CPR-GCN: Conditional Partial-Residual Graph Convolutional Network in Automated Anatomical Labeling of Coronary Arteries

Han Yang¹,*, Xingjian Zhen²,*, Ying Chi¹,†, Lei Zhang¹, and Xian-Sheng Hua¹
¹DAMO Academy, Alibaba Group  ²University of Wisconsin Madison
{xiuxian.yh,xinyi.cy,lei.zhang.lz,xiansheng.hxs}@alibaba-inc.com, xzhen3@wisc.edu

Abstract

Automated anatomical labeling plays a vital role in coronary artery disease diagnosing procedure. The main challenge in this problem is the large individual variability inherited in human anatomy. Existing methods usually rely on the position information and the prior knowledge of the topology of the coronary artery tree, which may lead to unsatisfactory performance when the main branches are confusing. Motivated by the wide application of the graph neural network in structured data, in this paper, we propose a conditional partial-residual graph convolutional network (CPR-GCN), which takes both position and CT image into consideration, since CT image contains abundant information such as branch size and spanning direction. Two majority parts, a partial-residual GCN and a conditions extractor, are included in CPR-GCN. The conditions extractor is a hybrid model containing the 3D CNN and the LSTM, which can extract 3D spatial image features along the branches. On the technical side, the partial-residual GCN takes the position features of the branches, with the 3D spatial image features as conditions, to predict the label for each branches. While on the mathematical side, our approach twists the partial differential equation (PDE) into the graph modeling. A dataset with 511 subjects is collected from the clinic and annotated by two experts with a two-phase annotation process. According to the five-fold cross-validation, our CPR-GCN yields 95.8% meanRecall, 95.4% meanPrecision and 0.955 meanF1, which outperforms state-of-the-art approaches.

1. Introduction

Cardiovascular disease is one of the leading causes of death worldwide [21]. Cardiac CT angiography (CCTA) image is widely adopted for the diagnosis of the cardiovascular disease because of its non-invasion and high sensitivity [15]. In the clinic, doctors need a series of manual labor to obtain the diagnostic report, which is a time-consuming effort. If a computer-aided diagnosis (CAD) system can generate the diagnostic report automatically, a huge amount of time can be saved. While automated anatomical labeling of the coronary artery tree extracted from CCTA image is a prerequisite step in the automated CAD system.

The coronary artery tree consists of two components, i.e., left domain (LD) and right domain (RD). Also they both originate from the aorta. According to [23], the main coronary arteries of interest are left main (LM), left descending artery (LAD), left circumflex artery (LCX), left ramus-intermedius (RI), obtuse margin (OM), diagonal artery (D), septal artery (S), right coronary artery (RCA), right posterior lateral branches (R-PLB), right posterior descending artery (R-PDA), right acute marginal artery (AM) (shown in Figure 1). From prior knowledge, we know that LAD, LCX, and RI are from LM. R-PLB, R-PDA, and AM are from RCA. Also S’s and D’s are from LAD, while OM’s are from LCX. This makes the whole coronary artery tree as structured data. Normally, RCA, LM, LAD, and LCX are treated as the main branches. Other branches are treated as the side branches.

Several anatomical labeling techniques have been developed for coronary arteries [23, 3, 22], brain arteries [1], abdominal arteries [18, 25] and airways [10]. However,
Our CPR-GCN and the hybrid model (i.e., 3D CNNs following BiLSTMs) can be jointly trained. We evaluate the CPR-GCN on a large private collected dataset. Our approach outperforms the state-of-the-art result.

2. Related Work

Most of the related work can be divided into two categories, i.e., traditional based and deep learning based. Traditional methods are based on the knowledge and the topology of the coronary arteries. Normally, they require two steps of registration and correction. With the development of deep learning, there are also some methods on the graph-based structural data. These methods extract features as nodes and train the model with the data they acquire. Also these methods heavily depend on the size and quality of the dataset.

2.1. Traditional Methods

Most traditional methods are based on registration. In [23], the author presented a two-step method. In the registration step, the main branches, LM, LAD, LCX, and RCA, are identified. Then the rest branches are matched afterward. In [3], they built the 3D models for both right dominant and left dominant. The 3D coronary trees from subjects are aligned with the 3D models to get the label of each segment. They also applied the logical rules to fulfill the clinical experience.

However, traditional methods highly rely on the main branches. If the main branches are missing in the automated segmentation system, the performance will deteriorate dramatically. Also, they require prior knowledge about how the coronary tree span. Thus if some sub-branches are missing, it will affect the topology and the performance. Finally, all these traditional methods have the human-interpretation, which means that if the topology is too complicated, the automatic system will raise the information that it is not capable of determining.

2.2. Deep Learning Based Methods

With the power of deep learning, the TreeLab-Net is developed in [22]. They combined the multi-layer perception encoder network and the bidirectional tree-structural LSTM to construct the TreeLab-Net. They used several selected features from positions and stacked the Child-Sum Tree LSTMs as the components. The left and right coronary arteries are trained independently.

In this method, the missing branches will cause a massive problem since this will change the layer index of the nodes inside the tree. With the tree-structured model, the message is only passing between nearby layers. The closer the branches are to the root nodes, the higher impact they will raise. The left and right are also classified with the rules and the thresholds. This method has a strong assumption that the branches will only bifurcate so that each node in the
tree-structure can only have two children. But in the coronary arteries, it’s quite reasonable that several sub-branches bifurcate from the points near to each other. So the node for the parent branch will return to have more than two children, which is beyond the capacity of the TreeLab-Net.

In a broad sense, the TreeLab-Net is a simplified version of the graph models. In [14], the author brought out the Graph Convolutional Networks (GCN) which operate directly on graphs. By projecting the graphs into the Fourier domain, the author defined the convolution operator and filter kernels in the Fourier domain using the Chebyshev polynomials. In [17], they modified the Graph Neural Networks (GNN) [19] to use gated recurrent units.

Even though these graph models are successful in molecular fingerprints [7] and protein interface prediction [8], they have not been used in labeling coronary arteries. Also graph models also suffer from shallow structure problems. Since stacking multiple GCN layers will result in over-smoothing [16].

### 3. Our Approach

In this section, we detail the CPR-GCN, which makes full use of both CCTA images and the position of the coronary artery centerlines. As shown in Figure 2, our method considers both CCTA images and the features from positions mentioned in [22]. The centerlines are extracted using the automated coronary artery tracking system [24]. Then, our CPR-GCN extracts the features from the centerlines within SCT($S^2$) block. Also, in the image domain, we use the sub-sampled control points on centerlines to get the moving cubes with a fixed radius $\gamma$ along each branches as the image domain data. The conditions for our CPR-GCN model are obtained with the 3D CNN and the BiLSTM. The detailed architecture of the CPR-GCN model is shown in Table 1.

#### 3.1. Position Domain Features

In [22], the author introduced a spherical coordinate transform 2D (SCT2D) which transforms the positions $P_k = [(x_i, y_i, z_i)]^\text{Length}_{h_k}$ in 3D into the azimuth and elevation angles $[(\varphi_i, \theta_i)]^\text{Length}_{h_k}$. They argued that this could normalize the variance of centerlines in the default Cartesian coordinate system. However, it is noticeable that $\varphi$ and $\theta$ have the periodic $2\pi$. So with limiting the range of the angle to be $[0, 2\pi)$, the small amount of shaking, due to the noise, will return to be $2\pi$ difference near the angle 0.

The similar idea of using spherical coordinate transform is applied in our approach. Since each branches are processed separately, so as to get the azimuth and elevation angles, we need to define the origin and the $x, y, z$ axes for each branches. For each branches, the first control point is chosen as the origin. The direction pointing from the first point to the second point is defined as $z$ axis. The vector from the first point to the last point of the centerline lies in $y-z$ plane.

To overcome the instability due to the periodic, we use the $S^2$ manifold to represent $\varphi, \theta$. The $S^2$ manifold is the
sphere with unit radius in $\mathbb{R}^3$. A trivial method is to use the $2 \times 2$ matrix $M = \begin{bmatrix} \sin\theta & \sin\varphi \\ \cos\theta & \cos\varphi \end{bmatrix}$. Because of the periodicity of $\sin(\cdot)$ and $\cos(\cdot)$, the matrix $M$ is stable on the whole manifold $S^2$. This kind of spherical coordinate transformation is called SCTS$^2$ in the rest of this paper. The Cartesian coordinate and the $S^2$ manifold transformation is in Eq. 1.

$$
\begin{align*}
x &= r \sin\theta \cos\varphi \\
y &= r \sin\theta \sin\varphi \\
z &= r \cos\theta
\end{align*}
$$

We use the similar features mentioned in [22]: (1) The $S^2$ projection and the normalized 3D positions of the first point, center point, and the last point. (2) Directional vector between first and last points and the tangential direction at the start point in both 3D and $S^2$.

### 3.2. Image Domain Conditions

Most of the medical images, including Magnetic Resonance Imaging (MRI) and the CCTA we use, lie in 3D. The branches, unlike other images, have the sequential dependency. Thus we use the 3D CNN to extract the spatial features and use the BiLSTM afterward to summarize the tubular sequential features. An example of processing the R-PLB branch is shown in Figure 3. The $z$ dimension of CCTA images are the slices, which might have different spacing from $x$ and $y$ dimensions. So we resample all the CCTA images to have the same spacing $v$ along $x$, $y$, $z$ dimensions.

Using the automatic segmentation method, we can get the centerlines of all branches ($P_1$, $P_2$, ..., $P_n$) which build into a coronary artery tree. We separate the branches where the centerlines end or bifurcate. If the starting points of the children branches are close to each other, we consider all these children branches are from the same point on the parent branch. The control points of the centerlines are smoothed using the Catmull-Rom spline [20]. Finally, the control points are sub-sampled with the same length.

The image domain data is the cubes $I_k$ with the fixed radius $\gamma$ around each control points of $P_k$. Three layers of 3D CNN and 3D maxpooling are used to extract the features of $I_k$. The weights of CNN are shared among the segments. In order to train the model in a mini-batch manner, these feature vectors are padded to the $\max_{k=1}^n$ as the input of the multi-layer bidirectional LSTM [9]. The last hidden state is treated as the final conditions $y$, which represents the image information of this branch. We treat this as the conditional information since the images are of less importance than the position domain features.

### 3.3. Partial-Residual Block of GCN

The layer-wise propagation rule for the traditional GCN is Eq. 2. In the multi-layer GCN, the features $X$ of nodes is the input for the first layer, $X \in \mathbb{R}^{n \times d_0}$. Here, the $n$ is the number of nodes. The $d_0$ is the dimension of the features for each nodes. The $A$ is the adjacency matrix for the graph. The $W^{l}$ is the layer-wise trainable weights, $W^l \in \mathbb{R}^{d_l \times d_{l+1}}$. The $\sigma(\cdot)$ is the activation function. In this paper, we choose ReLU(·) as our activation function to include the nonlinear ability.

$$
H^{l+1} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^l W^{l})
$$

The $\tilde{A}$ is the adjacency matrix $A$ added the self-loop identity matrix $I_N$. The input for the first layer is $H^0 = X$.

Our conditional partial-residual block requires both the position features $x$ and the CCTA image domain conditions $y$. The combination of the features and conditions from two domains is used as the representative of the nodes in the graph model. The edges are defined as the parent-children relationship. As we mentioned above, the topology of the whole coronary tree is collected from ($P_1, P_2, ..., P_n$). Whenever the branches bifurcate, we treat these parent-child branches as three (or more) nodes. The edges are from the parent branch to the children branches. This will build the graph of the subject with the adjacency matrix
As shown in Figure 4, the RCA first bifurcates the AM and then bifurcates the R-PLB and R-PDA. So the extracted graph has 4 edges and 5 nodes.

In [11], the author argued that the deep residual learning framework (Figure 5 (a)) can help in the performance. Instead of directly learning the map \( \mathcal{H}(x) \), the neural network learns the residual part \( \mathcal{F}(x) := \mathcal{H}(x) - x \) assuming the input and the output having the same dimension. When changing the output dimension, it’s a straightforward method to add the linear projection \( W_s \) by the shortcut connection:

\[
z = \mathcal{F}(x, W_i) + W_s x \quad (3)
\]

Also in [2], the authors extended the idea of residual connection into the Residual Gated Graph ConvNets. The updated propagation rule is:

\[
H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^l W^l) + H^l \quad (4)
\]

If we view the residual connection in Eq. 4 as a continuous function of \( l \) and add enough number of layers. In the limit, [6] parameterized the continuous dynamics of hidden units using the ordinary differential equation (ODE):

\[
dH(l) = f_A(H(l), W(l), l) \quad (5)
\]

In our setup, we have two kinds of input: position domain features \( x(l) \) and the CCTA image domain conditions \( y \). If we treat \( l \) as the layer index number and \( x \) is the function of layer \( l \) in Eq. 5, we have the partial differential equation (PDE) on \( x, y \):

\[
\nabla H^l(x(l), y) = \frac{\partial H^l(x, y)}{\partial x} \, dx + \frac{\partial H^l(x, y)}{\partial y} \, dy = \frac{\partial H^l(x, y)}{\partial x} \, dx + \frac{\partial H^l(x, y)}{\partial y} \, dy = \frac{\partial H^l(x, y)}{\partial x} \, dl \quad (6)
\]

\[
= GCN_A(x, y)dl \quad (7)
\]

Eq. 7 is based on the fact that we treat \( y \) as the conditions. We use the trainable \( GCN_A(x, y) \) to approach the partial differential \( \frac{dH^l(x, y)}{dx} \). If we take \( H^0 = x(0) = X \) and approximate \( dl = 1 \), we have

\[
H^1 = \nabla H^1(x, y)dl + H^0
\]

\[
= GCN_A(x, y) + X \quad (9)
\]

as the discrete numerical estimates. In our case, as shown in Figure 5 (b), the discrete partial-residual block of the CPR-GCN takes the weighted \( X \) because of the change of the channel sizes. If we push our PDE a bit further, we can have:

\[
H^k = \int \nabla H^k(x, y)dx + H^0
\]

\[
= \int_0^k GCN_A(x, y)dl + X \quad (10)
\]

Also, the \( X \) here should be weighted for the flexibility of the channels.

3.4. Data Flow

The algorithm of our CPR-GCN is shown in Alg. 1. CCTA image \( I \), centerlines \( P_1, P_2, \ldots, P_n \) and reference labels \( GT_1, GT_2, \ldots, GT_n \) compose the training sample of the model. We first build the graph \( A \) via the bifurcation of \( P_1, P_2, \ldots, P_n \). The position domain feature \( x \) and the image domain feature \( y \) are extracted via SCTS\(^2\) and a hybrid network (i.e., 3D CNN following BiLSTMs). In addition, we concatenate \( x \) and \( y \) as the input of GCN layers with \( x \) as the shortcut in residual connection. At last, a fully connected layer predicts the final labeling \( L_1, L_2, \ldots, L_n \) and our object is to minimize the cross-entropy of these two distributions, i.e., \( GT_1, GT_2, \ldots, GT_n \) and \( L_1, L_2, \ldots, L_n \).

**Algorithm 1:** Training procedure of our approach

<table>
<thead>
<tr>
<th>Data: CCTA image ( I ), centerlines ( P_1, P_2, \ldots, P_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground truth: ( GT_1, GT_2, \ldots, GT_n )</td>
</tr>
<tr>
<td>( A \leftarrow \text{Build Graph}(P_1, P_2, \ldots, P_n); )</td>
</tr>
<tr>
<td>( x_1, x_2, \ldots, x_n \leftarrow \text{SCTS}^2(P_1, P_2, \ldots, P_n); )</td>
</tr>
<tr>
<td>( I_1, I_2, \ldots, I_n \leftarrow I \odot (P_1, P_2, \ldots, P_n); )</td>
</tr>
<tr>
<td>( y_1, y_2, \ldots, y_n \leftarrow \text{BiLSTM} \leftarrow 3D \text{CNN}(I_1, I_2, \ldots, I_n); )</td>
</tr>
<tr>
<td>( H \leftarrow GCN_A(x, y) + x; )</td>
</tr>
<tr>
<td>( L_1, L_2, \ldots, L_n \leftarrow FC(H); )</td>
</tr>
</tbody>
</table>
| loss\(
=\text{Cross_Entropy}(L, GT); \)

4. Experimental Results

4.1. Dataset and Evaluation Metrics

To our best knowledge, a public dataset with CCTA image and annotation of coronary artery labeling is not available until now. Previous works [23, 3, 22] all collected a private experimental dataset from clinic. For instance, conventional methods [23, 3] only used 58 and 83 subjects respectively, while deep learning based method [22] used a
larger dataset with 436 subjects. In this study, we collected the largest relevant dataset from clinic. All vessel centerlines are first extracted using [24]. This dataset contains 511 subjects and all of them are annotated by two experts with a two-phase annotation process. These two experts give a label to every branch alone in the first round. Then the annotation results are merged and experts take discussions on inconsistent ones to obtain a final label. These 511 subjects and corresponding annotation compose our experimental dataset. The average number of branches in each subject is 9.65, with standard deviation (std) 2.13. The largest number of branches is 15, and the smallest number is 3. After bifurcated and separated, the average number of segments is 13.23. The detail is mentioned in Table 2. The edges in the table represent the relationship in the graph we build for each subject. The average number of edges in each graph is 11.18.

The evaluation is performed on all branch segments by the predicted label and the ground truth label. The recall rate for each segments is calculated by $Recall = \frac{tp}{tp+fp}(\text{label} = i)$. The precision is $Precision = \frac{tp}{tp+fn}(\text{label} = i)$. The F1 score is $F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}(\text{label} = i)$. Since the numbers of segments in are imbalance, we also use the mean metrics of all classes for the segments. meanRecall = $\frac{1}{n} \sum_{i=1}^{n} Recall$. It is similar for the other metrics.

### 4.2. Implementation Details

**Hyper-parameters selection** The CCTA images are scaled to $v = 0.5$ mm voxel spacing. Since the radius of branches usually ranges in $(0,3)$ mm, i.e., $(0,6)$ voxels. So the radius of the cube is chosen to be $\gamma = 12$ voxels to keep the angel, size, and the texture information. Thus the subsampling rate for the centerline positions is 10 voxels to have overlapping as well as keep the sequential information along the branches.

**Training** The dataset is randomly and equally divided into five subsets. In the training stage, we use a five-fold cross-validation strategy to evaluate all subjects in the dataset. The proposed CPR-GCN model has two trainable components, i.e., 3D CNN following LSTMs (3D CNN module), GCNs following a FC layer (GCN module). A series of 3D image cubes extracted along the vessel centerlines from the 3D CCTA image are the input of the 3D CNN module. The position domain features extracted from the vessel tree via SCTR2 are concatenated with the output of the 3D CNN module. The GCN module takes these combined features as input and predicts the label for every segment. In addition, the reference labels are needed to compute the cross-entropy loss with the predicted labels. The 3D image cubes, position domain features and the reference labels compose the training samples. Our CPR-GCN model is trained in an end-to-end manner.

The algorithm is implemented using PyTorch with an NVIDIA Tesla P100 GPU. We use Adam optimizer [13] with an initial learning rate of 0.001. Each mini-batch contains 8 coronary artery trees. For each training period, we train the CPR-GCN model up to 200 epochs which takes 2.7 hours. So the total training time for five-fold cross-validation is 13.5 hours.

**Testing** We first choose the best model in each fold according to the overall testing precision ignored classes. Then we use each model to evaluate the corresponding testing data. During inference, the average time spent on the CPR-GCN model is 0.045 s per case, which is greatly important in the clinical utilization.

**Results** Since there is not a public dataset in this field and conventional methods only evaluated their performance on a small private dataset. We also reproduced the conventional method and the deep learning based TreeLab-Net. Considering conventional methods [23, 3] mainly rely on registration and prior knowledge, we only reproduce [3] which has a improvement on [23]. Table 3 reports the detail performance on our dataset. Our CPR-GCN achieves the highest meanRecall of 0.958, meanPrecision of 0.954 and meanF1 of 0.955, which outperforms other methods with a large margin.

All the models have performed well on the main branches. But the side branches are also a crucially important part of the automated anatomical labeling in the CAD system. In our approach, we treat the main branches and side branches equally. So compared with other two-step methods, most of the side branches, such as OM and R-PDA, performs better in our approach. But since the number of segments of the main branches like LM and RCA is relatively large, the performance of these main branches is better. For the side branches, especially for the D and OM, the number of samples in the dataset is relatively small. So the performance is slightly worse than the main branches.

### 4.3. Ablation Study

To make sure that all the components of the CPR-GCN perform well, we design the ablation study experiments.

**Image domain condition** One of the major differences between our approach and the other methods is that we use the extra information from CCTA images domain. As shown in

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Avg.(std)</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branches</td>
<td>4929</td>
<td>9.65 (2.13)</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>Segments</td>
<td>6760</td>
<td>13.23 (3.55)</td>
<td>22</td>
<td>3</td>
</tr>
<tr>
<td>Edges</td>
<td>5714</td>
<td>11.18 (3.56)</td>
<td>20</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2. The basic information for the dataset we use. Segments are the branches after bifurcating. The edges are the relationship between segments.
Table 3. Comparisons of conventional method [3], deep learning based TreeLab-Net [22] and our CPR-GCN on our dataset. Recall, precision and F1 score are used as the evaluation metrics.

<table>
<thead>
<tr>
<th>Method</th>
<th>Metric</th>
<th>RCA</th>
<th>R-PDA</th>
<th>R-PLB</th>
<th>AM</th>
<th>LM</th>
<th>LAD</th>
<th>LCX</th>
<th>RI</th>
<th>D</th>
<th>OM</th>
<th>S</th>
<th>Avg(std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional [3]</td>
<td>Recall</td>
<td>0.918</td>
<td>0.850</td>
<td>0.852</td>
<td>0.893</td>
<td>0.984</td>
<td>0.911</td>
<td>0.832</td>
<td>0.848</td>
<td>0.799</td>
<td>0.720</td>
<td>0.835</td>
<td>0.859±0.066</td>
</tr>
<tr>
<td></td>
<td>Precision</td>
<td>0.925</td>
<td>0.835</td>
<td>0.860</td>
<td>0.871</td>
<td>0.991</td>
<td>0.929</td>
<td>0.810</td>
<td>0.803</td>
<td>0.781</td>
<td>0.739</td>
<td>0.865</td>
<td>0.855±0.069</td>
</tr>
<tr>
<td></td>
<td>F1</td>
<td>0.922</td>
<td>0.842</td>
<td>0.856</td>
<td>0.882</td>
<td>0.987</td>
<td>0.920</td>
<td>0.821</td>
<td>0.825</td>
<td>0.789</td>
<td>0.730</td>
<td>0.850</td>
<td>0.857±0.067</td>
</tr>
<tr>
<td>TreeLab-Net [22]</td>
<td>Recall</td>
<td>0.950</td>
<td>0.858</td>
<td>0.818</td>
<td>0.871</td>
<td>0.996</td>
<td>0.948</td>
<td>0.913</td>
<td>0.770</td>
<td>0.816</td>
<td>0.805</td>
<td>0.862</td>
<td>0.873±0.067</td>
</tr>
<tr>
<td></td>
<td>Precision</td>
<td>0.948</td>
<td>0.823</td>
<td>0.842</td>
<td>0.871</td>
<td>0.970</td>
<td>0.937</td>
<td>0.936</td>
<td>0.714</td>
<td>0.841</td>
<td>0.807</td>
<td>0.859</td>
<td>0.868±0.072</td>
</tr>
<tr>
<td></td>
<td>F1</td>
<td>0.949</td>
<td>0.840</td>
<td>0.830</td>
<td>0.871</td>
<td>0.983</td>
<td>0.942</td>
<td>0.924</td>
<td>0.741</td>
<td>0.829</td>
<td>0.807</td>
<td>0.860</td>
<td>0.871±0.069</td>
</tr>
<tr>
<td>Our CPR-GCN</td>
<td>Recall</td>
<td>0.994</td>
<td>0.930</td>
<td>0.944</td>
<td>0.991</td>
<td>0.994</td>
<td>0.990</td>
<td>0.982</td>
<td>0.921</td>
<td>0.936</td>
<td>0.896</td>
<td>0.954</td>
<td>0.958±0.033</td>
</tr>
<tr>
<td></td>
<td>Precision</td>
<td>0.987</td>
<td>0.946</td>
<td>0.947</td>
<td>0.983</td>
<td>0.984</td>
<td>0.986</td>
<td>0.971</td>
<td>0.896</td>
<td>0.883</td>
<td>0.933</td>
<td>0.974</td>
<td>0.954±0.035</td>
</tr>
<tr>
<td></td>
<td>F1</td>
<td>0.990</td>
<td>0.938</td>
<td>0.945</td>
<td>0.987</td>
<td>0.989</td>
<td>0.988</td>
<td>0.976</td>
<td>0.909</td>
<td>0.909</td>
<td>0.914</td>
<td>0.964</td>
<td>0.955±0.032</td>
</tr>
</tbody>
</table>

the second column of Table 5, if we only remove the image domain conditions, the metrics, i.e., meanRecall, meanPrecision and meanF1 will drop over 2.0%. For instance, the meanF1 score is 0.934 compared with 0.955 in our CPR-GCN.

Residual GCN connection Also, our approach brought the partial-residual connection in the graph model. In this part, we only remove the residual connection in the GCN block and keep other setting the same as the CPR-GCN. The third column in Table 5 reports that the F1 score will drop to 0.947. This argues that with the help from residual, our model can absorb the features from both position and image domain as well as keep the original position domain features.

Undirected graph In this part, we build the undirected graph, which means adding the opposite edges from the original graph. Detailed results are illustrated in Table 5, the fourth column. Although the average metrics among classes (i.e, meanRecall, meanPrecision and meanF1) are slightly worse than our best result, several classes (e.g, LM, LAD, R-PDA, etc.) for undirected graph have higher precision than directed graph. It is worth to note that we select directed graph to achieve higher average metrics. In clinical practice, we can make choice according to the requirement of the doctors.

Repeated GCN blocks [16] reports that stacking multiple GCN layers will result in over-smoothing. We also conduct the experiment to answer how the number of GCN blocks influences the performance. Considering the connection complexity in coronary tree graph, we respectively evaluate our CPR-GCN with 1, 2, 3, 4 GCN layers. Figure 6 illustrates that we can improve the performance by stacking GCN blocks, especially from 1 GCN block to 2 GCN blocks. However, the model with 4 GCN blocks has no obvious improvement compared with 3 GCN blocks. Therefore we use 3 GCN blocks in our CPR-GCN. The depth of the graph for the coronary artery tree mostly is less than 4, which might lead to the evaluation results.

4.4. Synthetic "Data Attack"

We hold the opinion that our CPR-GCN is more robust when main branch in the vessel tree is missing. So we first build a synthetic dataset from our original dataset. 20% LM and RCA branches is randomly removed. Most of the other side branches (e.g., LCX, LAD, RI, AM, etc.) directly originate from those two branches. In this new synthetic dataset, 295 RCA and LM branches are removed. 1123 of 6760 vessel segments directly connect with these 295 missing branches. Since conventional methods [23, 3] strictly rely on the main branches. We only evaluate the trained CPR-GCN and TreeLab-Net [22] on this synthetic dataset. As shown in Table 4, our CPR-GCN drops almost 2.6% while the TreeLab-Net drops almost 6.7% in three average metrics. This demonstrates that our method is more robust.

Table 4. Experimental results of TreeLab-Net [22] and our CPR-GCN on the original dataset and synthetic dataset. 20% RCA and LM is randomly removed in the synthetic dataset.

<table>
<thead>
<tr>
<th>Method</th>
<th>Dataset</th>
<th>meanRecall</th>
<th>meanPrecision</th>
<th>meanF1</th>
</tr>
</thead>
<tbody>
<tr>
<td>TreeLab-Net [22]</td>
<td>Original</td>
<td>0.873</td>
<td>0.868</td>
<td>0.871</td>
</tr>
<tr>
<td></td>
<td>Synthetic</td>
<td>0.806</td>
<td>0.799</td>
<td>0.802</td>
</tr>
<tr>
<td>Our CPR-GCN</td>
<td>Original</td>
<td>0.958</td>
<td>0.954</td>
<td>0.955</td>
</tr>
<tr>
<td></td>
<td>Synthetic</td>
<td>0.931</td>
<td>0.928</td>
<td>0.929</td>
</tr>
</tbody>
</table>

Figure 6. Ablation study of repeated GCN blocks. 1 GCN, 2 GCN, 3 GCN and 4 GCN represent stacking corresponding number of the GCN blocks.
Our CPR-GCN can make use of the enormous size of data. Robustness

Our CPR-GCN is purely driven by data. So as to stress the importance of the features from position domain, we introduced the partial-residual block on the position domain features.

Image domain information  As far as we know, we are the first to include the image domain information as the conditions for our approach. So as to stress the importance of the features from position domain, we introduced the partial-residual block on the position domain features.

Partial-Residual block  One of the main contributions of our approach is that we brought the partial-residual block. In the traditional residual block, all dimensions of the input \( x \) are treated equally. But in the coronary arteries labeling problem, the positions are proven to play an important role. So as to stress this importance as well as use the extra information inside the CCTA images, we use the partial-residual block to treat the image domain in formations as the extra conditions of the model. With the ablation study, we notice that this kind of structure can improve the metric from 0.947 to 0.955. Also, it can make the model more stable.

Robustness  Our CPR-GCN is purely driven by data. So the CPR-GCN can make use of the enormous size of data.

Without prior knowledge and the hard-coded "rule", the CPR-GCN is more robust to the noises. All the nodes in our graph, which represent different segments, weigh the same in the CPR-GCN. So the wrong classification of LM or other major branches has less chance to spread through all other branches. In order to prove this opinion, we conduct an interesting synthetic "Data Attack" experiment. The results show that missing main branch has less impact on our CPR-GCN than other deep leaning method.

Disadvantages and future work  There are still some future works for this problem. Since we treat every branches equally, the imbalance among branches is an issue. It's noticeable that the main branches perform better than side branches due to the number of samples are different. There will also be some tiny branches missing after the segmentation, which will increase the imbalance. Also in this model, we use the discrete approximation of the PDE. In the future, the model can be pushed forward to the continuous model.

6. Conclusion

In this paper, we propose the end-to-end conditional partial-residual graph convolutional network (CPR-GCN) model for automated anatomical labeling of coronary arteries, where few alternatives are available. Compared with the traditional methods and the recent deep learning based methods, our approach achieves the state-of-the-art result. We show that with the conditional partial-residual block, both information in position domain and CCTA image domain can be taken into consideration. On the experiment side, we show that our CPR-GCN is more robust and flexible compared with others. The result also shows that the CCTA image domain matters in coronary arteries labeling. Importantly, we show that our algorithmic contributions facilitate the CAD system.
References


