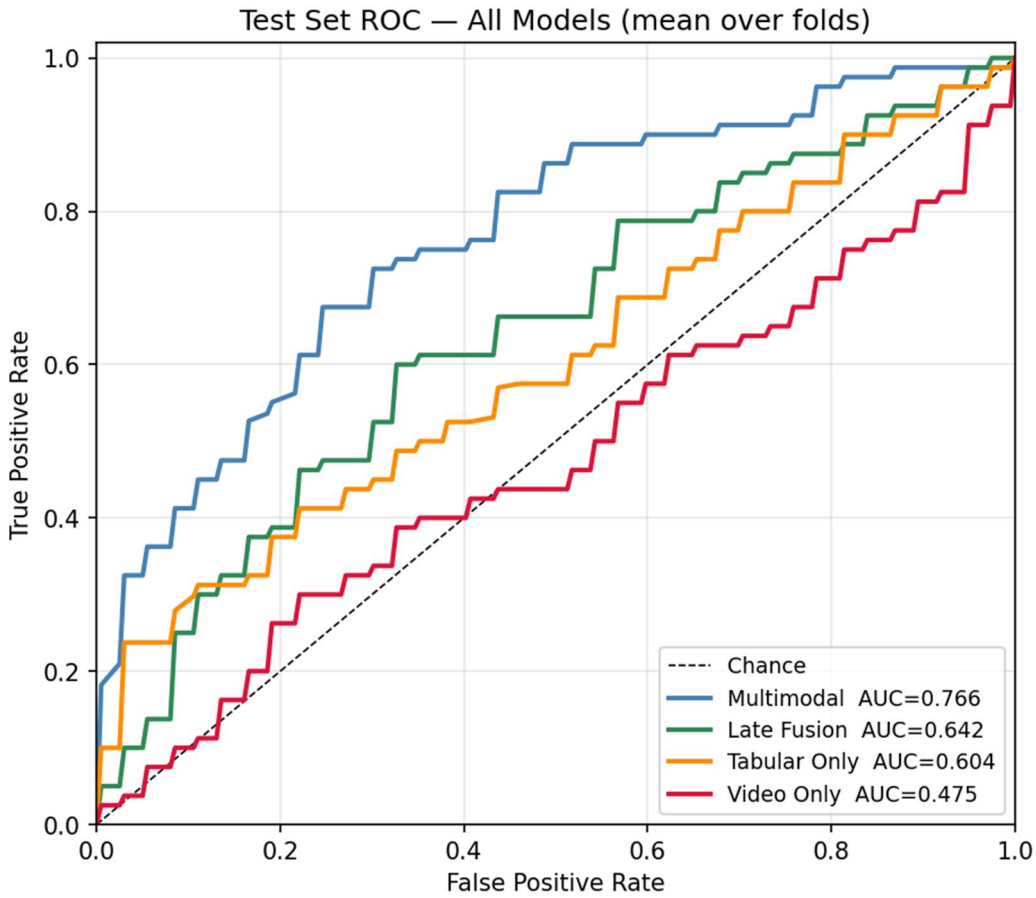

Figure 6.6. Test Set ROC with selected blood parameters for best seed. Abbreviations: AUC, area under curve, ROC, receiver operating characteristic.

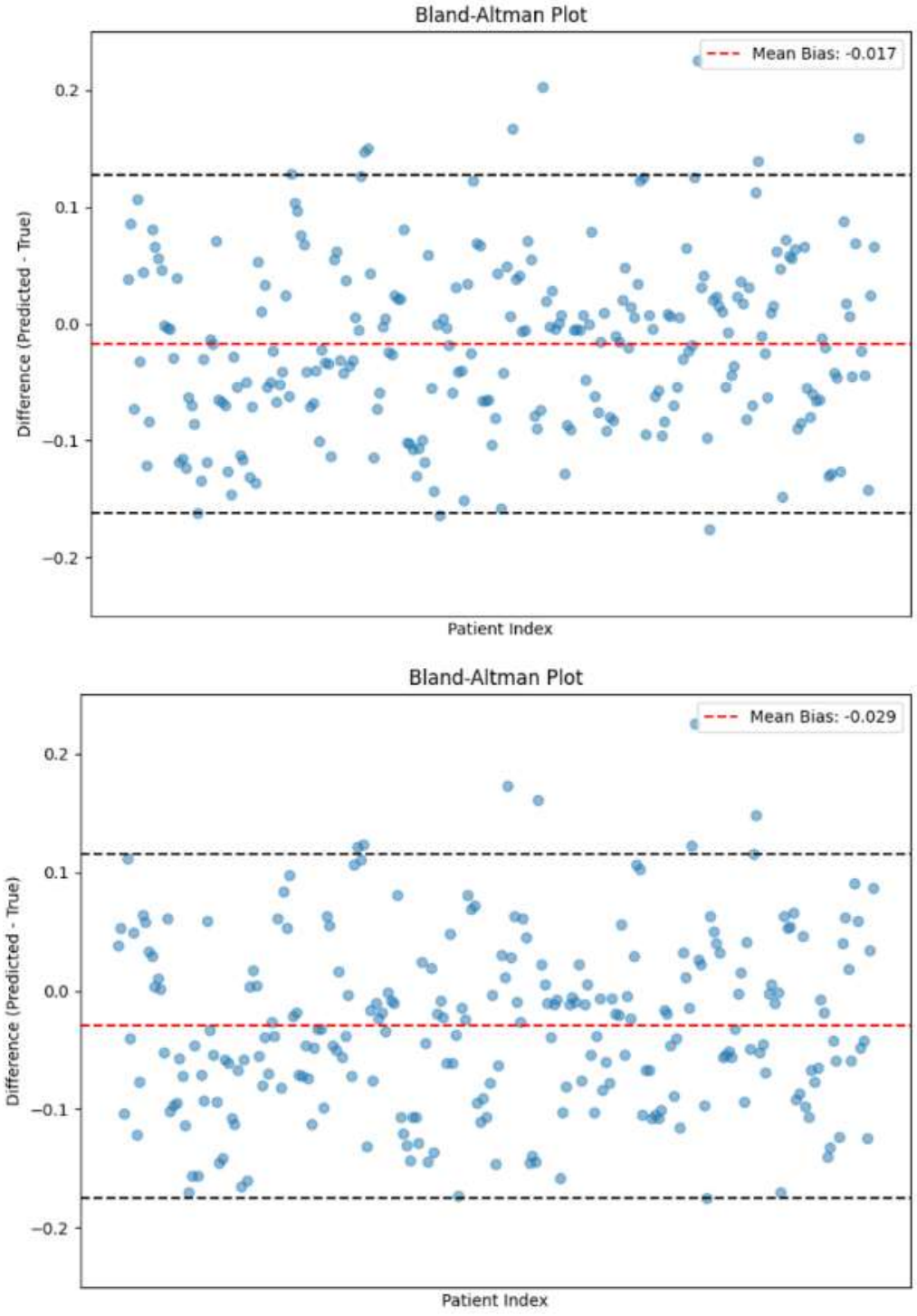

Figure 6.7. Bland-Altman plot for multimodal FiLM model (top) and multimodal late fusion model (bottom) for all seeds

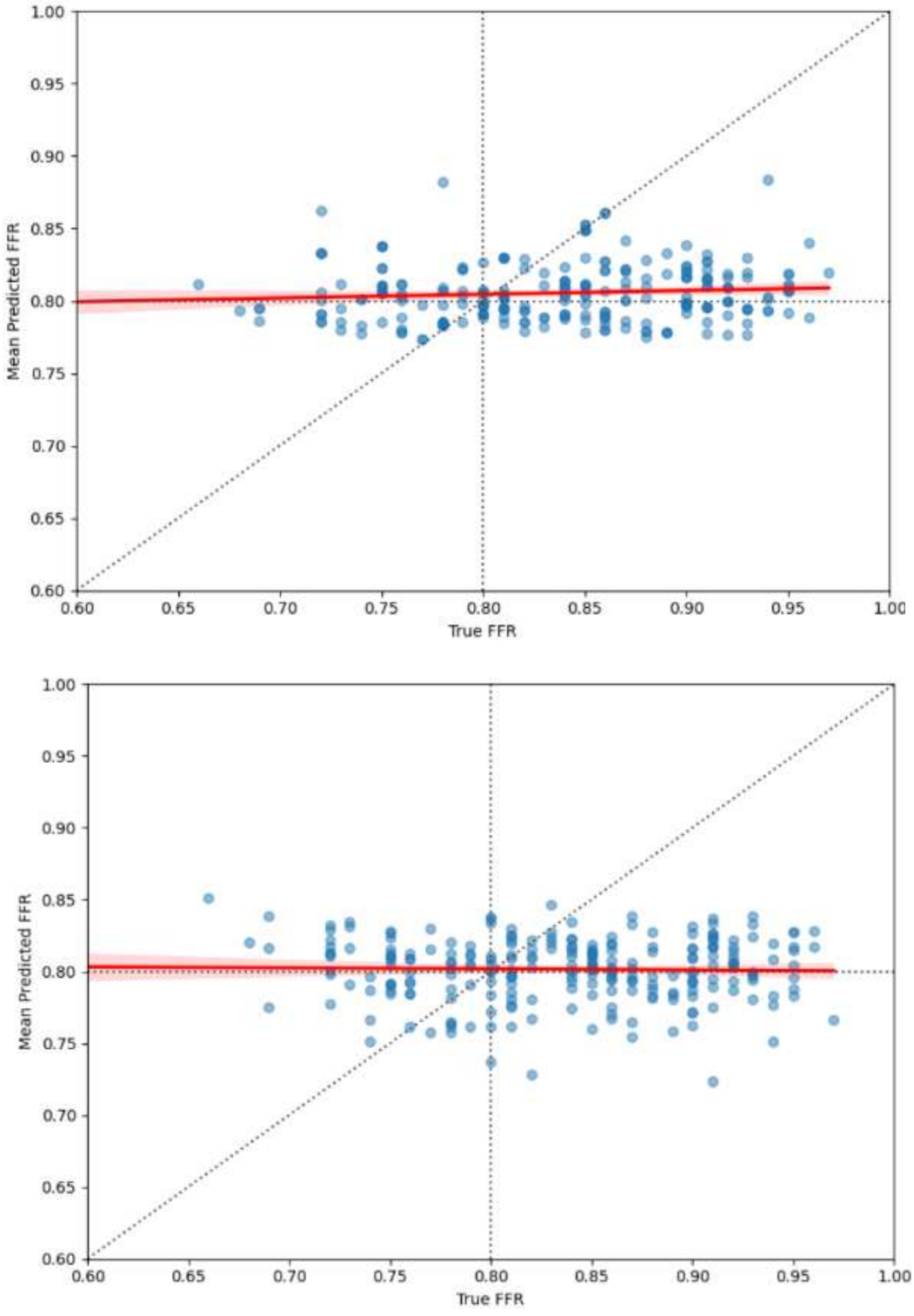

Figure 6.8. Scatter plots of: tabular-only model (top), video-only model (bottom)

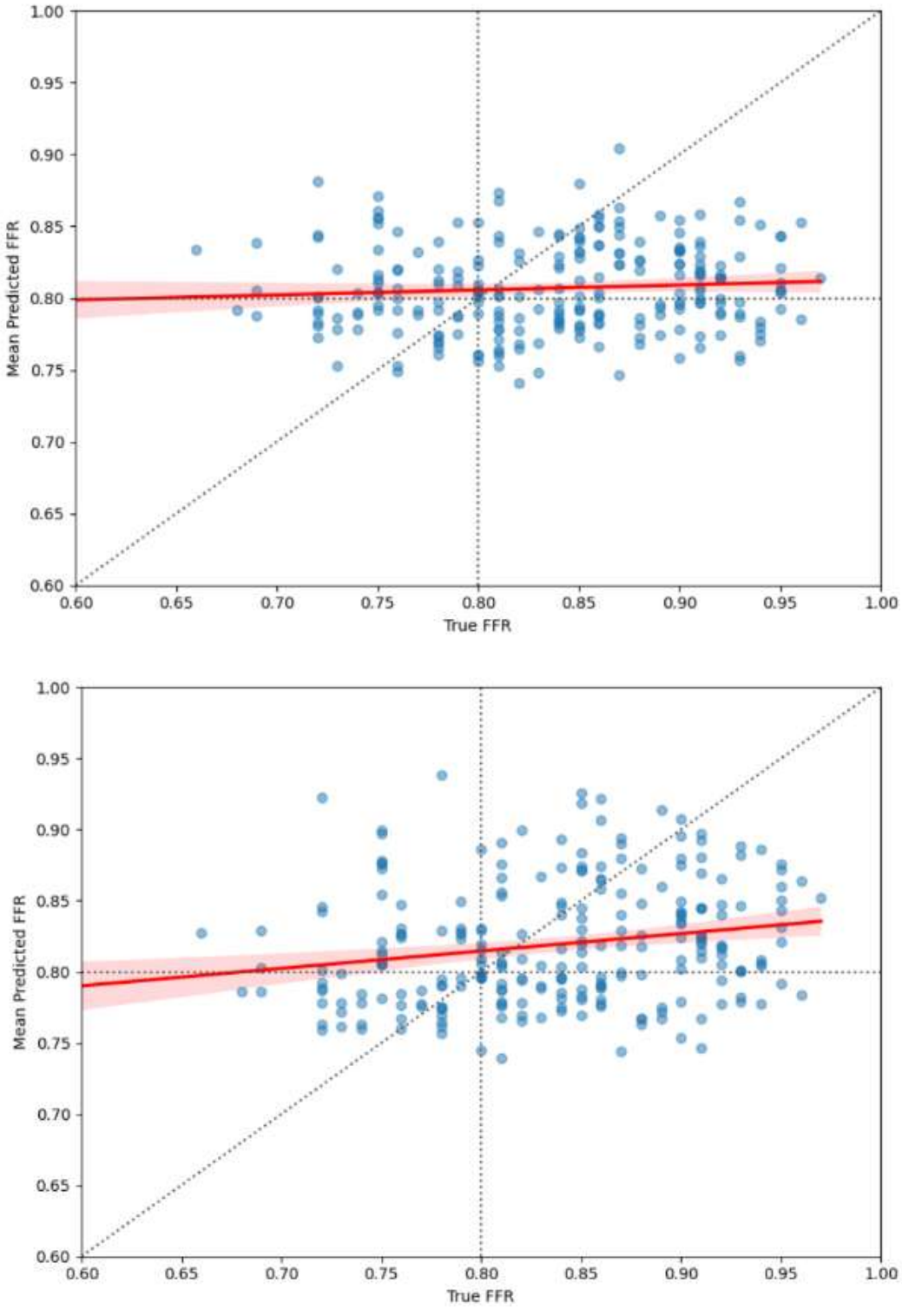

Figure 6.9. Scatter plots of: multimodal late fusion model (top) and multimodal FiLM model (bottom) for all seeds

Table 6.6 Comparison of Cross-Validation Modality Performance with all blood parameters

Model	*Tabular only*	*Multimodal FiLM*	*Multimodal Late Fusion*	*Video only*
AUC	0,587±0,047	**0,628±0,034**	0,592±0,045	0,600±0,041
MAE	0,120±0,006	0,123±0,005	0,118±0,006	**0,117±0,005**
Sensitivity	0,711±0,087	0,665±0,060	0,886±0,038	**0,981±0,032**
Specificity	0,670±0,062	0,818±0,025	0,947±0,037	**0,978±0,041**
Balanced Accuracy	0,597±0,037	**0,620±0,027**	0,601±0,026	0,583±0,027

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

Table 6.7 Comparison of Cross-Validation Modality Performance with selected blood parameters

Model	*Tabular only*	*Multimodal FiLM*	*Multimodal Late Fusion*	*Video only*
AUC	0,620±0,025	**0,674±0,024**	0,638±0,031	0,600±0,041
MAE	0,118±0,004	0,120±0,007	**0,115±0,006**	0,117±0,005
Sensitivity	0,761±0,084	0,747±0,057	0,945±0,045	**0,981±0,032**
Specificity	0,640±0,071	0,793±0,040	0,951±0,026	**0,978±0,041**
Balanced Accuracy	0,602±0,023	**0,646±0,027**	0,626±0,022	0,583±0,027

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

Table 6.8 Comparison of Test Set Modality Performance with all blood parameters

Model	*Tabular only*	*Multimodal FiLM*	*Multimodal Late Fusion*	*Video only*
AUC	0,524±0,067	**0,567±0,077**	0,542±0,068	0,480±0,045
MAE	**0,137±0,017**	0,152±0,015	0,162±0,021	0,170±0,024
Sensitivity	**0,537±0,113**	0,429±0,117	0,449±0,124	0,451±0,170
Specificity	0,507±0,081	**0,657±0,099**	0,585±0,115	0,519±0,158
Balanced Accuracy	0,522±0,062	**0,543±0,058**	0,517±0,045	0,485±0,034

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

Table 6.9 Comparison of Test Set Modality Performance with selected blood parameters

Model	*Tabular only*	*Multimodal FiLM*	*Multimodal Late Fusion*	*Video only*
AUC	0,561±0,066	**0,606±0,079**	0,547±0,065	0,480±0,045
MAE	**0,135±0,015**	0,154±0,016	0,158±0,021	0,170±0,024
Sensitivity	**0,589±0,087**	0,547±0,125	0,512±0,116	0,451±0,170
Specificity	0,484±0,092	**0,590±0,109**	0,533±0,147	0,520±0,158
Balanced Accuracy	0,536±0,049	**0,569±0,061**	0,522±0,041	0,485±0,034

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

6.5.4 GradCAM Visualization

Figure 6.10 and Figure 6.11 shows the GradCAM saliency map frame overlays of a positive and a negative class angiography video for the best multimodal fold of a specific seed. Top rows are the original video frame, continued by video only, late fusion and multimodal FiLM model frame overlays. Frames are ordered by frame number, selected by the greatest s value, the mean saliency value computed across all computed models. For Figure 6.10, for this specific seed, test set AUC values are 0.668 for tabular only, 0.597 for multimodal FiLM, 0.544 for late fusion and 0.420 for video only. All three video models can clearly identify vessels, but FiLM model gives the best overall result in AUC and balanced accuracy. For Figure 6.11, for this specific seed, test set AUC values are 0.576 for tabular only, 0.586 for multimodal FiLM, 0.453 for late fusion and 0.434 for video only. Video only and late fusion models miss most of the vessel structure. Multimodal FiLM model identifies most of the vessel structure, and outperforms other models.

6.6 Discussion

6.6.1 Video Features Alone Are Insufficient

The pressure drop across coronary stenosis depends on lesion geometry, vascular conditions, and fluid resistance. While invasive ICA visualizes anatomy, it lacks data on blood properties that determine ischemic risk. This explains why our Video-Only model, based solely on vessel images, performed poorly (AUC 0.480), showing that morphology alone can't reliably estimate fractional flow reserve (FFR). Previous work by [64] achieved 82% accuracy by combining geometric and clinical data, highlighting that raw images don't capture functional signals without additional physiological context. Similarly, [65] used ResNet18 on angiogram patches and got an AUC of 0.81, despite human accuracy being 58%. This shows CNNs can detect subtle geometric patterns but treat vessels as rigid, ignoring systemic physiology.

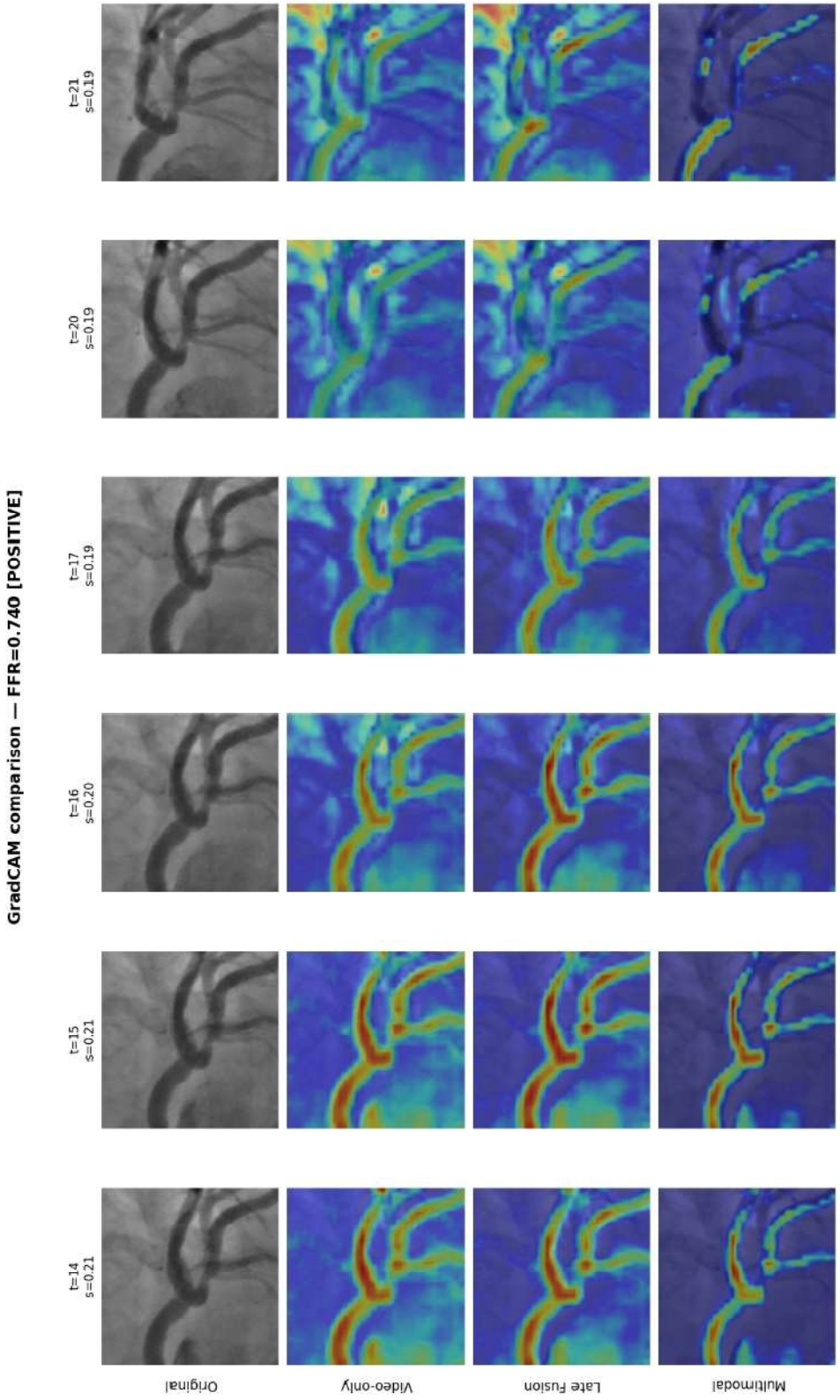

Figure 6.10. GradCAM comparison of a positive class video

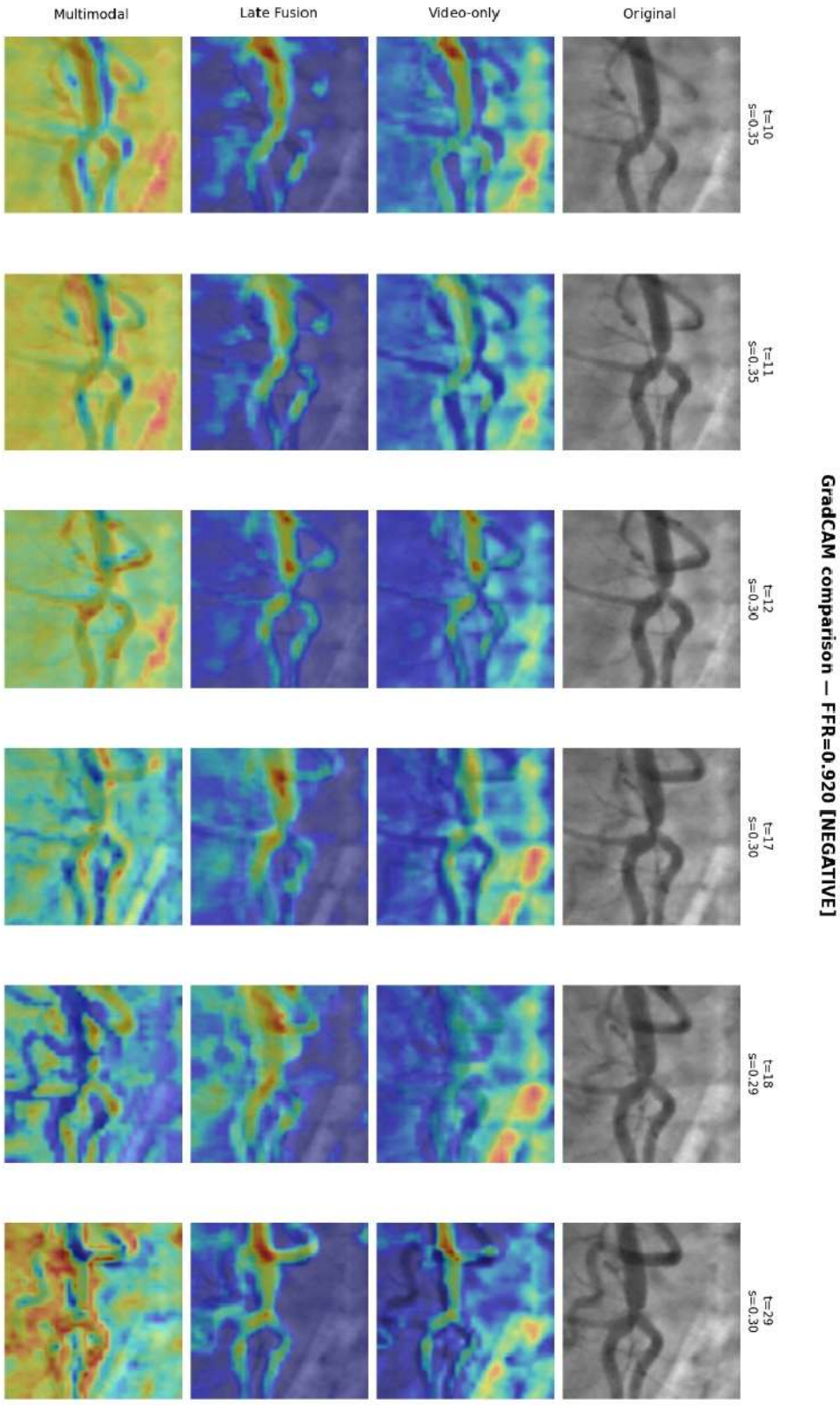

Figure 6.11. GradCAM comparison of a negative class video

DL model of [59] also plateaued at about 80% accuracy, indicating single signals can't solve the complex variables of hyperemic flow. Our Video-Only model's failure stems from these biomechanical constraints, as [76] explained that models assuming rigid vessels neglect critical factors like plaque composition and microcirculation. Without rheological or cellular data, a 3D CNN cannot infer fluid dynamics or microvascular resistance, leading to failure in identifying significant stenoses. Although [76] suggested a sample size over 150 patients is enough for validation, our dataset has 179 patients and 256 angiograms. The real issue is physiological, so we enhanced the vision model with blood property data, providing the biophysical context missing from raw frames. Multimodal approaches are crucial for accurate, personalized ischemia predictions.

6.6.2 Blood Parameters Carry Independent FFR Signal

The Tabular-Only model had an AUC of 0.561, indicating that peripheral hematological parameters independently signal FFR estimation. Our framework sees coronary blood flow as a complex, non-Newtonian cellular suspension influenced by systemic conditions, comorbidities, and genetics, reflected in changes in erythrocytes and platelets. Because these cells last 120 and 7-10 days, respectively, they serve as chronic, personalized sensors of vascular stress and endothelium wear. They form a feedback loop where shear forces remodel cells and influence blood flow.

It is a critical question whether blood parameters, which may vary by time and conditions, can accurately predict localized coronary indices like FFR. We view the peripheral blood profile as a stable hemorheological vector that represents baseline ischemic potential, rather than a transient variable. CBC metrics, once viewed as broad benchmarks, are now understood, based on a study by Brody H. Foy et al. [80], to cluster around individual, stable hematological setpoints—distinguishing individuals from 98% of the population and remaining stable for 20 years in healthy adults. This indicates that a CBC sample taken before ICA reflects a conserved

physiological blueprint rather than a snapshot. The blood profile also captures interconnected biological signals that outperform traditional demographic metrics. Consequently, chronological age was excluded from our model, as biological aging and inflammation are thought to be embedded in these vectors, validated by [81], who showed that CBC indices can predict health and biological age, enabling DL models to replace chronological age with high-dimensional indices.

Patient-specific hemorheological phenotypes influence trans-stenotic pressure drops via non-Newtonian fluid mechanics. [82] showed blood's rheological traits—affected by erythrocyte deformability, aggregation, and structure—determine local fluid resistance in various vessels. Increased flow resistance raises shear stress, damaging endothelium and lowering pressure in arterial narrowing. This study introduces EFI, a metric combining RDW and MCHC that captures how size heterogeneity and cellular rigidity disrupt blood flow in the coronary microenvironment. High RDW indicates cellular vulnerability, causing chaotic flow and increased collisions, viscosity, and endothelial strikes. Elevated MCHC signifies denser hemoglobin, impairing deformability. Rigid cells under shear force cause friction and endothelial injury. [26] showed altered erythrocytes adhere more to endothelium, promoting vasopathology. [83] linked higher RDW with increased coronary calcium, indicating erythrocyte size variability correlates with macrovascular atherosclerosis.

On the other hand, while literature links higher MPV with adverse cardiovascular outcomes [84], our study found that higher baseline MPV was associated with preserved FFR. Since our population's MPV values (8.2 fL for men, 8.5 fL for women) are within normal limits, this positive indicator reflects preserved platelet structural integrity, vital for endothelial function. Platelets maintain vascular barrier function by tightening endothelial belts without causing inflammation [85]. Under hemorheological stress or hemolysis, increased adhesion molecules promote harmful interactions between erythrocytes and endothelium [29]. Strong, unfragmented platelets protect endothelial cells from erythrocyte damage, maintaining microcirculatory stability. Fragmentation weakens this shield, increases damage, and

allows small platelets to infiltrate inter-erythrocyte gaps further disrupting the flow. Thus, physiologic MPV indicates preserved platelet integrity, crucial for vascular barrier and flow in microvascular stresses.

6.6.3 FiLM Outperforms Late Fusion

A key experimental question was whether the benefit of multimodal integration depends on the location of fusion — early/intermediate (FiLM) vs. late (concatenation). The Late Fusion model (mean test AUC = 0.547) substantially underperforms FiLM (mean test AUC = 0.606) despite having access to the same information. This gap is consistent with the theoretical motivation: FiLM allows blood parameters to gate which spatiotemporal features the CNN attends to at each convolutional stage, effectively telling the network to focus on the vessel characteristics given the patient's hematological profile. Late fusion can only weight the final representations, not reshape the feature extraction process itself.

6.6.4 Cross-Validation vs. Test-Set Metrics

A systematic gap exists between CV and test metrics (CV AUC = 0.674 vs. test AUC = 0.606 for multimodal FiLM). This result reflects the difficulty of selecting the best validation fold under the class-wise stratified K-Fold iterator variant with non-overlapping groups (StratifiedGroupKFold) constraint: with ≈143 development patients divided into 5 folds, each validation set contains ≈29 patients, and a single misclassification shifts fold AUC by approximately 0.035. Multiple seed runs and confidence interval estimation is done to help quantify this sampling variance.

6.6.5 Composite Output Layer

Our model bridges the gap between continuous regression (FFR value estimation) and binary classification (ischemia detection). By implementing a composite

objective that calibrates regression outputs with classification probabilities using the ratio output, we enforced a clinical prior: that the two heads should function as a coherent pair. This internal consistency is vital for clinical reliability, as it minimizes the risk of diagnostic ambiguity near the 0.80 FFR threshold. The heatmap saliency maps generated by our model further reinforce this, demonstrating that the model's attention is physiologically focused on relevant vascular segments, rather than arbitrary background artifacts.

6.6.6 Comparison with similar work

Cho et al. [64] used angiography key frames instead of videos and measured vessel diameters and extracted 24 features of the stenosis. They added patient age, gender, body surface area, and stenosis vessel to these features and used XGBoost for FFR greater than or equal to 0.80 and less than 0.80. They used XGBoost twice, once to select the best performing parameters, and once for output classification. With this much preprocessing, the best accuracy was 82%, the best sensitivity was 84%, and the best specificity was 80%. Since they only checked for the FFR cut-off value, they did not provide an MAE value.

Studies [67] and [68] process single view angiography videos to predict FFR. Despite our faster and lighter model, we have a limited dataset size. Langlais et al. [67] had a massive dataset over 50.000 videos organized in 4 datasets. They used calculated QFR and expert opinion for ground truth. Best MAE result was 0.077 for 568 LMCA videos at stenosis percentage prediction. Moreover, inference with our model takes only 16 milliseconds as opposed to tens of seconds. Mineo et al. [68] used 341 single view angiography videos paired with FFR values training for both regression and classification. Best MAE for FFR regression was 0.063 and best AUC was 0.95.

Comparison of FFR predicting studies is given in Table 6.10. These studies had superior datasets compared to ours. Commercial products often use multiple view

videos to train and predict. Yet, there is no study that combines blood parameters with angiography videos, and a direct comparison is not possible. With limited dataset size and with the use of blood parameters, our model established a supporting link between FFR values, angiography videos and blood parameters.

Our main argument is that blood parameters enhance video processing. Other studies may have better results than ours, but our model is fundamental and our dataset is the smallest in size. Despite this, we were able to produce meaningful results.

6.7 Limitations and Future Work

With 179 patients, the per-fold validation sets are small (≈29 patients), leading to high variance in fold-level metrics. The reported results should be validated on a larger, multi-center cohort. Each video corresponds to a single coronary vessel measurement; multi-vessel disease patterns are not captured. Integration direct measurements of blood rheology (e.g., Viscosity, Fibrinogen, Von Willebrand Factor, or direct RBC deformability metrics) may improve results.

Table 6.10 Comparison with FFR predicting studies

Study	*Video*	*Data*	*Ground Truth*	*Size*	*Performance*
Ben-Assa et al. [50]	Yes	No	Measured FFR value	Commercial: >10.000 labeled frames	AUC: 0.929
Fearon et al. [55]	Yes	No	Measured FFR value	Commercial: At least two view videos	AUC: 0.944
Roguin et al. [57]	Yes	No	Measured FFR value	Commercial: At least two view videos	AUC: 0.91
Zhao et al. [66]	No	Stenosis features extracted	Calculated CFR	5913 videos	Best MAE: 0.12
Cho et al. [64]	No	Stenosis and clinical features	Measured FFR classification	1501 stenosis image	Best AUC: 0.87
Langlais et al. [67]	Yes	No	Calculated QFR	568 single view videos	Best MAE: 0.077
Mineo et al. [68]	Yes	No	Measured FFR value	341 single view videos	Best MAE: **0.063** Best AUC: **0.95**
Proposed method	Yes	Blood parameters	Measured FFR value	256 single view videos	Best MAE: 0.132 Best AUC: 0.767

Abbreviations: AUC, area under curve, CFR, coronary flow reserve; FFR, fractional flow reserve; MAE, mean absolute error; QFR, quantitative flow reserve

CHAPTER 7

CONCLUSION

This thesis has presented the development and clinical validation of a novel, multimodal DL framework designed for the non-invasive estimation of FFR directly within the cardiac catheterization laboratory. Traditional diagnostic pathways for assessing CA stenosis are fundamentally divided between anatomical evaluation via ICA and subsequent functional assessment using pressure-wire-derived FFR. While the former may lack ischemic accuracy, the latter adds cost, procedural time, and rare but severe risks, such as arterial dissection. To overcome these limitations, this research bridged the gap by training a high-dimensional neural network architecture that can operate simultaneously on heterogeneous clinical inputs.

The core experimental pipeline utilized a selected dataset comprising 179 unique patients and 256 high-resolution angiograms. The architecture was engineered to ingest and process two distinct streams of patient data: spatiotemporal representations extracted from single-view ICA videos using an optimized 3D CNN, and a comprehensive vector of peripheral blood parameters and hemorheological markers processed via a specialized tabular DL backbone. Rather than using conventional late fusion mechanisms, which simply concatenate high level features at the final decision boundary, we introduced an intermediate fusion strategy that leverages FiLM. The FiLM layers dynamically conditioned the convolutional feature maps of the vision model based on the patient's unique hematological profile, effectively acting as an intelligent physiological gate. Finally, a composite output layer was developed to simultaneously produce a continuous FFR regression estimate and a binary ischemia classification probability, with strict calibration to ensure diagnostic internal coherence at the critical FFR ≤ 0.80 clinical threshold. The consistency of the ratio score (p_{cls} / p_{reg}) as a composite predictor further reinforces the model's clinical reliability, ensuring internal coherence between regression and

classification tasks. Inference through the fully trained pipeline is exceptionally lightweight, executing in milliseconds.

The empirical performance of the isolated sub-models validated the study's primary biophysical hypotheses. The baseline Video-Only model exhibited a critical operational failure, yielding a test AUC of 0.480. This starkly underscores that purely visual, geometric representations derived from standard ICA are insufficient for predicting functional flow. While traditional deep learning models, such as the ResNet18 patch architecture used by previous authors, can capture subtle macrovascular narrowing, they treat the coronary vasculature as an arrangement of rigid, static pipes. In doing so, they ignore systemic fluid resistance, plaque composition, microvascular compliance, and the underlying rheology of the fluid itself. Conversely, the Tabular-Only model, operating strictly on peripheral blood metrics, achieved a test AUC of 0.561. This confirms that systemic hematological parameters carry a weak but statistically significant independent signal regarding localized coronary ischemia. A major contribution of this work is the mathematical formalization of the EFI, derived by multiplying RDW and MCHC. Elevated EFI scores captured increased erythrocyte size heterogeneity and membrane rigidity, which alter local blood viscosity, accelerate chaotic cell-to-cell collisions under high shear rates, and cause friction-induced endothelial injury. Furthermore, while the literature associates high MPV with adverse cardiovascular events, our population-specific findings revealed that physiologic, normal-range MPV values were associated with preserved FFR. This indicates that structurally intact, unfragmented platelets maintain vital endothelial barrier function and shield microvessels from erythrocyte-mediated mechanical stress, thereby preventing downstream microcirculatory resistance.

Architecturally, the intermediate FiLM fusion model achieved a superior test AUC of 0.606, significantly outperforming the Late Fusion concatenation alternative (AUC 0.547). This empirical gap confirms the theoretical validity of our gating mechanism: conditioning the early and intermediate layers of a 3D CNN with hemorheological vectors allows the network to dynamically adapt its spatial feature

extraction to the specific fluid properties of that patient, rather than extracting generic structural features. The systematic divergence observed between our cross-validation metrics (CV AUC = 0.674) and test metrics (Test AUC = 0.606) was rigorously analyzed and attributed to the small-sample-size constraint inherent in the stratified K-fold iterator variant. With approximately 29 validation patients per fold, a single misclassification skewed the fold metric by roughly 0.035, a sampling variance that we mitigated through exhaustive multiple seed runs and confidence interval tracking.

When placed in the context of contemporary literature, the developed architecture offers unique advantages. Traditional frameworks in the literature rely heavily on intensive manual preprocessing, the extraction of dozens of hand-crafted geometric stenosis features, or massive datasets to achieve high accuracy. For instance, Cho et al. achieved an accuracy of 82%, a sensitivity of 84%, and a specificity of 80% using a dual-stage XGBoost model; however, their approach required extensive manual measurement of vessel diameters and the extraction of 24 specific stenosis features from static keyframes, completely omitting continuous FFR regression and mean absolute error evaluation. Similarly, other single-view angiography video frameworks face distinct data scale dependencies. Langlais et al. utilized a massive dataset exceeding 50,000 videos across four cohorts to predict stenosis percentages, achieving a best mean absolute error of 0.077 for left main coronary artery videos; yet, their ground truth relied heavily on calculated fractional flow reserve surrogates and expert opinion rather than direct invasive wire measurements, and their system suffered from high computational latency, demanding tens of seconds for inference. In contrast, Mineo et al. trained pure vision-based pipelines on 341 single-view videos paired with direct invasive FFR values, yielding a superior regression mean absolute error of 0.063 and a classification AUC of 0.95.

While benchmark studies in the literature often demonstrate superior absolute predictive metrics due to their vastly superior dataset scales, a direct, parallel comparison with our framework is not conceptually or mathematically feasible. Commercial alternatives and physics-based models often mandate multi-view

angiography and demand substantial computational overhead, frequently requiring tens of seconds to operate. Crucially, there is currently no other study in the literature that combines peripheral blood parameters with angiography videos, making our multimodal paradigm fundamentally unique. Our core argument remains that systemic blood parameters directly enhance video processing pipelines. While other frameworks operate on data volumes that are orders of magnitude larger, our model establishes a foundational and structurally fundamental link between FFR values, angiography streams, and universally available blood parameters using the smallest dataset footprint in the published field. This demonstrates that incorporating simple, zero additional cost hemorheological vectors can effectively compensate for the limitations of small datasets and the lack of multi-view visual inputs.

Despite promising foundational results, several key limitations must be addressed in future research. First, the current sample size of 179 patients yields high statistical variance across validation folds, making it imperative to validate the framework on an extensive, multicenter external cohort to ensure universal clinical generalizability and eliminate institutional bias. Second, the vision architecture is currently constrained to single-coronary-vessel video inputs, thereby failing to capture the complex, compounding hemodynamic interactions present in multi-vessel disease patterns or diffuse sequential stenoses. Future work should follow two distinct, parallel tracks: biophysical expansion and imaging enhancement. On the biological track, the tabular profile should be augmented with direct rheological and biochemical measurements. On the imaging track, future models may integrate high-resolution plaque-type morphology and geometry derived from CCTA. By training the multimodal network to combine the localized impact of plaque composition on vessel compliance with the high-dimensional blood vector, the next generation of this architecture will achieve the detailed, personalized precision required to fully revolutionize noninvasive ischemic diagnostics.

In conclusion, this research marks a shift in the paradigm for non-invasive assessment of coronary physiology. By moving away from purely structural, geometry-driven vision models and physics-based simulations bound by restrictive

rigid-wall assumptions, this study has established the feasibility of anchoring DL prediction of fluid dynamics in the cellular properties of the patient's blood. The empirical superiority of the Multimodal FiLM framework over both unimodal baselines and conventional late fusion structures demonstrates that early-stage biophysical integration is vital for capturing the intricate interplay between macrovascular stenosis and systemic rheology. From a clinical perspective, the developed model offers an efficient, zero-cost, and zero-risk diagnostic support tool. Operating at millisecond-scale inference, it minimizes diagnostic ambiguity near the critical 0.80 threshold via its internally consistent composite output layer, offering a viable pathway toward real-time, personalized interventional planning during active catheterization.

REFERENCES

[1] W. H. Organization, «Cardiovascular diseases (CVDs),» [Çevrimiçi]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds). [Erişildi: 31 July 2025].

[2] T. İ. Kurumu, «Ölüm ve Ölüm Nedeni İstatistikleri, 2024,» [Çevrimiçi]. Available: https://data.tuik.gov.tr/Bulten/Index?p=Olum-ve-Olum-Nedeni-Istatistikleri-2024-54195. [Erişildi: 18 September 2025].

[3] R. A. Byrne, X. Rossello, J. J. Coughlan ve et al., «2023 ESC Guidelines for the management of acute coronary syndromes: Developed by the task force on the management of acute coronary syndromes of the European Society of Cardiology (ESC),» *European Heart Journal,* p. 3720–3826, 7 October 2023.

[4] G. Ceyhun ve O. Birdal, «Relationship between Whole Blood Viscosity and Lesion Severity in Coronary Artery Disease,» *International Journal of Angiology,* pp. 117-121, 19 January 2021.

[5] A. Akdi, E. H. Çetin, Ö. Ç. Karaaslan, M. A. Erdöl, M. O. Özilhan, O. Maden ve D. Aras, «The Role of Whole Blood Viscosity Estimated by De Simone's Formula in Evaluation of Fractional Flow Reserve,» *E Journal of Cardiovascular Medicine,* pp. 18-24, 1 March 2022.

[6] M. N. Kaçar, İ. Ulusoy ve Ç. Yayla, «A Review on the Estimation of Coronary Fractional Flow Reserve Using Artificial Intelligence,» *Archives of the Turkish Society of Cardiology,* cilt 53, no. 4, pp. 275-280, 2025.

[7] T. Swinn, D. O'Shea ve A. Dastidar, «A Year in Coronary Artery Disease: A Focussed 2024 Update,» *British Journal of Hospital Medicine,* pp. 1-28, 19 December 2025.

[8] A. Timmis, V. Aboyans, P. Vardas, N. Townsend, A. Torbica, M. Kavousi, G. Boriani, R. Huculeci, D. Kazakiewicz, D. Scherr, E. Karagiannidis, M. Cvijic, A. Kaplon-Cieslicka, B. Ignatiuk, P. Raatikainen, D. De Smedt, A. Wood, D. Dudek, E. Van Belle, F. Weidinger ve ESC National Cardiac Societies, «European Society of Cardiology: The 2023 Atlas of Cardiovascular Disease Statistics,» *European Heart Journal,* pp. 4019-4062, 7 October 2024.

[9] A. A. Armoundas, S. M. Narayan, D. K. Arnett, K. Spector-Bagdady, D. A. Bennett, L. A. Celi, P. A. Friedman, M. H. Gollob, J. L. Hall, A. E. Kwitek, E. Lett, B. K. Menon, K. A. Sheehan ve S. S. Al-Zaiti, «Use of Artificial Intelligence in Improving Outcomes in Heart Disease: A Scientific Statement From the American Heart Association,» *Circulation,* pp. 1028-1050, 2 April 2024.

[10] H. R. Schelbert, «Anatomy and physiology of coronary blood flow,» *Journal of Nuclear Cardiology,* pp. 545-554, 17 August 2010.

[11] S. Rehman, A. Khan ve A. Rehman, Physiology, Coronary Circulation, StatPearls Publishing, 2023.

[12] R. D. Shahjehan, S. Sharma ve B. S. Bhutta, Coronary Artery Disease, StatPearls Publishing, 2024.

[13] C. M. Allen, «Radiographic Distances,» %1 içinde *Digital Radiographic Exposure: Principles & Practice*, Missouri, Pressbooks, 2022.

[14] L. Calmac, N.-M. Popa-Fotea, V. Bataila, V. Ploscaru, A. Turea, I. A. Tache, D. Stoian, L. Itu, E. Badila, A. Scafa-Udriste ve M. Dorobantu, «Importance of visual estimation of coronary artery stenoses and use of functional evaluation for appropriate guidance of coronary revascularization—multiple operator evaluation,» *Diagnostics (Basel),* p. 2241, 11 2021.

[15] N. H. J. Pijls, B. DeBruyne, K. Peels, P. H. van der Voort, H. J. R. M. Bonnier, J. Bartunek ve J. J. Koolen, «Measurement of fractional flow reserve to assess the functional severity of coronary-artery stenoses,» *New England Journal of Medicine,* pp. 334(26):1703-1708, 1996.

[16] P. A. L. Tonino, B. De Bruyne, N. H. J. Pijls, U. Siebert, F. Ikeno, M. van't Veer, V. Klauss, G. Manoharan, T. Engstrom, K. G. Oldroyd, P. N. Ver Lee, P. A. MacCarthy ve W. F. Fearon, «Fractional flow reserve versus angiography for guiding percutaneous coronary intervention,» *N Engl J Med,* pp. 360(3):213-224, 2009.

[17] G. R. Heyndrickx ve G. G. Toth, «The FAME Trials: Impact on Clinical Decision Making,» *Interventional Cardiology Review,* pp. 11(2):116-119, 2016.

[18] L. Dean, Blood Groups and Red Cell Antigens, National Center for Biotechnology Information, 2005.

[19] J. El Brihi ve S. Pathak, Normal and Abnormal Complete Blood Count With Differential, StatPearls Publishing, 2024.

[20] G. De Simone, R. B. Devereux, S. Chien, M. H. Alderman, S. A. Atlas ve J. H. Laragh, «Relation of Blood Viscosity to Demographic and Physiologic Variables and to Cardiovascular Risk Factors in Apparently Normal Adults,» *Circulation,* pp. 107-117, January 1990.

[21] V. P. Demkin, S. V. Mel'nichuk, K. V. Zavadovskii, M. N. Khoryak, V. V. Rudenko, A. T. Suyundukova, D. N. Kukartseva, A. O. Tsibulina ve V. V. Udut, «The Influence of dynamic blood viscosity on coronary bloow flow in stenotic artery,» *Russian Physics Journal,* pp. 2364-2370, April 2022.

[22] T. W. Secomb, «Comprehensive Physiology,» *Hemodynamics,* cilt 6, no. 2, pp. 975-1003, 2016.

[23] A. A. Tokarev, A. A. Butylin, E. A. Ermakova, E. E. Shnol, G. P. Panasenko ve F. I. Ataullakhanov, «Finite Platelet Size Could Be Responsible for Platelet Margination Effect,» *Biophysical Journal,* cilt 101, no. 8, pp. 1835-1843, 2011.

[24] J. Cao, Z. Deng, L. Xiang, F. Wang ve T. Yang, «Nonlinear association between red cell distribution width/albumin ratio and peripheral arterial disease in the NHANES: a cross-section study with propensity score matching analysis,» *Frontiers in Cardiovascular Medicine,* cilt 12, 2025.

[25] B. Yousefi, S. Sanaie, A. A. Ghamari, S. Hassan, A. Karimian ve A. Mahmoodpoor, «Red Cell Distribution Width as a Novel Prognostic Marker in Multiple Clinical Studies,» *Indian Journal of Critical Care Medicine,* cilt 24, no. 1, pp. 49-54, 2020.

[26] M. H. A. M. Fens, R. van Wijk, G. Andringa, K. L. van Roojien, H. M. Dijstelbloem, J. T. Rasmussen, K. M. K. de Vooght, R. M. Schiffelers, C. A. J. M. Gaillard ve W. W. van Solinge, «A role for activated endothelial cells in red blood cell clearance: implications for vasopathology,» *Haematologica,* cilt 97, no. 4, pp. 500-508, 2012.

[27] C. Caruso, M. E. Fay, X. Cheng, A. Y. Liu, S. I. Park, T. A. Sulchek, M. D. Graham ve W. A. Lam, «Pathologic mechanobiological interactions between red blood cells and endothelial cells directly induce vasculopathy in iron deficiency anemia,» *iScience,* cilt 25, no. 7, 2022.

[28] A. Candreva, G. De Nisco, M. L. Rizzini, F. D'Ascenzo, G. M. De Ferrari, D. Gallo, U. Morbiducci ve C. Chiastra, «Current and Future Applications of Computational Fluid Dynamics in Coronary Artery Disease,» *Reviews in Cardiovascular Medicine,* 2022.

[29] M. Stiffler, E. Seibert, C.-W. Lin, S. Panigrahi, P. Moore, M. L. Moorman, R. A. Callcut, J. D. Reynolds, K. E. Remy ve J. T. Ross, «Free Hemoglobin

Increases Endothelial Adhesiveness to Red Blood Cells via a Toll-like Receptor 4-Dependent Mechanism,» *Journal of Surgical Research,* cilt 321, pp. 609-615, 2026.

[30] K. J. McCarthy, E. A. Larnard, C. K. Anderson, M. Chowdhury, R. Shalev, D. A. Hakim, K. J. Croce ve E. A. Osborn, «Automated Artificial Intelligence Mapping of Coronary Plaque Calcification: A Comparison with Manual Intravascular Image Analysis,» *Journal of Clinical Medicine,* 18 November 2025.

[31] J. Han, Z. Wang, T. Chen, S. Liu, J. Tan, Y. Sun, L. Feng, D. Zhang, L. Ma, H. Liu, H. Tao, C. Jin, C. Fang, H. Yu, M. Zeng, H. Jia ve B. Yu, «Artificial intelligence driven plaque characterization and functional assessment from CCTA using OCT-based automation: A prospective study,» *International Journal of Cardiology,* 1 June 2025.

[32] İ. Atlı ve O. S. Gedik, «3D reconstruction of coronary arteries using deep networks from synthetic X-ray angiogram data,» *Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering,* pp. 1-20, 30 June 2022.

[33] C. A. Taylor, T. A. Fonte ve J. K. Min, «Computational Fluid Dynamics Applied to Cardiac Computed Tomography for Noninvasive Quantification of Fractional Flow Reserve,» *Journal of the American College of Cardiology,* pp. 2233-2241, 4 June 2013.

[34] C. Tesche, C. N. De Cocco, S. Baumann, M. Renker, T. W. McLaurin, T. M. Duguay, 2. R. R. Bayer, D. H. Steinberg, K. L. Grant, C. Canstein, C. Schwemmer, M. Schoebinger, L. M. Itu, S. Rapaka, P. Sharma ve U. J. Schoepf, «Coronary CT Angiography-derived Fractional Flow Reserve: Machine Learning Algorithm versus Computational Fluid Dynamics Modeling,» *Radiology,* pp. 64-72, July 2018.

[35] G. Nannini, J. Suk, P. Rygiel, S. Saitta, L. Mariani, R. Maranga, A. Baggiano, G. Pontone, J. M. Wolterink ve A. Redaelli, «Learning hemodynamic scalar fields on coronary artery meshes: A benchmark of geometric deep learning models,» *Computers in Biology and Medicine,* September 2025.

[36] J. Steinfeldt, T. Buergel, L. Loock, P. Kittner, G. Ruyoga, J. U. Zu Belzen, S. Sasse, H. Strangalies, L. Christmann, N. Hollmann, B. Wolf, B. Ference, J. Deanfield, U. Landmesser ve R. Eils, «Neural network-based integration of polygenic and clinical information: development and validation of a prediction model for 10-year risk of major adverse cardiac events in the UK Biobank cohort,» *Lancet Digital Health,* pp. e84-e94, February 2022.

[37] S. F. Weng, J. Reps, J. Kai, J. M. Garibaldi ve N. Qureshi, «Can machine-learning improve cardiovascular risk prediction using routine clinical data?,» *Plos One,* 4 April 2017.

[38] D. Tran, H. Wang, L. Torresani, J. Ray, Y. LeCun ve M. Paluri, «A closer look at spatiotemporal convolutions for action recognition,» *Computer Vision and Pattern Recognition,* p. 6450–6459, 2018.

[39] W. Kay, J. Carreira, K. Simonyan, B. Zhang, C. Hillier, S. Vijayanarasimhan, F. Viola, T. Green, T. Back, P. Natsev, M. Suleyman ve A. Zisserman, «The Kinetics Human Action Video Dataset,» *arXiv,* 2017.

[40] S. Xie, C. Sun, J. Huang, Z. Tu ve K. Murphy, «Rethinking Spatiotemporal Feature Learning: Speed-Accuracy Trade-offs in Video Classification,» %1 içinde *European Conference on Computer Vision - ECCV*, 2018.

[41] S. Hochreiter ve J. Schmidhuber, «Long Short-Term Memory,» *Neural Computation,* pp. 1735-1780, 15 November 1997.

[42] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-k. Wong ve W.-c. Woo, «Convolutional LSTM Network: a machine learning approach for precipitation nowcasting,» *NIPS'15: Proceedings of the 29th International Conference on Neural Information Processing Systems,* pp. 802 - 810, 7 December 2015.

[43] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser ve I. Polosukhin, «Attention Is All You Need,» %1 içinde *31st Conference on Neural Information Processing Systems (NIPS)*, Long Beach, CA, 2017.

[44] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit ve N. Houlsby, «An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,» %1 içinde *International Conference on Learning Representations*, Austria, 2021.

[45] A. Arnab, M. Dehghani, G. Heigold, C. Sun, M. Lucic ve C. Schmid, «ViViT: A Video Vision Transformer,» %1 içinde *International Conference on Computer Vision - ICCV*, 2021.

[46] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin ve B. Guo, «Swin Transformer: Hierarchical Vision Transformer using ShiftedWindows,» %1 içinde *International Conference on Computer Vision - ICCV*, 2021.

[47] J. H. Friedman, «Greedy Function Approximation: A Gradient Boosting Machine,» *Annals of Statistics,* pp. 1189-1232, 2001.

[48] T. Chen ve C. Guestrin, «XGBoost: A Scalable Tree Boosting System,» *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* pp. 785-794, 13 August 2016.

[49] E. Perez, F. Strub, H. de Vries, V. Dumoulin ve A. Courville, «FiLM: Visual Reasoning with a General Conditioning Layer,» *AAAI'18/IAAI'18/EAAI'18: Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence,* pp. 3942-3951, 18 December 2017.

[50] E. Ben-Assa, A. Abu Salman, C. Cafri, A. Roguin, E. Hellou, E. Koifman, Y. Feld, E. Lev, G. Sheinman, E. Harari, A. Abu Dogosh, R. Beyar, H. M. Garcia-Garcia, J. Davies ve O. Ben-Yehuda, «Performance of a novel artificial intelligence software developed to derive coronary fractional flow reserve values from diagnostic angiograms,» *Coronary Artery Disease,* pp. 533-541, December 2023.

[51] L. O. Müller, F. E. Fossan, J. Sturdy, A. T. Braten, A. Jorgensen, R. Wiseth ve L. R. Hellevik, «Machine learning augmented reduced-order models for FFR-prediction,» *Computer methods in applied mechanics and engineering,* 24 April 2021.

[52] T. C. Villines, «Can CT-derived FFR better inform clinical decision-making and improve outcomes in stable ischaemic heart disease?,» *European Heart Journal,* pp. 3712-3714, 25 August 2018.

[53] Z. Yan, Z. Yao, W. Guo, D. Shang, R. Chen, J. Liu, X.-C. Cai ve J. Ge, «Impact of Pressure Wire on Fractional Flow Reserve and Hemodynamics of the Coronary Arteries: A Computational and Clinical Study,» *IEEE Transactions on Biomedical Engineering,* pp. 1683-1691, 5 May 2023.

[54] H. J. Lee, Y. W. Kim, J. H. Kim, Y.-J. Lee, J. Moon, P. Jeong, J. Jeong, J.-S. Kim ve J. S. Lee, «Optimization of FFR prediction algorithm for gray zone by hemodynamic features with synthetic model and biometric data,» *Computer Methods and Programs in Biomedicine,* June 2022.

[55] W. F. Fearon, S. Achenbach, T. Engstrom, A. Assali, R. Schlofmitz, A. Jeremias, S. Fournier, A. J. Kirtane, R. Kornowski, G. Greenberg, R. Jubeh, D. M. Kolansky, T. McAndrew, O. Dressler, A. Maehara, M. Matsumura, M. B. Leon ve B. DeBruyne, «Accuracy of Fractional Flow Reserve Derived From Coronary Angiography,» *Circulation,* pp. 477-484, 22 January 2019.

[56] H. Zhang, X. Song, R. Wu, N. Li, Q. Hou, J. Xie, Y. Hou ve A. Qiao, «A novel method for noninvasive quantification of fractional flow reserve based on the custom function,» *Frontiers in Bioengineering and Biotechnology,* 30 August 2023.

[57] A. Roguin, A. Abu Dogosh, Y. Feld, M. Konigstein, A. Lerman ve E. Koifman, «Early Feasibility of Automated Artificial Intelligence Angiography Based Fractional Flow Reserve Estimation,» *The American Journal of Cardiology,* pp. 8-14, 15 January 2021.

[58] W. F. Griffin, A. D. Choi, J. S. Riess, H. Marques ve e. al, «AI Evaluation of Stenosis on Coronary CTA, Comparison With Quantitative Coronary Angiography and Fractional Flow Reserve,» *JACC: Cardiovascular Imaging,* pp. 193-205, 16 February 2023.

[59] F. M. Zimmermann, T. P. Mast, N. P. Johnson, I. Everts, B. Hennigan, C. Berry, D. T. Johnson, B. De Bruyne, W. F. Fearon, K. Oldroyd, N. H. J. Pijls, P. A. L. Tonino ve M. v. Veer, «Deep learning for prediction of fractional flow reserve from resting coronary pressure curves,» *Eurointervention,* pp. 51-58, September 2020.

[60] A. Lin, P. A. van Diemen, M. Motwani, P. McElhinney, Y. Otaki, D. Han, A. Kwan, E. Tzolos, E. Klein, K. Kuronuma, K. Grodecki, B. Shou, R. Rios, N. Manral, S. Cadet, I. Danad, R. S. Driessen, D. S. Berman, B. L. Norgaard, P. J. Slomka, P. Knaapen ve D. Dey, «Machine Learning From Quantitative Coronary Computed Tomography Angiography Predicts Fractional Flow

Reserve–Defined Ischemia and Impaired Myocardial Blood Flow,» *Circulation: Cardiovascular Imaging,* pp. 710-720, 13 October 2022.

[61] S. Candemir, R. D. White, M. Demirer, V. Gupta, M. T. Bigelow, L. M. Prevedello ve B. S. Erdal, «Automated coronary artery atherosclerosis detection and weakly supervised localization on coronary CT angiography with a deep 3-dimensional convolutional neural network,» *Computerized Medical Imaging and Graphics,* July 2020.

[62] F. E. Fossan, L. O. Müller, J. Sturdy, A. T. Braten, A. Jorgensen, R. Wiseth ve L. R. Hellevik, «Machine learning augmented reduced-order models for FFR-prediction,» *Computer Methods in Applied Mechanics and Engineering,* 30 May 2021.

[63] C. Zhou, T. V. Dinh, H. Kong, J. Yap, K. K. Yeo, H. K. Lee ve K. Liang, «Automated Deep Learning Analysis of Angiography Video Sequences for Coronary Artery Disease,» *arXiv.org,* 29 January 2021.

[64] H. Cho, J.-G. Lee, S.-J. Kang, W.-J. Kim, S.-Y. Choi, J. Ko, H.-S. Min, G.-H. Choi, D.-Y. Kang, P. H. Lee, J.-M. Ahn, D.-W. Park, S.-W. Lee, Y.-H. Kim, C. W. Lee, S.-W. Park ve S.-J. Park, «Angiography-Based Machine Learning for Predicting Fractional Flow Reserve in Intermediate Coronary Artery Lesions,» *Journal of the American Heart Association,* 4 January 2019.

[65] T. Mahendiran, D. Thanou, O. Senouf, D. Meier, N. Dayer, F. Aminfar, D. Auberson, O. Raita, P. Frossard, M. Pagnoni, S. Cook, B. De Bruyne, O. Muller, E. Abbe ve S. Fournier, «Deep learning-based future myocardial infarction using invasive coronary angiography: a feasibility study,» *Openheart,* 3 January 2023.

[66] Q. Zhao, C. Li, M. Chu, J. L. Gutierrez-Chico ve S. Tu, «Angiography-based coronary flow reserve: The feasibility of automatic computation by artificial intelligence,» *Cardiology Journal,* pp. 369-378, 2023.

[67] E. L. Langlais, D. Corbin, O. Tastet, A. Hayek, G. Doolub, S. Mrad, J.-C. Tardiff, J.-F. Tanguay, G. Marquis-Gravel, G. H. Tison, S. Kadoury, W. Le, R. Gallo, F. Lesage ve R. Avram, «Evaluation of stenoses using AI video models applied to coronary angiography,» *Digital Medicine 7:138,* 23 May 2024.

[68] R. Mineo, F. P. Salanitri, G. Bellitto, I. Kavasidis, O. De Filippo, M. Millesimo, G. M. De Ferrari, M. Aldinucci, D. Giordano, S. Palazzo, F. D'Ascenzo ve C. Spampinato, «A Convolutional-Transformer Model for FFR and iFR Assessment From Coronary Angiography,» *IEEE Transactions on Medical Imaging,* pp. 2866-2877, August 2024.

[69] B. Xie, X. Liu, H. Zhang, C. Xu, T. Zeng, Y. Yuan, G. Yang ve Z. Gao, «Conditional Physics-Informed Graph Neural Network for Fractional Flow Reserve Assessment,» %1 içinde *Medical Image Computing and Computer Assisted Intervention*, Vancouver, 2023.

[70] J. M. Carson, N. K. Chakshu, I. Sazonov ve P. Nithiarasu, «Artificial intelligence approaches to predict coronary stenosis severity using non-invasive fractional flow reserve,» *Proceedings of the Institution of Mechanical Engineers,* pp. 1337-1350, 30 June 2020.

[71] H. Sun, J. Liu, Y. Feng, X. Xi, K. Xu, L. Zhang, J. Liu, B. Li ve Y. Liu, «Deep learning-based prediction of coronary artery stenosis resistance,» *American Journal of Physiology-Heart and Circulatory Physiology,* p. H1194–H1205, 21 October 2022.

[72] G. Muscogiuri, M. Chiesa, A. Baggiano, P. Spadafora, R. De Santis, M. Guglielmo, S. Scafuri, L. Fusini, S. Mushtaq, E. Conte, A. Annoni, A.

Formenti, M. E. Mancini, F. Ricci, F. P. Ariano, L. Spiritigliozzi, M. Babbaro, R. Mollace, R. Maragna, C. M. Giacari, D. Andreini, A. I. Guaricci, G. I. Colombo, M. G. Rabbat, M. Pepi, F. Sardanelli ve G. Pontone, «Diagnostic performance of deep learning algorithm for analysis of computed tomography myocardial perfusion,» *European Journal of Nuclear Medicine and Molecular Imaging,* pp. 3119-3128, 23 February 2022.

[73] J. Xue, J. Li, D. Sun, L. Sheng, Y. Gong, D. Wang, S. Zhang, Y. Zou, J. Shi, W. Xu, M. An, C. Dai, W. Li, L. Zheng, A. Vinograd, G. Liu, Y. Kong ve Y. Li, «Functional evaluation of intermediate coronary lesions with integrated computed tomography angiography and invasive angiography in patients with stable coronary artery disease,» *Journal of Translational Internal Medicine,* pp. 255-263, 10 June 2022.

[74] M. Farajtabar, M. M. Larimi, M. Biglarian, D. Sabour ve M. Miansari, «Machine Learning Identification Framework of Hemodynamics of Blood Flow in Patient-Specific Coronary Arteries with Abnormality,» *Journal of Cardiovascular Translational Research,* pp. 722-737, 18 November 2022.

[75] J.-J. Cha, N.-L. Nguyen, C. Tran, W.-Y. Shin, S.-G. Lee, Y.-J. Lee, S.-J. Lee, S.-J. Hong, C.-M. Ahn, B.-K. Kim, Y.-G. Ko, D. Choi, M.-K. Hong, Y. Jang, J. Ha ve J.-S. Kim, «Assessment of fractional flow reserve in intermediate coronary stenosis using optical coherence tomography-based machine learning,» *Frontiers in Cardiovascular Medicine,* 25 January 2023.

[76] B. Li, J. Liu, Y. Yang, H. Sun, C. Zhu, L. H. Liang, Z. Wang, L. Fan, S. Huang, X. Li, G. Li, L. Zhang, Y. Zhang, J. Dong, B. Mao, C. Hou, J. Liu, W. Tan ve Y. Liu, «Non-invasive Computational Techniques for Diagnosing Myocardial Ischemia: Challenges and Future of FFRCT/iFRCT,» *Annals of Biomedical Engineering,* 2026.

[77] Z. Teed ve J. Deng, «RAFT: Recurrent All-Pairs Field Transforms for Optical Flow,» %1 içinde *European Conference on Computer Vision (ECCV)*, 2020.

[78] S. Ji, W. Xu, M. Yang ve K. Yu, «3D Convolutional Neural Networks for Human Action Recognition,» *IEEE Transactions on Pattern Analysis and Machine Intelligence,* pp. 221-231, 6 March 2012.

[79] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh ve D. Batra, «Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,» *International Journal of Computer Vision,* cilt 128, no. 2, pp. 336-359, 2020.

[80] B. H. Foy, R. Petherbridge, M. T. Roth, C. Zhang, D. C. De Souza, C. Mow, H. R. Patel, C. H. Patel, S. N. Ho, E. Lam, C. E. Powe, R. P. Hasserjian, K. J. Karczewski, V. Tozzo ve J. M. Higgins, «Haematological setpoints are a stable and patient-specific deep phenotype,» *Nature,* cilt 637, no. 8045, pp. 430-438, 2025.

[81] S. Hernandez-Orozco, A. Uthamacumaran, F. H. Quiroz, K. Saeb-Parsy ve H. Zenil, «Integrative adaptive indexes from noisy routine haematological markers can predict and discriminate health status and biological age,» *Computers in Biology and Medicine,* cilt 208, 2026.

[82] T. Alexy, J. Detterich, P. Connes, K. Toth, E. Nader, P. Kenyeres, J. Arriola-Montenegro, P. Ulker ve M. J. Simmonds, «Physical Properties of Blood and their Relationship to Clinical Conditions,» *Frontiers in Physiology,* cilt 13, 2022.

[83] W. Jian, X. Shen, Z. Zheng, Z. Wu, Y. Shi ve J. Liu, «Association Between Red Blood Cell Distribution Width and Coronary Calcification in Patients

Referred for Invasive Coronary Angiography,» *Angiology,* cilt 76, no. 7, pp. 690-698, 2025.

[84] E. Hosseini, S. M. S. Pezeshki, A. Ghasemzadeh ve M. Ghasemzadeh, «Indications of platelet indices in coronary artery disease (CAD): more insights into their diagnostic and prognostic importance,» *Journal of Thrombosis and Thrombolysis,* cilt 58, no. 7, pp. 806-820, 2025.

[85] «Platelets maintain vascular barrier function in the absence of injury or inflammation,» *Journal of Thrombosis and Haemostasis,* cilt 19, no. 5, pp. 1145-1148, 2021.

APPENDICES

A. Minimum and Maximum values of Modality Performances

Table A.1 Minimum values of Cross-Validation Modality Performance with all blood parameters

Model	*Tabular only*	*Multimodal*	*Late Fusion*	*Video only*
AUC	0,503	0,551	0,524	0,525
MAE	0,110	0,111	0,110	0,108
Sensitivity	0,569	0,543	0,829	0,890
Specificity	0,578	0,767	0,874	0,857
Balanced Accuracy	0,545	0,573	0,553	0,527

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

Table A.2 Minimum values of Cross-Validation Modality Performance with selected blood parameters

Model	*Tabular only*	*Multimodal*	*Late Fusion*	*Video only*
AUC	0,572	0,634	0,583	0,525
MAE	0,113	0,107	0,105	0,108
Sensitivity	0,609	0,652	0,821	0,890
Specificity	0,528	0,741	0,911	0,857
Balanced Accuracy	0,556	0,603	0,579	0,527

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

Table A.3 Maximum values of Cross-Validation Modality Performance with all blood parameters

Model	*Tabular only*	*Multimodal*	*Late Fusion*	*Video only*
AUC	0,683	0,677	0,680	0,687
MAE	0,128	0,134	0,129	0,126
Sensitivity	0,838	0,739	0,948	1
Specificity	0,780	0,849	1	1
Balanced Accuracy	0,674	0,663	0,660	0,625

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

Table A.4 Maximum values of Cross-Validation Modality Performance with selected blood parameters

Model	*Tabular only*	*Multimodal*	*Late Fusion*	*Video only*
AUC	0,651	0,712	0,680	0,687
MAE	0,125	0,131	0,126	0,126
Sensitivity	0,919	0,831	1	1
Specificity	0,791	0,860	0,985	1
Balanced Accuracy	0,632	0,687	0,665	0,625

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

Table A.5 Minimum values of Test Set Modality Performance with all blood parameters

Model	*Tabular only*	*Multimodal*	*Late Fusion*	*Video only*
AUC	0,402	0,398	0,449	0,420
MAE	0,114	0,129	0,130	0,138
Sensitivity	0,353	0,235	0,175	0,116
Specificity	0,411	0,456	0,321	0,263
Balanced Accuracy	0,421	0,401	0,452	0,414

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

Table A.6 Minimum values Test Set Modality Performance with selected blood parameters

Model	*Tabular only*	*Multimodal*	*Late Fusion*	*Video only*
AUC	0,442	0,478	0,435	0,420
MAE	0,113	0,132	0,123	0,137
Sensitivity	0,453	0,344	0,263	0,116
Specificity	0,329	0,388	0,318	0,263
Balanced Accuracy	0,441	0,453	0,457	0,414

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

Table A.7 Maximum values of Test Set Modality Performance with all blood parameters

Model	*Tabular only*	*Multimodal*	*Late Fusion*	*Video only*
AUC	0,624	0,661	0,701	0,588
MAE	0,173	0,176	0,198	0,216
Sensitivity	0,738	0,610	0,713	0,693
Specificity	0,657	0,817	0,740	0,794
Balanced Accuracy	0,656	0,626	0,589	0,551

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

Table A.8 Maximum values of Test Set Modality Performance with selected blood parameters

Model	*Tabular only*	*Multimodal*	*Late Fusion*	*Video only*
AUC	0,668	0,767	0,657	0,588
MAE	0,167	0,185	0,205	0,216
Sensitivity	0,747	0,825	0,676	0,693
Specificity	0,685	0,764	0,810	0,794
Balanced Accuracy	0,631	0,680	0,596	0,551

Abbreviations: AUC, area under curve, FiLM, feature wise linear modulation; MAE, mean absolute error

CURRICULUM VITAE

Kaçar, Mehmet Nazir

EDUCATION

Degree	Institution	Year of Graduation
MS	METU Electrical-Electronics Engineering, Ankara	2001
BS	METU Electrical-Electronics Engineering, Ankara	2003

FOREIGN LANGUAGES

Advanced English

PUBLICATIONS

1. M. N. Kaçar, İ. Ulusoy, and Ç. Yayla, 'A Review on the Estimation of Coronary Fractional Flow Reserve Using Artificial Intelligence', *Turk Kardiyol Dern Ars*, vol. 53, no. 4, pp. 275–280, 2025.
2. M. N. Kaçar, İ. Ulusoy, Ç. Yayla, and D. Kaçar, 'Feature-based Fusion of Angiography Videos and Blood Parameters for Fractional Flow Reserve (FFR) Prediction', in preparation.