面向半监督血管造影视频的不确定性感知概念与运动分割

Yu Luo[1] *、Guangyu Wei[2,5] *、Yangfan Li[3] (✉)、Jieyu He[4], 和 Yueming Lyu[5]

[1] 中国海洋大学数学科学学院，中国[2] 中国海洋大学海德学院，中国[3] 中南大学计算机科学与工程学院，中国 liyangfan37@csu.edu.cn[4] 中南大学湘雅二医院，中国[5] 南京大学智能科学与技术学院，中国

arXiv:2603.00881v1 [cs.CV] 1 Mar 2026

摘要。 从 X 射线冠状动脉造影（XCA）序列中分割主要冠状动脉，对冠状动脉疾病的诊断至关重要。然而，由于边界模糊、辐射对比度不一致、运动模式复杂以及缺乏训练所需的标注图像等问题，这项任务面临挑战。尽管半监督学习（SSL）可以减轻标注负担，但传统方法难以应对复杂的时间动态，也难以可靠地量化不确定性。为应对这些挑战，我们提出了基于 SAM3 的师生框架，结合运动感知一致性与渐进式置信度正则化（SMART），这是一种面向 X 射线血管造影视频的半监督血管分割方法。首先，我们的方法利用 SAM3 独特的可提示概念分割设计，并创新性地构建基于 SAM3 的师生框架，以充分发挥教师和学生模型的性能潜力。其次，我们通过整合血管掩膜变形技术和运动一致性损失来建模复杂的血管动态，从而提升分割效果。针对边界模糊、对比度极低导致教师预测不可靠的问题，我们进一步提出渐进式置信度感知一致性正则化，以降低产生不可靠输出的风险。在来自不同机构的三个 XCA 序列数据集上进行的大量实验表明，SMART 仅需显著更少的标注便能达到最先进性能，因此在标注数据稀缺的真实临床应用中尤具价值。我们的代码可在此处获取：https://github.com/qimingfan10/SMART。

关键词： 半监督学习 · 冠状动脉造影 · 一致性正则化 · 可提示概念分割。

1 引言

冠状动脉疾病（CAD）仍是全球致病和死亡的首要原因，且大多数心血管疾病死亡病例发生在中低收入国家和地区 [22]。X射线冠状动脉造影（XCA）被视为临床决策中的黄金标准 [26]。对于血管疾病的自动诊断，从连续的XCA图像序列中精确分割冠状动脉至关重要 [30]。

近年来，用于 XCA 序列分割的基于深度学习的方法已展现出良好前景 [7, 24, 8]。然而，获取标注数据极其昂贵且耗时

* 这些作者对本研究的贡献相同。

在临床场景中，导致未标注数据的体量显著大于标注样本。这一制约引发了人们对半监督学习（SSL）方法的浓厚兴趣 [6, 5, 31], ，这些方法利用大量未标注数据以及有限的标注数据来提升模型性能。近来，Segment Anything Model（SAM） [14]及其后续模型 SAM2 [21] 展现出出色的少样本分割能力 [2]„，为标注稀缺场景中生成伪标签提供了一种有前景的方法。然而，SAM 无法直接应用于医学图像分割场景。为解决这一局限，现有研究提出将精确的预测点、边界框或可学习特征提示融入 SAM 系列，以提高未标注数据伪标签的准确性 [18, 33, 32, 9, 19]。

由于不同机构使用的 X 射线成像系统各不相同，依赖几何线索或可学习特征提示的方法往往难以在多样化的临床场景中有效泛化。近期，创新模型 [3, 10, 12] 开始探索整合文本信息以辅助分割任务。值得注意的是，SAM3 [1] 通过聚焦于概念提示（即指代视觉对象的语义短语），引入了架构上的转变。这种设计确保系统能够真正理解视觉结构的语义。然而，直接将 SAM3 应用于 XCA 序列会忽略时间依赖关系和跨模态对齐，可能导致分割结果在时间上不一致，并使血管结构的识别更加复杂 [20, 28, 16]。XCA 固有的局限性进一步加剧了这些挑战，包括非自主器官运动导致目标形态和尺度随时间显著变化，以及对比度极低、信噪比低等问题 [4, 25, 15]。

为应对前文提到的挑战，我们提出 SMART（基于 SAM3 的不确定性感知置信度正则化与运动一致性教师-学生架构），这是一种用于 X 射线视频血管分割的半监督均值教师方法。我们首先利用 SAM3 在文本描述引导下进行可提示概念分割的能力，摆脱对几何提示的依赖，从而能够以更高的准确度捕捉局部细节和复杂边界。考虑到冠状动脉对比度较低且存在运动引起的模糊，伪标签的生成本质上可能不可靠且含有噪声。为解决这一问题，我们设计了一种新颖的渐进式置信度感知一致性正则化方法。该方法使框架能够在不同训练阶段动态调整对未标注数据监督信号的利用，同时增强预测的不变性。因此，即使使用可靠性较低的伪标签进行学习，也能确保获得更可靠、稳定的分割结果。鉴于心脏运动和血管血流模式的独特特点，我们引入了掩码变形操作、运动一致性损失和光流连贯性损失，从而确保分割结果在时间上的一致性和细节丰富度，同时有效地区分血管前景与背景。我们在三个 X 射线冠状动脉造影视频数据集上对所提方法进行了广泛评估，结果表明，该方法优于现有方法，尤其是在标注数据极其有限的情况下。具体而言，仅使用 16 个标注视频、且每个视频仅有一或两帧标注时，与多种强基线方法相比，我们的方法取得了超过 6% 的 Dice 系数提升。

2 方法

概述。 半监督学习分割的目标是利用有限的标注数据 $D_l = (\mathbf{x}_i,\ \mathbf{y}_i)_{i=1}^{N}$ 和大量未标注数据$D_u = (\mathbf{x}_i)_{i=N+1}^{N+M}$，使性能接近全监督模型，其中 $N \ll M$。这里，$\mathbf{x}_i \in \mathbb{R}^{H\times W}$表示一幅 $H\times W$图像，而$\mathbf{y}_i \in \{0,1\}^{H\times W}$

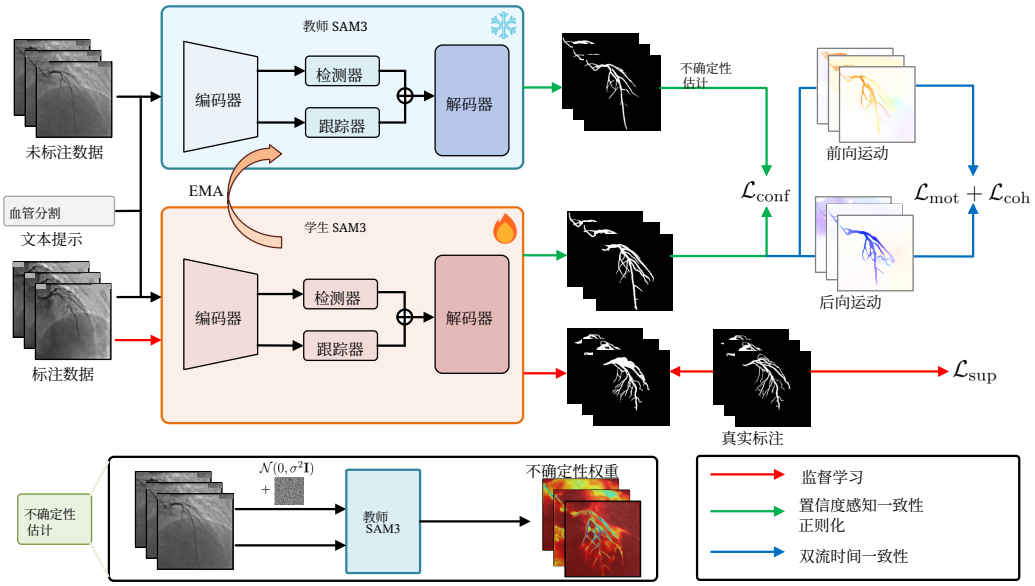


图1：所提出方法的概览。SMART（基于 SAM3 的教师-学生架构运动感知置信度正则化）模型通过置信度感知一致性损失 $\mathcal{L}_{\text{conf}}$提升教师模型输出的可靠性。该模型通过双流时间一致性损失 $\mathcal{L}_{\text{opti}}$ 和 $\mathcal{L}_{\text{coh}}$融入对视频序列的时间理解，同时借助监督损失 $\mathcal{L}_{\text{sup}}$有效利用已标注数据。

表示对应标注数据的标注。图 1 展示了 SMART 架构的概览，该架构采用教师模型 f_{Θ_T} 和学生模型 f_{Θ_S}，两者均利用文本提示进行分割。训练过程分为两个依次进行的阶段：文本驱动的分割微调和运动感知半监督学习。具体而言，在第一阶段，我们对标注数据 D_l上的教师模型 f_{Θ_T} 采用微调策略，在融入医学领域知识的同时保留通用的视觉与语言特征。在第二阶段，冻结微调后的教师模型，并由其指导学生模型处理未标注数据 D_u。关键的是，我们提出了双流时间一致性策略、运动一致性损失以及置信度感知的一致性正则化，以增强学生模型对视频序列中时间信息的有效利用，并提高其对不可靠教师输出的鲁棒性。在推理阶段，我们仅使用学生模型 f_{Θ_S} 进行预测。

2.1 文本驱动的分割微调

未标注数据的利用是半监督学习（SSL）的关键组成部分，教师模型生成的伪标签为学生模型提供了重要指导，使其能够从有限的标注数据中有效学习。尽管 SAM3 在自然图像领域取得了显著成功，但它缺乏对医学领域特定语义的理解 [16]。在第一阶段，我们采用一项基础技术来满足这一需求：将视觉指令微调的概念 [11] 转移到 SAM3。通过利用标注数据进行初步优化，这一步使教师

模型能够更有效地适应目标任务。教师 SAM3 保留其原有架构，仅对文本提示进行训练。具体而言，我们使用与血管分割相关的多样化提示，对 SAM3 的图像编码器、文本编码器和检测器进行微调，同时冻结其他组件。训练期间，我们优化这些文本提示，以分割感兴趣的目标对象。微调优化过程定义如下：

$$\mathcal{L}_{\text{ft}}(p(\mathbf{x}_i^l), \mathbf{y}_i^l) = \lambda_1 \mathcal{L}_{\text{Dice}}(p(\mathbf{x}_i^l), \mathbf{y}_i^l) + \lambda_2 \mathcal{L}_{\text{Bce}}(p(\mathbf{x}_i^l), \mathbf{y}_i^l), \tag{1}$$

其中 $\mathrm{p}\ \ p(\mathbf{x}_i^l) = f_{\Theta_T}(\mathbf{x}_i^l))$ 表示教师 SAM3 对标注图像的预测，$\mathbf{y}_i^l$ 为对应的真实掩膜。 $\mathcal{L}_{\text{Dice}}$ 和 $\mathcal{L}_{\text{Bce}}$ 分别表示 Dice 损失和交叉熵损失。

2.2 置信度感知一致性正则化

利用教师生成的伪标签，可以应对借助未标注数据提升半监督学习性能这一挑战。该方法成功的关键在于为未标注数据生成可靠的预测。然而，冠状动脉造影图像固有的局限性，如低信噪比和低对比度，可能导致 SAM3 输出不可靠，表现为低质量区域中准确率低、方差高。为缓解这一问题，关键在于增强不同噪声条件下的输出一致性，并在训练的不同阶段自适应地降低不确定区域的权重。在理想情况下，目标血管的预测应在不同噪声扰动下保持稳定，同时随着训练的进行，模型逐步加强对不确定区域（例如低对比度远端血管）的学习。基于上述考虑，我们提出一种新颖的置信度感知一致性正则化。具体而言，教师首先接收视频帧 $\mathbf{X} \in \mathbb{R}^{C\times H\times W}$ 的多个噪声扰动版本。扰动采样为 $\boldsymbol{\epsilon}^{(i)} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$，其中 $i = 1, \ldots, N$，从而生成 N 个独立预测，记为 $\mathcal{P} = \{\hat{}p_i\}_{i=1}^N = \{f_{\Theta_T}(\mathbf{X} + \epsilon^{(i)})\}_{i=1}^N$。这种噪声注入模拟了冠状动脉造影图像中常见的质量变化。在这些扰动下，预测 $\hat{}p_i$ 应在不同噪声实例之间保持一致。由于多个预测的集成通常能产生更稳健的结果，我们采用平均预测 $\bar{}\mathbf{P} = 1/N \sum_{i=1}^N \hat{}p_i$ 作为学生模型的可靠指导。

同时，为确保模型在不确定性较高的区域加强监督，我们将不确定性权重计算为$\mathcal{U} = 1/N \sum_{i=1}^N \left(f_{\Theta_T}(\mathbf{X}^{(i)}) - \bar{}\mathbf{P}\right)^2$。随后，根据学生模型的输出$\mathbf{S} = f_{\Theta_S}(\mathbf{X})$和平均预测$\bar{}\mathbf{P}$得出基本一致性距离，可表示为：

$$\mathcal{D}(x, y) = \left(\sigma(S(x, y)) - \sigma(\bar{P}(x, y))\right)^2, \tag{2}$$

其中，$\sigma(\cdot)$ 表示 Sigmoid 函数。通过使用推导出的不确定性权重对基本一致性距离进行加权，SMART 实现了可调节的监督强度，为不确定性较高的区域分配更高的重要性，从而推动性能提升。置信度感知一致性损失可表示为：

$$\mathcal{L}_{\text{conf}} = \frac{\sum_{x=1}^{H}\sum_{y=1}^{W} \mathcal{D}(x, y)\mathcal{U}(x, y)}{\sum_{x=1}^{H}\sum_{y=1}^{W} \mathcal{U}(x, y) + N\eta} + \frac{\beta}{N}\sum_{x=1}^{H}\sum_{y=1}^{W} \mathcal{U}(x, y), \tag{3}$$

其中，β 是不确定性正则化权重，η 是一个较小的稳定因子（例如 10^{-6}），用于避免除以零。

2.3 双流时间一致性

与静态医学图像相比，XCA视频展现出更为复杂的时间动态。为增强血管分割的时间连续性，并防止相邻帧之间的预测发生突变，我们提出了一种基于光流的双流时间一致性策略。该方法有效利用了视频序列中固有的时空信息。

为减轻运动建模中与单向光流相关的确认偏误，我们采用双流估计策略。具体而言，我们利用预训练的光流估计器SEA-RAFT [27], 计算连续帧$\mathbf{F}_{t\to t+1}$ 和$\mathbf{F}_{t+1\to t}$ 的前向光流与后向光流$\mathbf{X}_t$以及$\mathbf{X}_{t+1}$。由于当前XCA视频 [29],中的帧数有限，血管的拓扑结构在这一短时间段内保持不变。基于估计得到的光流，我们引入运动一致性损失函数，以确保预测掩膜在连续帧之间保持时间连贯性。该过程可表述为：

$$\mathcal{L}_{\text{opti}} = \frac{1}{2N}\sum_{x=1}^{H}\sum_{y=1}^{W}\left(\left[\mathbf{S}_t(x,y) - \mathcal{W}(\mathbf{S}_{t+1},\mathbf{F}_{t\to t+1})(x,y)\right]^2 + \left[\mathbf{S}_{t+1}(x,y) - \mathcal{W}(\mathbf{S}_t,\mathbf{F}_{t+1\to t})(x,y)\right]^2\right),$$

其中 $\mathcal{W}(\mathbf{S},\mathbf{F})(x,y) = \mathbf{S}\big(x+\mathbf{F}_u(x,y),\ y+\mathbf{F}_v(x,y)\big)$ 表示掩码变形操作。 (4)

尽管运动一致性损失确保了像素级时间对齐，但在血管边界处仍面临固有的歧义。为了解决这一问题，我们提出了一种光流连贯性损失，使模型能够根据运动模式的差异区分前景与背景。该方法会惩罚边界点偏离血管主体主导运动的情况。此过程可表示为：

$$\begin{cases} \bar{\Phi} = \dfrac{\sum_{x=1}^{W}\sum_{y=1}^{H}\hat{S}_t(x,y)\cdot F_{t\to t+1}(x,y)}{\sum_{x=1}^{W}\sum_{y=1}^{H}\hat{S}_t(x,y)+\epsilon}, \\ \mathcal{L}_{\text{coh}} = \dfrac{\lambda_{\text{coh}}}{\sum_{x=1}^{H}\sum_{y=1}^{W}\hat{S}_t(x,y)}\sum_{x=1}^{H}\sum_{y=1}^{W}\hat{S}_t(x,y)\left\|\mathbf{F}_{t\to t+1}(x,y)-\bar{\mathbf{\Phi}}\right\|_2^2, \end{cases} \tag{5}$$

其中，$\hat{}\ S_t(x,y)$ 表示第 t 帧的学生预测掩膜。

在本研究中，监督分割损失由 Dice 系数损失和交叉熵（CE）损失组成，可定义为 $\mathcal{L}_{\text{seg}}(y,\ \hat{}y) = \lambda_{\text{Dice}}\mathcal{L}_{\text{Dice}}(y,\ \hat{}y) + \lambda_{\text{Bce}}\mathcal{L}_{\text{Bce}}(y,\ \hat{}y)$，其中 $\hat{}y$ 和 y 分别表示预测和真实标注。总之，半监督学习的整体优化目标可写为：

$$\mathcal{L}_{\text{all}} = \lambda_{\text{Dice}}\mathcal{L}_{\text{Dice}} + \lambda_{\text{Bce}}\mathcal{L}_{\text{Bce}} + \lambda_{\text{conf}}\mathcal{L}_{\text{conf}} + \lambda_{\text{opti}}\mathcal{L}_{\text{opti}} + \lambda_{\text{coh}}\mathcal{L}_{\text{coh}}. \tag{6}$$

表 1：XCAV 和 CAVSA 数据集上的定量结果。

方法	XCAV						CAVSA					
	#实验室	Dice相似系数↑	NSD↑	clDice↑	特异度↑	Sen↑	#Lab	Dice相似系数↑	NSD↑	clDice↑	特异度↑	Sen↑
UNet [23]	16/111	70.80	60.12	69.24	97.75	66.33	16/1061	64.19	54.25	70.27	**99.08**	53.36
MedSAM2 [18]	16/111	22.37	21.07	3.57	84.14	32.01	16/1061	33.13	22.05	33.03	86.30	44.31
SAM3 [1]	16/111	42.73	30.86	34.51	98.11	35.59	16/1061	30.82	25.25	30.14	92.99	30.18
Denver [29]	16/111	73.30	51.50	70.40	98.50	65.60	16/1061	76.53	68.39	79.17	98.90	75.28
KnowSAM [9]	16/111	72.76	77.67	77.08	98.07	66.77	16/1061	72.76	60.57	76.12	98.07	66.77
CPC-SAM [19]	16/111	77.90	81.97	79.15	**98.53**	72.09	16/1061	77.90	65.34	78.28	98.53	72.09
SMART（我们的方法）	16/111	**84.39**	**89.13**	**83.01**	86.25	**81.83**	16/1061	**91.00**	**91.71**	**97.73**	91.70	**96.30**

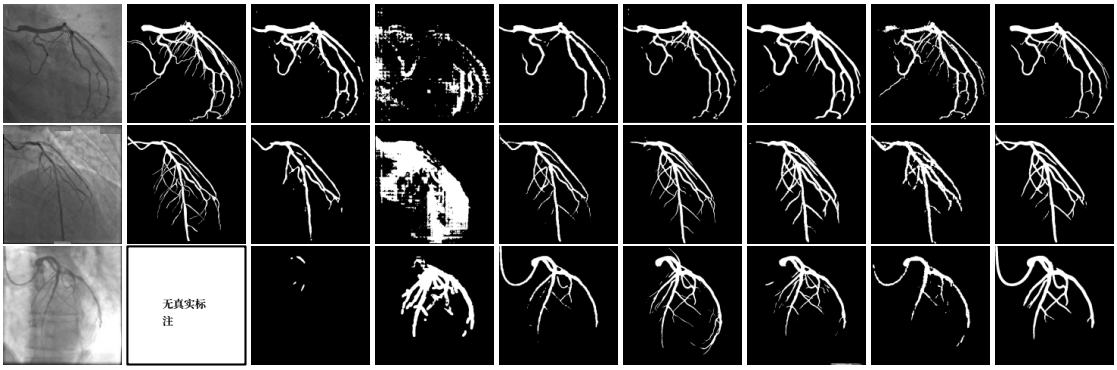


图 2：血管分割的可视化结果。

3 实验

3.1 实验设置

数据集。我们在两个公开数据集（XCAV数据集 [29] 和 CADICA数据集 [13], ）以及一个名为 CAVSA（冠状动脉造影视频分割）数据集的私有数据集上评估所提出的方法。XCAV数据集包含来自 59 名患者的 111 段完整 X 射线冠状动脉造影视频记录，每段视频在造影剂最显著的时刻标注一到两帧；而 CADICA数据集包含来自 42 名患者的 662 段完整记录，没有对应的真实标注，用于评估泛化性能。我们收集了来自 121 名患者的 1,061 段完整冠状动脉 X 射线造影视频序列。每段序列平均包含 79 帧，每帧的高分辨率为 512×512 像素。我们邀请经验丰富的放射科医生标注血管区域，为 16 段视频序列提供完整标注，并在其他序列中造影剂最显著的时刻标注一到两帧。遵循先前的工作，我们按 8:2 的比例将 XCAV 和 CAVSA 数据集划分为训练集和测试集。

实现细节和评估指标 SMART由PyTorch实现，并在NVIDIA A40 GPU上训练。我们采用AdamW optimizer[17] ，初始学习

表 2：关键组件的消融研究。

设置			XCAV					CAVSA				
TPT	CCR	DSTC	Dice相似系数↑	NSD ↑	clDice ↑	特异度 ↑	敏感度 ↑	Dice相似系数↑	NSD ↑	clDice ↑	特异度 ↑	敏感度 ↑
	✓	✓	82.38	83.42	79.84	**99.16**	77.26	78.87	67.35	81.17	99.42	73.80
✓		✓	76.71	78.29	79.86	97.46	78.69	25.82	21.24	32.65	**99.77**	19.38
✓	✓		76.24	77.27	78.53	98.05	74.33	47.77	40.71	50.37	99.59	38.46
✓	✓	✓	**84.39**	**89.13**	**83.01**	86.25	**81.83**	**91.00**	**91.71**	**97.73**	91.70	**96.30**



图 3：消融实验的可视化比较。

率为1e-4，权重衰减为0.01。批次大小设为4，迭代次数设为6k。随后，我们逐步引入 $\mathcal{L}_{\text{Dice}}$ ($\lambda_1 = 0.05$)、 $\mathcal{L}_{\text{Bce}}$ ($\lambda_2 = 0.95$)、 $\mathcal{L}_{\text{Dice}}$ ($\lambda_{\text{Dice}} = 0.5$)、$\mathcal{L}_{\text{Bce}}$ ($\lambda_{\text{Bce}} = 0.5$)、 $\mathcal{L}_{\text{conf}}$ ($\lambda_{\text{conf}} = 0.5$)、 $\mathcal{L}_{\text{opti}}$ ($\lambda_{\text{opti}} = 0.3$)和 $\mathcal{L}_{\text{coh}}$ ($\lambda_{\text{coh}} = 0.2$)。不确定性估计中，对每张图像施加的噪声扰动次数设为 $N = 8$。我们对教师模型和学生模型采用不同的数据增强策略。学生模型接受较弱的增强：旋转($\pm 5°$)、缩放(0.95–1.05)和低噪声($\sigma = 0.01$)。相比之下，教师接受较强的增强：旋转($\pm 15°$)、缩放(0.85–1.15)和较高噪声(σ =0.03)。在预处理阶段，我们使用SEA-RAFT计算光流 [27] 。遵循先前的研究 [29], ，我们采用五项评估指标，包括Dice相似系数(DSC)、归一化表面距离(NSD)、中心线 Dice(clDice)、特异度(Spe)和敏感度(Sen)。

3.2 与最先进方法的比较。

我们将所提出的方法与最先进（SOTA）的基于 SAM 的半监督学习（SSL）方法进行比较，包括 KnowSAM [9],和 CPC-SAM [19]。此外，我们还与仅使用标注数据训练的有监督对应方法进行比较，即 UNet [23], 、 MedSAM2 [18], 和 SAM3 [1]。此外，我们还使用专为 XCA 视频分割设计的方法 Denver [29], 开展实验。表 1 展示了定量比较结果。在 XCAV数据集上的实验结果清楚地表明，与多种方法相比，我们提出的 SMART 方法具有优势。仅使用 14% 的标注视频时，SMART 的 Dice相似系数达到 84.39%，clDice 达到 83.01%，分别比次优方法 CPC-SAM 高出 6.49% 和 3.86%。仅使用 1.5% 的标注数据时，我们的方法在 CAVSA 数据集上的 Dice相似系数提升了 13.1%。图 2 中的可视化比较进一步验证了我们方法的优越性。

我们进一步在 CAD- ICA 数据集上评估了 SMART 与对比方法的泛化能力，如图 2 所示。纯监督方法受训练数据领域的限制，因而难以有效泛化。基于 SAM 的半监督方法依赖几何先验或可学习特征提示，其泛化性能同样有限。相比之下，SMART 利用 SAM3 的可提示概念分割能力，在多样的临床场景中展现出更优的灵活性与泛化能力。

表 3：不同数量的噪声扰动对XCAV数据集的影响。

噪声	Dice相似系数(%)↑	NSD(%)↑	clDice(%)↑	特异度(%)↑	敏感度(%)↑
2	83.59	88.45	81.82	85.38	79.91
4	83.73	87.53	82.02	85.14	81.76
6	83.79	88.40	82.51	85.47	79.41
8	**84.39**	**89.13**	**83.01**	**86.25**	**81.83**

3.3 消融实验

文本驱动分割微调的有效性。我们在半监督学习之前对教师模型进行微调。为证明可提示分割微调的有效性，我们将本方法与点提示SAM3基线进行比较，该基线沿血管骨架以固定间隔提取中心点作为几何先验提示。如表 2 和图 3（a）所示，在 XCAV 和 CAVSA 数据集上，Dice系数分别从 82.38% 和 78.87% 提升至 84.39% 和 91.00%。

置信度感知一致性正则化的有效性。 我们进行了一项实验，通过从训练流程中移除置信度感知一致性损失来评估其效果。如表 2所示，分割性能显著下降，Dice 系数降低了43.23%。这表明，若不对该教师 SAM3的输出进行正则化，分割结果会受到不利影响。图 3（b）中的对比进一步证明了该正则化项的有效性。

双流时间一致性的有效性。 我们在表 2中进一步验证了双流时间一致性损失 $\mathcal{L}_{\text{opti}}$ 和 $\mathcal{L}_{\text{coh}}$ 的有效性。具体而言，clDice得分提高了约39%，表明分割的空间连通性获得了显著提升。结果清楚地表明，利用视频中的时间信息能够有效增强分割连续性，并有助于避免不连通或过度分割的预测。

噪声扰动数量的影响。表3分析了不同噪声扰动数量对 SMART 的影响，评估了其对置信度感知一致性正则化的贡献。结果表明，扰动数量发挥着关键作用：较多的扰动能够更准确地估计教师输出中的不确定性，从而直接影响评估指标。具体而言，随着扰动数量的增加，XCAV数据集上的 clDice 系数从 81.82% 提升至 83.01%。

4 结论

在本研究中，我们开发了一种基于SAM3的师生框架，用于半监督医学图像分割，并利用SAM3的可提示概念分割能力。首先，我们实施视觉指令微调策略，通过微调使教师编码器的解剖语义与解码器生成的分割掩码对齐，从而为学生 SAM3提供有效指导。接下来，我们的框架引入置信度感知一致性与双流时间一致性，以应对教师输出的不可靠性，并利用视频序列中的运动信息来提升分割性能。通过对三个数据集的对比实验，我们验证了所提出方法的有效性与泛化能力，展示了其增强临床适用性的潜力。

参考文献

1. Carion, N., Gustafson, L., Hu, Y.T., Debnath, S., Hu, R., Suris, D., Ryali, C., Alwala, K.V., Khedr, H., Huang, A., 等：Sam 3：基于概念的万物分割。arXiv预印本 arXiv:2511.16719（2025） 2. Chen, C., Miao, J., Wu, D., Zhong, A., Yan, Z., Kim, S., Hu, J., Liu, Z., Sun, L., Li, X., Liu, T., Heng, P.A., Li, Q.：Ma-sam：面向三维医学图像分割的模态无关的 SAM 适配。《医学图像分析》**98**，103310 103310 3 2024 10 1016 2024（）。https://doi.org/ . /j.media. . . Du, Y., Bai, F., Huang, T., Zhao, B.：Segvol：通用交互式体积医学图像分割。载于：Globerson, A., Mackey, L., Belgrave, D., Fan, A., Paquet, U., Tomczak, J., Zhang, C.（编）神经信息处理系统进展。第37卷，第110746–110783页。Curran Associates, Inc.（2024 3516 4 10 52202 079017 ）。https://doi.org/ . / - . Felfelian, B., Fazlali, H.R., Karimi, N., Soroushmehr, S.M.R., Samavi, S., Nallamothu, B., Najarían, K.：低对比度X 射线血管造影图像中的血管分割。载于：2016年IEEE 国际图像处理会议（ICIP）。第375 7532382 5 379 2016 10 1109 2016 –（）。https://doi.org/ . /ICIP. . . Gu, Y., Sun, Z., Chen, T., Xiao, X., Liu, Y., Xu, Y., Najman, L.：面向半监督医学图像分割的双结构感知图像滤波。《医学图像分析》**99**，103364（2025）。https://doi.org/10 103364 6 1016 2024 . /j.media. . . Han, K., Sheng, V.S., Song, Y., Liu, Y., Qiu, C., Ma, S., Liu, Z.：面向医学图像分割的深度半监督学习：综述。《专家系统与应用》**245**，123052（2024）。https://doi.org/10 123052 7 1016 2023 . /j.eswa. . . Hao, D., Ding, S., Qiu, L., Lv, Y., Fei, B., Zhu, Y., Qin, B.：基于深度通道注意力网络的序列血管分割。《神经网络》**128**，172–187（2020）。https://doi.org/10 005 8 1016 2020 05 . /j.neunet. . . . He, H., Banerjee, A., Choudhury, R.P., Grau, V.：利用时间信息的基于深度学习的X 射线血管造影冠状动脉血管分割。《医学图像分析》**102**，103496（2025）。https://doi.org/10 103496 9 1016 2025 . /j.media. . . Huang, K., Zhou, T., Fu, H., Zhang, Y., Zhou, Y., Gong, C., Liang, D.：面向半监督医学图像分割的可学习提示SAM诱导知识蒸馏。IEEE 医学影像汇刊**44**3530097 10 5 2295 2306 2025 10 1109 2025（），–（）。https://doi.org/ . /TMI. . . Huang, Z., Jiang, Y., Zhang, R., Zhang, S., Zhang, X.：Cat：协调解剖学文本提示以进行多器官与肿瘤分割。载于：神经信息处理系统进展。第37卷，0119 11 3588 3610 2024 10 52202 079017 第 –（）。https://doi.org/ . / - . Jia, M., Tang, L., Chen, B.C., Cardie, C., Belongie, S., Hariharan, B., Lim, S.N.：视觉提示调优。载于：计算机视觉——ECCV 2022：第17届欧洲会议，以色列特拉维夫，2022年10月23–27日，会议论文集，第XXXIII部分。第709–727页。Springer-Verlag（2022）。https://doi.org/10.1007/978-3-031- 19827-4 41¯12 . Jiang, Y., Huang, Z., Zhang, R., Zhang, X., Zhang, S.：Zept：通过查询解耦与自提示实现零样本泛肿瘤分割。载于：IEEE/CVF 计算机视觉与模式识别会议（CVPR）论文集。第11386–11397页（2024年6月） 13. Jiménez-Partinen, A., Molina-Cabello, M.A., Thurnhofer-Hemsi, K., Palomo, E.J., Rodríguez-Capitán, J., Molina-Ramos, A.I., Jiménez-Navarro, M.：Cadica：一个利用侵入性冠状动脉造影检测冠状动脉疾病的新数据集。《专家系统》**41**(12)，e13708（2024）。https://doi.org/10 13708 14 1111. /exsy. . Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., 等：万物分割。载于：IEEE/CVF 国际计算机视觉会议论文集。第4015–4026页（2023） 15. Lin, C.Y., Ching, Y.T.：利用电影式 X 射线血管造影图像提取冠状动脉树。《生物医学工程：应用、基础与通信》**17**，111–120（2005） 16. Liu, A., Xue, R., Cao, X.R., Shen, Y., Lu, Y., Li, X., Chen, Q., Chen, J.：Medsam3：深入探究结合医学概念的万物分割。arXiv预印本 arXiv:2511.19046（2025）

17. Loshchilov, I., Hutter, F.: 解耦权重衰减正则化。载于：第七届国际学习表征会议，ICLR 2019，美国路易斯安那州新奥尔良，2019 年 5 月 6—9 日。OpenReview.net（2019），https://openreview.net/forum?id = 6 7 Bkg RiCqY。Ma, 18 J., Yang, Z., Kim, S., Chen, B., Baharoon, M., Fallahpour, A., Asakereh, R., Lyu, H., Wang, B.: Medsam2：在三维医学图像和视频中实现万物分割。arXiv 预印本 arXiv:2504.03600（2025）19. Miao, J., Chen, C., Zhang, K., Chuai, J., Li, Q., Heng, P.A.: 面向半监督医学图像分割的交叉提示一致性与 Segment Anything 模型。载于：医学图像计算与计算机辅助干预——MICCAI 2024 论文集。第 LNCS 15011 卷，第 167–177 页（2024 年 10 月）20. Painchaud, N., Duchateau, N., Bernard, O., Jodoin, P.M.: 通过强制时间一致性实现超声心动图分割。《IEEE 医学影像汇刊》**41**(10), 2867–2878（2022）。https://doi.org/10 3173669 21 1109 2022 . /TMI. . . Ravi, N., Gabeur, V., Hu, Y.T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., 等：Sam 2：对图像和视频进行万物分割。arXiv 预印本 arXiv:2408.00714（2024）22. Rehman, S., Rehman, E., Ikram, M., Jianglin, Z.: 心血管疾病（CVD）：评估、预测与政策影响。《BMC 公共卫生》**21**(1), 1299（2021）23. Ronneberger, O., Fischer, P., Brox, T.: U-net：用于生物医学图像分割的卷积网络。载于：医学图像计算与计算机辅助干预国际会议。第 234–241 页。Springer（2015）24. Tang, X., Zhang, H., Xie, B., Liu, X.: X 射线冠状动脉造影序列中主要冠状动脉的时间一致性分割。《专家系统与应用》**271**, 126591（2025）。https://doi.org/10 126591 25 1016 2025 . / j.eswa. . . Toth, G.G., Hamilos, M.I., Pyxaras, S., Mangiacapra, F., Nelis, O., Vroey, F.D., Serafino, L.D., Muller, O., Mieghem, C.V., Wyffels, E., Heyndrickx, G.R., Bartunek, J., Vanderheyden, M., Barbato, E., Wijns, W., de Bruyne, B.: 血管造影概念的演变：4000 例冠状动脉狭窄中的血流储备分数不一致。《欧洲心脏杂志》**35**(40), 2831–2838（2014）26. Wan, T., Feng, H., Tong, C., Li, D., Qin, Z.: 基于 X 射线血管造影的冠状动脉狭窄自动识别与分级。《生物医学中的计算机方法与程序》**167**, 13–22（2018）27. Wang, Y., Lipson, L., Deng, J.: Sea-raft：用于光流的简单、高效、准确 RAFT。载于：计算机视觉——ECCV 2024：第 18 届欧洲会议，意大利米兰，2024 年 9 月 29 日至 10 月 4 日，论文集，第 VII 部分。第 36 54 2024 10 1007 978 3 031 72667 5 _3 – 28 页。柏林、海德堡（）。https://doi.org/ . / - - - - . Wei, H., Cao, H., Cao, Y., Zhou, Y., Xue, W., Ni, D., Li, S.: 通过外观与形状协同学习实现超声心动图的时间一致性分割。载于：医学图像计算与计算机辅助干预——MICCAI 2020：第 23 届国际会议，秘鲁利马，2020 年 10 月 4—8 日，论文集，第 II 部分。第 623–632 页。柏林、海德堡（2020）。https://doi.org/10.1007/978- 3-030-59713-9 6029 ¯ . Wu, C.H., Chen, S.H., Hu, C.Y., Wu, H.Y., Chen, K.H., Chen, Y.Y., Su, C.H., Lee, C.K., Liu, Y.L.: Denver：用于无监督视频血管分割的可变形神经血管表示。载于：IEEE/CVF 计算机视觉与模式识别会议（CVPR）论文集。第 15682–15692 页（2025 年 6 月）30. Xia, S., Zhu, H., Liu, X., Gong, M., Huang, X., Xu, L., Zhang, H., Guo, J.: X 射线冠状动脉造影图像序列的血管分割。《IEEE 生物医学工程汇刊》**67**(5), 1338–1348（2020）31. Yang, Y., Sun, G., Zhang, T., Wang, R., Su, J.: 通过由弱到强扰动一致性与边缘感知对比表征实现半监督医学图像分割。《医学图像分析》**101**, 103450 103450 32 2025 10 1016 2024（）。https://doi.org/ . / j.media. . . Zhang, Y., Lv, B., Xue, L., Zhang, W., Liu, Y., Fu, Y., Cheng, Y., Qi, Y.: Semisam+：在基础模型时代重新思考半监督医学图像分割。《医学图像分析》（2025）33. Zhang, Y., Yang, J., Liu, Y., Cheng, Y., Qi, Y.: Semisam：通过 SAM 辅助一致性正则化增强半监督医学图像分割。载于：2024 IEEE 生物信息学与生物医学国际会议（BIBM）（2024）