

拓扑几何约束管状结构分割的动态蛇卷积

Yaolei Qi, Yuting He, Xiaoming Qi, Yuan Zhang, Guanyu Yang

摘要—精确分割拓扑管状结构例如血管和道路，对各个领域至关重要，可确保下游任务的准确性和效率。然而，许多因素使任务变得复杂，包括细小脆弱的局部结构和复杂多变的全局形态。在这项工作中，我们注意到管状结构的特殊特征，并利用这一知识来引导我们的 DSCNet 在三个阶段同时增强感知：特征提取、特征融合和损失约束。首先，我们提出了一种动态蛇形卷积，通过自适应地聚焦于细长和曲折的局部结构，来准确捕捉管状结构的特征。随后，我们提出了一种多视角特征融合策略，在特征融合期间从多个角度补充对特征的关注，确保从不同全局形态中保留重要信息。最后，提出新的基于持续同调的连续性约束损失函数，以更好地约束分割的拓扑连续性。我们的方法在 2D 和 3D 数据集上均有实验证明，与经典的几种方法相比，我们的 DSCNet 在管状结构分割任务上提供了更好的准确性和连续性。代码可在 <https://github.com/YaoleiQi/DSCNet> 中找到。

Index Terms—先验知识融合，动态蛇形卷积，多视角特征融合，持续同调，管状结构分割。

I. 引言

在各个领域中，对拓扑管状结构的准确分割对于确保下游任务的精确性和效率至关重要。在临床应用中，清晰勾画的血管是计算血液动力学仿真的关键先决条件，并协助放射科医生定位和诊断病变 [12], [15]。在遥感应用中，完整的道路分割为路径规划提供了坚实的基础。无论在哪个领域，这些结构都拥有细长和曲折的共性特征，同时由于其在图像中的占比较小，这使得关键特征难以被模型捕捉。因此，迫切需要增强对细长管状结构的感知。

然而，由于以下困难，这仍然是一个具有挑战性的问题：(1) **细小脆弱的局部结构**。如图 1 所示，细小结构仅占整体图像很小比例，像素构成有限。这些结构容易受到复杂背景的干扰，使模型难以精确辨别目标微妙的变化。因此，模型可能在区分这些结构时遇到困难，

这项研究得到了国家重点研发计划政府间合作项目 (2022YFE0116700) 和江苏省研究生科研与实践创新计划、中央高校基本科研业务费专项资金 (KYCX220239) 的支持。

YL Qi, YT He, XM Qi, Y Zhang 和 GY Yang 在江苏南京 210096 东南大学新一代人工智能技术及其交叉应用教育部重点实验室。GY Yang 在江苏南京 210096 东南大学医学信息处理国际联合实验室。GY Yang (yang.list@seu.edu.cn) 是通讯作者。这项工作是我们会议论文 [21] 的延伸。

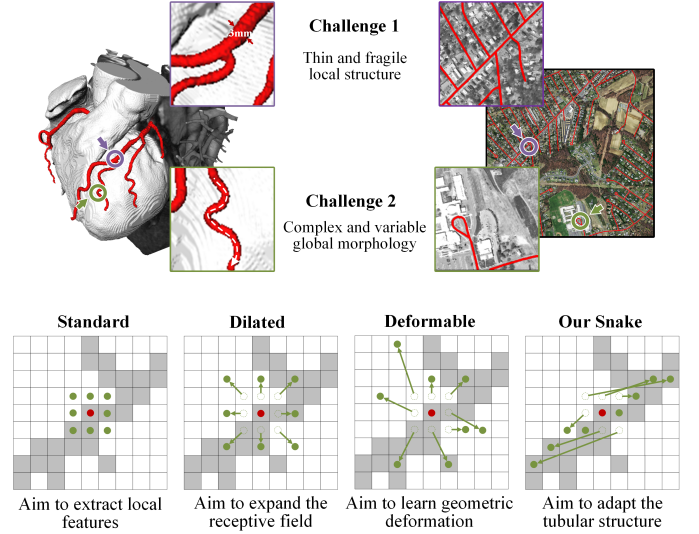


图 1: **挑战**。上图展示了一个 3D 心脏血管数据集和一个 2D 遥感道路数据集。这两个数据集的目标都是提取管状结构，但由于细长脆弱的局部结构和复杂多变的全局形态，这项任务面临着挑战。**动机**。标准卷积核旨在提取局部特征。在此基础上，可变形卷积核丰富其应用并适应不同目标的几何变形。然而，由于前述的挑战，很难有效地将注意力集中在细长的管状结构上。

导致分割出现断裂。(2) **复杂多变的全局形态**。图 1 展示了细长管状结构的复杂和多变的形态，即使在同一图像中也是如此。不同区域的目标呈现不同的形态上的变化，变化取决于分支数量、分叉点位置和管状结构的路径长度。当数据展现出前所未见的形态结构时，模型可能过度拟合于已见过的特征，导致弱的泛化能力。

最近，许多研究提出了将领域知识（例如几何拓扑和树形结构等）纳入模型，以更好地引导模型感知管状结构的独特特征，从而专注于提高局部分割的准确性并保持全局形态的连续性。现有方法可以粗略分为三类：(1) 基于网络的方法 [6], [7], [11], [13], [27], [30] 根据管状结构的特点设计特定的网络架构，引导模型关注关键特征。然而，由于管状结构所占比例较小，网络可能不可避免地失去对相应结构的感知。(2) 基于特征

的方法 [14], [19], [20], [32], [34] 通过为模型提供额外的特征表示, 增强对管状结构的特定几何和拓扑特征的理解。然而, 一些冗余的特征表示加重了计算负担, 同时对模型没有积极影响。(3) 基于损失的方法 [1], [24], [28], [29] 在训练过程中通过测量方法补充约束, 通常通过损失函数。这些方法加强了对分割的严格约束。在此基础上, 从拓扑角度结合结构化损失和连续性约束将有可能进一步提高管状结构分割的准确性。

为了应对上述挑战, 我们提出了一个全新的框架, DSCNet, 包括一个管状感知的动态蛇形卷积核、多视角特征融合策略和拓扑连续性约束损失函数。(1) 为了应对模型难以集中注意力处理细小脆弱的局部结构这一问题, 我们提出了动态蛇形卷积 (DSConv) 来通过自适应地聚焦于管状结构的细小且弯曲的局部特征, 增强对几何结构的感知。与可变形卷积 [5] 不同, 可变形卷积让网络完全自由学习几何变化, 因此导致感知区域漫游, 特别是在细小管状结构上。我们的 DSConv 考虑到管状结构的蛇形形态, 并通过约束补充自由学习过程, 从而有针对性地增强对管状结构的感知。(2) 为了应对复杂多变的全局形态的挑战, 我们提出了一种多视角特征融合策略。在这种方法中, 我们基于 DSConv 生成多个形态学核模板, 从不同角度观察目标的结构特征, 并通过总结典型的关键特征实现高效的特征融合。(3) 为了应对管状结构分割容易出现断裂的问题, 我们提出了一种基于持久同调 (PH) 的拓扑连续性约束损失函数 (TCLoss)。PH [4], [8], [18] 响应了拓扑特征从出现到消失的过程, 从嘈杂的高维数据中获取充分的拓扑信息。相关的贝蒂数是描述拓扑空间中连接性的一种方式。与 [9], [29] 不同, 我们的 TCLoss 将 PH 与点集相似性结合, 引导网络关注具有异常像素/体素分布的断裂区域, 从拓扑角度实现连续性约束。

总结一下, 我们的工作提出了一个解决细长管状结构困难的知识融合的新框架, 具体贡献有三个方面: (1) 我们提出了一种动态蛇形卷积, 自适应地关注细长曲折的局部特征, 实现了在 2D 和 3D 数据集上对准确管状结构的分割。我们的模型经过了充分的内部和外部测试数据验证。(2) 我们提出了一种多视角特征融合策略, 以补充对关键特征的多方面关注。(3) 我们提出了基于点集相似性的持久同调的拓扑连续性约束损失函数, 更好地约束了分割的连续性。

II. 相关工作

基于网络设计的方法。为了通过设计特定的网络架构和模块来更好地适应管状结构的形态, 人们提出了各种方法。(1) 基于卷积核设计的方法, 以著名的扩张卷积 [31] 和可变形卷积 [5] 为代表, 旨在处理卷积神经网络中固有的有限几何变换, 对于复杂的检测和分割任务表现出色。这些方法 [6], [11], [30], [33] 也被设计成动态感知对象的几何特征, 以适应具有可变形态的结构。例如, 在 [11] 中提出的 DUNet 将可变形卷积整合到 U 形架构中, 并根据血管的尺度和形状自适应调整感受野。(2) 基于网络架构设计的方法旨在学习管状结构的特殊几何拓扑特征。PointScatter [27] 提出使用点集来表示管状结构, 这是用于提取管状结构的分割模型的一种替代方法。在 [13] 中, 提出了一种树形结构的卷积门控循环单元, 以显式地建模冠状动脉的拓扑结构。与上述允许模型完全自由学习几何变化的思想不同, 考虑到过多的随机性可能导致难以收敛的困难, 以及模型可能会集中精力于目标以外的区域, 我们的工作特征提取过程中整合了关于管状结构形态的领域知识, 以稳定地增强对管状结构的感知。

基于特征融合的方法。基于特征融合的方法 [14], [19], [20], [32], [34] 通过为模型补充额外的特征信息来增强对管状结构的表示。考虑到管状结构的拓扑结构和稀疏性, [34] 提出了一种在两个网络之间执行的跨网络多尺度特征融合方法, 以有效支持高质量的血管分割。在 [14] 中, 通过深浅分层特征融合研究了全局变换器和双重局部注意网络, 以同时捕捉全局和局部特征。在 [32] 中, 提出了融合解剖学上下文信息和血管拓扑结构的方法, 以实现准确的管状结构分割。在我们的工作中, 我们提出了一种多视角特征融合策略, 以补充对关键特征的多方面关注。在这个策略中, 我们基于我们的 DSConv 生成了许多形态学核模板, 从多个角度观察目标的结构特征, 并通过总结基本的标准特征来实现特征融合, 从而提高我们模型的性能。

基于损失函数的方法。基于损失函数的方法 [1], [24], [28] 引入测量方法以在训练过程中补充约束。这些方法加强了对管状结构分割的强约束。[24] 引入了一种称为中心线 Dice 的相似度测量, 该测量是在分割掩模和骨架的交叉点上计算的。[28] 提出了一种几何感知的管状结构分割方法, Deep Distance Transform (DDT), 它结合了经典距离变换用于骨架化和管状结构分割的直觉。这些方法关注管状结构分割的连续性, 但骨架的不

准确性和偏移会影响约束的精度。在 [1] 中, 提出了一种相似性指数, 捕捉了预测分割的拓扑一致性, 并设计了一个基于形态闭合运算符的损失函数, 用于管状结构分割。在 [29] 中, 拓扑数据分析方法与几何深度学习模型结合, 用于三维对象的精细分割。这些方法将捕捉拓扑对象的特征。在这方面得到启发, 我们的工作提出了一种基于拓扑连续性与点集相似度约束损失函数 (TCLoss), 它从拓扑的角度更好地约束分割的连续性。我们的 TCLoss 在训练过程中逐渐引入基于持久同调的约束并用点击相似度评估, 引导网络关注断裂区域并实现连续性。

III. 方法

我们的方法旨在同时处理细长管状结构的 2D 和 3D 特征图。为简化起见, 我们的模块以 2D 描述, 并在我们的开源项目中提供了详细的 3D 扩展。

A. 动态蛇形卷积

在本节中我们将讨论如何执行动态蛇形卷积 (Dynamic Snake Convolution, DSConv) 以提取管状结构的局部特征。给定标准的 2D 卷积坐标 K , 中心坐标为 $K_i = (x_i, y_i)$ 。一个 3×3 的卷积核 K 表示为:

$$K = \{(x-1, y-1), (x-1, y), \dots, (x+1, y+1)\} \quad (1)$$

为了赋予卷积核更多灵活性, 使其能够聚焦于目标的复杂几何特征, 受到 [5] 的启发, 我们引入了变形偏移 Δ 。然而, 如果模型被完全自由地学习变形偏移, 感知场往往会偏离目标, 特别是在处理细长管状结构的情况下。因此, 我们采用了一个迭代策略 (图 3), 依次选择每个要处理的目标的下一个位置进行观察, 从而确保关注的连续性, 不会由于大的变形偏移而将感知范围扩散得太远。

在动态蛇形卷积中, 我们将标准卷积核在 x 轴和 y 轴方向都进行了直线化。我们考虑一个大小为 9 的卷积核, 以 x 轴方向为例, K 中每个网格的具体位置表示为: $K_{i\pm c} = (x_{i\pm c}, y_{i\pm c})$, 其中 $c = 0, 1, 2, 3, 4$ 表示距离中心网格的水平距离。卷积核 K 中每个网格位置 $K_{i\pm c}$ 的选择是一个累积过程。从中心位置 K_i 开始, 远离中心网格的位置取决于前一个网格的位置: K_{i+1} 相对于 K_i 增加了偏移量 $\Delta = \{\delta | \delta \in [-1, 1]\}$ 。因此, 偏移量需要进行累加 Σ , 从而确保卷积核符合线性形态结构。图 3 中 x 轴方向的变化为:

$$K_{i\pm c} = \begin{cases} (x_{i+c}, y_{i+c}) = (x_i + c, y_i + \Sigma_{i+c}^{\Delta} \Delta y), \\ (x_{i-c}, y_{i-c}) = (x_i - c, y_i + \Sigma_{i-c}^{\Delta} \Delta y), \end{cases} \quad (2)$$

在 y 轴方向的变化为:

$$K_{j\pm c} = \begin{cases} (x_{j+c}, y_{j+c}) = (x_j + \Sigma_j^{j+c} \Delta x, y_j + c), \\ (x_{j-c}, y_{j-c}) = (x_j + \Sigma_j^{j-c} \Delta x, y_j - c), \end{cases} \quad (3)$$

由于偏移量 Δ 通常是小数, 然而坐标通常是整数形式, 因此采用双线性插值, 表示为:

$$K = \Sigma_{K'} B(K', K) \cdot K' \quad (4)$$

其中, K 表示方程 2 和方程 3 的小数位置, K' 列举所有整数空间位置, B 是双线性插值核, 可以分解为两个一维核, 即:

$$B(K, K') = b(K_x, K'_x) \cdot b(K_y, K'_y) \quad (5)$$

如图 3 所示, 由于二维 (x 轴, y 轴) 的变化, 我们的动态蛇形卷积核在变形过程中覆盖了 9×9 的感受野可选择范围。动态蛇形卷积核旨在基于动态结构更好地适应细长的管状结构, 以更好地感知关键特征。

B. 多视角融合策略

本节讨论实施多视角特征融合策略, 以引导模型从多个角度补充对关键特征的关注。对于每个 K , 从 x 轴和 y 轴提取来自第 l 层的两个特征图 $f^l(K_x)$ 和 $f^l(K_y)$, 表示为:

$$f^l(K) = \underbrace{\{\Sigma_i w(K_i) \cdot f^l(K_i)\}}_{f^l(K_x)}, \underbrace{\{\Sigma_j w(K_j) \cdot f^l(K_j)\}}_{f^l(K_y)} \quad (6)$$

其中, $w(K_i)$ 表示位置 K_i 处的权重, 第 l 层卷积核 K 提取的特征采用累积方法计算。

基于方程 6, 我们提取 m 组特征作为 T^l , 其中包含 DSConv 的不同形态:

$$T^l = \underbrace{(f^l(K_x), f^l(K_y))}_{T_1^l}, \underbrace{(f^l(K_x), f^l(K_y))}_{T_2^l}, \dots, \underbrace{(f^l(K_x), f^l(K_y))}_{T_m^l} \quad (7)$$

多个模板的特征融合将不可避免地引入冗余噪声。因此, 在训练阶段引入了一种随机丢弃策略 r^l (图 4), 以提高我们模型的性能并防止过拟合, 而不增加额外的计算负担, 此时方程 7 变为:

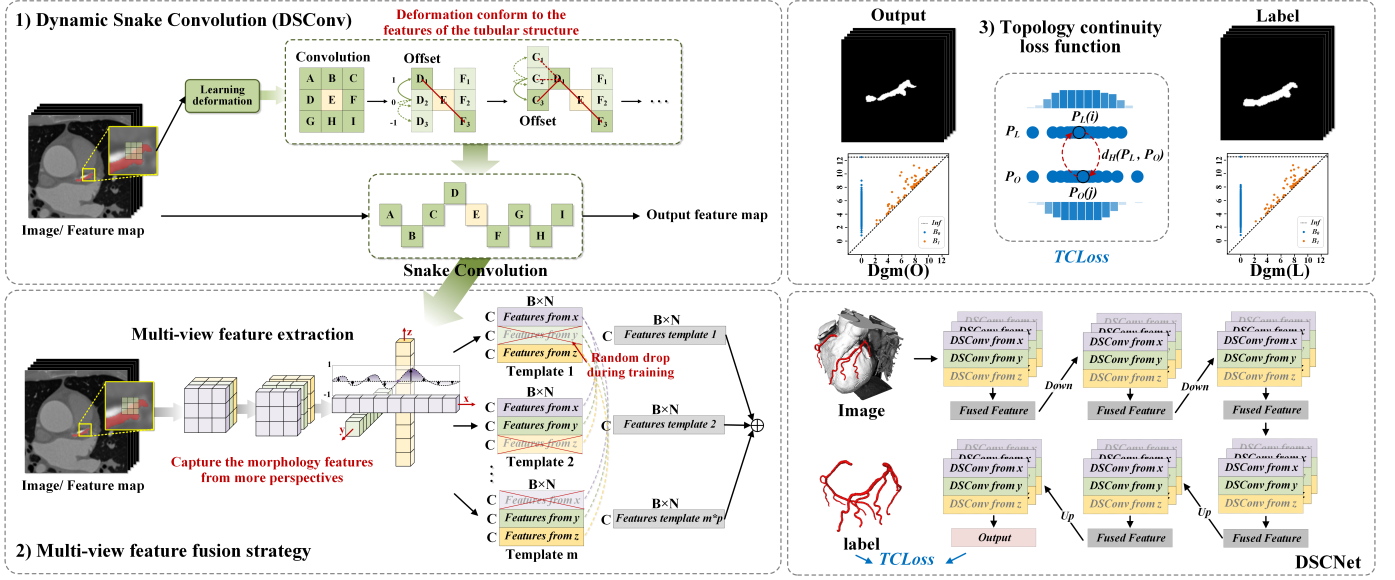


图 2: 方法。我们提出的方法在 3D 冠状动脉分割的示例上进行了示意性概述。我们的方法包括三个部分: (1) 动态蛇形卷积 (DSCConv), 根据输入特征图学习变形, 以适应管状结构形态的知识, 自适应地聚焦于细长且弯曲的局部特征。(2) 多视角特征融合策略, 基于我们的 DSCConv 生成多个形态学核模板, 用于从多个角度观察目标的结构特征。(3) 损失函数 (TCLoss), 基于持续同调与点集相似性, 引导网络关注具有异常低像素/体素分布的断裂区域, 并实现连续性约束。

$$\begin{cases} r^l \sim \text{Bernoulli}(p) \\ \hat{T}^l = r^l \cdot T^l \\ f^{l+1}(K) = \Sigma^{[m \times p]} \hat{T}^l \end{cases} \quad (8)$$

其中, p 是随机丢弃的概率, r^l 符合伯努利分布。在训练阶段保存最佳的丢弃策略, 并在测试阶段引导模型融合关键特征。

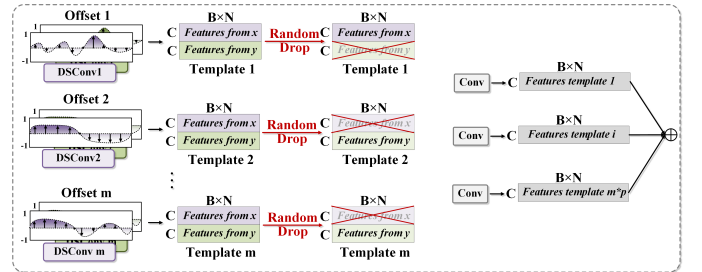


图 4: 多视角融合策略

C. 拓扑连续性约束损失

在这一部分, 我们讨论如何基于持久同调实现拓扑连续性约束损失 (TCLoss) 以约束分割的连续性。在复杂结构中, 几何和拓扑信息是帮助模型理解连续结构的

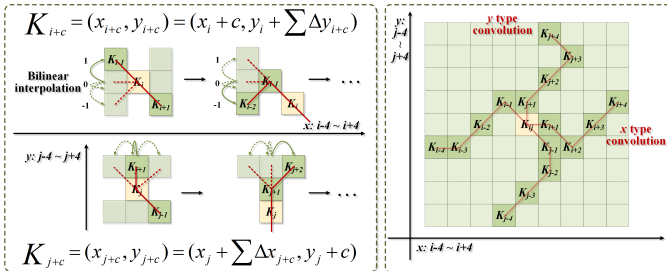


图 3: 左图: 动态蛇形卷积核坐标计算的示意图。右图: 动态蛇形卷积核可选的感受野。

关键线索。采用了拓扑数据分析工具, 以提取复杂管状结构中隐藏的关键特征。

我们的目标是构建数据的拓扑结构并提取复杂管状结构中的高维关系, 表示为持续条形图和持续同调 (PH), 如图 5 所示。

给定 G , 它的 N 维拓扑结构, 同调类 [8], [18] 是 N 维流形的等价类, 这些流形可以在 G 内彼此变形, 其中 0 维和 1 维是连接成分和手柄。PH 用于计算拓扑特征的演变, 保持拓扑特征出现时间 b 和消失时间 d 之间的周期 [29]。这些周期以简洁的格式总结为持续图 (PD), 其中包含一组点 (b, d) 。每个点 (b, d) 表示在 b 时出现并在 d 时消失的第 d 同调类。让 $PD = \text{dgm}(\cdot)$ 表示从 groundtruth L 和输出 O 中获得的持久同调。我们考

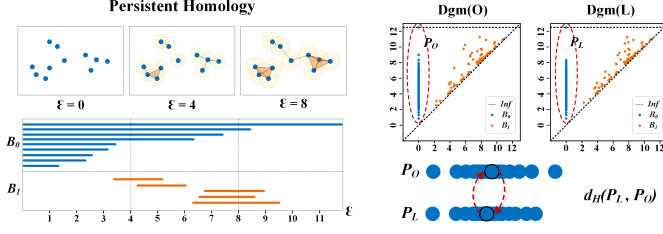


图 5: Illustration of the Persistent Homology and our TCLoss.

虑在复杂管状结构中的拓扑信息，其中包含确定断裂存在的关键线索，这些线索在 0 维和 1 维同调特征的同伦特性中表现出来。现有的方法 [4], [9], [29] 使用优化的 Wasserstein 距离计算输出生成的点与 groundtruth 生成的点之间的最佳匹配，而没有最佳配对的离群点与对角线匹配，并且不参与损失计算。然而，在我们的任务中，离群点代表异常的出现或消失时间，并暗示了错误的拓扑关系，起到了重要作用。因此，我们考虑使用 Hausdorff 距离 [26] 来衡量两组点之间的相似性：

$$\begin{cases} d_H(P_O, P_L) = \max_{u \in P_O} \min_{v \in P_L} \|u - v\| \\ d_H(P_L, P_O) = \max_{v \in P_L} \min_{u \in P_O} \|v - u\| \\ d_H^* = \max\{d_H(P_O, P_L), d_H(P_L, P_O)\} \end{cases} \quad (9)$$

其中， $P_O \in \text{Dgm}(O)$ ， $P_L \in \text{Dgm}(L)$ ， d_H^* 表示双向 Hausdorff 距离，它是用 n 维点来计算的。我们使用的 Hausdorff 距离对离群值敏感。如方程 9 所示，如果两组点相似，所有的点都完全重叠，除了 P_O 中的一个点，它远离 P_L 中的任何点，那么 Hausdorff 距离由该点决定，且较大 [10]。

然后对所有维度 ($n = 0, 1, 2, \dots, N$) 进行求和，得到 $\mathcal{L}_{PH}$ ，整个 TCLoss 与交叉熵损失 $\mathcal{L}_{CE}$ 集成为最终的损失函数 $\mathcal{L}_{TC} = \mathcal{L}_{CE} + \sum_{n=0}^N d_H^*$ 。

最终，通过两个损失函数的联合作用，拓扑和准确性得到了约束，有助于实现连续的管状结构分割。

IV. 实验设置

A. 数据集

我们使用了包含两个公共数据集和一个内部数据集的三个数据集来验证我们的框架。在 2D 方面，我们评估了 DRIVE 视网膜数据集 [25] 和马萨诸塞州道路数据集 [16]。在 3D 方面，我们使用了名为 Cardiac

CCTA 的数据集。有关实验设置的详细信息可以在补充材料中找到。

B. 评估指标

我们进行了比较实验和消融研究，以展示我们提出的框架的优势。我们将经典的分割网络 U-Net [3] 和 2021 年提出的用于血管分割的 CS²-Net [17] 与之比较，以验证准确性。为了验证网络设计性能，我们比较了 2022 年提出的用于视网膜血管分割的 DCU-net [30]。为了验证特征融合的优势，我们比较了 2021 年提出的用于医学图像分割的 Transunet [2]。为了验证损失函数的约束，我们比较了 2021 年提出的 cIDice [24] 和基于 Wasserstein 距离的 TCLoss $\mathcal{L}_{WTC}$ [29]。这些模型在相同的数据集上进行了训练，使用相同的实现，并通过以下指标进行了评估。所有指标都是针对每个图像计算的并进行了平均处理。

- 1) 体积分数: 平均 Dice 系数 (Dice), 相对 Dice 系数 (RDice) [22], cIDice 系数 (cIDice) [24], 准确度 (ACC) and AUC 是被用于评估结果表现的指标。
- 2) 拓扑误差: 我们遵循 [24], [27] 并且计算了基于拓扑连续性的分数包括 Betti 误差用于贝蒂数 β_0 和 β_1 。同时，为了客观验证冠状动脉分割的连续性，使用了直到第一个错误的重叠 (OF) [23] 来评估提取的中心线的完整性。
- 3) 距离误差: Hausdorff 距离 (HD) [26] 广泛用于描述两组点集的相似性，常用于评估细小管状结构。

V. 结果与讨论

在这一段中，我们将通过以下三种方式评估和分析我们提出的框架的有效性：(1) 我们将使用以下指标比较和验证我们提出的方法在薄管状结构分割任务中的性能。同时展示不同方法的视觉效果。(2) 我们分析了我们提出的 DSConv 引导模型专注于管状结构的有效性，以及 TCLoss 对分割拓扑的约束帮助。(3) 我们提供了以 DRIVE 数据集为例的综合性实验，包括消融研究。此外，由于篇幅限制，我们突出了其他数据集上一些最重要的比较实验的结果。结果显示我们的方法在 2D 和 3D 领域都表现出色。

A. 定量评估

我们的方法在每个指标上的优势在表 I 中得到了展示，结果显示我们提出的 DSCNet 在 2D 和 3D 数据集上都取得了更好的结果。

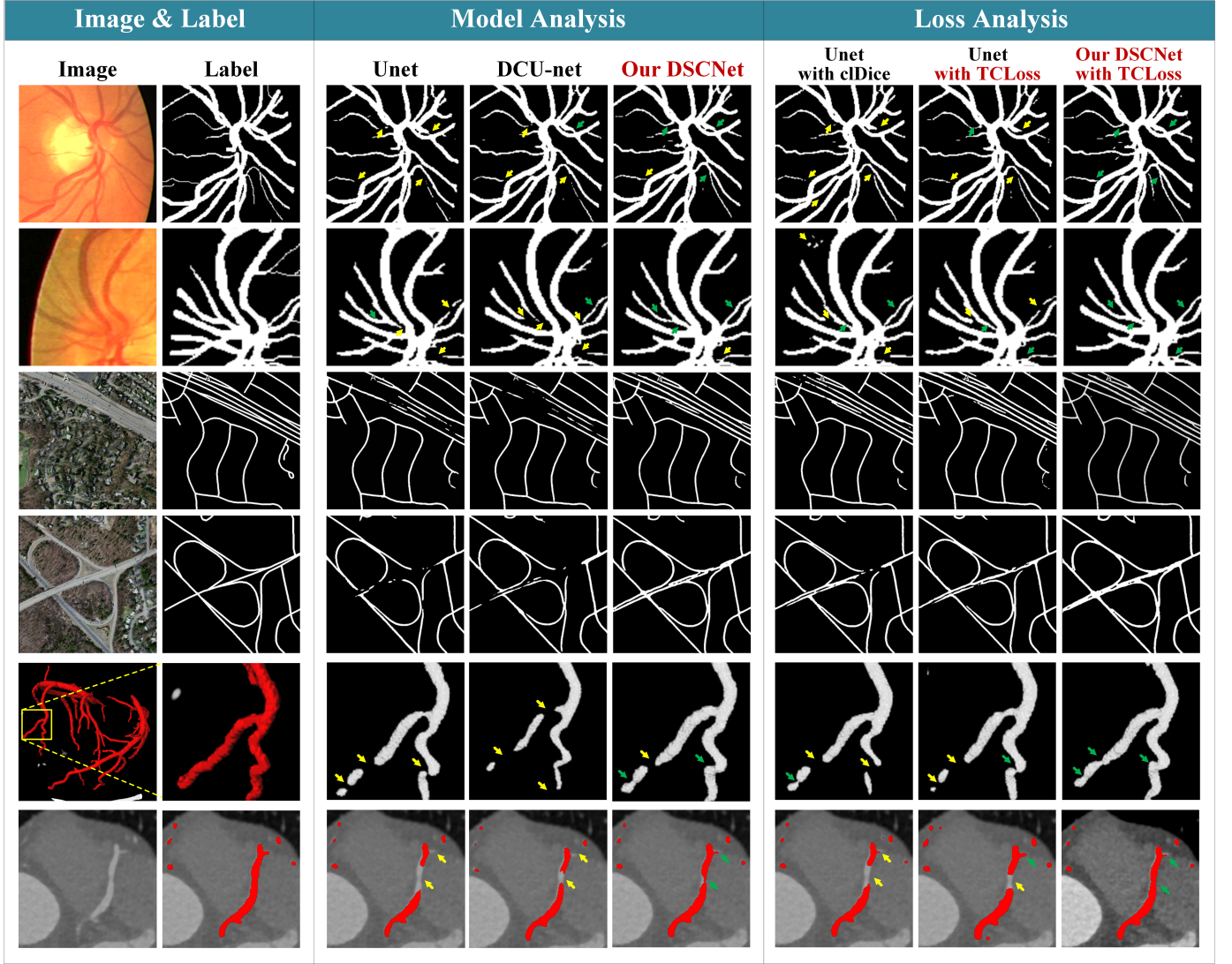


图 6: 视觉效果。为了更客观、高效地验证我们方法的性能，我们从每个数据集中选择了代表性的难以分割的区域。从上到下，我们展示了 DRIVE 数据集、Massachusetts 道路数据集和我们内部的冠状动脉数据集的两行结果。从左到右，显示了原始图像、地面真相以及来自经典 UNet、DCU-net、我们的 DSCNet、带有 clDice 的 UNet、带有我们提出的 TCLoss 的 UNet 以及带有 TCLoss 的我们的 DSCNet 的结果。结果表明，我们的 DSCNet 和 TCLoss 在分割准确性和拓扑连续性方面优于其他模型。黄色箭头指示分割中断的区域，而绿色箭头指示分割效果良好的区域。

DRIVE 数据集上的评估。在 DRIVE 数据集上，我们的 DSCNet 在分割准确性和拓扑连续性方面优于其他模型。在表 I 中，我们提出的 DSCNet 在体积准确性的角度上与其他方法相比取得了最佳的分割结果，Dice 达到 82.06%，RDice 达到 90.17%，clDice 达到 82.07%，ACC 达到 96.87%，AUC 达到 90.27%。与此同时，从拓扑的角度来看，我们的 DSCNet 在拓扑连续性方面也优于其他方法， β_0 误差为 0.998， β_1 误差为 0.803。结果表明，我们的方法更好地捕捉了细小管

状结构的特定特征，并呈现出更准确的分割性能和更连续的拓扑。正如表 I 的第六行到第十二行所示，通过添加我们的 TCLoss，不同的模型在分割的拓扑连续性上都有所改善。结果表明，我们的 TCLoss 能够准确约束模型专注于细小管状结构，以保持拓扑连续性。

ROADS 数据集上的评估。在 Massachusetts Roads 数据集上，我们的 DSCNet 同样取得了最佳结果。如表 I 所示，我们提出的带有 TCLoss 的 DSCNet 在 Dice (78.21%)、RDice (85.85%) 和 clDice (87.64%) 等指

Dataset	Network	Loss	Volumetric (%) $\uparrow$					Topology $\downarrow$		Distance $\downarrow$
			Dice	RDice	clDice	ACC	AUC	β_0	β_1	HD
DRIVE	UNet	$\mathcal{L}_{CE}$	80.73 $\pm$ 1.77	87.94 $\pm$ 3.32	79.66 $\pm$ 4.00	96.74 $\pm$ 0.28	88.57 $\pm$ 2.44	1.209 $\pm$ 0.342	0.883 $\pm$ 0.135	6.86 $\pm$ 0.56
	Transunet	$\mathcal{L}_{CE}$	80.56 $\pm$ 2.14	87.14 $\pm$ 3.82	79.02 $\pm$ 5.05	96.75 $\pm$ 0.32	88.02 $\pm$ 2.79	1.210 $\pm$ 0.309	0.844 $\pm$ 0.157	6.83 $\pm$ 0.52
	CS ² -Net	$\mathcal{L}_{CE}$	77.53 $\pm$ 2.94	82.55 $\pm$ 4.10	74.88 $\pm$ 5.27	96.46 $\pm$ 0.36	84.73 $\pm$ 2.82	1.391 $\pm$ 0.331	0.906 $\pm$ 0.177	6.90 $\pm$ 0.48
	DCU-net	$\mathcal{L}_{CE}$	80.83 $\pm$ 1.99	87.73 $\pm$ 3.60	80.19 $\pm$ 4.80	96.77 $\pm$ 0.31	88.45 $\pm$ 2.67	1.104 $\pm$ 0.327	0.817 $\pm$ 0.166	6.84 $\pm$ 0.58
	DSCNet(ours)	$\mathcal{L}_{CE}$	81.85 $\pm$ 1.74	88.93 $\pm$ 3.36	81.16 $\pm$ 4.54	96.91 $\pm$ 0.28	89.38 $\pm$ 2.54	1.094 $\pm$ 0.301	0.780 $\pm$ 0.162	6.68 $\pm$ 0.49
	UNet	$\mathcal{L}_{TC}$ (ours)	80.93 $\pm$ 1.97	88.00 $\pm$ 3.41	80.28 $\pm$ 4.41	96.78 $\pm$ 0.30	88.63 $\pm$ 2.56	1.117 $\pm$ 0.286	0.797 $\pm$ 0.151	6.88 $\pm$ 0.53
	Transunet	$\mathcal{L}_{TC}$ (ours)	80.79 $\pm$ 2.11	87.78 $\pm$ 3.80	79.86 $\pm$ 4.90	96.76 $\pm$ 0.32	88.48 $\pm$ 2.82	1.176 $\pm$ 0.295	0.818 $\pm$ 0.176	6.83 $\pm$ 0.51
	CS ² -Net	$\mathcal{L}_{TC}$ (ours)	79.69 $\pm$ 2.31	86.14 $\pm$ 3.82	77.72 $\pm$ 5.09	96.64 $\pm$ 0.32	87.25 $\pm$ 2.76	1.308 $\pm$ 0.334	0.848 $\pm$ 0.160	6.93 $\pm$ 0.45
	DCU-net	$\mathcal{L}_{TC}$ (ours)	81.18 $\pm$ 1.90	87.89 $\pm$ 3.43	80.60 $\pm$ 4.54	96.83 $\pm$ 0.31	88.59 $\pm$ 2.57	1.076 $\pm$ 0.313	0.817 $\pm$ 0.167	6.80 $\pm$ 0.56
	UNet	clDice	80.77 $\pm$ 1.92	87.53 $\pm$ 3.42	79.93 $\pm$ 4.48	96.77 $\pm$ 0.31	88.29 $\pm$ 2.52	1.199 $\pm$ 0.303	0.833 $\pm$ 0.157	6.93 $\pm$ 0.54
	UNet	$\mathcal{L}_{WTC}$	80.89 $\pm$ 1.95	87.85 $\pm$ 3.55	80.03 $\pm$ 4.75	96.78 $\pm$ 0.29	88.53 $\pm$ 2.64	1.144 $\pm$ 0.339	0.814 $\pm$ 0.176	6.79 $\pm$ 0.47
	DSCNet(ours)	$\mathcal{L}_{TC}$ (ours)	82.06 $\pm$ 1.44	90.17 $\pm$ 3.04	82.07 $\pm$ 4.35	96.87 $\pm$ 0.24	90.27 $\pm$ 2.32	0.998 $\pm$ 0.312	0.803 $\pm$ 0.179	6.78 $\pm$ 0.51
ROADS	UNet	$\mathcal{L}_{CE}$	76.90 $\pm$ 6.30	84.07 $\pm$ 6.46	86.87 $\pm$ 6.59	97.97 $\pm$ 1.27	98.29 $\pm$ 1.24	1.107 $\pm$ 0.551	1.505 $\pm$ 0.467	8.11 $\pm$ 2.42
	Transunet	$\mathcal{L}_{CE}$	75.82 $\pm$ 6.83	81.50 $\pm$ 6.65	86.04 $\pm$ 7.40	97.97 $\pm$ 1.28	98.23 $\pm$ 1.15	1.105 $\pm$ 0.615	1.570 $\pm$ 0.663	8.11 $\pm$ 2.53
	DCU-net	$\mathcal{L}_{CE}$	77.24 $\pm$ 6.30	84.26 $\pm$ 6.37	86.98 $\pm$ 6.53	98.03 $\pm$ 1.14	98.34 $\pm$ 1.19	1.085 $\pm$ 0.653	1.474 $\pm$ 0.497	8.04 $\pm$ 2.53
	UNet	$\mathcal{L}_{TC}$ (ours)	77.70 $\pm$ 6.07	84.80 $\pm$ 5.96	87.47 $\pm$ 6.31	98.03 $\pm$ 1.23	98.41 $\pm$ 1.13	1.072 $\pm$ 0.631	1.401 $\pm$ 0.496	8.04 $\pm$ 2.72
	UNet	clDice	77.37 $\pm$ 5.57	84.18 $\pm$ 5.99	87.05 $\pm$ 6.34	98.03 $\pm$ 1.22	98.40 $\pm$ 1.12	1.079 $\pm$ 0.613	1.407 $\pm$ 0.603	8.08 $\pm$ 2.46
	DSCNet(ours)	$\mathcal{L}_{CE}$	78.04 $\pm$ 5.77	85.35 $\pm$ 5.42	87.74 $\pm$ 6.02	98.05 $\pm$ 1.21	98.39 $\pm$ 1.19	1.118 $\pm$ 0.641	1.441 $\pm$ 0.523	7.96 $\pm$ 2.43
	DSCNet(ours)	$\mathcal{L}_{TC}$ (ours)	78.21 $\pm$ 5.77	85.85 $\pm$ 5.56	87.64 $\pm$ 5.99	98.05 $\pm$ 1.21	98.46 $\pm$ 1.08	1.053 $\pm$ 0.523	1.396 $\pm$ 0.456	7.34 $\pm$ 2.48

表 I: 本表显示了 DRIVE 视网膜数据集和马萨诸塞州道路数据集 (ROADS) 的定量结果。所有实验都在三个层面验证了我们方法的性能: 体积准确性、拓扑连接性和距离误差。在所有指标中, 我们的 DSCNet 和 TCLoss 取得了最具竞争力的结果。

Dataset	Network	Loss	Volumetric (%) $\uparrow$			Topology OF $\uparrow$			Distance $\downarrow$
			Dice	RDice	clDice	LAD	LCX	RCA	HD
CORONARY	UNet	$\mathcal{L}_{CE}$	76.87 $\pm$ 5.38	84.48 $\pm$ 4.55	81.43 $\pm$ 6.02	0.806 $\pm$ 0.252	0.847 $\pm$ 0.239	0.849 $\pm$ 0.267	7.727 $\pm$ 3.30
	Transunet	$\mathcal{L}_{CE}$	76.70 $\pm$ 6.65	83.23 $\pm$ 6.72	78.71 $\pm$ 6.93	0.810 $\pm$ 0.274	0.694 $\pm$ 0.307	0.816 $\pm$ 0.303	8.580 $\pm$ 4.11
	DCU-net	$\mathcal{L}_{CE}$	78.33 $\pm$ 5.00	85.67 $\pm$ 4.29	82.29 $\pm$ 5.31	0.833 $\pm$ 0.219	0.746 $\pm$ 0.296	0.835 $\pm$ 0.300	7.331 $\pm$ 3.06
	UNet	clDice	77.86 $\pm$ 5.25	84.42 $\pm$ 4.65	82.37 $\pm$ 5.54	0.817 $\pm$ 0.256	0.845 $\pm$ 0.234	0.859 $\pm$ 0.265	7.412 $\pm$ 3.68
	DSCNet(ours)	$\mathcal{L}_{CE}$	79.92 $\pm$ 5.26	85.98 $\pm$ 4.60	84.95 $\pm$ 5.76	0.858 $\pm$ 0.198	0.853 $\pm$ 0.241	0.862 $\pm$ 0.267	6.326 $\pm$ 2.85
	DSCNet(ours)	$\mathcal{L}_{TC}$ (ours)	80.27 $\pm$ 4.67	86.37 $\pm$ 4.16	85.26 $\pm$ 4.98	0.866 $\pm$ 0.195	0.885 $\pm$ 0.210	0.882 $\pm$ 0.250	5.787 $\pm$ 2.99

表 II: 3D 心脏 CCTA 数据集的定量结果。实验指标辅以 OF 指标以验证拓扑连接性。LAD、LCX 和 RCA 是冠状动脉的三个主要分支, 具有重要的临床价值。

标上取得了最佳的分割结果。与经典分割网络 UNet 的结果相比, 我们的方法在 Dice、RDice 和 clDice 上的改进分别为 1.31%、1.78% 和 0.77%。结果表明, 相比其他模型, 我们的模型在结构复杂、形态多变的道路数据集上表现良好。

冠状动脉 CCTA 数据集上的评估。在冠状动脉 CCTA 数据集上, 我们验证了我们的 DSCNet 在 3D 中仍然取得了最佳的细小管状结构分割结果。如表 II 所示, 我们提出的 DSCNet 在 Dice (80.27%)、RDice (86.37%) 和 clDice (85.26%) 等指标上取得了最佳的分割结果。与经典分割网络 UNet 的结果相比, 我们的方法在 Dice、RDice 和 clDice 上的改进分别为 3.40%、1.89% 和 3.83%。同时, 我们使用 OF 指标来评估分

割的连续性。使用我们的方法, LAD、LCX 和 RCA (LAD、LCX 和 RCA 是冠状血管的三个主干) 的 OF 指标分别提高了 6.00%、3.78% 和 3.30%。血管连续性的提高在临床中起着至关重要的作用。

消融实验分析。以 DRIVE 数据集为例, 消融实验证明了我们的 DSCNet 和 TCLoss 的重要性。(1) 证明我们的 DSCNet 的有效性。表 I 的前五行结果显示, 我们的方法更适合细长管状结构的分割。结果表明我们提出的 DSConv 在模型中起着至关重要的作用, 有助于网络更好地捕捉细长管状结构的关键特征。(2) 证明我们的 TCLoss 的有效性。如表 I 的第六至第九行所示, 加入我们的 TCLoss 后, 不同的模型在分割的拓扑连续性方面都有所改善。结果说明我们的 TCLoss 能够准确

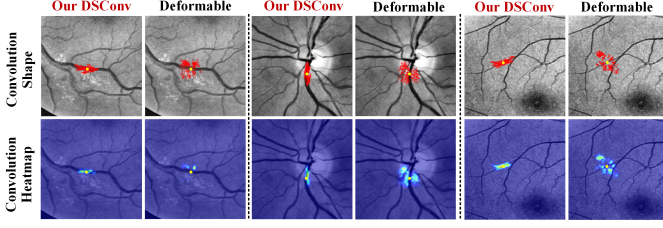


图 7: 上图: 我们在每个图像上叠加了包含 729 个点 (红色) 的 3 个图层, 以展示给定点 (黄色) 的卷积核位置和形状。下图: 热图显示了每个卷积的感兴趣区域。

约束模型关注失去拓扑连续性的细长管状结构。

B. 定性评估

我们的 DSCNet 和 TCLoss 在任意方面都具有决定性的视觉优势 (图 6)。(1) 为了展示我们的 DSCNet 的有效性。从左到右, 第三至第五列显示了不同网络在分割准确度方面的表现。由于我们的 DSCConv 能够自适应地感知细长管状结构的关键特征, 相比其他方法, 我们的模型更准确地聚焦于特殊的管状特征, 因此在管状结构分割方面表现更好。(2) 为了展示我们的 TCLoss 的有效性。从左到右, 第六至第八列显示了不同损失函数对细长管状结构分割的拓扑连续性的影响。加入我们提出的 TCLoss 后, 分割在难以处理的区域的连续性得到了显著改善。结果证实了我们的方法在复杂和多形态结构中能够提供更稳定的分割性能, 特别是在马萨诸塞州道路数据集上, 我们的模型在相邻的直道或曲线道路上实现了良好的可视化效果。更多可视化结果可在附加材料中找到。

C. 模型分析

我们的 DSCConv 能够动态地适应管状结构的形状, 而注意力也很好地匹配目标。(1) 适应管状结构的形状。图 7 的顶部显示了卷积核的位置和形状。可视化结果显示, 我们的 DSCConv 能够很好地适应管状结构并保持形状, 而可变形卷积则会偏离目标。(2) 关注管状结构的位置。图 7 的底部显示了给定点的注意力热图。结果显示, 我们的 DSCConv 最亮的区域集中在管状结构中, 这表示我们的 DSCConv 对管状结构更敏感。

D. 未来工作

我们提出的框架很好地处理了对薄壁管状结构的分割, 并成功地将形态特征与拓扑知识相结合, 引导模

型适应分割。然而, 其他形态目标是否会以类似的范例获得更好的性能仍然是一个令人激动的课题。与此同时, 更多的研究将探讨将其他类型的领域知识或拓扑分析纳入, 以进一步提高分割性能的可能性。此外, 更多的实验证明和理论验证将丰富这个主题。

VI. 结论

在这项研究中, 我们专注于管状结构的特殊特征, 并利用这一知识来引导模型在三个阶段同时增强感知: 特征提取、特征融合和损失约束。首先, 我们提出了动态蛇形卷积, 以自适应地聚焦于细长且弯曲的结构, 从而准确捕捉管状结构的特征。其次, 我们引入了一种多视角特征融合策略, 在特征融合过程中从多个角度补充对特征的关注, 确保保留来自不同全局形态的重要信息。最后, 我们提出了一种拓扑连续性约束损失, 以约束分割的拓扑连续性。我们的方法在 2D 和 3D 数据集上得到验证, 结果表明与几种方法相比, 我们的方法在管状结构分割任务上提供了更好的准确性和连续性。

参考文献

- [1] R. J. Araújo, J. S. Cardoso, and H. P. Oliveira. Topological Similarity Index and Loss Function for Blood Vessel Segmentation. *IEEE TRANSACTIONS ON MEDICAL IMAGING*, XX:1, jul 2021. 2, 3
- [2] J. Chen, Y. Lu, Q. Yu, et al. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306*, 2021. 5
- [3] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, et al. 3d u-net: Learning dense volumetric segmentation from sparse annotation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 424–432. Springer, 2016. 5
- [4] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, et al. Lipschitz functions have l_p -stable persistence. *Foundations of computational mathematics*, 10(2):127–139, 2010. 2, 5
- [5] J. Dai, H. Qi, Y. Xiong, et al. Deformable convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Oct 2017. 2, 3
- [6] S. Dong, Z. Pan, Y. Fu, et al. Deu-net 2.0: Enhanced deformable u-net for 3d cardiac cine mri segmentation. *Medical Image Analysis*, 78:102389, 2022. 1, 2
- [7] Y. He, R. Ge, J. Wu, et al. Thin semantics enhancement via high-frequency priori rule for thin structures segmentation. In *2021 IEEE International Conference on Data Mining (ICDM)*, pages 1096–1101. IEEE, 2021. 1
- [8] E. Herbert and H. John. *Computational topology: an introduction*. American Mathematical Soc, 2010. 2, 4
- [9] X. Hu, F. Li, D. Samaras, et al. Topology-preserving deep image segmentation. *Advances in neural information processing systems*, 32, 2019. 2, 5

- [10] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge. Comparing images using the hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence*, 15(9):850–863, 1993. [5](#)
- [11] Q. Jin, Z. Meng, T. D. Pham, et al. Dunet: A deformable network for retinal vessel segmentation. *Knowledge-Based Systems*, 178:149–162, 2019. [1](#), [2](#)
- [12] B. S. Ko, J. D. Cameron, R. K. Munnur, et al. Noninvasive ct-derived ffr based on structural and fluid analysis. *10(6):663–673*, 2017. [1](#)
- [13] B. Kong, X. Wang, J. Bai, et al. Learning tree-structured representation for 3d coronary artery segmentation. *Computerized Medical Imaging and Graphics*, 80:101688, 2020. [1](#), [2](#)
- [14] Y. Li, Y. Zhang, J.-Y. Liu, et al. Global transformer and dual local attention network via deep-shallow hierarchical feature fusion for retinal vessel segmentation. *IEEE Transactions on Cybernetics*, pages 1–14, 2022. [2](#)
- [15] J. K. Min, J. Leipsic, M. J. Pencina, et al. Diagnostic Accuracy of Fractional Flow Reserve From Anatomic CT Angiography. *JAMA*, 308(12):1237–1245, 09 2012. [1](#)
- [16] V. Mnih. Machine learning for aerial image labeling. University of Toronto (Canada), 2013. [5](#)
- [17] L. Mou, Y. Zhao, H. Fu, et al. Cs2-net: Deep learning segmentation of curvilinear structures in medical imaging. *Medical Image Analysis*, 67:101874, 2021. [5](#)
- [18] J. R. Munkres. *Elements of algebraic topology*. CRC press, 2018. [2](#), [4](#)
- [19] X. Qi, Y. He, G. Yang, et al. Mvsgan: Spatial-aware multi-view cmr fusion for accurate 3d left ventricular myocardium segmentation. *IEEE Journal of Biomedical and Health Informatics*, 26(5):2264–2275, 2021. [2](#)
- [20] X. Qi, G. Yang, Y. He, et al. Contrastive re-localization and history distillation in federated cmr segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 256–265, 2022. [2](#)
- [21] Y. Qi, Y. He, X. Qi, Y. Zhang, and G. Yang. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 6070–6079, October 2023. [1](#)
- [22] Y. Qi, H. Xu, Y. He, et al. Examinee-examiner network: Weakly supervised accurate coronary lumen segmentation using centerline constraint. *IEEE Transactions on Image Processing*, 30:9429–9441, 2021. [5](#)
- [23] M. Schaap, C. T. Metz, T. van Walsum, et al. Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms. *Medical Image Analysis*, 13(5):701–714, 2009. [5](#)
- [24] S. Shit, J. C. Paetzold, A. Sekuboyina, et al. cldice - a novel topology-preserving loss function for tubular structure segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 16560–16569, June 2021. [2](#), [5](#)
- [25] J. Staal, M. D. Abramoff, M. Niemeijer, et al. Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging*, 23(4):501–509, 2004. [5](#)
- [26] A. A. Taha and A. Hanbury. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. *BMC medical imaging*, 15(1):1–28, 2015. [5](#)
- [27] D. Wang, Z. Zhang, Z. Zhao, et al. Pointscatter: Point set representation for tubular structure extraction, 2022. [1](#), [2](#), [5](#)
- [28] Y. Wang, X. Wei, F. Liu, et al. Deep distance transform for tubular structure segmentation in ct scans. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. [2](#)
- [29] C.-C. Wong and C.-M. Vong. Persistent homology based graph convolution network for fine-grained 3d shape segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 7098–7107, October 2021. [2](#), [3](#), [4](#), [5](#)
- [30] X. Yang, Z. Li, Y. Guo, et al. DCU-net: a deformable convolutional neural network based on cascade U-net for retinal vessel segmentation. *Multimedia Tools and Applications*, 81(11):15593–15607, may 2022. [1](#), [2](#), [5](#)
- [31] F. Yu, V. Koltun, and T. Funkhouser. Dilated residual networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017. [2](#)
- [32] X. Zhang, J. Zhang, L. Ma, et al. Progressive deep segmentation of coronary artery via hierarchical topology learning. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, pages 391–400, Cham, 2022. Springer Nature Switzerland. [2](#)
- [33] C. Zhao, W. Zhu, and S. Feng. Superpixel guided deformable convolution network for hyperspectral image classification. *IEEE Transactions on Image Processing*, 31:3838–3851, 2022. [2](#)
- [34] G. Zhao, K. Liang, C. Pan, et al. Graph Convolution Based Cross-Network Multi-Scale Feature Fusion for Deep Vessel Segmentation. *IEEE Transactions on Medical Imaging*, pages 1–1, 2022. [2](#)