Threshold Quantification of Coronary Artery Blockage Using A Cascaded Ensemble U-Net-Based Pipeline

by

Shadman Sakib
24141034
Momtaheena Chowdhury
21201307
Sumaiya Saba Nitu
23241129
Munira Akter Mysha
24341133

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

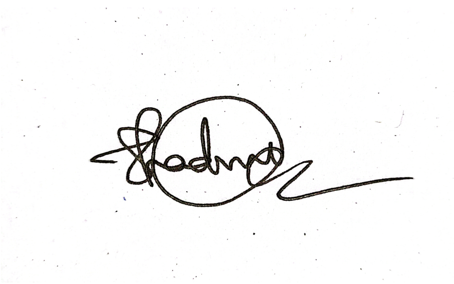
Department of Computer Science and Engineering
Brac University
June 2025

Declaration

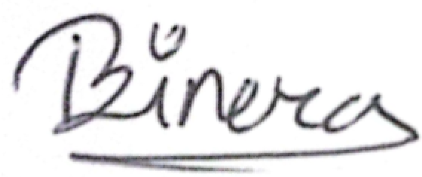
It is hereby declared that

1. The thesis submitted is our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

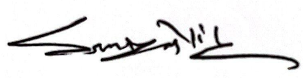


Shadman Sakib
24141034



Momtaheena Chowdhury
21201307

Sumaiya Saba Nitu
23241129



Munira Akter Mysha
24341133

Approval

The thesis titled "An Efficient Machine Learning Approach to Predict Cardiovascular and Chronic Diseases Using Feature Engineering" submitted by

1. Shadman Sakib(24141034)
2. Momtaheena Chowdhury(21201307)
3. Sumaiya Saba Nitu(23241129)
4. Munira Akter Mysha(24341133)

Of Summer, 2024 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on October 20, 2024.

Examining Committee:

Supervisor:
(Member)

Dr. Md. Ashraful Alam
Associate Professor
Department of Computer Science and Engineering
Brac University

Co-Supervisor:
(Member)

Mr. Dibyo Fabian Dofadar
Lecturer
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

Coronary artery disease remains a primary driver of morbidity and mortality worldwide, with precise quantification of arterial stenosis essential for risk stratification and treatment planning. We propose a fully automated, cascaded–ensemble deep learning framework for 3D CTA analysis that integrates coarse whole-volume segmentation with multi-model patch-level refinement and skeleton-based stenosis assessment. First, a 3D U-Net coarse network localizes the coronary tree across the entire volume. Regions of interest are then cropped via dilated centerline extraction and passed through an ensemble of four patch-level architectures—3D U-Net, 3D ResU-Net, Attention U-Net, and U-Net++—to capture fine vessel details. Predicted centerlines are skeletonized and compared against ground truth to compute percentage stenosis, leveraging distance-transform radii measurements. We evaluate performance over 40 training epochs, reporting AUC, Dice, precision, recall, F1 for training, validation, and test cohorts, alongside bootstrapped 95% CIs, Wilcoxon p-values, and correlation matrices. Visual comparisons of all model outputs on the same Frangi-enhanced slices demonstrate complementary strengths, while ensemble voting yields more robust boundary delineation. On a 1000-case CTA dataset, this research's approach achieves a mean Dice >0.80 and significantly less stenosis error, illustrating its potential and reliability than other prior researches and 2D Unet Models.

Keywords: Coronary, quantification, stenosis.

Acknowledgement

All appreciation to the Almighty Allah who has provided us the strength and endurance to successfully go through with our thesis.

We humbly appreciate our Supervisor, Dr. Ashraful Alam sir, who helped us sufficiently and maintained endless leadership and support throughout the research, and our co-supervisor, Dibyo Fabian Dofadar sir, for his kind assistance and encouragement throughout the research.

And lastly to our parents without their throughout support it might not be possible. Through their kind help and prayer we are now on the threshold of our graduation.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Global burden of cardiovascular disease

Cardiovascular disease continues to be the leading cause of death globally, fueled by preventable lifestyle factors. According to the World Health Organization, cardiovascular diseases (CVD) contribute to 17.9 million deaths annually , which is about 32% of all deaths globally. In Australia, according to the Australian Institute of Health and Welfare's (AIHW) report, in 2018, CVD was the top cause of death. It was responsible for 42% of all fatalities [7]. Abnormal coronary artery stenosis reduces the blood flow to the heart muscle, which causes hypoxia damage to myocardial cells. As a result, infarction happens among those myocardial cells, which consequently makes coronary heart disease the most common among all kinds of CVD [2].Besides, those myocardial infractions can lead to untimely cardiac arrest.To start timely interventions that may substantially decrease patient morbidity and mortality, early and precise diagnosis of coronary artery blockages is essential [63].

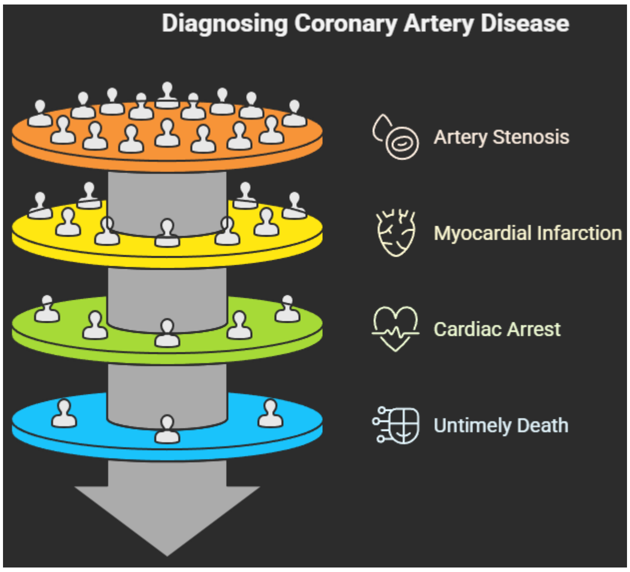


Figure 1.1: Coronary artery disease pipeline

This makes the need to detect artery stenosis early and accurately very obvious. One of the best medical practices for this purpose is Computed Tomography Angiography

(CTA). This is a non-invasive scan that produces clear 3D images of the coronary muscles and arteries, which is why it is widely used by doctors in diagnosis and planning treatments. [31]. This technique exceeds accuracy while diagnosing coronary artery diseases than other processes like magnetic resonance imaging (MRI) [6]. Currently, doctors (radiologists) process these images by hand, outlining the coronary arteries. Then they measure how narrow the arteries are. But this manual work takes a lot of time and can be susceptible to errors [67]. On top of that, the amount of medical imaging data is growing rapidly both in number and nature , which is making manual work less realistic in terms of time, cost, and consistency [36].

As a result, there's been a growing interest in using artificial intelligence, especially deep learning, to automatically find and outline coronary arteries. However, this isn't easy at all for multiple reasons. Everyone's heart structure is a bit different. Some people have arteries surrounded by fat, while others have them deep inside the heart muscle. Again, the image quality can also be affected by noise, resolution, or other issues. On top of that, coronary arteries are complicated in shape and build-up. They are long, thin, and branched out in many directions, which makes them difficult to segment and especially in cross-section views [64].

To solve these problems, researchers have mainly used two approaches. They are traditional machine learning (ML) and deep learning (DL). Traditional ML methods depend on manually chosen features and knowledge of coronary anatomy. These are often divided into pixel-based [3], [4], [11], [20] and structure-based methods [5], [15], [21], [25]. These techniques can work promisingly well, but their performance drops significantly when arteries are badly deformed. Deep learning, on the other hand, doesn't require manual feature design. And thus, has shown better effectiveness in segmenting coronary arteries [32], [38], [43], [45], [49], [54], [57], [65]

Among deep learning models, the U-Net architecture and its improved versions like U-Net++ and U-Net+ are functioning particularly well, especially in segmenting coronary arteries from 3D medical scans [61],[68]. Previously, there was one key problem that there was no particular numerical medical threshold globally to decide when a segmented artery is considered dangerously narrow. But according to the National Heart, Lung, and Blood Institute's article, coronary artery disease is considered obstructive when the artery is occupied 50% of its diameter by fat deposits; which information was used in this research.

Another issue is that current research often lacks consistent and fair comparisons. Some studies only compare their models to general deep learning models and not similar ones [50]. Others use different datasets or label things differently, which makes comparisons unreliable [19], [23]. This is partly because there are very few large, public datasets available for coronary segmentation; only one exists, with just 18 training images [13]. Also, many researchers didn't share their code. This makes it difficult for other researchers to have a decent comparison of their work.

1.2 Research Motivation

Due to the data driven approaches of hierarchical feature learning in deep learning frameworks, these advances can be translated to medical images without much difficulty. One of the most common tasks in medical imaging is semantic segmentation. Cardiovascular disease has become one of the most widespread causes of death in the world today.Among this is coronary artery disease which is in fact one of the most widespread forms.It comes into occurrence when the arteries that are carrying blood to the heart are interrupted by the fat distribution on the artery walls.This is call stenosis.Less severe stenosis in its early stages does not attract any attention since it has minor imaging manifestations. Even mild cases will otherwise result in deadly blockages in case they are not correctly diagnosed. These could be detected earlier with more sensitivity by the use of automated tools as compared to human interpretation..This possibility of early detection by using technology is what has driven the work of this research. The existing possibilities to enhance the accuracy of segmentation following the spread of popularity of deep learning with the use of models like U-Net is quite large .These models have the ability to automatically infer the characteristics based on the data and thus, it is appropriate to apply them to the problem of coronary artery segmentation that is challenging because the anatomy of patients may vary significantly.Measuring the arterial stenosis is important to ensure that the patients with coronary disease are classified properly and receive appropriate treatment without any adverse outcomes related to them. The fact that the illness becomes more and more widespread proves the necessity in obtaining higher quality diagnostic methods . As radiologists perform manual segmentation of the pictures ,It requires much time and may be contaminated with human error.This poor efficiency has brought the need to have automatic procedures that would not only be accurate but also leave the healthcare workers with less to do since they will have to face reduced workload. Coronary artery rotation is one such procedure that has been simplified using Computed Tomography Angiography (CTA) that results in clear 3D images. The advancement exposes a possibility of applying deep learning to autonomous CTA picture extraction.

1.3 Research Problem

Proper segmentation of coronary arteries forms a fundamental requirement in different applications of medical imaging like detection of stenosis,3D reconstruction, and addressing of cardiac dynamics. Quantitative evaluation performed on real angiography images shows that the proposed segmentation method identifies about 80% of the total coronary artery tree in relatively easy images and 70% in challenging cases with a mean precision of 82% and outperforms others segmentation methods in terms of sensitivity. The research problem that is presented in this paper is based on the issues and the constraints of segmenting of the coronary arteries during the medical imaging, in specific Computed Tomography Angiography (CTA). Now, manual tagging of the coronary arteries by the radiologists is not only time consuming but it is also prone to human error. Such discrepancy may contribute to a false evaluation of arterial narrowing, which is crucial in the identification of coronary artery disease and the development of a treatment plan .The coronary arteries are anatomically complex and exhibit high degrees of variation which presents a challenge to the process of segmentation. Segmentation is not an easy one to get good results in varying patients because factors like the depth of the arteries in heart muscle, fat around it, and different anatomy of the heart makes it complicated to get a good outcome.The available datasets, although publicly available in large quantities, are not always of high quality and may not be very suitable to train at and validate segmentation models. The current datasets usually have few number of picture images and this limits the progress and image deep learning having effective training to provide ample generalization of unseen pictures.Most of the research on the topic does not employ uniform measurement of the evaluations and it becomes hard to compare the performance of various segmentation methods. This is not standardized and may give a faulty conclusion on effectiveness of different methods.Since there is a huge growth in the amount of medical imaging data very rapidly there is a required need to have automatic solutions which can process and analyse the available data effectively. Based on manually designed features, the traditional machine learning approaches can rarely include complex patterns in the images of coronary arteries.

1.4 Research Objective

The main purpose of this study is to develop and apply an automated deep learning pipeline that could effectively segment coronary arteries in 3D CTA scans and properly measure the stenosis of arteries. Our current research project is carefully designed to overcome the limitations posed by manual segmentation that would feature its subjective nature, huge time-consumption, and vulnerability to misinterpretation by the human being, through adhering to the cascaded ensemble structure of advanced forms of U-Net architecture. The aim is to increase the precision of segmentation using a two-staged approach performing crude localization of the coronary arteries and then proceed to do patch-level segmentation on multiple deep learning models for accurate stenosis quantification.

1.5 Research Contribution

The present research has several important contributions to the coronary artery segments issue, especially, using innovative methods of deep learning.This research aims to develop a cascaded-ensemble deep learning framework that can mitigate the error that occurs during the interpretation of CTA images conducted by humans.This framework includes volumetric segmentation of the datasets using coarse segmentation which is followed by patch -based segmentation ,All these steps provides the possibility of great accuracy and efficiency of coronary artery segmentation over conventional approaches.With an ensemble of four various patch-level architectures such as 3D U-Net, U-Net++, 3D ResU-Net and Attention U-Net, the paper achieves better visualisation of fine vessels in detail. The method is associated with a multi-model solution providing thorough examination of the coronary tree as compared to single-model theories. The study establishes a new procedure of stenosis evaluation on skeletonized centerlines and then compares them to truth with reference to a prediction. It uses the distance-transform radii to calculate the percentage stenosis by a more accurate measures based on the narrowing of the arteries

Chapter 2

Related Work

This section has well reviewed previous literature related to the coronary artery segmentation. This discussion was mainly concentrated on two different types of approaches that had been dealt with in the field, which is the traditional machine learning (ML) based methods and deep learning (DL) based methods. The intention was to report an overview of the developed methodologies and their characteristics. The course of the review considered the emergence and use of these various methods sequentially in the background of coronary artery segmentation. This involved the examination of how the conventional approaches had remained dependent on hand-designed characteristics and sector-specific priorizations in segmenting. Subsequently, the section explored deep learning based approaches which had became center of intense interests because they could handle complex data without explicit field engineering. Diverse methods in both of the types were discussed with references to their role in the sphere of coronary artery segmentation.

2.1 Traditional ML-based approach

Traditional machine learning methods for coronary artery segmentation often rely on hand-crafted features and domain-specific priors. For example, this paper used geometric and image features calculated from expert annotations to guide the segmentation[8]. Similarly, this other paper proposed a level set approach that uses smooth processing to improve segmentation accuracy[10] .
Building on this, this research combined the two approaches (level sets and 3D implicit model of vessel structures) to obtain more precise results[12]. In this paper, the centerline extracted from the coronary vessels were used as the basis of the segmentation[14]. Also this paper a three-step pipeline for segmentation was introduced, although they were not developed with details in this context [9]. An intensive ray-casting technique was applied along with the learning-based boundary detector for better detection of the contour of the lumen in [16]. Model-guided segmentation using Markov Random Field formulation was discussed in this paper[17]. Furthermore, vessel segmentation was also identified as an iterative tracking problem, and a Bayesian tracking algorithm using particle filtering was proposed in this research to capture the coronary artery structures effectively[25].
Further, this paper reported using flow simulation to analyze fluid behavior within the vascular system, giving insight into segmentation of a PVE on the coronary tree[27]. Also, this research used an active search method to reconnect seemingly

Work	Year	Data quantity	Compared Traditional ML methods		Compared Deep Learning methods		Average Dice score	Code available	Data available
			General	Specific	General	Specific			
[8]	2011	54(40)	[42]	NA	NA	NA	NA	No	No
[10]	2012	24¶	NA	NA	NA	NA	70–73	No	[13]
[12]	2012	42(18)	NA	NA	NA	NA	68–72	No	[13]
[9]	2012	42(18)	NA	NA	NA	NA	66	No	[13]
[41]	2013	48(18)	NA	NA	NA	NA	65	No	[13]
[16]	2014	48(18)	NA	[10] [[12]] [[14]]	NA	NA	72–74	No	[13]
[17]	2014	48(18)	NA	[10] [16]] [14] [12]]	NA	NA	75–77	No	[13]
[19]	2015	10(6)	NA	NA	NA	NA	84	No	No
[25]	2016	61(10)	[77] [15]	NA	NA	NA	86.2	No	No
[24]	2016	32(8)	NA	[70]	NA	NA	< 84.3	No	[31]
[27]	2017	48(18)	NA	[10] [16]	NA	NA	69–74	No	[13]
[41]	2017	50¶	NA	NA	NA	NA	93–95*	No	No
[53]	2021	100¶	NA	NA	[29] [43]	NA	82	No	No

Table 2.1: Comparison of traditional ML vs. deep-learning methods on coronary artery segmentation.[67]

disconnected vessel segments [24]. Another research applied the Hough transform in the extraction of the aorta and location of the coronary root to allow greater accuracy in the initial step of segmentation[41].

2.2 DL-Based Approach

In these present days, deep learning has created such amazing opportunities in various fields that it has already grabbed huge attention. As this method is already vastly popular, there are various related communities where deep learning is making its way up to the top. Currently, it resides there in about roughly five technique trends such as direct segmentation [52], [48], [54], patch based segmentation, [39], [40], [33], [42], [56], [59], pixel based segmentation , [26], tree data based [51], [34], [35],[55], [60], [39], [58], [66], segmentation [50], graph data-based segmentation [45]. Among the researchers, direct segmentation has become quite popular as it does not require many preprocessing steps to get the desired output results. Another reason for this segmentation method to be popular among researchers is due to the rise of the U-Net model . This approach is used more frequently for segmented imaging models as it is not time-consuming due to having fewer steps. Another paper uses a method where the process starts with a pre-shaped stencil that is similar to any coronary artery structure, and after getting the selected patient's information, this pre-shaped stencil gets aligned according to the actual artery. This uses a spatial transformer network(STN) that is used during the whole process while it's navigating for the actual alignment. Another work shows that two different types of approaches have been adopted to track the artery's path & to monitor whether the tracked path is the diseased path or not using a tracker & discriminator. Provide a combined deep learning and conventional level set approach framework. This works as a hybrid model that improves the segmentation accuracy. The model used in the deep learning method eventually extracts and uses the global features of the provided image, so it compares the results and finds out the local details, which is essential to get more robust segmentation results [48]. This method is even useful when the provided image might be of low quality. As in the first process, the model captures the essence of 2D CNN that makes the process faster, while 3D CNN are more vast and take more time and space. But this strategy raises the bar of getting more accurate results, which is more efficient in the field of diagnosis. One method proposed a different method by using U-net, which uses attention mechanisms. These mechanisms focus on the most relevant area that makes this process faster, which makes it easier to identify between the regions that are very similar in nature[55]. For a precise detection of artery boundaries, a model is used by creating a U-shaped network that uses a spatiotemporal fusion approach. Basically this approach uses multiple slices of the same image and identifies the different features using a different sizing system within the same given image[66]. This research used a unique approach for segmentation of the arteries by using a 2D UNet model that is fully automated[52]. This paper provided a method where the model they used uses their learning ability and distinguishes between the differences of pulmonary vessels from arteries[47]. This helps the segmentation to detect the required area more smoothly. This method uses Mask R-CNN. Even though direct segmentation has done a tremendous job in the field of deep learning methods, it has some boundaries as well. For instance, to use direct segmentation on an image, it uses some sort of low resolution, which can give false results in some cases.

Work	Year	Data quantity	Traditional ML General[1]	Traditional ML Specific[2]	Deep Learning General	Deep Learning Specific	Avg. Dice score	Code available	Data available
[26]	2016	10(6)	NA	NA	NA	NA	$< 65^{\dagger}$	No	No
	2017	42(18)	NA	NA	NA	NA	78	No	[13]
[40]	2018	50(40)	[20]	NA	NA	NA	80	No	No
[30]	2018	44(35)	NA	NA	NA	NA	86	No	No
[33]	2018	52(45)	NA	[10] [14]	NA	[26]	83†	No	[13]
[43]	2019	70(50)	NA	NA	[47]	NA	90†	No	No
[42]	2019	50(40)	NA	NA	NA	NA	92	No	No
[45]	2019	42(18)	NA	[17] [9]	NA	NA	73–75	No	[13]
	2019	548(274)	NA	NA	[8]	NA	85	No	No
[46]	2019	8(7)	[50]	NA	NA	NA	$< 99^{\dagger}$	No	[54]
[47]	2020	25(20)	NA	NA	NA	NA	90	No	No
[51]	2020	48(18)	NA	NA	NA	NA	83	No	[13]
[51]	2020	70(50)	NA	NA	[22]	NA	91	No	No
[49]	2020	916(733)	NA	NA	[11]	NA	85	No	No
[59]	2021	48(18)	NA	NA	NA	[26] [39]	91	No	[13]
[54]	2021	59(34)	NA	NA	[50]	NA	87	No	No
[65]	2021	30(24)	NA	NA	[8]	NA	88	Yes	No
[55]	2021	300(200)	NA	NA	[22]	NA	89	No	No
[58]	2021	70(50)	NA	NA	[22]	NA	94†	No	No
[56]	2021	70(50)	NA	NA	[22]	NA	94	No	No
[52]	2021	69(44)	NA	NA	[93]	NA	89	No	No

Table 2.2: Work Details Comparison (traditional ML vs. deep learning). [67]

Also, in that chase, patch-based segmentation is more preferable where the images are too big to process all at once, rather than direct segmentation. Patch-based segmentation focuses on the region of interest only, so it is faster and does not require much time or memory to process. This paper classified its voxels to determine which are the actual part of the blood vessels by using a multi-scale 3D CNN. One work showed the transformation of CTA images by dividing them into small patch-based segments to get more precise results faster[33]. This paper provided a method where the model gets a vesselness map that helps it to learn and identify the shapes that are the arteries [39]. By using a convolutional neural network (CNN) with a recurrent neural network, it is able to identify the provided slices and differentiate between arteries and other parts of the given image[42]. This paper introduced a new system where the 3D images are easily monitored by the Dense U-net model that uses various functions in order to find the actual pixels that are required for the identification of the vessels[56]. This helps to eliminate all the unnecessary pixels and lets the model focus only on the required parts. Another important part of the deep learning method is pixel-based segmentation. One of the very first methods that uses CNN to segment the coronary arteries. In pixel-based segmentation, this system was a huge stepping stone for the deep learning system as it was the first time a segmentation method was classifying each pixel of the provided image.

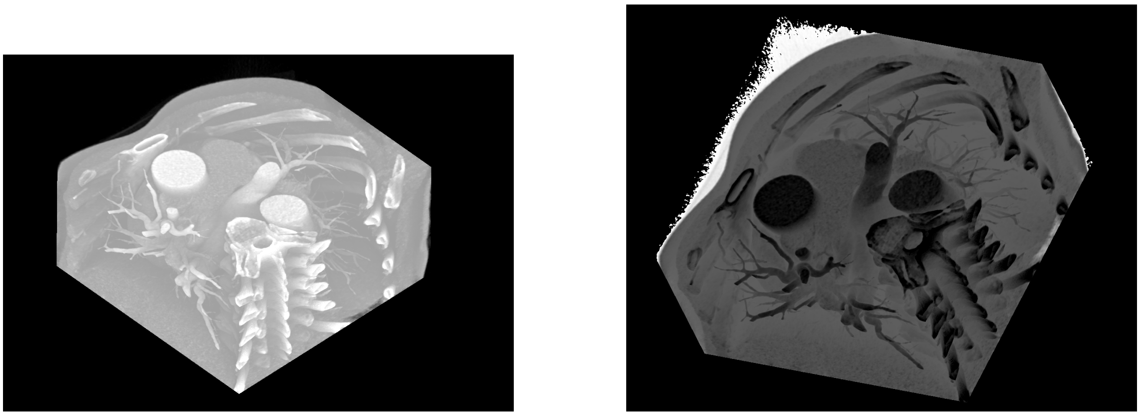
In conclusion, these reviews further suggested that such processes were prone to errors and were time consuming in their manual nature. Several ways of scanning with manual features were presented, but then deep learning which promised to automatically outline and identify coronary arteries without manual feature construction was brought up. In particular, it discussed the U-Net architecture and its derivative models as very useful in segmenting coronary arteries of 3D medical scans. Several trends in deep learning were discussed, such as direct segmentation, patch-based segmentation, pixel-based segmentation, and the graph data scale segmentation. Lastly, the chapter noted that challenges were observed in the current research, including the absence of a consistent fair comparison of different models. It further adds that other studies had varied datasets used or labeled differently which made the comparison not reliable. It was also referred as one of the causes to these issues by the limited availability of large, publicly available, datasets to be used in segmentation of coronary.

Chapter 3

Proposed Methodology

3.1 Dataset Description

For this research purpose, ImageCAS (A-Large-Scale-Dataset-and-Benchmark-for-Coronary-Artery-Segmentation-based-on-CT) dataset was used [67].



(a) (b)

Figure 3.1: A 3D CTA Volume

The ImageCAS dataset contained 1000 3D CTA images, shown in Figure, of patients who were diagnosed with coronary artery disease previously. The original dimension of the image is 512*512*(206-275) voxels with a planar resolution and spacing of 0.29-0.43 mm square and 0.25-0.45 mm consequently [67].

3.2 Experimental Setup

The experiment of this research was extremely thought out to guarantee the best functioning of deep learning models applied in supervised segmentation of the coronary arteries.The complete pipeline has been adopted in terms of PyTorch framework that is most commonly known to be flexible and efficient in context of deep learning applications.The PyTorch framework was implemented for the whole pipeline on an NVIDIA RTX 4090 GPU with 24GB VRAM. As our input size was fixed, for faster runtime, cuDNN benchmark was used [70]. Each of the neural networks in the pipeline got a total of 40 epochs of training. The learning rate was fixed to

0.002 and this is a frequently used option that is a trade- off concept between fast convergence and a stable training process. This setting was necessary in order to make sure that the models could learn on the data without the danger of overfitting it.

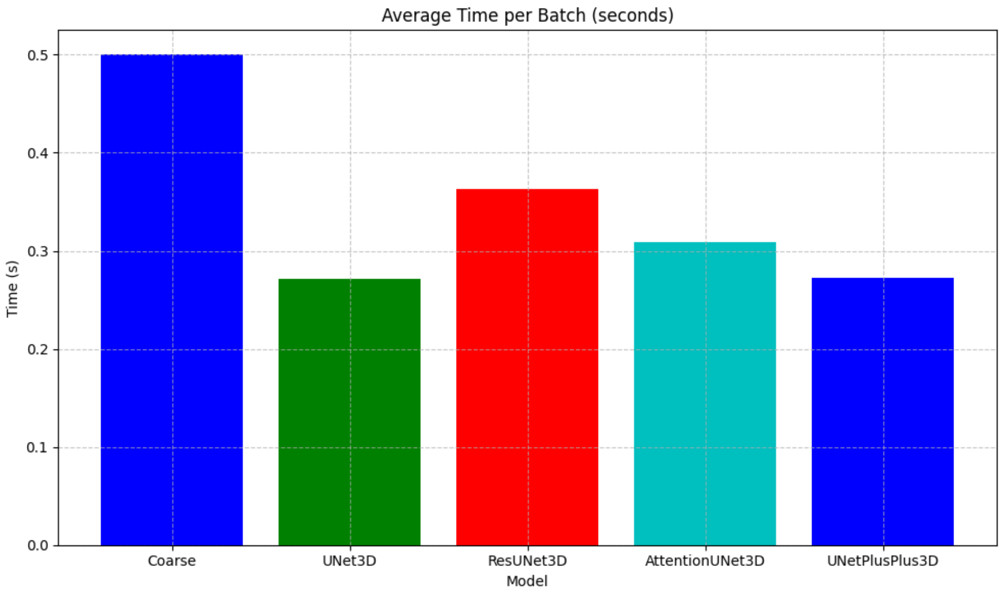


Figure 3.2: Average Time Per Batch (seconds)

During training, the batch sizes were smaller with coarse and patch based networks because of the shortage of CUDA memory on GPU. The smaller batch was 3 in the case of the coarse-based network. This was done to accommodate the higher volumes of input which are processed by the coarse network. Conversely, the patch-based networks used a larger batch size of 60 which made training using smaller localized sections of the data more efficient. This strategic alternation of the batch sizes played a very essential role in well-utilization of available GPU resources without losing a well working training dynamics (3.2). For running the coarse based model, every epoch took 8-10 minutes and for patch based model, each epoch took 3-6 minutes to perform.

3.3 Preprocessing

Preprocessing was done individually for images and masks with various steps (3.3). Every steps are being described below :

3.3.1 Resizing

To support the implementation of a similar and efficient processing of the 3D CTA images, it was important that all volumetric scans were adjusted to be resized into a standardized spatial resolution before the actual processing was performed. Such normalization guarantees that all images, regardless of the size of acquisition, and parameters, will be able to be entered into the deep learning models with standard size of input, simplified the design of networks architecture and decreased variability

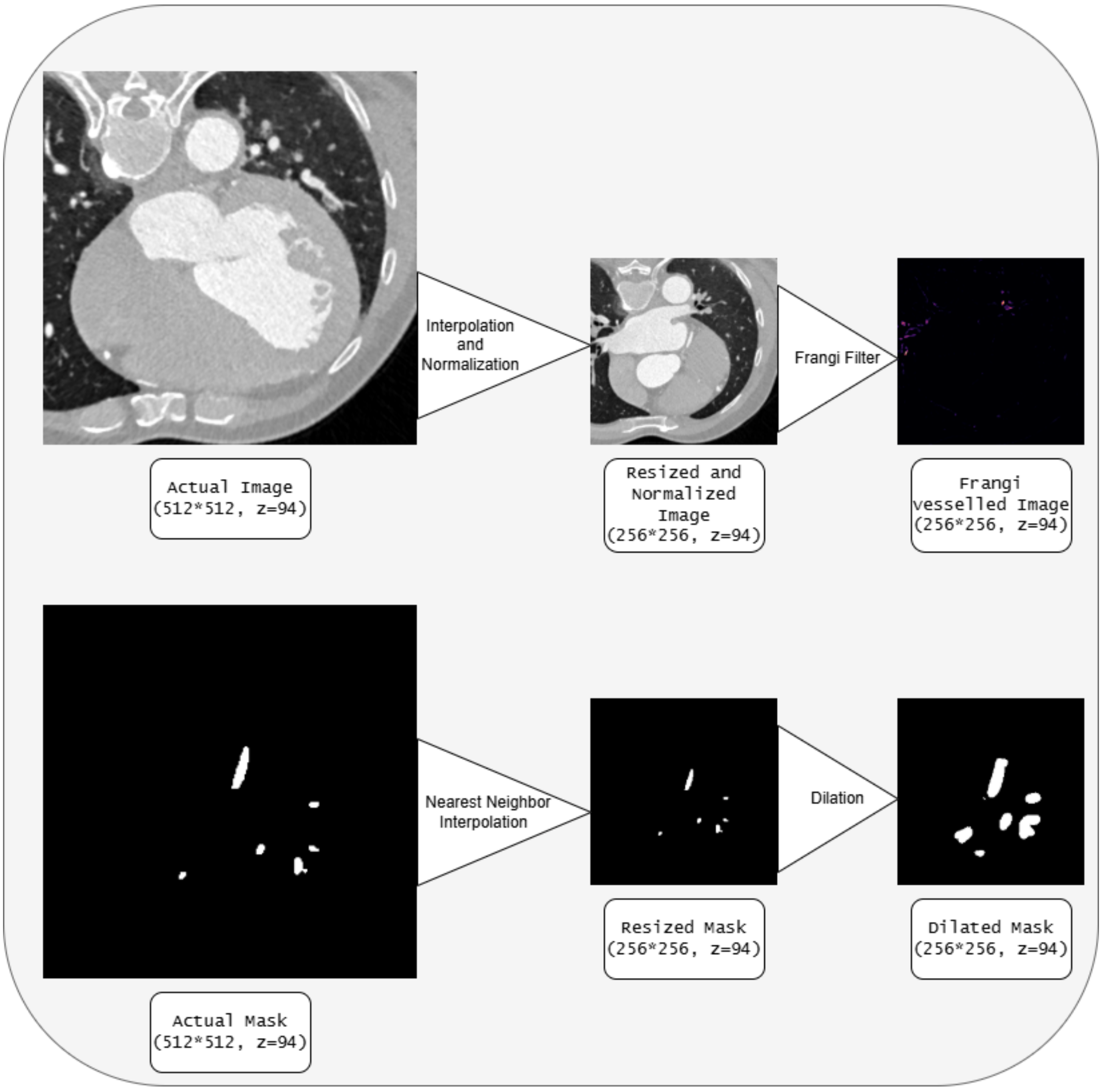


Figure 3.3: Preprocessing Workflow

in computation. In order to do so, the interpolation method called linear interpolation was used to resize the images: new voxel values were recalculated to occupy a new, evenly spaced grid. In particular, the whole image was subsampled to rectangular dimension of 256*256*128 voxels. This will entail interpolating the original image data along the axes to as close as the desired resolution, and will result in either stretching or compressing the image in such a way that does not necessarily change the structural integrity of the image and anatomical features. Notably, the adopted approach was selected to preserve as much information of the original image as possible, but at the same time result in minimal to none artifacts or distortions that may distract successful segmentation. After embracing linear interpolation, the process balances between effectiveness and fidelity in the sense that the resampled images capture the spatial relationships as they were in the original data. Further with the respective binary masks as ground truth the nearest neighbor resampling over the whole field was adopted so as to maintain voxel-wise label precision to avoid induction of non-specified answers or interpolated answers. This diligent resampling allows maintaining the alignment between image data and their labels at the highest level without losing the trustworthiness of training and evaluation of the model.

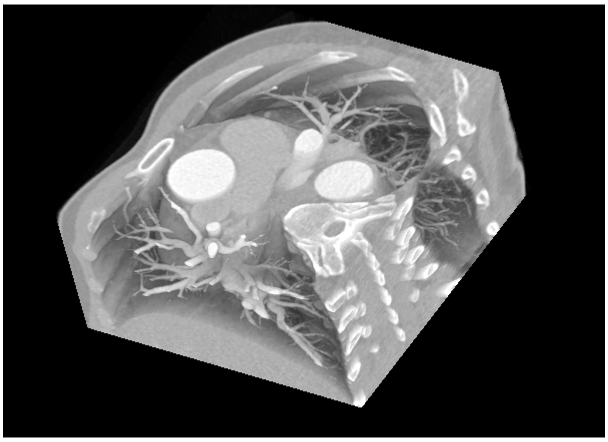


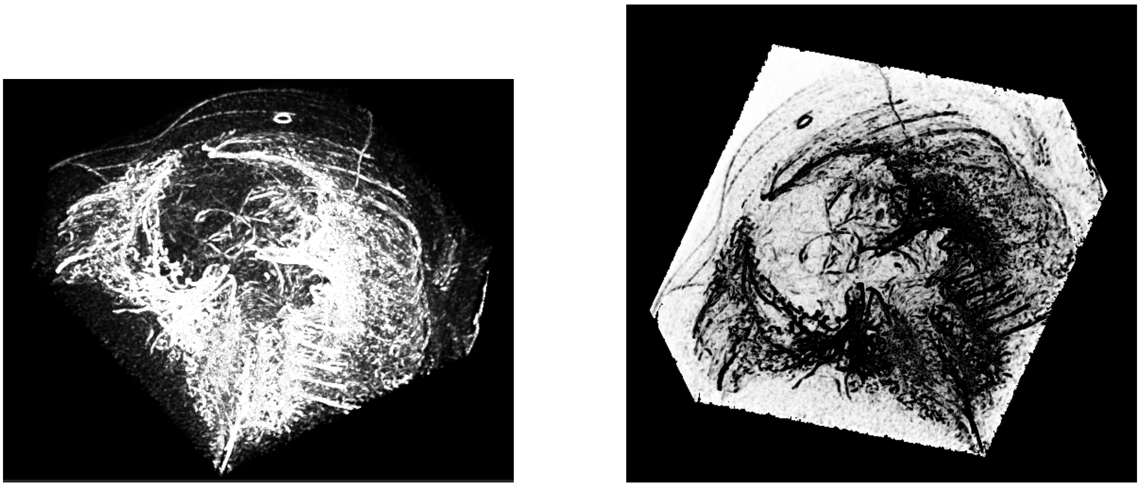
Figure 3.4: Resized and Normalized Image

3.3.2 Intensity Normalization

Once the resampling of the images to a consistent spatial resolution was performed, the process of standardizing the intensity over the images was the next important step of the preprocessing, which is central to the consistent model performance throughout the dataset. These Computed Tomography Angiography (CTA) images have diverse and broad admission of the Hounsfield Unit (HU) values and are prone to variations because of factors like the types of scanners, the acquisition techniques, the specific tissue of the patient. Such variability may be problematic to deep learning algorithms that are highly sensitive to the scale and distribution of the input features. In order to reduce this effect, each 3D image had the individual intensity values clipped into the 1st and 99th percentiles of the corresponding HU value distributions. It is a technique of examining the allocation of voxel intensities in every picture and ascertaining the threshold intensities at the percentiles, which is a strategy of eliminating the extreme outliers and the skew values that might be developed by the noise, artifacts or non-tissue compounds. Reducing the range to the central 98 percent of the data makes it possible that the most interest tissue structures, and in particular the vessels and the building tissues and vessels around the muscles need to be adequately contrasted to be emphasized without being interfered with by the anomalous high intensities of the data and anomalous low intensities of the data. Then, within this clipped range, the intensities were scaled using methods commonly referred to as normalization to come within a standardized range (0 to 1).This invariant scaling enables neural networks to learn better as they concentrate on what is really important in the images, the common attributes among them, and are not left confused by varying values of absolute intensities. Moreover, such normalization increased the strength of the model implemented in this research pipeline, making it less prone to inter-scanner and inter-patient differences, which is also essential when implementing the model in various clinical scenarios. (3.4).

3.3.3 Vessel Enhancement

The patch based models were made to perform their calculations more on the substantially important contents of the image, that is the sections where the coronary arteries could be found by extracting smaller, defined ROIs. ROI extraction is an important procedure in the proposed framework that helps to narrow down to the area of interest in further studies. It was done following a rough estimate of vessel areas which was obtained based on a rough approximated segmenting model. The dilated mask (obtained during the earlier preprocessing stages) as well as original mask independently undergo this ROI selection procedure, also. It made sure that the ground truth and other masks were also subjected to the pre-processing and fit the obtained ROI that could be uniformly used during the training and assessing of the patch-based models. ROI extraction helped the patch-based models to focus on the areas where the coronary arteries could be mostly seen. Such specificity allowed such models to have improved accuracy in performing the diagnosis and description of arterial stenosis In consequence, a multi-scale Frangi vesselness map was computed on the normalized CTA for extracting tubular structures accurately.



(a) (b)

Figure 3.5: Frangi Vesselled 3D Volume

This 3D Frangi filter emphasized keeping the vessel's anatomical structure more highlighted by suppressing the background (3.5).

3.3.4 Dilation

To resolve the issue of coarse or incomplete boundaries, which was frequently produced during mask interpolation, an operation of dilation was enacted as a means of pre-processing. In particular, the segmentation masks were dilated with spherical voxel radius of 5 (so-called morphological dilation). This dilation is the process of increasing the size of the borders, that is, the segmented vessel structures by including the voxels around the existing vessel boundaries in a prescribed radius, in this pipeline five voxels and at the same time thickening the vessel boundaries in the process. The main idea behind using dilation in such case is to produce a soft border zone around the vessels' masks that would act as a compromise to resolve the problem of misalignment or tiny boundaries or minor inaccuracies in the resulting segmentation. With such a dilation, the masks now tolerate small misrepresentations and boundary variations to better and facilitate an easier transition between the vessel and tissue-margins. The presence of such a soft border zone helped to minimize the effects of the segmentation errors in the area of the borderlines affecting later stages in this pipeline, such as skeletonization, stenosis quantification. It also made sure that the most central areas of the vessel were well represented both in the training and during inferences, and gave some form of padding that added presence of the vessel in the input data fed to the model. Moreover, dilation was more conducive to overall robustness in owning the downstream processing so that the model could cope with minor differences, or artifacts, that could otherwise compromise accuracy of boundaries. It led to the more stable feature extraction and increased reliability of the segmentation results in this study. The five-voxel radius had a fair trade-off between adequate boundary smoothing and detailed vessel morphology. This dilation step is a general step of the pipeline in order to enhance the overall quality and accuracy of vessel segmentation, in particular, to address boundary uncertainties due to the presence of many imaging or preprocessing artifacts.

(a) (b)

Figure 3.6: Dilated Mask vs Original Mask

In this step of preprocessing, dilation with a radius of 5 (5 voxels) was applied on the masks to create a soft border zone to mitigate the unaligned border.

3.3.5 Dataset Splitting

As the dataset of this research was too large, cross-validation during model training was too time costly, even for 3-fold CV. To mitigate this situation, the dataset was split by a stratification logic. The dataset was categorized into two parts: vessel-rich and vessel-sparse, by calculating the percentage of vessels an image contained and their mean.

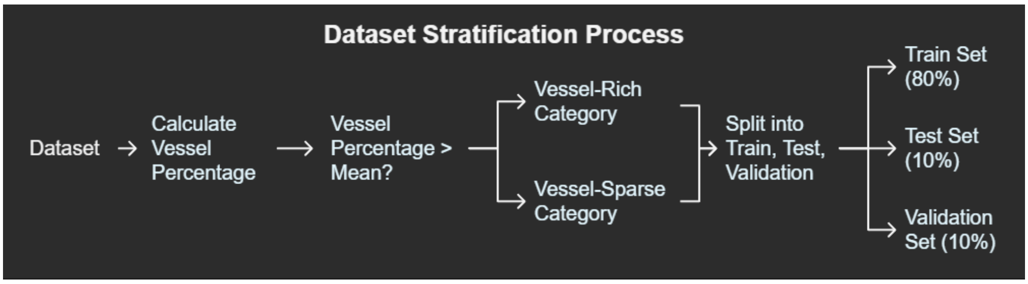


Figure 3.7: Data Splitting Pipeline

After that, the train, test, and validation sets of data got consecutively 80%, 10%, and 10% from those 2 categories where the count of vessel percentages was higher than the mean and equal to or lower than the mean equally.
After all the preprocessing (3.3), the images were input in 2 channels, where 1 channel contained the normalized images and the other the Frangi-vesselled image. In the case of masks, the dilated masks and the original masks were both input separately into the next steps of the pipeline.

3.4 Model descriptions

3.4.1 UNet++

UNet++ employs re-designed skip pathways composed of nested dense convolutional blocks that bridge the semantic gap between encoder and decoder feature maps, thereby facilitating more accurate medical image segmentation [37]."U-net++ is also known as the nested U-net model. To improve the accuracy of image segmentation, the more elevated version of U-Net known as U-Net++ was introduced .There are five architectural steps that work in the process of U-Net++, such as Encoder path, Decoder path, Nested skip pathways, Deep supervision, and final segmentation.

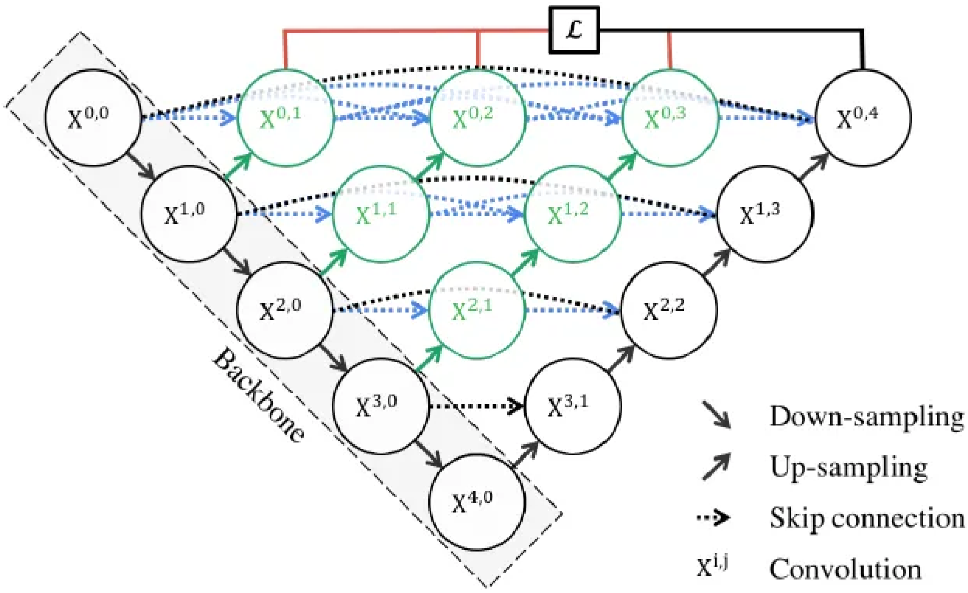


Figure 3.8: Unet++ Architecture [44]

Basically, this model starts with an encoder that is used to extract the features from the given image and then prepares it for the nested skip connection pathways. UNet++ modifies these connections into nested dense convolutional blocks. After that decoder path combines the results of two previous steps so that the output segmentation is reconstructed. For the final part, the segmentation gives the output of refined features. UNet++ was announced in an effort to improve the accuracy of the segmentation task by eliminating the disparities between the encoder and decoder features, as well as providing the model with more information at various levels.[37]

3.4.2 ResUNet

The ResUNet is a deep learning model with a network architecture that integrates the U-Net structure and residual learning to perform efficiently and robustly with image segmentation tasks, especially in biomedical and medical imaging.The limitations faced by the traditional U-Nets, such as gradients vanishing and deficient feature reuse, are addressed in ResUNet. ResUnet follows a similar encoder-decoder structure as U-Net architecture.ResUNet shown performing better than other models in various segmentation tasks, specifically in biomedical imaging, for example, liver and tumor segmentation from CT scans, retinal vessel segmentation, and brain tumor detection.[62]

3.4.3 Attention based 3D U-Net

The Attention based 3D U-Net is a smarter version of the regular 3D U-Net. It is specially made to analyze medical scans like coronary CT angiography (CTA). These scans create 3D images of the heart. They use many thin slices in the process. Thus, using a 3D model helps capture the full structure clearly. One limitation of the regular U-Net is that it looks at every part of the image equally. Even the parts that don't help with diagnosis are observed. The attention mechanism improves this by allowing the model to focus more on the important regions. Such as- areas with blocked or narrowed arteries. The model is made of three main parts. They are- the encoder, the decoder and the skip connections. Skip connections now include attention gates. The encoder works like a feature extractor. It takes the 3D image input. Then passes it through multiple blocks of 3D convolutional layers, 14 ReLU activation functions and batch normalization. Each block is followed by a downsampling layer (3D max-pooling). It reduces the image size while increasing the depth of features. These layers are used to detect textures or patterns that are possibly vital in predicting diseases. The decoder does the opposite. It starts with the smallest and most detailed features. Then slowly upsamples them using 3D transposed convolutions. Each upsampling step is followed by more 3D convolutions and activations. These help sharpen the final segmented image. At every stage, the decoder is helped by the encoder through skip connections. Here's where the attention gates come in. In a regular U-Net, skip connections pass all the encoder features to the decoder. But in this improved model, attention gates act like smart filters between the encoder and decoder. They look at both the encoder features (x_l) and a signal from the decoder (g). Then calculate which regions are worth focusing on. The gate produces an attention map (α) using this equation:

$$\alpha = \sigma \left(W_0 \cdot \mathrm{ReLU}(W_x x_l + W_g g + b) \right) \tag{3.1}$$

The filtered output $\hat{x}_l$ is:

$$\hat{x}_l = \alpha \cdot x_l \tag{3.2}$$

This means that only the important features from the encoder are allowed to pass to the decoder. It helps the model ignore irrelevant background details. Then focus on areas like plaques or calcifications. Before sending features to the next step, both the encoder part (x) and decoder part (g) are made smaller using tiny 1×1×1 filters. Then some basic clean-up (called instance normalization) is being done. If they are not the same size, the decoder part is resized. It is done to match the encoder part. Then they are added together. This goes through a ReLU (a kind of switch). Then another small filter and more cleaning. Thus finally, this goes through a special function called sigmoid. This gives a map that tells the model what to focus on. The values in this map go from 0 to 1. The map is used to multiply with the encoder features. This keeps the important parts and removes the extra noise. It helps the model see only what matters. Because this model is fully 3D, it can understand the shape and location clearly. Especially the heart structures across all slices. This makes its segmentation output more accurate and realistic compared to 2D models.

Studies show that using attention in this way helps improve detection accuracy. It also reduces errors in segmenting complex areas like blood vessels.

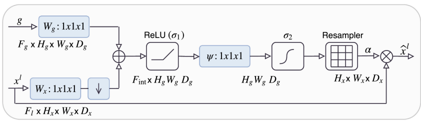


Figure 3.9: Attention Gate Functionality

The model uses MaxPool3D to shrink images when going down. This makes the image smaller. Then uses upsampling to make it bigger again. At the end, it uses a tiny filter (1×1×1) to give the final result. This model looks at small parts of the image known as regions of interest (ROIs). But the model does not look at the whole image at once. It picks important areas like- blood vessels. This helps it work better and find even the tiny vessel parts.
By combining smart attention with full 3D awareness, this model gives doctors much clearer views of the heart and arteries. This can lead to earlier and more confident diagnoses of cardiovascular disease. And this is critical for effective treatment planning.

3.4.4 3D Unet

The 3D CNN(convolutional neural network) architecture which is primarily designed for 3D(volumetric) medical image segmentation. It generally extends the known 2D U-net by integrating 3D convolutions, allowing it to process 3D medical imaging data such as MRI,microscopy volumes and CT scan. The main difference between the 2D and 3D U-Net is in the implementation of operations: the former engages in 2D operations, whereas the latter uses their 3D counterparts. In the encoder path, features are captured by 3D convolution layers, followed by 3D max-pooling applied for downsampling.

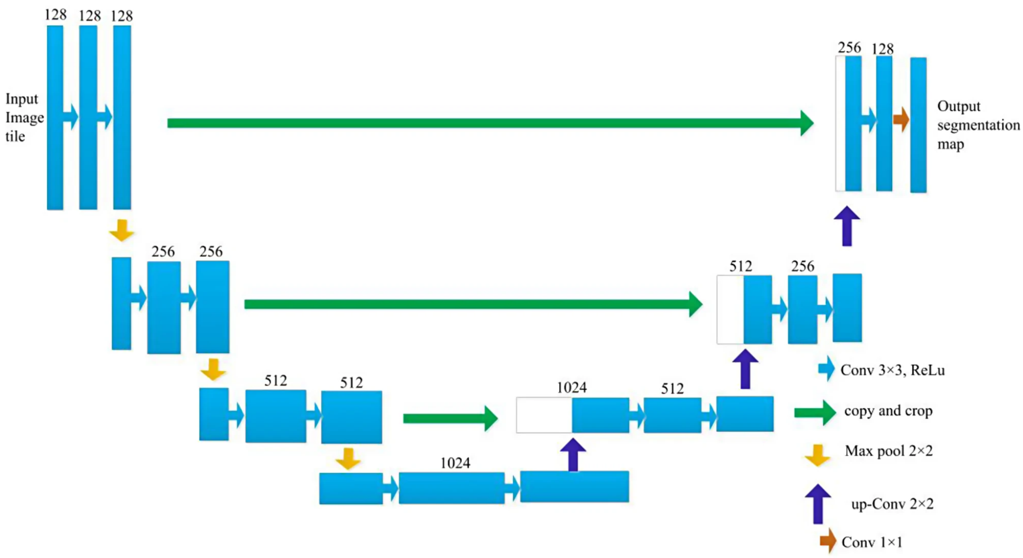


Figure 3.10: 3D Unet Pipeline

From the deeper part of the network abstract representations of inputs are extracted. In the decoder path, the 3D up-convolutions or transposed convolutions are applied for the upsampling of the feature maps. The skip connections connect each layer in the encoder with the corresponding layer in the decoder. These connections are important when recovering fine-level spatial information that is lost during down-sampling.[34] states that 3D-UNet comprises contracting and expanding pathways that intend to build up the bottleneck part in its very center through convolution and pooling operations. After this bottleneck, somehow, the image is reconstructed from a mixture of convolutions and upsampling. Skip connections are added to take care of supporting gradient backward flow for better training.

3.5 Proposed pipeline for Training and Testing

A cascaded pipeline was proposed for model training with ensemble learning (3.11). On the proposed pipeline, first of all, the image dataset and the dialated mask went through coarse based segmentation (3.5.1). After getting the rough prediction mask output from this coarse based segmentation, ROI was selected from the original images and masks based on that output (3.5.2). These ROIs went through four different patch based models as a fine tune factor and output accurate anatomically structured prediction masks (3.5.3). Then the original masks were skeletonized and from that centerline, the arteries radii of each mask were calculated and with the 80th percentile, a reference radius was selected for further comparison. At the same time, the radius of predicted masks from patch-based model output was calculated in the same process. Then radii calculated from the output predicted masks were than compared with the reference radius to decide whether there is stenosis present or not. If stenosis remains present, stenosis percentage was calculated in the process (3.5.4).

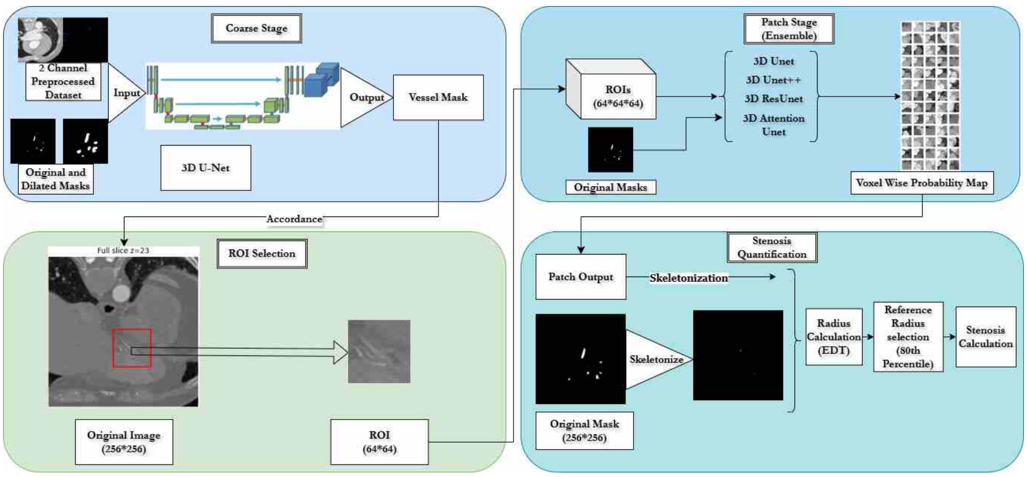


Figure 3.11: Proposed Pipeline

3.5.1 Coarse-Based Segmentation

As per coarse-based segmentation, 3D-Unet, a single neural network, processed the whole 3D image. The development of the coarse-based segmentation model was designed to attain success in the localization of the coronary tree within three-dimensional medical imaging. As our research goal was to find out the volumetric

area of the arteries, the convolution layers of this model assisted in capturing that successfully [22]. This model took our two-channel preprocessed volume as input, and as output, it provided a rough vessel mask, which helped to locate the Region of Interest(ROI) for later usage in the pipeline. This type of activity is a critical step since it increases the quality of the information feeding in the network enabling the model to learn the anatomical features contained in the images at a higher level. The two channels were of different characteristics extracted out of the original images thus contributing to better performance of the model during training.
A smoothed dice loss was used to train this network :

$$\text{DiceLoss} = 1 - \frac{2\sum_i p_i m_i + \varepsilon}{\sum_i p_i + \sum_i m_i + \varepsilon}, \quad p_i = \sigma(\ell_i),\ m_i \in \{0,1\},\ \varepsilon = 1 \times 10^{-5}. \tag{3.3}$$

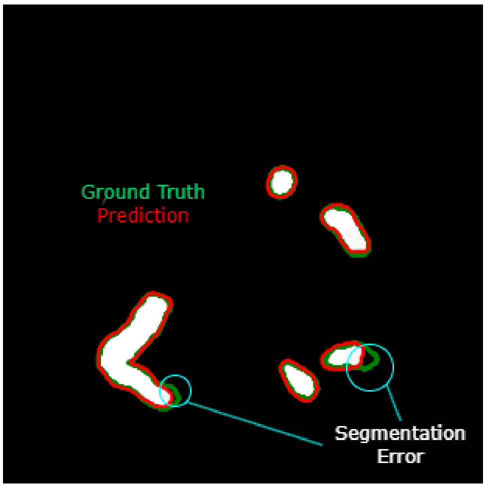
where $\mathbf{i}$ index of voxels, $\ell_{\mathbf{i}}$ predicted probability of model, $\mathbf{p_i}$ and $\mathbf{m_i}$ are predicted probability at voxel i and mask value of voxel i, ε is the smoothing constant to remove division by zero error.
Then a positive-class weight was computed from all dilated mask voxels for addressing class imbalance. Formula for positive-class weight was :

$$\text{pos\_weight} = \min\Bigl(\frac{N_{\text{vox}} - N_{\text{pos}}}{N_{\text{pos}}},\ \text{MAX\_POS\_WEIGHT}\Bigr). \tag{3.4}$$

where $\mathbf{N}_{\text{vox}}$ and $\mathbf{N}_{\text{pos}}$ referred to total number of voxels and positive number of voxels across the dilated masks.
then binary cross entropy loss was calculated :

$$\text{BCEWithLogits}(\ell, m; w) = -\frac{1}{N}\sum_{i=1}^{N}\Bigl[\, w\, m_i\ \log\bigl(\sigma(\ell_i)\bigr)\ +\ (1 - m_i)\ \log\bigl(1 - \sigma(\ell_i)\bigr)\Bigr], \tag{3.5}$$

where ℓ referred to the logit of coarse network, $\sigma(\ell_{\mathbf{i}})$ referred to sigmoid of predicted probability, $\mathbf{m_i}$ is the GT mask label, $\mathbf{w}$ is positive class weight and $\mathbf{N}$ is total number of voxels
After that, the coarse loss was calculated using these adding these 2 losses. These losses assisted the model to perform better [28]. Although this model was able to capture large and mid-scale anatomical context, it could not focus on the boundary properly, as shown in the figure (4.7).



(a)

(b)

Figure 3.12: Coarse Prediction

3.5.2 Region Of Interest Extraction

After getting the approximate prediction of vessel regions (4.7) from the coarse model, a smaller volume of dimension consisting of 64*64*64 voxels is selected as the region of interest(ROI) to run patch-based models (3.11). The center of ROI was selected by calculating the mean of the coordinates of vessel voxels predicted by the coarse model. The dilated mask and the original mask also went through the ROI selection process. The patch based models were made to perform their calculations more on the substantially important contents of the image, that is the sections where the coronary arteries could be found by extracting smaller, defined ROIs. ROI extraction was an important procedure in the proposed framework that helped to narrow down to the area of interest in further implementations.
It was done following a rough estimate of vessel areas which was obtained based on a rough approximated segmenting model. The dilated mask obtained during the earlier preprocessing stage as well as original mask independently underwent this ROI selection procedure, also. It made sure that the ground truth and other masks were also subjected to the pre-processing and fit the obtained ROI that could be uniformly used during the training and assessing of the patch-based models. ROI extraction helped the patch-based models to focus on the areas where the coronary arteries could be mostly seen. Such specificity allowed models to have improved accuracy in performing the diagnosis and description of arterial stenosis.

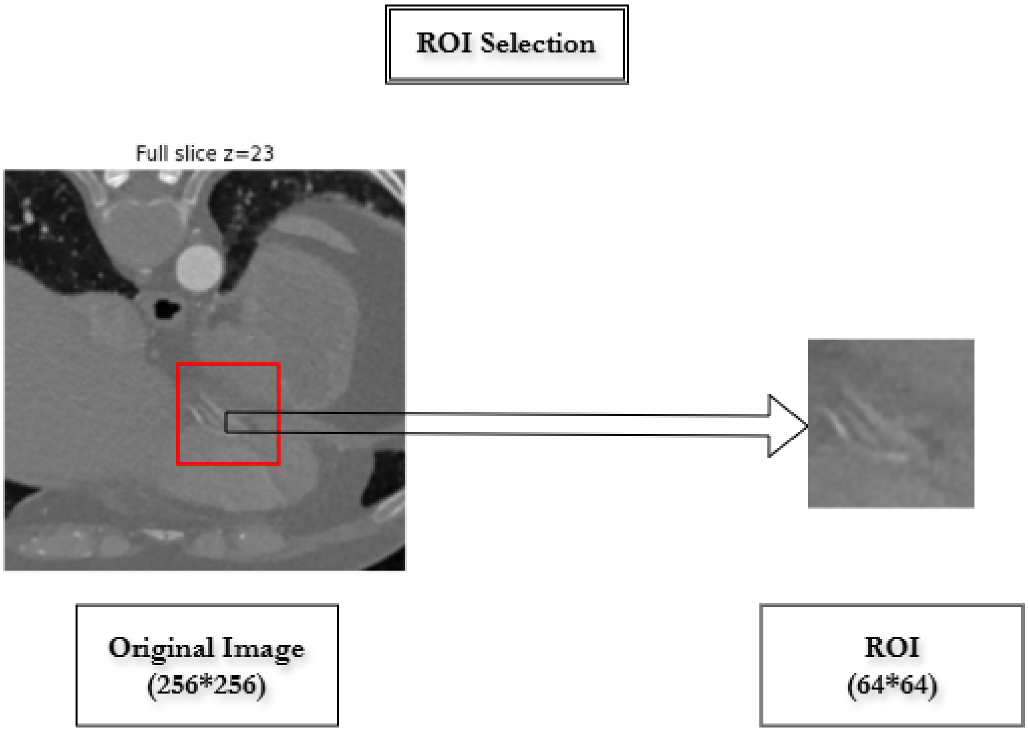


Figure 3.13: ROI selection

3.5.3 Patch-Based Segmentation

On the ROI extracted volumes, four distinct architectures: 3D Unet, 3D ResUnet, 3D AttentionUnet, and 3D Unet++ were trained.
While training the patch based model, the ROI volume and the original mask with ROI clipped were taken as input. Training on the selected patches allowed these models to focus on the boundary with utmost effectiveness. Besides, some models are more deeper and complex like Attention Unet, Unet++; so it was hard to run them coarsely as that would be time sensitive, therefore, they were applied on patches so that without processing full image but only an important portion of image, they could provide richer outcomes. Patch level segmentations also used the same loss functions as coarse level one did (3.3), (3.5). This patch-based segmentation technique of coronary vessels analysis were conducted in a procedural manner where four different deep learning architectures, such as 3D U-Net, U-Net ++, Attention-Based 3D U-Net and ResU-Net, were used.The models were distinctive in their contribution to the segmentation process, and obtained some fine details of the regions of interest (ROIs).To achieve a good segmentation framework through the improvement of the accuracy in the detection and characterization of arterial stenosis,the working principle and architecture of the models were contributory. This procedure was important because it enabled the models to target the vital regions of the images including the images of vessels, improving the accuracy of vessel detail capture to a large extent. While training the patch based model, the ROI volume and the original mask with ROI clipped were taken as input and later those were run by models such as UNet++ , ResUNet, Attention based 3D U-Net and 3D UNet. By focusing on the smaller parts that were relevant to the entire image, the models can be trained to detect complex structures, which are critical towards the segmentation process.Training on the selected patches allowed

these models to focus on the boundary with utmost effectiveness.The batch size of the patch-based networks was initialized to 60. The increase in this batch size was favorable because it allowed quicker training over the smaller patches. Applying many patches in parallel did not only increase the overall training speed but also helped the model to learn a rich variety of examples on each training step, which was a crucial aspect of generalization.The results of the coarse-based segmentation were used as a starting point to the patch-based refinement .
After the patch stage's completion, there were some voxel probability overlapping because of neighboring patch overlap which was solved by calculating mean of their probabilities; which consequently provided refined predicted mask as patch output.

3.5.4 Stenosis quantification

For translating the output got from the patch level models. At first, the original mask got straight from preprocessing, was skeletonized and was selected centerline of the artery. Then from this centerline voxels, the distance of the nearest background was calculated which referred to the radius of the vessel. After that, with 80th percentile calculation for all centerline radii, a reference lumen size was decided. At the same time, this radii calculation process was applied on the output mask from patch models too. After that the percentage of stenosis was calculated using this formula [1]:

$$\text{Percent Stenosis} = \left(1 - \frac{r_{\text{lesion}}}{r_{\text{ref}}}\right) \times 100\% \tag{3.6}$$

where 'lesion' refers to the narrowest part of the predicted mask's vessel, 'ref' refers to the reference radius found out after calculation on original mask. After these processes, the metrics was calculated on the model performances and stenosis calculation (3.6).

Chapter 4

Results and Discussions

The cascaded ensemble pipeline study demonstrated robust performance across the whole pipeline and also yielded viable results clinically. The selected models chosen for this pipeline were evaluated on different kinds of metrics, such as Dice score, precision-recall, F1-score, AUC score, and stenosis percentage. The Dice Score, or the Dice Similarity Coefficient (DSC), is usually a range in spatial indices, segmentation, and binary classification problems. It indicates the level of similarity between predicted and actual positive sets and is thus an excellent metric for judging the overlap between classifications. In the context of image segmentation, for instance, the Dice Score can be utilized to evaluate the degree of similarity between the predicted segmentation mask and the ground truth segmentation mask. [69]

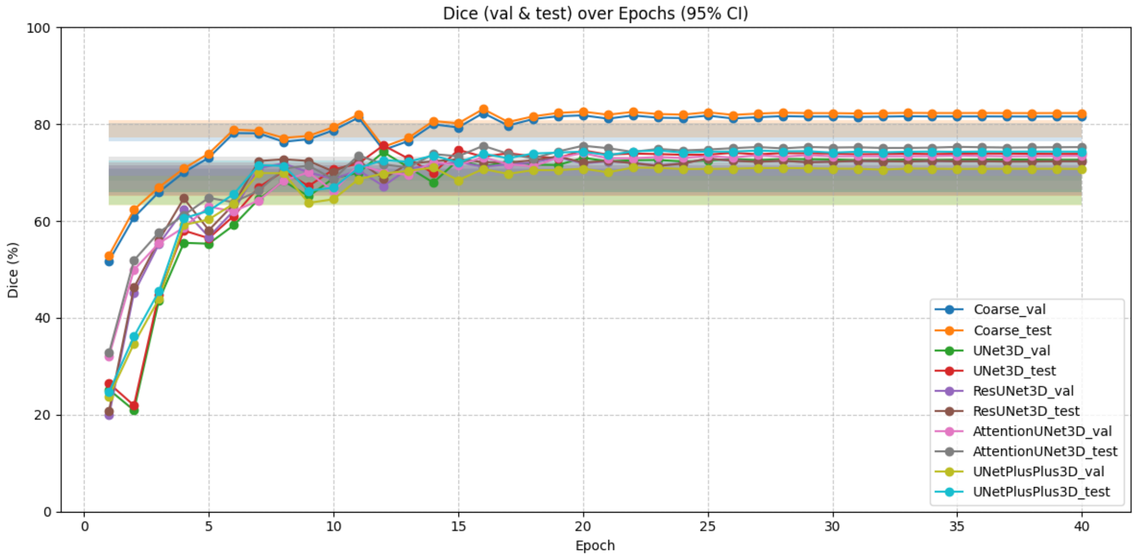


Figure 4.1: Dice Score for all the networks

According to Dice score, the coarse-based model showed the highest performance with a dice score of 82% (4.1). Although the coarse-based model performed well in terms of Dice, it still cannot segment the arteries with utmost precision, as the mask used in this case was dilated to facilitate the initial localization of vessels. Besides, among the patch level models, Attention Unet 3D and Unet++3D models showed best convincing performances with a dice score of 75.3% and 74.3% cosecutively on the test data set (4.1). The remaining patch-based models, Eg, ResUnet3D and

Unet 3D, lacked in terms of this metric with dice score of 73% and 74%. These metrics showed better results than previous research using 2D Unet.
For this research, precision and recall combination is crucial, as the indirect goal was to differentiate between voxel classes. Precision tells us how many of the people the model says have heart disease do. In other words, it checks how correct the model's positive predictions are. Precision = TP/(TP+FP) For example, if the model says 10 people are sick but only 7 truly are, then the precision is 70%. Models like UNet need to be careful not to over-predict disease. High precision means the model is good at avoiding these mistakes.

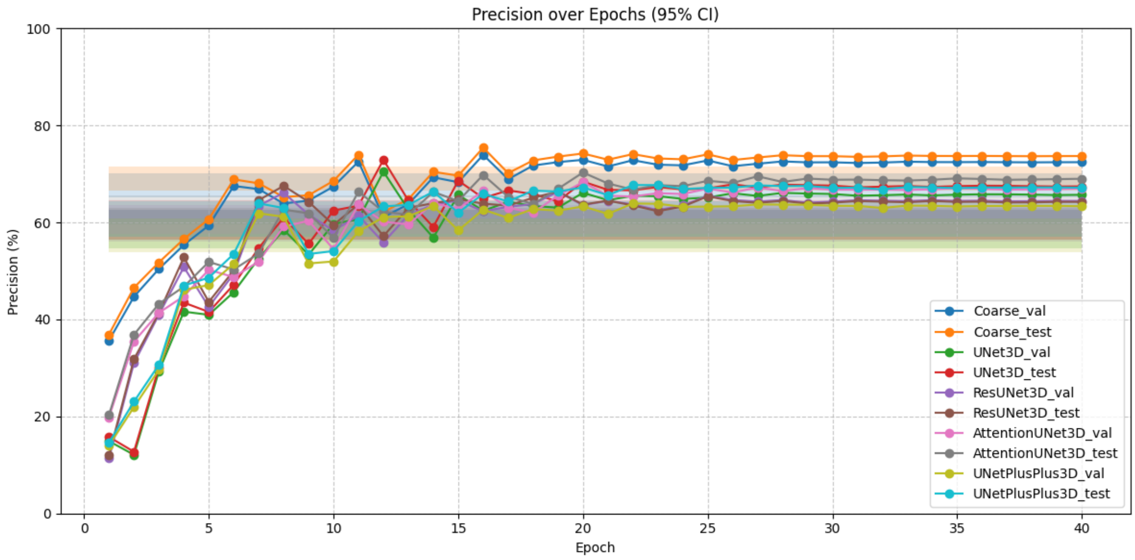


Figure 4.2: Precision for all the networks

This becomes even more important when we have fewer sick cases than healthy ones in the dataset. In that case, precision helps check if the model is still reliable. Most of the dataset voxels were background, thus the precision graph (4.2) was upwards and the recall graph (4.3) was downwards. For precision, a patch-based model performed better with a peak value of 70% (4.2) from Attention Unet 3D, as well as its recall value was 85.6% (4.3). This was the lowest distance between precision and recall among all other models. Among other models, Unet++3D also showed somewhat consistency with precision and recall value of 69% and 86.5% concurrently (4.1). The other models e.g. ResUnet3D and Unet3D lacked behind as the precision and recall distance was a bit far (4.1).
When analyzing coronary computed tomography (CTA) images to predict heart disease, recall is a vital measure. It tells us how well our model spots actual problem areas. Like diseased parts of the heart's arteries, by calculating: Recall= TP/(TP+FN) In simple terms, high recall means we're catching most of the trouble spots. Thus, such as- plaques or narrowed arteries. Which is crucial to avoid missing something that could lead to serious health issues [18].
The coarse-based Unet 3D model resulted in better recall than precision, which shows that the result was biased towards positive predictions. Mostly, the patch-based models showed compelling outcomes as their precision and recall values were closer than the coarse-based model, as shown in the (4.2),(4.3).
F1 score is an alternative machine learning assessment metric used to measure a

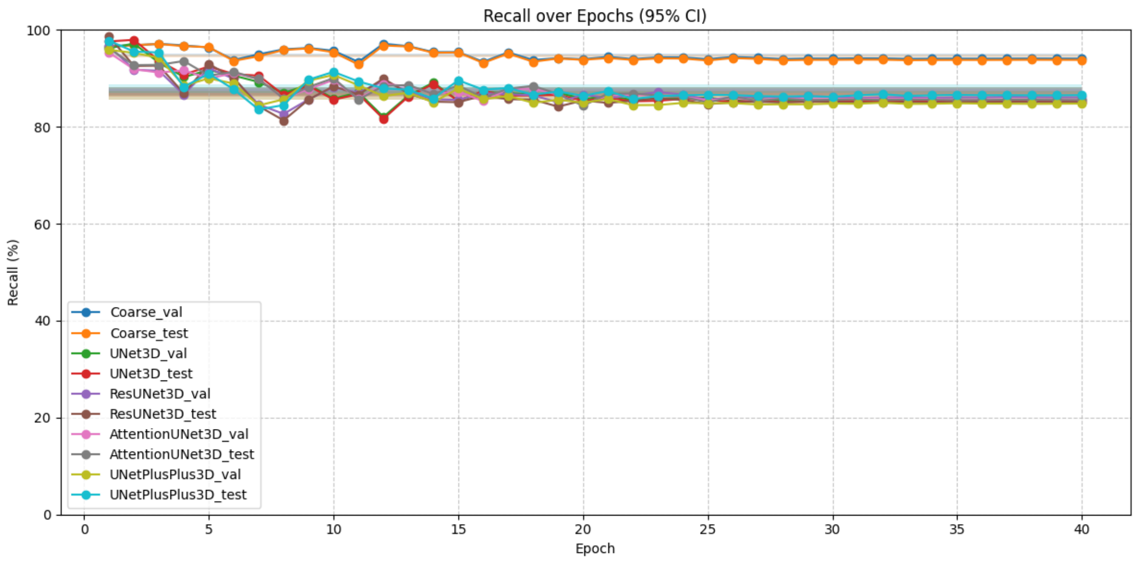


Figure 4.3: Recall for all the networks

model's skill in prediction by specifying the class-wise performance of that predictive model instead of giving an overall performance index like accuracy. F1 score combines two antagonistic measurements, being the precision and the other being the recall score of a model. For this reason the Attention Unet 3D model

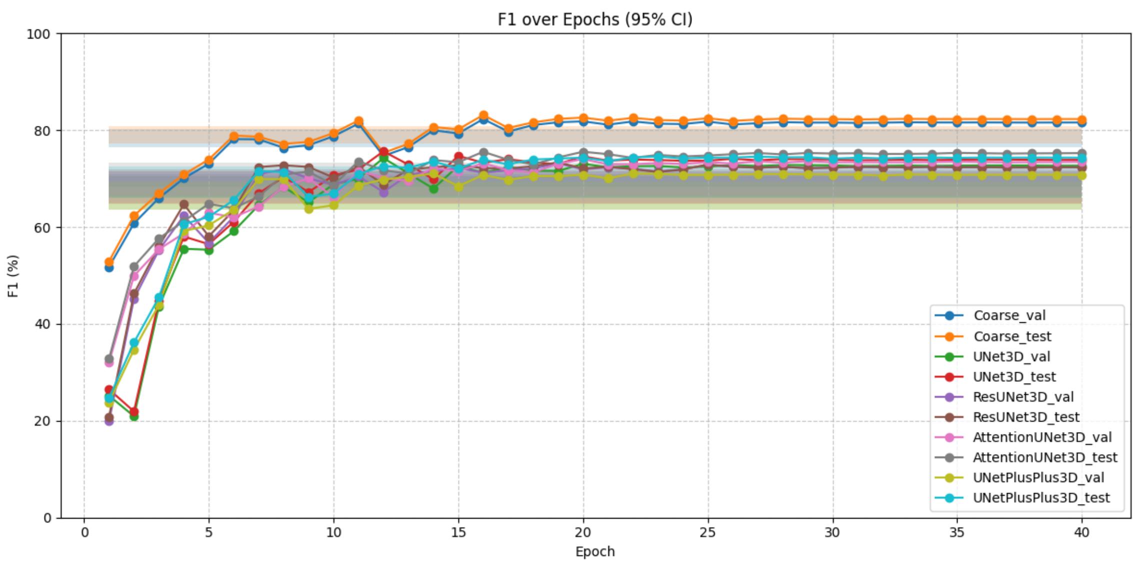


Figure 4.4: F1 score for all the networks

showed best promising performance with an F1 score of 75.3%(4.4). Among the other patch-based models, Unet++3D resulted better F1 with a score of 74.3%. The coarse-based 3D Unet model exceeds this result with a F1 score of 83%, but it was biased in terms of recall (4.4).
The AUC score for all the models was between 98% and 99%, which means that all of the models were able to differentiate between voxel types almost perfectly (4.5).
According to the box plot of stenosis percentage error across all the models (4.6), the error between the ground truth and the predicted stenosis percentage for all

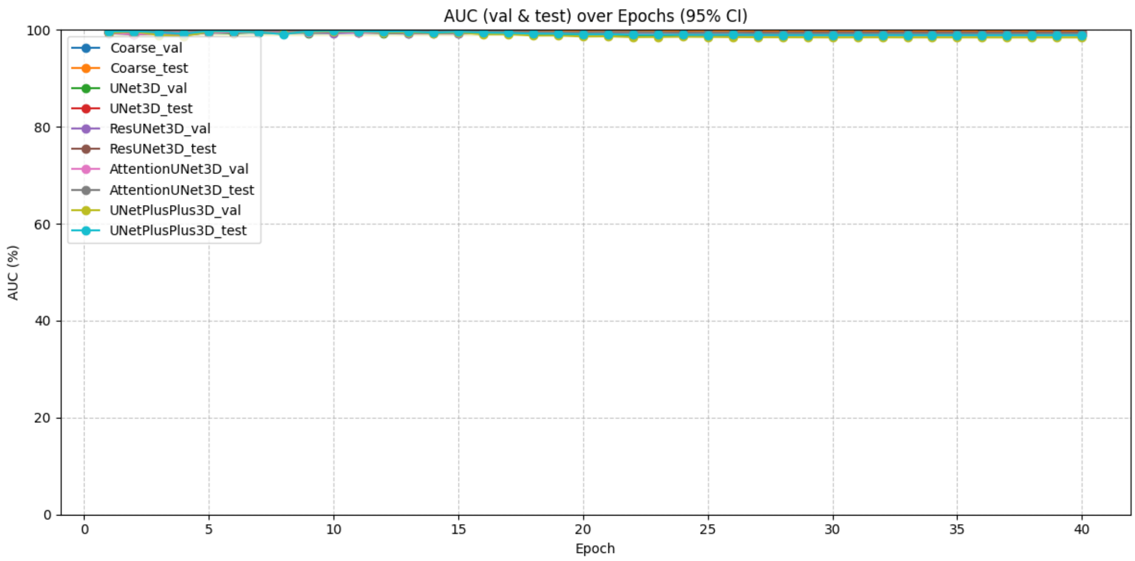
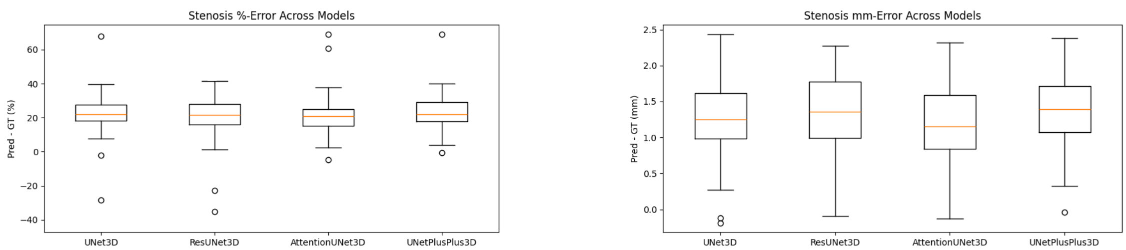


Figure 4.5: AUC score for all the networks

the models were around -20% to 20% (4.6a) with some outliers. The predictions are almost centered near ground truth, and their median error is close to 0%, which suggests that the models were quite accurate in their estimation. As Attention Unet 3D has the smallest IQR,lowest median error and insignificant amount of outliers, this model has the best consistent performance.

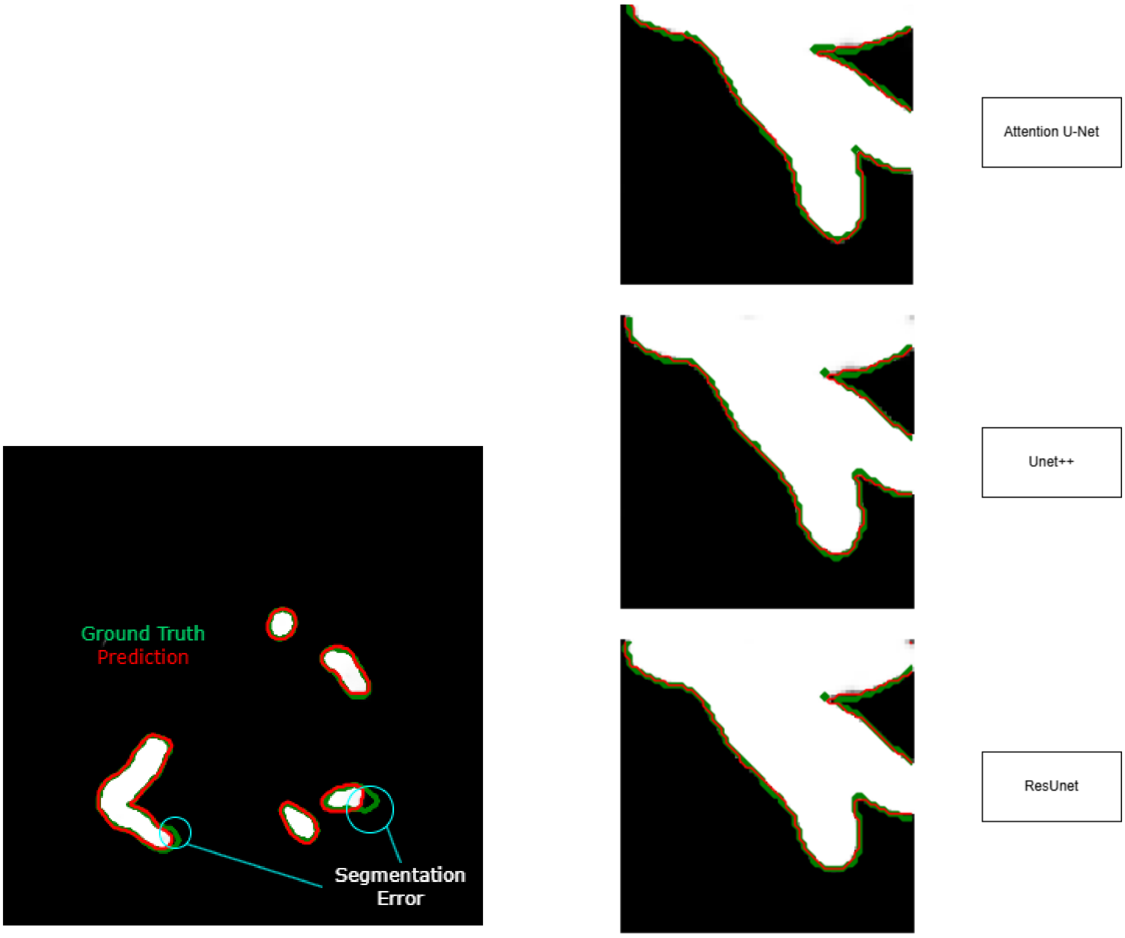


(a) Stenosis Percentage Predictions for all models

(b) Stenosis mm Predictions for all models

Figure 4.6: Stenosis Prediction

The ResUnet shows the widest IQR, which indicates the variability among the predictions at a large scale in its predictions. The same results go for the stenosis percentage graph (4.6b), where also Attention Unet 3D showed best performance. Among all the patch-based models, **Attention Unet 3D** showed the best and clinically consistent results across all metrics (4.1) and stenosis calculation, which is being considered as the best model for this unique pipeline. Despite this factor, according to the metrics, all of the models of this pipeline resulted in high recall values,with an average value of 85%, which indicates that the vessel detection had strong sensitivity. But, the precision value was a bit inconsistent in this case, having 70% on average, which indicates that the model is having challenges while reducing false positive predictions. Summing up, the coarse model's prediction had some error which was addressed and solved by patch-based models, especially Attention Unet 3D (4.7), (4.1).



(a) Stenosis Percentage Predictions for all models

(b) Stenosis mm Predictions for all models

Figure 4.7: Cross Sectional View for Coarse vs Patch Prediction errors

Table 4.1: Performance Comparison of Best Models Based on Recall, Dice, Precision, AUC, and F1 *[Best Epoch = 40]*

Models	**Recall (%)**	**Dice (%)**	**Precision (%)**	**AUC (%)**	**F1 (%)**
UNet++3D	86.5	74.3	69	98.9	74.3
AttentionUNet3D	85.6	75.3	70	99.2	75.3
ResUNet3D	86.1	73	64.4	99.4	72.3
Unet3D	85.7	74	68	99.1	74

Chapter 5

Future Improvement and Conclusion

5.1 Limitations and Future Improvements

As a coarse based model only 3D Unet was run. Due to computational complexity and prolonged runtime, the other dense models like attention Unet, ResUnet and Unet++ 3D were not run coarsely. Besides, A cascaded multitask learning pipeline could be designed for better segmentation. Moreover, The 3D image volumes were not evenly sliced, for which different criteria was used to form a global dimension among all the volumes.But this arises anatomical breach, which is insignificant but still a drawback. The models could be run more than 40 epochs to get more generalized metrics, but for time complexity made it harder to run for more epochs. Besides, the dataset should be cross validated which was also tough to accomplish as there was time limit for this experiment.

5.2 Real-Life Implementation

This achievement of the cascade ensemble U-Net framework on the segmentation of coronary arteries has shown great capabilities to be implemented in the clinical area, and it can be seen as an applicative instrument in the early diagnosis and risk prediction of coronary artery disease. In order to enable the system in real-life settings, the introduction of this system into clinical workflows should be smooth. In this, it would include the development of user friendly graphical user interfaces (GUIs) that would be easily handled by clinicians that have little technical know-how to facilitate easy upload of imaging data and display of the output results including detailed segmentation maps and stenosis dimensions. Moreover, it is also important to integrate with the hospital information systems (HIS) and picture archiving and communication systems (PACS) to facilitate an automated process of data retrieval, processing, and reporting with reduced manual assistance and error possibilities. In a real-life scenario, the system should be super-fast and efficient to deliver near-real time results, which are crucial in emergency-like and routine diagnostics scenarios. This includes the use of optimal computational algorithms and even the use of hardware acceleration mechanisms in the form of GPU processing, that would allow completing the analysis fast without affecting the accuracy. Besides, the system needs to be deeply validated on the data of various patients, with varied demographic backgrounds, varieties of imaging machines, and disease conditions, to certify its strength and universality. This will make it easier to determine possible biases or restrictions, with a view of improving as time goes by. Moreover, since such information as patient history, biomarkers and any other clinical information regarding diagnosis is significant, its addition to the model can in turn add to the overall clinically representative predictive ability of the model and overall clinical decision-making. There is also a possibility of future developments to accommodate cloud-based deployment to enable remote consultation and access to expert analysis, particularly in resource-scarce environments. Overall, successfully implementing this framework could revolutionize cardiovascular diagnostics by enabling precise, automated, and rapid assessments of coronary artery stenosis. This would not only improve diagnostic accuracy and reduce inter-observer variability but also support timely treatment interventions, ultimately leading to better patient outcomes. Continued research, collaboration with healthcare professionals, and technological advancements will be essential to translate this promising research into a widely adopted clinical tool that can be validated in diverse healthcare environments.

5.3 Conclusion

Quantifying stenosis in coronary arteries plays a critical role in the diagnosis and management of cardiovascular diseases. In this study, we present a cascaded ensemble U-Net based framework. It is designed to accurately estimate the severity of arterial blockages from medical imaging data. The cascaded structure allows successive U-Net models to progressively refine segmentation outputs. This is while the ensemble strategy integrates their predictions to enhance overall robustness and precision. Evaluations conducted on a comprehensive dataset demonstrated that the proposed method outperformed existing approaches. This achieved a Dice similarity score of 83%. Although these results are promising, there is still potential for improvement through further tuning and optimization of the pipeline. This approach offers a valuable tool for clinical decision-making, with the potential to support timely and accurate diagnosis. Future work will aim to incorporate relevant clinical variables, explore real-time application, and validate performance across diverse populations and imaging modalities to ensure broader clinical applicability.

Bibliography

[1] K. L. Gould, K. Lipscomb, and G. W. Hamilton, “Physiologic basis for assessing critical coronary stenosis: Instantaneous flow response and regional distribution during coronary hyperemia as measures of coronary flow reserve”, *The American journal of cardiology*, vol. 33, no. 1, pp. 87–94, 1974.

[2] R. Cooper, J. Cutler, P. Desvigne-Nickens, S. P. Fortmann, L. Friedman, R. Havlik, G. Hogelin, J. Marler, P. McGovern, G. Morosco, *et al.*, “Trends and disparities in coronary heart disease, stroke, and other cardiovascular diseases in the united states: Findings of the national conference on cardiovascular disease prevention”, *Circulation*, vol. 102, no. 25, pp. 3137–3147, 2000.

[3] S. Doyle, A. Madabhushi, M. Feldman, and J. Tomaszeweski, “A boosting cascade for automated detection of prostate cancer from digitized histology”, in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006: 9th International Conference, Copenhagen, Denmark, October 1-6, 2006. Proceedings, Part II 9*, Springer, 2006, pp. 504–511.

[4] A. Tabesh, M. Teverovskiy, H.-Y. Pang, V. P. Kumar, D. Verbel, A. Kotsianti, and O. Saidi, “Multifeature prostate cancer diagnosis and gleason grading of histological images”, *IEEE transactions on medical imaging*, vol. 26, no. 10, pp. 1366–1378, 2007.

[5] C. Gunduz-Demir, M. Kandemir, A. B. Tosun, and C. Sokmensuer, “Automatic segmentation of colon glands using object-graphs”, *Medical image analysis*, vol. 14, no. 1, pp. 1–12, 2010.

[6] G. M. Schuetz, N. M. Zacharopoulou, P. Schlattmann, and M. Dewey, “Meta-analysis: Noninvasive coronary angiography using computed tomography versus magnetic resonance imaging”, *Annals of internal medicine*, vol. 152, no. 3, pp. 167–177, 2010.

[7] D. P. Zhang, “Some article title”, *Some Journal*, vol. 10, pp. 100–110, 2010.

[8] Y. Zheng, M. Loziczonek, B. Georgescu, S. K. Zhou, F. Vega-Higuera, and D. Comaniciu, “Machine learning based vesselness measurement for coronary artery segmentation in cardiac ct volumes”, in *Medical Imaging 2011: Image Processing*, Spie, vol. 7962, 2011, pp. 489–500.

[9] A. Broersen, P. Kitslaar, M. Frenay, and J. Dijkstra, “Frenchcoast: Fast, robust extraction for the nice challenge on coronary artery segmentation of the tree”, in *Proc. of MICCAI Workshop” 3D Cardiovascular Imaging: a MICCAI segmentation Challenge*, 2012.

[10] B. Mohr, S. Masood, and C. Plakas, “Accurate lumen segmentation and stenosis detection and quantification in coronary cta”, in *Proceedings of 3D Cardiovascular Imaging: a MICCAI segmentation challenge workshop*, 2012.

[11] K. Nguyen, A. Sarkar, and A. K. Jain, “Structure and context in prostatic gland segmentation and classification”, in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2012, pp. 115–123.

[12] C. Wang, R. Moreno, and Ö. Smedby, “Vessel segmentation using implicit model-guided level sets”, in *MICCAI Workshop" 3D Cardiovascular Imaging: a MICCAI segmentation Challenge", Nice France, 1st of October 2012.*, 2012.

[13] H. Kirişli, M. Schaap, C. Metz, A. Dharampal, W. B. Meijboom, S.-L. Papadopoulou, A. Dedic, K. Nieman, M. A. de Graaf, M. Meijs, *et al.*, “Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography”, *Medical image analysis*, vol. 17, no. 8, pp. 859–876, 2013.

[14] R. Shahzad, H. Kirişli, C. Metz, H. Tang, M. Schaap, L. van Vliet, W. Niessen, and T. van Walsum, “Automatic segmentation, detection and quantification of coronary artery stenoses on cta”, *The international journal of cardiovascular imaging*, vol. 29, pp. 1847–1859, 2013.

[15] H. Fu, G. Qiu, J. Shu, and M. Ilyas, “A novel polar space random field model for the detection of glandular structures”, *IEEE transactions on medical imaging*, vol. 33, no. 3, pp. 764–776, 2014.

[16] F. Lugauer, J. Zhang, Y. Zheng, J. Hornegger, and B. M. Kelm, “Improving accuracy in coronary lumen segmentation via explicit calcium exclusion, learning-based ray detection and surface optimization”, in *Medical Imaging 2014: Image Processing*, SPIE, vol. 9034, 2014, pp. 993–1002.

[17] F. Lugauer, Y. Zheng, J. Hornegger, and B. M. Kelm, “Precise lumen segmentation in coronary computed tomography angiography”, in *Medical Computer Vision: Algorithms for Big Data: International Workshop, MCV 2014, Held in Conjunction with MICCAI 2014, Cambridge, MA, USA, September 18, 2014, Revised Selected Papers 4*, Springer, 2014, pp. 137–147.

[18] C. E. Rochitte, R. T. George, M. Y. Chen, A. Arbab-Zadeh, M. Dewey, J. M. Miller, H. Niinuma, K. Yoshioka, K. Kitagawa, S. Nakamori, *et al.*, “Computed tomography angiography and perfusion to assess coronary artery stenosis causing perfusion defects by single photon emission computed tomography: The core320 study”, *European heart journal*, vol. 35, no. 17, pp. 1120–1130, 2014.

[19] Y. Chi, W. Huang, J. Zhou, L. Zhong, S. Y. Tan, K. Y. J. Felix, L. C. S. Sheon, and R. S. Tan, “A composite of features for learning-based coronary artery segmentation on cardiac ct angiography”, in *Machine Learning in Medical Imaging: 6th International Workshop, MLMI 2015, Held in Conjunction with MICCAI 2015, Munich, Germany, October 5, 2015, Proceedings 6*, Springer, 2015, pp. 271–279.

[20] K. Sirinukunwattana, D. R. Snead, and N. M. Rajpoot, "A novel texture descriptor for detection of glandular structures in colon histology images", in *Medical Imaging 2015: Digital Pathology*, SPIE, vol. 9420, 2015, pp. 186–194.

[21] ——, "A stochastic polygons model for glandular structures in colon histology images", *IEEE transactions on medical imaging*, vol. 34, no. 11, pp. 2366–2378, 2015.

[22] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3d u-net: Learning dense volumetric segmentation from sparse annotation", in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, Springer, 2016, pp. 424–432.

[23] D. Han, H. Shim, B. Jeon, Y. Jang, Y. Hong, S. Jung, S. Ha, and H.-J. Chang, "Automatic coronary artery segmentation using active search for branches and seemingly disconnected vessel segments from coronary ct angiography", *PLoS One*, vol. 11, no. 8, e0156837, 2016.

[24] ——, "Automatic coronary artery segmentation using active search for branches and seemingly disconnected vessel segments from coronary ct angiography", *PLoS One*, vol. 11, no. 8, e0156837, 2016.

[25] D. Lesage, E. D. Angelini, G. Funka-Lea, and I. Bloch, "Adaptive particle filtering for coronary artery segmentation from 3d ct angiograms", *Computer Vision and Image Understanding*, vol. 151, pp. 29–46, 2016.

[26] P. Moeskops, J. M. Wolterink, B. H. Van Der Velden, K. G. Gilhuijs, T. Leiner, M. A. Viergever, and I. Išgum, "Deep learning for multi-task medical image segmentation in multiple modalities", in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, Springer, 2016, pp. 478–486.

[27] M. Freiman, H. Nickisch, S. Prevrhal, H. Schmitt, M. Vembar, P. Maurovich-Horvat, P. Donnelly, and L. Goshen, "Improving ccta-based lesions' hemodynamic significance assessment by accounting for partial volume modeling in automatic coronary lumen segmentation", *Medical Physics*, vol. 44, no. 3, pp. 1040–1049, 2017.

[28] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations", in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer International Publishing, 2017, pp. 240–248, ISBN: 9783319675589. DOI: 10.1007/978-3-319-67558-9_28. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-67558-9_28.

[29] L. Yu, J.-Z. Cheng, Q. Dou, X. Yang, H. Chen, J. Qin, and P.-A. Heng, "Automatic 3d cardiovascular mr segmentation with densely-connected volumetric convnets", in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part II 20*, Springer, 2017, pp. 287–295.

[30] F. Chen, Y. Li, T. Tian, F. Cao, and J. Liang, "Automatic coronary artery lumen segmentation in computed tomography angiography using paired multi-scale 3d cnn", in *Medical imaging 2018: Biomedical applications in molecular, structural, and functional imaging*, SPIE, vol. 10578, 2018, pp. 652–658.

[31] C. Collet, Y. Onuma, D. Andreini, J. Sonck, G. Pompilio, S. Mushtaq, M. La Meir, Y. Miyazaki, J. de Mey, O. Gaemperli, *et al.*, "Coronary computed tomography angiography for heart team decision-making in multivessel coronary artery disease", *European Heart Journal*, vol. 39, no. 41, pp. 3689–3698, 2018.

[32] W. Huang, L. Huang, Z. Lin, S. Huang, Y. Chi, J. Zhou, J. Zhang, R.-S. Tan, and L. Zhong, "Coronary artery segmentation by deep learning neural networks on computed tomographic coronary angiographic images", in *2018 40th Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, IEEE, 2018, pp. 608–611.

[33] ——, "Coronary artery segmentation by deep learning neural networks on computed tomographic coronary angiographic images", in *2018 40th Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, IEEE, 2018, pp. 608–611.

[34] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-denseunet: Hybrid densely connected unet for liver and tumor segmentation from ct volumes", *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.

[35] X. Li, Q. Dou, H. Chen, C.-W. Fu, X. Qi, D. L. Belavỳ, G. Armbrecht, D. Felsenberg, G. Zheng, and P.-A. Heng, "3d multi-scale fcn with random modality voxel dropout learning for intervertebral disc localization and segmentation from multi-modality mr images", *Medical image analysis*, vol. 45, pp. 41–54, 2018.

[36] Z. Li, X. Zhang, H. Müller, and S. Zhang, "Large-scale retrieval for medical image analytics: A comprehensive review", *Medical image analysis*, vol. 43, pp. 66–84, 2018.

[37] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation", in *Deep learning in medical image analysis and multimodal learning for clinical decision support: 4th international workshop, DLMIA 2018, and 8th international workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, proceedings 4*, Springer, 2018, pp. 3–11.

[38] Y.-C. Chen, Y.-C. Lin, C.-P. Wang, C.-Y. Lee, W.-J. Lee, T.-D. Wang, and C.-M. Chen, "Coronary artery segmentation in cardiac ct angiography using 3d multi-channel u-net", *arXiv preprint arXiv:1907.12246*, 2019.

[39] ——, "Coronary artery segmentation in cardiac ct angiography using 3d multi-channel u-net", *arXiv preprint arXiv:1907.12246*, 2019.

[40] Y. Duan, J. Feng, J. Lu, and J. Zhou, "Context aware 3d fully convolutional networks for coronary artery segmentation", in *Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges: 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers 9*, Springer, 2019, pp. 85–93.

[41] Z. Gao, X. Liu, S. Qi, W. Wu, W. K. Hau, and H. Zhang, "Automatic segmentation of coronary tree in ct angiography images", *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 8, pp. 1239–1247, 2019.

[42] P. Mirunalini, C. Aravindan, A. T. Nambi, S. Poorvaja, and V. P. Priya, "Segmentation of coronary arteries from cta axial slices using deep learning techniques", in *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*, IEEE, 2019, pp. 2074–2080.

[43] Y. Shen, Z. Fang, Y. Gao, N. Xiong, C. Zhong, and X. Tang, "Coronary arteries segmentation based on 3d fcn with attention gate and level set function", *Ieee Access*, vol. 7, pp. 42 826–42 835, 2019.

[44] S.-H. Tsang, *Review: Unet++ — a nested u-net architecture (biomedical image segmentation)*, https://sh-tsang.medium.com/review-unet-a-nested-u-net-architecture-biomedical-image-segmentation-57be56859b20, Accessed: 2025-06-24, 2019.

[45] J. M. Wolterink, T. Leiner, and I. Išgum, "Graph convolutional networks for coronary artery segmentation in cardiac ct angiography", in *Graph Learning in Medical Imaging: First International Workshop, GLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Proceedings 1*, Springer, 2019, pp. 62–69.

[46] H. Yang, J. Chen, Y. Chi, X. Xie, and X. Hua, "Discriminative coronary artery tracking via 3d cnn in cardiac ct angiography", in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22*, Springer, 2019, pp. 468–476.

[47] Y. Fu, B. Guo, Y. Lei, T. Wang, T. Liu, W. Curran, L. Zhang, and X. Yang, "Mask r-cnn based coronary artery segmentation in coronary computed tomography angiography", in *Medical Imaging 2020: Computer-Aided Diagnosis*, SPIE, vol. 11314, 2020, pp. 1047–1052.

[48] J. Gu, Z. Fang, Y. Gao, and F. Tian, "Segmentation of coronary arteries images using global feature embedded network with active contour loss", *Computerized Medical Imaging and Graphics*, vol. 86, p. 101 799, 2020.

[49] B. Kong, X. Wang, J. Bai, Y. Lu, F. Gao, K. Cao, J. Xia, Q. Song, and Y. Yin, "Learning tree-structured representation for 3d coronary artery segmentation", *Computerized Medical Imaging and Graphics*, vol. 80, p. 101 688, 2020.

[50] ——, "Learning tree-structured representation for 3d coronary artery segmentation", *Computerized Medical Imaging and Graphics*, vol. 80, p. 101 688, 2020.

[51] Y. Lei, B. Guo, Y. Fu, T. Wang, T. Liu, W. Curran, L. Zhang, and X. Yang, "Automated coronary artery segmentation in coronary computed tomography angiography (ccta) using deep learning neural networks", in *Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications*, SPIE, vol. 11318, 2020, pp. 279–284.

[52] W. K. Cheung, R. Bell, A. Nair, L. J. Menezes, R. Patel, S. Wan, K. Chou, J. Chen, R. Torii, R. H. Davies, *et al.*, "A computationally efficient approach to segmentation of the aorta and coronary arteries using deep learning", *Ieee Access*, vol. 9, pp. 108 873–108 888, 2021.

[53] H. Du, K. Shao, F. Bao, Y. Zhang, C. Gao, W. Wu, and C. Zhang, "Automated coronary artery tree segmentation in coronary cta using a multiobjective clustering and toroidal model-guided tracking method", *Computer Methods and Programs in Biomedicine*, vol. 199, p. 105 908, 2021.

[54] L. Gu and X.-C. Cai, "Fusing 2d and 3d convolutional neural networks for the segmentation of aorta and coronary arteries from ct images", *Artificial intelligence in medicine*, vol. 121, p. 102 189, 2021.

[55] R. Liang, J. Ma, G. Ma, K. Wang, *et al.*, "3d u-net with attention and focal loss for coronary tree segmentation", 2021.

[56] L.-S. Pan, C.-W. Li, S.-F. Su, S.-Y. Tay, Q.-V. Tran, and W. P. Chan, "Coronary artery segmentation under class imbalance using a u-net based architecture on computed tomography angiography images", *Scientific reports*, vol. 11, no. 1, p. 14 493, 2021.

[57] F. Tian, Y. Gao, Z. Fang, and J. Gu, "Automatic coronary artery segmentation algorithm based on deep learning and digital image processing", *Applied Intelligence*, vol. 51, no. 12, pp. 8881–8895, 2021.

[58] ——, "Automatic coronary artery segmentation algorithm based on deep learning and digital image processing", *Applied Intelligence*, vol. 51, no. 12, pp. 8881–8895, 2021.

[59] Q. Wang, W. Zhao, X. Yan, H. Che, K. Ye, Y. Lu, Z. Li, and S. Cui, "Geometric morphology based irrelevant vessels removal for accurate coronary artery segmentation", in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2021, pp. 757–760.

[60] Y. Lin, H. Yao, Z. Li, G. Zheng, and X. Li, "Calibrating label distribution for class-imbalanced barely-supervised knee segmentation", in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2022, pp. 109–118.

[61] S. Ono, M. Komatsu, A. Sakai, H. Arima, M. Ochida, R. Aoyama, S. Yasutomi, K. Asada, S. Kaneko, T. Sasano, *et al.*, "Automated endocardial border detection and left ventricular functional assessment in echocardiography using deep learning", *Biomedicines*, vol. 10, no. 5, p. 1082, 2022.

[62] H. Rahman, T. F. N. Bukht, A. Imran, J. Tariq, S. Tu, and A. Alzahrani, "A deep learning approach for liver and tumor segmentation in ct images using resunet", *Bioengineering*, vol. 9, no. 8, p. 368, 2022.

[63] C. W. Tsao, A. W. Aday, Z. I. Almarzooq, A. Alonso, A. Z. Beaton, M. S. Bittencourt, A. K. Boehme, A. E. Buxton, A. P. Carson, Y. Commodore-Mensah, *et al.*, “Heart disease and stroke statistics—2022 update: A report from the american heart association”, *Circulation*, vol. 145, no. 8, e153–e639, 2022.

[64] H. Zhu, S. Song, L. Xu, A. Song, and B. Yang, “Segmentation of coronary arteries images using spatio-temporal feature fusion network with combo loss”, *Cardiovascular Engineering and Technology*, pp. 1–12, 2022.

[65] ——, “Segmentation of coronary arteries images using spatio-temporal feature fusion network with combo loss”, *Cardiovascular Engineering and Technology*, pp. 1–12, 2022.

[66] ——, “Segmentation of coronary arteries images using spatio-temporal feature fusion network with combo loss”, *Cardiovascular Engineering and Technology*, pp. 1–12, 2022.

[67] A. Zeng, C. Wu, G. Lin, W. Xie, J. Hong, M. Huang, J. Zhuang, S. Bi, D. Pan, N. Ullah, *et al.*, “Imagecas: A large-scale dataset and benchmark for coronary artery segmentation based on computed tomography angiography images”, *Computerized Medical Imaging and Graphics*, vol. 109, p. 102 287, 2023.

[68] S. Li, W. Zhang, and Q. Chen, “Multimodal knowledge distillation-guided artery-vein segmentation network for oct/octa images”, in *Sixteenth International Conference on Graphics and Image Processing (ICGIP 2024)*, SPIE, vol. 13539, 2025, pp. 220–229.

[69] OECD AI, *Dice score*, Accessed: 2025-06-19, 2025. [Online]. Available: https://oecd.ai/en/catalogue/metrics/dice-score.

[70] PyTorch Contributors. (2025). torch.backends.cudnn — pytorch 2.7 documentation. Accessed June 16, 2025, [Online]. Available: https://pytorch.org/docs/stable/backends.html#module-torch.backends.cudnn (visited on 06/16/2025).