

冠状动脉先验建模与训练数据表示

陈奕好

2025 年 11 月 24 日

1 冠状动脉数据与树结构表示

设有若干病例（病人），用索引 $i = 1, \dots, N_{\text{case}}$ 表示。对第 i 个病例，给定：

- CT 体数据

$$V^{(i)} : \Omega^{(i)} \subset \mathbb{R}^3 \rightarrow \mathbb{R},$$

其中 $\Omega^{(i)}$ 为该病例的体素域；

- 冠状动脉分割（GT 掩膜）

$$M^{(i)} : \Omega^{(i)} \rightarrow \{0, 1\}.$$

经过预处理（重采样、裁剪、标准化）后，假定仍记为 $V^{(i)}, M^{(i)}$ 。

1.1 左右冠状动脉的分离

对每个病例，假定我们可以将掩膜划分为左、右冠状动脉两部分：

$$M^{(i)} = M^{(i,L)} \cup M^{(i,R)}, \quad M^{(i,L)} \cap M^{(i,R)} = \emptyset,$$

其中 L 表示左冠（LCA）， R 表示右冠（RCA）。

1.2 中心线与树结构

对每个冠状动脉系统 $s \in \{L, R\}$ ，从掩膜 $M^{(i,s)}$ 中提取中心线骨架，得到一族中心线分支（branches）：

$$\mathcal{B}^{(i,s)} = \{b_1^{(i,s)}, b_2^{(i,s)}, \dots, b_{N_b^{(i,s)}}^{(i,s)}\}.$$

每一条分支 b 是一条从某个分叉点或起始点到分叉点或终末点的中心线曲线，记为

$$c_b : [0, L_b] \rightarrow \mathbb{R}^3,$$

其中 $L_b > 0$ 是该分支的物理弧长（单位 mm）， $c_b(l)$ 给出弧长参数 $l \in [0, L_b]$ 处的中心线点坐标。

整个冠状动脉系统可抽象为一个有向树图：

$$T^{(i,s)} = (\mathcal{B}^{(i,s)}, \mathcal{E}^{(i,s)}),$$

其中边集 $\mathcal{E}^{(i,s)}$ 描述分支之间的父子关系（哪条分支从哪条分支上分出）。

Centerline tree (Matplotlib 3D)

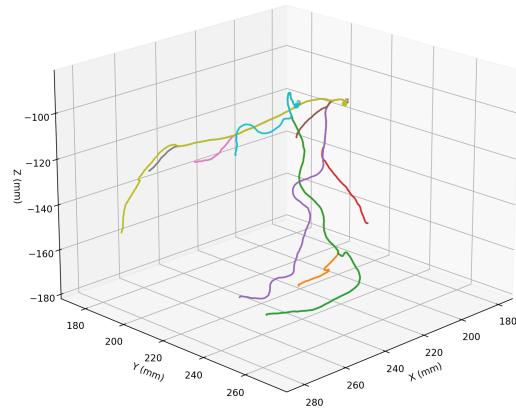


图 1: 中轴线结构

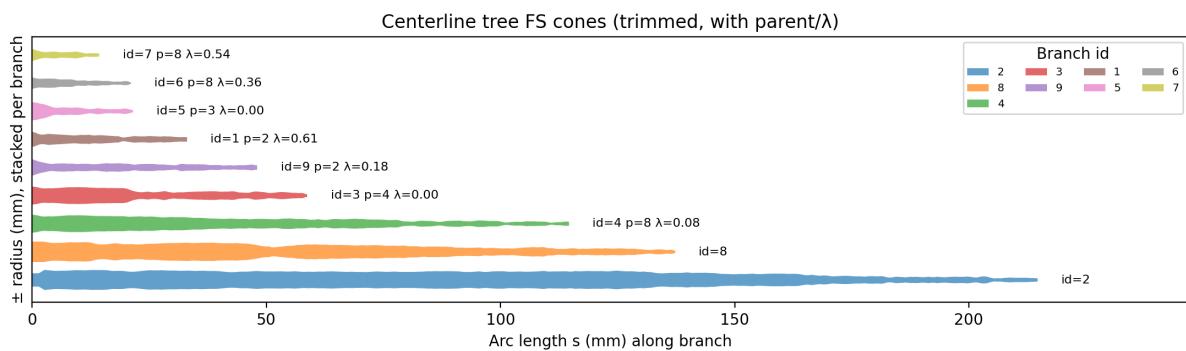


图 2: 每一分枝长度

2 单支血管 (branch) 的参数化表示

2.1 归一化中心线参数

为了在不同长度的血管之间建立统一表示，对每条分支 b 定义归一化参数

$$t = \frac{l}{L_b} \in [0, 1].$$

于是中心线可写为

$$\tilde{c}_b : [0, 1] \rightarrow \mathbb{R}^3, \quad \tilde{c}_b(t) = c_b(L_b t).$$

2.2 主干身份与分叉位置 λ

假设每一条分支 b 都有一个“主干”归属（解剖学 LAD, LCX, RCA 等）。设

$$\mathcal{T} = \{\text{LAD}, \text{LCX}, \text{RCA}, \dots\},$$

则

$$\tau_b \in \mathcal{T}$$

表示分支 b 所属的主干类别（或代号）。

若 b 是从某条主干分支 b_{trunk} 上分叉出来，记主干的中心线为

$$\tilde{c}_{\text{trunk}} : [0, 1] \rightarrow \mathbb{R}^3,$$

在主干上的分叉点对应弧长参数 l_{bifur} ，则其相对分叉位置定义为

$$\lambda_b = \frac{l_{\text{bifur}}}{L_{\text{trunk}}} \in [0, 1].$$

λ_b 表示该分支在主干上的相对位置（从近心端到远端）。

3 截面采样与截面形状编码

3.1 沿中心线等间距采样 K 个位置

在归一化参数 $t \in [0, 1]$ 上选取固定的 K 个采样点

$$t_k = \frac{k}{K-1}, \quad k = 0, 1, \dots, K-1.$$

对应的物理弧长为

$$l_k = t_k L_b,$$

对应中心线点为

$$x_{b,k} = \tilde{c}_b(t_k) \in \mathbb{R}^3.$$

3.2 局部坐标系与角度采样

在每个中心线点 $x_{b,k}$ 处构造局部直角坐标系。首先定义切向向量

$$T_{b,k} = \frac{\frac{d\tilde{c}_b}{dt}(t_k)}{\left\| \frac{d\tilde{c}_b}{dt}(t_k) \right\|},$$

再在与 $T_{b,k}$ 垂直的平面内选取两个正交单位向量 $N_{b,k}, B_{b,k}$, 从而得到局部坐标系

$$\{T_{b,k}, N_{b,k}, B_{b,k}\}.$$

在截面平面 (由 $N_{b,k}, B_{b,k}$ 张成) 上, 采用固定的 $M = 2n$ 个角度进行采样:

$$\theta_m = \frac{2\pi m}{M}, \quad m = 0, 1, \dots, M-1.$$

通过选定统一的参考方向 (例如主干方向在截面平面上的投影), 可以保持所有分支、所有截面的角度 θ_m 具有一致的物理含义。

3.3 半径函数 $r_b(t_k, \theta_m)$ 的定义

在每个采样点 $x_{b,k}$ 处, 对每个角度 θ_m , 定义截面的单位方向向量

$$u_{b,k,m} = \cos \theta_m N_{b,k} + \sin \theta_m B_{b,k}.$$

沿该方向从中心线点出发

$$x_{b,k,m}(r) = x_{b,k} + r u_{b,k,m}, \quad r \geq 0,$$

定义半径为“从中心线到血管边界的最小距离”:

$$r_b(t_k, \theta_m) = \inf \{r \geq 0 \mid x_{b,k,m}(r) \notin \text{lumen}(M^{(i,s)})\},$$

其中 $\text{lumen}(M^{(i,s)})$ 表示掩膜中对应血管腔的区域。

于是, 对每条分支 b 和每个采样位置 t_k , 得到在角度上离散化的截面半径函数

$$\{r_b(t_k, \theta_m)\}_{m=0}^{M-1}.$$

3.4 单支血管的形状张量表示

综合上述采样, 对每条分支 b 最终获得一个 $K \times M$ 的半径张量

$$R_b \in \mathbb{R}^{K \times M}, \quad R_b[k, m] = r_b(t_k, \theta_m).$$

4 每支血管的完整训练样本表示

对病例 i 、冠状系统 $s \in \{L, R\}$ 的一条分支 $b \in \mathcal{B}^{(i,s)}$, 将其表示为一个训练样本

$$\mathcal{D}_b = (\text{meta}_b, R_b),$$

其中

$$\text{meta}_b = (\tau_b, \lambda_b, L_b, \text{parent_id}_b, \text{其他可选属性}),$$

包含该分支的主干类别 τ_b 、在主干上的相对分叉位置 λ_b 、物理长度 L_b 等元信息； $R_b \in \mathbb{R}^{K \times M}$ 则为统一维度的截面半径张量，编码了这条血管沿中心线与角度方向的形状。

5 训练数据集的整体形式

在所有病例、所有冠状系统、所有分支上汇总，得到训练数据集

$$\mathcal{D} = \{ \mathcal{D}_b \mid b \in \mathcal{B}^{(i,s)}, i = 1, \dots, N_{\text{case}}, s \in \{L, R\} \}.$$

该数据集可用于后续学习：

- 基于 $\{R_b\}$ 的树枝形状先验 (Branch Shape Prior)，建模血管截面随弧长变化的统计规律；
- 基于 $\{\tau_b, \lambda_b, L_b, \text{parent_id}_b, \dots\}$ 的树结构先验 (Tree Prior)，建模分支拓扑与分叉模式。

6 一些问题和想法

在上面的建模框架下，其实我自己还有不少地方是“想明白大方向，但细节没完全想透”的，这里简单把目前的疑惑和备选思路记下来，后面可以一边做实验一边修正。

6.1 从分割到左右支：要不要一开始就区分得很漂亮？

现在假设是“从 CT + ground truth 分割开始”，理论上可以直接得到一整块冠脉掩膜。但在真正实现时，其实有几个选择：

- 一种做法是：先不纠结左右支，直接把整块掩膜当成一个整体，做中心线提取、断点修复和树结构建模，之后在树层面再去区分 LCA / RCA；
- 另一种做法是：一开始就用两个分割模型或者对现有模型做 fine-tune，把左右支分得比较干净，这样后面每棵树都更“纯净”。

直觉上我更倾向于先把整个 pipeline 跑通，再考虑左右支精细分割的问题，否则一上来就会卡在分割模型上，偏离了“先建好先验”的主线。后面如果发现某些统计量确实对左右支非常敏感，再回头做针对性的 fine-tune 可能更合理。

6.2 中心线断点修复：规则、学习还是模板化？

我现在对中心线断点的看法是：断点概率是不可避免的，尤其是远端小血管，显影一差就断了。但血管本身是很有“指向性”的结构，所以从几何上看，有不少可利用的信息：

- 空间上，两个端点的距离不能太远；

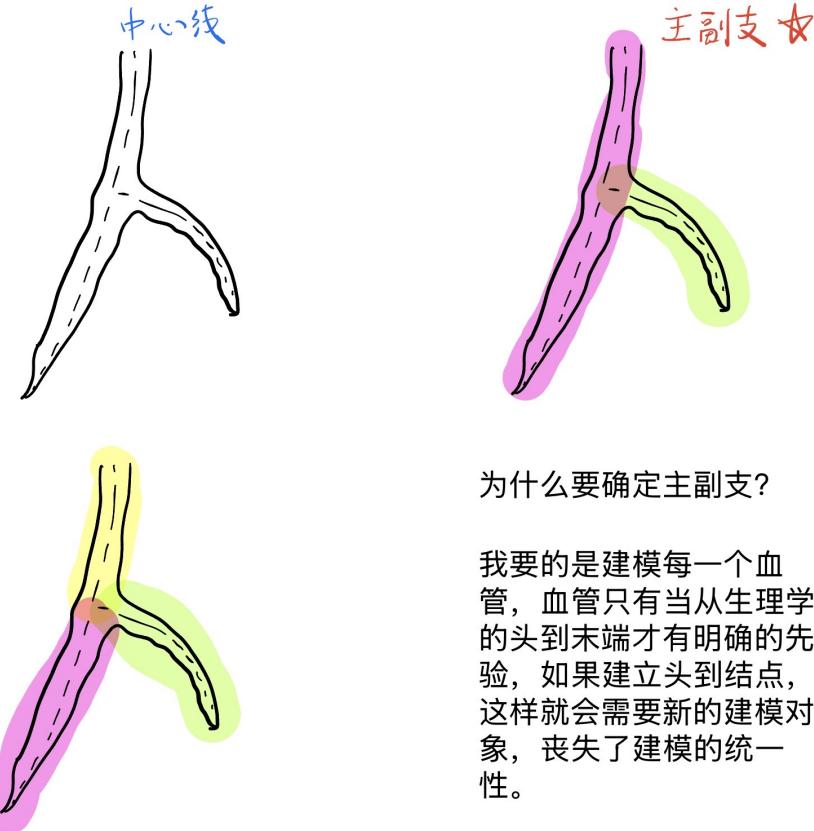


图 3: 主副支

- 方向上，延长后的切向应该大致对得上；
- 尺度上，半径变化不能太突兀。

目前的想法有三条路线：

1. 先用几何规则 + 三次样条把明显的断点修一修，作为 baseline；
2. 在积累一些“修得比较满意”的案例之后，尝试训练一个小模型，专门学“两段中心线是不是同一根血管”，把断点修复变成一个二分类问题；
3. 更激进一点的做法是做“模板化”：先把很多完整的小血管统计出来，学一个中心线的形状先验，再用这个先验去“吸附”断掉的中心线。

现在还不确定最后会走到哪一步，但可以先从最简单的规则方法做起，边可视化，边看几何规则到底有多“靠谱”。

6.3 主干 / 副支划分的原因？以及结点附近怎么处理？

为什么要确定主副支？

我要的是建模每一个血管，血管只有当从生理学的头到末端才有明确的先验，如果建立头到结点，这样就会需要新的建模对象，丧失了建模的统一性。

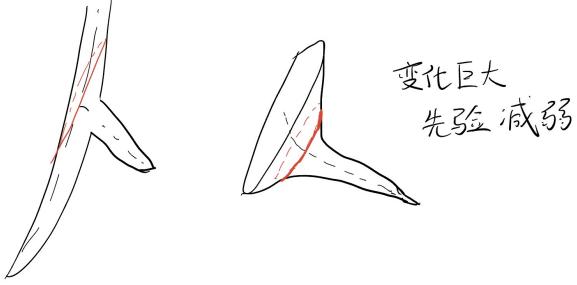


图 4: 结点处

直觉上，如果把“头到结点”也作为一个独立对象来建模，整个空间会变得非常碎，每个模型都要重新考虑分叉模式，会丧失统一性。所以我更倾向于坚持：

一支血管 = 从生理学意义的起点，一路走到末端，不在结点处硬切断。

这就引出了结点附近怎么切分的问题：如果所有 branch 都从结点开始读数据，截面会从“非常粗（带着部分主干横截面）”一下子跳到很细，容易造成建模不稳定。所以我的临时想法是：

- 先在主干上往前退一小段距离，确定一个“过渡区”；
- 在这个过渡区内，对主干和副支都做平滑拼接，而不是正好在结点处硬切；
- 具体的距离可以先用固定值（比如几毫米或者若干个体素），后续再根据实验调整。

这一块我现在是“有图像感，但没有完全形式化”，后面需要结合可视化结果做更多尝试。

6.4 中心线先验：“流苏 / 珊瑚”图景与解剖先验的差距

在脑子里，我对最终的中心线先验有一个比较形象的图景：只要给出每个树枝的相对长度、在哪个位置分叉、分叉角度，就可以生成一棵“流苏 / 珊瑚”一样的中心线树，然后再往上贴截面形状，就变成一棵完整的三维冠脉。

问题在于：这种先验未必严格满足所有解剖学约束，它更像是一个“统计意义上的冠脉树”。这带来两个问题：

- 在数据增强、仿真等任务中，这种“流苏式”的树可能已经够用；
- 但在更严格的解剖应用中，可能需要加入更多基于解剖知识的硬约束（例如某些分支一定要出现，某些角度范围不能超过多少）。

短期内，我会优先把“统计先验”搞清楚，让模型先学会生成“大致合理”的中心线树。后面如果有必要，再往里加入更强的解剖约束。

6.5 拼接和平滑：怎样对齐主副支？

最后一个让我有点纠结的问题是拼接和对齐。现在的粗略想法是：

- 在每个结点附近，先在中心线层面实现光滑拼接：弧长参数是连续的，切向方向不要出现明显的折角；
- 在截面层面，以中心线局部坐标系为基准，约定副支的 0° 方向对齐主干方向，这样主干和副支的截面在拼接时有一个统一的参考；
- 对截面半径场 $r(t, \theta)$ 做一定程度的平滑过渡，避免在结点附近出现特别尖锐的形状跳变。

这一块可能需要和可视化工具配合起来，肉眼看一看“拼起来到底是不是一根看着顺眼的血管”，在此基础上再去调节具体的数学形式。

总体来说，这些问题现在都还处在“边想边做”的阶段。好处是大方向已经比较清晰：从 CT 分割出发，先建树、再建枝、最后再拼起来；难点主要集中在细节：断点怎么连、主副支怎么切、结点附近怎么平滑。这些都可以在后续实验中逐步迭代。