

一种用于图谱构建的高效模型驱动逐步方法

Ziwei Zou^{1,3,4}, Bei Zou¹, Xiaoyan Kui^{1†}, Wenqi Lu², Haoran Dou^{3,4}, Arezoo Zakeri^{3,4}, Timothy Cootes³, Alejandro F Frangi^{3,4}, Jinming Duan^{3,4†}

¹ School of Computer Science and Engineering, Central South University, China

² Department of Computing and Mathematics, Manchester Metropolitan University, UK

³ Division of Informatics, Imaging and Data Sciences, University of Manchester, UK

⁴ Centre for Computational Imaging and Modelling in Medicine, University of Manchester, UK

Abstract

图谱构建是医学图像分析的基础，为如群体级解剖建模等任务提供了标准化的空间参考。尽管基于数据驱动的配准方法在成对设置中最近显示出有潜力的结果，但其对大型训练数据集的依赖、有限的泛化性以及成组内上下文缺乏真实推理阶段，限制了其实际应用。相比之下，基于模型驱动的方法提供了无需训练、理论基础扎实且数据高效的替代方案，尽管在应用于大型 3D 数据集时它们通常面临可扩展性和优化的挑战。在这项工作中，我们介绍了 DARC（通过坐标下降的微分同胚图谱配准），这是一种用于图谱构建的新型模型驱动组间配准框架。DARC 支持广泛范围的图像不相似性度量，并且可以高效处理任意数量的 3D 图像，而不会产生 GPU 内存问题。通过坐标下降策略和中心化约束激活函数，DARC 生成了偏差小、解剖保真度高的微分同胚图谱。除了图谱构建，我们展示了两个关键应用：(1) 一次性分割，其中仅在图谱上标注的标签通过逆变形传播到受试者身上，超越了最先进的少样本方法；以及 (2) 形状合成，其中通过使用合成的微分同胚变形场对图谱网格进行扭曲生成新的解剖变体。总体而言，DARC 为图谱构建和应用提供了灵活、可推广且资源高效的框架。

1 介绍

图谱是一种代表性的医学图像，概括了从个体主体群体中提取的平均解剖结构。在医学图像分析中，图谱作为标准化的空间参考，促进了一系列后续任务，如自动分割、区域定量分析、病变检测和群体水平分析。解剖图谱的构建通常依赖于群体间图像配准，这一过程本质上与成对配准技术相关联。

在成对图像配准的背景下，基于深度学习的数据驱动方法在配准精度上已与依赖迭代优化的传统模型驱动方法媲美 (Thorley et al. 2021; Jia et al. 2021; Duan et al. 2023; Jia et al. 2025)。数据驱动方法的一个关键优势在于其在推断时的计算效率；一旦训练完成，这些模型能比迭代的模型驱动方法显著更快地执行配准任务 (Balakrishnan et al. 2019; Jia et al. 2021; Zhang et al. 2024)。然而，这种效率是有代价的，并伴随一些限制。数据驱动的配准框架通常需要大量的训练数据才能实现稳健的性能 (Cheng et al. 2024; Jia et al. 2022)。通过神经网络回归估算形变或位移场本质上易受近似误差的影响，这可能会降低配准精度，特别是在具有复杂解剖变化的区域 (Cheng et al. 2025; Jia et al. 2023a)。此外，数据驱动方法的泛化能力往往有限，尤其是应

用于与训练时不同的数据分布时。相比之下，模型驱动方法无需训练，因此具有很高的数据效率。它们通过迭代算法直接优化形变或位移场，避免了由神经网络引入的中间近似误差。此外，模型驱动方法在很大程度上不受数据分布变化的影响，使其本质上更具通用性并易于应用于新的数据集或临床环境中。

在通过群体配准进行图谱构建的背景下，数据驱动的方法在计算效率方面并未显示出对传统模型驱动方法的显著改进。虽然训练好的神经网络在成对配准任务中可以实现近乎瞬时的推理，这一优势并不能直接应用于群体配准环境。最近的数据驱动群体配准框架通常采用所谓的基于神经网络的优化 (Dalca et al. 2019a; He and Chung 2025)，其中网络通过在训练过程中优化其参数来学习构建模板图像 (Sinclair et al. 2022; Dannecker et al. 2024; Dey et al. 2021; Ding and Niethammer 2022)。换句话说，这些方法缺乏推理阶段。针对不同的图像组，甚至是相同组但不同数量的图像时，学习到的网络参数必须重新训练或进行微调来构建新模板。这种对微调或重新训练的依赖削弱了数据驱动方法的核心优势，即快速推理。因此，数据驱动的群体配准的实用性仍然有限，尤其是在需要效率和跨异质数据集的泛化能力的应用中。尽管最近的一些研究 (例如 (Abulnaga et al. 2025)) 提出了数据驱动的方法，可以在无需微调或重新训练的情况下通过一次前向传递生成图谱，但这些方法是启发式的，通常缺乏理论基础，可能无法很好地泛化到不同的解剖结构或成像协议。

因此，模型驱动的图谱构建仍然是一个活跃的研究领域。这些方法无需训练，理论上可解释，并且通常在不同的图像组中具有稳健性。然而，优化可能具有挑战性，特别是在结合复杂约束（如中心性和微分同胚）时。此外，从大规模数据集中计算图谱仍然在计算上要求高，因为在同时处理多个高分辨率三维体积图像时，GPU 内存占用可能呈指数增长。为了解决这些挑战，我们提出了一种用于图谱构建的模型驱动的组内配准方法，并引入了一种新颖的坐标下降算法进行优化。我们将这种方法称为 DARC（通过坐标下降的微分同胚图谱配准），其具有三个主要创新点：

- DARC 包含一个通用的不相似性数据项，支持 MSE、 L_1 、NCC、SSIM，或者理论上任何可以测量图谱图像与每个变形输入图像之间距离的度量。对于 MSE 和 L_1 ，我们展示了用于更新图谱图像的闭式解决方案是可用的。对于 NCC 和 SSIM，我们建议使用随机梯度下降来进行更新。这个设计使得 DARC 能够在不触及 GPU 内存限制的情况下高效处理任意数

† Corresponding authors.

量的 3D 图像。此外，我们引入了一个激活函数，以有效地执行零位移约束（中心性），确保生成的图谱图像相对于任何单个输入图像都是无偏的。DARC 使用 PyTorch 实现，实现了 GPU 加速和自动微分，使其既简单又高效。DARC 与最先进的模型驱动和数据驱动方法 (Jia et al. 2010; Avants et al. 2010; Dalca et al. 2019a; Ding and Niethammer 2022; He and Chung 2025) 进行了比较，展示了其卓越的性能。此外，DARC 的结果支持下面描述的两个新应用。

- 我们利用从 DARC 构建的图谱进行一次性分割。在这种方法中，仅需要对图谱图像进行标注，然后通过逆形变场传播到每个对象。由于 DARC 实现了精确的图谱构建，我们的一次性分割方法优于现有的分割方法 (Cao et al. 2022; Islam et al. 2019; Ronneberger, Fischer, and Brox 2015)。
- 基于图谱网格和 DARC 产生的位移场，我们提出了一种新颖的方法，通过合成形变场并将其应用于图谱网格，隐式地生成一个新的虚拟解剖学人口。该方法确保生成的形状自然地保持逐点对应，从而促进群体层面的形状分析。此外，基础的变形仍然是微分同胚的，从而避免了网格折叠并保持解剖学的合理性。得益于 DARC 实现的对齐，我们证明了使用线性 PCA 模型的形状合成表现可以与先进的扩散模型相媲美。

2 相关工作

基于迭代优化的模型驱动方法长期以来一直是图谱构建的经典方法 (Cootes et al. 2004)。早期的群体配准方法关注于变形的合理性和拓扑保留，采用的技术包括用于刚性和仿射对齐的迭代最近点算法 (Besl and McKay 1992)、多尺度 B 样条方法 (Rueckert, Frangi, and Schnabel 2001) 以及大变形微分同胚度量映射 (LDDMM) 算法 (Glaunes et al. 2008)。在 LDDMM 的基础上，提出了各种扩展，包括用于大变形的无偏图谱构建 (Avants and Gee 2004; Lorenzen, Davis, and Joshi 2005)、黎曼几何中位数图谱 (Fletcher, Venkatasubramanian, and Joshi 2009)、多模态贝叶斯婴儿大脑图谱 (Lorenzen et al. 2006)、基于 EM 的图谱估计 (Bhatia et al. 2007) 以及 PDE 约束的 LDDMM (Hernandez 2019)。(Avants et al. 2010) 随后引入了对称规范化 (ANTs)，这是第一个对称、无偏的微分同胚图谱构建方法。另一方向是分层群体配准 (Wang et al. 2010; Jia et al. 2010; Wu et al. 2011; Ying et al. 2014; Jia et al. 2011)，它利用局部流形结构来提高可扩展性、鲁棒性和效率。然而，模型驱动方法难以优化，特别是在具有复杂约束或不相似性度量（如 NCC）的情况下，并且在没有封闭形式损失函数的情况下难以扩展到大型数据集。

基于神经网络的数据驱动方法为图谱构建提供了一种现代替代方案。AtlasMorph (Dalca et al. 2019a) 首次使用卷积神经网络进行无条件条件和条件图谱生成。在此基础上，(Yu, Dalca, and Sabuncu 2020) 建模了属性依赖的形态变化为特定条件脑图谱的微分同胚变换，而 (Dannecker et al. 2024) 使用隐式神经表示来学习连续的时空嵌入，从而在任意时间点和分辨率下生成胎儿脑图谱。(Ding and Niethammer 2022) 提出了一种端到端框架，消除了仿射预配准的需求，并引入了一种新的成对相似性损失来提高配准精度。生成模型也被应用于

图谱构建。(Dey et al. 2021) 将可变形配准和条件图谱估计表述为对抗游戏，开发了一种基于 GAN 的属性指导图谱生成方法。(Starck et al. 2024) 使用潜在扩散模型从通用图谱合成特定条件的形变场。(He and Chung 2025) 引入了 InstantGroup，一个基于 VAE 的框架，在潜在空间中进行插值和算术运算，以任意尺度构建图谱。然而，数据驱动方法通常需要针对不同的数据集或未见过的主体进行微调或重新训练。因此，它们在推理时间上相对于模型驱动的迭代方法的优势尚不明显。

给定一组图像 $\mathcal{I} = \{I_i \in \mathbb{R}^{M \times N \times H}\}_{i=1}^n$ ，其中 $M \times N \times H$ 是图像大小， n 是组中图像的数量，我们的目标是构建一个图谱 $A \in \mathbb{R}^{M \times N \times H}$ ，使得所有样本通过其对应的形变场 $\Phi = \{\phi_i \in (\mathbb{R}^{M \times N \times H})^3\}_{i=1}^n$ 被映射到一个共同的参考空间中。这个过程可以被表述为如下的约束优化问题：

$$\min_{\phi_1, \dots, \phi_n, A} f(\phi_1, \dots, \phi_n, A) \text{ s.t. } \sum_{i=1}^n \phi_i = 0, \quad (1)$$

其中

$$f(\phi_1, \dots, \phi_n, A) = \sum_{i=1}^n \mathcal{L}(I_i \circ \phi_i, A) + \lambda \|\nabla \phi_i\|_2^2. \quad (2)$$

在 (2) 中， $\mathcal{L}$ 表示不相似性数据项，第二项是扩散正则项 (Fischer and Modersitzki 2003)， $\circ$ 代表空间变换（扭曲）， λ 是正则化超参数。此外，(1) 中的中心性约束 $\sum \phi_i = 0$ 确保生成的图谱图像 A 对组中任何单个图像无偏倚。为了为整个数据集 $\mathcal{I}$ 建立模板或图谱，在使用梯度下降时，需要在每次迭代中共同优化 A 和所有形变场 Φ 。然而，随着三维图像数量的增加，GPU 和 CPU 内存使用可能呈指数增长，最终超出可用内存。

坐标下降法：为了有效缓解内存限制，我们引入了一种基于坐标下降法的异步图集更新策略，如下所示：坐标下降算法详细描述在 (??) 中，通过在一次优化一个变量（或坐标）时保持其他变量不变，从而最小化多变量损失函数 (1)，并通过迭代依次循环更新每个变量以降低损失函数值直到收敛。同样值得注意的是，无论使用的数据项为何，形变场 $\Phi = \{\phi_i\}_{i=1}^n$ 的优化可以并行进行，因为它们不是耦合变量。

微分同胚：为了确保形变场 Φ 是微分同胚的，我们使用静止速度场 $\mathcal{V} = \{v_i\}_{i=1}^n$ (Ashburner 2007) 对 Φ 进行参数化。微分同胚变形是可微且可逆的，从而保留拓扑结构，这对于我们在第四节详细介绍的一次性分割和形状合成来说是至关重要的。在群论中，静止速度场属于李代数，我们可以通过 $\phi_i = \text{Exp}(v_i)$ 对其进行指数运算以获得微分同胚变形。在本文中，我们进行了七次 (Ashburner 2007; Dalca et al. 2019b) 的缩放和平方以实现微分同胚。请注意，实现微分同胚需要在 (??) 中优化 v_i ，而不是直接优化 ϕ_i 。

中心激活函数：(??) 中倒数第二个方程作为激活函数严格执行中心约束（零位移），确保图集 A 是无偏的。这种激活函数是可微的，因此可以在 PyTorch 中使用其自动微分功能。为此，许多方法 (Dalca et al. 2019a; Sinclair et al. 2022; Dey et al. 2021) 在优化过程中将“平均速度场”的范数作为正则化项纳入，并与其他能量项一起最小化。然而，这引入了一个需要仔细调整的额外超参数，并且该约束可能无法严格满足。

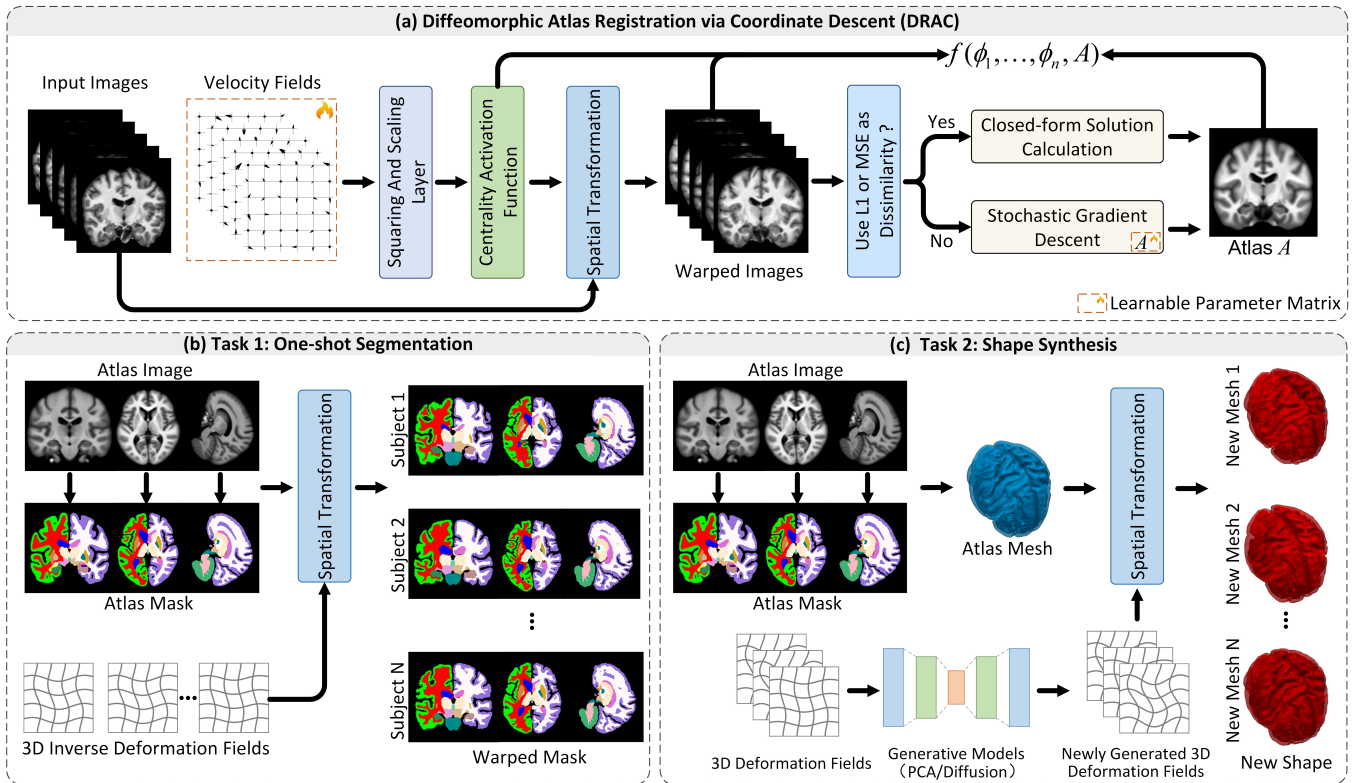


Figure 1: 关于所提出的 DARC 算法及其两个关键应用的概述: (a) DARC 是一个模型驱动的框架, 可以联合优化速度场和图集图像; (b) DARC 在单次分割中的应用; (c) DARC 在新形状合成中的应用。

不相似度度量: 如果在 (2) 中选择不相似数据项不同, 则应单独考虑 (??) 中的最后优化问题。为此, 我们考虑四种不相似度度量: 均方误差 (MSE)、 L_1 、归一化互相关 (NCC) 和结构相似性 (SSIM)。然而, 原则上任何可微分度量用于测量图谱图像和每个变形输入图像之间的距离都兼容于所提出的算法。MSE 和 L_1 都是点状度量, 而 NCC 和 SSIM 是块状度量。对于 MSE, (??) 中的最后一个优化问题为

$$A^{k+1} = \underset{A}{\operatorname{argmin}} \sum_{i=1}^n \|I_i \circ \hat{\phi}_i^{k+1} - A\|_2^2 = \frac{1}{n} \sum_{i=1}^n I_i \circ \hat{\phi}_i^{k+1}. \quad (3)$$

在这种情况下, 图谱图像在迭代 $k+1$ 时的封闭形式解由所有变形图像的平均值给出。对于 L_1 数据项, 图谱图像也有一个封闭形式解, 由下式给出

$$A^{k+1} = \operatorname{median}(I_1 \circ \hat{\phi}_1^{k+1}, I_2 \circ \hat{\phi}_2^{k+1}, \dots, I_n \circ \hat{\phi}_n^{k+1}), \quad (4)$$

其中 median 表示在迭代 $k+1$ 时所有变形图像的通道中位值。注意, 使用闭形式解时, 计算非常廉价, 因此对于 MSE 和 L_1 数据项都不存在内存不足问题。

由于 NCC 和 SSIM 缺乏封闭形式的解决方案, 最小化关于 (??) 中的图集图像的最后优化问题变得不平凡。这是因为优化要求所有扭曲的 3D 图像加载到内存中, 这带来了显著的内存挑战。为了解决这一问题, 我们建议使用小批量随机梯度下降来更新图集。一旦变形场 Φ 被更新, 且扭曲图像 $I \circ \Phi$ 被计算, 我们就从这些扭曲图像中采样一个子集 (一个小批量), 以计算相对于 A 的梯度, 并在 A 上进行更新。这个设计使得

算法能够高效地处理任意数量的 3D 图像, 而不会遇到 GPU 内存限制。完整的算法在算法 1 中给出。

3 基于图谱建模的应用

图谱构建过程生成了一个图谱图像, 以及 n 扭曲图像, n 前向变形场和 n 后向变形场。这些结果进一步用于两个新颖的下游应用, 即单次分割和形状合成, 分别如图 1 (b) 和 (c) 所示。

一次性分割: 每个主体图像的分割可以通过一个模板图像的单个分割掩码完成, 并通过相应的反向变形场传播到所有单独的主体。用于此一次性分割的方程如下所示:

$$I_i^{(S)} = A^{*(S)} \circ \operatorname{Exp}(-v_i^*), \quad (5)$$

其中, $I_i^{(S)}$ 表示主体图像 I_i 的结果分割掩码, 而 $A^{*(S)}$ 表示优化后的模板图像 A^* 的分割掩码。 $\operatorname{Exp}(-v_i^*)$, 即 $\operatorname{Exp}(v_i^*)$ 的逆, 表示将模板图像映射到第 i 个主体图像的变形场, 其中 v_i^* 是使用 DARC 优化的 i th 输出速度场。请注意, 该过程只需要模板图像的注释。

形状合成: 利用图谱图像的分割掩膜, 可以使用网格细分算法提取由顶点和三角面组成的表面网格。我们建议将合成变形场应用于图谱表面网格, 以生成具备点对点对应关系的虚拟解剖结构群体 (脑和心脏)。为此, 我们考虑两种类型的生成模型: 一种是基于主成分分析 (PCA) 的简单模型, 另一种是基于扩散过程的高级模

型。首先，对于 PCA，生成变形使用的公式为：

$$\phi^{New} = Exp \left(\tilde{v} + \sum_{j=1}^p \alpha_j U_{\{:,j\}} \right), \alpha_j \sim \mathcal{N}(0, \lambda_j).$$

这里， $\tilde{v}$ 表示通过平均所有生成的速度场 $\mathcal{V}^* = \{v_i^*\}_{i=1}^n$ 来计算的平均速度场， p 是用于生成的主成分 (PCs) 数量。 U 的列对应于从 $\mathcal{V}^*$ 派生的协方差矩阵的排序后的 PCs， λ_j 表示协方差矩阵的第 j 个特征值。

接下来，用于使用扩散模型生成变形场的逆方程如下所示：

$$\begin{cases} P_\theta(x^{0:T}) = \mathcal{N}(x^T; 0, I) \prod_{t=1}^T \mathcal{N}(x^{t-1}; \mu_\theta(x^t, t), \beta^t I) \\ \phi^{New} = Exp(x^0) \end{cases},$$

其中， $x \in \{v_1^*, \dots, v_n^*\}$ 表示多通道速度图像， $t \in [1, \dots, T]$ 表示扩散时间步长， $P_\theta(x^{0:T})$ 是在网络参数 θ 下轨迹 $\{x^0, x^1, \dots, x^T\}$ 的联合概率。术语 $\mathcal{N}(x^T; 0, I)$ 表示 x^T 上的高斯先验，而 $\mathcal{N}(x^{t-1}; \mu_\theta(x^t, t), \beta^t I)$ 表示用于从当前噪声输入 x^t 采样略微去噪版本 x^{t-1} 的高斯分布，其均值由神经网络预测，且具有固定方差 $\beta^t I$ 。

最后，形状合成通过使用从 PCA 或扩散模型获得的生成（前向）变形场 ϕ^{New} 对图集表面网格进行扭曲来完成。该方法有三大新颖之处。首先，生成的变形场是微分同胚的，这防止了网格折叠，从而确保了解剖上的合理性。其次，由于来自图集的一致扭曲，所有生成的形状本质上保持点对点的对应关系。第三，整个生成过程在图像空间中进行，不涉及表面网格，因此可以直接应用为图像设计的生成模型。

4 实验

数据集和实现细节：我们使用 DARC 方法在来自不同器官的两个图像数据集上评估了图谱构建：脑数据集 OASIS (Marcus et al. 2007) 和心脏数据集 (Duan et al. 2019)。OASIS 包括 416 份三维 T1 加权脑扫描，其尺寸为 $192 \times 160 \times 224$ 。心脏数据集包含 220 份三维心脏 MRI 体积，均经过裁剪或填充至 $96 \times 128 \times 128$ 。我们使用 Adam 优化器以固定学习率 $1e-2$ 训练 DARC，以更新两个数据集中的速度场和图谱图像。对于心脏数据集的速度场估计和随机梯度下降 (SGD)，使用了批量大小为 10；对于 OASIS 数据集，使用了批量大小为 4。对于 L_1 和 MSE，正则化参数为 $\lambda = 0.5$ ，而对于 NCC 和 SSIM 则为 $\lambda = 8$ 。外部迭代次数为 $K_1 = 10$ ，更新速度场的内部迭代为 $K_2 = 300$ 。在使用 NCC 或 SSIM 作为相似性度量时，使用 SGD 的图谱更新循环为 $K_3 = 20$ 。所有模型均在具有 48GB RAM 的单个 NVIDIA RTX A6000 GPU 上训练。

评估指标：我们使用三个指标评估了使用 DARC 进行图谱构建的准确性：Dice 准确性、变形中心性和负 Jacobian 行列式值的百分比 ($|J|_{<0}\%$)，这反映了折叠率。对于一次分割，我们报告了 Dice 得分。对于形状合成，我们采用了一些指标，例如特异性误差（生成网格点到用于训练的原始表面的平均距离）(Davies et al. 2009)。覆盖率（通过最近邻匹配由生成网格覆盖的真实网格的比例）、最小匹配距离（在最佳点匹配下的最

Table 1: 针对两个数据集，对不同相异度度量的图谱构建结果的定量比较。

Dataset	Metric	Dice (%, $\uparrow$)	$ J _{<0}\%$ ($\downarrow$)	Centrality ($\times 10^{-3}$ $\downarrow$)	
				Before	After
Heart	MSE	86.28 $\pm$ 4.37	0.26 $\pm$ 0.20	1.2 $\pm$ 0.93	0.0 $\pm$ 0.0
	L_1	74.34 $\pm$ 5.09	0.66 $\pm$ 0.28	6.0 $\pm$ 6.29	0.0 $\pm$ 0.0
	NCC	74.21 $\pm$ 17.63	0.82 $\pm$ 2.50	0.9 $\pm$ 1.02	0.0 $\pm$ 0.0
	SSIM	80.71 $\pm$ 5.23	0.0 $\pm$ 0.0	0.4 $\pm$ 0.46	0.0 $\pm$ 0.0
OASIS	MSE	79.54 $\pm$ 4.83	0.06 $\pm$ 0.04	0.5 $\pm$ 0.28	0.0 $\pm$ 0.0
	L_1	66.08 $\pm$ 5.61	0.02 $\pm$ 0.05	0.2 $\pm$ 0.13	0.0 $\pm$ 0.0
	NCC	82.44 $\pm$ 2.79	0.01 $\pm$ 0.01	0.3 $\pm$ 0.19	0.0 $\pm$ 0.0
	SSIM	72.10 $\pm$ 4.53	0.0 $\pm$ 0.0	0.2 $\pm$ 0.01	0.0 $\pm$ 0.0

小表面误差)、形状 JSD (生成形状和真实形状分布之间的 Jensen-Shannon 发散) 以及 1-NNA (单一最近邻准确性，其中接近 50 的值% 表示不可区分性和良好的分布覆盖) (Yang et al. 2019)。总体而言，这些指标反映了所提出的 DARC 方法的准确性和有效性。

不同不相似性度量的影响：首先，我们比较了不同不相似性度量之间的性能：MSE、 L_1 、NCC 和 SSIM。定量和定性结果分别在表格 1 和图 3 中给出。从表 1 中可以看出，所有这四种指标都产生了无折叠的模板构建，但在 Dice 精度、中心误差和定性性能上表现出不同的行为。MSE 由于其二次惩罚，过度平滑了变形场，但在心脏上产生了最高的 Dice 分数。 L_1 由于对中心位数的敏感性，减少了极端偏差，但在 Dice 和集中化上表现最差。NCC 在对比丰富的 OASIS 脑部 MRI 上达到了最佳的 Dice 和适中的中心误差。SSIM 优先考虑结构相似性，保留细致的心肌和皮质纹理，并产生最具视觉细节的结果，尽管其全局 Dice 略低。通过中心激活函数，所有方法均满足零位移约束，确保模板无偏见。图 3 显示，像 MSE 和 L_1 这样的逐像素指标产生过于平滑的结果（其中 L_1 最平滑），而像 NCC 和 SSIM 这样的逐贴片指标更好地保留了解剖细节，SSIM 产生最具视觉细节的结构。值得注意的是，观察到 MSE 在心脏数据集上表现更好，而 NCC 在脑部数据集上表现突出，与之前的研究 (Jia et al. 2023b) 一致。

Table 2: 来自两个数据集的最先进方法的图谱构建结果的定量比较。

Datasets	Method	Dice (%, $\uparrow$)	$ J _{<0}\%$ ($\downarrow$)	Centrality ($\times 10^{-3}$ $\downarrow$)
Heart	AtlasMorph	81.35 $\pm$ 2.32	0.13 $\pm$ 0.15	0.5 $\pm$ 0.81
	Aladdin	80.41 $\pm$ 6.13	0.0 $\pm$ 0.0	957.0 $\pm$ 407.84
	Ours	86.28 $\pm$ 4.37	0.26 $\pm$ 0.20	0.0 $\pm$ 0.0
OASIS	AtlasMorph	80.27 $\pm$ 11.03	1.14 $\pm$ 1.29	0.25 $\pm$ 0.04
	Aladdin	77.39 $\pm$ 6.46	0.0 $\pm$ 0.0	907.5 $\pm$ 423.78
	InstantGrp	74.16 $\pm$ 0.38	—	168 $\pm$ 3.0
	ANTs	73.55 $\pm$ 0.44	—	239 $\pm$ 10.0
	ABSORB	73.50 $\pm$ 0.55	—	743 $\pm$ 57.0
	Ours	82.44 $\pm$ 2.79	0.01 $\pm$ 0.01	0.0 $\pm$ 0.0

与最新技术的对比：我们将 DARC 与四种最近的数据驱动图谱构建方法进行了比较：Multimorph (Abulnaga et al. 2025)、AtlasMorph (Dalca et al. 2019a)、Aladdin (Ding and Niethammer 2022) 和 InstantGrp (He and Chung 2025)。我们还包括与两种模型驱动方法的比较，即 ANTs (Avants et al. 2010) 和 ABSORB (Jia et al. 2010)。从它们的 GitHub 仓库获取，我们在本文使用的两个数据集上重新训练了 AtlasMorph¹ 和 Aladdin

¹<https://github.com/voxelmorph/voxelmorph>

² 的官方实现。对于 AtlasMorph, 我们使用了 0.1 的正则化权重, 而对于 Aladdin, 我们将正则化权重设置为 10,000, 并将相似性损失权重固定为 10。两种方法均使用批量大小为 4。由于 Multimorph 没有发布完整的训练和评估代码, 我们只能使用其预训练权重运行图谱构建的推理代码³。InstantGrp 没有提供公开代码, 因此我们报告了其论文中的大脑结果, 因为我们使用了相同的数据集进行比较。请注意, 与其他基线方法不同的是, Multimorph 在训练过程中结合了掩膜信息作为一种弱监督形式。

表 2 显示, 在两个数据集上的 Dice 值中, 我们的 DRAC 方法都超过了所有竞品方法, 实现了接近 0 的折叠和严格为 0 的中心性误差, 这证明了其在精确对齐和拓扑保真度之间的最佳权衡。通过施加极强的正则化, Aladdin 抑制了变形场的振荡以减少网格折叠; 然而, 其中心性误差在数百的数量级上, 导致了显著的整体偏离零中心的位移, 造成明显的解剖偏差。图 4 直观地展示了每种方法生成的图谱。在心脏数据集上, AtlasMorph 生成的图谱过于平滑, 而 Aladdin 的大中心性误差导致过度变形。相较之下, 我们的 DARC 图谱表现出色。

我们进一步在每个数据集的随机抽样子集上评估了图谱构建, 每个模型使用 [5, 10, 20, ..., 50] 张图像进行训练。如图 5 所示, DARC 始终优于其他两种数据驱动的方法, 尤其是在训练图像数量有限的情况下。由于 DARC 是模型驱动的, 因此无论数据集大小如何, 其性能都保持稳定。附录中的图 1 进一步证实了这一趋势。总体而言, DARC 在具有不同解剖结构的数据集中优于现有方法, 同时提供了更可靠和更具普遍性的解决方案, 特别是在数据有限的情况下。

一次性分割: 我们比较了我们的单次分割方法 (5) 的 Dice 得分与几次分割基线的得分, 包括 U-Net (Ronneberger, Fischer, and Brox 2015)、3D 注意力 U-Net (Islam et al. 2019) 和 Swin UNETR (Cao et al. 2022)。结果在表 ?? 中报告。我们的单次分割在单次和五次设置中始终优于这些方法, 同时保持稳健的分割准确率。请注意, 分割方法在较简单的解剖结构上通常表现更好, 例如心脏, 相较于更复杂的解剖结构比如大脑。附录图 2 中的视觉示例进一步说明了我们方法在有限标注设置下的优势。

形状合成: 表 ?? 显示, 扩散模型和 PCA 在特异性误差、最小匹配距离、形状 JSD 以及多样性/分布对齐指标上的合成性能相似, 其边际直方图和散点图在图 6 中进一步支持了这一点。虽然 PCA 在其线性子空间假设下有效捕捉到主要形状变化, 但在极端或非高斯区域则表现不佳, 限制了它在表示稀有形态及微妙非线性相关性上的能力。相比之下, 扩散模型更好地捕捉尾部分布和结构一致性, 同时不影响生成准确性, 使得形状合成更为鲁棒。

在图 7 中, 我们比较了两种方法生成的网格样本, 结果显示我们的形状合成方法产生了多样且拓扑稳定的表面。由于所有形状都进行了点对点对应, 因此我们对特定解剖结构进行了 PCA 分析, 以探索群体级别的变化。附录图 3 说明, 对于左心室, 第一个主成分反映了大小, 第二个主成分描述了拉长和倾斜, 第三个主成分代表了拉长和圆度。对于大脑, 第一个模式对应整体膨

胀或收缩, 整个大脑大小均匀变化; 第二个模式揭示了区域变化, 强调了不同皮质区域的对立形态变化; 第三个模式捕捉了不对称性和扭曲, 表明局部的剪切变化。

我们提出了 DARC, 一种用于图谱构建的模型驱动的组注册方法。DARC 包含了一个通用的相异性数据项, 支持任何度量标准来评估图谱图像与每个变形输入图像之间的距离。我们还引入了一个激活函数, 以确保生成的图谱图像保持无偏。此外, 我们提出了一个基于 DARC 构建图谱的一次性分割框架, 减少了对手动注释图像的依赖。最后, 利用 DARC 生成的图谱网格和变形场, 我们开发了一种新颖的形状合成方法, 该方法通过合成变形场并将其应用于图谱网格来隐式生成新的虚拟解剖学人群。在心脏和 OASIS 数据集上的实验结果表明, DARC 是多功能的、准确的且无偏。此外, 我们验证了形状合成方法的有效性, 并证明了所提出的一次性分割在分割准确性上优于现有方法。

References

- Abulnaga, S. M.; Hoopes, A.; Dey, N.; Hoffmann, M.; Fischl, B.; Guttag, J.; and Dalca, A. 2025. MultiMorph: On-demand Atlas Construction. In Proceedings of the Computer Vision and Pattern Recognition Conference, 30906–30917.
- Ashburner, J. 2007. A fast diffeomorphic image registration algorithm. *NeuroImage*, 38(1): 95–113.
- Avants, B.; and Gee, J. C. 2004. Geodesic estimation for large deformation anatomical shape averaging and interpolation. *NeuroImage*, 23: S139–S150.
- Avants, B. B.; Yushkevich, P.; Pluta, J.; Minkoff, D.; Korczykowski, M.; Detre, J.; and Gee, J. C. 2010. The optimal template effect in hippocampus studies of diseased populations. *NeuroImage*, 49(3): 2457–2466.
- Bai, W.; Shi, W.; de Marvao, A.; Dawes, T. J.; O’Regan, D. P.; Cook, S. A.; and Rueckert, D. 2015. A bi-ventricular cardiac atlas built from 1000+ high resolution MR images of healthy subjects and an analysis of shape and motion. *Medical Image Analysis*, 26(1): 133–145.
- Balakrishnan, G.; Zhao, A.; Sabuncu, M. R.; Guttag, J.; and Dalca, A. V. 2019. Voxelmorph: a learning framework for deformable medical image registration. *IEEE Transactions on Medical Imaging*, 38(8): 1788–1800.
- Besl, P. J.; and McKay, N. D. 1992. Method for registration of 3-D shapes. In *Sensor Fusion IV: Control Paradigms and Data Structures*, volume 1611, 586–606. Spie.
- Bhatia, K. K.; Aljabar, P.; Boardman, J. P.; Srinivasan, L.; Murgasova, M.; Counsell, S. J.; Rutherford, M. A.; Hajnal, J. V.; Edwards, A. D.; and Rueckert, D. 2007. Groupwise combined segmentation and registration for atlas construction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 532–540. Springer.
- Cao, H.; Wang, Y.; Chen, J.; Jiang, D.; Zhang, X.; Tian, Q.; and Wang, M. 2022. Swin-unet: Unet-like pure transformer for medical image segmentation. In

²<https://github.com/uncbiag/Aladdin>

³<https://github.com/mabulnaga/multimorph>

- European Conference on Computer Vision, 205–218. Springer.
- Cheng, X.; Jia, X.; Lu, W.; Li, Q.; Shen, L.; Krull, A.; and Duan, J. 2024. Winet: Wavelet-based incremental learning for efficient medical image registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 761–771. Springer.
- Cheng, X.; Zhang, T.; Lu, W.; Meng, Q.; Frangi, A. F.; and Duan, J. 2025. SACB-Net: Spatial-awareness Convolutions for Medical Image Registration. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 5227–5237.
- Cootes, T. F.; Marsland, S.; Twining, C. J.; Smith, K.; and Taylor, C. J. 2004. Groupwise diffeomorphic non-rigid registration for automatic model building. In *European Conference on Computer Vision*, 316–327. Springer.
- Dalca, A.; Rakic, M.; Guttag, J.; and Sabuncu, M. 2019a. Learning conditional deformable templates with convolutional networks. *Advances in Neural Information Processing Systems*, 32.
- Dalca, A. V.; Balakrishnan, G.; Guttag, J.; and Sabuncu, M. R. 2019b. Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. *Medical Image Analysis*, 57: 226–236.
- Dannecker, M.; Kyriakopoulou, V.; Cordero-Grande, L.; Price, A. N.; Hajnal, J. V.; and Rueckert, D. 2024. CINA: conditional implicit neural atlas for spatio-temporal representation of fetal brains. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 181–191. Springer.
- Davies, R. H.; Twining, C. J.; Cootes, T. F.; and Taylor, C. J. 2009. Building 3-D statistical shape models by direct optimization. *IEEE Transactions on Medical Imaging*, 29(4): 961–981.
- Dey, N.; Ren, M.; Dalca, A. V.; and Gerig, G. 2021. Generative adversarial registration for improved conditional deformable templates. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3929–3941.
- Ding, Z.; and Niethammer, M. 2022. Aladdin: Joint atlas building and diffeomorphic registration learning with pairwise alignment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 20784–20793.
- Duan, J.; Bello, G.; Schlemper, J.; Bai, W.; Dawes, T. J.; Biffi, C.; de Marvao, A.; Doumoud, G.; O’Regan, D. P.; and Rueckert, D. 2019. Automatic 3D bi-ventricular segmentation of cardiac images by a shape-refined multi-task deep learning approach. *IEEE Transactions on Medical Imaging*, 38(9): 2151–2164.
- Duan, J.; Jia, X.; Bartlett, J.; Lu, W.; and Qiu, Z. 2023. Arbitrary order total variation for deformable image registration. *Pattern Recognition*, 137: 109318.
- Fischer, B.; and Modersitzki, J. 2003. Curvature based image registration. *Journal of Mathematical Imaging and Vision*, 18(1): 81–85.
- Fletcher, P. T.; Venkatasubramanian, S.; and Joshi, S. 2009. The geometric median on Riemannian manifolds with application to robust atlas estimation. *NeuroImage*, 45(1): S143–S152.
- Glaunes, J.; Qiu, A.; Miller, M. I.; and Younes, L. 2008. Large deformation diffeomorphic metric curve mapping. *International Journal of Computer Vision*, 80(3): 317–336.
- He, Z.; and Chung, A. C. 2025. InstantGroup: Instant Template Generation for Scalable Group of Brain MRI Registration. *IEEE Transactions on Image Processing*.
- Hernandez, M. 2019. PDE-constrained LDDMM via geodesic shooting and inexact Gauss–Newton–Krylov optimization using the incremental adjoint Jacobi equations. *Physics in Medicine & Biology*, 64(2): 025002.
- Hoogendoorn, C.; Duchateau, N.; Sanchez-Quintana, D.; Whitmarsh, T.; Sukno, F. M.; De Craene, M.; Lekadir, K.; and Frangi, A. F. 2012. A high-resolution atlas and statistical model of the human heart from multislice CT. *IEEE Transactions on Medical Imaging*, 32(1): 28–44.
- Islam, M.; Vibashan, V.; Jose, V. J. M.; Wijethilake, N.; Utkarsh, U.; and Ren, H. 2019. Brain tumor segmentation and survival prediction using 3D attention UNet. In *International MICCAI Brainlesion Workshop*, 262–272. Springer.
- Jia, H.; Wu, G.; Wang, Q.; and Shen, D. 2010. ABSORB: Atlas building by self-organized registration and bundling. *NeuroImage*, 51(3): 1057–1070.
- Jia, H.; Yap, P.-T.; Wu, G.; Wang, Q.; and Shen, D. 2011. Intermediate templates guided groupwise registration of diffusion tensor images. *NeuroImage*, 54(2): 928–939.
- Jia, X.; Bartlett, J.; Chen, W.; Song, S.; Zhang, T.; Cheng, X.; Lu, W.; Qiu, Z.; and Duan, J. 2023a. Fourier-net: Fast image registration with band-limited deformation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 1015–1023.
- Jia, X.; Bartlett, J.; Zhang, T.; Lu, W.; Qiu, Z.; and Duan, J. 2022. U-net vs transformer: Is u-net outdated in medical image registration? In *International Workshop on Machine Learning in Medical Imaging*, 151–160. Springer.
- Jia, X.; Lu, W.; Cheng, X.; and Duan, J. 2025. Decoder-only image registration. *IEEE Transactions on Medical Imaging*.
- Jia, X.; Thorley, A.; Chen, W.; Qiu, H.; Shen, L.; Styles, I. B.; Chang, H. J.; Leonardis, A.; De Marvao, A.; O’Regan, D. P.; et al. 2021. Learning a model-driven variational network for deformable image registration. *IEEE Transactions on Medical Imaging*, 41(1): 199–212.

- Jia, X.; Thorley, A.; Gomez, A.; Lu, W.; Kotecha, D.; and Duan, J. 2023b. Fourier-net+: Leveraging band-limited representation for efficient 3d medical image registration. *arXiv preprint arXiv:2307.02997*.
- Lorenzen, P.; Davis, B. C.; and Joshi, S. 2005. Unbiased atlas formation via large deformations metric mapping. In *International conference on Medical Image Computing and Computer-Assisted Intervention*, 411–418. Springer.
- Lorenzen, P.; Prastawa, M.; Davis, B.; Gerig, G.; Bullitt, E.; and Joshi, S. 2006. Multi-modal image set registration and atlas formation. *Medical Image Analysis*, 10(3): 440–451.
- Marcus, D. S.; Wang, T. H.; Parker, J.; Csernansky, J. G.; Morris, J. C.; and Buckner, R. L. 2007. Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *Journal of Cognitive Neuroscience*, 19(9): 1498–1507.
- McWhinney, S. R.; Hlinka, J.; Bakstein, E.; Dietze, L. M.; Corkum, E. L.; Abé, C.; Alda, M.; Alexander, N.; Benedetti, F.; Berk, M.; et al. 2024. Principal component analysis as an efficient method for capturing multivariate brain signatures of complex disorders—ENIGMA study in people with bipolar disorders and obesity. *Human Brain Mapping*, 45(8): e26682.
- Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234–241. Springer.
- Rueckert, D.; Frangi, A. F.; and Schnabel, J. A. 2001. Automatic construction of 3D statistical deformation models using non-rigid registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 77–84. Springer.
- Sinclair, M.; Schuh, A.; Hahn, K.; Petersen, K.; Bai, Y.; Batten, J.; Schaap, M.; and Glocker, B. 2022. Atlas-ISTN: joint segmentation, registration and atlas construction with image-and-spatial transformer networks. *Medical Image Analysis*, 78: 102383.
- Starck, S.; Sideri-Lampretsa, V.; Kainz, B.; Menten, M. J.; Mueller, T. T.; and Rueckert, D. 2024. Diff-def: Diffusion-generated deformation fields for conditional atlases. *arXiv preprint arXiv:2403.16776*.
- Thorley, A.; Jia, X.; Chang, H. J.; Liu, B.; Bunting, K.; Stoll, V.; de Marvao, A.; O’Regan, D. P.; Gkoutos, G.; Kotecha, D.; et al. 2021. Nesterov accelerated ADMM for fast diffeomorphic image registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 150–160. Springer.
- Wang, Q.; Chen, L.; Yap, P.-T.; Wu, G.; and Shen, D. 2010. Groupwise registration based on hierarchical image clustering and atlas synthesis. *Human Brain Mapping*, 31(8): 1128–1140.
- Wu, G.; Jia, H.; Wang, Q.; and Shen, D. 2011. Sharp-Mean: groupwise registration guided by sharp mean image and tree-based registration. *NeuroImage*, 56(4): 1968–1981.
- Yang, G.; Huang, X.; Hao, Z.; Liu, M.-Y.; Belongie, S.; and Hariharan, B. 2019. Pointflow: 3d point cloud generation with continuous normalizing flows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4541–4550.
- Ying, S.; Wu, G.; Wang, Q.; and Shen, D. 2014. Hierarchical unbiased graph shrinkage (HUGS): a novel groupwise registration for large data set. *NeuroImage*, 84: 626–638.
- Yu, E. M.; Dalca, A. V.; and Sabuncu, M. R. 2020. Learning conditional deformable shape templates for brain anatomy. In *International Workshop on Machine Learning in Medical Imaging*, 353–362. Springer.
- Zhang, T.; Zheng, S.; Cheng, J.; Jia, X.; Bartlett, J.; Cheng, X.; Qiu, Z.; Fu, H.; Liu, J.; Leonardis, A.; et al. 2024. Structure and Intensity Unbiased Translation for 2D Medical Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Figure 2:

算法 1: DARC 的 PyTorch 风格伪代码

```

def diffeo_pairwise_reg(
     $I$ ,  $A$ ,  $K$ ,  $\lambda$ 
):
    #  $I$ : moving image;  $K$ : number of iterations
    #  $A$ : atlas;  $\lambda$ : smooth parameter
     $v = \text{initialize}()$ 
    for  $k$  in range( $K$ ):
         $\phi = \text{Exp}(v)$  # squaring and scaling
        data_term = dissimilarity( $I \circ \phi$ ,  $A$ )
        loss = data_term +  $\lambda \|\nabla \phi\|_2^2$ 
        loss.backward() # compute gradient
        optimizer.step() # optimization
    return  $\phi$ 

def SGD_for_atlas(
     $\mathcal{I}$ ,  $\Phi$ ,  $K$ 
):
    #  $\mathcal{I} = \{I_i\}_{i=1}^n$ : all subject images;
    #  $\Phi = \{\phi_i\}_{i=1}^n$ : deformation fields;  $K$ : epochs
     $A = \text{initialize}()$ 
    for  $k$  in range( $K$ ):
        for  $\mathcal{B}_{\mathcal{I}}$ ,  $\mathcal{B}_{\Phi}$  in dataloader: #batch
            loss = dissimilarity( $\mathcal{B}_{\mathcal{I}} \circ \mathcal{B}_{\Phi}$ ,  $A$ )
            loss.backward()
            optimizer.step()
    return  $A$ 

def main():
     $A = \text{initialize}()$ 
    for  $k$  in range( $K_1$ ):
        # step 1: compute deformations
         $\phi_i = \text{diffeo\_pairwise\_reg}(I_i, A, K_2, \lambda)$   $\forall i \in \{1, \dots, n\}$ 
        # step 2: centrality activation
         $\hat{\phi}_i = \phi_i - \frac{1}{n} \sum_j \phi_j \quad \forall i \in \{1, \dots, n\}$ 
         $\hat{\Phi} = \{\hat{\phi}_1, \dots, \hat{\phi}_n\}$  #collect deformations
        # step 3: update atlas
        if dissimilarity( $A$ ,  $L_1$ , MSE):
             $A = \text{closed\_form}(\mathcal{I} \circ \hat{\Phi})$ 
        else:
             $A = \text{SGD\_for\_atlas}(\mathcal{I}, \hat{\Phi}, K_3)$ 
    return  $A$ ,  $\hat{\Phi}$ 

```

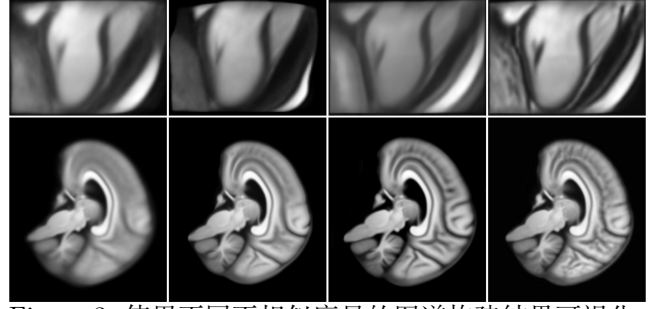
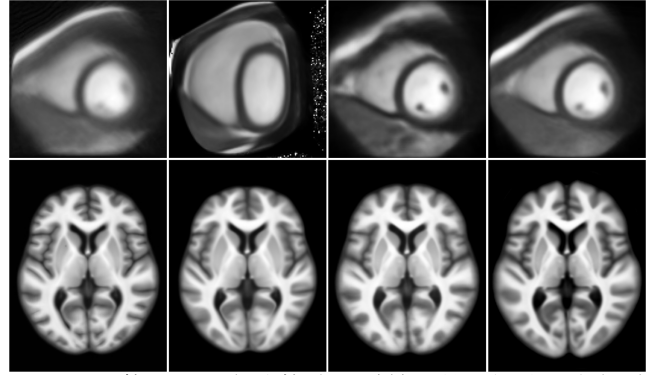
Figure 3: 使用不同不相似度量的图谱构建结果可视化。从左到右，结果分别对应于 MSE, L_1 , NCC, 和 SSIM。

Figure 4: 使用不同方法构建图谱结果的可视化。从左到右，结果分别对应 AtlasMorph、Aladdin、MultiMorph 和 DARC。

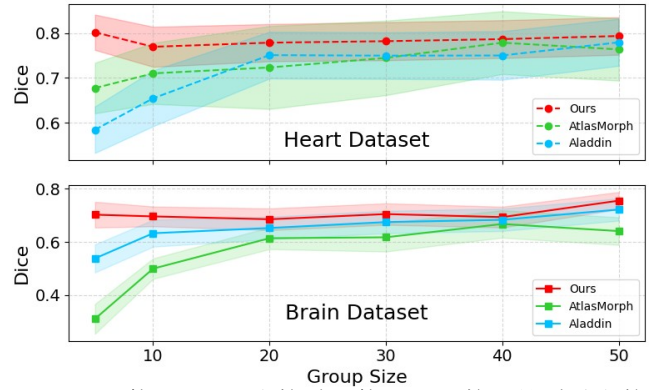


Figure 5: 使用 Dice 分数对比使用不同数量的受试者构建的图谱的性能。

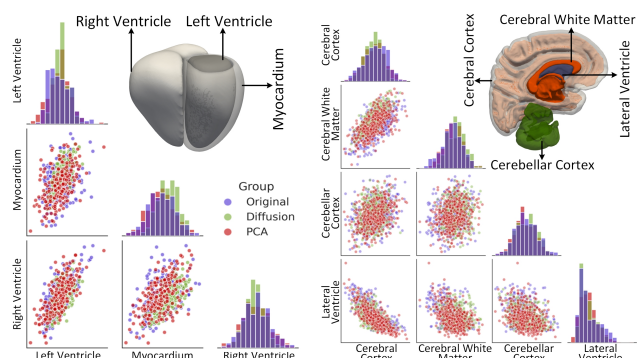


Figure 6: 视觉比较心脏和大脑数据集上由 PCA 和扩散模型生成的合形状分布。左侧面板比较了原始数据集的分布，并叠加了由 PCA 和扩散模型生成的分布，分别针对三个心脏结构。右侧面板对左脑的四个关键区域进行了类似的比较。

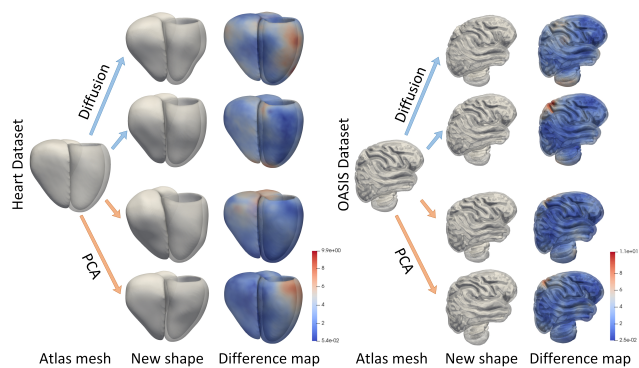


Figure 7: 使用 PCA/扩散模型生成的心脏和脑部结构的合形状的视觉示例。

5.1 大脑图谱构建方法的比较

图 7 对比了使用不同方法和不同主体数量的图谱构建结果。我们的 DARC 在小规模数据集上相较于其他两种最先进的方法，展示了更优越的稳定性和可视化质量。另一方面，对于大规模数据集，AtlasMorph 生成的图谱过于平滑，而 Aladdin 的显著居中误差则导致过度变形。相比之下，我们的 DARC 图谱展现了更佳的表现。

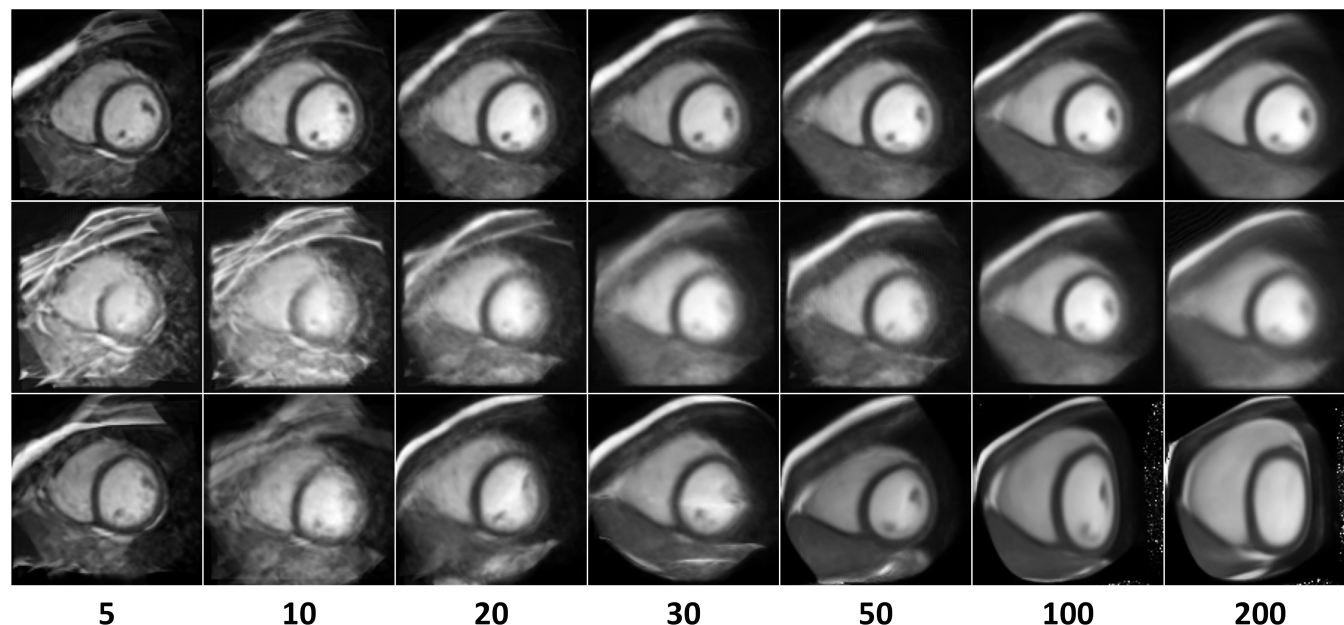


Figure 8: 使用不同人数的个体和不同方法构建的图谱可视化。自上而下的行分别代表来自 DARC、AtlasMorph 和 Aladdin 的结果，而自左至右的列则表示增加的个体数量。

px

5.2 一次性分割

图 8 显示了各种方法的单次分割结果，我们的 DARC 方法在心脏和 OASIS 数据集的样本上实现了最高精度和最视觉准确的结果。

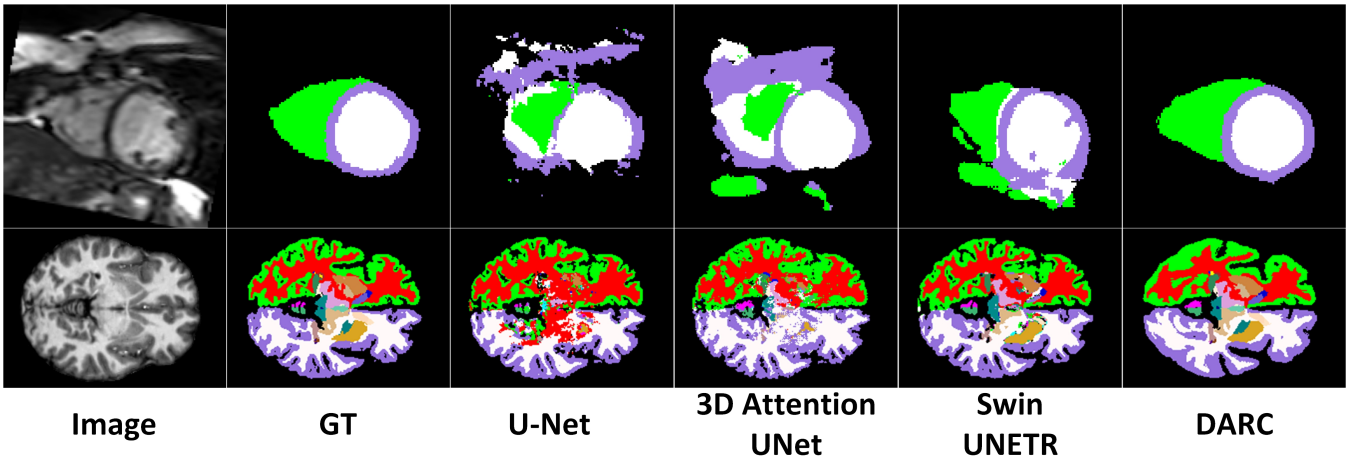


Figure 9: 不同方法一次性分割结果的可视化。

图 9 展示了所有方法的 Dice 准确率箱线图。我们的 DARC 方法在两个数据集上都获得了最高的中位数分数，且显示出最小的四分位距范围和极少的离群值，这突显了其卓越的准确性和稳健性。

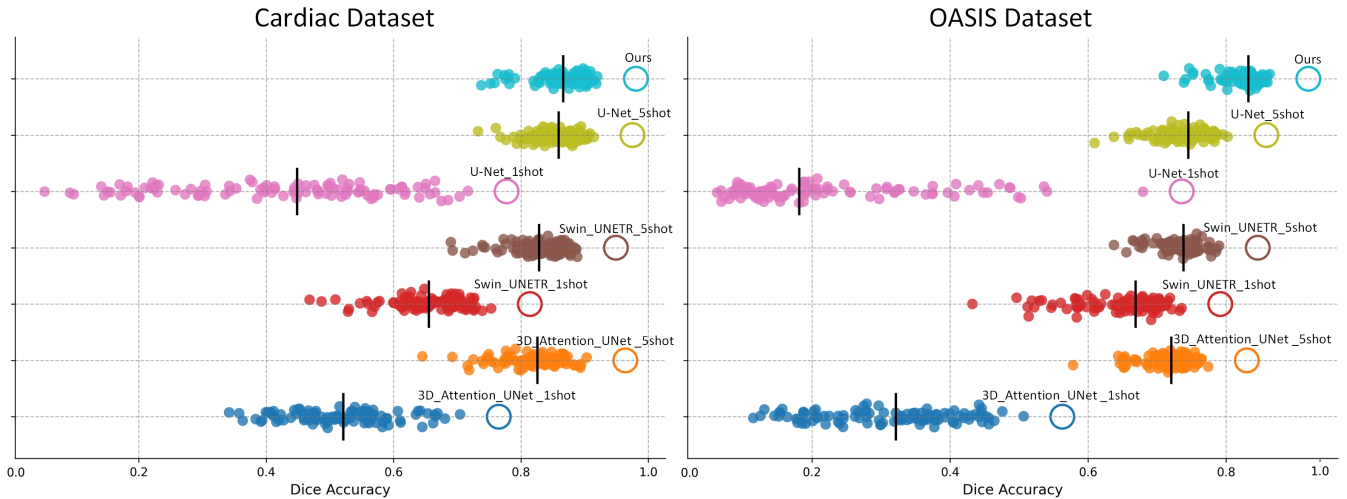


Figure 10: 我们的一次性分割方法与少样本分割方法在两个数据集上的箱形图比较。

图 10 展示了右脑白质的主成分分析 (PCA)，第一个主成分（模式）对应于全局扩展或收缩，整个大脑尺寸均匀变化；第二个主成分揭示了区域变化，突出了不同皮质区域的形状变化相反；第三个主成分捕捉到了不对称和扭曲，表明局部的剪切现象。这个观察结果与 (McWhinney et al. 2024) 中报道的结果一致。

图 11 展示了左心室主成分分析，第一主成分反映了大小，第二主成分捕捉了伸长和倾斜，第三主成分表示伸长和圆度。这个观察结果与 (Bai et al. 2015; Hoogendoorn et al. 2012) 中报道的研究结果一致。

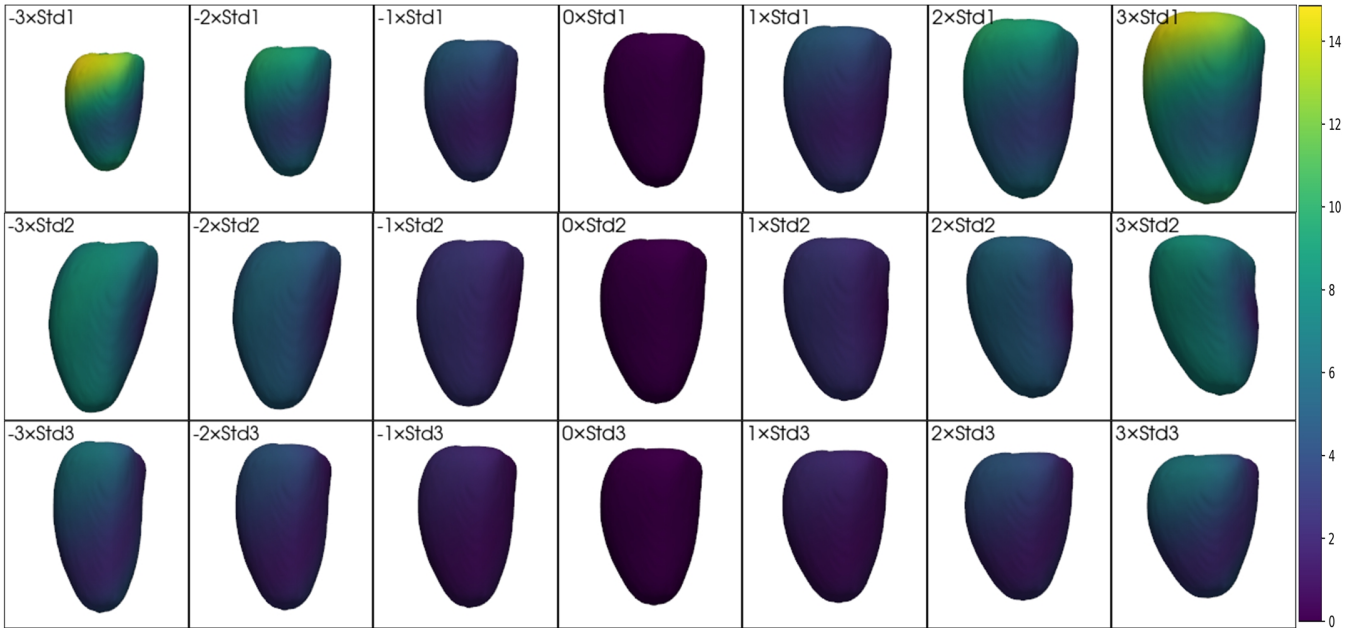


Figure 11: 左心室的 PCA 可视化，顶部至底部的行分别表示第一、第二和第三主成分。