University of New South Wales

Thesis C

Automated Labeling of Coronary Arteries through Neural Networks

Qie Shang Pua

School of Mechanical and Manufacturing Engineering

Supervised by

Dr Susann Beier

School of Mechanical and Manufacturing Engineering

6 November 2020

Acknowledgments

First and foremost, I would like to express my utmost gratitude to my family members for their constant emotional support, especially in weathering the Covid-19 pandemic. I also would like to thank Dr. Susann for providing guidance, encouragement and checking on my progress in the weekly meetings – this has certainly motivated me to perform better. Not to mention, thank you Ramtin for being a constant source of knowledge. I really appreciate your technical advice, research papers suggestions and installation of tricky VMTK package. Last but not least, I would like to thank my colleague and friend, Jun Wen, for helping me to navigate in this complex but interesting realm of machine learning and artificial intelligence.

Abstract

An automatic labelling algorithm for coronary arteries is essential to facilitate the diagnostic process for physicians. The main challenge in this problem is the large individual variability in coronary artery tree, which stems from geometrical discrepancies in vessels, poor CT imaging quality and segmentation errors. Existing literature have focused on knowledge based and machine learning based approaches to address this challenge. Motivated by the wide application of convolutional neural network in classification applications, in this research, a novel 1D Convolutional Neural Network is proposed. Centreline points of coronary artery tree are used as input, with vessel spatial locations and diameter as features. Four different neural network frameworks (i.e. single-output CNN, multi-output CNN, Bi-LSTM, and my implementation of TreeLab-Net) are developed. The best-performing model, a single-output 1D Convolutional Neural Network framework, demonstrated a mean precision, recall and F1-score of 83.3% and 83.6% and 83.2% respectively. Moreover, this research outperforms the state-of-the-art model in labelling arteries such as RCA, acutes and diagonals. Overall, the experimental results demonstrate that the 1D Convolutional Neural Network is robust and effective method to automatically label coronary artery trees.

Contents

List of Figures

List of Tables

Acronyms

CAD Coronary Artery Disease.

CCTA or CT Coronary Computed Tomography Angiography.

CNN Convolutional Neural Network.

LAD Left Descending Artery.

LCIM Remus Intermedius.

LCX Left Circumflex Artery.

LM Left Main Artery.

LSTM Long Short-Term Memory.

MLP Multi Layer Perceptron.

OM Obtuse Marginal.

RCA Right Coronary Artery.

1 Introduction

Coronary Artery Disease (CAD) is the single leading cause of death globally [1] and early detection methods are urgently needed. Currently, Coronary Computed Tomography Angiography (CCTA or CT) remains as the gold standard for non-invasive assessment of CAD [2]. By interpreting the CT scans based on CT image guidelines, radiologists are able to report any occlusions, calcification, and stenosis in a patient's coronary vessel. Due to CT's widespread practice, identification of all individual vessels in a CT image becomes increasingly crucial. In the past, physicians and medical experts have manually labelled the coronary vessels by referring to a standard coronary anatomy model by Dodge et al [3]. However due to the complex and tortuous nature of the coronary arteries, manual labelling becomes a time-consuming task. Hence, a novel automatic labelling process for

coronary arteries is required to facilitate the diagnostic process and to visualize regions of interest.

1.1 Coronary Atlas

Dodge et al [3] proposed a standardized definition of the segmented coronary anatomy, which is depicted in Fig. 1 below. Approximately 86% of the human population possess a Right Dominant (RD) Coronary System, 9.2% of the population have a Left Dominant (LD) coronary systems and the remaining 4.2% have a balanced coronary system [4].

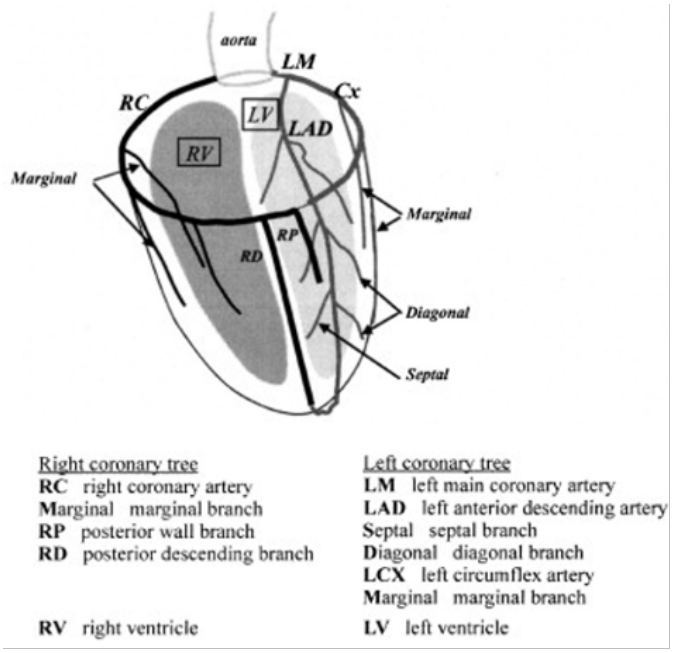

Figure 1: Schematic depiction of the coronary anatomy with a right dominant distribution (86% of cases) [4]

In a right dominant anatomy as displayed in Fig. 1, the left coronary tree supplies oxygen and nutrients to the left part of the myocardial muscles. This supply begins when two main branches exits from the aorta - the Left Main Artery (LM) and Right Coronary Artery (RCA). The LM further splits into two branches called Left Descending Artery (LAD) and Left Circumflex Artery (LCX). On rare occasions, the LM trifurcates, splitting into LAD, LCX and an additional branch

called Remus Intermedius (LCIM). LAD has several branches called Septals and Diagonals, whereas LCX subbranches are generally called Obtuse Marginal (OM). On the other hand, the Right Coronary Artery (RCA) also has subbranches called acute marginal (AM). The main vessels and its subbranches can be classified in Table 1 below.

Table 1: Main Coronary Branches and its side branches

Main Coronary Vessels	Subbranches
LAD	Septals, Diagonals
LCX	Obtuse Marginals
RCA	Acute Marginals, Crux

Left Dominant Coronary System which exists in minority of human population has a slightly different structure as compared to the Right Dominant. For instance, LCX completely supplies the posterior wall and has fewer or no marginal branches compared to a Right Dominant anatomy.

As RCA, LAD and LCX are long in length, these vessels can be divided into proximal, middle and distal sections for further medical analysis. Labelling these sections is useful as physicians could analyse the severity of the coronary artery disease and subsequently provide the appropriate prognosis. For instance, an occlusion at the proximal section would be more dangerous than at the distal section. All these considerations are included while developing a robust labelling system that is accurate for different types of heart structures.

2 Literature Review

2.1 Previous Works

Table 2: Knowledge-based methods

Author	Year	Modality	Contribution
Ezquerra et al	1998	X-ray	Converts patient's coronary tree into a directed acyclic graph, where its graph nodes are matched with a generalized 3D coronary model.
Chalopin et al	2000	X-ray	Employs a Graph Matching approach by comparing a patient's coronary tree with a reference model based on features such as eccentricity of arcs and nodes.
Yang et al	2011	CT	Uses point-set registration to align extracted centerlines to a generic coronary model. All segments are labeled based on hierarchical relationship and a priori knowledge from the generic model.
Gulsun et al	2014	CT	Incorporates relative position of detected cardiac structures during the point-set registration process. Also presented a methodology to account for missing side branches during matching.
Cao et al	2017	CT	Uses point-set registration to identify all segments of coronary tree. Also developed a labeling methodology for Left Dominant Coronary Model.

Table 3: Machine Learning based Methods

Author	Year	Modality	Contribution
Akinyemi et al [10]	2009	CT	All plausible labelled trees are generated and the most likely labeling is chosen based on closeness to parameters defined by a Multivariate Gaussian Classifier.

Table 4: Neural Network Based Methods

Author	**Year**	**Modality**	**Contribution**
Wu et al	2019	CT	Developed a Deep Neural Network framework by combining multi-layer perceptron encoder network with a bidirectional tree-structural long short term memory. Spherical coordinate transform is used to normalize spatial variations in vessels.
Yang et al [12]	2020	CT	Developed a conditional partial-residual graph convolutional network (CPR-GCN) to label vessel segments, which takes both position and CT image into consideration. 3D CNN and the Bi-LSTM model were employed to learn the conditions in the image domain for the partial-residual block.

2.1.1 Knowledge-Based Optimisation Methods

Many automatic or semi-automatic labeling methods have been developed for coronary arteries, airway trees, cerebral arteries and colon arteries which are listed in Table 2 and 3. The main challenge, especially for coronary arteries, is the large variation in the individual vessels. Vessels features such as the location, length, diameter and tortuosity vary from one person to another, resulting in different branching patterns. Moreover, noise, poor imaging quality and segmentation errors may result in missing side branches, further contributing to the variation. Hence, the ability to handle these variations becomes the primary focus for automated labeling.

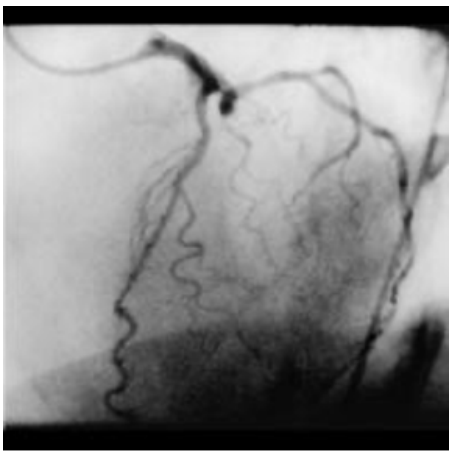

Figure 2: Typical X-ray angiogram of coronary arteries [5]

As shown in Table 2, one of the pioneering works in automated coronary arteries labeling is conducted on X-ray images. Ezquerra et al [6] demonstrated this by segmenting a coronary tree from a X-ray image as shown in Fig. 2 above and extracting the centrelines from the segmented model. Then, a graph-matching approach is employed to filter out noise, artefacts and competing structures. In this method, a generalized 3D model created by Dodge et al [3] is used as a reference model. Similarly, Chalopin et al [5] also labels the coronary tree by matching an initial graph to a list of reference graphs based on features such as inertia, inertia axis orientation and eccentricity of arcs and nodes.

Since X-ray is a different imaging modality than CT, automated labeling of coron-

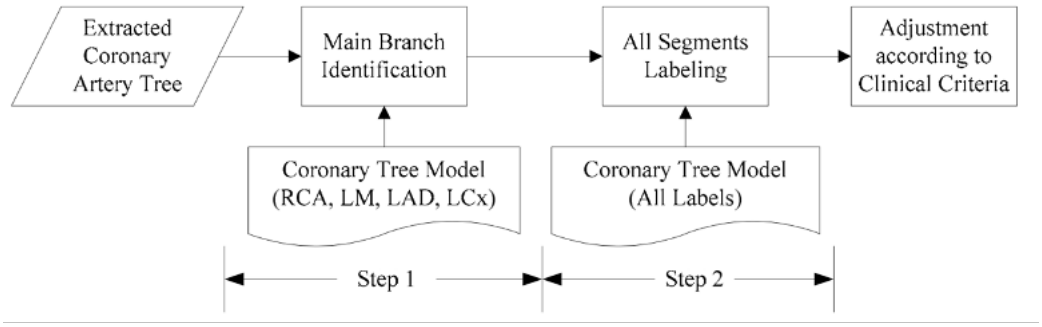

Figure 3: Labeling Procedure for a typical knowledge-based method. This approach is adopted in the works of Yang et al [6] and Cao et al [7]

ary arteries in CT images poses different challenges. It is much more difficult to extract tiny side branches from CT images compared to X-ray due to the spatial resolution of CT images, motion artifacts and presence of stenoses and occlusions. Being one of the pioneers in labelling CT modality, Yang et al [7] uses Point-set registration to align a targeted model with a generic coronary model by Dodge et al [3]. The entire workflow is displayed in Fig. 3. Then, major branches are assigned based on proximity to the ostium. Third, the side branches are labeled based on iteratively reducing a cost function which penalizes unlabeled centreline points, as shown in Fig 4.

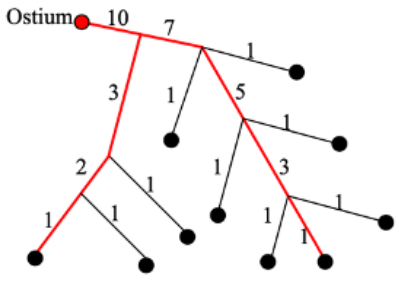

Figure 4: Edges closer to the ostium have higher weight. The path with the minimal distance to the main branch is assigned to the corresponding label [6]

A major drawback in this work is the exclusion of datasets that contains severe lesions at the proximal parts of the main branches, thus reducing the robustness of the model. Furthermore, if side branches cannot be extracted due to partial volume effect or motion artifacts, that dataset is omitted. Hence, this method is

only accurate if a high quality CT images is used and severe occlusions are absent, which is rarely the case in clinical datasets. Despite the flaws, significant progress was achieved because not only the individual branches as a whole are successfully labelled, but the proximal, middle and distal sections of the main branches are labelled as well. This work achieved an **overall overlap rate or precision of 91.41%** for 741 vessels segments in 58 CT datasets.

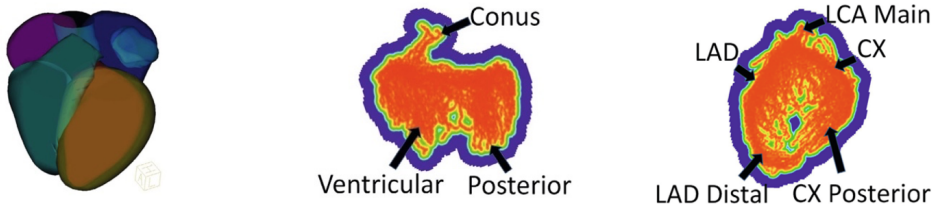

Figure 5: (a) Four chamber models are detected in a patient's coronary system to assist in labeling. (b) Right and (c) left coronary territories with certain regions depicted. Blue indicates low probability, red high probability [8]

Gulsun et al. [8] sought to overcome the rigidness of Yang et al's model [7] by incorporating a coronary tree's geometric information during the registration process. Firstly, the geodesic paths between a patient's coronary tree model and standard model are calculated using Dijkstra's shortest path algorithm. Second, the branches are ordered based on their position relative to the heart surface. Specifically, all four heart chambers are identified to compute the likelihood of where a particular coronary artery tends to lie, as shown in Fig. 5. For example, LAD is highly likely to lie on the left ventricle. One major contribution of this work is the ability to account for missing side branches, which is achieved by considering equivalent partitioned representations of source and target tree-shapes. Despite being theoretically sound, Gulsun et al's [8] method has a lower precision compared to Yang et al's [7] method – achieving an **overlap rate of 86.5% (87% for the right coronary artery and 86%** for the left coronary artery) if centrelines are automatically detected. However, if centrelines are manually annotated instead of automatically detected, the success rate increases to 94% in a Right Dominant (RD) system. A downside to this method is that finding geodesic paths in Quotient Euclidean Distance (QED) space is computationally very expensive. For this method, 37 CT datasets are used.

Another significant work on automated labeling was done by Cao et al [9]. The methodology is very similar to Yang et al's [7] work, which is through a point set registration method followed by evaluating the matching costs for every vessel segment (refer Fig. 3 for the entire workflow). Cao et al's [9] method is more superior than Gulsun et al's [8] method due to two reasons: (1) labelling take less than 3 seconds whereas Gulsun et al's [8] method takes 3 minutes. (2) only the coronary centrelines points were needed unlike Gulsun et al's [8] method which requires the anatomical prior location such as position of four chambers. A significant contribution of this work is the development of a left dominant coronary model as this was not accomplished in previous works. In this work, two experts were requested to assess the precision or the overlap rate. The results demonstrated an **overall precision of 92% and 90.7% for right dominant (RD) and left dominant (LD) coronary system, respectively**. The overall precision is slightly higher compared to Yang et al [7] and Gulsun et al [8], which is 91.41% and 86.5% respectively.

2.1.2 Machine Learning Based Methods

As aforementioned, most of the early works adopts a knowledge based approach, where a generalized model such as Dodge et al's [3] model is used as reference. However, graph matching approach may be flawed as the cost functions do not account for the multivariate nature of the coronary vessels. For example, if a septal vessel is located far from the expected location or more curved than expected, point-set registration is unable to match them effectively, hence this approach would yield a false negative i.e. there is no septal. As a result, clinical data set with anomalies may fail. To address the challenge of large variation in individual vessels, several studies have proposed to adopt a machine learning approach.

The only study that employs a machine learning approach to label coronary arteries is conducted by Akinyemi et al [10]. In this study, centrelines of coronary arteries and its geometric features such as length, branch angle, tortuosity, and diameter are used to train a multivariate Gaussian classifier. As shown in Fig. 7, an unlabeled coronary tree is assigned with all permutations of labels, and only those that fulfill the topological and geometrical rules are retained. The topological

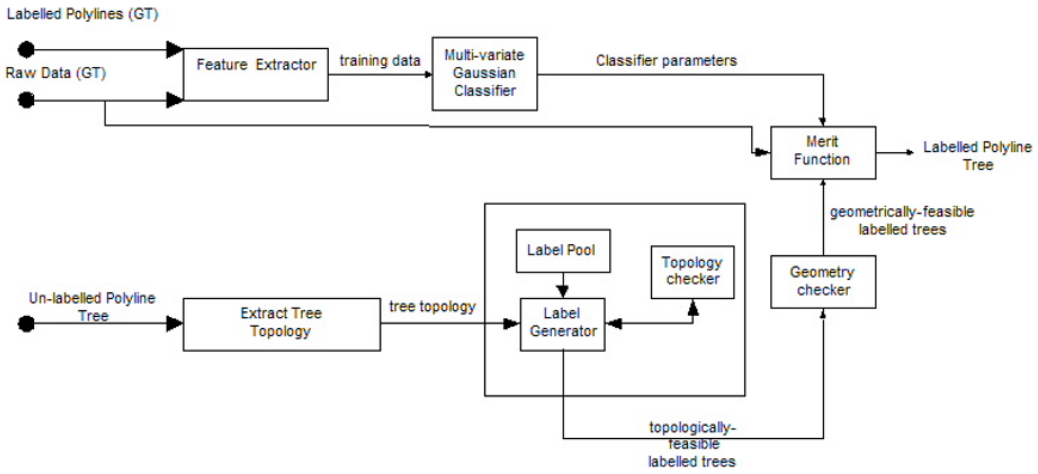

Figure 6: Workflow Diagram by Akinyemi et al [9]

and geometrically feasible labelled tree possibilities are then fed into the merit function. The classified tree with the highest merit is returned as labeled polyline tree. The Multivariate Classifier uses 42 CT datasets of varying image quality, yielding an overall **accuracy of 76.47%** when automatically segmented vessels are used as input. A popular source of error here is the Left LCIM segments being misclassified as Left Diagonal and Left OM due to logic of the topological rules.

2.1.3 Neural Networks and Deep Learning

Prior to 2019, the machine learning based anatomical labelling techniques [10, 13, 14] are used to provide a likelihood of a certain label based on extracted geometrical features. However in 2019, a ground-breaking journal article by Wu et al [11] demonstrated that anatomical labelling can be achieved by Neural Networks. The proposed framework is called TreeLab-Net, which has the following methodology:

- Spherical Coordinate Transform - To normalise the large anatomical variation across various subjects, Cartesian coordinates of the centreline points are converted to spherical coordinates.
- Feature Selection - the centrelines are split into multiple vessel segments, where a segment's features such as coordinates, parent-class angles and directional vectors are fed into a Recurrent Neural Network (RNN) based ar-

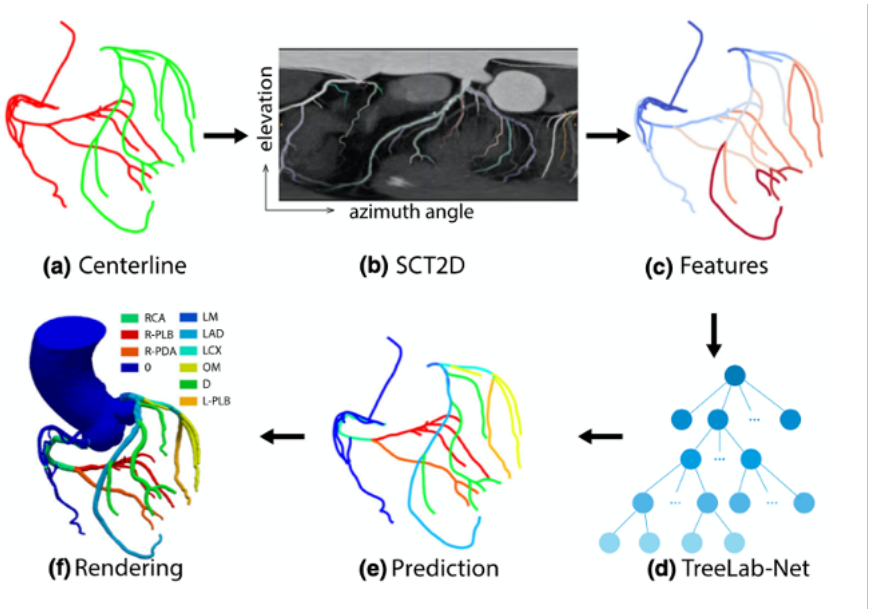

Figure 7: An overview of the proposed TreeLab-Net framework. From (a) centrelines of the coronary arteries are extracted as inputs, (b) Spherical coordinates are transformed to 2D space for (c) feature extraction. Then, the labels for vessel segments were predicted by (d) TreeLab-Net and stores in the (e) centrelines. Finally (f) a 3D coronary tree model with the predicted labels are rendered for visualization [10]

chitecture called TreeLab-Net.

- Train and Test Tree-Lab Net – Tree-Lab Net consists of a Multi Layer Perceptron (MLP) and bidirectional tree structural Long Short-Term Memory (LSTM). The MLP encodes the feature vector for each vessel segment into a compact fixed length vector, which is fed into Bi-LSTM to predicts the labels.

The validation results were promising; **the precision, recall and F1 rates are 0.866, 0.879 and 0.871 respectively**, and is considered to be the state-of the-art in this research domain. Most of the inaccuracies are due to the error in labelling tiny side branches such as right posterior descending artery. Branches with high clinical importance such as LAD, LCX and RCA achieved an accuracy that is greater than 0.9. Compare to previous works, which have a small number of subjects of less than 100, this work has achieved a high accuracy with 436 datasets.

Although TreeLab-Net is a significant breakthrough in the field of coronary arteries labelling, Yang et al [12] argued that TreeLab-Net is modelled based on a flawed assumption - that each main branch will only bifurcate into two sub-branches, hence each node in a tree will only have two children. However, in practice, it is possible to have more than two sub-branches bifurcating near to each other, therefore undermining the accuracy of TreeLab-Net. Yang et al [12] sought to solve this problem through conditional partial-residual graph convolutional network (CPR-GCN), with the steps featured in the Fig. 9 below.

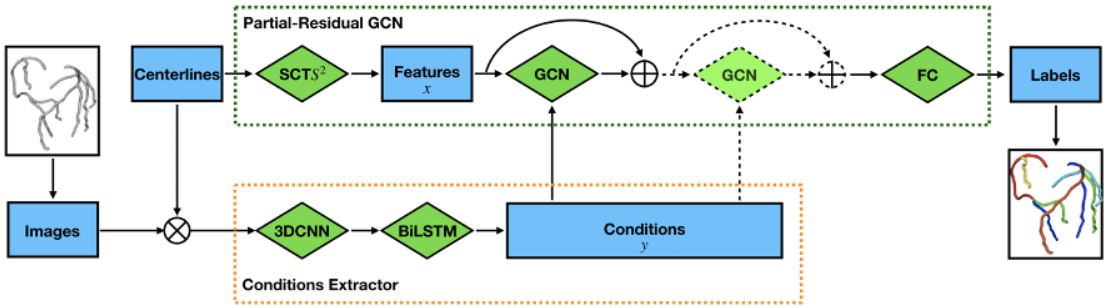

Figure 8: Framework of CPR-GCN, which consists of two parts – partial-residual GCN (in green box) and conditions extractor (in orange box). 3D images of the coronary map and centreline points are fed into both parts, where the GCN eventually generates the labels for every 3D segment of coronary vessels. [11]

In essence, CPR-GCN consists of two parts:

- Conditions Extractor – The spatial features of the 3D images and centreline points are extracted by 3D Convolutional Neural Network (CNN) model, and analysed in a sequential manner using Bi-LSTM model. The last hidden state of the Bi-LSTM model becomes the final conditions, which represents the image information of a single branch.
- Partial-Residual Block of GCN – Image information from the condition extractor and centrelines points are fed into the block, which builds edges and nodes of a graph. The graph then predicts a label for every node in the graph, representing each coronary vessel segment.

This model achieves a **mean F1-score of 0.955, precision of 0.954 and recall**

of 0.958. Despite outperforming Wu et al's TreeLab-Net, this paper is relatively new and not widely adopted in the scientific community compared to TreeLab-Net, hence the validity of the results remains in question.

2.2 Research Gap - Convolutional Neural Networks

In recent years, there is a clear shift towards Deep Learning models for image classification, object detection, segmentation, registration, and other tasks in the biological field [16]. With the development of computational capability, deep neural networks have won numerous contests in image processing and machine learning. In particular, Convolutional Neural Network (CNN) is widely adopted for segmentation and extraction of coronary centreline [16, 17]. A CNN workflow is generally composed as shown in Fig. 10 below.

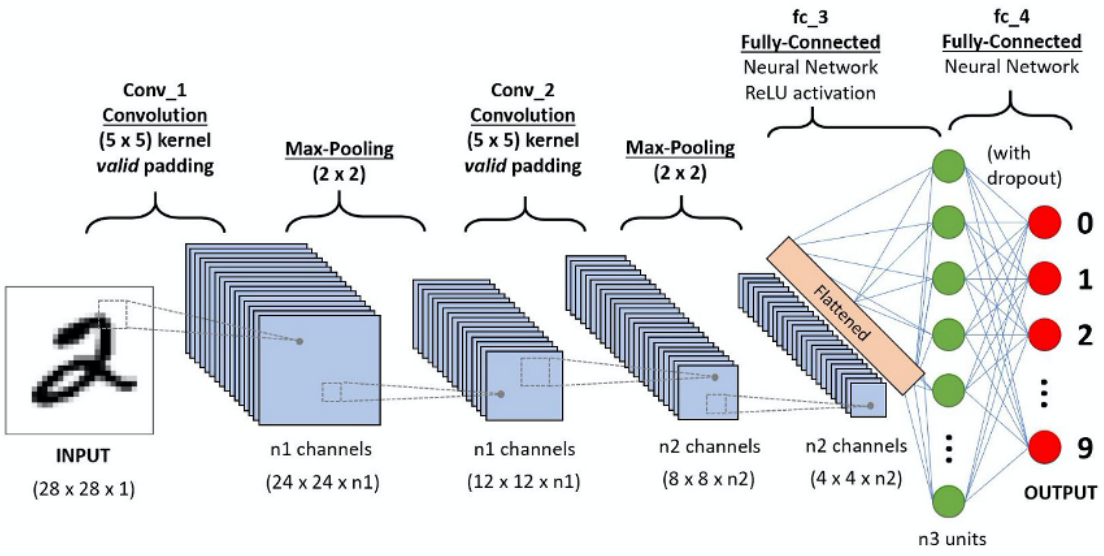

Figure 9: Layers of Convolutional Neural Network which takes a 2D handwritten image as input. [12]

2.2.1 Types of Layers

As displayed in Fig. 11, the **convolutional layer** uses filters that perform convolution operations as it is scanning the input with respect to its dimensions. Its

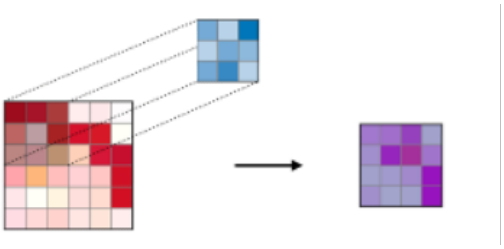

Figure 10: The input (red) is filtered with a 3x3 filter/kernel (in blue) and the corresponding output is a feature map (purple)

hyperparameters include the filter (also known as kernel) size and stride. The resulting output is called feature map or activation map.

The **pooling layer** is a down sampling operation, typically applied after a convolution layer, which does some spatial invariance. In particular, max and average pooling are special kinds of pooling where the maximum and average value is taken from a kernel, respectively. Pooling layers aims to decrease computational power required to process the data.

The input is then flattened into a column vector or one-hot vector. In a feed-forward neural network, the **fully connected layer** operates on a flattened input where each input is connected to all neurons. If present, fully connected layers are usually found towards the end of CNN architectures and can be used to optimize objectives such as class scores. Backpropagation is applied to update the weights in this process. Lastly, the SoftMax step is a generalized logistic function that takes as input a vector of scores and outputs a vector of output probability.

The role of CNN is to reduce data such as images, large chunks of text into a form which is easier to process, without losing features which are critical for getting a good prediction. This is crucial when architecture is required to be not only good at learning feature, but is also scalable to massive datasets. Despite the popularity and success of CNN, there is only a single literature by Yang et al [12] that employs CNN to label coronary vessels. However, that paper uses a Graph-based, 3-dimensional CNN which is computationally expensive and takes a long time. To reduce the time taken to label coronary arteries among patients, lower dimensional CNN models presents an opportunity to be explored.

3 Methodology

3.1 Data Collection and Preprocessing

3.1.1 Data Collection and Ground Truth Labeling

A total of 323 CT scans are segmented with an in-house segmentation software to generate coronary artery trees. As seen in Fig. 12 below, each coronary artery tree comes in two forms: (1) a 3D segmented structure in STL form and (2) a n×4 tensor, where n represents the number of centreline points of the coronary artery tree. A centreline point is the middle of a coronary artery when the thickness of the branch is equal to 1 pixel. In the n×4 tensor, the four columns represent the Cartesian coordinates (x, y, z) of the centreline point and the branch radius. Due to the variation of heart anatomy among patients, every tensor would usually have 5 to 16 coronary branches.

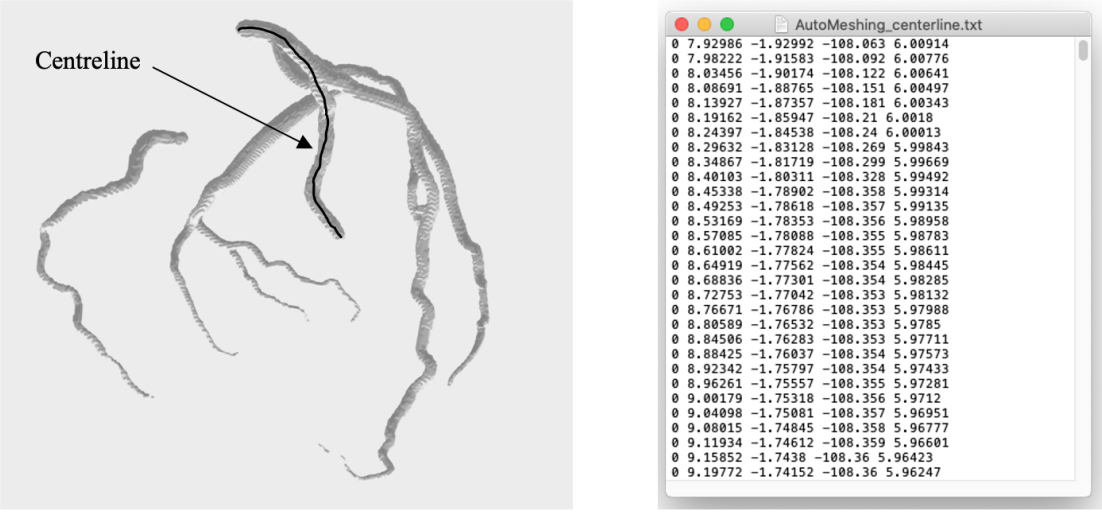

Figure 11: Raw Data is a coronary artery tree in (a) STL format and (b) n×5 tensor of centreline points

Between the two candidates for input, 2D tensors are computationally less expensive to train in neural network models, hence n×4 tensors are selected for further preprocessing. As the a n×4 tensors are unlabelled, the ground truth labels are first obtained from experts and assigned to the relevant branches. In this research,

there are a total of 10 classes of coronary branches, namely (1) LAD; (2) Diagonals; (3) Septals; (4) LCX; (5) OM; (6) Atrials; (7) LCIM (8) Acutes; (9) Crux. These classes are based on Dodge et al's reference model [3] as shown in Fig. 1. Refer to the acronyms section for the complete name of the coronary vessel.

3.1.2 Spherical Coordinate Transformation

In the tensor, the location of all centreline points are based on the Cartesian coordinate system, which may vary greatly due to segmentation errors and noise. To normalize the spatial variations, the Cartesian coordinates of the centreline points are converted into spherical coordinates, defined by the radius, elevation angle and azimuth angle. The origin of the spherical coordinates system is located by determining the bounding box center of all the centreline points in the n×4 tensor. Transformation to spherical coordinates will generate a 2D image with azimuth angle and elevation angle to act as the new x- and y- axis, which is analogous to a world map view of the coronary artery tree.

3.1.3 Feature Selection

The features that will be used to train the model are the Cartesian or spherical coordinates of the centreline points and the corresponding radius of the vessel segment.

3.1.4 Fixed Shape Tensor as Input for Convolutional Neural Networks

For CNN models, the input should ideally be a fixed-size tensor if padding is absent. Hence for every coronary branch in the n×4 tensor, 100 points from each branch are uniformly selected to represent itself. For instance, if a LAD branch has 3000 centreline points, 1 in every 30 points is selected to achieve a final row number of 100. This ensures that every coronary branch will be represented by a fixed-size tensor of 100×4 before feeding into the CNN model.

3.2 Model Creation and Optimization

In this thesis, four different Neural Network Models have been implemented to determine the best-performing model for automatically labelling coronary centrelines. The four models are as follows:

- Single-Output CNN Model
- Multi-output CNN Model
- Bidirectional Long Short-Term Memory (Bi-LSTM) Model
- My implementation of TreeLab-Net

3.3 Single-output CNN Model

Table 5: Model Architecture of Single-Output CNN Model

Layer	**Details**
1D CNN	kernel size = 3, in channel = 4, out channel = 16, maxpooling size = 2 kernel size = 3, in channel = 16, out channel = 64, maxpooling size = 2 kernel size = 3, in channel = 64, out channel = 256, maxpooling size = 2 kernel size = 3, in channel = 256, out channel = 512, maxpooling size = 2
Fully Connected	out features = 128 out features = 32 out features = Number of classes (9)

Essentially, this CNN model analyses a single branch in the coronary artery tree and predicts an output for it. The input for this model is a single coronary branch, represented by a 100×4 tensor as mentioned in Section 3.1.4. Features such as Cartesian coordinates and the radius are fed into the channels of the convolutional layers. The input tensor then undergoes multiple convolutions by a kernel, maximum pooling and eventually flattens itself for the fully connected layer. After passing through the fully connected layers, a LogSoftMax function followed by a logarithm is used to predict the final class for the coronary branch. A simplification of the LogSoftMax function is as follows

$$LogSoftMax(x_i) = log\left(\frac{e^{x_i}}{\sum_{j=1}^{N} e^{x_j}}\right),$$

where x_i is the input vector, and x_j is the output vector, and N is the number of classes.

Dropout layers and batch normalisation is incorporated in this CNN model to prevent over-fitting. ReLU activation is chosen because it is a linear function, and models are easier to optimise if their behaviour is closer to linear . Adam optimiser is used with a learning rate of 1×10^{-4} and weight decay of 1×10^{-4}. The mini-batch size is set to 16.

3.4 Multi-output CNN Model

So far, the single-output CNN model takes in a single coronary branch as an input and predicts its class. Instead of predicting one single branch at a time, the multi-output CNN model intends to predict all branches in a patient's coronary artery tree at once. For example, if a patient has 12 coronary branches, the CNN models will generate predictions for all 12 branches at one go. Theoretically, this model will be superior than the single-output model, because the weights that are usually associated with a particular class is likely to be activated when same class are often detected by those weights. However, a potential issue arises in this method as each patient has a different number of coronary branches, but the CNN model can only accept a fixed number of branches as input. To mitigate this problem, it is assumed that each patient will only have a maximum of 16 branches and a new class is created for unknown branches. Hence, if a patient has 10 coronary branches, but only 6 of them have ground truth labels, the remaining 4 unknown branches will be labelled as "unknown".

Overall, the model architecture and optimiser used in this model are similar to that of single-output CNN Model, but with a slight difference – all coronary branches on a single patient are taken in as inputs whereas the outputs are generated all at once. Apart from that, the mini-batch size is set to 4, indicating that coronary artery trees of 4 patients are being fed into the model at once.

Table 6: Model Architecture of of Bi-LSTM Model

Layer	**Details**
Bi-LSTM	Layer= 2, hidden size = 128
Fully Connected	out features = 64 out features = Number of classes (9)

3.4.1 Bi-LSTM Model

Apart from CNN, Recurrent Neural Networks are also investigated to solve this problem. Recurrent Neural Networks have loops that allows information to persist, acting in a manner that is like human memory. Naturally, these networks are mainly used for data that consist of sequences or lists, as the order of the information matters. Long Short-Term Memory (LSTM) is a special kind of Recurrent Neural Networks, capable of learning long-term dependencies and has been widely used. A bidirectional LSTM or Bi-LSTM consist of two LSTM models: one taking the input in a forward direction and the other in a backwards direction. Bi-LSTM typically performs better than a LSTM because propagating from both directions provides more information for the network from both directions, thus improving the context available to the algorithm.A LSTM is made up of three gates: (1) Input (2) Forget and (3) Output and these gates are calculated through the following equations

$$
\begin{aligned}
i_t &= \sigma(w_i[h_{t-1}, x_t] + b_i) \\
f_t &= \sigma(w_f[h_{t-1}, x_t] + b_f) \\
o_t &= \sigma(w_o[h_{t-1}, x_t] + b_o)
\end{aligned}
$$

where
i_t = input gate
f_t = forget gate
σ = sigmoid function
w_x = weights of the respective gates (x) neurons
h_{t-1} = output from the previous LSTM block

x_t = input at current timestamp
b_x = biases for the respective gates (x)

The equations for the cell state, candidate cell state and the final output of the LSTM cell is calculated as follows:

$$\tilde{c_t} = tanh(w_c[h_{t-1}, x_t] + b_c)$$
$$c_t = f_t * c_{t-1} + i_t * \tilde{c_t}$$
$$h_t = o_t * tanh(c_t)$$

where
c_t = cell state(memory) at timestamp t
$\tilde{c_t}$ = represents candidate for cell state at timestamp t

The centreline points of the coronary artery tree is first processed to become sequences of length 16. These sequences are fed into the Bi-LSTM model and the fully connected layer, and a SoftMax layer is used to predict the branch label of the next centreline point.

3.4.2 My implementation of TreeLab-Net

This model is inspired by Wu et al's [11] Neural Network architecture called TreeLab-Net. As mentioned in Section 7.2.1, their model consists of a MLP and a Tree-structural Bi-LSTM model. Vessel spatial locations and directions are selected as input features for their model, and the spatial variation are normalised via spherical coordinate transform.

MLP, colloquially known as "vanilla" neural networks, are essentially feed-forward Artificial Neural Networks, and these networks have the same characteristics as fully connected linear layers. In Wu et al's model, MLP acts as a feature encoder, encoding the vessel segments of varying length into a compact fixed-length vector. A similar approach is adopted in my implementation, with a modification – a graph-based convolutional layer is used instead of the MLP. This decision was

made based on computational complexity; CNN models have smaller and shared weights, making it easier and faster to train compared to MLP.

Prior to training my TreeLab-Net model, the centreline points in the n×5 tensor are converted to a directed graph data structure, as shown in Fig. 13 below. Each node in the graph represents a bifurcation or the terminal points of a vessel segment, and the edges represents the vessel segment between two bifurcations or terminal points. The nodes store features of the bifurcation such as the Cartesian coordinates and the radius.

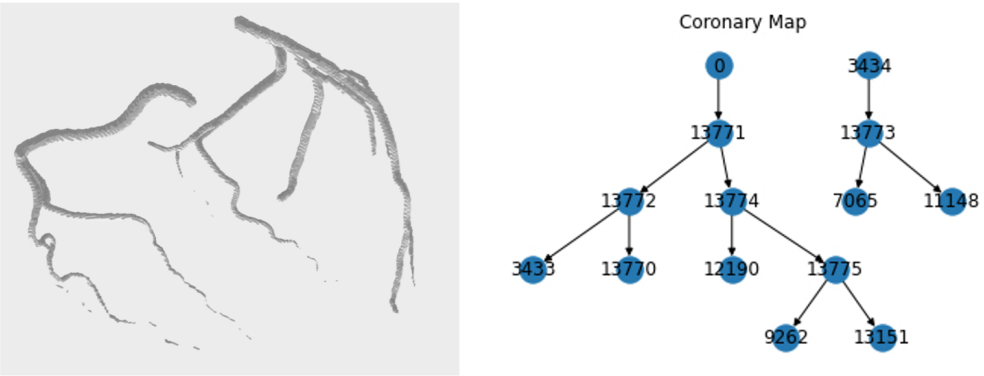

Figure 12: Coronary artery tree (left) is converted to a directed graph structure (right)

After converting a patient's coronary artery tree to a graph data structure, it is fed into TreeLab-Net as input. The graphs first enters a CNN layer, where it encodes not only the features of the nodes, but the parent-child relationship among vessel segments, therefore providing more context to the model. Then, Bi-LSTM uses these high-level representations of the graph input to predict the labels of the nodes or bifurcations. The fully connected layers flatten the output of the Bi-LSTM and the SoftMax function predicts the label for each bifurcation. The complete architecture is displayed in Table 7 below.

In this model, Adam optimiser is used, with a learning rate of 0.001 and a weight decay of 1x10-5. The mini-batch size is set to 16. Dropout layers are incorporated in this model to prevent over-fitting. ReLU activation is used for the CNN and fully connected layers.

Table 7: Model Architecture of my implementation of TreeLab-Net

Layer	**Details**
Graph CNN	In channels = 4, out channels = 32 In channels = 32, out channels = 16
Bi-LSTM	Number of layers = 1, Hidden size= 30,
	out features = 64 out features = Number of classes (9)

3.5 Training and Testing Models

After developing all four models, the 323 patient datasets are split into training, validation and testing data in a ratio of 0.8:0.1:0.1. Through a CUDA-enabled Graphical Processing Unit (GPU), training data is used to update the weights in the models. Then, for each model, validation data is used to tune the hyperparameters and metaparameters, namely learning rate, momentum, weights initialization, batch size and optimiser.

After tuning the hyperparameters, all four models are tested with unseen data. The metrics of each model – precision, recall and F1-score - are recorded. K-fold cross validation is performed to obtain an average of those metrics.

4 Results

4.1 Experimental Results for all models

Fivefold cross-validation is implemented for every model, where the number of training samples to testing samples are 260:63 in each fold. The evaluation is performed by comparing the predicted label and the ground truth label. Precision, recall and F1 scores were calculated for each class, based on the following equations:

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$
$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}$$
$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Table 8: Comparison of evaluation metrics between models in this research

Vessel	Single-Output CNN			Multi-Output CNN			My TreeLab-Net		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
LAD	0.885	0.885	0.885	0.759	0.759	0.759	0.762	0.755	0.759
Diagonals	0.842	0.941	0.889	0.645	0.667	0.656	0.541	0.625	0.580
Septal	0.750	0.600	0.667	0.125	0.333	0.182	0.667	0.250	0.364
LCX	0.800	0.750	0.774	0.828	0.857	0.842	0.547	0.565	0.556
OM	0.500	0.556	0.526	0.571	0.632	0.6	0.545	0.514	0.529
Atrial	0	0	0	0	0	0	0	0	0
LCIM	0.500	0.286	0.364	0	0	0	0.200	0.143	0.167
Acutes	0.913	0.840	0.875	0.852	0.852	0.852	0.840	0.808	0.824
Crux	0.926	0.955	0.940	0.93	0.816	0.87	0.924	0.953	0.740
Overall	**0.833**	**0.836**	**0.832**	0.736	0.724	0.729	0.737	0.740	0.736

Table 9: Comparison between my single-output model with the state-of-the-art models

Method	Knowledge based method			Wu et al's TreeLab-Net			My Single Output CNN		
Metric	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
RCA	0.925	0.918	0.922	0.948	0.95	0.949	-	1	-
Acutes	0.871	0.893	0.882	0.871	0.871	0.871	0.913	0.955	0.875
LAD	0.929	0.911	0.92	0.937	0.948	0.942	0.855	0.855	0.855
LCX	0.81	0.832	0.821	0.936	0.913	0.924	0.8	0.75	0.774
LCIM	0.803	0.848	0.825	0.714	0.77	0.741	0.5	0.286	0.364
Diagonals	0.781	0.799	0.789	0.841	0.816	0.829	0.842	0.941	0.889
OM	0.739	0.72	0.73	0.841	0.816	0.829	0.5	0.556	0.526
Septal	0.865	0.835	0.85	0.859	0.862	0.86	0.75	0.6	0.667
Average	0.855	0.859	0.857	0.868	0.873	0.871	0.833	0.836	0.832

Table 10: Comparison of evaluation metrics between Cartesian coordinates and spherical coordinates in Single-Output CNN Model

Vessel	Single-Output CNN Model					
	Cartesian Coordinates			Spherical Coordinates		
	Precision	Recall	F1-score	Precision	Recall	F1-score
LAD	0.885	0.885	0.885	0.625	0.938	0.750
Diagonals	0.842	0.941	0.889	0.548	0.500	0.523
Septal	0.75	0.6	0.667	0.750	1	0.857
LCX	0.8	0.75	0.774	1	0.200	0.333
Obtuse Marginal	0.5	0.556	0.526	0.425	0.850	0.567
Atrial	0	0	0	0	0	0
LCIM	0.5	0.286	0.364	0.500	0.083	0.143
Acutes	0.913	0.840	0.875	0.833	0.952	0.889
Crux	0.926	0.955	0.940	0.941	0.865	0.901
Overall	**0.833**	**0.836**	**0.832**	0.729	0.667	0.632

Table 11: Accuracy of Predicting the Main Branches based a single-output CNN Model

Main Vessels	**Side Branches**	**Accuracy**
LAD	Septals, Diagonals	0.942
LCX	LCX, Obtmar	0.981
RCA	Acutes, Crux	1.0

The metrics of Bi-LSTM model is not displayed in this table due to the poor results compared to the three models above, where the recall ranges from 0.6 to 0.65. The accuracy and loss graphs of the Bi-LSTM model are featured in Appendix A.

4.2 Confusion Matrix

Atrials class is omitted in this confusion matrix because atrial samples is absent in the test set; it is a very rare coronary vessel.

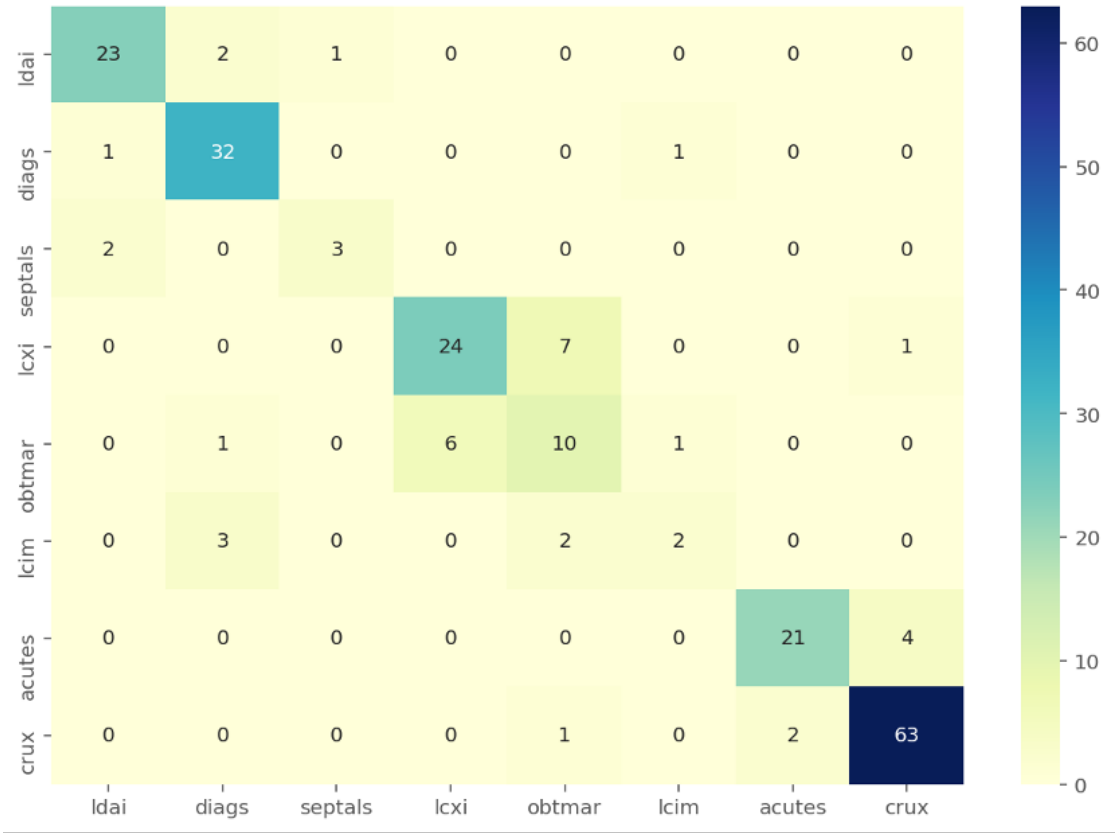

Figure 13: Confusion Matrix of Single Output CNN Model

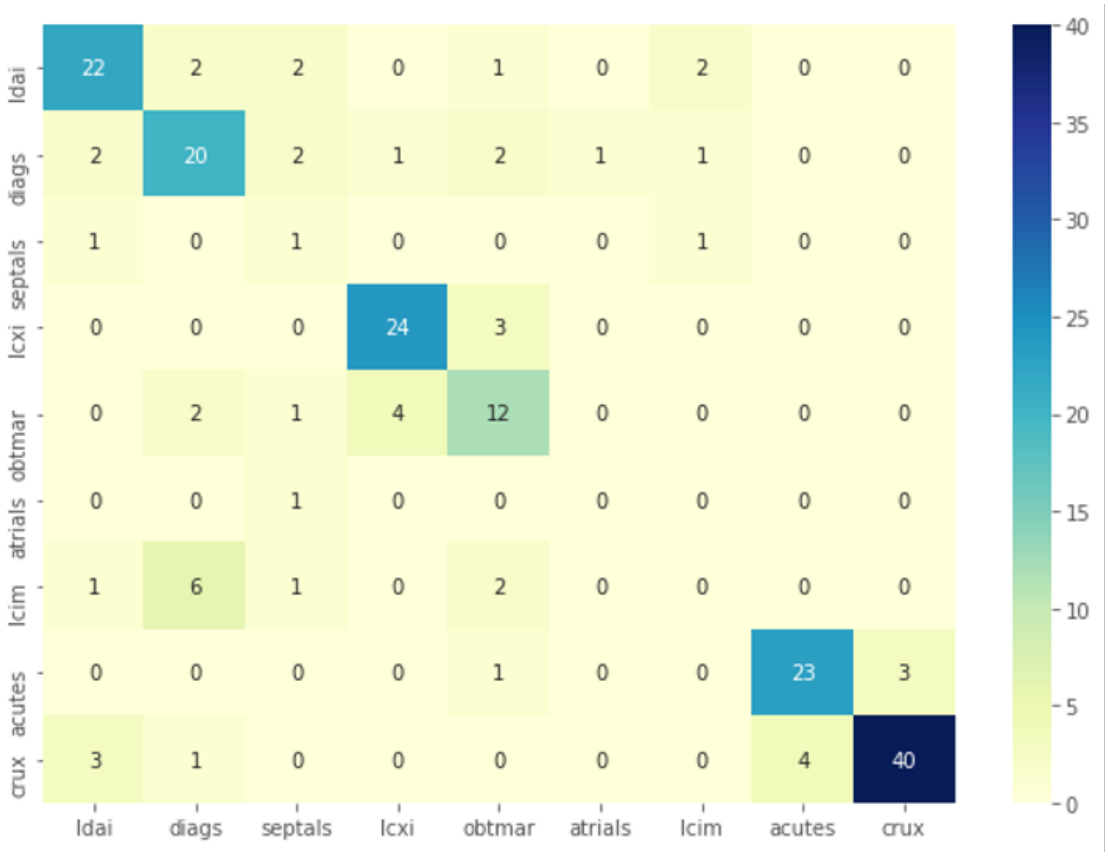

Figure 14: Confusion Matrix of Multi Output CNN Model

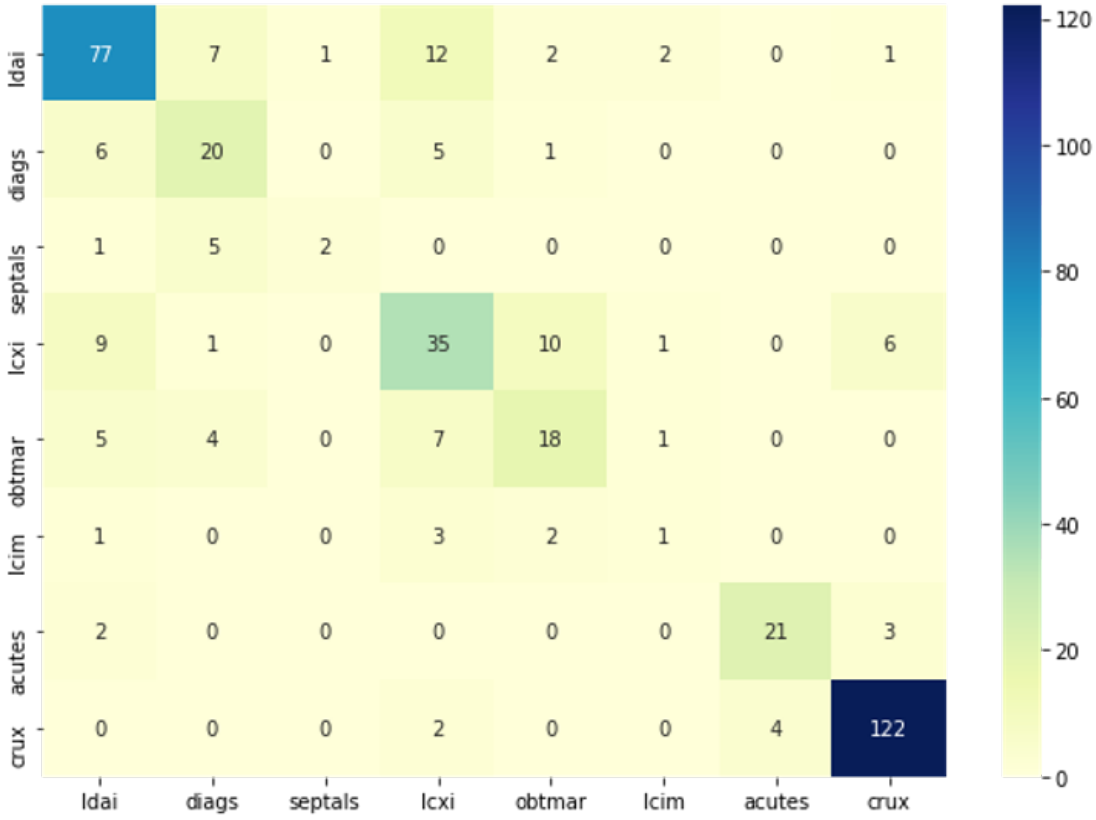

Figure 15: Confusion Matrix of my implementation of TreeLab-Net

5 Discussion

The evaluation results for single output CNN model, multi-output CNN model and TreeLab-Net are featured in Table 2. Out of the three models, single-output CNN model has the best performance. In particular, the single-output CNN model achieved at least 85% accuracy in labelling arteries that have high clinical significance such as LAD and diagonals. Other major arteries such as LCX and crux have also a high accuracy. Small side branches such as Septals and trials have low accuracy. partly due to large variance and small sample size

To further break down the classification results of all three models, confusion matrices are plotted, shown in Fig. 4 to 6. For all methods, high misclassifications are found in two groups: (1) LCXI and OM; (2) Crux and Acutes. LCXI tends to be mistaken for OM, and vice versa because OM usually bifurcate from LCXI. Anatomically, LCXI and OM are not easily distinguished spatially even by experts, as mentioned by Cao et al [9]. Similarly, crux and acutes tends to be mistaken for one another because both vessels are side branches of the Right Coronary Artery (RCA), thus having similar spatial features. Although the CNN model tends to mislabel the side branches, it is still capable of labelling the three main branch classes (i.e. LAD, LCX, and RCA) with an accuracy of 94

Another source of error in all models is possibly due to an unbalanced dataset; all classes are not represented equally. In the current dataset of 323 samples, there are very few atrial and LCIM samples compared to LAD, diagonals and crux, which might cause the neural network to neglect classes with fewer samples. This effect is shown in Table 1, where vessel classes with small sample sizes, such as atrial an LCIM, tend to perform worse.

Interestingly, having spherical coordinates as features did not yield better overall results compared to Cartesian coordinates, as shown in Table 3. As mentioned in Section, the purpose of spherical coordinate transformation is to normalise the huge variance in 3D spatial features. As observed in Table 3, it achieved a higher recall score for certain vessels such LAD and Obtuse Marginals, whereas it achieved lower recall scores for diagonals and LCX. This is because spherical coordinate transform tends to normalise two vessels that have similar spatial locations and features, for

instance a LAD and diagonals, causing both vessels to be even more difficult to distinguish. Hence, the CNN model is unable to detect the subtle differences between both vessels. Spherical coordinate transform is not recommended if the goal is to accurately label not one but all arteries.

As for my implmentation of TreeLab-Net, the accuracy is expected to be higher as its implementation is inspired by Wu et al's [11] Neural Network framework, which is the current state-of-the-art model. The discrepancy in accuracies between this research and Wu et al's is due to two factors (1) Replacement of MLP with a Graph Convolutional Layer; (2) The absence of spherical coordinate transform in my implmentation. Spherical coordinate transformation is not considered in this research due to the lack of time.

Based on the Table 1, the single-ouput CNN model in this research performs better than other literature when labelling RCA, acutes and diagonals. However, the remaining coronary arteries have a similar or worse performance. The mean precision, recall and F1-score of this research is also slightly lower than the current state-of-the-art model. Overall, the single-output CNN model is robust and effective enough to automatically label a large dataset with great variance.

6 Recommendations

Padding Variable-Sized Input

In the single-output CNN model, the input is a single coronary artery branch, which is set to be a fixed-sized tensor of 100x4. Doing so may result in data loss, as only selected points are chosen to train. Instead of feeding 100 centreline points per branch as input, it is recommended to feed all centreline points from a branch and add padding to ensure that the variable size branches have the same dimensions.

Balanced Classes

One of the reasons why side branches have poor labeling accuracy is due to imbalanced classes, as aforementioned. To mitigate this issue, it is recommended to increase the number of dataset, therefore more samples of side branches will be added. Apart from that, it is suggested by experts to generate synthetic samples by randomly sampling the attributes from instances in the minority class. A very popular algorithm to achieve this is the Synthetic Minority Over-Sampling Technique (SMOTE).

Spherical Coordinates Transformation for my TreeLab-Net

As mentioned before, a major reason why my implementation of TreeLab-Net is inferior compared to Wu et al's is due to the absence of spherical coordinate transformation. Wu et al has demonstrated that spherical coordinate transformation is capable of normalising the large variance among patient's coronary artery tree, potentially boosting the accuracy of the model.

Get Creative As Neural Network is a relatively new technology in this research domain, it is worth incorporating the current state-of-the-art technologies into my current models. For instance, Residual Network (ResNets) are networks with shortcut or skip connections and these networks have proven to be very accurate in various classification problems. Furthermore, one of the most powerful tools in deep learning – Self-Attention – has seen breakthroughs in the classification of medical images like X-Ray. Although CNN models are still the golden standard in

computer vision application, the CNN models could be guided with an Attention Layer to see if it increases the accuracy of the existing model.

In addition to experimenting different models, it is also suggested to select new features for the input. Previous works have used parent-child segment angles in 3D, length and tortuosity [10, 11]for as features in the input hence trial and error is required to see which features are most useful in increasing the classification accuracy.

7 Conclusion

In this research, a novel deep learning framework is proposed to automatically label coronary arteries from CT images. Early works have mainly focused on a knowledge based approach; extracted centreline from a patient's coronary tree is matched with a generic model via point-cloud registration. Other researches use machine learning algorithms such as Tree-Structural Bi-LSTM and Multivariate Gaussian Classifier to predict the labels. In this research, the input data consists of a tensor that contains all centreline points of coronary artery tree, where the vessel spatial locations and the radius are selected as features. The centreline points are converted to a 2D spherical coordinate system to normalize the spatial variance among different patients. Four different neural network, single-output CNN, multi-output CNN, Bi-LSTM, and my implementation of TreeLab-Net, have been developed, trained and cross-validated. The best performing model, single-output CNN model, demonstrated a mean precision, recall and F1-score of 83.3% and 83.6% and 83.2% respectively. These metrics are slightly lower than the current state-of-the-art – Wu et al's TreeLab-Net. Moreover, my single-output CNN model in this research performs better than state-of-the-art model when labelling arteries such as RCA, acutes and diagonals. The labelling accuracy of the remaining coronary arteries are on par or worse than the state-of-the-art model. Further action will be undertaken to augment the prediction accuracy of all branches in a coronary tree, especially the side branches. This includes using a padding a variable-sized input, having balanced classes, developing new models and selecting new features for the input. The current results demonstrated that the single-output CNN model

is robust and effective for anatomical labelling problems in coronary artery tree, and this application can be further extended for other tree-like anatomies such as airway trees, colon arteries and cerebral arteries.

8 References

[1] C. J. McAloon, L. M. Boylan, T. Hamborg, N. Stallard, F. Osman, P. B. Lim and S. A. Hayat, 'The changing face of cardiovascular disease 2000–2012: An analysis of the world health organisation global health estimates data', *International Journal of Cardiology*, vol. 224, pp. 256–264, Dec. 2016, ISSN: 01675273. DOI: 10.1016/j.ijcard.2016.09.026. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167527316321842 (visited on 12/04/2020).

[2] M. C. Williams, J. H. Reid, G. McKillop, N. W. Weir, E. J. R. van Beek, N. G. Uren and D. E. Newby, 'Cardiac and coronary CT comprehensive imaging approach in the assessment of coronary heart disease', *Heart*, vol. 97, no. 15, p. 1198, 1 Aug. 2011. DOI: 10.1136/heartjnl-2011-300037. [Online]. Available: http://heart.bmj.com/content/97/15/1198.abstract.

[3] J. T. Dodge, B. G. Brown, E. L. Bolson and H. T. Dodge, 'Intrathoracic spatial location of specified coronary segments on the normal human heart. applications in quantitative arteriography, assessment of regional risk and contraction, and anatomic display.', *Circulation*, vol. 78, no. 5, pp. 1167–1180, Nov. 1988, ISSN: 0009-7322, 1524-4539. DOI: 10.1161/01.CIR.78.5.1167. [Online]. Available: https://www.ahajournals.org/doi/10.1161/01.CIR.78.5.1167 (visited on 12/04/2020).

[4] C. Chalopin, G. Finet and I. E. Magnin, 'Modeling the 3d coronary tree for labeling purposes', *Medical Image Analysis*, vol. 5, no. 4, pp. 301–315, Dec. 2001, ISSN: 13618415. DOI: 10.1016/S1361-8415(01)00047-0. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1361841501000470 (visited on 12/04/2020).

[5] N. Ezquerra, S. Capell, L. Klein and P. Duijves, 'Model-guided labeling of coronary structure', *IEEE Transactions on Medical Imaging*, vol. 17, no. 3, pp. 429–441, Jun. 1998, ISSN: 02780062. DOI: 10.1109/42.712132. [Online]. Available: http://ieeexplore.ieee.org/document/712132/ (visited on 12/04/2020).

[6] G. Yang, A. Broersen, R. Petr, P. Kitslaar, M. A. de Graaf, J. J. Bax, J. H. Reiber and J. Dijkstra, 'Automatic coronary artery tree labeling in coronary computed tomographic angiography datasets', p. 4,

[7] Q. Cao, A. Broersen, M. A. de Graaf, P. H. Kitslaar, G. Yang, A. J. Scholte, B. P. F. Lelieveldt, J. H. C. Reiber and J. Dijkstra, 'Automatic identification of coronary tree anatomy in coronary computed tomography angiography', *The International Journal of Cardiovascular Imaging*, vol. 33, no. 11, pp. 1809–1819, Nov. 2017, ISSN: 1569-5794, 1573-0743. DOI: 10.1007/s10554-017-1169-0. [Online]. Available: http://link.springer.com/10.1007/s10554-017-1169-0 (visited on 12/04/2020).

[8] M. A. Gülsün, G. Funka-Lea, Y. Zheng and M. Eckert, 'CTA coronary labeling through efficient geodesics between trees using anatomy priors', in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014*, P. Golland, N. Hata, C. Barillot, J. Hornegger and R. Howe, Eds., vol. 8674, Cham: Springer International Publishing, 2014, pp. 521–528, ISBN: 978-3-319-10469-0 978-3-319-10470-6. DOI: 10.1007/978-3-319-10470-6_65. [Online]. Available: http://link.springer.com/10.1007/978-3-319-10470-6_65 (visited on 12/04/2020).

[9] A. Akinyemi, S. Murphy, I. Poole and C. Roberts, 'Automatic labelling of coronary arteries', p. 5,

[10] D. Wu, X. Wang, J. Bai, X. Xu, B. Ouyang, Y. Li, H. Zhang, Q. Song, K. Cao and Y. Yin, 'Automated anatomical labeling of coronary arteries via bidirectional tree LSTMs', *International Journal of Computer Assisted Radiology and Surgery*, vol. 14, no. 2, pp. 271–280, Feb. 2019, ISSN: 1861-6410, 1861-6429. DOI: 10.1007/s11548-018-1884-6. [Online]. Available: http://link.springer.com/10.1007/s11548-018-1884-6 (visited on 12/04/2020).

[11] H. Yang, X. Zhen, Y. Chi, L. Zhang and X.-S. Hua, 'CPR-GCN: Conditional partial-residual graph convolutional network in automated anatomical labeling of coronary arteries', *arXiv:2003.08560 [cs, eess]*, 17 Apr. 2020. arXiv: 2003.08560. [Online]. Available: http://arxiv.org/abs/2003.08560 (visited on 21/10/2020).

[12] S. Saha. (17 Dec. 2018). A comprehensive guide to convolutional neural networks — the ELI5 way, Medium. Library Catalog: towardsdatascience.com, [Online]. Available: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53 (visited on 19/04/2020).

9 Appendices