

采用分割一切模型的弱监督互学习三维医学图像配准方法

苏渝钦¹ 潘杭镇¹ 李锦江¹ 刘欢² 李钟毓³ 刘杨^{4,5}

(1. 西安交通大学软件学院, 710049, 西安; 2. 西安交通大学计算机科学与技术学院, 710049, 西安; 3. 上海交通大学计算机学院, 200032, 上海; 4. 西安交通大学网络空间安全学院, 710049, 西安; 5. 西安交通大学智能网络与网络安全教育部重点实验室, 710049, 西安)

摘要: 针对三维医学图像配准中无监督方法配准精度不足、弱监督方法依赖高质量分割标签且分割基础模型生成标签难以直接稳定指导配准训练的问题, 提出了一种采用分割一切模型的弱监督互学习三维医学图像配准方法。首先, 利用分割一切模型从不同视角提取主要解剖区域掩码, 并通过融合构建交集掩码与并集掩码, 以表征不同粒度的结构先验信息。其次, 在弱监督配准框架中引入互学习机制, 通过结合可信度与位置信息模块, 促进不同掩码监督下配准模型之间的知识蒸馏与协同优化, 从而实现互补信息的有效利用并提升训练稳定性。最后, 在三组脑图像数据集上的实验结果表明, 所提方法能够在充分利用标签信息的同时保持稳定训练, 与现有三维图像配准方法相比, 在区域重叠度等指标上表现更优, 其中重叠系数提升约5%, 体现出更高的配准精度与可靠性。

关键词: 医学图像配准; 分割一切模型; 弱监督方法; 多视角; 互学习方法

中图分类号: TP391.41 **文献标志码:** AA

文章编号: 0253-987X(XXXX)XX-0001-17

Foundation Model Guided 3D Medical Registration Via Weakly-Supervised Mutual Learning

SU Yuqin¹, PAN Hangzhen¹, LI Jinjiang¹, LIU Huan², LI Zhongyu³, LIU Yang^{4,5}

- (1. School of Software Engineering, Xi'an Jiaotong University, Xi'an 710049, China;
2. School of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China;
3. School of Computer Science, Shanghai Jiao Tong University, Shanghai 200032, China;
4. School of Cyber Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China;
5. MOE Key Lab for Intelligent and Network Security, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: To address the problems that unsupervised methods for 3D medical image registration suffer from limited accuracy, weakly supervised methods rely on high-quality segmentation labels, and labels

收稿日期: 2025-11-17。

generated by foundation segmentation models cannot directly and stably guide registration training, a weakly supervised mutual-learning method for 3D medical image registration guided by a foundation segmentation model is proposed. First, masks of major anatomical regions are extracted from different views using segment anything model, and intersection and union masks are constructed through mask fusion to represent structural prior information at different granularities. Second, a mutual-learning mechanism is introduced into the weakly supervised registration framework, where a confidence and position information module is incorporated to promote knowledge distillation and collaborative optimization between registration models under different mask supervision, thereby enabling effective utilization of complementary information and improving training stability. Finally, experimental results on three brain image datasets demonstrate that the proposed method can effectively exploit label information while maintaining stable training, and outperforms existing 3D image registration methods on multiple evaluation metrics, showing better registration accuracy and reliability.

Keywords: medical image registration; segment anything model; weakly-supervised method; multi-view; mutual learning

随着医学影像技术与相关交叉学科的持续发展,医学图像分析研究已逐步进入精细化、标准化的发展阶段。在医学图像配准任务中,标准图谱与模板的构建为不同模态、不同个体之间的空间对齐提供了重要参照。以脑影像分析为例,2020年Allen脑科学研究所发布了小鼠标准脑图谱CCFv3^[1],为小鼠大脑的空间配准提供了重要参照。在此基础上,诸如mBrainAligner^[2]与D-LMBmap^[3]等代表性工作相继提出,实现了多种全脑影像与标准模板之间的高精度配准,为大规模医学图像分析任务,尤其是全脑尺度下的单细胞神经元形态分析与类型鉴定奠定了技术基础。然而,复杂解剖结构带来的形变差异以及对配准精度不断提高的要求,仍然对现有医学图像配准方法提出了严峻挑战。

近年来,基于深度学习的无监督配

准方法已成为医学图像配准领域的主流技术^[4-10]。相较于传统基于迭代优化的方法^[11-13],这类方法在配准速度上实现了数量级的提升;然而,其在配准精度方面的改进仍然有限,尤其在解剖结构复杂的医学图像数据上,性能仍有待提升。为了进一步提高配准精度,部分研究转向弱监督范式,通过引入分割标签作为监督信号来引导配准过程,并取得了显著效果^[14-18]。但此类方法依赖于专家进行人工标注,成本高昂且难以大规模获取。

随着分割一切模型(segment anything, SAM)^[19]的发布,研究者开始探索其在配准中的应用,并涌现出如SAMReg模型^[20]、SAT-Morph模型^[21]等一系列新模型。这些工作尝试直接利用分割基础模型的分割结果指导配准,展现出良好潜力。然而,这些模型的性能在很大程度上受限于基础模型本身的分

割精度,且普遍忽视了所生成伪标签的多样性与可靠性问题,因而其泛化能力与鲁棒性面临挑战。

为解决上述问题,提出了一种由分割一切模型引导的弱监督互学习三维医学图像配准方法。首先,针对三维医学图像的成像特点,采用多视角策略对SAM模型进行微调,以生成主要解剖结构的伪分割标签。进一步地,为增强伪标签的多样性与可靠性,设计了一种多视角融合机制,通过对不同视角下生成的标签进行交集与并集操作,有效整合多源空间信息。最后,基于所生成的多样化伪标签,构建了一种基于互学习的配准框架,并引入基于可靠性的损失模块,以确保模型能够充分、稳健地利用伪标签中的解剖结构信息。本文的贡献主要如下:针对弱监督配准中解剖标注稀缺的问题,引入SAM模型以生成初始伪标签,并进一步提出一种多视角融合策略,充分利用三维图像的立体信息以增强伪标签的多样性与鲁棒性;提出了一种基于互学习的配准框架,该框架能促使配准网络之间通过协同训练,充分挖掘并利用图像中的解剖结构信息,从而共同提升配准性能;在多个数据集上的实验结果验证了所提方法的有效性。结果表明,基于本文生成的多样化伪标签进行互学习,能够实现有效的知识蒸馏,显著提升配准精度。这一工作为互学习范式在图像配准领域的拓展与应用提供了新的思路。

1 相关工作

1.1 弱监督图像配准方法

图像配准旨在建立不同图像间的空间对应关系,其方法可分为刚性配准与形变(非刚性)配准两类^[22-23]。从表示方式来看,配准变换通常可基于参数化函数或密集位移场来表征。参数化方法(如刚性、仿射及基于样条的变换^[24]),对关键位置之间的映射进行建模^[25];而密集位移场则为每个体素定义其位移向量,通常由基于学习的配准模型直接输出^[26]。

早期的学习型方法大多采用监督学习的策略,其监督信息来源于人工标注或传统方法估计的变形场^[27-28]。VoxelMorph模型^[4]首次将3D-UNet网络^[29]和空间变形块^[30]结合,实现了端到端的无监督配准,为此后一系列基于深度学习的先进方法奠定了基础。

弱监督配准方法的核心思想是利用不完整或稀疏的标注信息(如部分关键点、分割掩码、解剖先验等)引导模型学习,从而在降低对全标注数据依赖的同时,仍能获得有效的配准效果。胡等人^[6]开创性地利用器官分割标签作为弱监督信号;Blendowski等人^[31]则进一步提出了基于多模态特征的弱监督学习方法。尽管这些方法在性能上显著优于无监督方法,但其依赖的分割标签仍需专业医学知识进行人工标注,成本高昂。

为缓解标注负担,李等人^[32]提出一种半监督的弱监督配准框架,结合少量标注数据与大量未标注数据进行协同训

练。与上述方法不同,本文提出借助在有限标注数据上微调分割一切模型的方式生成高质量伪标签,仅需少量标注即可完成高效训练,为降低标注依赖提供了新的技术路径。

1.2 视觉基础模型

近年来,基础模型在计算机视觉领域发展迅速,通过在海量多样化数据集上进行预训练,展现出强大的泛化能力与广泛的下游任务适应性。SAM模型^[19]作为首个通用分割基础模型,在多种零样本分割任务中表现卓越。此后,研究者陆续提出了一系列基于SAM模型的改进模型,如SAT模型^[33]、MedSAM模型^[34]等,推动了视觉基础模型在专业领域的应用。

随着医学分割基础模型的发展,将其用于辅助图像配准逐渐成为可能。例如,徐等人^[21]提出SAT-Morph模型,直接使用SAT模型^[33]输出的分割结果替代原始图像进行配准;黄等人^[20]则利用SAM模型^[19]生成的分割掩码作为感兴趣区域,并通过区域匹配实现图像配准。然而,由于此类基础模型通常以二维图像作为输入,现有方法大多将三维数据沿单一视角切片后分别处理,导致丢失重要的三维空间上下文信息。此外,其配准性能高度依赖于基础模型的分割质量,而未能充分考虑所生成伪标签的可靠性与多样性问题。

与传统的可靠性评估方法不同,本文基于“伪标签越可靠,其引导的配准

精度越高”的基本假设,提出以配准精度本身作为衡量伪标签可靠性的依据,从而更直接地关联分割质量与配准目标。

1.3 互学习方法

在相互学习的概念提出之前,主流深度学习模型通常采用独立训练方式,即每个模型在固定数据集上通过最小化损失函数完成训练,模型之间不存在交互。Hinton等人^[35]提出的知识蒸馏,技术改变了这一范式,通过将训练完备的大型“教师模型”中的知识迁移至轻量级“学生模型”,在维持模型性能的同时显著提升了推理效率。然而,该方法中的知识传递是单向的,学生模型的学习效果高度依赖于教师模型的质量。

为克服上述限制,张等人^[36]提出相互学习的概念,其核心思想是让多个模型在训练过程中通过知识蒸馏相互引导、共同优化,并在图像分类任务中取得了显著效果。此后,互学习机制逐步扩展至物体检测、图像分割与光流估计等多个视觉任务^[37-39]。在图像配准领域,刘等人^[40]构建了一种基于互学习的配准框架,通过师生网络之间的相互参数更新以提升性能。然而,该方法中师生网络均基于相同的预训练数据与网络结构初始化,导致模型间差异性有限,从而制约了互学习的潜力。一般而言,传统互学习方法中不同学生模型之间的差异主要依赖于随机初始化或超参数设置,这种差异来源较为单一,难以充分

激发模型间的互补性。

为突破以上局限,本文提出使用多样化的分割伪标签分别指导不同的配准网络,从而在输入层面引入显著差异性,增强模型学习的多样性。在此基础上,通过互学习机制促进各网络间有效互补与知识融合,实现对解剖结构特征更充分、更可靠的学习,最终提升配准性能。

2 方法

2.1 问题描述

在典型的弱监督配准场景中,给定一对图像 I_m 和 I_f , 其中 I_m 为移动图像 (moving image), I_f 为固定图像 (fixed image), 以及它们对应的分割掩码, 目标是求得一个形变场 ϕ , 将 I_m 变换以与 I_f 配准。本文将 $I_m \circ \phi$ 视为由 I_m 变换得到的变形图像。本文旨在解决在缺乏分割掩码的情况下如何提升配准精度的问题。为此, 本文利用分割一切模型生成伪分割掩码, 并将其引入互学习框架, 以充分利用标签信息并增强模型的配准能力。在此基础上, 本文提出一种结合 SAM 微调与互学习的联合配准框架 SAM-MutReg。

2.2 基于 SAM 模型微调的多视角掩码获取

SAM-MutReg 由 SAM 模型微调、预训练和互学习 3 个阶段组成, 其整体流程见图 1。

在 SAM 微调阶段, 设计了一种多视角融合的三维医学图像分割方法。对于配准任务中的固定图像 $I_f \in \mathbb{R}^{H \times W \times D}$, 其中 H 、 W 和 D 分别表示图像的高度、宽度和深度。考虑到所采用的三维医学图像数据在水平面与冠状面两个视角下具有较好的成像质量, 分别从这两个视角将三维图像分解为二维切片 $S_h \in \mathbb{R}^{W \times D}$ 和 $S_d \in \mathbb{R}^{H \times W}$ 。以两组切片作为输入, 分别对 SAM 模型进行微调, 从而获得具有不同视角特征的分割模型。随后, 将数据按照对应视角切片后输入相应模型, 得到基于不同视角的分割掩码伪标签。

在方法设计过程中, 对水平面、冠状面和矢状面 3 种正交视角下的伪分割生成策略进行了探索。实验结果表明, 对于本研究所选取的 5 个解剖区域, 矢状面视角在部分样本中易产生切片间不连续和噪声较大的伪分割结果; 同时, 三视角融合在整体分割质量上未表现出稳定优势, 反而会对后续配准产生不利影响。基于上述分析, 最终选取结构连续性和成像稳定性更优的水平面与冠状面视角用于伪标签生成, 并进一步研究两视角伪标签的互补融合与互学习机制, 如图 1(a) 所示。

为了充分融合多视角特征, 本文进一步设计了掩码融合模块。设两组伪标签分别为 S_1 和 S_2 , 定义并运算和交运算如下

$$(S_1 \cup S_2)(p) = \begin{cases} 1, & \max(S_1(p), S_2(p)) = 1 \\ 0, & \text{其他.} \end{cases} \quad (1)$$

$$(M_1 \cup M_2)(p) = \begin{cases} 1, & \min(M_1(p), M_2(p)) = 1 \\ 0, & \text{其他.} \end{cases} \quad (2)$$

式中: p 表示像素点。通过对两视角伪标签进行交与并运算, 可分别获得两种具有互补特征的融合伪标签 M_u 和 M_i 。在实际实验中, 本文从三维脑影像中提取了5个主要脑区, 包括脑干 (BS)、大脑皮层 (CTX)、海马体 (HPF)、小脑 (CB) 以及尾壳核 (CP)。

2.3 互学习配准框架

在第二阶段 (见图1 (b)), 基于第一阶段生成的交集和并集伪标签, 按照常规弱监督配准范式分别预训练交集监督模型 ISM_1 (intersection-supervised

model) 和并集监督模型 USM_1 (union-supervised model)。在配准的预训练阶段, 本文使用第一阶段SAM生成多视角下多个区域的分割结果, 并采取基于交集和并集的融合策略生成对应的伪标签, 以此作为弱监督标签, 预训练阶段不依赖任何人工标注或现成的配准结果。在此基础上, 借鉴互学习 (mutual learning) 的思想, 将 ISM_1 与 USM_1 均视为学生模型, 使二者在后续阶段中实现相互学习与协同优化。

如图2 (c) 所示, 在第 $t(t \geq 1)$ 轮训练中, 首先固定模型 ISM_t 的参数, 更新模型 USM_t 的参数, 使模型 USM_t 能够充分吸收 ISM_t 的知识。为了确保知识传递的有效性与稳定性, 本文设计了一个基于置信度与掩码位置混合判断的置信模块。

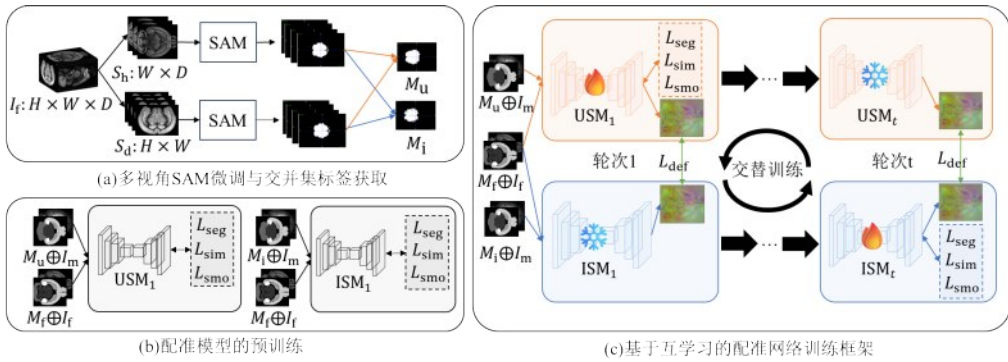


图1 本文提出的配准整体框架

Fig. 1 Overview of the proposed registration framework

置信度损失计算图如图2所示, 其中 $\phi_{1,2}$ 表示模型输出的形变场, $W_{1,2}$ 表示将形变场作用于移动图像后得到的配

准后图像, $N_{1,2}$ 表示对应的置信度图。由于配准任务的核心输出为形变场, 因此置信度反映了形变场在每个像素位置

上的可靠程度。该可靠性可通过强度相似性函数进行度量,即形变后的图像与固定图像之间的相似性越高,对应位置的置信度越高。本文使用归一化互相关(normalized cross-correlation, NCC)作为相似性度量函数。对于图像 I_1 和 I_2 在位置 p 处的NCC计算公式如下

$$ncc(I_1, I_2; p) = \frac{\sum_{q \in \Omega_p} S_1(q) S_2(q)}{V_1(q) V_2(q)} \quad (3)$$

$$S_k(q) = I_k(q) - I_k(p) \quad (4)$$

$$V_k(q) = \sqrt{\sum_{q \in \Omega_p} S_k(q)^2} \quad (5)$$

式中: Ω_p 是以 p 为中心的局部窗口, $\bar{I}_1(p)$ 和 $\bar{I}_2(p)$ 是局部窗口内的算术平均值, $k \in \{1, 2\}$ 。根据置信度判断结果可生成置信度掩码,其中掩码值为1的部分代表该像素点上形变场 ϕ_1 的置信度更高,值为0的部分则相反。置信度掩码公式如下

$$M_{\text{conf}} = \text{conf}(ncc(W_1, I_f), ncc(W_2, I_f)) \quad (6)$$

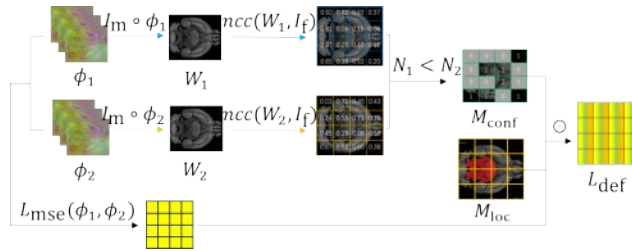


图2 置信度损失计算图

Fig. 2 Overview of the confidence loss calculation module

在复杂度方面,本文方法的改进主要体现在训练策略层面,配准网络本身的结构与参数规模并未增加。互学习阶

$$\text{conf}(A, B) = \begin{cases} 1, & B > A \\ 0, & B \leq A \end{cases} \quad (7)$$

式中: W_1 和 W_2 为模型得到的形变场作用在移动图像上得到的配准后图像, I_f 为固定图像。同时,依据分割掩码伪标签可得到基于掩码位置的掩码 M_{loc} ,以进一步关注解剖区域的配准可靠性。通过位置掩码损失可以更加关注分割区域的置信度。本文通过计算两个形变场之间的均方误差(MSE)作为基础损失,并将其与置信度掩码及位置掩码进行加权融合,得到最终的形变场损失函数

$$L_{\text{def}} = L_{\text{mse}} \odot M_{\text{conf}} \odot M_{\text{loc}} \quad (8)$$

式中: L_{mse} 为形变场之间的均方差损失; $\odot$ 符号为矩阵之间的哈达玛乘积。通过上述设计,模型间的互学习可在置信度模块的引导下实现可靠且高效的知识传递,从而提升配准精度与训练稳定性。从模型 USM_t 训练得到 USM_{t+1} 后,固定模型 USM_{t+1} 的参数然后按照相同方法训练模型 ISM_t 得到 ISM_{t+1} , 并重复该过程直至模型收敛。

段虽然同时使用两个结构相同的配准模型,但每次迭代仅更新其中一个模型的参数,另一模型保持冻结,因此参与梯

度更新的可学习参数量与单模型训练一致。在推理阶段,本文方法仅使用单个配准模型进行前向计算,不再涉及多视角伪标签生成与互学习过程,推理流程与标准配准模型完全一致,多视角伪标签仅作为训练阶段的辅助信息,不会影响测试阶段的配准效率。

2.4 损失函数

在预训练过程中,按照一般弱监督配准过程,损失函数可表如下

$$L_{\text{sup}} = L_{\text{sim}} + \alpha L_{\text{smo}} + \beta L_{\text{seg}} \quad (9)$$

式中: L_{sim} 表示固定图像和配准后图像的相似性损失; L_{smo} 表示位移场的正则化损失; L_{seg} 表示固定图像分割标签和配准后图像分割标签之间的相似性损失; α 和 β 表示权重参数。对于 L_{sim} , 使用区域归一化互信息作为相似性度量, 对于 L_{seg} , 使用区域重叠系数 (Dice) 相似性损失。对于 L_{smo} , 使用位移场的梯度的 L_2 范数作为正则化

$$L_{\text{smo}}(\phi) = \frac{1}{|\Omega|} \sum_{p \in \Omega} \|\nabla \phi(p)\|^2 \quad (10)$$

式中: $\nabla \phi(p)$ 表示在体素 p 处的梯度。在互学习过程中, 损失函数如式所示

$$L_{\text{mut}} = L_{\text{sim}} + \alpha' L_{\text{smo}} + \beta' L_{\text{seg}} + \gamma' L_{\text{def}} \quad (11)$$

式中: L_{sim} , L_{smo} , L_{seg} 与式 (9) 中相同, L_{def} 如式 (8) 所示。其中 α' , β' , γ' 为权重参数。通过上述多项损失的联合优化, 置信度评估不会在早期阶段对模型产生主导性影响。随着训练推进和配准精度逐步提升, 置信度评估的可靠性也随之增强, 从而在后期更有效地引导

互学习过程。模型能够在保证形变场平滑性的同时, 充分利用分割标签信息和互学习机制, 从而显著提升配准精度与模型稳定性。

2.5 评价指标

为了综合评价本文提出的配准框架在三维小鼠大脑图谱配准中的性能, 采用区域重叠系数 (S_{Dice}), 95% 豪斯多夫距离 ($D_{\text{HD95\%}}$) 和负雅可比行列式比例 (J) 这 3 项指标进行定量评价。其中, S_{Dice} 和 $D_{\text{HD95\%}}$ 用于衡量配准的精度, 而 J 用来评估形变场的平滑性和可逆性。各指标的计算公式如下

$$S_{\text{Dice}} = \frac{2|I_1 \cap I_2|}{|I_1| + |I_2|} \quad (12)$$

$$D_{\text{HD95\%}} = P_{95}(D_x \cup D_y) \quad (13)$$

$$D_x = \{ \min_{y \in Y} d(x, y) \mid \forall x \in X \} \quad (14)$$

$$D_y = \{ \min_{x \in X} d(x, y) \mid \forall y \in Y \} \quad (15)$$

$$J(p) = \det(\nabla \phi(p)) \quad (16)$$

$$J = \frac{1}{N} \sum_p [J(p) < 0] \quad (17)$$

式中: I_1 和 I_2 分别表示固定图像与配准后图像的分割结果; $P_{95}(\cdot)$ 表示选取集合中距离的 95% 分位值; ϕ 为形变场; $J(p)$ 表示其在体素 p 处的雅可比行列式。在性能评价中, S_{Dice} 越高, $D_{\text{HD95\%}}$ 值越低表示配准精度越高; 而 J 值越低, 说明形变场越平滑、形变结果越合理。

3 实验与结果

3.1 数据集

为全面验证本文提出的配准算法性能,实验在3组显微成像下的三维脑图像数据集上进行,分别为Soma-CCFv3数据集、P21数据集和P28数据集。

其中,Soma-CCFv3数据集包含了光片荧光显微镜成像的12个成年小鼠大脑三维图像(Soma)以及通过串行双光子断层扫描成像的1个小鼠大脑三维图谱(CCFv3)。P21数据集包含了30个出生后第21天的鼠大脑三维图像。P28数据集则包含了30个出生后第

28天的鼠大脑三维图像,均基于光片荧光显微镜成像采集。

为了保证数据一致性,所有原始图像均被统一裁剪至 $80 \times 114 \times 132$ 的体素尺寸。实验中,Soma-CCFv3数据集使用CCFv3小鼠脑图谱作为固定图像进行配准。P21和P28数据集则从各自的30个样本中随机选取一例作为固定图像,其余作为移动图像。

如图3所示,展示了CCFv3小鼠脑三维图像及其对应的5个主要脑区的三维可视化结果,主要脑区包括脑干(BS)、大脑皮层(CTX)、海马体(HPF)、小脑(CB)和尾壳核(CP)。

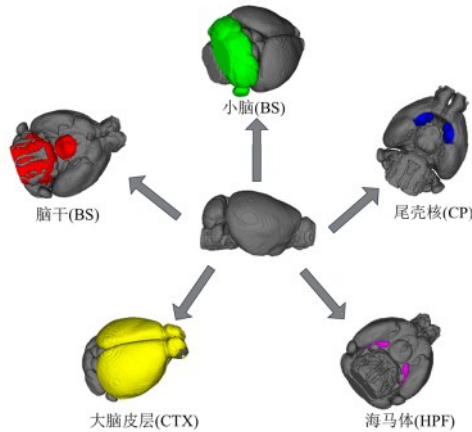


图3 本研究所涉及的5个主要脑区

Fig. 3 The five major brain regions involved in this study

3.2 模型训练

本文所有的模型的训练和测试均在NVIDIA RTX 3090 GPU上进行,实验环境基于PyTorch深度学习框架。在预训练阶段,采用自适应矩阵估计优化器

(Adam)进行网络参数更新,训练轮次设置为50。学习率使用“预热—固定步长衰减”策略,从0逐渐升至 1×10^{-3} ,随后衰减至 1×10^{-5} 。批量大小(batch size)设为1。超参数设置为 $\alpha=0.1$,

$\beta=1$ 。在互学习训练阶段,学习率固定为 1×10^{-3} 。此外损失函数中引入了额外的互学习约束项,其对应的权重参数 $\alpha'=0.1$ 、 $\beta'=1$ 、 $\gamma'=1$ 。上述超参数通过验证集上的经验调优确定,在保证配准精度提升的同时避免形变场不稳定。

3.3 对比与消融试验

为了验证本文方法在三维脑图像配准中的有效性,本文将其与多个代表性配准算法进行了对比实验。对比方法包括公认性能优良且被广泛采用的传统基于优化的配准模型 SyN^[41]与 Elastix^[42],以及近年来提出的深度学习无监督配准模型 VoxelMorph (VM)^[4], TransMorph (TM)^[6], H-ViT^[7] 和 LessNet^[8]。此外,考虑到弱监督配准方法通常通过在既有无监督配准框架中引入分割或解剖标签约束来实现,本文进一步基于上述深度模型构建了使用由 SAM 模型生成伪分割标签的弱监督版本(如 VM+SAM、TM+SAM 等),以在相同模型结构和相同伪标签条件下评估不同方法的效果,从而保证对比实验的公平性。实验结果如表1所示,其中 S_{Dice} 和 $D_{HD95\%}$ 指标为5个主要脑区的平均值,表中加粗结果表示该指标下的最优性能。

从表1的结果可以看出,相较于深度学习的无监督配准方法,引入 SAM 模型生成伪标签的弱监督方法在配准性能上获得了显著提升。这表明伪标签所提供的结构先验特征能够有效促进模型

的配准过程。然而,直接使用伪标签的弱监督方法并未充分考虑伪标签的可靠性问题。在训练过程中,模型可能错误地利用了伪标签中存在噪声或分割误差的区域,导致训练过程不稳定,并受限于伪标签自身的精度,最终配准性能提升有限。相比之下,本文提出的方法通过引入互学习机制,不仅充分挖掘了伪标签中的有效特征信息,还在知识蒸馏过程中动态过滤了不可靠区域,从而显著提升了训练稳定性与配准精度。在 J 指标上,本文方法未取得最优值。这主要源于本文在训练过程中更加关注解剖结构一致性与互学习过程的稳定性,在形变场平滑性约束上采取了相对保守的策略,从而在一定程度上限制了该指标的进一步优化。然而,该策略有效避免了严重的形变场折叠,并在 S_{Dice} 与 $D_{HD95\%}$ 等核心配准精度指标上取得了更优表现,体现了配准精度与形变合理性之间的权衡关系。

为了验证本文所提出配准框架中各个模块的有效性,本文分别在 SomaCCFv3、P21 和 P28 数据集上进行了消融实验,结果如表2所示。

作为脑图像配准中最为广泛使用的算法,本文将使用伪标签的 VM 模型^[4]作为基线模型(Base),并在其上逐步加入不同模块进行对比分析。其中, T1 和 T2 表示互学习过程中循环训练的轮数(实验中设定为 T1=1、T2=2), Confidence Module (CM) 表示在互学习蒸馏过程中引入的可靠性评估模块。

通从表2中可以观察到,比较第1、2、3行结果,在未引入可靠性模块的情况下,互学习机制仍能带来一定的配准性能提升。这主要得益于模型在不同伪标签间的交互学习,使其能够更充分地利用标签信息。与此同时,形变场指标表明该互学习策略在一定程度上减轻了形变场折叠现象,但随着循环次数增加,错误折叠数量略有上升,说明伪标签中不可靠区域会对互学习过程造成干扰。进一步比较第2、3行结果可见,性能提升幅度趋于平稳,表明模型已基本收敛。当引入可靠性模块(第4、5行)后,模型在所有评价指标上均取得明显提升,说明该模块能够有效识别并抑制伪标签中的不可靠区域,从而提高互学习过程的稳定性与整体配准性能。以上结果充分验证了本文所提出可靠性模块在提升模型鲁棒性和配准精度方面的有效性。

为进一步分析多视角伪分割掩码在不同融合策略下的特性,本文在Soma-CCFv3数据集上对交集与并集掩码在多个脑区上的信息完整性与可靠性进行了定量评估。结果如表3所示,可以观察到,交集掩码在各脑区上均表现出较高的精确率(Precision),但召回率(Recall)相对较低,表明其能够有效减少错误伪标签,但同时会遗漏部分真实解剖结构;相比之下,并集掩码显著提高了召回率,但精确率明显下降,说明其在覆盖更多解剖区域的同时引入了更多噪声。上述现象在不同脑区上具有

一致性,反映了交集与并集掩码在可靠性与信息完整性之间的固有权衡关系。同时我们还将仅使用交集和并集掩码作为伪标签进行配准的结果(Dice(Reg))进行了对比,指标在多个脑区上呈现出明显差异,侧面验证了交并集掩码的多样化特征。

基于上述分析,单独依赖任一类伪标签均难以兼顾信息完整性与可靠性。本文通过互学习机制引导由不同伪标签监督的配准模型相互学习,并结合置信度模块动态抑制不可靠区域,从而更充分、稳健地利用解剖结构信息。

此外,为验证本文方法在分割一切模型预测质量下降情况下的鲁棒性,我们设计了伪分割掩码退化实验,通过逐步降低可用伪分割掩码比例来模拟伪标签质量退化的情形,为便于表示记 r 为分割掩码的随机保留采样比例,并且在不同比例下重复实验5次记录配准精度的平均值(蓝色线条)和波动范围(蓝色阴影)。如图4所示,基线方法是不使用伪分割掩码的无监督方法。随着伪分割信息的减少,本文方法的配准性能呈现出渐进式下降趋势,并在中度退化条件下仍显著优于基线方法,表明所提出的多视角融合、互学习及置信度机制在分割质量退化场景下具有一定的鲁棒性。

3.4 可视化分析

为了更加直观地比较和分析不同配准方法的效果,本文将在Soma-CCFv3

数据集上的实验结果进行可视化展示，如图 5 所示。

表 1 不同模型配准性能对比
Table 1 Registration performance
among different models

数据集	模型	S_{Dice}	$D_{HD95\%}$	J
Soma-CCFv3	SyN ^[41]	0.7281	6.8750	0.0426
	Elastix ^[42]	0.6943	7.6434	0.0149
	VM ^[4]	0.765	3.3761	0.0139
	TM ^[6]	0.7491	3.6763	0.0004
	H-ViT ^[7]	0.7649	2.9934	0.0020
	LessNet ^[8]	0.7655	2.7863	0.0021
	VM+SAM	0.8213	1.8617	0.0342
	TM+SAM	0.8341	1.7215	0.0121
	H-ViT+SAM	0.8412	1.6432	0.0426
	LessNet+SAM	0.8370	1.7423	0.0359
P21	本文	0.8816	1.5464	0.0222
	SyN ^[41]	0.8496	3.2315	<1e-04
	Elastix ^[42]	0.8511	3.5479	0.0072
	VM ^[4]	0.8604	1.5760	0.0005
	TM ^[6]	0.8778	1.4180	0.0013
	H-ViT ^[7]	0.8747	1.4262	0.0006
	LessNet ^[8]	0.8648	1.5338	0.0009
	VM+SAM	0.9155	1.3349	0.1259
	TM+SAM	0.9138	1.3041	0.0799
	H-ViT+SAM	0.9191	1.2245	0.1819
P28	LessNet+SAM	0.9074	1.4059	0.0132
	本文	0.9394	0.8309	0.0098
	SyN ^[41]	0.8110	4.0228	0.0001
	Elastix ^[42]	0.8097	3.2606	0.0012
	VM ^[4]	0.8254	1.8882	0.0018
	TM ^[6]	0.8324	1.7077	0.0019
	H-ViT ^[7]	0.8321	1.7387	0.0093
	LessNet ^[8]	0.8300	1.7461	0.0013
	VM+SAM	0.8802	1.4146	0.0109
	TM+SAM	0.8753	1.5799	0.0053
P28	H-ViT+SAM	0.8827	1.3991	0.0201
	LessNet+SAM	0.8779	1.4942	0.0059
	本文	0.9292	0.8414	0.0092

表 2 不同模块配准性能对比
Table 2 Registration performance
among different modules

数据集	模块	S_{Dice}	$D_{HD95\%}$	J
Soma-CCFv3	Base	0.8213	1.8617	0.0342
	Base+T1	0.8555	1.6701	0.0224
	Base+T2	0.8589	1.6232	0.0312
	Base+CM+T1	0.8641	1.5931	0.0191
	Base+CM+T2	0.8816	1.5464	0.0223
P21	Base	0.9155	1.3349	0.1259
	Base+T1	0.9213	1.1728	0.0205
	Base+T2	0.9258	1.0937	0.0311
	Base+CM+T1	0.9347	0.9112	0.0064
	Base+CM+T2	0.9394	0.8309	0.0098
P28	Base	0.8802	1.4146	0.0109
	Base+T1	0.9019	1.2511	0.0083
	Base+T2	0.9051	1.1946	0.0132
	Base+CM+T1	0.9134	1.0653	0.0061
	Base+CM+T2	0.9292	0.8414	0.0092

表 3 交集与并集伪分割掩码特性及其对配准性能的影响

Table 3 Characteristics of Intersection and Union Pseudo-Segmentation Masks and Their Impact on Registration Performance

	指标	BS	CB	CP	CTX	HPF
交集	Precision	0.95	0.91	0.96	0.94	0.95
	Recall	0.74	0.52	0.79	0.70	0.69
	Dice(Reg)	0.84	0.80	0.90	0.82	0.83
并集	Precision	0.68	0.48	0.60	0.74	0.48
	Recall	0.92	0.82	0.92	0.93	0.90
	Dice(Reg)	0.83	0.81	0.84	0.85	0.86
本文方法	Dice(Reg)	0.87	0.82	0.92	0.90	0.87

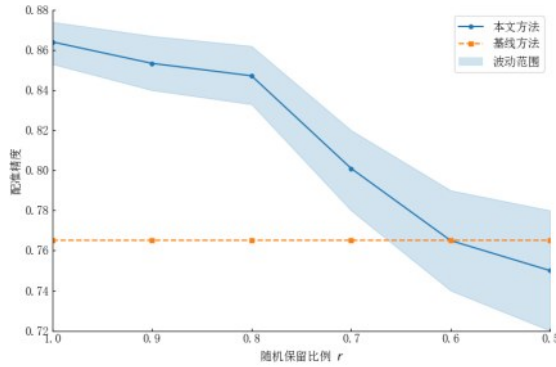


图4 分割质量退化条件下的配准鲁棒性分析

Fig. 4 Robustness analysis under degraded pseudo-segmentation quality

图5中展示了各方法在水平面视角下不同切片的脑区轮廓,其中黄色线条表示固定图像的脑区轮廓,用于直观评估配准后的结构重合程度,紫色和蓝色方框放大了切片中较复杂的脑区轮廓,用于明显可视化不同方法的配准结果。配准结果中,脑区结构与轮廓线的重合度越高,说明配准精度越高。从图5可以看出,传统迭代优化方法 SyN 和 Elastix 的配准精度有限,在多个区域出现明显偏移。深度学习方法在大尺度脑区结构上能够较好配准,但在结构复杂或边界模糊的区域中与真实轮廓的吻合度较低。相比之下,本文提出的方法借助分割伪标签对配准过程进行辅助,有效提高了关键脑区的结构配准精度。同时,引入的置信度模块能够在学习过程中动态聚焦于可靠性较高的区域,从而在复杂脑区边界处也能保持良好的配准效果。

4 结论

本文针对复杂三维大脑结构对高精度配准的需求,提出了一种结合分割一切模型与互学习机制的弱监督三维脑图像配准方法。通过微调分割一切模型并采用多视角融合策略,本文获得了具有多样性和结构一致性的伪分割掩码,从而有效解决了弱监督配准中分割标签难以获取的问题,并为互学习提供了丰富而可靠的特征先验。在此基础上,设计的互学习框架能够充分挖掘多视角伪标签的互补信息,并借助置信度模块动态过滤不可靠区域,从而显著提升了训练过程的稳定性与配准性能。实验结果表明,该方法在多组三维小鼠脑图像数据集上均取得了优于现有无监督及弱监督方法的性能,验证了利用分割一切模型辅助弱监督配准的有效性与潜力。本文方法在一定程度上依赖分割一切模型生成的伪分割掩码质量,但通过多视角融合、互学习机制及置信度模块的协同设计,并经伪分割退化实验验证,在分割

质量逐步下降的情形下仍能够保持较为平缓的性能退化趋势。未来工作将进一步探索更鲁棒的伪标签生成策略以及分

割与配准的联合优化框架,以提升方法在复杂成像条件和跨模态场景下的适应能力。

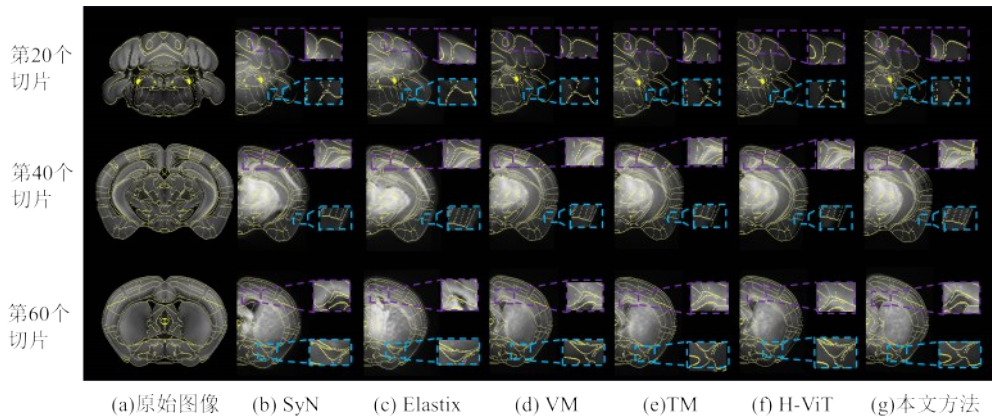


图5 不同方法配准后图像在水平面视角下的不同截面脑区轮廓图

Fig. 5 Brain region contours from different registration methods in horizontal cross-sections.

参考文献

- [1] WANG Quanxin, DING Songlin, LI Yang, et al. The Allen mouse brain common coordinate framework: a 3D reference atlas [J]. Cell, 2020, 181(4): 936-953.
- [2] PENG Hanchuan, XIE Peng, LIU Lijuan, et al. Morphological diversity of single neurons in molecularly defined cell types [J]. Nature, 2021, 598(7879): 174-181.
- [3] LI Zhongyu, SHANG Z, LIU J, et al. D-LMBmap: a fully automated deep-learning pipeline for whole-brain profiling of neural circuitry [J]. Nature Methods, 2023, 20(10): 1593-1604.
- [4] BALAKRISHNAN G, ZHAO A, SABUNCU M R, et al. Voxelmorph: a learning framework for deformable medical image registration [J]. IEEE transactions on medical imaging, 2019, 38(8): 1788-1800.
- [5] 余升林, 吴彤, 葛明锋, 等. 循环一致性的无监督可变形图像配准方法 [J]. 计算机辅助设计与图形学学报, 2023, 35(04): 516-524.
- YU Shenglin, WU Tong, GE Mingfeng, et al. Unsupervised deformable image registration method with cyclic consistency [J]. Journal of Computer-Aided Design & Computer Graphics, 2023, 35(4): 516-524.
- [6] CHEN Junyu, FREY E C, HE Y, et al. Transmorph: Transformer for unsupervised medical image registration [J]. Medical image analysis, 2022, 82: 102615.
- [7] GHAREMANI M, KHATERI M, JIAN B, et al. H-vit: A hierarchical vision transformer for deformable image registration [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern

- Recognition. 2024: 11513-11523.
- [8] JIA Xi, LU W, CHENG X, et al. Decoder-only image registration [J]. IEEE Transactions on Medical Imaging, 2025.
 - [9] KIM B, KIM D H, PARK S H, et al. CycleMorph: cycle consistent unsupervised deformable image registration [J]. Medical image analysis, 2021, 71: 102036.
 - [10] ZHU Yongpei, LU S. Swin-voxelmorph: A symmetric unsupervised learning model for deformable medical image registration using swin transformer [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer Nature Switzerland, 2022: 78-87.
 - [11] 赵海峰,姚丽莎,罗斌.改进的人工鱼群算法和Powell法结合的医学图像配准[J].西安交通大学学报,2011,45(04):46-52.
ZHAO Haifeng, YAO Lisha, LUO Bin. Registration of multi-resolution medical images using a modified artificial fish-swarm algorithm combined with Powell's method [J]. Journal of Xi'an Jiaotong University, 2011, 45(4): 46-52.
 - [12] 高军,李学伟,张建,等.基于模板匹配的图像配准算法[J].西安交通大学学报,2007,(03):307-311.
GAO Jun, LIXuewei, ZHANGJian, et al. Image registration algorithm based on template matching [J]. Journal of Xi'an Jiaotong University, 2007, 41(3): 307-311.
 - [13] 吴攀超,王宗义,林欣堂.采用局部差分模型描述的彩色图像配准技术[J].西安交通大学学报,2011,45(10):30-37.
WU Panchao, WANGZongyi, LINXintang. Color image registration technique based on local differential model [J]. Journal of Xi'an Jiaotong University, 2011, 45(10): 30-37.
 - [14] YEW Z J, LEE G H. 3dfeat-net: Weakly supervised local 3d features for point cloud registration [C]//Proceedings of the European conference on computer vision (ECCV). 2018: 607-623.
 - [15] 顾峰.基于分割辅助和级联的医学图像配准[D].西安电子科技大学,2021.DOI: 10.27389/d.cnki.gxadu.2021.002996.
GU Feng. Medical image registration based on segmentation assistance and cascade [D]. Xidian University, 2021. DOI: 10.27389/d.cnki.gxadu. 2021.002996.
 - [16] HU Yipeng, MODAT M, GIBSON E, et al. Label-driven weakly-supervised learning for multimodal deformable image registration [C]//2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE, 2018: 1070-1074.
 - [17] WU Menglin, HE X, LI F, et al. Weakly supervised volumetric prostate registration for MRI-TRUS image driven by signed distance map [J]. Computers in biology and medicine, 2023, 163: 107150.
 - [18] XU Zhenlin, NIETHAMMER M. DeepAtlas: Joint semi-supervised learning of image registration and segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer International Publishing, 2019: 420-429.
 - [19] KIRILLOV A, MINTUN E, RAVI N, et al. Segment anything [C]//Proceedings of the IEEE/CVF international conference on computer vision. 2023: 4015-4026.
 - [20] HUANG Shiqi, XU T, SHEN Z, et al. One registration is worth two segmentations [C]//International Conference on Medical Image Computing and Computer-Assisted

- Intervention. Cham: Springer Nature Switzerland, 2024: 665-675.
- [21] XU Hao, XUE T, LIU D, et al. Sat-morph: Unsupervised deformable medical image registration using vision foundation models with anatomically aware text prompt [C]//International Workshop on Foundation Models for General Medical AI. Cham: Springer Nature Switzerland, 2024: 71-80.
- [22] 冯林, 管慧娟, 孙焘, 滕弘飞. 医学图像非刚性配准研究进展 [J]. 北京生物医学工程, 2006, 25(4): 437-440.
- FENG Lin, GUAN Huijuan, SUNTao, TENG Hongfei. Research Progress of Non-Rigid Registration in Medical Images [J]. Beijing Biomedical Engineering, 2006, 25(4): 437-440.
- [23] 孟勋. 典型医学图像配准方法及优化策略的研究进展 [J]. 中国医疗设备, 2015, 30(8): 64-69.
- MENG Xun. Research on the Typical Medical Image Registration Method and Its Optimization Strategies [J]. China Medical Devices, 2015, 30(8): 64-69.
- [24] ROHR K, STIEHL H S, SPRENGEL R, et al. Landmark-based elastic registration using approximating thin-plate splines [J]. IEEE Transactions on medical imaging, 2002, 20(6): 526-534.
- [25] 靳峰, 冯大政. 利用空间序列描述子的快速准确的图像配准算法 [J]. 西安交通大学学报, 2014, 48(6): 6. DOI: 10.7652/xjtuxb201406004.
- JIN Feng, FENG Dazheng. Fast and accurate image registration algorithm using spatial sequence descriptor [J]. Journal of Xi'an Jiaotong University, 2014, 48(6): 6. DOI: 10.7652/xjtuxb201406004.
- [26] DETONE D, MALISIEWICZ T, RABINOVICH A. Deep image homography estimation [J]. arXiv preprint arXiv: 1606.03798, 2016.
- [27] ROHÉ M M, DATAR M, HEIMANN T, et al. SVF-Net: learning deformable image registration using shape matching [C]//International conference on medical image computing and computer-assisted intervention. Cham: Springer International Publishing, 2017: 266-274.
- [28] 马露凡, 罗凤, 严江鹏, 徐哲, 罗捷, 李秀. 深度医学图像配准研究进展: 迈向无监督学习 [J]. 中国图象图形学报, 2021, 26(9): 2037-2057.
- MA Lufan, LUOFeng, YAN Jiangpeng, XU Zhe, LUO Jie, LI Xiu. Deep-learning based medical image registration pathway: towards unsupervised learning [J]. Journal of Image and Graphics, 2021, 26(9): 2037-2057.
- [29] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]//International Conference on Medical image computing and computer-assisted intervention. Cham: Springer international publishing, 2015: 234-241.
- [30] JADERBERG M, SIONYAN K, ZISSERMAN A. Spatial transformer networks [J]. Advances in neural information processing systems, 2015, 28.
- [31] BLENDOWSKI M, HANSEN L, HEINRICH M P. Weakly-supervised learning of multi-modal features for regularised iterative descent in 3D image registration [J]. Medical image analysis, 2021, 67: 101822.
- [32] LI Yiwen, FU Y, GAYO I J M B, et al.

- Semi-weakly-supervised neural network training for medical image registration [J]. arXiv preprint arXiv:2402.10728, 2024.
- [33] ZHAO Ziheng, ZHANG Y, WU C, et al. One model to rule them all: Towards universal segmentation for medical images with text prompts [J]. arXiv preprint arXiv:2312.17183, 2023.
- [34] MA Jun, HE Y, LI F, et al. Segment anything in medical images [J]. Nature Communications, 2024, 15(1): 654.
- [35] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network [J]. arXiv preprint arXiv: 1503.02531, 2015.
- [36] ZHANG Ying, XIANG T, HOSPEDALES T M, et al. Deep mutual learning [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 4320-4328.
- [37] PANG Youwei, ZHAO X, ZHANG L, et al. Multi-scale interactive network for salient object detection [C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 9413-9422.
- [38] YOU Chenyu, ZHOU Y, ZHAO R, et al. Simcvd: Simple contrastive voxel-wise representation distillation for semi-supervised medical image segmentation [J]. IEEE Transactions on Medical Imaging, 2022, 41(9): 2228-2237.
- [39] KONG Lintong, YANG Jie. MDFlow: Unsupervised optical flow learning by reliable mutual knowledge distillation [J]. IEEE Transactions on circuits and systems for video technology, 2022, 33(2): 677-688.
- [40] LIU Jun, WANG W, SHEN N, et al. Mutualreg: mutual learning for unsupervised medical image registration [C]//ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2024: 1561-1565.
- [41] AVANTS B B, TUSTISON N J, SONG G, et al. A reproducible evaluation of ANTs similarity metric performance in brain image registration [J]. Neuroimage, 2011, 54(3): 2033-2044.
- [42] KLEIN S, STARING M, MURPHY K, et al. Elastix: a toolbox for intensity-based medical image registration [J]. IEEE transactions on medical imaging, 2009, 29(1): 196-205.