

文章编号:1007-2780(2026)04-0565-13

基于原型对齐的持续图像分割

高勇占^{1,2,3}, 黄凌霄^{1,2,3*}, 姚新波^{1,2,3}, 周开元^{1,2,3}, 徐海喆^{1,2,3}

- (1. 宁夏大学 信息工程学院, 宁夏 银川 750021;
2. 宁夏大学“东数西算”人工智能与信息安全重点实验室, 宁夏 银川 750021;
3. 宁夏大数据与人工智能省部共建协同创新中心, 宁夏 银川 750021)

摘要:针对持续图像分割中旧类知识灾难性遗忘、新类目标表征不足以及背景语义随任务推进动态偏移等问题, 本文提出一种面向类增量场景的原型对齐持续图像分割方法(Prototype-Aligned Incremental Segmentation, PAIS)。该方法从对象级表征稳定性的角度出发, 对持续学习过程中解码端对象查询的初始化、演化与历史语义补偿进行统一建模。该方法包含 3 个关键模块: 原型感知初始化(Prototype-Aware Initialization, PAI)通过从当前特征图中选择语义响应显著的位置生成对象查询, 增强模型对新类别的建模能力; 时序一致性学习(Temporal Consistency Learning, TCL)约束相邻学习阶段关键语义位置及其响应分布的一致性, 缓解特征漂移导致的对象性退化; 类别感知记忆(Category-Aware Memory, CAM)维护紧凑的类别级记忆表征, 在不依赖大量图像重放的条件下提升旧类别保持能力。在 ADE20K 数据集上的持续全景分割与持续语义分割实验结果表明, PAIS 在多种任务设置、数据划分及输入序列下均优于 ECLIPSE、BalConpas 和 CoMFormer 等代表性方法; 在覆盖 150 类的标注数据集上实现了 6.75% 的 PQ 指标提升, 同时将存储需求降低约 10 倍。实验结果表明, PAIS 能够在新类学习能力、旧类保持能力与存储效率之间取得较好的平衡。

关键词:持续图像分割; 原型对齐; 时序一致性学习; 类别感知记忆; 全景分割

中图分类号: TP391.4 **文献标识码:** A **doi:** 10.37188/CJLCD.2026-0050 **CSTR:** 32172.14.CJLCD.2026-0050

Continual image segmentation based on prototype alignment

GAO Yongzhan^{1,2,3}, HUANG Lingxiao^{1,2,3*}, YAO Xinbo^{1,2,3}, ZHOU Kaiyuan^{1,2,3}, XU Haizhe^{1,2,3}

- (1. School of Information Engineering, Ningxia University, Yinchuan 750021, China;
2. Key Laboratory of Artificial Intelligence and Information Security for “East Data West Computing” in Ningxia, Yinchuan 750021, China;
3. Collaborative Innovation Center for Big Data and Artificial Intelligence, Ministry of Education and Ningxia Province, Yinchuan 750021, China)

Abstract: Continual image segmentation suffers from catastrophic forgetting of previously learned classes, insufficient representation of newly introduced classes, and dynamic background semantic shifts in continual image segmentation. To address these issues, this paper proposes PAIS (Prototype-Aligned Incremental Segmentation), a prototype-aligned incremental segmentation method for continual image

收稿日期: 2026-03-08; 修订日期: 2026-03-15.

基金项目: 国家自然科学基金(No.12462027); 宁夏自然科学基金(No.2025AAC030205)

Supported by National Natural Science Foundation of China (No.12462027); Natural Science Foundation of Ningxia (No.2025AAC030205)

*通信联系人, E-mail: huanglx@nxu.edu.cn

segmentation. From the perspective of object-level representation stability, PAIS jointly models the initialization, evolution, and historical semantic compensation of decoder object queries during continual learning. Specifically, prototype-aware initialization (PAI) selects semantically salient positions from the current feature map to initialize object queries, thereby improving the modeling capability for new classes. Temporal consistency learning (TCL) constrains the consistency of key semantic positions and their response distributions across adjacent learning stages, which alleviates objectness degradation caused by feature drift. In addition, category-aware memory (CAM) maintains compact class-level memory representations to enhance old-class retention without relying on large-scale image replay. Experiments on continual panoptic segmentation and continual semantic segmentation on the ADE20K dataset demonstrate that the proposed method consistently outperforms representative methods such as ECLIPSE, BalConpas, and CoMFormer under multiple task settings, data splits, and input sequences. Under the setting covering 150 classes, the proposed method achieves a clear improvement in PQ while reducing storage requirements by approximately 10 times. These results show that PAIS provides a favorable balance among new-class adaptation, old-class retention, and memory efficiency.

Key words: continual image segmentation; prototype alignment; temporal consistency learning; category-aware memory; panoptic segmentation

1 引言

近年来,随着深度学习的快速发展,语义分割、实例分割及全景分割等密集预测任务取得了显著进展^[1-3]。与此同时,小样本语义分割和局部区域图像分割等相关研究也不断发展,在有限标注、复杂背景和细粒度目标条件下,构建稳定、鲁棒的分割模型已成为图像分割领域的重要研究方向。然而,绝大多数现有方法仍依赖离线监督学习范式,即假定训练阶段可一次性访问所有类别数据,并在固定类别集合上完成充分优化。这一假设在实际场景中往往难以成立:真实视觉系统通常需随时间顺序接收新任务与新类别。若每次引入新数据均需联合重训全部历史数据,将导致极高的存储、标注与计算开销^[4]。因此,如何在持续引入新类别的同时有效保留已有知识,成为持续视觉学习(Continual Visual Learning)的核心难题。

持续学习为上述问题提供了解决方案框架,其核心目标是在顺序任务流中持续更新模型,同时最大程度缓解灾难性遗忘^[5-8]。与图像分类相比,持续图像分割面临更为严峻的结构性挑战。一方面,像素级密集预测要求模型对旧类别的语义边界与区域结构保持高度稳定,因而对表征漂移^[9]极为敏感,且新类别常与旧类别共享局部纹

理、空间上下文或背景信息,在标注有限条件下易导致新类表征不足或类别边界模糊^[10-12]。另一方面,随着任务推进,背景区域的潜在语义分布持续演变,原有背景概念不再稳定,进一步加剧类别混淆与表征漂移。总体来看,持续图像分割不仅需要具备对新类别的快速适应能力,还需在跨阶段学习过程中尽可能维持旧类别语义结构和判别边界的稳定性,这使其较一般静态图像分割任务更具挑战性。

现有针对持续图像分割的研究主要沿样本重放、知识蒸馏、参数正则化以及结构扩展方向展开^[13-15]。样本重放通过存储历史图像或特征有效缓解遗忘,但随任务数增加,内存开销快速累积;知识蒸馏与参数正则化通过约束新旧模型输出或参数差异减少对重放的依赖,但在像素级高度复杂的语义分布下,单纯的响应对齐或权重约束^[16]难以稳定维持对象级结构信息;结构扩展通过引入任务特定分支或模块缓解新旧冲突,但往往导致模型规模膨胀及跨任务特征共享不足^[17]。总体而言,现有的方法在稳定性与可塑性之间仍难以取得理想平衡:约束过强抑制新类学习,约束过弱则无法有效抑制旧类遗忘。

近年来,基于 Transformer 的分割框架快速发展,其中对象查询(Object Query)已成为统一建模区域语义与实例结构的核心中介表征^[18]。

该机制不仅负责目标区域的定位与特征聚合^[19-21],还在很大程度上决定了模型对类别语义及对象边界的解析精度。在持续学习场景下,对象查询的初始化策略、跨阶段演化轨迹以及与历史语义的对齐机制直接影响新类适应与旧类保持的均衡^[22]。若初始化缺乏针对当前任务的有效语义先验,新类响应往往较弱;若缺少跨阶段关键位置与响应分布的稳定约束,查询表征将随特征漂移逐步退化;若仅依赖图像级重放保存历史知识,又将重新引入较大存储负担。

为此,本文针对持续图像分割任务,提出一种面向类增量场景的原型对齐持续图像分割(Prototype-Aligned Incremental Segmentation, PAIS)方法。该方法从对象级表征的跨阶段稳定性出发,对解码器中对象查询的初始化、演化与历史语义注入进行统一建模,有效缓解新旧表征冲突。具体而言,针对随机初始化导致的新类查询不稳定与响应不足问题,设计原型感知初始化(Prototype-Aware Initialization, PAI)模块,通过从深层特征图中自适应选取语义响应显著的位置生成初始对象查询,增强模型对新类别的感知与建模能力;针对增量训练中对象表征易发生漂移的问题,引入时序一致性学习(Temporal Consistency Learning, TCL)模块,通过跨阶段约束保持关键语义位置及其响应分布在相邻学习阶段间的一致性,从而抑制表征退化;针对传统样本重放带来的存储负担,设计类别感知记忆(Category-Aware Memory, CAM)模块,以轻量

级类别原型记忆替代图像级重放,并在解码过程中动态注入历史语义先验,实现对旧类别的有效保持而无需大规模样本存储。

本文的主要贡献如下:

(1)提出一种面向类增量场景的原型对齐持续图像分割方法(PAIS),从对象级表征稳定性的角度统一建模新类别学习与旧类别保持过程,为缓解灾难性遗忘和背景语义漂移提供了新的思路。

(2)设计了原型感知初始化、时序一致性学习和类别感知记忆3个关键模块,分别从对象查询初始化、跨阶段表征约束和轻量化历史记忆3个方面提升模型对新老类别的建模能力。

(3)在ADE20K数据集上的持续全景分割与持续语义分割任务中进行了系统实验。结果表明,PAIS在多种任务设置下均优于现有代表性方法,并在分割性能与存储开销之间取得了良好平衡。

2 模型方法

2.1 模型总体架构

针对持续图像分割任务中旧类知识易遗忘、新类适应能力不足以及跨阶段表征不稳定等问题,本文提出一种由原型感知初始化(Prototype-Aware Initialization, PAI)、时序一致性学习(Temporal Consistency Learning, TCL)和类别感知记忆(Category-Aware Memory, CAM)组成的持续图像分割方法,如图1所示。该方法在整体上遵循“查询初始化-阶段约束-记忆补偿”的设

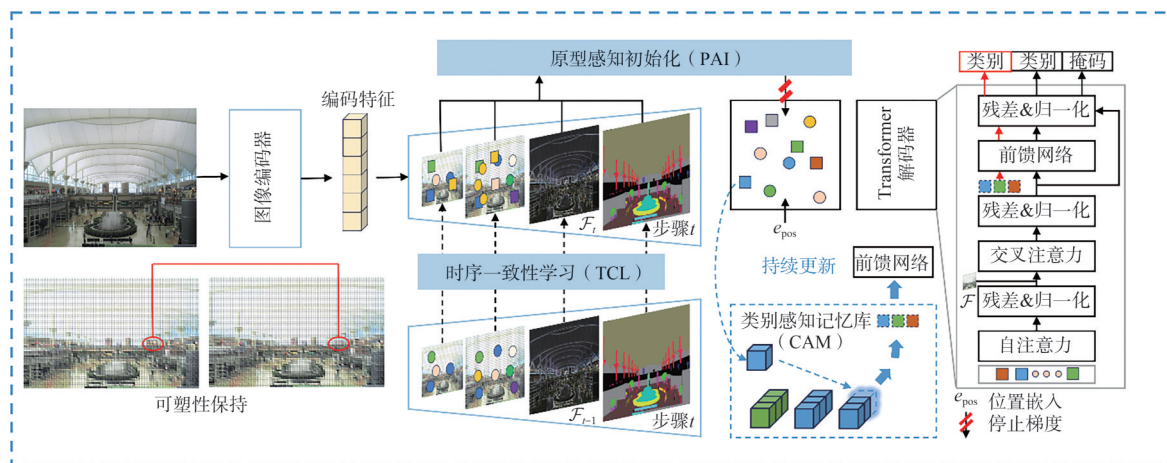


图1 PAIS模型的整体结构

Fig. 1 Overall structure of PAIS

计思路:首先利用PAI模块从当前特征图中选择语义响应显著的位置,对对象查询进行原型感知初始化,以增强模型对新类别的建模能力;随后通过TCL模块约束不同学习阶段间关键语义位置及响应分布的一致性,缓解特征漂移带来的对象性退化;最后引入CAM模块维护类别级紧凑记忆表征,在不依赖大量图像重放的条件下提升旧类别保持能力。

本文方法通过在查询初始化、表征约束和历史知识保持3个层面协同建模,实现了对持续图像分割任务中可塑性与稳定性矛盾的统一优化。

2.2 原型感知初始化

在基于查询的分割框架中,对象查询的初始化方式直接影响后续解码过程的对象建模能力。若仍采用固定可学习查询,则在持续学习阶段中难以随类别分布变化及时调整,容易造成新类别表征不足。为此,本文提出原型感知初始化(PAI)模块,通过类别原型引导,从当前特征图中选择具有较强语义判别性的特征位置来初始化对象查询。

在训练的第 t 步,对每个类别 $i \in C^t$ 维护一组可训练原型 P^i ,其中 $\{P^i | P^i \in \mathbb{R}^D\}$ 。通过将前一阶段的原型集合 P^{t-1} 与当前阶段的类别原型拼接,得到当前阶段的原型集合 P^t :为兼顾旧类先验与新类语义,当前阶段使用的原型集合由前一阶段原型集合与当前阶段新增类别原型拼接得到:

$$P^t = \text{concat}(P^{t-1}, \{P^i | i \in C^t\}), \quad (1)$$

其中, P^{t-1} 表示截至第 $t-1$ 阶段已保留的历史类别原型集合。

在得到原型集合后,对特征图 F^t 上每个空间位置 $(l, h, w) \in \Omega$ 计算其与所有类别原型的相似度,并以最大响应作为该位置的语义显著性评分。函数 $\text{topK}(X, Y)$ 返回 X 中最大的 Y 个值的索引, N 为对象查询 Q_N 的数量。由此选择与原型相似度最高的 N 个特征点以初始化 Q_N :

$$\mathcal{I}^t = \text{topK}\left(\left\{\max_{i \in C^t} S(F_{(l,h,w)}^i, P^i) \mid \forall (l, h, w) \in \Omega\right\}, N\right), \quad (2)$$

其中: $S(\cdot)$ 表示特征与原型之间的相似度函数,本文采用点积形式进行计算; $\mathcal{I}^t$ 表示第 t 阶段选中的关键位置索引集合。

依据索引集合 $\mathcal{I}^t$,从当前特征图中提取相应特征以构造对象查询:

$$Q_N = \mathcal{E}_m^{n=t} = \{F_i^{m=t} | i \in \mathcal{I}^{n=t}\}, \quad (3)$$

其中, $\mathcal{E}_m^n$ 表示在第 n 阶段依据所选索引从第 m 个特征表征中抽取的特征集合。由式(3)可知,PAI并非使用与输入无关的静态查询,而是利用当前图像中的高响应语义位置动态生成查询表征,从而在新类到来时保留更强的任务适应能力。

此外,为使原型能够更准确地刻画类别中心,可对原型集合 P^t 施加辅助分类监督;同时,为避免初始化过程对底层特征产生过强扰动,在实现时可对由 F^t 到 Q_N 的选择过程采用梯度隔离策略。这样既能保持原型驱动的语义选择能力,又有助于增强跨阶段对象表征的稳定性。

2.3 时序一致性学习

虽然PAI能够在每个阶段自适应地生成对象查询,但持续学习过程中模型参数与特征分布会不断更新,同一图像或相似语义区域在不同阶段可能对应不同的高响应位置,从而导致查询与对象区域之间的对应关系发生漂移。若缺乏有效约束,这种特征漂移将导致旧类高响应区域逐渐“扁平化”或被错分为背景,进而削弱模型对旧类的对象性建模能力。为此,本文引入时序一致性学习(TCL)模块,对跨阶段关键语义位置的响应分布进行约束。

设在第 $t-1$ 阶段中,PAI模块基于类别响应图选中的关键位置(即响应图的“峰值”)索引集合为 $\mathcal{I}^{t-1}$ 。在当前第 t 阶段训练时,可直接从当前特征图 F^t 中抽取这些位置对应的特征表征,记为:

$$\mathcal{E}_i^{t-1} = \{F_i^t | i \in \mathcal{I}^{t-1}\}. \quad (4)$$

相应地, $\mathcal{E}_{i-1}^{t-1}$ 表示前一阶段在同一索引集合上的原始特征表征。为减弱语义漂移,本文以历史原型集合 P^{t-1} 为参照,对上述两组特征分别计算类别响应分布,并通过散度度量其差异,定义时序一致性损失为:

$$L_{\text{tel}} = \frac{1}{|\mathcal{I}^{t-1}|} \sum_{i=1}^{|\mathcal{I}^{t-1}|} S(\mathcal{E}_{i-1}^{t-1}, P^{t-1}) \log \frac{S(\mathcal{E}_i^{t-1}, P^{t-1})}{S(\mathcal{E}_{i-1}^{t-1}, P^{t-1})}. \quad (5)$$

式(5)在形式上延续了阶段间一致性约束的思想,但在逻辑上与PAI的响应图保持了高度一致性解耦协同。具体而言,TCL的约束目标正

是PAI模块在上一阶段生成的响应图“峰值”集合。PAI负责在当前任务中“定位”潜在对象,TCL则确保这些定位点在时间维度上形成“锚点”。当模型学习新类别引发全局特征漂移时,旧类核心区域的特征往往会发生位移或响应下降,即对象性退化。TCL通过上述散度损失,强制要求同一组关键位置在相邻阶段中相对于历史原型的语义响应分布尽可能一致。

与直接对全局特征施加蒸馏不同,TCL仅作用于PAI筛选出的局部关键语义位置。这种分布对齐而非单纯的绝对数值对齐,确保了即使底层特征发生微调以适应新类,旧类核心区域作为“对象”的拓扑结构与相对响应模式依然不变。换言之,TCL强迫模型在更新参数时,不能抹除原先已被识别为对象的关键表征结构,从而在局部特征锚定的作用下,有效缓解了由特征漂移引起的旧类边界模糊和对象性退化问题,因此约束更加集中,既降低了额外计算开销,也减少了背景噪声对一致性学习的不利影响。该设计使模型能够在保持新类学习能力的同时,维持跨阶段语义选择的稳定性。其中, $\mathcal{E}_{t-1}^{-1}$ 为前一阶段选出的特征点集合, $\mathcal{E}_t^{-1}$ 为当前阶段对应的特征表征, $|\mathcal{I}^{t-1}|$ 为前一阶段选中的特征点数量。

2.4 类别感知记忆

仅依赖阶段间一致性约束仍不足以完全抑制旧类别遗忘,尤其是在当前阶段训练样本中不再包含旧类别标注时,分类分支容易出现类别边界模糊和旧类预测退化。针对这一问题,我们提出类别感知记忆(CAM)模块,通过维护紧凑的类别级记忆表征,辅助模型在后续阶段持续保留旧类判别信息。

如图2所示,在进入解码阶段前,模型首先构建多层级的特征金字塔以捕捉多尺度上下文。给定初始的高分辨率特征映射 f_1 ,我们通过空间下采样(SDS)模块降低其空间维度并与深层特征融合得到 f_2 。对于后续更深层的特征,则依次采用分辨率与维度下采样(RDS)模块进行特征提取与对齐,最终形成多尺度特征集合 $\{f_1, f_2, \dots, f_n\}$ 。这些多尺度特征将作为键(Key)和值(Value),与右侧解码器中的对象查询进行交互。对应于当前输入图像的对象查询 Q_n ,依次通过自注意力(Self-Attention)与交叉注意力(Cross-Attention)

模块,以实现当前场景中目标区域的精确定位与特征聚合。从类别感知记忆库中采样的历史记忆查询 Q_j 通过跳跃连接直接输入至前馈网络(FFN)。这确保了 Q_j 能够直接在深层特征空间中提供历史类别的判别性补偿,从而在有效维持旧类决策边界的同时,最大程度地保护模型对当前新类别的像素级解析能力。

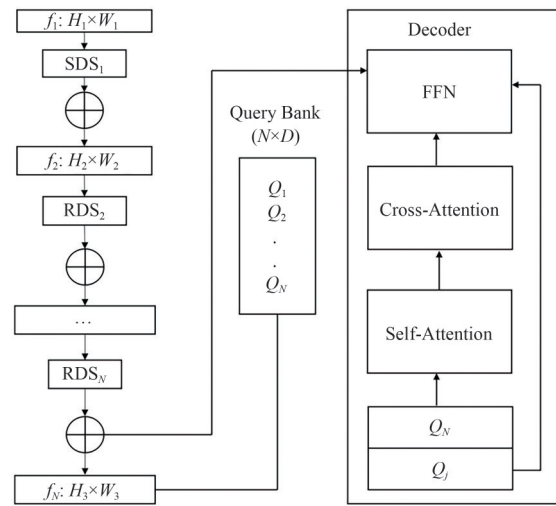


图2 类别感知记忆解码流程

Fig. 2 Category-aware memory decoding process

具体而言,在每一阶段训练完成后,对解码器最终输出中与真实标签成功匹配的查询特征进行收集,并按照类别存入对应的记忆队列中,解码流程如图2所示。设 Z_n 为解码器输出的查询表征, y 为真实标签,经二分匹配后可得到查询与标签的对应关系:

$$\begin{cases} \mathcal{I}_b = \text{Bipartite}(Z_n, y) \\ B_{\text{cam}} \leftarrow \text{Enqueue}_{\forall i=(i_q, i_y) \in \mathcal{I}_b} \left(Q_N(i_q), b_{\hat{y}^{(i_y)}} \right) \end{cases}, \quad (6)$$

其中: B_{cam} 表示类别感知记忆库, $b_{\hat{y}^{(i_y)}}$ 表示类别 $\hat{y}^{(i_y)}$ 对应的记忆队列, $\mathcal{I}_b$ 为匹配得到的索引对集合, i_q 和 i_y 分别表示查询索引和标签索引。通过上述方式,模型能够以类别为单位积累具有代表性的历史查询特征,而无需保存大量原始图像样本。

考虑到不同阶段旧类在当前数据中的隐式出现频率并不均衡,本文进一步依据前一阶段模型在当前数据集 D' 上的输出统计构造类别采样权重,以提升稀有旧类被访问的概率。参照原有符号体系,可将类别权重表示为:

$$w = \left\{ \left(\frac{\sum_{i=1}^m \sigma_i}{\sigma_j} \right)^{\frac{1}{2}} \right\}_{j=1}^m, \quad (7)$$

其中: σ_j 表示第 j 个旧类别在当前阶段伪统计中的样本数量, $m = |C^{1:t-1}|$ 表示旧类别总数。该权重能够在采样时适当提高长尾旧类别的重放概率, 从而缓解类别不平衡带来的记忆偏置问题。

在每次迭代中, 根据权重 w 从记忆库 B_{cam} 中采样 j 个类别记忆特征, 记为 $Q_j = \{\text{cam}_1, \text{cam}_2, \dots, \text{cam}_j\}$, 并与当前阶段由 PAI 生成的对象查询 Q_N 进行拼接, 得到扩展查询集合 $Q_{N+j} = \{q_1, q_2, \dots, q_N, \text{cam}_1, \text{cam}_2, \dots, \text{cam}_j\}$ 。考虑到记忆查询所对应的对象通常并不直接存在于当前输入图像中, 若将其与当前图像特征进行完全相同的交互, 可能对当前阶段的新类学习造成干扰。如图 2 所示, 在解码器中, 我们为记忆查询设计了跳跃注意力 (Skip Attention) 策略。由于记忆查询所代表的对象在当前图像中不存在, 为防止其自注意力 (Self-Attention, SA) 和交叉注意力 (Cross-Attention, CA) 过程中干扰 Q_N , 记忆查询绕过注意力层, 直接作用于前馈网络 (FFN) 层:

$$Q'_{N+j} = \text{FFN} \left(\text{concat} \left[\text{CA} \left(\text{SA} \left(Q_N + e_{\text{pos}}, F \right) \right), Q_j \right] \right), \quad (8)$$

其中: $\text{SA}(\cdot)$ 、 $\text{CA}(\cdot)$ 和 $\text{FFN}(\cdot)$ 分别表示自注意力、交叉注意力和前馈网络, e_{pos} 为位置编码。式 (8) 表明, 当前图像对应的对象查询 Q_N 仍参与标准解码流程, 而从记忆库中采样得到的类别记忆查询 Q_j 主要在分类相关表征层面提供补偿。基于此, CAM 重点作用于类别预测分支, 并通过类别监督损失维持旧类判别边界。

2.5 损失函数

综合上述 3 个模块, 本文在第 t 阶段的总体优化目标由分割主损失、时序一致性损失和记忆分类损失共同构成, 即

$$L = L_{\text{seg}} + \lambda_1 L_{\text{tcl}} + \lambda_2 L_{\text{cam}}, \quad (9)$$

其中: L_{seg} 表示当前阶段分割主任务损失, 可由类别损失、掩码损失及 Dice 损失等组成; L_{tcl} 为式 (5) 所定义的时序一致性损失; L_{cam} 表示基于类别感知记忆的辅助分类损失; λ_1 和 λ_2 为平衡系数。

通过式 (9) 的联合优化, PAI 负责提升新类别的查询表达能力, TCL 用于稳定跨阶段对象性表征, CAM 则补偿旧类别的判别信息缺失。三者相互配合, 使模型在持续图像分割场景下能够同时兼顾新类学习能力与旧类知识保持能力。超参数 λ_1 与 λ_2 协同权衡模型的稳定性与可塑性。 λ_1 通过时序一致性损失 L_{tcl} 约束特征漂移, 在防止旧类语义退化的同时保留新类适应空间; λ_2 通过类别感知记忆损失 L_{cam} 正则化决策边界, 平衡新旧类别的梯度贡献以缓解预测偏差。为确定最优取值, 本文在 ADE20K 验证集上对 $\{0.1, 0.3, 0.5, 0.8, 1.0\}$ 空间进行网格搜索, 最终确定 $\lambda_1 = 0.5, \lambda_2 = 0.3$ 。

3 实验结果和分析

3.1 实验设置

为验证所提方法在持续图像分割任务中的有效性, 本文在 ADE20K 数据集上开展实验, 并分别对持续全景分割和持续语义分割任务进行评估。ADE20K 是场景理解领域常用的标准数据集, 包含丰富的室内外场景, 共标注 150 个语义类别, 其中包括 100 个 thing 类和 50 个 stuff 类, 能够较好地满足持续学习场景下类别逐步扩展的实验需求。

在评价指标方面, 持续全景分割任务采用全景质量 (Panoptic Quality, PQ) 指标作为主要评价指标, 持续语义分割任务采用平均交并比 (mean Intersection over Union, mIoU) 作为评价指标。PQ 指标综合考虑了目标识别结果与分割掩膜质量, 能够较全面地反映模型的整体分割性能, 其定义为:

$$\text{PQ} = \frac{\sum_{(p,g) \in \text{TP}} \text{IoU}(p,g)}{|\text{TP}| + \left(\frac{1}{2}\right)|\text{FP}| + \left(\frac{1}{2}\right)|\text{FN}|}, \quad (10)$$

其中: TP 表示预测结果与真实标注之间成功匹配的目标对集合, FP 表示未与任何真实目标匹配的预测目标数量, FN 表示未被正确匹配的真实目标数量, $\text{IoU}(p, g)$ 表示预测目标 p 与真实目标 g 之间的交并比。通常, 当 $\text{IoU}(p, g) > 0.5$ 时, 认为预测目标与真实目标匹配成功。

进一步地, PQ 可分解为分割质量 (SQ) 与识

别质量(RQ)两部分,即:

$$PQ = SQ \times RQ, \quad (11)$$

其中,SQ与RQ计算方式如下:

$$SQ = \frac{\sum_{(p,g) \in TP} IoU(p,g)}{|TP|}, \quad (12)$$

$$RQ = \frac{|TP|}{|TP| + \left(\frac{1}{2}\right)|FP| + \left(\frac{1}{2}\right)|FN|}, \quad (13)$$

其中:SQ反映匹配目标在区域分割上的平均精度,RQ反映模型对目标的识别与匹配能力。因此,PQ能够从识别和分割两个层面对持续全景分割模型的性能进行综合评价。为全面反映模型在持续学习过程中的综合表现,本文在每个增量阶段结束后,分别统计基础类别、新增类别以及全部类别上的性能。同时,引入联合训练(Joint Training)^[23-25]结果作为性能上界,用于衡量持续学习模型与理想充分监督条件之间的差距。

在实验实现上,持续全景分割实验采用预训练ResNet-50作为主干网络,持续语义分割实验采用预训练ResNet-101作为主干网络。输入图像分辨率分别设置为 640×640 和 512×512 。模型按照类增量方式进行优化:首先在基础类别集合 C^1 上训练,随后依次引入新增类别 $C^2, C^3, \dots, C^T$ 。为保证对比公平,优化器、学习率策略、批量大小和数据增强方式均与主流对比方法保持一致。

在实验设置方面,本文采用持续图像分割研

究中常用的类增量任务划分A-B设置,其中A表示基础阶段包含的类别数,B表示后续每个阶段新增的类别数。对于持续语义分割的类增量实验,采用100-5、100-10和100-50三种设置;对于持续全景分割的类增量实验,除100-5、100-10和100-50外,还进一步测试50-10、50-20和50-50任务。该设置既能够考察模型在长序列小步长增量场景下的抗遗忘能力,也能够评价其在大步长类别扩展场景中的快速适应能力。

3.2 结果分析

表1~表3给出了本文方法与现有代表性方法在ADE20K持续全景分割与持续语义分割任务上的实验结果,其中:加粗数字表示除Joint外各持续学习方法中的最优结果,下划线数字表示次优结果;-表示原文献未报告该结果;Joint表示联合训练上界,不参与最优和次优比较。

对比方法包括基于微调的FT、基于蒸馏或正则化的方法MiB、PLOP、CoMFormer,以及引入查询调优、重放或记忆机制的方法ECLIPSE、BalConpas和CoMasTRe^[26-28]等。从整体结果来看,本文方法在多种任务划分下均取得了较优的综合表现,尤其在新增类别性能和全部类别平均性能方面优势较为明显。这表明所提出的PAI、TCL和CAM三个模块能够有效协调持续学习中的“新类适应”与“旧类保持”两方面需求^[29-31]。

对于持续全景分割任务,在100-5设定下,由于增量阶段较多,模型需要频繁应对类别分布变化,因此更容易出现累积遗忘现象。实验结果表明,本文方法在新增类别上的PQ明显优于对比

表1 在ADE20K数据集连续全视分割结果

Tab. 1 Continual panoptic segmentation results on ADE20K dataset

Method	100-5 (11 tasks)				100-10 (6 tasks)				100-50 (2 tasks)			
	1-100	101-150	all	avg	1-100	101-150	all	avg	1-100	101-150	all	avg
FT	0.0	2.2	0.7	4.7	0.0	4.8	1.6	8.9	0.0	32.4	10.8	26.8
MiB ^[25]	2.3	0.0	1.5	13.4	6.8	0.2	4.6	19.1	23.3	14.9	20.5	31.7
PLOP ^[26]	31.1	11.9	24.7	31.3	37.7	23.3	32.9	37.8	42.4	23.7	36.2	39.5
SSUL ^[27]	30.2	7.9	22.8	27.9	31.6	11.9	25.0	30.3	35.9	18.1	30.0	33.8
CoMFormer	34.4	15.9	28.2	34.0	36.0	17.1	29.7	35.3	41.1	27.7	36.7	38.8
BalConpas	36.1	<u>20.3</u>	30.8	35.8	40.7	<u>22.8</u>	<u>34.7</u>	38.8	<u>42.8</u>	<u>25.7</u>	<u>37.1</u>	40.0
ECLIPSE	<u>41.1</u>	16.6	<u>32.9</u>	—	<u>41.4</u>	18.8	33.9	—	41.7	23.5	35.6	—
Our PAIS	42.7	22.4	35.9	39.3	42.6	30.7	38.5	40.9	44.8	31.4	40.8	42.9
joint	43.6	34.2	40.4	—	43.6	34.2	40.4	—	43.6	34.2	40.4	—

表 2 PQ 在 ADE20K 数据集上的持续全景式分割结果

Tab. 2 Continual panoptic segmentation results on ADE20K dataset in PQ

Method	50-10 (11 tasks)			50-20 (6 tasks)			50-50 (3 tasks)		
	1-50	1-50	all	1-50	1-50	all	1-50	1-50	all
FT	0.0	1.7	1.1	0.0	4.4	2.9	0.0	12.0	8.1
MiB	34.9	7.7	16.8	38.8	10.9	20.2	42.4	15.5	24.4
PLOP	39.9	15.0	23.3	43.9	16.2	25.4	45.8	18.7	27.7
CoMFormer	38.5	15.6	23.2	42.7	17.2	25.7	45.0	19.3	27.9
ECLIPSE	<u>45.9</u>	17.3	26.8	46.4	19.6	28.6	46.0	20.7	29.2
BalConpas	44.6	<u>24.8</u>	31.4	<u>49.2</u>	<u>28.2</u>	35.2	<u>51.2</u>	<u>26.5</u>	34.7
Our PAIS	49.4	30.6	36.7	51.9	32.5	38.9	52.6	31.5	38.7
joint	51.1	35.1	40.4	51.1	35.1	40.4	51.1	35.1	40.4

表 3 在 ADE20K 数据集上通过 mIoU 测量的连续语义分割结果

Tab. 3 Continual semantic segmentation results on the ADE20K dataset, measured by mIoU

Model	100-5 (11 tasks)				100-10 (6 tasks)				100-50 (2 tasks)			
	1-100	101-150	all	avg	1-100	101-150	all	avg	1-100	101-150	all	avg
FT	0.0	0.3	0.1	5.6	0.0	0.1	0.0	9.1	0.0	3.2	1.1	26.3
MiB	36.0	5.7	26.0	—	31.8	14.1	25.9	—	37.9	27.9	34.6	—
PLOP	39.1	7.8	28.8	35.3	40.5	14.1	31.6	36.6	41.9	14.9	32.9	37.4
SSUL	42.9	<u>17.8</u>	34.6	—	42.9	17.7	34.5	—	42.8	17.5	34.4	—
EW ^[28]	41.4	13.4	32.1	—	41.5	16.3	33.2	—	41.2	21.3	34.6	—
CoMFormer	39.5	13.6	30.9	36.5	40.6	15.6	32.3	37.4	39.5	26.2	38.4	41.2
ECLIPSE	<u>43.3</u>	16.3	<u>34.2</u>	—	43.4	17.4	34.6	—	45.0	21.7	37.1	—
BalConpas	42.1	17.2	33.8	41.3	<u>47.3</u>	<u>24.2</u>	<u>38.6</u>	43.6	<u>49.9</u>	<u>30.1</u>	<u>43.3</u>	47.4
CoMasTRe	40.8	15.8	32.6	38.6	42.3	18.4	34.4	38.4	45.7	26.0	39.2	41.6
Our PAIS	47.5	23.6	39.4	47.9	49.8	28.1	42.9	50.1	55.8	36.7	49.5	52.8
Joint	57.1	39.1	51.2	—	57.1	39.1	51.2	—	57.1	39.1	51.2	—

方法,同时基础类别性能下降较小,说明 PAI 能够通过原型感知初始化增强新类别查询表达,而 TCL 则有助于稳定跨阶段对象表征,缓解旧类对象性退化问题。在 100-10 和 100-50 设定下,随着每阶段新增类别数增加,本文方法仍保持较强竞争力,表明其不仅适用于长序列小步长增量场景,也能够适应单阶段大规模类别扩展场景。对于持续语义分割任务,本文方法在 100-5、100-10 和 100-50 三种设置下同样取得了较优的 mIoU 结果,尤其在基础类别和全部类别指标上表现稳定。相较于仅依赖蒸馏或参数约束的方法,本文方法能够在学习新类别时更有效地保留旧类别的判别边界。这一结果说明,TCL 对关键语义位置施加的时序一致性约束有助于减轻持续训练中的像素级语义漂移,而 CAM 所提供的类别级紧凑记忆进一步增强了模型对旧类语义的保持

能力。

综合持续全景分割与持续语义分割的实验结果可以看出,本文方法在不同任务形式和不同增量设置下均具有较好的稳定性和泛化能力。这表明围绕“查询初始化-阶段一致性-类别记忆”构建的联合机制能够更有效地服务于持续图像分割任务。

图 3 展现了本文模型在极端空间遮挡与长尾类别组合下的局限性。在面对结构极度细碎且空间布局高度重叠的场景时,解码端生成的对象查询难以精确锚定各个细粒度目标的边界。这导致预测结果出现了明显的语义崩塌现象。这一失败案例表明,在持续学习过程中,当新任务中引入具有极端长宽比或高度镂空拓扑结构的新类别时,单纯的特征响应对齐仍不足以完全维持像素级的锐利边界。



图3 分割失败情况实例
Fig. 3 Examples of segmentation failures

3.3 消融实验与存储开销分析

为进一步验证各组成模块的有效性,本文在 ADE20K 数据集上开展消融实验。实验以基础分割框架为基线,依次加入 PAI、TCL 和 CAM 3 个模块,相关结果如表 4 所示。当仅采用基线模型时,模型在持续训练过程中表现出较明显的旧类遗忘和新类适应不足问题,基础类别与全部类别性能均处于较低水平。加入 PAI 后,模型在新增类别上的性能获得明显提升,说明原型感知初始化能够从当前特征图中选取更具判别性的

语义位置生成对象查询,从而增强模型对新类别的建模能力;与此同时,基础类别性能也有所改善,表明更合理的查询初始化能够提升解码阶段的稳定性。在此基础上进一步引入 TCL 后,基础类别指标和总体性能继续提高,这说明对关键语义位置施加时序一致性约束可以有效减轻跨阶段特征漂移问题,尤其有助于稳定旧类别的语义响应。与传统全局蒸馏方式相比,TCL 仅对 PAI 筛选出的关键位置进行约束,因此更加聚焦,能够以较小代价获得更好的阶段稳定性。当引入 CAM 后,模型在全部类别上的性能进一步提升,尤其在旧类别保持方面表现更为突出。这说明类别感知记忆能够为分类分支提供持续的旧类判别信息补偿,从而降低当前阶段训练对历史类别造成的覆盖效应。完整模型在基础类别、新增类别和全部类别上均取得最佳结果,证明三者之间具有较好的互补性,其中 PAI 主要提升新类学习能力,TCL 主要增强跨阶段稳定性,CAM 则主要缓解旧类遗忘问题。

表 4 所提模块的消融实验结果

Tab. 4 Ablation study of the proposed modules

PSD	PAI	TCL	CAM	Panoptic100-5			Semantic 100-5		
				1-100	101-150	all	1-100	101-150	all
✓				31.6	21.3	28.2	15.6	8.5	13.2
✓	✓			30.7	22.3	27.9	37.4	16.7	30.5
✓	✓	✓		35.7	24.0	31.8	43.2	17.0	34.5
✓	✓		✓	35.1	23.3	31.2	42.5	19.5	34.8
✓	✓	✓	✓	42.1	21.9	35.4	46.7	22.8	38.7

除性能比较外,本文还分析了不同记忆方式下的存储开销。如表 5 所示,与直接保存原始图像或标注样本的重放策略相比,CAM 仅保存类别级紧凑记忆特征,无需存储原始图像及像素级标注信息,因此在较低存储成本下即可获得稳定的旧类保持收益。实验结果表明,在相近甚至更低的存储预算下,CAM 相比传统图像重放方法能够取得更优或至少相当的持续分割性能,说明类别级记忆表征在效率和效果之间达到了较好的平衡。图 4(a)表明,CAM 在更小内存下即可达到更高的性能。CAM 在 5.9 MB 时达到 42.1 (Base)与 35.4(All),而图像重放在约 11.8 MB 时为 38.5(Base)与 33.7(All),这表明 CAM 不需要更大的回放存储来获得更好结果。图 4(b)进

一步以 0 MB(无重放)作为统一基线,比较旧类保持增益。在各内存点上,CAM 的增益整体更高,这表明 CAM 在低重放预算下对旧类保持更

表 5 不同重放方式及其存储开销比较

Tab. 5 Comparison of different replay types and storage requirements

Type	Samples	Memory/MB	base	all
Image	0 (×20)	0.0	35.7	31.8
	75 (×20)	3.4	38.9	33.4
	150 (×20)	6.1	38.9	34.0
	300 (×20)	11.8	38.5	33.7
Category-Aware Memory (CAM)	0 (×150)	0.0	35.7	31.8
	20 (×150)	1.5	40.6	34.6
	40 (×150)	3.0	40.4	34.1
	80 (×150)	5.9	42.1	35.4

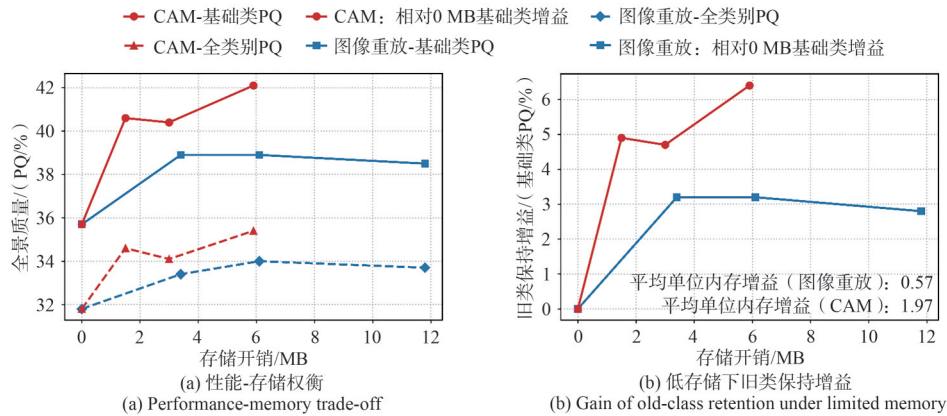


图4 不同重放策略在存储受限条件下的性能与旧类保持对比

Fig. 4 Comparison of performance and old-class retention for different replay strategies under storage-constrained conditions

有效。此外,CAM更高的单位存储收益表明其提升主要来自类别感知记忆机制本身,而不是依赖大规模图像回放。

3.4 定性分析

为更直观地展示本文方法的有效性,图5展示了若干ADE20K样本的原始图像及其对应的分割结果。从可视化结果可以看出,本文方法在不同场景条件下均能够较好地完成图像语义区域的划分,

并保持分割结果与原始场景内容之间较高的一致性。对于目标类别较多、空间布局较复杂的场景,模型能够较准确地识别主要语义区域;对于存在尺度变化和一定遮挡的场景,模型同样能够对关键目标区域形成有效响应,表明本文方法具有较好的场景适应能力。综合来看,定性可视化结果与前文定量实验分析相一致,进一步说明了所提方法在持续图像分割任务中的有效性和鲁棒性。

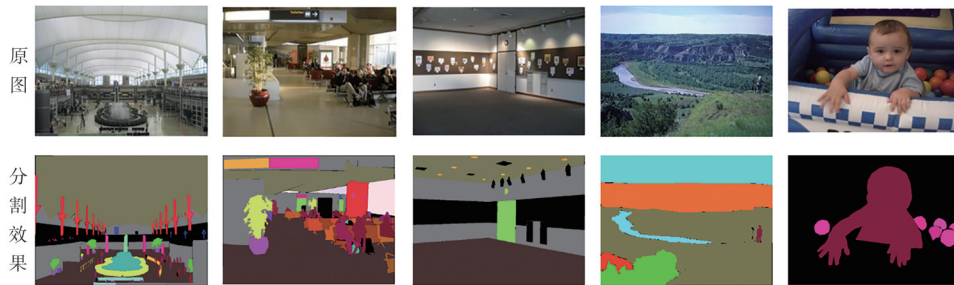


图5 ADE20K 100-5持续全景分割设置下 PAIS 的定性结果

Fig. 5 Qualitative results of PAIS under the ADE20K 100-5 continual panoptic segmentation setting

为全面验证模型在不同密集预测任务下的泛化能力,图6展示了PAIS在持续语义分割任务下的定性仿真结果。与全景分割相比,语义分割更侧重于全局区域的语义一致性与像素级密集分类边界。如图6所示,得益于时序一致性学习对关键语义锚点的稳定约束,PAIS能够精准识别大面积的背景区域并保持极高的内部平滑度。同时,类别感知记忆的引入使得模型在应对零散分布的新旧类别交界处时,依然能够输出锐利的语义边界,避免了类别间的相互渗透与吞噬。这一定性结果充分证明了本文框架在持续

语义分割任务中的卓越性能。

图7(a)展示了不同模块组合下的整体性能变化。可以观察到,随着PAI、TCL和CAM的逐步引入,模型在持续全景分割(all PQ)与持续语义分割(all mIoU)上均呈提升趋势,表明查询初始化质量提升、跨阶段稳定性增强与历史语义补偿三者具有互补作用。图7(b)展示了跨阶段查询漂移趋势:相比不使用TCL的情况,引入TCL后查询表征在连续阶段中的漂移幅度更小,说明时序一致性约束能够有效抑制增量学习中的表征偏移,从而降低旧类对象性退化风险。图7(c)展示了存储



图6 ADE20K 100-5持续语义分割设置下PAIS的定性结果

Fig. 6 Qualitative results of PAIS under the ADE20K 100-5 continuous semantic segmentation setting

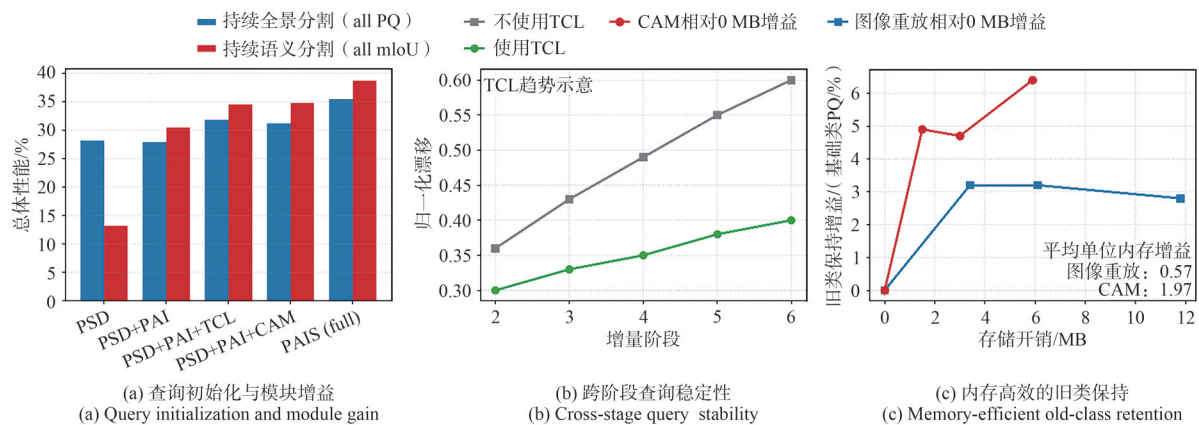


图7 查询演化仿真结果分析

Fig. 7 Analysis of evolution simulation results

受限条件下旧类保持增益与内存开销关系。以0 MB为基线时,CAM在更低存储预算下获得更高的旧类保持增益,且单位存储收益(gain/MB)优于图像重放,说明其性能提升主要来自类别感知记忆机制,而非依赖大规模图像回放。

4 结 论

本文提出一种面向类增量场景的原型对齐持续图像分割方法PAIS。针对持续学习过程中旧类知识易遗忘、新类表征不足以及背景语义动态偏移等问题,本文设计了原型感知初始化、时序一致性学习和类别感知记忆3个模块,分别从

查询初始化、跨阶段表征稳定和旧类知识补偿3个层面进行协同优化。在ADE20K数据集的多种任务设定中,PAIS在新类和基础类的PQ与mIoU指标上均显著优于现有方法,如在100-10持续全景分割任务中新类PQ提升13.0%,在100-50任务中基础类PQ表现超过联合训练基准。消融实验进一步验证了各模块在提升可塑性、稳定性与存储效率方面的互补作用,同时方法对任务顺序的性能波动控制在1.2%以内,体现出较强的鲁棒性。未来将进一步研究多模态信息融合、自适应增量机制以及更高效的训练与部署策略,以提升该方法在真实开放环境中的泛化能力与应用价值。

参 考 文 献:

- [1] CSURKA G, VOLPI R, CHIDLOVSKII B, *et al.* Semantic image segmentation: two decades of research [J]. *Foundations and Trends in Computer Graphics and Vision*, 2022, 14(1/2): 1-162.
- [2] PARISI G I, KEMKER R, PART J L, *et al.* Continual lifelong learning with neural networks: a review [J]. *Neural Networks*, 2019, 113: 54-71.

- [3] KIRILLOV A, MINTUN E, RAVI N, *et al.* Segment anything [C]//*Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision*. Paris: IEEE, 2023: 3992-4003.
- [4] CHAUDHRY A, DOKANIA P K, AJANTHAN T, *et al.* Riemannian walk for incremental learning: understanding forgetting and intransigence [C]//*Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018: 556-572.
- [5] DOUILLARD A, CORD M, OLLION C, *et al.* PODNet: pooled outputs distillation for small-tasks incremental learning [C]. 16th European Conference on Computer Vision. Glasgow: Springer, 2020: 86-102.
- [6] OSTAPENKO O, PUSCAS M, KLEIN T, *et al.* Learning to remember: a synaptic plasticity driven framework for continual learning [C]//*Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019: 11313-11321.
- [7] DOUILLARD A, RAMÉ A, COUAIRON G, *et al.* DyTox: transformers for continual learning with dynamic token expansion [C]//*Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE, 2022: 9275-9285.
- [8] HOU S H, PAN X Y, LOY C C, *et al.* Learning a unified classifier incrementally via rebalancing [C]//*Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019: 831-839.
- [9] SHIN H, LEE J K, KIM J, *et al.* Continual learning with deep generative replay [C]//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach: Curran Associates Inc., 2017: 2994-3003.
- [10] 唐璐, 万良, 王婷婷, 等. DECANet: 基于改进 DeepLabv3+ 的图像语义分割方法[J]. 激光与光电子学进展, 2023, 60(4): 0410002.
TANG L, WAN L, WANG T T, *et al.* DECANet: image semantic segmentation method based on improved DeepLabv3+ [J]. *Laser & Optoelectronics Progress*, 2023, 60(4): 0410002. (in Chinese)
- [11] 刘腊梅, 宗佳旭, 肖振久, 等. 流形正则化的交叉一致性语义分割算法[J]. 中国图象图形学报, 2022, 27(12): 3542-3552.
LIU L M, ZONG J X, XIAO Z J, *et al.* Cross-consistent semantic segmentation algorithm based on manifold regularization [J]. *Journal of Image and Graphics*, 2022, 27(12): 3542-3552. (in Chinese)
- [12] 杨静, 何瑶, 李斌, 等. 基于门控机制与重放策略的持续语义分割方法[J]. 电子与信息学报, 2024, 46(7): 2908-2917.
YANG J, HE Y, LI B, *et al.* A continual semantic segmentation method based on gating mechanism and replay strategy [J]. *Journal of Electronics & Information Technology*, 2024, 46(7): 2908-2917. (in Chinese)
- [13] BAEK D, OH Y, LEE S, *et al.* Decomposed knowledge distillation for class-incremental semantic segmentation [C]//*Proceedings of the 36th International Conference on Neural Information Processing Systems*. New Orleans: Curran Associates Inc., 2022: 754.
- [14] CERMELLI F, CORD M, DOUILLARD A. ComFormer: Continual learning in semantic and panoptic segmentation [C]//*Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023: 3010-3020.
- [15] GONG Y Z, YU S Y, WANG X Y, *et al.* Continual segmentation with disentangled objectness learning and class recognition [C]//*Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2024: 3848-3857.
- [16] KIM B, YU J, HWANG S J. ECLIPSE: efficient continual learning in panoptic segmentation with visual prompt tuning [C]//*Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2024: 3346-3356.
- [17] JIA M L, TANG L M, CHEN B C, *et al.* Visual prompt tuning [C]. 17th European Conference on Computer Vision. Tel Aviv: Springer, 2022: 709-727.
- [18] CHEN J P, CONG R M, LUO Y X, *et al.* Strike a balance in continual panoptic segmentation [C]. 18th European Conference on Computer Vision. Milan: Springer, 2024: 126-142.
- [19] 张龙. 基于 Transformer 的多目标跟踪算法研究[D]. 南宁: 广西大学, 2025.
ZHANG L. Research on multi-object tracking algorithm based on Transformer [D]. Nanning: Guangxi University,

2025. (in Chinese)
- [20] YU Q H, HE J, DENG X Q, *et al.* Convolutions die hard: open-vocabulary segmentation with single frozen convolutional CLIP [C]//*Proceedings of the 37th International Conference on Neural Information Processing Systems*. New Orleans: Curran Associates Inc. , 2023: 1399.
- [21] XU M D, ZHANG Z, WEI F Y, *et al.* Side adapter network for open-vocabulary semantic segmentation [C]//*Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023: 2945-2954.
- [22] 王伏增,汪西原,宋佳乾. 基于形态学滤波的标记分水岭脑肿瘤图像分割[J]. 宁夏大学学报(自然科学版), 2017, 38(1): 27-32.
WANG F Z, WANG X Y, SONG J Q. Brain tumor image segmentation using marker-controlled watershed based on morphological filtering [J]. *Journal of Ningxia University (Natural Science Edition)*, 2017, 38(1): 27-32. (in Chinese)
- [23] CARION N, MASSA F, SYNNAEVE G, *et al.* End-to-end object detection with transformers [C]. 16th European Conference on Computer Vision. Glasgow: Springer, 2020: 213-229.
- [24] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 770-778.
- [25] CERMELLI F, MANCINI M, BULÒ S R, *et al.* Modeling the background for incremental learning in semantic segmentation [C]//*Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 9230-9239.
- [26] DOUILLARD A, CHEN Y F, DAPOGNY A, *et al.* PLOP: learning without forgetting for continual semantic segmentation [C]//*Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021: 4039-4049.
- [27] CHA S M, KIM B, YOO Y, *et al.* SSUL: semantic segmentation with unknown label for exemplar-based class-incremental learning [C]//*Proceedings of the 35th International Conference on Neural Information Processing Systems*. Curran Associates Inc. , 2021: 835.
- [28] XIAO J W, ZHANG C B, FENG J K, *et al.* Endpoints weight fusion for class incremental semantic segmentation [C]//*Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver: IEEE, 2023: 7204-7213.
- [29] 刘玉,郭迎春,朱叶,等. 基于双交叉注意力Transformer网络的小样本图像语义分割[J]. 液晶与显示, 2024, 39(11): 1494-1505.
LIU Y, GUO Y C, ZHU Y, *et al.* Dual cross-attention Transformer network for few-shot image semantic segmentation [J]. *Chinese Journal of Liquid Crystals and Displays*, 2024, 39(11): 1494-1505. (in Chinese)
- [30] 刘玉,于明,朱叶. 基于双重聚合和自合并网络的小样本图像语义分割[J]. 液晶与显示, 2024, 39(10): 1421-1430.
LIU Y, YU M, ZHU Y. Bi-aggregation and self-merging network for few-shot image semantic segmentation [J]. *Chinese Journal of Liquid Crystals and Displays*, 2024, 39(10): 1421-1430. (in Chinese)
- [31] 李更生,刘国军. 融合统计范数度量的局部区域图像分割[J]. 光学精密工程, 2024, 32(16): 2550-2563.
LI G S, LIU G J. Local region image segmentation by fusion of statistical norm metrics [J]. *Optics and Precision Engineering*, 2024, 32(16): 2550-2563. (in Chinese)

作者简介:



高勇占,男,硕士研究生,2024年于长春师范大学获得学士学位,主要从事计算机视觉、图像处理等方面的研究。
E-mail: 2575624391@qq.com



黄凌霄,男,博士,副教授,2019年于宁夏大学获得博士学位,主要从事计算机视觉、计算水力学和河流动力学方面的研究。E-mail: huanglx@nxu.edu.cn