

基于超像素自监督的训练无注释的小样本医学图像分割

徐圆斐

摘要

本文提出一种新的医学图像的自监督 FSS 框架，用来消除训练期间对于注释的要求。并且生成基于超像素的伪标签来提供监督。将“自适应局部原型池模块”插入到原型网络，解决医学图像分割中常见的前景背景不平衡的问题。使用了三种不同的任务证明对医学图像的普适性，腹部器官的 CT 和 MRI 以及心脏的 MRI。

关键词：自监督学习，小样本分割，超像素，医学图像分割。

1 介绍

医学图像分割是大量的临床处理和医学成像研究的关键步骤，其中包括疾病的诊断和随访，治疗的规划以及人群研究。当在大量标记的数据上进行训练时，全监督的基于深度学习分割模型可以获得不错的结果。但是全监督的方法在医学图像方面往往不切实际，由于进行注释需要大量的临床专业知识，因为这耗费大量的成本和时间，缺乏足够多的专家来对数据进行注释以便于训练。图像采集程序的差异也进一步加剧了这一问题，比如跨医院、跨医疗设备，通常导致数据集只包含很少的手动标记的图像。此外，可分割的目标也不计其数（比如不同的解剖结构，不同类型的病变）。因此，模型训练之后很难将模型应用于新类的预测。

为了解决上述两个问题，小样本学习被提出，我们可以仅从几个标记的示例——支持集中提取新的类别来对未标记的示例进行预测——查询集。即，将小样本应用于医学图像中就可以仅使用几个标记的示例就可以有效地分割罕见的或者新颖的病变。

然而，训练现有的用于医学图像的小样本语义分割模型（FSS），在过去并没有取得太大的成功，这是因为大多数小样本语义分割方法都依赖于具有许多注释类的大型训练数据集，以避免过度拟合。为了绕过这一为满足的注释需求。我们建议通过自监督学习（一种无监督的技术，通过解决精心设计的任务学习可推广的图像表示），在未标记的图像上训练 FSS 模型 [26–33]。许多最先进的 FSS 网络架构面临的另一个挑战是在其学习表示中丢失空间变化类中的局部信息。由于在医学图像中普遍存在极端的前景背景不平衡，这一问题在医学图像上尤其严重。如图 1（b）所示，背景类别大且空间不均匀，而前景类别（紫色）小且均匀。在这种情况下，如果背景中不同局部区域（或者说，部分）的不同外观信息被不合理地平均化，则可能发生前景背景边界预测中的模糊。这种类内局部信息存在于许多最近的工作当中。为适应这一问题，我们鼓励网络通过提取每一个类的局部来保存类内局部信息。

为了打破训练数据稀缺的僵局并提高分割精度，我们提出了 SSL—ALPNet，这是一种用于医学成像的自我监督的少镜头语义分割框架。所提出的框架利用基于超像素的自监督学习（SSL），使用超像素消除对手动注释的需要，以及自适应局部原型池增强原型网络（ALPNet），通过在学习表示中保留局部信息来提高分割精度。如图 1（a）所示，为了确保通过自我监督学习的图像表示能够很好地推广到真实的语义类，我们使用超像素生成伪语义标签，超像素是语义对象的紧凑构建块 [35–37]。

此外，为了提高学习图像表示的辨别能力，我们将自监督任务制定为一个超像素，以防止分割。此外，为了增强支持和查询之间表示的不变性（这对于现实世界中的少镜头分割至关重要），我们通过在支持和查询间应用随机几何和强度变换来合成形状和强度的变体。在我们的实验中，我们观察到，通过纯粹使用 SSL 进行训练，我们的网络在相当大的幅度上优于使用手动注释类进行训练的网络。此外，如图 1 所示，为了提高分割精度，我们设计了自适应局部原型模块（ALP），用于在原型表示中保留每个类的局部信息。这是通过提取局部表示原型的集合来实现的，每个原型都聚焦于不同的区域。值得注意的是，原型的数量由网络基于每个类的空间大小自适应地分配。

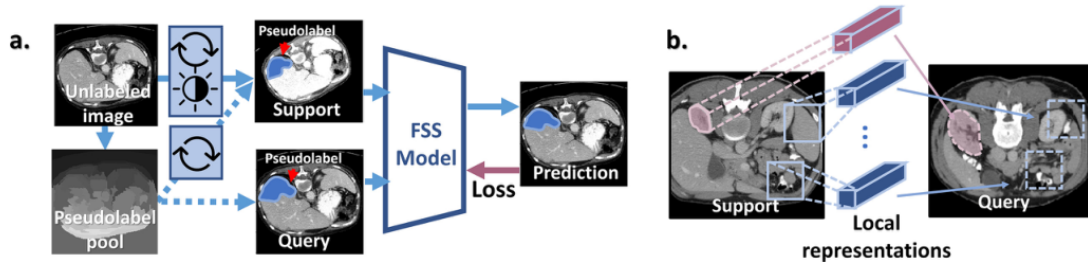


图 1: (a) 提出的基于超像素的自监督学习。对于每个未标记的图像，在超像素上生成伪标签。在训练期间的每次迭代中，随机选择一个伪标签和原始图像作为支持和查询的候选。然后，在支持和查询之间应用随机转换 (在蓝框中标记)。自我监督任务被设计为参考支持对查询上的伪标签进行分割，尽管在两者之间应用了转换。(b). 提出的 ALPNet 通过自适应提取大型背景类的多个局部表示 (蓝色部分) 解决了类不平衡引起的歧义问题。每一个只代表一个局部的背景区域。

本文的贡献总结如下：

提出了 SSL ALPNet，这是我们所知的第一项研究小样本医学图像分割的自监督学习的工作。它通过仅对未标记的图像进行训练而优于通常需要使用手动注释进行训练的同级 FSS 方法。

提出了自适应局部原型池，这是一种局部表示计算模块，可显著提高医学图像上最先进原型网络的性能。

首次评估了不同成像模式、分割类别和患者病理情况下的 FSS。既定的评估策略不仅突出了我们工作的广泛适用性，而且有助于未来在更现实的情况下评估 FSS 的工作。

2 相关工作

2.1 小样本语义分割

最近的工作首次将自监督学习引入到小样本图像分类中。然而，小样本分割通常更具挑战性：密集预测需要在像素级上执行。为了充分利用有限支持数据中的信息，目前流行的 FSS 方法大多直接将支持集作为引导信号注入网络，或者从支持集构造判别表示，参考段查询。开创性工作学习从支持中生成分类器权重；将权重生成扩展到多尺度。而不是直接使用支持条件分割查询融合他们的特征映射。最近的工作利用网络组件，如注意力模块和图网络，提高了分割精度，并使 FSS 具有粗级监督。利用基于学习的优化，将元学习与 FSS 结合起来。然而，几乎所有这些方法都假设有丰富的注释 (包括弱注释) 训练数据可用，这使得它们很难转化为医学成像中的分割场景。

FSS 的一个主流称为原型网络，重点是利用从支持中提取的语义类的表示原型。这些原型用于对查询进行基于相似性的预测，或优化查询的表示。最近，原型对齐网络 (PANet) 在自然图像上取得了最先进的性能。

这是简单地实现了一个通用的卷积网络和对齐正则化。相比之下，我们的工作重点是通过开发创新的训练策略和伪标签来利用无标签的医学图像进行训练。然而，由于 PANet 是最先进的技术之一，而且概念简单，所以我们采用这种方法作为我们的基线来强调我们的自我监督学习作为一种通用的训练策略。

在医学影像中，目前关于少镜头分割的工作大多集中在数据量较少的训练上。这些方法在应用到未知类之前通常仍然需要重新训练，因此它们不在我们的讨论范围之内。没有对未知类进行再训练，SE-Net 引入挤压和激发块。据我们所知，这是第一个专门为医学图像设计的 FSS 模型，我们在实验中与我们的方法进行了比较。

2.2 语义分割中的自监督学习

提出了一系列用于语义分割的自我监督任务。

这些工作主要集中在直观的手工监督任务，包括空间变换预测、图像绘制、补丁重新排序、图像着色、差异检测、运动插值等。类似的方法也应用于医学图像。然而，这些工作中的大多数仍然需要在初始化后进行第二阶段的微调，使用从自我监督中学习到的权重。此外，从手工任务中学习到的特征可能无法充分推广到语义分割，因为两个任务可能没有很强的相关性。相比之下，在我们的工作中，分割基于超像素的伪标签与分割真实对象直接相关。

这是因为超级像素是真实对象语义掩码的紧凑构建块。最近在医学影像方面的工作依赖于二阶优化。这些工作与我们的工作重点方法和任务上有所不同。

我们提出的 SSL 技术在方法论上与（或者说，与最近一些关于对比学习）具有相似的精神。这两种方法都通过有意地创建变量来鼓励图像表示的不变性。虽然他们关注的是视觉信息聚类，但我们关注的是实用但具有挑战性的少样本医学图像分割问题。

2.3 超像素分割

超像素是小而紧凑的图像片段，通常是分段平滑的。超像素是通过使用统计模型对低层次图像特征进行聚类生成的。这些模型包括高斯混合和图切。在这项工作中，我们采用了现成的、高效的、无监督的基于图切的算法。使用这些超像素进行训练直观地提高了网络对各种形状的未见类的泛化能力。

3 方法

首先介绍了少样本语义分割 (FSS) 的问题表述。

然后，介绍了 ALPNet 体系结构，重点介绍了自适应本地原型池和相应的推理过程。最后，我们介绍了所提出的 SSL 技术下的整体端到端训练目标。

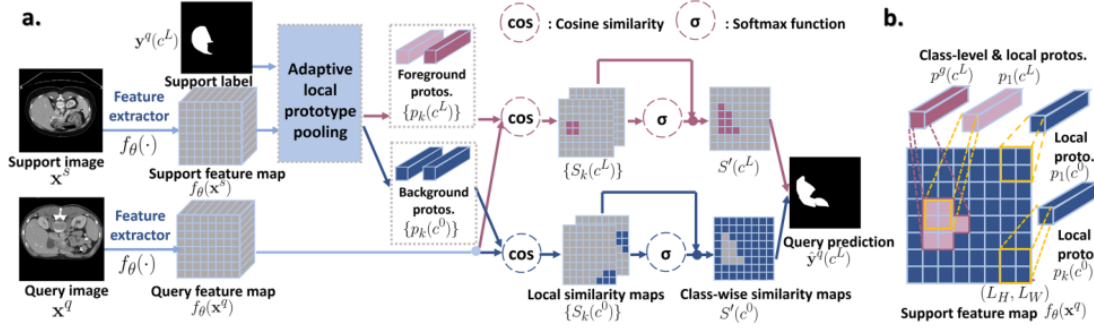


图 2: (a). 网络的工作流程: 特征提取器 f_θ 以支持图像和查询图像为输入, 生成支持的特征映射 $f_\theta(x^s)$ 和查询的特征映射 $f_\theta(x^q)$ 。该自适应局部原型池模块以支持特征映射和支持标签为输入, 得到表示原型 $p^k(c^j)$ 的集合。这些原型被用作与查询特征映射 $f_\theta(x^q)$ 比较的参考。通过这些比较生成的相似图被融合在一起形成最终的分割。该图演示了一个单向分割设置, 其中 c^L 是前景类, c^0 是背景类。(b). 自适应局部原型池化模块说明: 通过池化窗口 (橙色框) 内的空间平均支持特征映射计算局部原型; 类级原型在整个支持标签 (紫色区域) 下平均。

3.1 问题公式化

少样本分割的目的就是获得一个可以分割未知语义类的模型, 通过在推理过程中只用少量的该未知类别的标记, 而不需要重新训练这个模型。我的理解就是先通过别的类别训练好这个分割模型——这个过程中可能用到大量的别的类别的标签样本, 然后如果我们要分割一个未知语义类别, 只需要少量的这个未知类别的标记的标签样本就可以让模型很好分割, 而不需要重新训练出一个新的专门用于分割这个未知类的模型。详细解释如下: 训练集 D_{tr} 包含一些用于训练的语义类 C_{tr} : 如, 肝脏、脾脏、脊柱, 的图像。测试集 D_{te} 包含一些测试未知类 C_{te} : 如, 心脏和肾脏。 $C_{tr} \cap C_{te} = \emptyset$ 这个任务就是在 D_{tr} 上训练一个分割模型, 然后就可以在 D_{te} 上分割 C_{te} 的语义类, D_{te} 上只需要很少量的有标签的 C_{te} 的图像。我们不用为了分割 C_{te} 而重新训练一个模型。结合例子来说就是: 任务就是在训练集上训练出一个分割模型 (用被标记的肝脏脾脏脊柱), 这个模型可以分割训练集 (心脏肾脏) 只需要少量的被标记的心脏肾脏的样本。

3.2 网络结构

概述: 我们的网络由: (a) 一个通用的特征提取网络 $f_\theta(\cdot) : \mathcal{X} \rightarrow \mathcal{E}$ 由 θ 参数化, 其中 \mathcal{E} 是表示空间 (即, 特征空间), 对其进行分割。(b) 提出的自适应局部原型池化模块 (ALP) $g(\cdot, \cdot) : \mathcal{E} \times \mathcal{Y} \rightarrow \mathcal{E}$ 由于从支持特征和标签中提取表示原型; (c) 与基于相似的的分类器 $\text{sim}(\cdot) : \mathcal{E} \times \mathcal{E} \rightarrow \mathcal{Y}$ 用于通过比较原型和查询特征进行分段。

如图 2 所示, 在推理中, 特征提取器网络 $f_\theta(\cdot)$ 为 ALP 提供特征映射, 通过将 \mathbf{x}_l^{s1} 和 \mathbf{x}^q 映射到特征空间 \mathcal{E} , 得到特征映射 $\{(f_\theta(\mathbf{x}^q), f_\theta(\mathbf{x}_l^s))\} \in \mathcal{E}$ 。ALP 以每个 $(f_\theta(\mathbf{x}_l^s), \mathbf{y}_l^s(c^j))$ 座位输入计算语义类 c^3 和背景 c^0 。这些原型稍后将用作每个类的引用, 用于分割查询图像。所有的 c^j 原型形成一个原型集合 $\mathcal{P} = \{p_k(c^j)\}, j = 0, 1, 2, \dots, N$ 其中 k 为原型指数 $k \geq 1$ 分类器 $\text{sim}(\cdot, \cdot)$ 使用这个原型集合来对查询图像进行分割, 即 $\hat{\mathbf{y}}^q = \text{sim}(\mathcal{P}, f_\theta(\mathbf{x}^q))$ 。这是通过首先测量每个 $p_k(c^j)$'s 和查询特征映射 $f_\theta(\mathbf{x}^q)$ 之间的相似性来实现的, 然后将这些相似性融合到一起。

自适应局部原型池: 与之前工作 [17, 18, 34], 不同的是, 在语义掩码下, 类内局部信息被不合理地在空间上平均, 我们建议通过引入自适应局部原型池模块 (ALP) 来保留原型中的局部信息。在 ALP 中, 每个本地原型只在覆盖在支持上的本地池窗口内计算, 并且只表示感兴趣的对象的一部分。

具体来说，我们在每个 $f_\theta(\mathbf{x}_l^s) \in \mathbb{R}^{D \times H \times W}$ 上使用池化窗口大小 (L_H, L_W) 。其中， (H, W) 是空间大小 D 通道深度。注意 (L_H, L_W) 决定了每个局部原型在表示空间 \mathcal{E} 中计算的空间范围。

$$p_{l,mn}(c) = \text{avgpool}(f_\theta(\mathbf{x}_l^s))(m, n) = \frac{1}{L_H L_W} \sum_h \sum_w f_\theta(\mathbf{x}_l^s)(h, w),$$

$$\text{where } mL_H \leq h < (m+1)L_H, nL_W \leq w < (n+1)L_W.$$

为了确定 $p_{l,mn}(c)$ 的每个 c 类，我们将前景类 $p_{l,mn}(c)$ 的二值掩码平均到相同大小 $(\frac{H}{L_H}, \frac{W}{L_W})$ 。设 $y_{l,mn}^a$ 为位置 (m, n) 的 c 类池化后的值 $\mathbf{y}_l^s(c^{\hat{j}})$ ：

$$c = \begin{cases} c^0 & y_{l,mn}^a < T \\ c^{\hat{j}} & y_{l,mn}^a \geq T \end{cases} \quad \text{where } y_{l,mn}^a = \text{avgpool}(\mathbf{y}_l^s(c^{\hat{j}}))(m, n).$$

T 为前景的下界阈值，经验设置为 0.95。

为了确保为小于池窗口 (L_H, L_W) 的对象生成至少一个原型，我们还计算了一个类级原型 $p_l^g(c^{\hat{j}})$ 用来做平均池化掩码：

$$p_l^g(c^{\hat{j}}) = \frac{\sum_{h,w} \mathbf{y}_l^s(c^{\hat{j}})(h, w) f_\theta(\mathbf{x}_l^s)(h, w)}{\sum_{h,w} \mathbf{y}_l^s(c^{\hat{j}})(h, w)}$$

端到端的训练：该网络是端到端的训练每次迭代以 $(\mathcal{S}_i, \mathcal{Q}_i)$ 作为输入。利用交叉熵损失函数： $\mathcal{L}_{\text{seg}}^i$ 每次迭代的分割损失函数如下：

$$\mathcal{L}_{\text{seg}}^i(\theta; \mathcal{S}_i, \mathcal{Q}_i) = -\frac{1}{HW} \sum_h \sum_w \sum_{j \in \{0,p\}} \mathcal{T}_g(\mathbf{y}_i^r(c^j))(h, w) \log(\hat{\mathbf{y}}_i^r(c^j)(h, w))$$

其中 $\mathbf{y}_i^r(c^p)$ 表示的是查询位掩码的预测，还使用了原型对齐正则化；将预测反过来作为支持集，按理说也是可以被正确分割的。

$$\mathcal{L}_{\text{reg}}^i(\theta; \mathcal{S}_i', \mathcal{S}_i) = -\frac{1}{HW} \sum_h \sum_w \sum_{j \in \{0,p\}} \mathbf{y}_i^r(c^j)(h, w) \log(\bar{\mathbf{y}}_i^r(c^j)(h, w))$$

将二者综合起来得到：

$$\mathcal{L}^i(\theta; \mathcal{S}_i, \mathcal{Q}_i) = \mathcal{L}_{\text{seg}}^i + \lambda \mathcal{L}_{\text{reg}}^i$$

在这里还可以继续进行改进可以用到 Dice 损失函数，将交叉熵损失函数和 Dice 损失函数进行结合。使其既有交叉熵的稳定性又有 Dice 损失不受等级平衡影响的特性。引入 Dice 损失，算法不受前景像素与背景像素比例的影响，对预测结果和全局值之间的相似性进行评估。

$$D = \frac{2 \sum_h \sum_w p_{ij} g_{ij}}{\sum_h \sum_w p_{ij} + \sum_h \sum_w g_{ij}}$$

$$L^i(\theta; \mathcal{S}_i; \mathcal{Q}_i) = (1 - \alpha) (L_{\text{seg}}^i + \lambda L_{\text{seg}}^i) - \alpha \log(D)$$

4 复现细节

作者所给出的代码是不完整的，只有他所提出的自适应局部原型池化模块的关键部分，由于文章是基于 PANet 写的，于是我将 PANet 的代码进行了改进，引入本文的创新方法，进行了实验验证，

然后又在这改进的基础上引入了我的想法，并再次进行了相关实验验证。

4.1 实验数据集

为了证明我们提出的方法在不同成像模式、分割类别和受试者健康状况下的一般适用性，我们在三种场景下进行了评估:CT 和 MRI 的腹部器官分割 (Abd-CT 和 Abd-MRI) 和 MRI 的心脏分割 (Card-MRI)。

- **Abd-CT** 来自 MICCAI 2015 多图集腹部标记挑战赛。它包含 30 个三维腹部 CT 扫描。这是一个临床数据集，包含各种病理和扫描之间强度分布变化的患者。

- **Abd-MRI** 来自 ISBI 2019 联合健康腹部器官分割挑战赛。它包含 20 个 3D T2-SPiR MRI 扫描。

- **Card-MRI** 来自 MICCAI 2019 多序列心脏 MRI 分割挑战赛，共有 35 个的临床三维心脏 MRI 扫描。

4.2 实验设置

为了统一实验设置，将所有图像重新格式化为 2D 轴向 (AbdCT 和 Abd-MRI) 或 2D 短轴 (Card-MRI) 切片，并将大小调整为 256×256 像素。预先占有的应用遵循一般惯例。每个 2D 切片在信道维度上重复三次，以适应网络。

为了比较不同形状、位置和纹理的特殊病理、不均匀的 Abd-CT 和完全健康、均匀的 Abd-MRI 的结果，我们构建了一个包含左肾、右肾、脾和肝的共享标签集; 对于 Card-MRI，标签集包含左室血池 (LV-BP)、左室心肌 (LV-MYO) 和右室 (RV)。在所有实验中，我都进行了五次交叉验证。

4.3 实现细节

本网络基于 PANet 来实现，用 Pytorch。为了在特征途中获得高分辨率 $f_{\theta}(\cdot)$ 被分配为此案城的全卷积 ResNet101，它在 MS-COCO 上进行部分预训练来获得更高的分割性能。以 $3 \times 256 \times 256$ 的图像作输入，生成 $256 \times 32 \times 32$ 的特征图。原型局部池化窗口设置为 4×4 来进行训练。学习率为 0.001，每 1000 次迭代的步进衰减率为 0.98。在单个 Nvidia RTX 2080Ti GPU 上进行自我监督训练需要约 3 小时，占用 2.8 gb 内存。

4.4 评估

为了衡量预测和地面真相之间的重叠，使用常用于医学图像分割研究的 Dice 评分 (0 - 100,0: 不匹配;100: 完全匹配)。

为了评估对未见测试类的泛化能力，前人建立的医学图像标准少镜头分割实验设置 (设置 1) 之外，测试类可能作为训练数据的背景出现，我引入了一个设置 2。在设置 2 中，我们通过从训练数据集中删除包含测试类的任何图像来强制测试类 (甚至是未标记的) 完全不可见。

因此，标签根据监管的设置和类型进行了不同的划分。在设置 1 中，使用 SSL 进行训练时，不需要进行标签划分。当使用带注释的图像进行训练时，我们每次都上一节课进行测试，其余的课进行训练。为了观察习得的表征是否能编码像左和右这样的空间概念，我们故意将左肾和右肾分组。

为了模拟临床实践中标记数据的稀缺性，所有实验都在单向 1-shot 设置下进行。

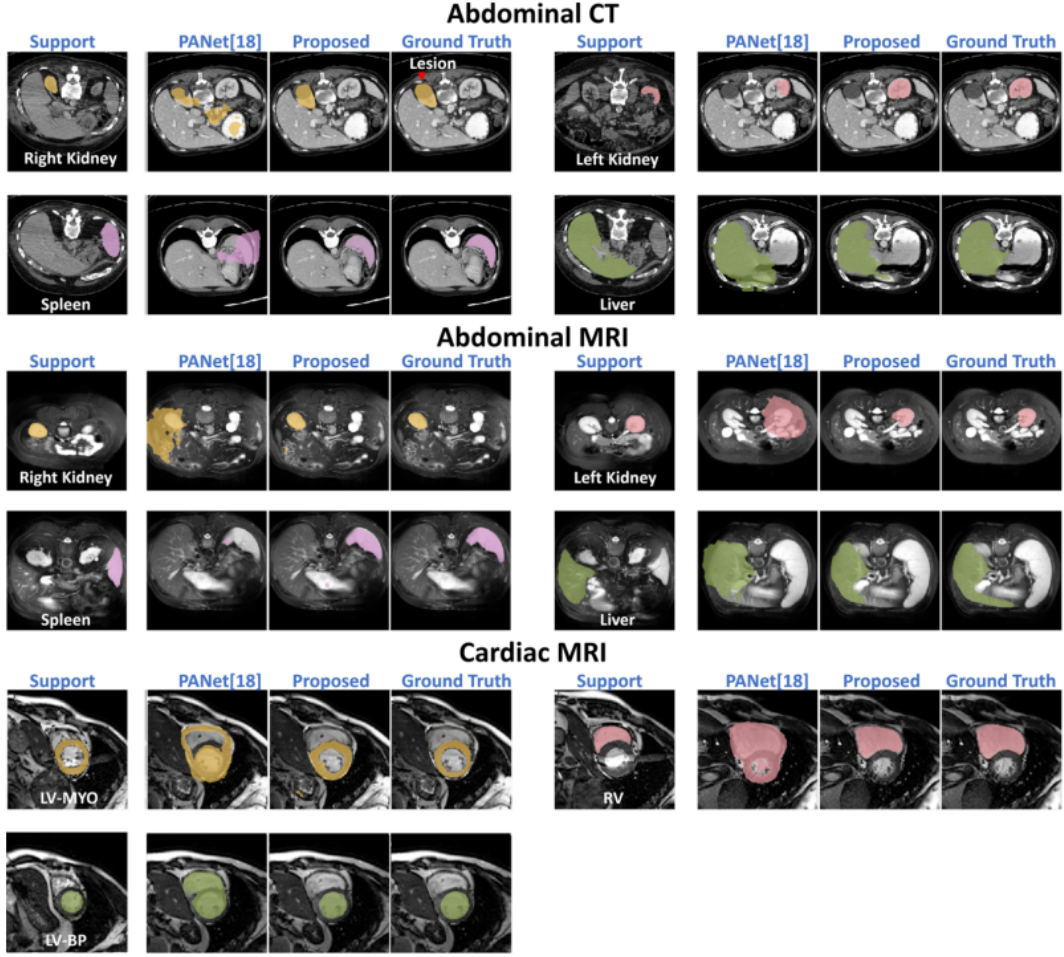


图 3: 复现论文时在三种成像模式下还原作者的实验产生的定性结果

4.5 创新点

我将在这一部分来展开讲解主要创新点。

基于相似度的分割: 基于相似度的分类 $\text{sim}(\cdot, \cdot)$ 是利用 \mathcal{P} 中的局部图像信息对查询进行密集预测的, 它首先将每个原型与查询中对应的局部区域进行匹配, 然后将局部相似度融合在一起。具体来说 $\text{sim}(\cdot, \cdot)$ 首先以查询特征映射 $f_\theta(\mathbf{x}^q)$ 和原型集合 $\mathcal{P} = \{p_k(c^j)\}$ 为输入。分别计算 $f_\theta(\mathbf{x}^q)$ 和所有 $p_k(c^j)$ 之间的局部相似度映射 $S_k(c^j)$ 。对应于 $S_k(c^j)$ 的空间位置 (h, w) 的每一项 $f_\theta(\mathbf{x}^q)$ 表示为:

$$S_k(c^j)(h, w) = \alpha p_k(c^j) \odot f_\theta(\mathbf{x}^q)(h, w),$$

\odot 表示余弦相似性, 是有界的 $a \odot b = \frac{\langle a, b \rangle}{\|a\|_2 \|b\|_2}$, $a, b \in \mathbb{R}^{D \times 1 \times 1}$, α 是一个乘数有助于梯度在训练中反向传播。为了获得关于每个类 c^j 作为整体的相似度映射 (未归一化) 将每个类的局部相似度映射 $S_k(c^j)$ 分别融合为类间相似度 $S'(c^j)$, 这是通过一个 softmax 函数完成的:

$$S'(c^j)(h, w) = \sum_k S_k(c^j)(h, w) \text{softmax}_k [S_k(c^j)(h, w)].$$

$\text{softmax} [S_k(c^j)(h, w)]$ 是指将所有 $S_k(c^j)(h, w)$ 沿信道维度进行叠加, 然后沿信道计算 softmax 函数。为了获得最终的密集预测, 最后, 按类别的相似性被归一化为概率:

$$\hat{\mathbf{y}}^q(h, w) = \text{softmax} [S'(c^j)(h, w)].$$

为了获得准确和稳健的结果, 基于相似性的分类器非常需要两个属性。对于每个类, 表示应该聚

类，以便在相似度度量下具有鉴别性；同时，这些表示在图像之间应该是不变的，以确保预测的鲁棒性。

这两个特性是基于超像素的自监督学习 (SSL) 所提倡的。由于真实语义类的注释不可用，SSL 利用伪标签在超像素级上强制集群。这通过基于余弦相似度的分类器反向传播分割损失来实现。超像素级的聚类属性可以转移到真实的语义类中，因为一个语义掩码通常由几个超像素组成。此外，为了鼓励表示对图像之间的形状和强度差异保持不变，我在支持和查询之间执行几何和强度转换。这是因为形状和强度是医学图像变化的最大来源。

所提出的 SSL 框架包括两个阶段：离线伪标签生成和在线训练。整个流程如图 4 所示。

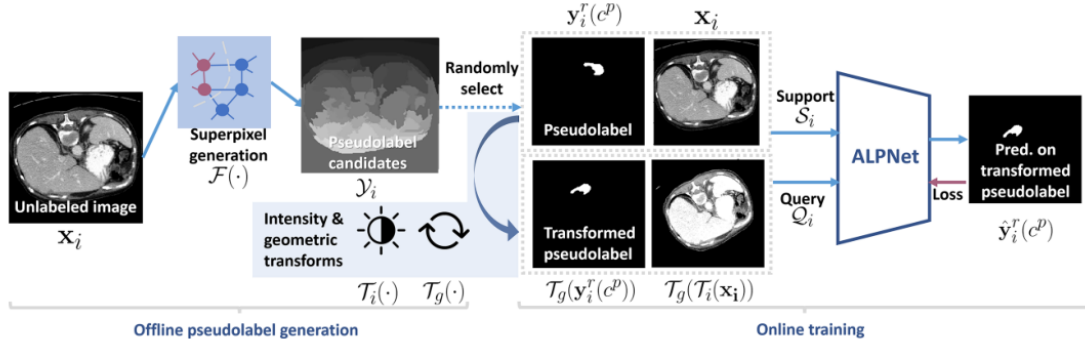


图 4: 基于超像素的自监督学习技术的工作流程

伪标签生成: 为了获得伪标签的候选对象, 对每幅图像 \mathbf{x}_i 生成超像素集合 $\mathcal{Y}_i = \mathcal{F}(\mathbf{x}_i)$ 生成超像素集合. 线上训练: 对于每一个 episode, 一幅图像 \mathbf{x}_i 和一个随机选择的超像素 $\mathbf{y}_i^r(c^p) \in \mathcal{Y}_i$ 用于形成支持集 $\mathcal{S}_i = \{(\mathbf{x}_i, \mathbf{y}_i^r(c^p))\}$. 这里的 $\mathbf{y}_i^r(c^p)$ 表示带索引 $r = 1, 2, 3, \dots, |\mathcal{Y}_i|$ 的二值掩码, c^p 表示伪标签类 (对应的背景掩码 $\mathbf{y}_i^r(c^0)$ 用 $1 - \mathbf{y}_i^r(c^p)$ 表示). 同时查询集 $\mathcal{Q}_i = \{(\mathcal{T}_g(\mathcal{T}_i(\mathbf{x}_i))), \mathcal{T}_g(\mathbf{y}_i^r(c^p))\}$ 通过对支持集应用随即几何变换和强度转换来构建. 这样每一个 $(\mathcal{S}_i, \mathcal{Q}_i)$ 就成了 1-way 1-shot 的分割问题. 其中, $\mathcal{T}_g(\cdot)$ 包括仿射变换和弹性变换, $\mathcal{T}_i(\cdot)$ 是伽马变换。

图 5 为改进后的不同角度的实验效果图。可见对于脑部 MRI 的分割效果是不错的。

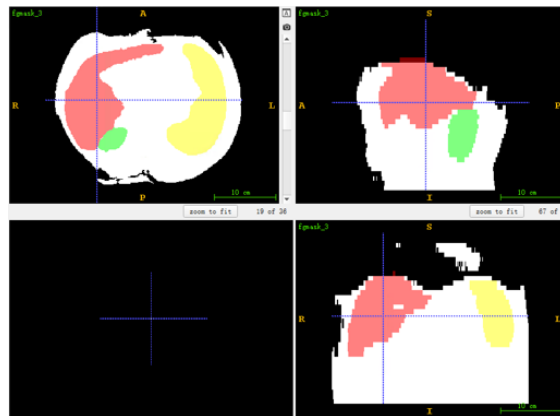


图 5: 改进后对脑部 MRI 不同角度的分割效果图

下表为同样在腹部 CT 和腹部 MRI 中实验的 Dice 分数，对比不同的方法可以看出，改进后的 Dice 分数要更高。

表 1: 腹部图像的实验结果 (以 Dice 分数表示)

Method	Manual Anno.?	Abdominal-CT					Abdominal-MRI				
		Kidneys		Spleen	Liver	Mean	Kidneys		Spleen	Liver	Mean
		LK	RK				LK	RK			
SE-Net [43]	✓	24.42	12.51	43.66	35.42	29.00	45.78	47.96	47.30	29.02	42.51
Vanilla PANet [18]	✓	20.67	21.19	36.04	49.55	31.86	30.99	32.19	40.58	50.40	38.53
ALPNet	✓	29.12	31.32	41.00	65.07	41.63	44.73	48.42	49.61	62.35	51.28
SSL-PANet	×	56.52	50.42	55.72	60.86	57.88	58.83	60.81	61.32	71.73	63.17
SSL-ALPNet	×	72.36	71.81	70.96	78.29	73.35	81.92	85.18	72.18	76.10	78.84
Zhou et al. [75]	Ful. Sup	95.3	92.0	96.8	97.4	95.4	-				
Isensee et al. [76]	Ful. Sup	-					-				94.6

5 Conclusion

在这项工作中，我们提出了一个自监督的医学图像分割框架。所提出的方法在不需要任何人工标记进行训练的情况下，成功地优于目前最先进的方法。在我的有关基于相似度分割部分进行改进以及对于最终损失函数的改进下，实验效果有所提高。此外，该模型在实验中对未见语义类具有较强的泛化能力。提出的基于超像素的自监督技术为图像表示学习提供了一种有效的方法，为未来的半监督和无监督图像分割工作开辟了新的可能。

参考文献

- [1] Pham, D.L., Xu, C., Prince, J.L.: Current methods in medical image segmentation. Annu. Rev. Biomed. Eng. 2(1), 315–337 (2000)
- [2] Sharma, N., Aggarwal, L.M.: Automated medical image segmentation techniques. J. Med. Phys. Assoc. Med. Phys. India 35(1), 3 (2010)
- [3] Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D., Alzheimer’ s Disease Neu- roimaging, et al.: Multimodal classification of Alzheimer’ s disease and mild cogni- tive impairment. Neuroimage 55(3), 856–867 (2011)
- [4] Ei Naqa, I., et al.: Concurrent multimodality image segmentation by active con- tours for radio- therapy treatment planning a. Med. Phys. 34(12), 4738–4749 (2007)
- [5] Zaidi, H., El Naqa, I.: Pet-guided delineation of radiation therapy treatment vol- umes: a survey of image segmentation techniques. Eur. J. Nucl. Med. Mol. Imaging 37(11), 2165–2187 (2010)
- [6] De Leeuw, F., et al.: Prevalence of cerebral white matter lesions in elderly people: a popula- tion based magnetic resonance imaging study. the rotterdam scan study. J. Neurol. Neurosurg. Psychiatry 70(1), 9–14 (2001)
- [7] Petersen, S.E., et al.: Imaging in population science: cardiovascular magnetic reso- nance in 100,000 participants of UK biobank-rationale, challenges and approaches. J. Cardiovasc. Magn. Reson. 15(1), 46 (2013)
- [8] Snell, J., Swersky, K., Zemel, R.: Prototypical networks for few-shot learning. In: Advances in Neural Information Processing Systems, pp. 4077–4087 (2017)

- [9] Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P.H., Hospedales, T.M.: Learning to compare: relation network for few-shot learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1199–1208 (2018)
- [10] Rakelly, K., Shelhamer, E., Darrell, T., Efros, A., Levine, S.: Conditional networks for few-shot semantic segmentation. In: 6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, 30 April–3 May, 2018, Workshop Track Proceedings (2018)