

增强不确定性引导均值教师模型的HER2组织病理图像半监督分割模型

姚博¹, 赵晶², 韩天晓³, 谢怡宁²

¹东北林业大学计算机与控制工程学院, 黑龙江 哈尔滨

²东北林业大学机电工程学院, 黑龙江 哈尔滨

³中电科东方通信集团有限公司北京分公司, 北京

收稿日期: 2026年1月13日; 录用日期: 2026年2月17日; 发布日期: 2026年2月25日

摘要

在乳腺癌诊断中, 为了解决大量需求HER2组织病理染色图像专家标注的问题, 传统半监督模型会过早收敛为自训练, 导致生成低置信度为标签。为了应对这些挑战, 本文提出一种新的增强不确定性引导的半监督分割网络, 用于HER2组织病理染色图像的细胞膜分割。首先, 该网络在协同均值教师模型中引入一个不确定目标选择模块, 该模块通过构建联合置信度矩阵, 利用学生模型和教师模型之间的预测差异, 有效识别出由于系统性误差导致的伪标签噪声区域。其次, 将识别出的不确定区域加入到一致性损失中, 引导教师学生模型在不确定目标区域预测中达成共识, 增强模型像素级目标选择能力。此外, 该方法无需修改主干网络, 也不增加额外的可训练参数, 具备良好的泛化能力和可扩展性。我们的实验在HER2组织病理染色图片数据集中得到验证, 结果表明, 我们的方法在Dice评价指标下得到73.09%的分数, 相较于基线模型有明显提升, 同时我们的方法也超越了最新的SSL方法, 也优于经典的全监督学习方法。该方法不仅在HER2组织病理染色图片数据集分割任务中表现优异, 在宫颈癌细胞核分割任务中Dice评价指标也达到85.02%, 也具备作为半自动标注工具在其他疾病领域的广泛应用。

关键词

HER2, 图像分割, 半监督学习

Enhanced Uncertainty-Guided Mean Teacher Model for Semi-Supervised Segmentation Model of HER2 Histopathological Images

Bo Yao¹, Jing Zhao², Tianxiao Han³, Yining Xie²

¹College of Computer and Control Engineering, Northeast Forestry University, Harbin Heilongjiang

²College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin Heilongjiang

³Beijing Branch of CETC Orient Communications Group Co., Ltd., Beijing

文章引用: 姚博, 赵晶, 韩天晓, 谢怡宁. 增强不确定性引导均值教师模型的 HER2 组织病理图像半监督分割模型[J]. 计算机科学与应用, 2026, 16(2): 488-504. DOI: 10.12677/csa.2026.162078

Abstract

In breast cancer diagnosis, the extensive demand for expert annotations of HER2-stained histopathological images poses a major bottleneck. Conventional semi-supervised learning frameworks tend to converge prematurely into self-training, which often leads to the generation of low-confidence pseudo-labels. To address these challenges, we propose a novel enhanced uncertainty-guided semi-supervised segmentation network for cell membrane segmentation in HER2-stained histopathological images. Specifically, an uncertainty-aware target selection module is introduced into a collaborative mean teacher framework. This module constructs a joint confidence matrix and exploits the prediction discrepancies between the student and teacher models to effectively identify pseudo-label noisy regions caused by systematic errors. Subsequently, the identified uncertain regions are incorporated into the consistency loss, guiding the teacher-student models to reach consensus on uncertain targets and thereby enhancing pixel-level target selection capability. Notably, the proposed method does not require any modification of the backbone network nor introduce additional trainable parameters, demonstrating strong generalization and scalability. Extensive experiments conducted on the HER2-stained histopathological image dataset validate the effectiveness of the proposed approach. Our method achieves a Dice score of 73.09%, showing a substantial improvement over the baseline model, while also outperforming recent state-of-the-art semi-supervised learning methods and even classical fully supervised approaches. Furthermore, the proposed framework exhibits strong generalization performance on cervical cancer nucleus segmentation, attaining a Dice score of 85.02%. These results indicate that our method is not only effective for HER2 histopathological image segmentation but also holds significant potential as a semi-automatic annotation tool for a wide range of diseases.

Keywords

HER2, Image Segmentation, Semi-Supervised Learning

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

乳腺癌在全世界女性中非常常见(肺癌后第二流行的癌症) [1]-[4]。乳腺癌评分大多是按照 Nottingham 评分系统[5]完成的。临床医生和病理学家通常使用免疫组织化学(IHC)研究来衡量乳腺癌的严重程度。IHC 筛查分析了活检组织切片中细胞内和细胞表面的蛋白质, 以确定组织切片的分子表型和组织学亚型。IHC 诊断标记物, 如雌激素受体、孕激素受体、Ki-67 和人表皮生长因子受体 2 (HER2), 被广泛用于识别转移性和良性肿瘤、肿瘤分级以及确定癌组织的起源。在临床诊断中, HER2 阳性定位于细胞膜, 诊断标准为视野内阳性细胞数在全部肿瘤细胞的占比, 因此, 在临床诊断中标记出染色的细胞膜是诊断的重要的一步。

在常规临床实践中, 病理学家在显微镜下直观地分析活检组织载玻片。这种目视检查通常是主观的且容易出错。最终, 在资源有限、病理学家无法及时获得有效信息的领域, 医学图像分割已经成为强大的工具[6]。然而, 组织病理图像的切片中有大量的组织细胞和组织液, 因此区分出组织细胞膜工作量极大且难以辨认, 对于全监督语义分割, 需要制作像素级注释的标签, 而制作像素级注释的标签既困难又

耗时,因此越来越多的临床研究使用半监督学习(SSL),它通过少量标记图像和大量未标记图像进行训练,但半监督语义分割的效果是否可以达到全监督语义分割的水平是很重要的问题。

为应对上述挑战,我们进一步优化了不确定性引导的协同 Mean Teacher 架构。具体而言,我们引入了两个歧义共识模块[7],以充分挖掘未标记数据中的互补信息线索。在此基础上,我们从标签噪声自识别的视角,系统提出了一种可即插即用的不确定性目标选择策略,该策略无需引入额外的可训练参数或对主干网络进行结构性修改。此外,我们认为歧义共识区域蕴含更丰富的语义信息,因此将其评价映射纳入一致性损失计算,以引导两个模型在这些不确定区域中的预测趋于一致。本质上,该方法旨在从未标记数据中识别并聚焦于最具信息价值、训练难度较高的区域,从而提升模型对复杂结构的建模能力。

简而言之,我们的主要贡献包括:

(1) 我们提出一个新的有效的 HER2 染色细胞膜分割方法,该方法利用大量未标记的图像结合标签的噪声来提高学习模型的泛化性。

(2) 我们通过解决了一致性目标的选择的问题,即在不确定但具有信息量的区域中鼓励扰动稳定性可以促使模型从未标注数据中学习到有用的表示。

(3) 我们在 HER2 染色组织切片图像数据集和宫颈癌细胞核公开数据集上进行了广泛实验,测试结果证明了我们的模型优于最先进的模型,其中在无标签数据占比更多的实验中我们的模型效果提升更明显。

2. 相关工作

2.1. 半监督学习

半监督学习(Semi-Supervised Learning, SSL)旨在结合少量标注样本与大量未标注数据,从而在标注成本受限的条件下实现接近全监督模型的性能。目前,主流的半监督学习方法主要可归为两大类:一致性正则化(Consistency Regularization)方法和伪标签(Pseudo-labeling)方法。前者通过引导模型在不同扰动下保持预测结果的一致性,从而增强模型的泛化能力;后者则通过生成高置信度的伪标签,将未标注样本纳入监督信号,以逐步提升模型的判别能力。

一致性正则化方法主要基于平滑性假设(smoothness assumption),即对输入施加轻微扰动不应显著影响模型的预测结果[8]。据此,该方法通过在不同层级引入微小扰动,以促使模型在扰动前后的预测保持一致性。根据扰动施加的位置,一致性正则化方法可进一步细分为数据级别扰动[9]、特征级别扰动[10][11]以及模型级别扰动[12]。在训练过程中,学生模型通过一致性约束被引导其输出尽可能贴近教师模型的预测结果。教师模型通常通过对学生模型参数进行指数移动平均(EMA)更新,从而具备更强的稳定性。此外,部分研究引入 Monte Carlo Dropout 以对预测不确定性进行量化,并在高置信区域加入一致性损失,从而增强模型在不确定区域的鲁棒性与泛化能力。

伪标签方法则主要采用自训练(self-training)策略,基于聚类假设(cluster assumption),即同一类别的样本在特征空间中应形成紧凑聚类,并呈现低熵分布。该类方法通常先在无标注数据上生成伪标签,再将其与原始标注数据联合用于模型再训练。例如, Kalluri 等人引入熵正则模块,引导模型在无标注数据上生成低熵预测;在此基础上, Sohn 等人进一步提出使用弱增强生成的伪标签来监督增强样本的训练,从而提高伪标签的鲁棒性与泛化能力。

2.2. 半监督语义分割

医学图像的语义分割任务通常依赖于像素级别的精细标注,但获取此类高质量标签的过程耗时且成本高昂,严重限制了大规模有监督方法在临床实践中的应用。为此,半监督语义分割(Semi-Supervised Semantic Segmentation, S4)通过融合少量标注数据与大量未标注图像,在降低标注成本的同时,有望达到甚

至超过全监督方法的性能表现。

近年来,研究者提出了多种有效的半监督语义分割方法。例如,均值教师(Mean Teacher)方法[13][14]借助教师模型(teacher)对学生模型(student)进行监督,引导其在未标注图像上生成更为稳定的预测,从而增强模型的泛化能力。同时,教师模型通过对学生模型参数的指数移动平均(EMA)计算生成一致性目标,有效改进了传统 Π 模型的训练机制[15]。深度协同训练(Deep Co-training) [16]则引入多个模型或多分支架构,利用模型间的互相监督为无标注图像提供更可靠的伪标签或一致性约束。交叉伪监督(Cross Pseudo Supervision) [17]则通过设置两个网络分支或训练两个独立模型,让其在无标注样本上相互生成伪标签并进行交叉监督,从而降低伪标签的不确定性。在半监督语义分割中,伪标签生成与一致性正则化构成了两类主流方法。其中,伪标签策略是一种直观的自训练机制,先利用有限的标注样本训练初始模型以生成未标注图像的伪标签,再将伪标签与标注数据联合用于模型迭代优化[18]-[20]。而一致性学习方法则在相同图像上施加不同增强操作,迫使网络对增强前后的输入输出一致,从而提升模型对输入扰动的鲁棒性。然而,当模型泛化能力较弱时,不同增强方式可能导致图像在特征空间中相距较远,从而削弱一致性约束对模型优化的有效性[21]。为应对该问题,已有研究从多个层面提出一致性扰动策略,包括输入层扰动[9]、特征层扰动[10]以及模型结构扰动[5][17][22][23],以增强模型在不同上下文中的一致性表现。然而,这些方法存在过早收敛至共识的问题,易退化为传统自训练机制,限制了模型性能的进一步提升。有研究提出,基于高置信度伪标签引导的协同训练在增强模型间分歧的同时,可有效提升半监督学习效果。

因此,本文提出一种新的半监督学习框架,旨在生成更高质量的伪标签并维持模型间适度分歧,从而进一步提升医学图像分割任务中的协同训练效果与整体性能。

2.3. 不确定性引导的半监督语义分割

不确定性分析在机器学习与计算机视觉领域中受到了广泛关注[24]。在半监督语义分割任务中,不确定性估计为模型从伪标签中提取可靠信息提供了重要依据。当前主要存在两类主流策略用于评估模型预测的不确定性:一类是通过 Monte Carlo Dropout 技术对模型进行多次前向传播并计算预测方差[25];另一类则直接衡量不同模型预测之间的统计差异,如方差估计等方法[26]。在半监督语义分割领域,已有研究将不确定性信息用于重加权训练损失[27][28],以在训练过程中动态区分“可信”与“不可信”的像素,并优先关注高置信度区域,以提升监督信号的质量。此外,不确定性还被用于对比学习中的样本选择[27],通过区分相似与不相似样本,进一步增强模型对语义一致性区域的表征能力。也有方法引入专门的不确定性估计策略以过滤质量较低的伪标签[29]。然而,许多方法依赖于人工设置的置信度阈值,不仅对数据集敏感,而且难以统一调参,导致部署复杂性增加。

为此,本研究提出了一种去除阈值依赖的不确定性估计策略,旨在更稳健地评估伪标签的可信程度及其对模型训练的贡献。尽管当前主流方法已在性能上取得进展,但我们强调,提升模型的一致性目标仍具有重要意义。与本研究最为相关的是 UA-MT 方法[27],该方法基于 Kendall 与 Gal (2017)提出的 Monte Carlo Dropout 框架对教师模型的不确定性进行建模,并仅对高置信度区域施加一致性损失,以提升伪标签的可靠性。我们则进一步通过优化无监督一致性目标的选择,强调从可靠伪标签中高效学习的重要性,从而在多个半监督语义分割任务中实现了显著性能提升。

3. 方法

3.1. 方法概述

本文提出了一种基于均值教师(Mean Teacher)框架的加强不确定选择的均值教师模型。该方法在传统均值教师结构的基础上,引入协同学习机制,并设计了两个不确定性选择模块,以增强模型对高训练价

值目标的筛选能力。EUGMT 的实现流程如表 1 所示。具体而言，我们将置信度阈值引入至经典混淆矩阵的构建中，形成联合置信矩阵，以缓解类别不平衡与模型过度置信预测等常见问题。通过不确定性选择模块，模型能够筛选出高价值的目标像素，并据此计算一致性损失，从而提升模型在 HER2 染色组织病理切片图像及其他复杂医学图像中对复杂结构的建模与表达能力。整体框架如图 1 所示：

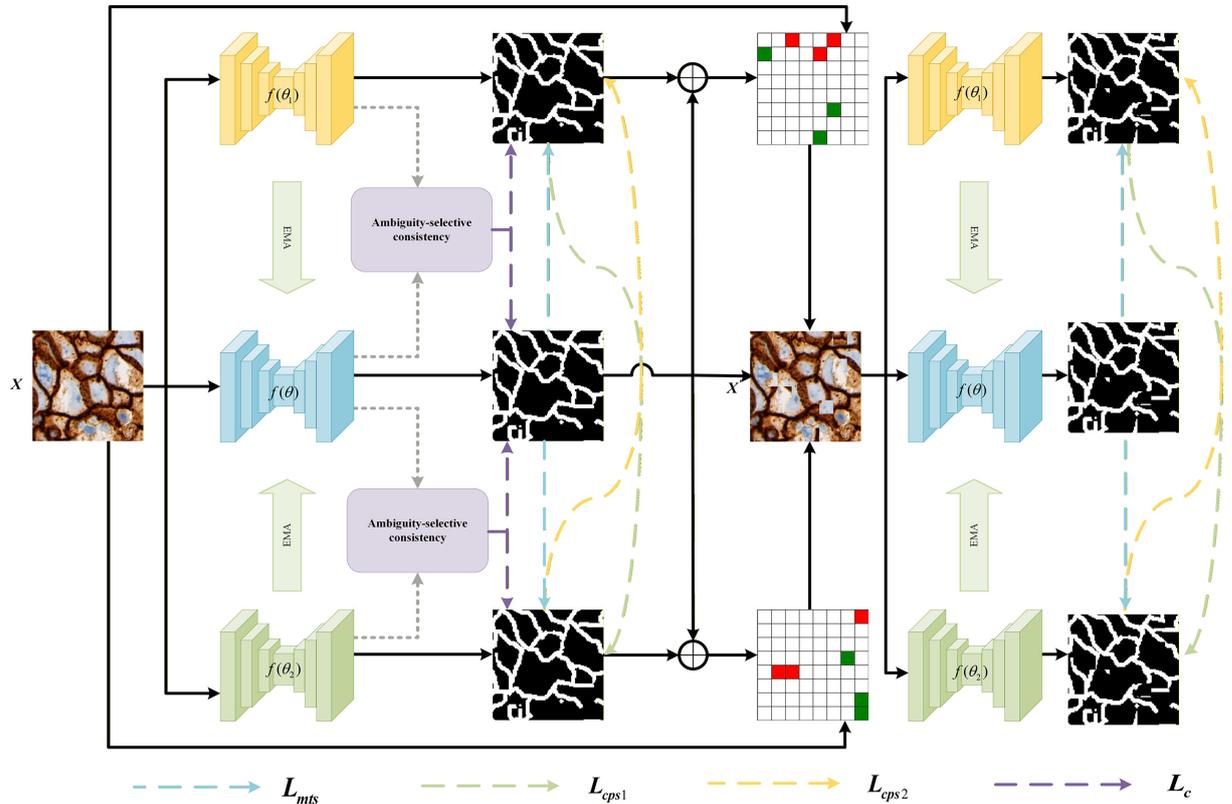


Figure 1. Architecture of the EUGMT network
图 1. EUGMT 网络结构图

3.2. 背景：均值教师模型

设有标签数据集记为 D_L ，无标签数据集记为 D_U ，其中 $D_U = \{(X_j)_{j=N+1}^M\}$ 表示包含 $M - N$ 个无标签样本的集合， $D_L = \{(X_i, Y_i)_{i=1}^N\}$ 表示包含 N 个有标签样本的集合， Y_i 为样本 X_i 对应的真实标签。在遵循经典 Mean Teacher (MT) 框架的前提下，该半监督学习任务可形式化地表示如下：

$$L = L_s + \lambda L_u \tag{1}$$

其中， L_s 是 D_L 的监督损失， L_u 是 D_U 的无监督损失， λ 是一个正则化参数，用于平衡监督损失与无监督损失的权重。我们采用高斯渐增函数逐步提升该系数的值，以实现训练过程中无监督分支的渐进式引导，其具体形式如下：

$$\lambda(t) = \lambda_m \times \exp \left[-5 \left(1 - \frac{t}{t_m} \right)^2 \right] \tag{2}$$

其中， λ_m 控制加权函数的最大值， t 表示当前迭代次数， t_m 为训练过程中的最大迭代次数。

其次，在 MT 模型中，教师的权重通过学生权重的指数移动平均(EMA)来更新，其公式为：

$$\tilde{\theta}_t = \alpha \tilde{\theta}_{t-1} + (1 - \alpha) \theta_t \quad (3)$$

其中， α 是 EMA 的衰减率，一般设置为 0.99， θ_t 和 $\tilde{\theta}_t$ 分别是在训练步 t 时的学生模型和教师模型参数权重。

3.3. 加强不确定性选择用于半监督分割

如图 2 所示，其中 $f(\theta_1), f(\theta_2)$ 表示学生模型， $f(\theta)$ 表示教师模型。 P_s 和 \hat{Y}_s 分别表示学生模型的预测概率和基于 argmax 的伪标签； P_t 表示自蒸馏教师模型的预测概率。 L_c 是无标签数据的不确定掩码一致性损失。为提升学生模型与教师模型在预测结果上的一致性，并增强其对不确定区域的识别能力，本文设计了一种加强不确定性选择机制(Ambiguity-selective)。该机制通过引入目标选择策略，有效强化学生模型和教师模型在高不确定区域的学习能力，并将所筛选出的高价值目标区域加入一致性损失计算中，从而提升模型在无标签数据中的信息利用效率。不确定性通常源于像素自身的模糊特性，为此，我们构建了一种基于类别条件的系统性标签噪声识别策略，以有效识别和表征这类由数据本身特性引发的不确定性。值得强调的是，该策略为模型无关设计，既无需对主干网络结构进行任何修改，也不引入额外的训练参数，因而具备良好的通用性和可扩展性，可适用于多种主干网络以适应不同的医学图像分析任务。

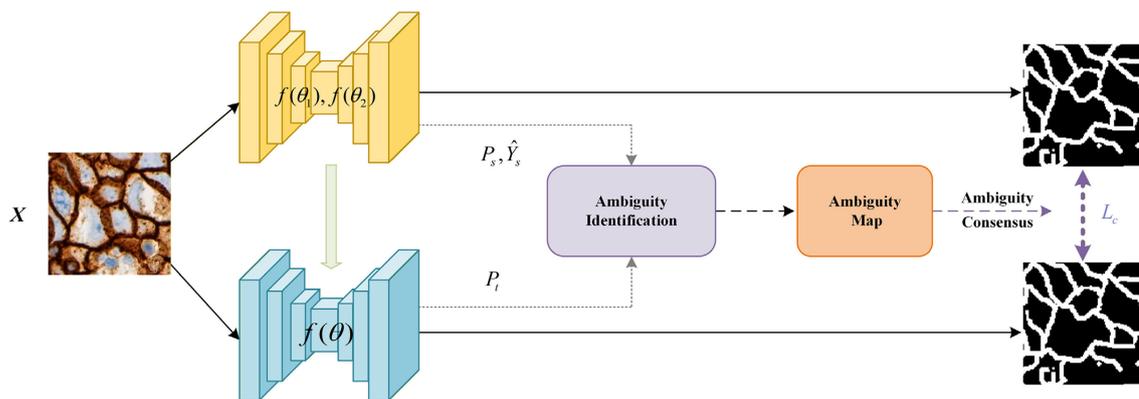


Figure 2. Method overview of the uncertainty selection module

图 2. 不确定选择模块的方法概述

Table 1. EUGMT algorithm pseudocode

表 1. EUGMT 算法伪代码

算法 1 EUGMT 算法

输入：标记数据： $D_l = \{(x_i, y_i)\}_{i=1}^N$ ，未标记数据： $D_u = \{x_j\}_{j=1}^M$ ，网络参数： $\theta_1, \theta_2, \theta_t$ ，最大训练轮数 E ，Top-K patch 数 K

训练完成的教师模型： $f(\cdot; \theta_1), f(\cdot; \theta_2), f(\cdot; \theta_t)$

1. for $epoch = 1$ to $Num\ epochs$ do
2. 从 D_l 采样标注样本 (x_i, y_i) ，从 D_u 采样未标注样本 x_u
3. 监督训练过程：
4. $\hat{y}_1 = f(x_i; \theta_1)$ ， $\hat{y}_2 = f(x_i; \theta_2)$ ， $\hat{y}_t = f(x_i; \theta_t)$
5. $L_{sup} = L(\hat{y}_1, y_i) + L(\hat{y}_2, y_i)$
6. 无监督协同训练：

续表

7. $x_u^1, x_u^2 = \text{StrongAugment}(x_u)$
8. $p_1 = f(x_u^1; \theta_1), p_2 = f(x_u^2; \theta_2), p_t = f(x_u; \theta_t)$
9. 生成伪标签:
10. $\tilde{y}_1 = \arg \max(p_1), \tilde{y}_2 = \arg \max(p_2), \tilde{y} = \arg \max(p_t)$
11. 构造噪声样本 x_u^{noise}
12. $p_t^{\text{soft}} = \text{softmax}(f(x_u; \theta_t)), p_t^{\text{noise}} = f(x_u^{\text{noise}}; \theta_t)$
13. 利用 Cleanlab 检测像素噪声并生成置信掩码 M_1, M_2
14. $L_{\text{con}} = \sum M \cdot \|p_{\text{student}} - p_t^{\text{noise}}\|^2$
15. CPS 损失: $L_{\text{cps}} = \frac{1}{2}[L(p_1, \tilde{y}_2) + L(p_2, \tilde{y}_1)]$
16. CMT 损失: $L_{\text{cmt}} = \frac{1}{2}[L(p_1, \tilde{y}) + L(p_2, \tilde{y})], L_u = \alpha(L_{\text{cps}} + L_{\text{cmt}}) L = L_{\text{sup}} + \lambda L_u + L_{\text{con}}$
17. 反向传播更新 $\theta_1, \theta_2, \theta_t = \text{EMA}(\theta_t, \theta_2)$
18. 不确定性引导 Patch 交换: 计算预测熵图 U_1, U_2
19. 在两个学生之间交换对应 patch 及标签
20. 重构新样本 (x', y')
21. 重新计算 CPS 与 CMT 损失, 更新 $\theta_1, \theta_2, \theta_t = \text{EMA}(\theta_t, \theta_2)$
22. end for
23. return $f(:, \theta_A), f(:, \theta_B), f(:, \theta_C)$

3.3.1. 类别条件的系统性标签噪声识别

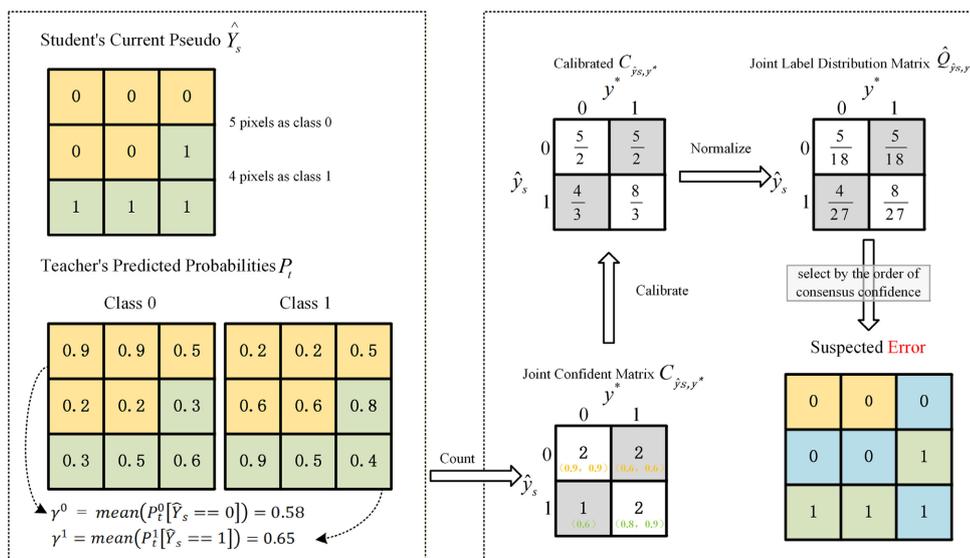


Figure 3. Schematic illustration of the category label self-identification strategy
图 3. 类别标签自识别策略方法示意图

我们引入了一种类别标签自识别策略，该策略借鉴了置信学习中关于类别条件标签噪声识别的核心

思想。在此策略中，联合置信矩阵(Joint Confidence Matrix, JCM)展现出在应对类别不平衡及模型过度置信预测方面的稳健性，其归一化后的统计量可有效表征由潜在标签噪声引发的随机性不确定性。为了适应我们的图像分割任务，我们对这个思想进行了调整，即假设每个像素 $x \in X$ 有一个真实(潜在)的标签 y^* ，而每个类别 $j \in L$ 的标签可以以独立概率 $p(y=i | y^*=j)$ 被错误标记为类别 $i \in L$ 。我们可以借鉴随机游走的想法来计算联合概率 $p(y=i, y^*=j)$ ，从而计算图像 X 中可能被错误标注的像素数量。该策略依赖外部预测结果，因此能够有效识别系统性标注误差。例如，在 HER2 染色的组织切片图像中，部分与任务无关的组织区域可能被系统性地误标为细胞膜区域。在此基础上，我们利用教师模型 $f(\theta)$ 生成的加权平均预测 P_t 作为具有集成特性的可靠参考，同时，将学生模型的预测 \hat{Y}_s 视作受噪声影响的观测标签，用以构建联合置信矩阵。综上所述，教师模型通过不断提供自集成的概率预测信息，为识别学生模型 $f(\theta_1)$ 和 $f(\theta_2)$ 预测 \hat{Y}_s 中的错误标签提供了稳定支撑。

其次，如图 3 所示。我们用共识置信度来表示一个像素 x 属于学生模型的伪标签 $\hat{y}_s = i$ 的概率 $p_i^t(x)$ (如图中类别 0 黄色部分所示 $P_t^0[\hat{Y}_s = 0]$)。如果一个像素 x 属于学生模型的伪标签 $\hat{y}_s = i$ ，但其满足 $\hat{p}_j^t(x) \geq \gamma_j$ ，则该像素的潜在真实标签 y^* 可能是 j 而不是 i 。我们希望得到图像 X 的每类阈值，这需要计算每个类别的平均置信度，其公式如下：

$$\gamma_j = \text{mean}\left(P_t^j[\hat{Y}_s = j]\right) \quad (4)$$

基于该阈值，我们进一步估计联合置信矩阵 $C_{\hat{y}_s, y^*}$ ，即在图像 X 统计满足每个类别阈值的像素数量所得矩阵，其公式如下：

$$C_{\hat{y}_s, y^*}[i][j] = \left| \hat{X}_{\hat{y}_s, y^*}^i \right| = \left| \left\{ x \in X_{\hat{y}_s=i} : p_i^t(x) \geq \gamma_j, j = \arg \max_{l \in L, p_l^t(x) \geq \gamma_l} p_l^t(x) \right\} \right| \quad (5)$$

其中， L 表示语义类别。由于阈值过滤掉了大多数像素，我们需要对联合置信矩阵 $C_{\hat{y}_s, y^*}$ 进行校准，然后对校准后的矩阵进行归一化得到联合标签分布矩阵 $\hat{Q}_{\hat{y}_s, y^*}[i][j]$ ，其计算方法如下：

$$\hat{Q}_{\hat{y}_s, y^*}[i][j] = \frac{C_{\hat{y}_s, y^*}[i][j]}{\sum_{j \in L} C_{\hat{y}_s, y^*}[i][j]} \cdot |X_{\hat{y}_s=i}| \quad (6)$$

$$\sum_{i \in L, j \in L} \left(\frac{C_{\hat{y}_s, y^*}[i][j]}{\sum_{j \in L} C_{\hat{y}_s, y^*}[i][j]} \cdot |X_{\hat{y}_s=i}| \right)$$

其中， $|X_{\hat{y}_s=i}|$ 表示伪标签 $\hat{y}_s = i$ 像素数量。

然后借鉴随机游走的方法， $\hat{Q}_{\hat{y}_s, y^*}$ 的非对角元素反映了不同类别被错误标注的可能性，而低共识置信度可以作为用来判断某个像素是否可能被错误标注的启发式指标。因此选择每个类别 i 中共识置信度最低的错误像素，其数量为：

$$\text{round}\left(|X| \cdot \sum_{j \in L, j \neq i} \hat{Q}_{\hat{y}_s, y^*}[i][j]\right) \quad (7)$$

其中 $|X|$ 表示图像 X 中像素总数。如图 3 蓝色区域所示，有三个被标注为背景的像素点被可能是错误标注，然后，还有可能存在错误的目标像素点。这样，学生模型的预测可以划分为错误组和正确组，接着如图 2 所示，将错误组作为不确定像素进行下一步损失计算。

3.3.2. 不确定性一致损失函数设置

为了同时利用标注数据和未标记数据，我们将损失函数分为两部分，即监督损失和无监督损失。其

损失函数表示如下：

$$L = L_s + \lambda(L_{cps} + L_{ms} + L_c) \quad (8)$$

在这里， λ 是正则化参数， L_s 表示监督损失， L_c 表示不确定选择模块筛选的出目标像素的一致性损失。监督损失 L_s 计算公式如下：

$$L_s = \frac{1}{N} \sum_{i=1}^N \{L_{seg}(f(\theta_1), Y_i) + L_{seg}(f(\theta_2), Y_i)\} \quad (9)$$

其中， Y_i 表示第 i 个真实标签， L_{seg} 是任何监督语义分割损失，例如交叉熵损失和 Dice 损失。在我们的实验中使用了 Dice 损失。

我们的方法的无监督损失由 L_{cps} ， L_{ms} 和 L_c 三部分组成，其中， L_{cps} 表示交叉伪监督，这个损失函数是由两个学生模型得到的，所以其定义为：

$$L_{cps} = L_{cps1} + L_{cps2} \quad (10)$$

其中， L_{cps1} 和 L_{cps2} 分别是两个学生模型双向交互得到的，其定义式如下：

$$L_{cps[1/2]} = \frac{1}{M-N} \sum_{j=1}^{M-N} L_{seg}(f(X_j; \theta_{[1/2]}), \hat{Y}_{j=1}^{[2/1]}) \quad (11)$$

其中， X_j 是无标签数据样本， \hat{Y}_j^1 和 \hat{Y}_j^2 分别是学生模型 $f(\theta_1)$ 和 $f(\theta_2)$ 对无标签数据 X_j 的预测作为伪标签。 L_{ms} 表示教师模型 $f(\theta)$ 指导学生模型的均值教师监督，目的是为了避免两个学生模型朝错误的方向共同训练， L_{ms} 通过两个分支得到，其表达式如下：

$$L_{ms} = L_{ms1} + L_{ms2} \quad (12)$$

其中， $L_{ms[1/2]}$ 表示为：

$$L_{ms[1/2]} = \frac{1}{M-N} \sum_{i=1}^{M-N} L_{seg}(f(X_j; \theta_{[1/2]}), \hat{Y}_j^0) \quad (13)$$

其中， \hat{Y}_j^0 是通过教师模型 $f(\theta)$ 得到的预测。

L_c 表示不确定选择模块筛选出的目标像素的一致性损失。为了加强在可靠但信息量较少的区域进行一致性约束，其次，为了鼓励学生模型和教师模型的预测保持一致，我们将一致性损失函数设计为不确定性掩码下的均方误差， L_c 计算公式如下：

$$L_c(f_t, f_s, E) = \frac{\sum_v \left[E_v \sum_c \left(\|f_{t,v,c} - f_{s,v,c}\|^2 \right) \right]}{\sum_v E_v} \quad (14)$$

其中， E_v 是第 v 个像素二值不确定指示器； $f_{t,v,c}$ 和 $f_{s,v,c}$ 分别是教师模型和学生模型在第 v 个像素上对每个类别 $c \in C$ 上的 softmax 之前的预测值。

4. 实验与结果

4.1. 数据集和评价指标

4.1.1. HER2 染色组织病理切片图像分割数据集

我们构建了一套用于 HER2 免疫组化(IHC)诊断的细胞膜分割数据集。该数据集共包含 126 张 HER2 染色组织病理全视野图像(Whole Slide Image patches)，每张原始图像分辨率为 1536×1536 。所有图像均来自于临床 HER2 IHC 染色切片，并在专业病理学家指导下完成精细化人工标注。在标注过程中，以 HER2

阳性表达区域对应的棕色细胞膜作为主要标注对象, 严格遵循临床 HER2 评分标准并且在专家的指导下, 仅对具有明确膜染色特征的肿瘤细胞膜进行逐像素勾画, 背景区域及非特异性染色区域不纳入标注范围。该标注结果主要用于支持后续 HER2 表达状态判定与辅助诊断分析。为适配深度学习模型训练需求, 采用滑窗裁剪策略将原始大尺寸图像划分为大小为 256×256 的小块图像。最终共获得 2905 张图像块, 其中 2350 张用于模型训练, 其余 555 张用于测试评估, 训练集与测试集在患者层面进行划分, 以避免数据泄露并保证评估的客观性。该数据集以细胞膜分割任务为核心, 通过精确提取 HER2 染色细胞膜结构, 为后续 HER2 定量分析与诊断决策提供可靠的形态学依据。

4.1.2. 宫颈癌细胞核分割数据集

我们除了使用 HER2 染色组织图片作为实验的基准, 我们还使用了宫颈癌细胞核分割数据集作为评估我们的实验的数据集。该数据集来自[30], 在这个数据集中, 核图像是在复杂条件下收集的核图像, 聚类核的聚类分割数据集(ClusterSeg)。我们使用 1500 张图片用于训练, 454 张图片用于测试, 这个数据集从不同的角度综合评价宫颈癌核细胞分割方法。对于这个数据集的癌细胞的细胞核分割是具有挑战性的, 以便在后续步骤中对良性和恶性癌细胞进行分类。

4.2. 基线方法与实验细节

我们的实验使用配置 ResNet50 的 Deeplabv3+作为基线架构。我们所有图像均调整为 256×256 的分辨率。在实验中, 我们按照经验将参数设置为 $\lambda_m = 1$ 、 $k = 2$ 、 $r = 16$ 、 $\alpha = 0.99$ 、 $\beta = 0.99$ 。优化器使用 AdamW, 学习率为 10^{-4} , 批次大小设置为 6, 训练 50 个 epoch。在保证训练数据一致的情况下, 我们的实验设置多个 SSL 的方法, 其中包括 Duo-SegNet [31], MMS [32], crcfp [12], CCT [12]。除了和我们实验方法相关的 SSL 方法, 我们还和一些经典的全监督方法做对比, 其中包括: TransUnet [33], SwinUnet [34], TransFuse [35], FAFuse [36]。

4.3. HER2 染色组织图片分割实验

4.3.1. 消融实验分析

本节设计了三组消融实验, 以系统评估各关键模块对模型性能的贡献。第一组实验以 DeepLabv3+为基础网络, 构建不确定性引导的协同均值教师(Uncertainty-guided Collaborative Mean Teacher, UCMT)模型, 作为半监督分割的基本框架。第二组实验在 UCMT 架构上进一步引入不确定性选择共识模块与一致性损失函数, 形成增强不确定选择的均值教师模型(Enhanced Uncertainty-Guided Mean Teacher, EUGMT), 以验证在扰动区域强化预测一致性对分割性能的影响。第三组实验以 U-Net 替代 EUGMT 中的 DeepLabv3+主干网络, 评估主干结构差异对半监督分割效果的影响。

上述三组实验均在极端低标注比例(2%)的条件下进行, 实验结果如表 2 所示。为进一步验证所提出方法在更低标注成本场景下的适应能力, 我们在相同设置下补充了 5% 标注数据的实验。除标注比例外, 其余训练配置保持一致, 对应实验结果汇总于表 3。

Table 2. Ablation study with 2% labeled data

表 2. 2% 标注数据消融实验

方法	Dice	IoU	ASD	HD
DeepLabv3+ (SupOnly)	0.6951	0.5327	60.55	342.32
UCMT	0.7206	0.5633	6.23	29.00
EUGMT (DeepLabv3+)	0.7309	0.5760	5.71	27.31
EUGMT (Unet)	0.7092	0.5494	6.32	32.52

Table 3. Ablation study with 5% labeled data**表 3.** 5% 标注数据消融实验

方法	Dice	IoU	ASD	HD
DeepLabv3+ (SupOnly)	0.7008	0.5394	60.56	342.42
UCMT	0.7274	0.5717	5.21	34.52
EUGMT (DeepLabv3+)	0.7241	0.5677	5.18	28.16
EUGMT (Unet)	0.7037	0.5429	5.56	39.01

(1) Deeplabv3+作为网络的不确定性引导的协作均值教师实验

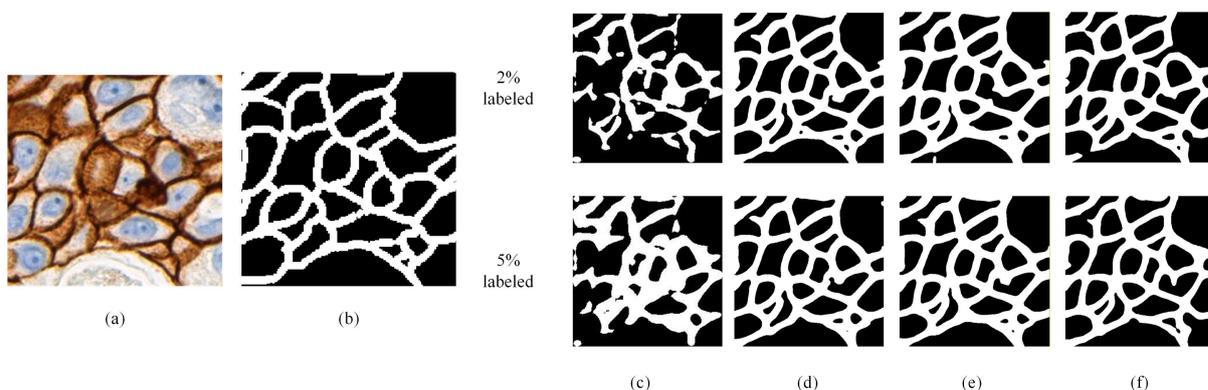
本小节旨在评估以 DeepLabv3+作为主干网络的不确定性引导协同均值教师(UCMT)模型在半监督语义分割任务中的性能表现。作为经典的图像分割架构，DeepLabv3+在多尺度上下文建模及高分辨率特征保留方面具备显著优势，其模块化结构也为模型在多种场景中的适配与扩展提供了良好基础。将 DeepLabv3+集成至 UCMT 框架中，不仅增强了模型对复杂结构的表达能力，还可通过模型间预测分歧有效提取未标记数据中的互补信息。实验结果表明，与基线模型相比，UCMT 模型的 Dice 系数提升了 2.55%，验证了该方法在缺少标签条件下提升分割性能的有效性与可行性。

(2) 加入不确定性选择共识模块和一致性损失函数的实验

本小节旨在评估在第一阶段的学生-教师模型中引入两个不确定性选择共识模块与两个一致性损失函数的有效性。需要指出的是，上述模块仅在第一阶段模型中进行集成，旨在探索其对初始特征提取与伪标签质量的影响。如表 2 所示，引入这些模块后模型整体性能显著提升，其中 Dice 系数相较于原始 UCMT 模型提升了 1.03%，较基线模型提升了 3.58%。该结果充分验证了所提出的不确定性选择共识模块与一致性损失函数在优化模型预测一致性、提升半监督分割性能方面的有效性与实用性。

(3) 使用 Unet 网络替换 EUGMT 中 Deeplabv3+网络实验

作为医学图像分割领域最经典的架构之一，U-Net 同样可被集成至增强不确定选择的均值教师模型(Enhanced Uncertainty-Guided Mean Teacher, EUGMT)中。为评估主干网络结构对模型性能的影响，我们将原 EUGMT 中的 DeepLabv3+替换为 U-Net，并在相同实验设置下进行对比。实验结果如表 2 所示，采用 U-Net 的模型在性能上未能超越基线模型，且相比采用 DeepLabv3+的 EUGMT 模型表现出明显下降。该结果表明，DeepLabv3+在 HER2 染色组织病理图像数据集上，凭借其更强的多尺度上下文感知与高分辨率特征保持能力，作为半监督框架的主干网络具备更优的适应性与实用性。

**Figure 4.** Ablation study results on the HER2-stained histopathological image segmentation dataset**图 4.** HER2 染色组织病理切片图像分割数据集消融实验结果

综上所述, 本文提出的增强不确定性选择共识模块在 HER2 染色组织病理图像中对细胞膜区域的分割表现最为优越。该模块有效提升了模型对复杂边界结构的识别与定位能力, 进一步验证了其在高挑战性医学图像任务中的适用性与优势。具体分割结果可视化如图 4 所示。其中, (a) images, (b) ground truth, (c) supervised baseline, (d) UCMT, (e) EUGMT(DeepLabv3+), (f) EUGMT(Unet)。

(4) 使用 5% 标注数据消融实验

本小节旨在探讨所提出模型在更低标注数据比例条件下的鲁棒性与适应性。如表 3 所示, 所有实验在标注样本进一步减少的情况下均较表 2 呈现不同程度的性能下降。然而, 即便在极端稀缺的标注条件下, 模型依然保持了较优的分割性能, 充分体现了所提出方法在低资源环境中的有效性与泛化能力。其具体分割效果可视化如图 4 所示。

4.3.2. 对比试验分析

本节选用 HER2 染色组织病理图像作为评估数据集, 系统比较不同模型在细胞膜分割任务中的性能表现。为全面验证所提出方法的有效性, 我们在仅使用 2% 标注数据的条件下, 将本文提出的模型与当前主流的非监督分割方法(包括 Duo-SegNet、MMS 和 CCT)以及医学图像分割领域中广泛采用的全监督方法(如 TransUnet、SwinUnet、FAFuse 和 TransFuse)进行了对比。表 4 汇总了各方法在 HER2 染色图像分割任务中的性能评估结果。实验结果显示, 尽管标注数据极为有限, 所提出方法依然在多个评价指标上优于现有方法, 体现出其在低标注场景下的强大建模能力与实用性。

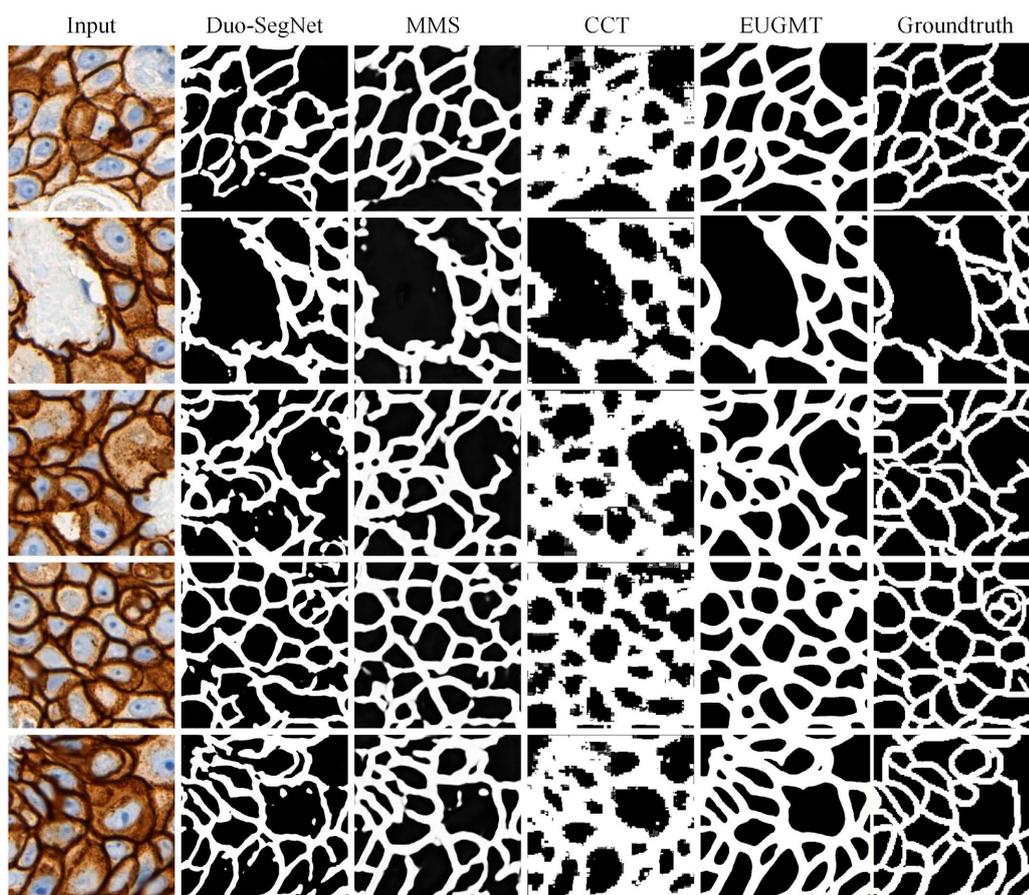


Figure 5. Comparison of different models at multiple locations on the HER2-stained histopathological image segmentation dataset
图 5. HER2 染色组织病理切片图像分割数据集多个位置的不同模型对比

从实验结果可见, 在仅使用 2% 标注数据的设定下, 本文方法在分割性能上显著优于现有方法。具体而言, 与经典的全监督方法 TransUnet 相比, 所提出模型在 Dice 指标上提升了 2.51%; 相比当前先进的半监督方法 MMS, 则提升了 7.2%。此外, 从图 5 中的可视化结果可以观察到, EUGMT 在细长的细胞膜区域展现出更完整且更贴合真实标签的分割效果, 尤其在结构边界细节处理上表现更为精准。总体而言, 2% 标注数据的实验设置本身具有极高挑战性, 但本文提出的模型在此条件下不仅明显优于最新的半监督学习方法, 同时也超越了多个全监督方法, 进一步验证了其在极低标注数据条件下的鲁棒性与优越性。

Table 4. Performance comparison of various models on the HER2-stained histopathological image dataset

表 4. 各类模型分割 HER2 染色组织图片数据集的性能对比

方法	Dice	IoU	ASD	HD
Duo-SegNet	0.6921	0.5291	55.48	328.74
MMS	0.6589	0.4912	2.57	19.64
CCT	0.5394	0.4079	5.64	38.89
TransUnet	0.7058	0.5455	2.76	60.03
SwinUnet	0.6885	0.5249	2.85	62.34
FAFuse	0.7288	0.5767	5.68	30.61
TransFuse	0.7193	0.5651	6.28	33.55
DeepLabv3+ (100% 标注)	0.6752	0.5097	6.35	30.58
UCMT	0.7206	0.5633	5.21	34.52
EUGMT (DeepLabv3+)	0.7309	0.5760	5.18	28.16

4.4. 宫颈癌细胞核分割数据集分割实验

4.4.1. 宫颈癌细胞核分割消融实验和对比试验分析

我们注意到, 宫颈癌病理图像通常存在显著的颜色差异、染色不均以及复杂的病理结构变化, 此外, 大量细胞核粘连或重叠的现象进一步增加了分割任务的难度。如表 5 所示, 在消融实验中, 采用全监督 Deeplabv3+ 模型的 Dice 系数仅为 0.6995, 这可能表明在未标注数据中仍蕴含着尚未利用的关键特征信息。相比之下, 在仅使用 2% 标注数据的条件下, 我们提出的模型以 Deeplabv3+ 作为主干网络, Dice 系数可达 0.8502, 显著优于全监督 Deeplabv3+ 模型, 同时也超过了 UCMT 模型。图 6 展示了对应实验条件下宫颈癌细胞核分割的可视化结果, 进一步支持模型在复杂结构下的适应能力。

此外, 如表 6 所示, 在使用 5% 标注数据时, 全监督 Deeplabv3+ 模型的 Dice 值略有提升, 达到 0.7024, 而我们的方法在采用 U-Net 作为主干网络时 Dice 值达到 0.8456, 表现出明显优势。这一结果说明, 我们提出的模型在有限标注条件下, 结合 U-Net 架构, 能够更有效地捕捉宫颈癌细胞核的结构特征。图 7 展示了使用 5% 标注数据下的分割可视化结果, 进一步验证了模型的有效性。

Table 5. Ablation study on cervical cancer nucleus segmentation with 2% labeled data

表 5. 2% 标注数据宫颈癌细胞核分割消融实验

方法	Dice	IoU	ASD	HD
DeepLabv3+ (SupOnly)	0.6951	0.5327	60.55	342.32
UCMT	0.7206	0.5633	6.23	29.00
EUGMT (DeepLabv3+)	0.7309	0.5760	5.71	27.31
EUGMT (Unet)	0.7092	0.5494	6.32	32.52

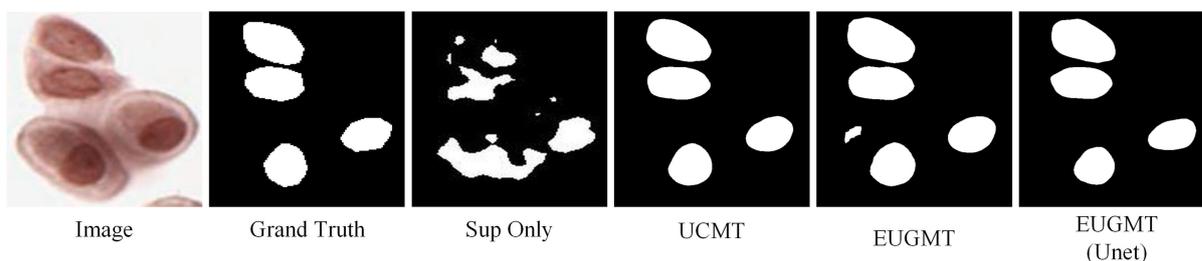


Figure 6. Visualization of ablation study results for cervical cancer nucleus segmentation with 2% labeled data

图 6. 2%标注数据宫颈癌细胞核分割消融实验效果图

综上所述结果表明, 本文提出的方法在面对标注数据极为有限且图像结构复杂的宫颈癌病理切片任务中, 依然能够实现高质量的细胞核分割, 具有良好的鲁棒性与应用潜力。

Table 6. Ablation study on cervical cancer nucleus segmentation with 5% labeled data

表 6. 5%标注数据宫颈癌细胞核分割消融实验

方法	Dice	IoU	ASD	HD
DeepLabv3+ (SupOnly)	0.7024	0.5415	9.56	24.55
UCMT	0.8371	0.7194	2.45	15.76
EUGMT (DeepLabv3+)	0.8433	0.7292	2.23	14.96
EUGMT (Unet)	0.8456	0.7322	1.56	10.89

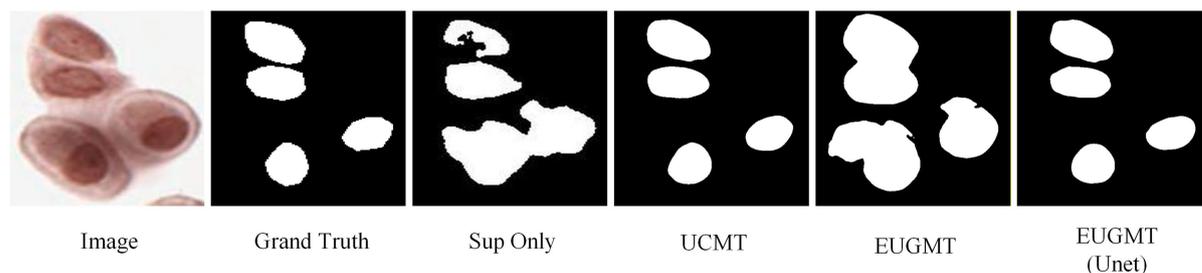


Figure 7. Visualization of ablation results for cervical cancer nucleus segmentation with 5% labeled data

图 7. 5%标注数据宫颈癌细胞核分割消融实验效果图

Table 7. Comparative experiments on cervical cancer nucleus segmentation

表 7. 宫颈癌细胞核分割对比试验

方法	Dice	IoU	ASD	HD
Duo-SegNet	0.6333	0.4634	9.37	91.00
MMS	0.6673	0.5007	4.95	83.25
CCT	0.5699	0.4709	3.27	18.78
TransUnet	0.7702	0.6264	2.95	19.31
SwinUnet	0.6491	0.4807	3.25	21.62
FAFuse	0.6665	0.5001	3.41	23.78
TransFuse	0.6750	0.5094	4.61	33.56
EUGMT (DeepLabv3+)	0.8433	0.7292	2.23	14.96
EUGMT (Unet)	0.8456	0.7322	1.56	10.89

如表 7 所示, 在 2% 标注数据的对比实验中, 部分经典的全监督模型在性能上超过了部分半监督方法。其中, CCT 模型的 Dice 值仅为 0.5699, 表现相对较差, 这可能是由于该模型并非专为医学图像分割任务设计, 难以有效提取细胞核的细粒度结构信息。而其他半监督模型未能取得理想性能的原因, 可能在于极少量的宫颈癌细胞核标注样本无法提供足够的监督信号, 从而限制了模型对关键特征的学习能力, 导致其在结构复杂、形态多样的病理图像中难以达到全监督模型的分割效果。

综合消融实验与对比实验的结果可以看出, 所提出的模型在跨数据集评估中依然保持了优越的性能表现, 证明其在其他类型医学图像数据上同样具备良好的泛化能力与适应性。

5. 结论与讨论

HER2 染色组织病理图像中细胞膜的准确分割在乳腺癌的辅助诊断中具有关键意义。为此, 本文提出了一种基于半监督学习的分割框架——增强不确定性引导的均值教师模型(Enhanced Uncertainty-Guided Mean Teacher, EUGMT), 用于实现 HER2 染色图像中细胞膜区域的精确分割。EUGMT 的核心优势体现在以下几个方面: 首先, 以均值教师(Mean Teacher)为基础构建半监督分割框架, 使得在极少标注样本(如 2%)的条件下, 模型性能仍可超越多个经典全监督分割方法。其次, 引入了不确定性选择模块, 该模块结合类别条件识别系统性标签噪声, 并对筛选出的高置信像素区域计算一致性损失, 从而提升模型对噪声扰动和复杂结构的鲁棒性。最后, 模型能够充分挖掘大量未标注数据中的潜在信息, 有效筛选并利用具有高训练价值的目标区域, 进一步提升分割效果。在私有数据集与公开数据集上的消融实验与对比实验充分验证了所提出方法的有效性, 表明 EUGMT 在极低标注数据条件下依然能够实现当前先进水平的分割性能。未来的研究将继续聚焦于高价值信息的筛选策略, 以进一步提升模型在半监督医学图像分割中的效率与泛化能力。

参考文献

- [1] Engel, R.H. and Kaklamani, V.G. (2007) HER2-Positive Breast Cancer. *Drugs*, **67**, 1329-1341. <https://doi.org/10.2165/00003495-200767090-00006>
- [2] Saha, M., Agarwal, S., Arun, I., Ahmed, R., Chatterjee, S., Mitra, P., et al. (2015) Histogram Based Thresholding for Automated Nucleus Segmentation Using Breast Imprint Cytology. In: Gupta, S., et al., Eds., *Advancements of Medical Electronics*, Springer India, 49-57. https://doi.org/10.1007/978-81-322-2256-9_5
- [3] Saha, M., Arun, I., Agarwal, S., Ahmed, R., Chatterjee, S. and Chakraborty, C. (2017) Imprint Cytology-Based Breast Malignancy Screening: An Efficient Nuclei Segmentation Technique. *Journal of Microscopy*, **268**, 155-171. <https://doi.org/10.1111/jmi.12595>
- [4] Banerjee, S., Saha, M., Arun, I., Basak, B., Agarwal, S., Ahmed, R., et al. (2017) Near-Set Based Mucin Segmentation in Histopathology Images for Detecting Mucinous Carcinoma. *Journal of Medical Systems*, **41**, Article No. 144. <https://doi.org/10.1007/s10916-017-0792-6>
- [5] Elston, C.W. and Ellis, I.O. (1991) Pathological Prognostic Factors in Breast Cancer. I. The Value of Histological Grade in Breast Cancer: Experience from a Large Study with Long-Term Follow-Up. *Histopathology*, **19**, 403-410. <https://doi.org/10.1111/j.1365-2559.1991.tb00229.x>
- [6] Liu, X., Song, L., Liu, S. and Zhang, Y. (2021) A Review of Deep-Learning-Based Medical Image Segmentation Methods. *Sustainability*, **13**, Article No. 1224. <https://doi.org/10.3390/su13031224>
- [7] Xu, Z., Wang, Y., Lu, D., Luo, X., Yan, J., Zheng, Y., et al. (2023) Ambiguity-Selective Consistency Regularization for Mean-Teacher Semi-Supervised Medical Image Segmentation. *Medical Image Analysis*, **88**, Article ID: 102880. <https://doi.org/10.1016/j.media.2023.102880>
- [8] Laine, S. and Aila, T. (2016) Temporal Ensembling for Semi-Supervised Learning. arXiv preprint arXiv:1610.02242.
- [9] Li, X., Yu, L., Chen, H., Fu, C., Xing, L. and Heng, P. (2021) Transformation-Consistent Self-Ensembling Model for Semisupervised Medical Image Segmentation. *IEEE Transactions on Neural Networks and Learning Systems*, **32**, 523-534. <https://doi.org/10.1109/tnnls.2020.2995319>
- [10] Ouali, Y., Hudelot, C. and Tami, M. (2020) Semi-Supervised Semantic Segmentation with Cross-Consistency Training. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 14-19 June 2020, 12674-

12684. <https://doi.org/10.1109/cvpr42600.2020.01269>
- [11] Wu, Z., Shi, X., Lin, G. and Cai, J. (2021) Learning Meta-Class Memory for Few-Shot Semantic Segmentation. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, 11-17 October 2021, 517-526. <https://doi.org/10.1109/iccv48922.2021.00056>
- [12] Xia, Y., Liu, F., Yang, D., Cai, J., Yu, L., Zhu, Z., *et al.* (2020) 3D Semi-Supervised Learning with Uncertainty-Aware Multi-View Co-Training. 2020 *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Snowmass Village, 1-5 March 2020, 3646-3655. <https://doi.org/10.1109/wacv45572.2020.9093608>
- [13] Tarvainen, A. and Valpola, H. (2017) Mean Teachers Are Better Role Models: Weight-Averaged Consistency Targets Improve Semi-Supervised Deep Learning Results. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 1195-1204.
- [14] You, C., Zhou, Y., Zhao, R., Staib, L. and Duncan, J.S. (2022) SimCVD: Simple Contrastive Voxel-Wise Representation Distillation for Semi-Supervised Medical Image Segmentation. *IEEE Transactions on Medical Imaging*, **41**, 2228-2237. <https://doi.org/10.1109/tmi.2022.3161829>
- [15] Samuli, L. and Timo, A. (2017) Temporal Ensembling for Semi-Supervised Learning. *International Conference on Learning Representations (LCLR)*, Vol. 4, 6.
- [16] Qiao, S., Shen, W., Zhang, Z., Wang, B. and Yuille, A. (2018) Deep Co-Training for Semi-Supervised Image Recognition. *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, 8-14 September 2018, 135-152.
- [17] Chen, X., Yuan, Y., Zeng, G. and Wang, J. (2021) Semi-Supervised Semantic Segmentation with Cross Pseudo Supervision. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 19-25 June 2021, 2613-2622. <https://doi.org/10.1109/cvpr46437.2021.00264>
- [18] Bai, W., Oktay, O., Sinclair, M., Suzuki, H., Rajchl, M., Tarroni, G., *et al.* (2017) Semi-Supervised Learning for Network-Based Cardiac MR Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention MICCAI 2017: 20th International Conference*, Quebec City, 11-13 September 2017, 253-260. https://doi.org/10.1007/978-3-319-66185-8_29
- [19] Zhang, W., Zhu, L., Hallinan, J., Zhang, S., Makmur, A., Cai, Q., *et al.* (2022) BoostMIS: Boosting Medical Image Semi-Supervised Learning with Adaptive Pseudo Labeling and Informative Active Annotation. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 20666-20676. <https://doi.org/10.1109/cvpr52688.2022.02001>
- [20] Zheng, H., Motch Perrine, S.M., Pitirri, M.K., Kawasaki, K., Wang, C., Richtsmeier, J.T., *et al.* (2020) Cartilage Segmentation in High-Resolution 3D micro-CT Images via Uncertainty-Guided Self-Training with Very Sparse Annotation. In: Martel, A.L., *et al.*, Eds., *International Conference on Medical Image Computing and Computer Assisted Intervention*, Springer International Publishing, 802-812. https://doi.org/10.1007/978-3-030-59710-8_78
- [21] Weng, L. (2019) From GAN to WGAN.
- [22] Wu, Y., Xu, M., Ge, Z., Cai, J. and Zhang, L. (2021) Semi-Supervised Left Atrium Segmentation with Mutual Consistency Training. In: de Bruijne, M., *et al.*, Eds., *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer International Publishing, 297-306. https://doi.org/10.1007/978-3-030-87196-3_28
- [23] Wu, Y., Ge, Z., Zhang, D., Xu, M., Zhang, L., Xia, Y., *et al.* (2022) Mutual Consistency Learning for Semi-Supervised Medical Image Segmentation. *Medical Image Analysis*, **81**, Article ID: 102530. <https://doi.org/10.1016/j.media.2022.102530>
- [24] Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., *et al.* (2021) A Review of Uncertainty Quantification in Deep Learning: Techniques, Applications and Challenges. *Information Fusion*, **76**, 243-297. <https://doi.org/10.1016/j.inffus.2021.05.008>
- [25] Gal, Y. and Ghahramani, Z. (2015) Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. Cornell University.
- [26] Zheng, Z. and Yang, Y. (2021) Rectifying Pseudo Label Learning via Uncertainty Estimation for Domain Adaptive Semantic Segmentation. *International Journal of Computer Vision*, **129**, 1106-1120. <https://doi.org/10.1007/s11263-020-01395-y>
- [27] Yu, L., Wang, S., Li, X., Fu, C. and Heng, P. (2019) Uncertainty-Aware Self-Ensembling Model for Semi-Supervised 3D Left Atrium Segmentation. In: Shen, D.G., *et al.*, Eds., *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer International Publishing, 605-613. https://doi.org/10.1007/978-3-030-32245-8_67
- [28] Wang, T., Lu, J., Lai, Z., Wen, J. and Kong, H. (2022) Uncertainty-Guided Pixel Contrastive Learning for Semi-Supervised Medical Image Segmentation. *Proceedings of the 31st International Joint Conference on Artificial Intelligence*, Vienna, 23-29 July 2022, 1444-1450. <https://doi.org/10.24963/ijcai.2022/201>
- [29] Rizve, M., Duarte, K., Rawat, Y. and Shah, M. (2021) In Defense of Pseudo-Labeling: An Uncertainty-Aware

-
- Pseudolabel Selection Framework for Semi-Supervised Learning.
- [30] Zhao, J., He, Y., Zhou, S., Qin, J. and Xie, Y. (2023) CNSeg: A Dataset for Cervical Nuclear Segmentation. *Computer Methods and Programs in Biomedicine*, **241**, Article ID: 107732. <https://doi.org/10.1016/j.cmpb.2023.107732>
 - [31] Peiris, H., Chen, Z., Egan, G. and Harandi, M. (2021) Duo-SegNet: Adversarial Dual-Views for Semi-Supervised Medical Image Segmentation. In: de Bruijne, M., *et al.*, Eds., *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer International Publishing, 428-438. https://doi.org/10.1007/978-3-030-87196-3_40
 - [32] Lou, A., Tawfik, K., Yao, X., Liu, Z. and Noble, J. (2023) Min-Max Similarity: A Contrastive Semi-Supervised Deep Learning Network for Surgical Tools Segmentation. *IEEE Transactions on Medical Imaging*, **42**, 2832-2841. <https://doi.org/10.1109/tmi.2023.3266137>
 - [33] Chen, J., Lu, Y., Yu, Q., *et al.* (2021) Transunet: Transformers Make Strong Encoders for Medical Image Segmentation.
 - [34] Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., *et al.* (2023) Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation. In: Karlinsky, L., *et al.*, Eds., *European Conference on Computer Vision*, Springer, 205-218. https://doi.org/10.1007/978-3-031-25066-8_9
 - [35] Zhang, Y., Liu, H. and Hu, Q. (2021) Transfuse: Fusing Transformers and CNNs for Medical Image Segmentation. In: de Bruijne, M., *et al.*, Eds., *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer International Publishing, 14-24. https://doi.org/10.1007/978-3-030-87193-2_2
 - [36] Xu, S., Xiao, D., Yuan, B., Liu, Y., Wang, X., Li, N., *et al.* (2023) FAFuse: A Four-Axis Fusion Framework of CNN and Transformer for Medical Image Segmentation. *Computers in Biology and Medicine*, **166**, Article ID: 107567. <https://doi.org/10.1016/j.compbiomed.2023.107567>