

学校代码	10699
分类号	O242
密 级	公开
学 号	xxx

题目 半监督遥感影像
 变化检测算法研究

作者 xxx

学科专业 计算机技术
指导教师 xxx
培养单位 计算机学院
申请日期 2025年3月

西 北 工 业 大 学

硕 士 学 位 论 文

题目：半监督遥感影像
变化检测算法研究

学科专业：计算机技术

作 者：xxx

指导教师：xxx

2025 年 3 月

Research on semi-supervised remote Sensing image change detection algorithm

By

Dongcheng Wen

Under the Supervision of Professor

Lingyan Ran

A Dissertation Submitted to

Northwestern Polytechnical University

In Partial Fulfillment of The Requirement

For The Degree of

Master of Computer technology

Xi'an, P.R. China

March 2025

学位论文评阅人和答辩委员会名单

学位论文评阅人名单

姓名	职称	工作单位
全盲评阅	无	无

答辩委员会名单

答辩日期	2023 年 x 月 y 日		
答辩委员会	姓名	职称	工作单位
主席	赵钱孙	教授	西北工业大学
委员	周吴郑	教授	西北工业大学
委员	冯陈褚	教授	西北工业大学
委员	蒋沈韩	教授	西北工业大学
委员	朱秦尤	教授	西北工业大学
委员	何吕施	教授	西北工业大学
委员	孔曹严	教授	西北工业大学
秘书	金魏陶	教授	西北工业大学

摘要

遥感影像变化检测是遥感领域的重要研究课题，在许多领域发挥着重要作用，例如，城市建设规划、森林环境保护、农村土地管理、自然灾害评估等。虽然基于深度学习的变化检测方法相比人工目视和传统机器学习方法已经在检测效率和准确性上取得了巨大的成功，但是其高度依赖于大量的标注训练数据，当有标注的训练样本减少时，模型的识别能力急剧下降；而且，变化检测任务的数据标注非常复杂，需要高精度几何图像配准和像素级精细标注，耗时耗力。而在如今非常成熟的对地观测技术之下，无标注的数据已经很容易获得，半监督学习可以有效利用这些大量的无标注样本进行训练。因此，本文着力研究在有限标注数据和大量无标注数据的情况下，从伪标签阈值、伪标签改造以及扰动一致性三个角度设计了半监督变化检测算法。本文的主要研究内容如下：

1) 针对目前基于伪标签和一致性正则化的方法中，采用固定阈值或者特定阈值调整方案可能无法更有效地利用未标记数据的问题。我们提出了基于自适应动态阈值的半监督变化检测算法——AdTSemiCD，算法根据模型的预测置信度来把握模型的学习状态，并据此自适应地调整阈值。此外，针对被筛选掉的标签，我们还设计了一个低置信度学习模块来学习其中可能包含的隐藏信息，进一步提高训练效率。大量的实验表明了这种自适应阈值的优越性，特别是在标记数据极其稀少的情况下。

2) 针对模型为具有不同难易程度的无标记样本生成的伪标签存在不同程度噪声，以及对模型训练的指导作用也存在差异性的问题，我们提出了基于伪标签评估的自适应半监督变化检测算法——AdaSemiCD。首先，我们提出了 AdaFusion 模块旨在单个样本水平上对不确定性高的样本区域进行改造，从而提高伪标签的准确性。其次提出了 AdaEMA 模块，引入了模型参数更新层面的自适应选择过程，使得模型能够充分集成更优参数。大量的实验结果表明了我们所提出的方法能够极大地改善伪标签的质量，从而提高变化检测的性能。

3) 针对目前基于一致性正则化的半监督变化检测算法中对所有样本以同样策略进行扰动会扩大模型的确认偏差的问题，以及基于硬伪标签自训练方法中存在的不可避免的噪声问题，我们提出了基于自适应特征扰动的半监督变化检测算法——AdPSemiCD。我们基于模型在无标记样本上的预测对样本进行评估，从而对难、易样本特征进行自适应扰动，包括扰动数量自适应选择，以及扰动强度自适应选择，通过这种样本级的定制化扰动方案，达到扩充特征空间和减小确认偏差的平衡，为模型训练提供更加有效的一致性约束。最后通过实验证明了这种自适应特征扰动方法的有效性。

关键词：变化检测；半监督学习；伪标签；自适应

Abstract

Remote sensing image change detection is an important research topic in the field of remote sensing, playing a vital role in various domains such as urban construction planning, forest environmental protection, rural land management, and natural disaster assessment. Although deep learning-based change detection methods have achieved significant success in terms of detection efficiency and accuracy compared to manual visual inspection and traditional machine learning methods, they heavily rely on large amounts of labeled training data. When the number of labeled training samples decreases, the model's recognition ability sharply declines. Furthermore, the data labeling for change detection tasks is very complex, requiring high-precision geometric image registration and pixel-level detailed labeling, which is time-consuming and labor-intensive. With the now highly mature earth observation technologies, unlabeled data is readily available, and semi-supervised learning can effectively leverage these large amounts of unlabeled samples for training. Therefore, this paper focuses on designing a semi-supervised change detection algorithm under the condition of limited labeled data and abundant unlabeled data, from three perspectives: pseudo-label threshold, pseudo-label refinement, and perturbation consistency. The main research contents of this paper are as follows:

1) To address the issue that methods based on pseudo-labels and consistency regularization, which use fixed or specifically adjusted thresholds, may not effectively utilize unlabeled data, we propose an adaptive dynamic threshold-based semi-supervised change detection algorithm—AdTSemiCD. The algorithm adjusts the threshold adaptively according to the model's prediction confidence, which allows the model to better grasp its learning status. Additionally, for the labels that are filtered out, we design a low-confidence learning module to capture the potential hidden information, further improving training efficiency. Extensive experiments demonstrate the superiority of this adaptive threshold, especially when labeled data is extremely scarce.

2) To solve the problem that pseudo-labels generated for unlabeled samples with different difficulty levels may contain varying amounts of noise and provide inconsistent guidance for model training, we propose an adaptive semi-supervised change detection algorithm based on pseudo-label evaluation—AdaSemiCD. First, we introduce the AdaFusion module, which aims to refine highly uncertain sample regions at the individual sample level, thus improving the accuracy of pseudo-labels. Second, we propose the AdaEMA module, which incorporates an adaptive selection process in model parameter updates, allowing the model to fully integrate better parameters. Extensive experimental results show that our proposed method significantly improves the quality of pseudo-labels, thus enhancing the performance of change detection.

3) To address the issue in current consistency regularization-based semi-supervised change detection algorithms, where applying the same perturbation strategy to all samples may exacerbate the model’ s confirmation bias, and the unavoidable noise in self-training methods using hard pseudo-labels, we propose an adaptive feature perturbation-based semi-supervised change detection algorithm—AdPSemiCD. We evaluate the samples based on the model’s predictions on unlabeled data, allowing for adaptive perturbation of features for difficult and easy samples. This includes the adaptive selection of the number of perturbations and the intensity of perturbations. This sample-level customized perturbation scheme aims to balance the expansion of the feature space and the reduction of confirmation bias, providing more effective consistency constraints for model training. Finally, experiments validate the effectiveness of this adaptive feature perturbation method.

Key Words: change detection; semi-supervised learning; pseudo-label; adaptive

目 录

摘要.....	I
ABSTRACT	III
图目录.....	VII
表目录.....	IX
第1章 绪论.....	1
1.1 研究背景与意义	1
1.2 国内外研究现状及趋势	3
1.2.1 半监督学习算法.....	3
1.2.2 基于深度学习的遥感影像变化检测算法.....	4
1.2.3 半监督遥感影像变化检测算法	7
1.3 本文主要内容及结构安排.....	8
第2章 相关技术	11
2.1 深度神经网络.....	11
2.1.1 ResNet	11
2.1.2 Vision Transformer.....	12
2.1.3 金字塔池化模型.....	14
2.1.4 平均教师模型	15
2.2 经典半监督变化检测算法	16
2.2.1 SemiCDNet.....	16
2.2.2 RCR.....	18
2.2.3 FPA	20
2.3 实验指标.....	23
2.4 实验数据集介绍	24
2.5 本章小结.....	27
第3章 基于自适应动态阈值的半监督变化检测算法	29
3.1 引言	29
3.2 基于自适应动态阈值的半监督变化检测框架.....	30
3.2.1 整体框架	30
3.2.2 全局自适应阈值.....	32
3.2.3 类别自适应阈值.....	33
3.2.4 低置信度学习	34

3.3 实验结果及分析	34
3.3.1 对比试验	34
3.3.2 消融实验	40
3.4 本章小结	43
第 4 章 基于伪标签评估的自适应半监督变化检测算法	44
4.1 引言	44
4.2 基于伪标签评估的自适应半监督变化检测框架	45
4.2.1 整体框架	45
4.2.2 伪标签评估指标设计	46
4.2.3 自适应样本融合机制	47
4.2.4 自适应师生模型参数更新机制	48
4.3 实验结果及分析	49
4.3.1 实验设置	49
4.3.2 对比试验	49
4.3.3 定性对比	54
4.3.4 消融实验	56
4.3.5 局限性分析	60
4.4 本章小结	60
第 5 章 基于自适应扰动的半监督变化检测算法	61
5.1 引言	61
5.2 基于自适应扰动的半监督变化检测框架	62
5.2.1 整体框架	62
5.2.2 自适应特征扰动	63
5.3 实验结果及分析	64
5.3.1 实验设置	64
5.3.2 对比实验	65
5.3.3 消融实验	72
5.4 本章小结	75
第 6 章 总结与展望	77
6.1 本文工作总结	77
6.2 未来研究展望	78
参考文献	79
致 谢	89
在学期间发表的学术成果和参加科研情况	91

图目录

1-1 各种应用领域中的变化检测	2
1-2 本文研究内容和组织结构对应关系图	8
2-1 不同深度的 ResNet 网络层次架构 ^[1]	12
2-2 两种残差块的网络结构 ^[1]	12
2-3 Vision Transformer 网络框架 ^[2]	13
2-4 金字塔池化模型架构 ^[3]	14
2-5 平均教师模型的网络结构 ^[4]	15
2-6 SemiCDNet 网络框架 ^[5]	16
2-7 RCR 网络框架 ^[6]	18
2-8 FPA 网络框架 ^[7]	21
2-9 FPA 使用的强增强操作列表 ^[7]	22
2-10 建筑物变化检测数据集部分样例图	25
2-11 多类变化检测数据集部分样例图	26
3-1 AdTSemiCD 算法的整体框架	30
3-2 训练过程中的全局阈值和两种类别的置信度变化情况	32
3-3 公开变化检测数据集的类别不平衡统计	33
3-4 AdTSemiCD 与 SOTA 方法在 10 个数据集上的性能对比折线图	35
3-5 5% 标记训练样本下 AdTSemiCD 与 SOTA 方法在 WHU-CD 测试集上的可视化对比图	38
3-6 5% 标记训练样本下 AdTSemiCD 与 SOTA 方法在 LEVIR-CD 测试集上的可视化对比图	39
3-7 不同标记率下 AdTSemiCD 在 LEVIR-CD 测试集上的可视化对比图	40
3-8 5% 标记训练样本下 AdTSemiCD 与 SOTA 方法在 CDD 测试集上的可视化对比图	40
3-9 5% 标记训练样本下 AdTSemiCD 与 SOTA 方法在 CL-CD 测试集上的可视化对比图	41
3-10 训练过程中的伪标签利用率	43
4-1 AdasemiCD 的整体框架图	45
4-2 AdaSemiCD 与 SOTA 方法在 10 个数据集上的性能对比折线图	50

4-3 AdaSemiCD 与 SOTA 方法在 LEVIR-CD 测试集上的变化检测结果可视化对比	54
4-4 AdaSemiCD 与 SOTA 方法在 WHU-CD 测试集上的变化检测结果可视化对比	55
4-5 AdaSemiCD 与 SOTA 方法在 CDD 测试集上的变化检测结果可视化对比	55
4-6 AdaSemiCD 与 SOTA 方法在 CL-CD 测试集上的变化检测结果可视化对比	56
4-7 类别重平衡对训练中未标记样本生成的伪标签质量的影响	58
4-8 AdASemiCD 的一些失效场景	59
5-1 AdPSemiCD 算法的整体框架	62
5-2 AdPSemiCD 与 SOTA 方法在 10 个数据集上的性能对比折线图	65
5-3 5% 标记训练样本下 AdPSemiCD 与 SOTA 方法在 LEVIR-CD 测试集上的可视化对比图	69
5-4 不同标记率下 AdPSemiCD 在 LEVIR-CD 测试集上的可视化对比图	69
5-5 5% 标记训练样本下 AdPSemiCD 与 SOTA 方法在 WHU-CD 测试集上的可视化对比图	70
5-6 5% 标记训练样本下 AdPSemiCD 与 SOTA 方法在 CDD 测试集上的可视化对比图	71
5-7 5% 标记训练样本下 AdPSemiCD 与 SOTA 方法在 CL-CD 测试集上的可视化对比图	72
5-8 LEVIR-CD 测试集上的 t-SNE 特征可视化对比	74
5-9 WHU-CD 测试集上的 t-SNE 特征可视化对比	74
5-10 CDD 测试集上的 t-SNE 特征可视化对比	75
5-11 CL-CD 测试集上的 t-SNE 特征可视化对比	75

表目录

2-1 本文所使用的公开数据集	24
3-1 不同 CD 方法在二值建筑物变化检测数据集上的平均定量性能对比	36
3-2 不同 CD 方法在多类变化检测数据集上的平均定量性能对比	37
3-3 在 LEVIR-CD 数据集上的自适应动态阈值消融研究	40
3-4 AdTSemiCD 在 LEVIR-CD 数据集上的消融实验	41
3-5 损失权重消融实验结果	42
4-1 不同 CD 方法在二值建筑物变化检测数据集上的平均定量性能对比	51
4-2 不同 CD 方法在多类变化检测数据集上的平均定量性能对比	53
4-3 AdaSemiCD 在 LEVIR-CD 数据集上的消融研究	57
4-4 在 LEVIR-CD 数据集上伪标签量化指标的消融研究	57
4-5 超参数敏感性消融实验结果	59
4-6 LEVIR-CD 数据集上不同方法的参数、计算复杂度和训练时间的比较	59
5-1 AdPSemiCD 使用的特征扰动方法	64
5-2 不同 CD 方法在建筑物变化检测数据集上的定量性能对比	67
5-3 不同 CD 方法在多类变化检测数据集上的平均定量性能对比	68
5-4 AdPSemiCD 在 WHU-CD 数据集上的消融实验	71
5-5 超参数敏感性消融实验结果	73

第1章 绪论

1.1 研究背景与意义

随着科技水平的不断进步，人类的生产、生活对于自然界和人类世界都以更快的速度发挥着更重要的影响力，在以往可能经过数十年乃至几个世纪间的自然演变过程才造成的地球地形地貌变化，例如河流改道、填海填湖，在如今或许被缩短至数年甚至数月，观测和把握这种变化对我们分析和指导生产活动是一件非常重要的事情。目前相关的对地观测技术也得到了飞速的发展，人们借助于高空无人机和遥感卫星的对地传感器能够轻松完成对地球表面的信息采集，并实时返回遥感影像数据，这已经逐渐成为了人们了解和观测地球的主要方式。并且现代遥感成像技术的成熟也使得采集到的遥感图像具有较高的空间分辨率，从而为人们动态检测地表变化提供了可能性和便利性。

遥感影像变化检测 (Remote sensing change detection, RSCD) 任务，本质上是一个二分类的问题，就是在无人机或者卫星对于同一区域在不同时间拍摄的双时相图像对中，识别出感兴趣的目标变化区域，比如建筑物、水域、植被和道路等，如图1-1中所示。该技术在许多领域中发挥着重要作用，例如，城市建设规划^{[8][9][10]}、森林环境保护^[11]、农村土地管理^{[12][13][14]}、自然灾害评估^{[15][16]}等。在智能影像分析技术的广泛应用之前，人类主要依靠人工目视法来进行变化检测并手动标注变化区域和类型。这种方法虽然可靠，但是依赖于专业研究人员的检测经验，并且在面对海量任务时，这种方法的可行性和经济性就受到了极大的挑战。伴随计算机科学的进步和机器学习的兴起，自动变化检测开始走入了人们的视线。早期广泛采用的基于传统机器学习算法的变化检测方法，主要包括：(1) 基于图像差分^[17]、图像回归^[18]、图像比例^[17]、变化向量分析^[19](Change vector analysis, CVA) 等代数方法；(2) 基于变换的方法，如主成分分析^[20](Principle component analysis, PCA)、独立成分分析^[21] (Independent Component Analysis, ICA) 多元变化检测^[22](Multivariate alteration detection, MAD)、慢特征分析^[23] (Slow feature analysis, SFA) 等，通过将高维特征投影到低维特征空间中，使特征分量去相关，从而突出重要的变化信息表示。(3) 基于分类方法。例如，后分类比较方法，该方法首先对前后时间图像进行独立分类，然后逐像素比较两幅图像的分类结果，既可以回答“哪里发生了变化”的问题，也可以回答专家感兴趣的另一个问题“发生了什么变化”，但缺点也很明显，即高度依赖高精度高配准的分类结果，实施起来难度极大。

目前，深度学习算法已经被广泛应用到各种计算机视觉任务中，比如图像分类、目标检测、深度估计，以及语义分割、实例分割等与变化检测相似的密集预测任务，并且表现出了远超传统方法的优异性能。然而，基于卷积神经网络^{[24][25][26]} (Convolutional Neural Network, CNN) 或者 Transformer 网络^{[27][28]} 的全监督的深度学习变化检测算法都极其依赖于大量的人工标注，当有标注的训练样本减少时，模型的识别能力急剧下降。而且，

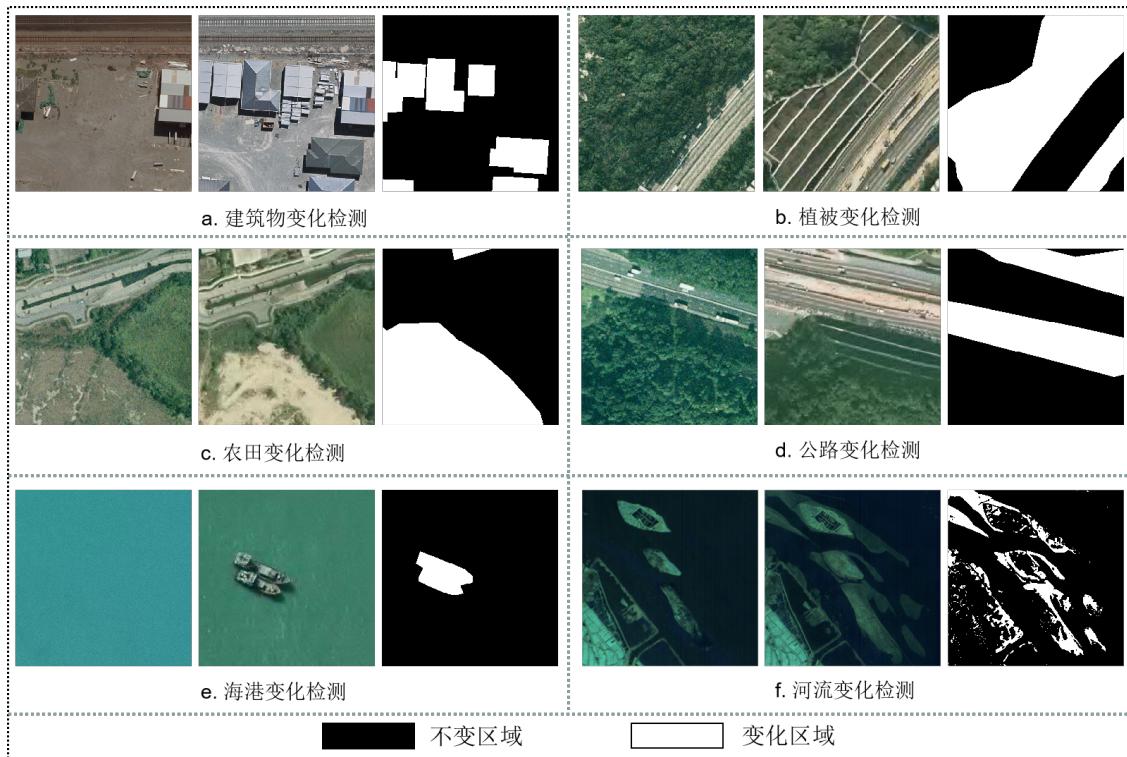


图 1-1 各种应用领域中的变化检测

变化检测任务的数据标注非常复杂，需要高精度几何图像配准和像素级精细标注，耗时耗力。这一点从目前可以获取到的公开数据集的数据规模对比就可以看出来，通常图像分类、目标检测的数据集能够达到数十万张，而变化检测的数据集规模仅在几百至几万张。为了应对这些挑战，研究人员已经探索了一系列方法，如弱监督变化检测^{[29][30][31]}（Weakly Supervised Change Detection, WSCD）、无监督变化检测^{[32][33][34]}（Unsupervised Change Detection, USCD）和半监督变化检测^{[5][6][7]}（Semi-supervised Change Detection, SSCD）。虽然弱监督变化检测具有一定的成本效益，但它依赖于不完整或不准确的标签，这可能会引入重大错误和不可预测的噪声。另一方面，无监督变化检测不使用任何标记数据，而是利用数据中存在的固有特性；然而，在处理分类或检测等特定任务时，它经常面临缺乏语义信息的挑战。还有一些方法采用样本生成策略，包括数据增强^{[35][36]}、生成对抗网络^[37]（Generative Adversarial Networks, GAN）和扩散模型^[38]，这些策略经常需要模拟或合成额外的数据。然而，当处理有限的可用样本时，由于生成的数据多样性不足，这些方法可能会遇到约束，这可能会降低模型的泛化能力。因此，半监督变化检测可能是一种更有效的解决方案。半监督学习^{[4][39][40]}（Semi-supervised Learning, SSL）旨在通过利用有限的可用标记数据和大量未标记样本来提高变化检测性能。通常，研究人员为未标记的数据生成伪标签，作为训练期间的指导，这些伪标签通常是具有较高概率的临时预测。最广泛使用的方法是平均教师^[4]（Mean-Teacher, MT）框架，它使用教师模型来生成伪标签，在训练过程中为学生模型提供指导。随后使用学生模型的指数移动平均^[41]（Exponential moving average, EMA）更新教师模型。学生模型受益于在有限

的标记数据和充足的伪标记数据的混合上进行训练，使其能够识别更重要的特征，从而显著提高性能。

在此背景之下，本文着力研究了半监督变化检测算法，针对模型为无标注样本生成的伪标签可能包含错误标签从而引入额外噪声这一问题，我们进行了一系列的研究，分别从伪标签筛选阈值、伪标签优化、扰动一致性等方面进行了有效的改进，从而提高了半监督变化检测算法的性能。

1.2 国内外研究现状及趋势

此前国内外学者已经在半监督变化检测算法领域进行了大量的研究，本小节将介绍与本文最相关的几个研究方向的研究现状，包括半监督学习算法，基于深度学习的遥感影像变化检测算法，以及半监督遥感影像变化检测算法的基本任务与方法和研究发展历程。

1.2.1 半监督学习算法

在实际应用场景中，无标签的数据易于获取，而有标签的数据收集起来通常很困难，标注过程也是一项极度劳动密集的工作。在这种情况下，半监督学习是一种克服样本标注困难问题的可行方法，近年来已经成为深度学习领域一个热门的研究方向，其旨在仅利用一小部分标记数据进行监督训练，学习到正确的语音信息，同时利用大量的无标注训练样本进行无监督训练，以提高模型的泛化性，减少过拟合现象。SSL 主要包含三种策略：一致正则化（Consistent Regularization, CR）、自训练（Self Training）、生成模型以及包括其中多种思想的整体方法。

一致性正则化方法基于平滑假设^[42] 和聚类假设^[43] 的思想，所谓的平滑假设，即假设如果两个输入样本相似，则它们的输出也应该相似。这意味着当两个输入样本属于同一类别并且位于同一个簇时，它们的预测输出应该相近。聚类假设则假设输入数据点形成簇，每个簇对应于一个输出类。如果点在同一个簇中，则它们可以认为属于同一类，它也可以被视为低密度分离假设，即给定的决策边界位于低密度地区。因此在一个未标记的数据上施加一定程度的扰动，则预测不应发生显著变化。

这种方法首先对输入数据施加不同程度的扰动，将模型在这些输入数据上的输出之间的一致性作为训练约束。目前的三个主流的一致性正则化训练框架包括： π -模型^[44]，时间集成模型^[44] 和平均教师模型^[4]。这几种框架都是以双分支网络作为基础架构，两个网络分支分别对扰动后的训练样本和原始训练样本进行前向推理，通过最小化两次推理的概率分布的均方误差（Mean Square Error, MSE）来优化模型参数。其中 π -模型的双分支网络之间是共享参数权重的，每次更新参数之后两个分支同步更新；时间集成模型合并时间序列上的所有输出结果，当前模型预测结果与历史预测结果的指数移动平均值做均方误差计算，这有效地保留历史了信息，消除了扰动并稳定了当前值；MT 模型从模型参数层面进行了平滑操作，学生模型的权重是历史模型参数的 EMA 集

成，该模型在随后的各个领域的半监督研究中得到了应用，例如半监督对象检测算法 Active-Teacher^[45]，半监督一般语义分割算法^{[46][47][48][49]}，半监督图像分类算法^[50]，半监督医学图像分割算法^{[51][52]}。此外有研究人员也在扰动设计方面进行了探索，谢等人^[53]和 Bandara 等人^[6]分别在一致性正则化中应用了图像级扰动和特征级扰动。

基于自训练的半监督学习算法思想是：首先使用一个预测模型或它的某些变体来为无标注样本生成伪标签，将这些伪标签和有标注的数据样本混合在一起进行监督训练，提供额外的训练信息。一个最为关键的问题就在于伪标签的可靠性，因为这直接影响了模型的训练效果，因此大量的研究从此处入手，研究了生成高质量伪标签的方法。赖等人^[54]采用一个设定的概率阈值作为选择标准，预测概率小于此阈值的类别标签直接被丢弃。ST++^[55]开发了一种多层自训练结构，每个阶段选择一批高质量的伪标签参与到训练，反复进行此挑选和训练过程，直到所有未标记的样本都得到了利用。U2PL^[56]计算出伪标签中每个像素的信息熵，使用一个恒定的熵值作为过滤阈值，排除那些不可靠的伪标签。此外还有一种协同训练（Co-training）框架^[57]，这种框架包含两个模型，分别为彼此提供伪标签。

基于生成模型的方法旨在利用生成模型对数据的分布进行建模，从而推断出未标注数据的潜在信息。半监督变分自编码器^[58]是变分自编码器（Variational Autoencoder, VAE）在半监督学习中的一种扩展方法，添加了一个分类器来学习语义信息，首先 VAE 从标注样本集中学习一个真实的数据分布，然后在大量的无标注样本上基于分类器分类结果去指导模型学习一个更为丰富的潜在分布。Ren 等人^[37]利用生成对抗策略训练了一个能够生成共配准图像的生成器，从原始图像高维特征空间中采样，从特征空间分布的上界和下界之间按照策略选择一些特征向量生成高质量配准图像，从而后续只需用通过特征融合方法即可找出变化区域，并且扩充了学习到的特征空间。Bandara 等人^[38]在大规模数据集上预训练了一个去噪扩散概率模型（Denoising Diffusion Probabilistic Models, DDPM），利用其具有强大表征能力的编码器进行双时图像的特征提取以提高变化检测性能。

通常情况下，实际应用中更多采用的是整体方法（Holistic Methods），即在一个框架中整合前述半监督学习中多种方法，从而获得更好的性能。其中最为经典的是 FixMatch^[50]，提出了一种基于伪标签和一致性正则化的简单而有效的方法。进一步，Yang 等人^[47]根据分类和分割任务的差异，增加了一种新的特征扰动前馈输入流，以及一个额外的强增强扰动输入流，通过多重一致性约束构造更广阔的扰动空间。此外，他们同期关于半监督语义分割的另外两项研究^{[48][49]}考虑了在半监督语义分割训练中加入样本级自适应调整机制，表现出了具有前景的性能。

1.2.2 基于深度学习的遥感影像变化检测算法

受深度学习在各领域取得的巨大成功所启发，近十年在变化检测领域也涌现出了许多经典的工作，推动了遥感影像变化检测研究的发展，下面本小节将按照大致的时间顺

序分别介绍基于 CNN 的变化检测方法、基于 Transformer 的变化检测方法以及近期出现的基于大模型的变化检测方法。

(1) 基于 CNN 的变化检测算法

CNN 引入了卷积和池化网络层，使其能够有效地捕获数据中的空间特征和局部关系，同时通过深度网络结构层层提取更抽象的高维特征，经典的 CNN 架构包括 LeNet-5^[59]、AlexNet^[60]、VGGNet^[61]、ResNet^[1] 以及 UNet^[62] 等。研究人员利用双分支孪生网络分别对两幅时相图像进行处理，提取高维特征，然后进行特征融合，识别出双时相图像对之间的差异。最为经典的工作是 Daudt 等人^[24] 使用全卷积网络构建了基于 UNet 的变化检测框架及其两个孪生变体，这三种变化检测框架分别是 FC-EF、FC-Siam-conc、FC-Siam-diff，每种框架都采用了不同的特征融合策略，其中 FC-EF 是在输入层面首先对双时相图像进行了图像级别的融合，另外两种都是对从双时相图像对中抽取的高维特征进行融合。Shi 等人^[63] 提出的 DSAMNet 在每个多尺度特征融合阶段添加了卷积块注意模块^[64]（Convolutional Block Attention Modules, CBAM），这种轻量级的注意力机制从空间和通道两个维度上对特征之间的关系进行了建模，动态地调整了特征图地权重。Zhang 等人^[65] 提出了一个互特征学习网络——MFNet（Mutual Feature-Aware Networks），提出的对称变化特征融合模块弥补了此前差分特征融合造成的信息丢失问题，同时通过在编码阶段提前引入差异感知，使得编码器更加聚焦于对潜在变化区域的特征学习。Fang 等人^[66] 将孪生网络和稠密连接的 NestedUNet 网络^[67] 结合到了一起，浅层特征和深层特征之间紧凑的信息传输减少了特征抽取中的位置信息丢失。Zheng 等人提出的 ChangeStar^[68] 通过构造伪配准图像对，以语义分割方式来处理两幅图像，以单时相图像来训练双时相图像对变化检测模型，减少了对特征融合和图像配准的依赖。总的来说，基于 CNN 的监督变化检测方法大多侧重于特征融合模块的设计或者精巧的编码器设计，以更加精确地表示变化特征。

(2) 基于 Transformer 的变化检测算法

Transformer 是一种基于自注意力机制的深度学习模型，最初由 Vaswani 等人在 2017 年提出^[69]，首次用于自然语言处理任务中的序列到序列（Seq2Seq）建模，尤其是在机器翻译任务中取得了革命性突破。其核心特性是完全抛弃了传统的循环神经网络^[70]（Recurrent Neural Network, RNN）和卷积神经网络，以自注意力机制和全连接网络为基础，显著提高了模型在长程依赖任务中的效率和性能。随着 Google Research 于 2020 年提出了 Vision Transformer (ViT)^[2]，首次将 Transformer 网络应用到了计算机视觉领域，其迅速成为了计算机视觉领域的新兴方向，在变化检测领域亦是如此。Chen 等人首先将 ViT 引入了变化检测任务，提出的 BIT^[27] 是一种基于原生 Transformer 架构的网络，旨在通过 Transformer 对双时相图像的时空上下文进行建模。然而，该模型直接采用原生 Transformer 的解码器，未充分利用浅层特征，这限制了其性能。相比之下，Li 等人提出的 TransUnet^[71] 引入了 UNet 风格的解码器，取代了 Transformer 原有的解码方式，

通过融合上一阶段的特征图逐步恢复至原始尺寸，从而显著提升了模型的表现。英伟达提出的 SegFormer^[72] 是一种基于 Transformer 的通用语义分割架构，利用轻量级编码器实现多尺度特征提取。受其启发，Bandara 等人设计了 ChangeFormer^[73]，该模型引入了针对变化检测任务的特定模块，通过卷积操作学习双时相特征图之间的变化关系。相比传统方法，ChangeFormer 展现了更高的精度和鲁棒性。Jiang 等人提出的 VcT^[74] 将每个像素作为一个图节点，利用图神经（GNN）网络对所有节点组成的结构化信息进行建模，挖掘具有共同背景信息的可靠 token，而不是像以往手动设置固定 token，更具针对性的特征学习提高了检测的效率和准确性。

一些研究发现，在变化检测中，由于数据有限，纯 Transformer 模型可能无法发挥其全部潜力。因此还有一些研究将 CNN 和 Transformer 结合在了一起，使得两种模型的优点能够互补，既可以利用 CNN 强大的局部信息捕捉能力，又能够利用 Transformer 兼顾对全局关系的建模，实现全局特征和局部特征学习的统一。比较经典的工作有 Jiang 提出的 MSFCTNet^[75] 和 Li 等提出的 MCTNet^[76] 和 ConvTransNet^[77]。这类方法总体思想基本一致，但这些研究人员分别从不同方向对 CNN 与 Transformer 的高效信息交互进行了改进和创新。

（3）基于大模型的变化检测算法

近年来，大模型（Large Models）的发展成为人工智能领域的核心热点之一。这些模型以大规模参数、海量数据和复杂架构为特征，在自然语言处理、计算机视觉、强化学习等领域展现出卓越性能。它们通常基于自注意力机制（Self-Attention）驱动的 Transformer 模型，通过在大规模数据集上边的自监督预训练，获得了强大的通用表征提取能力，虽然这些模型和上一小节同属 Transformer 类型，但由于其带来的变革性突破，本小节将单独分析基于大模型的变化检测研究。自从自然语言处理领域的 BERT、GPT 问世并引起广泛关注以后，计算机视觉领域的预训练大模型也迅速地出现在了人们的眼前。其中 SAM^[78]（Segment Anything Model）是图像分割通用大模型的开篇之作，以其强大的零样本推理能力和无需微调的交互式分割，被认为彻底颠覆了传统的深度学习方法。Li 等人^[79] 最早提出了结合大模型进行变化检测的新范式，提出了一个包含冻结基础模型（如 CLIP^[80]、SAM 等）、双时态适配分支（Bi-TAB）以及它们之间的桥接模块三部分的双时态适配网络（BAN），将大模型的大量先验知识注入到了变化检测模型。Ding 等人^[81] 通过训练一个轻量级的适配器网络（Adapter）来更具针对性地利用 SAM 在遥感影像场景中强大的视觉表示能力对双时相图像进行特征提取。Liu 等人^[82] 将大语言模型和变化检测模型结合起来，提出了一个变化检测智能体（Change-Agent），能够输入文本和图像两种模态的数据，按照输入指令交互式地检测感兴趣的变化区域。Dong 等人^[83] 同样基于 Clip 进行了多模态变化检测的研究，他们将图像-文本编码的结果与解码阶段的视觉特征相结合，从而增强了图像的语义。Zheng 等人^[84] 利用 SAM 强大的零样本扩展能力，基于构建的点提示以及 SAM 提取的特征空间在图像内和图像间的潜在

相似性，进行无需训练的零样本推理，首次提出并应用到了零样本变化检测任务。但是这和传统意义上的零样本任务任然有差别，就是需要人工对每幅图像进行点提示的标注。

1.2.3 半监督遥感影像变化检测算法

对大量图像进行变化检测的精细标注非常耗时，目前解决此问题的方法主要集中在半监督变化检测上。与第 1.2.1 小节的半监督算法研究类似，SSCD 也主要分为两大类别，其一是基于 CR 的方法，其二是基于 GAN 的方法。

在一致性正则化方面，在变化检测中最早引入平均教师模型是在 Bousias 等人^[85] 的研究工作中。然而，最初的实验结果并没有显示出相当大的潜力，因为与仅使用有限数量的标记数据进行完全监督学习的基准相比，这种 SSCD 方法存在不足。甚至随着真实标注数据越来越多，这种与全监督训练之间的性能差距仍在继续扩大。以此为基础，Mao 等人^[86] 分别对教师模型和学生模型的输入进行了强、弱增强操作。此外，他们制定了一个额外的教师虚拟对抗训练组件，以进一步减少伪标签噪音的负面影响。此外，有的其他半监督方法使用单模型或具有共享权重的双分支模型。如 Sun 等^[87] 引入了孪生网络。他们结合了基于伪标签的额外自训练，采用阈值过滤来消除低质量的伪标签。这种过滤背后的基本原理在于低置信度的伪标签引入的潜在噪声，这可能对自训练产生不利影响。Hafner 等人^[88] 提出了一种双任务 SSCD 框架，该框架结合了建筑物分割和变化检测这两个密切相关的下游任务。他们在孪生分割网络和变化检测网络产生的两个变化检测掩码之间设计了一种新的一致性约束。Bandara 等^[6] 探索了新的正则化项，即基于特征的扰动，在特征层面应用各种数据扰动来扩展一致性约束的分布空间。该方法充分利用了未标记样本中嵌入的信息，在最近的工作中，Zhang 等人^[7] 对未标记数据集施加了类一致性和特征一致性两个约束。通过将未标记样本在变化类和不变类上的特征表示对齐，模型可以从更接近真实分布的特征空间中学习，这使得他们极大地改进了 SSCD 的性能。

其他方法主要利用 GAN 这种生成模型，它最初是由 Goodfellow 在 2014 年提出的^[89]。一些方法使用 GAN 来学习接近真实标记数据的特征分布空间^{[5][90][91][92]}；另一部分使用 GAN 生成数据样本^{[93][37]}；在最近的工作中，比较值得注意的研究是 Wu 等人^[29] 提出了一种新的变化检测范式，将无监督、弱监督、区域监督和完全监督的变化检测统一到了一个端到端的框架中，在无监督变化检测中，主要目标被定义为最小化一个区域，使得生成网络可以在屏蔽该区域后产生与另一个时间的图像相似的图像。对于弱监督和新提出的区域监督变化检测任务，关键思想在于最小化一个区域，使判别网络在屏蔽该区域后无法区分真实不变图像对。虽然这些努力在 SSCD 中取得了巨大的成功，但 GAN 的高度不稳定训练使得超参数调整变得具有挑战性。此外，梯度消失的问题经常出现在训练阶段，以及如果没有实施额外的训练技术，鉴别器的强判别能力可能导致 GAN 的生成器和鉴别器之间的性能不平衡。因此，实现理想的最优场景是具有挑战性

的，这使得该方法的基本应用有些复杂。

1.3 本文主要内容及结构安排



图 1-2 本文研究内容和组织结构对应关系图

本文着力研究半监督变化检测任务，在少量有标注训练数据上学习正确的语义信息并在大量且易获得的无标注训练样本上学习到一个更加丰富的特征空间分布。其中一个关键问题就在于能否减少无标注样本训练中不可避免的噪声问题，这有两种解决方案：通过改善伪标签的质量来提高半监督变化检测的性能；从特征层面构造正则化项从而排除错误伪标签的重要影响。本文从这两点入手，研究内容和论文组织结构的对应关系如图1-2所示，其中主要研究内容如下：

2) 基于自适应动态阈值半监督变化检测。针对此前基于一致性正则化和自训练的变化检测算法中，使用固定阈值或者固定阈值调整方案筛选伪标签导致的无标记训练数据利用率低下的问题，我们提出了一种根据模型的学习状态自适应调整阈值的半监督变化检测算法，首先根据全局预测置信度随着训练进程自适应地增大整体阈值，并根据前景、背景类别置信度为两个类别分配更加精确的阈值，最后对于被筛选掉的标签，通过

低置信度学习模块，将其纳入模型训练之中，进一步提高训练样本利用率。实验表明我们提出的自适应阈值相比此前半监督变化检测中采用的阈值筛选方案，能够达到更好的变化检测性能。

2) 基于伪标签评估的自适应半监督变化检测。针对大量的无标注样本中不同样本个体之间存在很大的差异，模型为这些具有不同难易程度的样本生成的伪标签可靠性也不尽相同。本算法设计了一种自适应动态学习策略 AdaSemiCD，旨在提高伪标签的准确性并简化训练过程。我们的框架结合了传统的半监督训练方法，并辅以两个创新的功能模块 AdaFusion 和 AdaEMA。首先，我们利用 AdaFusion 在单个样本水平上对不确定性高的样本区域进行改造，从而提高伪标签的准确性。其次在 AdaEMA 模块中引入了模型级参数更新的自适应选择过程，使模型能够充分集成优越的参数。大量的实验结果表明了我们所提出的方法能够极大地改善伪标签的质量，使得变化检测性能更好。

3) 基于自适应特征扰动的半监督变化检测。针对可能对半监督学习的一致性原理和边缘假设造成挑战的两个问题：(1) 模型对未标记图像的预测并不总是足够准确的；(2) 对无标记图像施加的扰动并不总是对模型训练有效的。我们提出了一种根据样本质量进行自适应特征扰动的半监督变化检测算法，首先利用概率分布的对其替换单一硬标签学习，从而缓和伪标签的不确定性，并提高对软标签信息的利用；其次我们针对不同样本进行自适应地特征扰动，在随机扰动地基础上实现了一定的可控性，减少了模型的确认偏差。最后进行了大量的实验，验证了我们的设想的正确性和改进措施的有效性。

全文包括六章，具体的章节结构安排如下：

第一章首先系统性地阐述了变化检测任务的应用场景和研究价值，以及半监督变化检测算法的研究背景和意义，然后梳理了目前国内外关于半监督学习、深度变化检测、半监督变化检测的研究进展和主流方法，最后介绍了本文的主要研究内容和文章结构安排。

第二章首先介绍了本文研究工作中用到的一些骨干特征提取网络，之后对本研究工作中进行对比的几种经典半监督变化检测算法进行了阐述，最后详细介绍了本研究中采用的实验指标和用到的数据集。

第三章基于自适应动态阈值的半监督变化检测工作介绍。首先整体性的阐述了算法的框架和训练流程，随后详细介绍了基于预测概率分布的全局动态阈值和类别动态阈值，以及设计的低置信度学习模块。最后在实验部分报告了本章方法在十个公开数据集上取得的实验结果，从定性和定量两个角度进行了全方位的对比，以及全面的消融实验分析，验证了该工作的有效性。

第四章基于伪标签评估的自适应半监督变化检测工作介绍。首先介绍了模型的整体框架以及训练流程，随后详细介绍了设计的伪标签评估指标，以及基于此指标设计的两个自适应模块。最后在实验部分报告了该研究方法在十个公开数据集上取得的实验结

果，并与其他经典半监督变化检测算法在定性和定量上进行了公平的对比，最后通过消融实验证明了每个模块的有效性。

第五章基于自适应特征扰动的半监督变化检测工作介绍，首先从全局角度阐述了算法的框架和流程，随后具体地介绍了我们设计的基于无标记样本预测评估的自适应特征扰动方法。最后在实验部分报告了我们在十个公开数据集上对本章方法进行实验证取得的实验结果，包括定量实验指标和定性可视分析，此外还对本章提出的自适应扰动模块进行了消融分析，证明了本章研究的有效性。

第六章总结与展望。从全局角度总结了本文的主要研究工作，并对未来半监督变化检测领域可以继续改进创新之处做出了展望。

第2章 相关技术

本章将介绍本文基于深度学习的半监督变化检测算法研究中所用到的相关技术，为后续章节做好铺垫。具体包括：2.1小节中详细介绍本文几个研究工作中用到的深度神经网络，包括卷积神经网络，Vision Transformer以及Segment Anything Model的网络结构等；2.2小节对上一章节研究现状中概括性介绍的几个经典半监督变化检测算法进行了详细的阐述；2.3小节对本文所有实验所采用的实验指标进行了解释；2.4小节对本文所有进行实验的公开数据集进行了详细介绍。

2.1 深度神经网络

深度神经网络是一种以多层神经元结构为基础的机器学习模型，广泛应用于诸多领域，包括计算机视觉、自然语言处理、语音识别以及推荐系统等。基于任务需求，衍生出了多种具体架构，例如用于图像处理的卷积神经网络CNN，在序列建模中表现优异的循环神经网络RNN，以及基于多头自注意力机制对全局关系进行建模的Transformer架构。虽然神经网络并不是本文的主要研究方向，但是由于后续研究中会多次涉及卷积神经网络中比较经典的网络模型，因此在本小节将介绍后续研究工作所用到的几种骨干特征提取网络，即ResNet，Vision Transformer，以及基于Transformer预训练的通用分割大模型Segment Anything Model的模型结构。

2.1.1 ResNet

在深度学习的发展过程中，研究者们发现随着卷积神经网络层数的增加，模型的表达能力理论上会不断增强。然而，在实际训练深层网络时，却遇到了两个主要问题：（1）梯度消失和梯度爆炸问题；（2）退化问题。这些问题阻碍了深层神经网络的发展，最终何等人^[1]提出了残差网络（Residual Network, ResNet），通过在网络层之间引入残差连接使得前层的输出直接加到后层的输出上，如公式2-1所示，其中 x 代表输入特征，下标表示网络层序号， W_l 代表第 l 层的参数， $F(x_l, W_l)$ 表示卷积映射。通过这种方式，梯度在反向传播时可以通过跳跃路径直接传播到前层，避免了梯度逐层递减的问题，从而缓解了梯度消失，同时即使一些层对学习贡献较小，残差连接可以保证网络的性能至少不会发生退化现象。

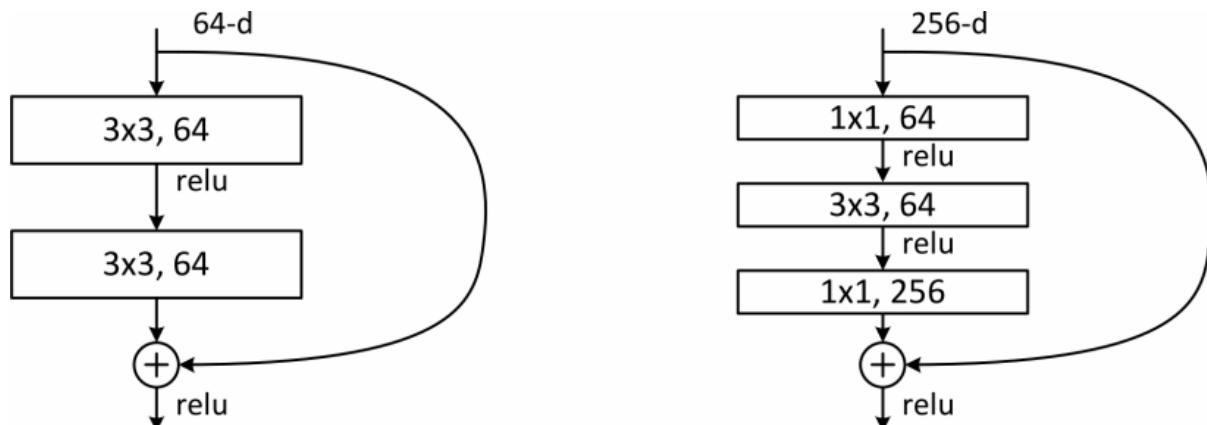
$$x_{l+1} = x_l + F(x_l, W_l) \quad (2-1)$$

此外残差结构还减少了优化深层网络的难度，是网络结构更加复杂、网络层数进一步增加成为了可能，从而显著提升了卷积神经网络的表征能力和性能，推动了深度学习向更深、更强的方向发展。比如何等人同时提出了具有不同网络层数的ResNet，包括ResNet-18、ResNet-34、ResNet-50、ResNet-101以及ResNet-152，它们的网络架构如图2-1所示。此外后续还有研究提出了一些ResNet变体，例如ResNetV2^[94]和ResNeXt^[95]。

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112			7×7, 64, stride 2		
				3×3 max pool, stride 2		
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1			average pool, 1000-d fc, softmax		
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

图 2-1 不同深度的 ResNet 网络层次架构^[1]

根据网络深度和任务需求, ResNet 中存在两种经典的残差块, 一种是用于浅层网络 (如 ResNet-18 和 ResNet-34) 的基本残差块 (Basic Residual Block), 包含两个 3×3 的卷积层和直接的跳跃连接, 一种是用于深层网络 (如 ResNet-50、ResNet-101 和 ResNet-152) 的瓶颈残差块 (Bottleneck Residual Block), 它将两个 3×3 的卷积层替换为 1×1 , 3×3 , 1×1 的卷积网络, 其中前后两个 1×1 卷积层分别用于降低和恢复特征通道数, 有效地减少了计算量和参数量, 网络结构如图2-2所示。本文研究工作中采用的是 ResNet-50, 一共包含 4 个卷积阶段, 每一个阶段中分别有 3, 4, 6, 3 个瓶颈残差块。

图 2-2 两种残差块的网络结构^[1]

2.1.2 Vision Transformer

虽然 CNN 在计算机视觉任务 (如图像分类、目标检测、语义分割等) 中占据主导地位, 然而其在捕捉长距离依赖和全局信息时存在很大局限性。与此同时, Transformer 在自然语言处理 (Natural Language Processing, NLP) 领域取得了显著成功, 尤其是在任务如机器翻译和文本生成中。Transformer 的核心优势是通过自注意力机制 (Self-Attention Mechanism) 捕捉输入数据中全局依赖关系。受此启发, Dosovitskiy 等人^[2]对 Transformer

进行改进使其能够适配计算机视觉任务，形成了 Vision Transformer (ViT) 架构。这种方法首次完全摆脱了卷积操作，在多个基准测试中展现出与甚至超越 CNN 的性能。它的核心思想是将图像分割为固定大小的图像块 (patches)，并将每个块视为一个独立的“词”，类似于 NLP 中的词嵌入，这些图像块嵌入向量通过 Transformer 的自注意力机制进行处理，从而实现对图像的全局理解。

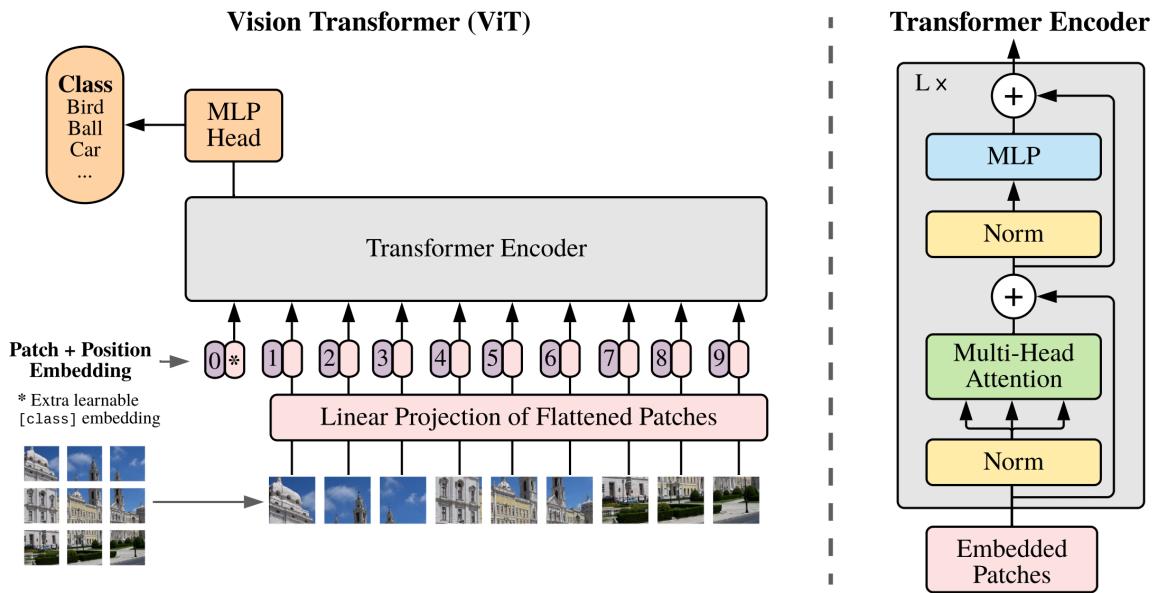


图 2-3 Vision Transformer 网络框架^[2]

ViT 的网络结构如图2-3所示，输入的图像首先被分割为大小为 $P \times P$ 的非重叠图像块 (patches)，每个块展平为一维向量并通过线性映射嵌入到 D 维特征空间，假设图像大小为 $H \times W$ 且通道数为 C，则切割后共生成 $N = \frac{H}{P} \times \frac{W}{P}$ 个图像块，这些块被视为 Transformer 的输入“词”。为保留块之间的空间顺序信息，ViT 为每个嵌入向量添加一个可学习的位置编码，此外，在输入序列的起始位置插入一个特殊的分类标识符 ([CLS] Token)，该标识符的嵌入向量用于最终的分类输出，如公式2-2。经过嵌入和位置编码后的输入序列被送入由 L 层组成的 Transformer 编码器，每层包含多头自注意力机制 (Multi-Head Self-Attention, MHSA) 和前馈网络 (Feed-Forward Network, FFN)。在 MHSA 中，每个图像块嵌入向量通过查询 (Query)、键 (Key) 和值 (Value) 矩阵生成注意力权重，用以捕捉图像块间的全局关系，表示为公式2-3。其中 Q、K、V 是从输入嵌入计算得到的查询、键和值矩阵， $\sqrt{d_k}$ 是注意力头的维度。每个 Transformer 层都配有残差连接和层归一化，用以提高训练稳定性。经过 L 层编码器后，分类标识符的嵌入向量 $z_0^{[CLS]}$ 被用作图像的全局表示，并通过一个全连接层映射到类别数的维度，最终通过 Softmax 函数生成分类概率，如式 (2-4)，其中 W_{head} 和 b_{head} 为分类头的参数。

$$z_0 = [z_0^{[CLS]}; z_0^1; z_0^2; \dots; z_0^N] + E_{\text{pos}}, \quad (2-2)$$

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^\top}{\sqrt{d_k}} \right) V, \quad (2-3)$$

$$\hat{y} = \text{Softmax} \left(z_L^{[CLS]} W_{\text{head}} + b_{\text{head}} \right) \quad (2-4)$$

ViT 摒弃了传统卷积操作，采用自注意力机制实现图像全局特征的高效建模，其模块化和可扩展的架构使其在大规模数据集和高效计算资源支持下展现出优异性能。然而，ViT 在小规模数据场景中仍面临一定挑战，主要体现在模型对训练数据的需求较高和计算效率的限制。因此，针对训练数据稀缺的情况，结合半监督学习算法是一种具有潜力的研究方向，通过充分利用未标注数据，有望显著提升 ViT 在小数据集场景下的表现。

2.1.3 金字塔池化模型

在传统的卷积神经网络（CNN）中，特征图的尺寸通常会随着卷积层的增加而逐渐减小。这对于处理不同尺寸的输入图像可能会带来问题，因为图像的空间信息可能在池化的过程中丢失。尤其是在语义分割、目标检测等任务中，捕捉多尺度的上下文信息对于提升模型性能至关重要。为了更好地捕捉图像的多尺度信息，金字塔池化模型（Pyramid Pooling Module, PPM）通过对特征图进行不同尺度的池化操作，避免了仅依赖单一尺度的局限性，增强了网络对多种尺度物体的感知能力。

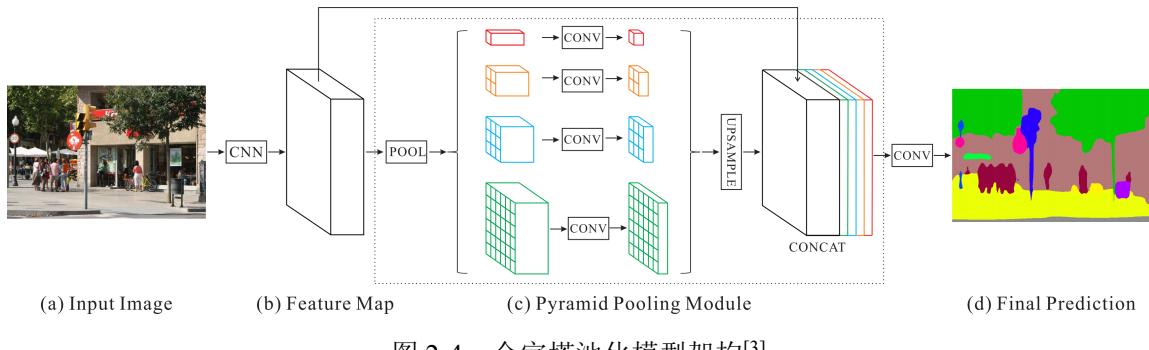


图 2-4 金字塔池化模型架构^[3]

金字塔池化模型的核心思想是对特征图进行不同尺度的池化操作，然后将池化后的结果拼接（concatenate）起来，以获得多尺度的信息。具体来说，PPM 的网络结构如图2-4所示，通常包括以下几个部分：

- (1) 多尺度池化：PPM 通过对输入的特征图进行不同尺寸的池化操作，例如：1x1、2x2、4x4 等。每个尺度的池化结果都有不同的感受野，能够捕捉到图像中不同大小的物体。
- (2) 上采样：每个池化层得到的特征图通过上采样（upsampling）操作恢复到相同的尺寸，以便可以与其他尺度的池化结果拼接。
- (3) 拼接（Concatenation）：将不同尺度池化后的特征图拼接在一起，生成包含多尺度信息的特征表示。拼接后的特征图具有丰富的全局上下文信息，同时保留了原始特征图的细节信息。

通过对特征图进行多尺度池化操作，金字塔池化模型有效地增强了网络的多尺度上下文信息表达能力，从而提升了图像处理任务的性能。它在很多领域中都有着广泛的应用，特别是在语义分割、目标检测等需要多尺度信息的任务中表现尤为出色，因此我们在本文的变化检测研究中引入了这种模型。

2.1.4 平均教师模型

平均教师模型是 Tarvainen 等人^[4]于 2018 年提出的一种半监督学习算法，该算法是针对时间集成模型存在的计算成本大这一缺陷（在每个 epoch 上更新一次目标标签）提出的改进算法，不同之处是时间集成模型基于时间序列的指数移动平均是在预测结果上，而平均教师模型则是在模型的权重上。在训练步骤上平均模型权重往往比直接使用最终权重更准确，在训练过程中利用这一点能够训练更优模型。

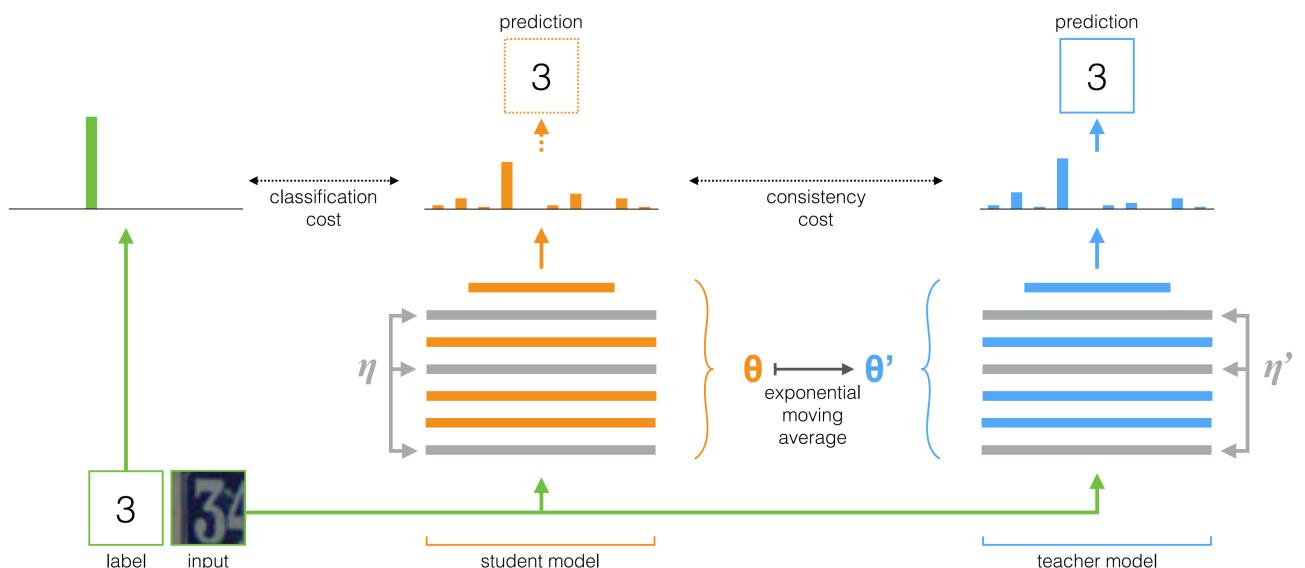


图 2-5 平均教师模型的网络结构^[4]

模型框架如图2-5所示，该半监督框架包含一个教师模型和一个学生模型，由于教师模型的权重是历史学生模型的平均值，故将其称为平均教师模型。核心思想为：使用教师模型用来生成学生模型学习的目标，指导学生模型从无标记样本进行学习，基于一致性正则化原理，认为输入数据添加一个微小的扰动噪声时，模型的预测结果不会发生变化。更正式地说，将一致性成本损失 J 定义为学生模型的预测（模型权重为 θ ，施加噪声 η ）与教师模型的预测（模型权重为 θ' ，施加噪声 η' ）之间的预期距离，如公式2-5。

$$J(\theta) = \mathbb{E}_{x,\eta',\eta} \left[\|f(x, \theta', \eta') - f(x, \theta, \eta)\|^2 \right] \quad (2-5)$$

π -模型、时间集成模型和平均教师模型之间的区别在于教师模型预测结果的生成过程。其中 π -模型中教师、学生权重共享，即 $\theta' = \theta$ ；时间集成模型通过连续预测的指数移动平均值来估计 $f(x, \theta', \eta')$ ；在平均教师模型中，将训练步骤 t 时刻的教师模型权重

定义为连续学生模型权重权重的 EMA，表示为如下：

$$\theta'_t = \alpha\theta'_{t-1} + (1 - \alpha)\theta_t \quad (2-6)$$

其中 α 为平滑系数超参数。三种算法之间的另一个区别是， π -模型直接将训练应用于 θ' ，而时间集成模型和平均教师模型将其视为另一个优化途径。

最终可以在每个训练步骤中随机采样噪声 η 和 η' ，来计算一致性成本 $J(\theta)$ ，并通过随机梯度下降来优化学生模型参数，在此方法中，作者使用均方误差损失作为一致性成本。

2.2 经典半监督变化检测算法

2.2.1 SemiCDNet

Peng 等人^[5]提出了 SemiCDNet 这种基于生成对抗网络 (GAN) 的半监督卷积网络。首先，将标记数据和未标记数据输入分割网络，生成初始预测和熵图。然后，为了充分利用未标记数据的潜力，采用两种鉴别器来加强标记数据与未标记数据之间的分割图和熵图的特征分布一致性。在对抗训练过程中，利用未标记信息不断地对生成器进行正则化，从而提高生成器的泛化能力。SemiCDNet 是 RCR 之前唯一可获得的开源 SOTA 方法，在本文的研究中作为对比方法之一，因此在本小节对其进行简要阐述。

(1) 网络框架

SemiCDNet 的网络框架如图2-6所示，主要包括一个生成器 G 和两个判别器 D_s 和 D_e ，分别用于对齐生成器在无标记数据上的预测概率分布和熵图。

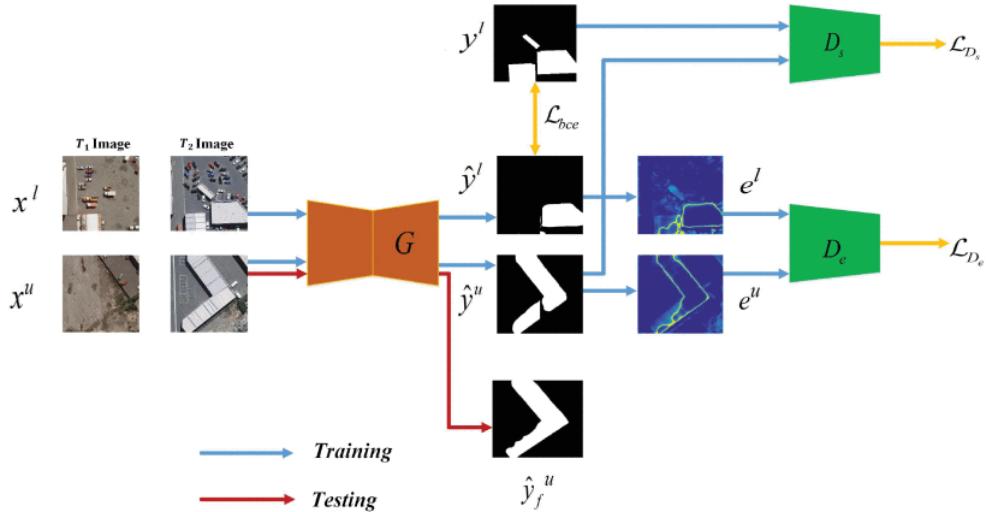


图 2-6 SemiCDNet 网络框架^[5]

(2) 损失函数

主要通过三种损失来训练生成器网络和鉴别器网络，即在标记数据上的加权二值交叉熵损失以及在无标记数据上的分割对抗损失和熵对抗损失。

加权二值交叉熵损失：为了克服变化检测中变化像素和不变像素之间的样本偏差问

题，采用简单的加权二值交叉熵损失。这是一种标准的监督的像素级的分割损失，仅在标记数据上计算这种损失。计算方式可以表示为：

$$\mathcal{L}_{\text{bce}} = -\frac{1}{M} \left[\beta \sum_{j \in y'_+} \log(P(y_j = 1)) + (1 - \beta) \sum_{j \in y'_-} \log(P(y_j = 0)) \right] \quad (2-7)$$

其中 β 和 $1 - \beta$ 分别表示标记图像的真实标签变化和不变的像素数， $P(\cdot)$ 是生成器在像素 j 上的预测输出。

分割对抗损失： D_s 用于判断分割图是来自生成器在无标记样本上的预测，还是来自有标注的真实标签，促使生成器 G 在无标记样本的预测 $G(x^u)$ 的特征分布与有标注样本的真实标签 y^l 的特征分布对齐。分割对抗损失计算公式如下：

$$\begin{aligned} \mathcal{L}_{D_s} = & \frac{1}{M} \sum_{x^l, y^l \in \mathcal{D}_C} \mathcal{L}_D(y^l \oplus x^l, 1) \\ & + \frac{1}{N} \sum_{x^u \in \mathcal{D}_U} \mathcal{L}_D(G(x^u) \oplus x^u, 0) \end{aligned} \quad (2-8)$$

其中 \oplus 表示堆叠拼接操作， \mathcal{L}_D 表示二元交叉熵损失函数，其目的是最小化预测 $G(x^u)$ 与真实 y^l 分布之间的差异。

此外，为了与鉴别器对抗，使用以下对抗损失对分割网络进行优化：

$$\mathcal{L}_{\text{adv}}^{D_s} = \frac{1}{N} \sum_{x^u \in \mathcal{D}_U} \mathcal{L}_D(G(x^u) \oplus x^u, 1) \quad (2-9)$$

熵对抗损失：通常生成器倾向于对标记数据进行低熵、高确定性的预测，而对未标记数据进行的预测具有高熵、低确定性。因此可以对无标记数据进行低熵值的约束。其中熵图（Entropy map）被定义为：

$$E(x) = G(x) \bullet \log[G(x)] \quad (2-10)$$

其中 \bullet 表示点积的操作。 D_s 被训练用于判断熵图是来自有标记数据还是无标记样本上的预测，通过对齐 $E(x^u)$ 和 $E(x^l)$ 之间的特征分布，还可以抑制无标记数据熵图的高不确定性。因此判别器的熵对抗损失通过以下公式计算：

$$\begin{aligned} \mathcal{L}_{D_e} = & \frac{1}{M} \sum_{x^l, y' \in \mathcal{D}_C} \mathcal{L}_D(E(x^l) \oplus x^l, 1) \\ & + \frac{1}{N} \sum_{x^u \in \mathcal{D}_U} \mathcal{L}_D(E(x^u) \oplus x^u, 0) \end{aligned} \quad (2-11)$$

生成器的熵对抗损失为：

$$\mathcal{L}_{\text{adv}}^{D_e} = \frac{1}{N} \sum_{x^u \in \mathcal{D}_U} \mathcal{L}_D(E(x^u) \oplus x^u, 1) \quad (2-12)$$

于是，SemiCDNet 的总体损失为：

$$\mathcal{L}_G = \mathcal{L}_{\text{bce}} + \lambda_s \mathcal{L}_{\text{adv}}^{D_s} + \lambda_e \mathcal{L}_{\text{adv}}^{D_e} \quad (2-13)$$

其中 λ_s 和 λ_e 分别表示分割对抗损失和熵对抗损失的自定义权重。

2.2.2 RCR

RCR (Revisiting Consistency Regularization) 是由 Bandara 等人^[6]提出的，它们基于聚类假设构建了一个更加广阔的特征扰动空间，这种特征扰动一致性正则化使得模型具有更加强大的泛化能力。本文的研究工作的代码是基于 RCR 的变化检测网络和训练框架以及前文提到的平均教师模型改进实现的，因而有必要在本小节详细介绍 RCR 方法。

(1) 网络框架

总体框架如图2-7所示，主要包括三个模块：1) 编码器，用于提取前时相图像和后时相图像的隐藏特征表示；2) 特征差分模块，用于获取变化前和变化后图像的隐藏特征表示 F_d 的差异；3) 解码器，从隐藏的差异特征表示预测变化掩膜。

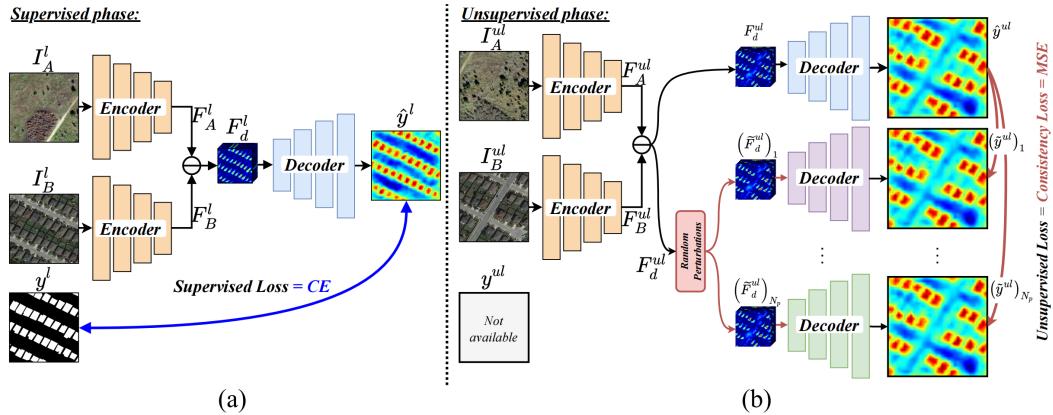


图 2-7 RCR 网络框架^[6]

编码器 f_e 。对于编码器，RCR 使用预训练的 ResNet50^[1]。编码器的输出是 2048 维的特征矩阵，空间分辨率为 $\frac{H}{4} \times \frac{W}{4}$ ，其中 H 和 W 分别为输入双时相图像 $\{I_A, I_B\}$ 的高度和宽度。在孪生网络架构中使用共享权重的双分支编码器，以分别提取两幅图像的隐藏特征表示 F_A 和 F_B ，数学公式表示为：

$$\begin{aligned} F_A &= f_e(I_A), \\ F_B &= f_e(I_B) \end{aligned} \quad (2-14)$$

特征差分模块。一旦从编码器获得给定双时相图像 $\{I_A, I_B\}$ 的隐藏特征表示 F_A 和 F_B ，RCR 通过简单地计算 F_A 和 F_B 之间的绝对差值以获得隐藏特征的差分表示 F_d ，随后通过特征金字塔池化模块 (Pyramid Pooling Module, PPM)^[3] f_{PPM} 对其进行处理，

以有效地获取不同尺度的变化。用数学方法将特征差分模块内部的过程表示为：

$$F_d = f_{\text{PPM}}(|F_A - F_B|_1) \quad (2-15)$$

编码器 f_d 。解码器的目的是从隐藏的差分特征 F_d 中估计输出的变化概率图 \hat{y} 。 j 解码器中通过一系列亚像素卷积上采样模块^[96]，直到特征达到输入双时图像的空间分辨率 $H \times w$ 。解码器内部的过程可以用数学表达为：

$$\hat{y} = f_d(F_d) \quad (2-16)$$

(2) 训练管道

训练以上变化检测的过程包含两个部分，一部分是在少量标注样本上进行的监督训练，另一部分则是在大量无标注样本上进行的无监督训练。

监督训练。在这一阶段，RCR 只预测标记训练数据所隐含的兴趣变化： $\mathcal{D}_l = \{\{I_{A,i}^l, I_{B,i}^l\}, y_i^l\}_{i=1}^{N_l}$ ，其中 $\{I_{A,i}^l, I_{B,i}^l\}$ 表示第 i 个双时相图像对， y_i^l 是相应的真实变化掩码标注， N_l 是标记数据集的大小。利用预测值与真实标签之间计算的交叉熵（Cross Entropy, CE）损失^[97]作为监督损失 L_{sup} ，如公式2-17所示：

$$\mathcal{L}_{\text{sup}} = \text{CE}(\hat{y}_i^l, y_i^l) \quad (2-17)$$

整个过程如图2-7-a 所示。

无监督训练。在这一阶段，除了标记数据 \mathcal{D}_l ，我们还使用未标记的双时相图像对 $\mathcal{D}_u = \{I_{A,i}^{ul}, I_{B,i}^{ul}\}_{i=1}^{N_{ul}}$ ，其中 $\{I_{A,i}^{ul}, I_{B,i}^{ul}\}$ 是第 i 个未标记的双时相图像对， N_{ul} 是未标记数据集的大小，通常假设它大于标记数据集的大小（即 $N_{ul} \gg N_l$ ）。为了有效地利用这些容易获得的未标记双时相图像来提高变化检测模型 $f_{CD}(\cdot)$ 的性能，RCR 提出了一个基于未标记数据的无监督损失 L_{unsup} ，它提供了一个额外的训练信号来优化 $f_{CD}(\cdot)$ 的参数。所提出的无监督损失基于半监督学习中的聚类假设，其中 RCR 强制 $f_{CD}(\cdot)$ 的预测在深度特征差分图 F_d 上施加不同随机扰动下依然保持一致，如图2-7-b 所示。

定义 $\left\{(\tilde{F}_{d,i}^{ul})_1, \dots, (\tilde{F}_{d,i}^{ul})_p, \dots, (\tilde{F}_{d,i}^{ul})_{p=N_p}\right\}$ 作为第 i 个未标记双时相图像对 $\{I_{A,i}^{ul}, I_{B,i}^{ul}\}$ 的隐藏差分特征 $F_{d,i}^{ul}$ 的随机扰动版本集合。接下来，RCR 通过主解码器 $f_d(\cdot)$ 对 $F_{d,i}^{ul}$ 进行处理，得到预测的变化概率图为：

$$\hat{y}_i^{ul} = f_d(F_{d,i}^{ul}) \quad (2-18)$$

以及通过一组与 $f_d(\cdot)$ 设计相似的辅助解码器处理隐藏特征差分图的每个扰动版本，得到它们对应的预测 $(\tilde{y}_i^{ul})_p$ 为：

$$(\tilde{y}_i^{ul})_p = f_d^p \left(\left(\tilde{F}_{d,i}^{ul} \right)_p \right), \text{ where } p = 1, \dots, N_p \quad (2-19)$$

接下来，RCR 通过定义无监督损失 L_{unsup} 来强制 $\left\{(\tilde{y}_i^{ul})_p\right\}_{p=1}^{N_p}$ 与 \hat{y}_i^{ul} 一致，如下公式所示：

$$\mathcal{L}_{unsup} = \sum_{p=1}^{N_p} \mathbf{d}\left((\tilde{y}_i^{ul})_p, \hat{y}_i^{ul}\right) \quad (2-20)$$

其中 $d(\cdot)$ 是距离度量，用于测量预测之间的不相似性，在 RCR 中 Bandara 等人使用均方误差（Mean Squares Error，MSE）作为 $d(\cdot)$ 。

(3) 扰动方式

在 RCR 中，在输入特征上采取的扰动方式有以下几种：

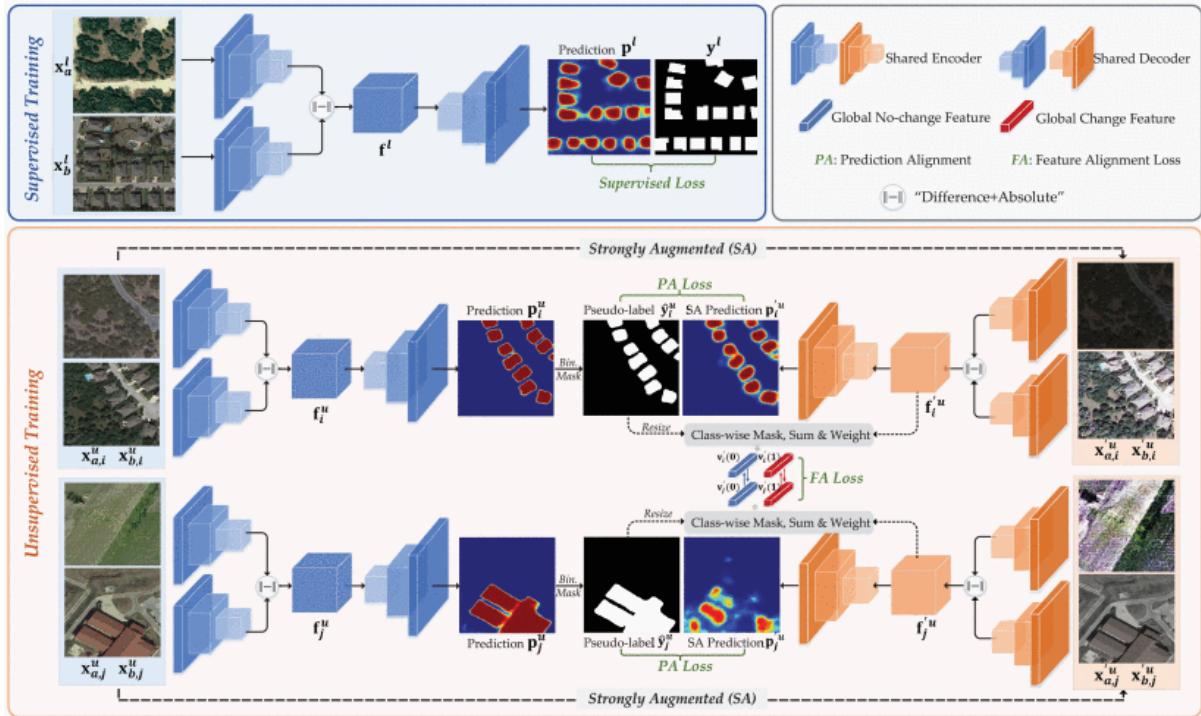
- 1) 随机特征噪声：随机生成一个三维噪声张量，然后根据 $F_{d,i}^{ul}$ 中值的大小对其进行缩放，并将其添加到隐藏特征差分图中，得到一个扰动版本。
- 2) 随机特性丢弃：首先通过阈值从特性差异图中选取出 10% 到 40% 的最需要的区域，生成掩膜，然后沿着通道维度，将潜在差分特征和该掩膜进行逐像素相乘，从而丢弃掉那些不需要的区域特征。
- 3) 引导特征剪切：我们基于预测的变化图，从差分特征图中随机取零，得到扰动特征图。
- 4) 内容和对象覆盖：基于变化检测网络的输出对变化类或不变类保持不变的假设，通过变化掩码掩盖隐藏差分特征图中的变化区域，或者通过不变掩膜掩盖隐藏差分特征图中的不变区域，创建隐藏差分特征图的两个扰动版本。
- 5) 特征 VAT^[98]：在变化最大的方向上对差分特征图应用对抗性扰动。

2. 2. 3 FPA

FPA (Feature-Prediction Alignment) 是张等人^[7]2023 年提出的半监督变化检测框架。FPA 提出了两种对齐策略，来有效地利用未标记的双时相图像对进行训练。首先，设计了一种类感知特征对齐 (Feature Alignment, FA) 策略，将从不同未标记图像对（即跨区域）中提取的区域级变化/无变化特征进行对齐，以减少同一类内的特征差异。其次，设计了一种像素级预测对齐 (Pixel-wise Prediction Alignment, PA) 方法，将强增强未标记图像对的像素级变化预测与弱增强对应的伪标签进行对齐，以降低各种具有物理意义的图像对变换的预测不确定性。其总体框架如图2-8所示。其中使用的变化检测网络和监督训练过程都和 RCR^[6] 相同，因此以下内容主要介绍其新颖的无监督训练部分。

(1) 类感知特征对齐

类感知特征对齐目的是在训练阶段实现小批量内不同未标记图像对的类内全局特征对齐。具体说来，首先对原始的无标记图像对施加弱增强操作，得到一个弱增强无标记图像对 $\{\mathbf{x}_a^u, \mathbf{x}_b^u\}$ ，并进一步继续对其施加强增强操作得到其对应的强增强图像对

图 2-8 FPA 网络框架^[7]

$\{\mathbf{x}'_a^u, \mathbf{x}'_b^u\}$:

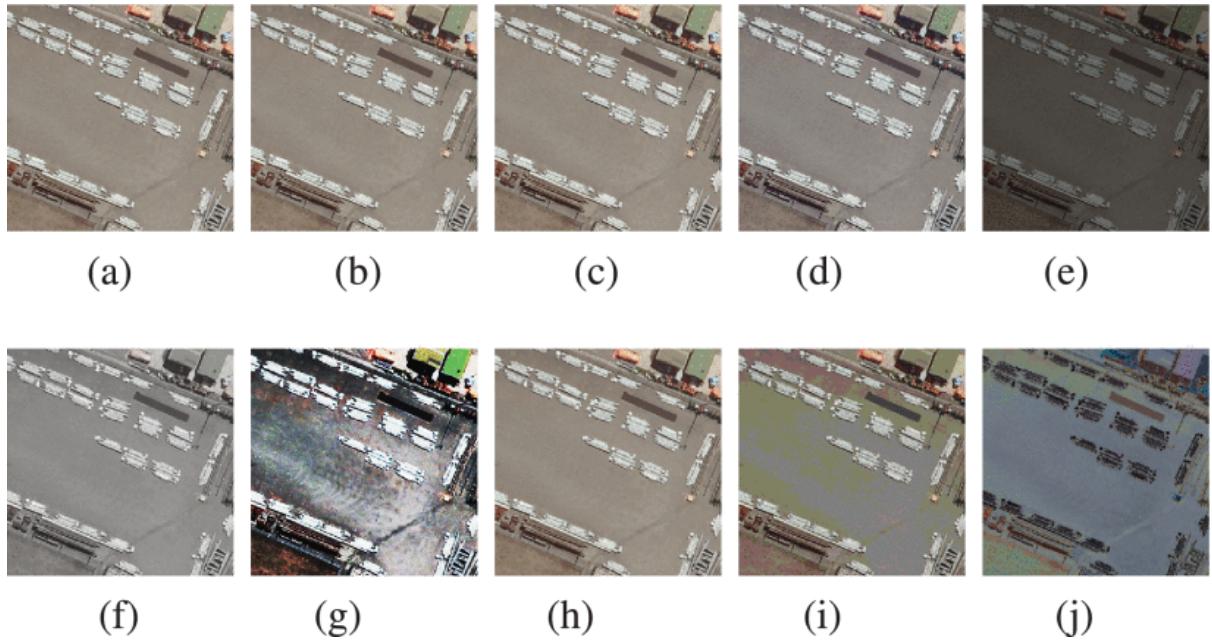
$$\begin{aligned} \mathbf{x}'_a^u &= \text{RandAugment}(\mathbf{x}_a^u), \\ \mathbf{x}'_b^u &= \text{RandAugment}(\mathbf{x}_b^u) \end{aligned} \quad (2-21)$$

其中 $\text{RandAugment}(\cdot)$ 表示从预定义的增强列表中随机抽样两个连接的强增强操作，增强列表如图2-9所示，包括 Identity, Contrast, Autocontrast, Brightness, Color, Equalize, Sharpness, Posterize 和 Solarize 等 9 种常用增强方式，其中 a-g 分别是 (a) 原始图像；(b) Identity；(c) Contrast；(d) Autocontrast；(e) Brightness；(f) Color；(g) Equalize；(h) Sharpness；(i) Posterize；(j) Solarize。

然后，通过编码器分别从 $\{\mathbf{x}_a^u, \mathbf{x}_b^u\}$ 和 $\{\mathbf{x}'_a^u, \mathbf{x}'_b^u\}$ 中提取特征映射 f^u 和强增强特征映射 f'^u 。为了模型训练的鲁棒性，以解码器从 f^u 中提取的弱增强预测映射 p^u 作为伪标签参考，指导在强增强特征图 f'^u 上进行特征对齐操作。为了实现这一步，通过固定阈值筛选掉 p^u 中部分含噪较高的像素预测值：

$$\mathbf{m}^{uf}(i, j, k) = \begin{cases} 1, & \text{if } \mathbf{p}^u(i, j, k) > \tau \\ 0, & \text{else} \end{cases} \quad (2-22)$$

接下来对 m^{uf} 进行最近邻下采样为 $\mathbf{m}^{uf} \in \mathbb{R}^{H/s \times W/s \times 2}$ 以适应强增强特征图 f'^u 的空间分辨率。如此一来便可以基于预测掩码 m^{uf} 和强增强特征映射 f'^u ，提取出分类全

图 2-9 FPA 使用的强增强操作列表^[7]

局特征向量，记为 $\mathbf{v}' \in \mathbb{R}^{2 \times C}$ ，以上过程用公式可表示为如下：

$$\mathbf{v}'(k) = \frac{1}{\mathbf{w}(k)} * \sum_{i=1}^{H/s} \sum_{j=1}^{W/s} \mathbf{f}^u(i, j) * \mathbf{m}^{uf}(i, j, k) \quad (2-23)$$

其中 $*$ 表示逐像素点积运算， $\mathbf{w}(k)$ 为类权像素和，即变化类别的像素总数，计算过程如2-24， $\epsilon = 1e - 8$ 是一个很小的余量，以避免每个类的权重可能为零。

$$\mathbf{w}(k) = \sum_{i=1}^{H/s} \sum_{j=1}^{W/s} \mathbf{m}^{uf}(i, j, k) + \epsilon \quad (2-24)$$

最后为了实现跨区域的类内特征对齐，FPA 通过增加它们的余弦相似度约束来使一个小批中的所有分类全局特征向量彼此对齐，无监督特征对齐损失计算如下：

$$\begin{aligned} \mathcal{L}_u^{FA} = & \frac{1}{2BB} \sum_{k=1}^2 \sum_{i=1}^B \sum_{j=1}^B \mathbb{I}(\mathbf{w}_i(k) > 0, \mathbf{w}_j(k) > 0) \\ & \cdot \frac{1}{2} \left(1 - \frac{\mathbf{v}'_i(k) * \mathbf{v}'_j(k)}{\|\mathbf{v}'_i(k)\| \|\mathbf{v}'_j(k)\| + \epsilon} \right) \end{aligned} \quad (2-25)$$

其中， $\mathbf{v}'_i(k)$ 和 $\mathbf{v}'_j(k)$ 分别表示当前小批量（ B 为小批量大小）中第 i 和第 j 个未标记图像对的第 k 类全局特征向量。 $(\mathbf{v}'_i(k) * \mathbf{v}'_j(k) / \|\mathbf{v}'_i(k)\| \|\mathbf{v}'_j(k)\|)$ 表示 $\mathbf{v}'_i(k)$ 与 $\mathbf{v}'_j(k)$ 之间第 k 类全局特征的余弦相似度，其本身取值范围为 $[-1, 1]$ ，为了便于训练 FPA 将其重映射到了 $[0, 1]$ 。

(2) 像素级预测对齐

像素级预测对齐的目的是使强增强图像对的输出与弱增强图像对的输出保持一致，

从而使模型获得鲁棒的特征提取能力。为此，FPA 中引入了^[50] 的置信度一致性学习策略，从半监督图像分类任务适用到SSCD的像素级任务。对于从弱增强图像对获取到的像素级的变化概率图 \mathbf{p}^u ，按以下方式生成伪标签映射 $\hat{\mathbf{y}}^u \in \mathbb{R}^{H \times W}$ ：

$$\hat{\mathbf{y}}^u(i, j) = \arg \max_{k=\{0,1\}} \mathbf{p}^u(i, j, k) \quad (2-26)$$

同样，为了减少噪声伪标签的干扰，对从 \mathbf{p}^u 生成的基于置信度的变化掩码 $\mathbf{m}^{up} \in \mathbb{R}^{H \times W}$ 基于一个固定阈值仅筛选出那些最为可信的预测像素。

$$\mathbf{m}^{up}(i, j) = \begin{cases} 1, & \text{if } \mathbf{p}^u(i, j, \hat{\mathbf{y}}^u(i, j)) > \tau \\ 0, & \text{else.} \end{cases} \quad (2-27)$$

其中 τ 是固定阈值的取值，在 FPA 中默认设置为 0.95。

因此，第 k 个图像对的像素级预测对齐损失，可以表示为公式2-28和公式2-29：

$$\mathcal{L}_u^{PA}(k) = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W \text{CE}(\mathbf{p}'^u(i, j), \hat{\mathbf{y}}^u(i, j)) * \mathbf{m}^{up}(i, j), \quad (2-28)$$

$$\mathcal{L}_u^{PA} = \frac{1}{B} \sum_{k=1}^B \mathcal{L}_u^{PA}(k) \quad (2-29)$$

最终整个无监督训练的损失函数即为类感知特征对齐损失和像素级预测对齐损失的求和：

$$\mathcal{L}_u = \mathcal{L}_u^{FA} + \mathcal{L}_u^{PA} \quad (2-30)$$

2.3 实验指标

为了更好地衡量所有模型的性能，我们引入了 2 个广泛使用的变化检测评价指标，包括交并比（Intersection over Union, IoU）和总体精确度（Overall Accuracy, OA）。IoU、OA 的取值范围均为 0-100 (%)。对于所有这些指标，该值越大，变化检测性能就越好，但是在变化检测任务中，由于二分类和不平衡的类别不平衡，OA 总体上都是一个很高的值。它们的计算表述如下：

$$IoU = \frac{TP}{TP + FP + FN}, \quad (2-31)$$

$$OA = \frac{TP + TN}{TP + TN + FN + FP} \quad (2-32)$$

其中 TP 和 TN 分别表示正识别的变化像素数和未变化像素数。相反， FP 表示未发生变化的像素被错误地分类为变化的像素的数量， FN 表示发生变化的像素被错误地识别为未变化的像素的数量。此外，由于我们更关注变化区域的预测性能，并且变化类和背景类极度不平衡，因此在实验中使用变化类别区域的平均交并比 (IoU^c) 作为评

价指标。

2.4 实验数据集介绍

本文中所有方法都在十个基准公开变化检测数据集上进行了实验，分别是 LEVIR-CD^[99]、LEVIR-CD+^[99]、WHU-CD^[100]、EGY-CD^[101]、HRCUS-CD^[102]、Change Detection Dataset (CDD)^[103]、GZ-CD^[5]、DSIFN-CD^[104]、SYSU-CD^[105] 和 CL-CD^[106]。其中，这些数据集涵盖了不同的分辨率（0.03m-2.0m）、不同的数据大小（2400 至 20000 对）、不同的标注类别（二值建筑物或多类）、不同的图像对时间跨度（1 年-16 年），汇总如表2-1所示，二值建筑物变化检测数据集部分样例展示如图2-10所示，多类变化检测数据集部分样例展示如图2-11所示。

表 2-1 本文所使用的公开数据集

变化类别	数据集	空间分辨率	大小	样本数量	时间跨度	链接
建筑物	LEVIR-CD ^[99]	0.5m	1024 × 1024	637	5 到 14 年	Link
	LEVIR-CD+ ^[99]	0.5m	1024 × 1024	985	5 到 14 年	Link
	WHU-CD ^[100]	0.2m	15354 × 32507	1	2012 年至 2016 年	Link
	GZ-CD ^[5]	0.55m	Varying	19	2006 年至 2019 年	Link
	EGY-BCD ^[101]	0.25m	256 × 256	6091	2015 年至 2022 年	Link
	HRCUS-CD ^[102]	0.5m	256 × 256	11388	混合	Link
多类	CDD ^[103]	0.03m-1.0m	256 × 256	16000	混合	Link
	DSIFN-CD ^[104]	Unknown	512 × 512	3940	未知	Link
	SYSU-CD ^[105]	0.5m	256 × 256	20000	2007 年至 2014 年	Link
	CL-CD ^[106]	0.5-2.0m	512 × 512	600	2017 年至 2019 年	Link

LEVIR-CD 数据集：该数据集是一个综合性的遥感建筑变化检测数据集，由 637 对超高分辨率的谷歌地球图像块组成，每个块的空间分辨率为 0.5m，尺寸为 1024 × 1024 像素。这些双时相图像来自美国德克萨斯州七个城市的 20 个不同地点，双时相图像拍摄于 2002 年至 2018 年间，时间跨度在 5 到 14 年。经过切割之后，该数据集的分布是这样的：训练集占比为 70% (7120 对)，验证集占比为 10% (1024 对)，测试集占另外 20% (2048 对)。

LEVIR-CD+ 数据集：LEVIR-CD+ 是对现有 LEVIR-CD 数据集的一个扩展版本，包含的样本数量扩充到 985 对 1024 × 1024 像素的图像对，其中来自 LEVIR-CD 的 637 对用于训练，我们将其中的 10% 用于验证，其余扩展的 385 对用于测试。

WHU-CD 数据集：原始数据集由单个双时相图像对组成，其中包括 2012 年和 2016 年拍摄的新西兰克赖斯特彻奇的两张航拍图像。以同样的切割方式，我们将其分成 7434 个不重叠的图像对，每个图像对的大小为 256 × 256 像素。训练、验证和测试数据集分别由 5947、743 和 744 对图像组成。

GZ-CD 数据集：GZ-CD 数据集的双时相图像收集于 2006 年至 2019 年，覆盖中国

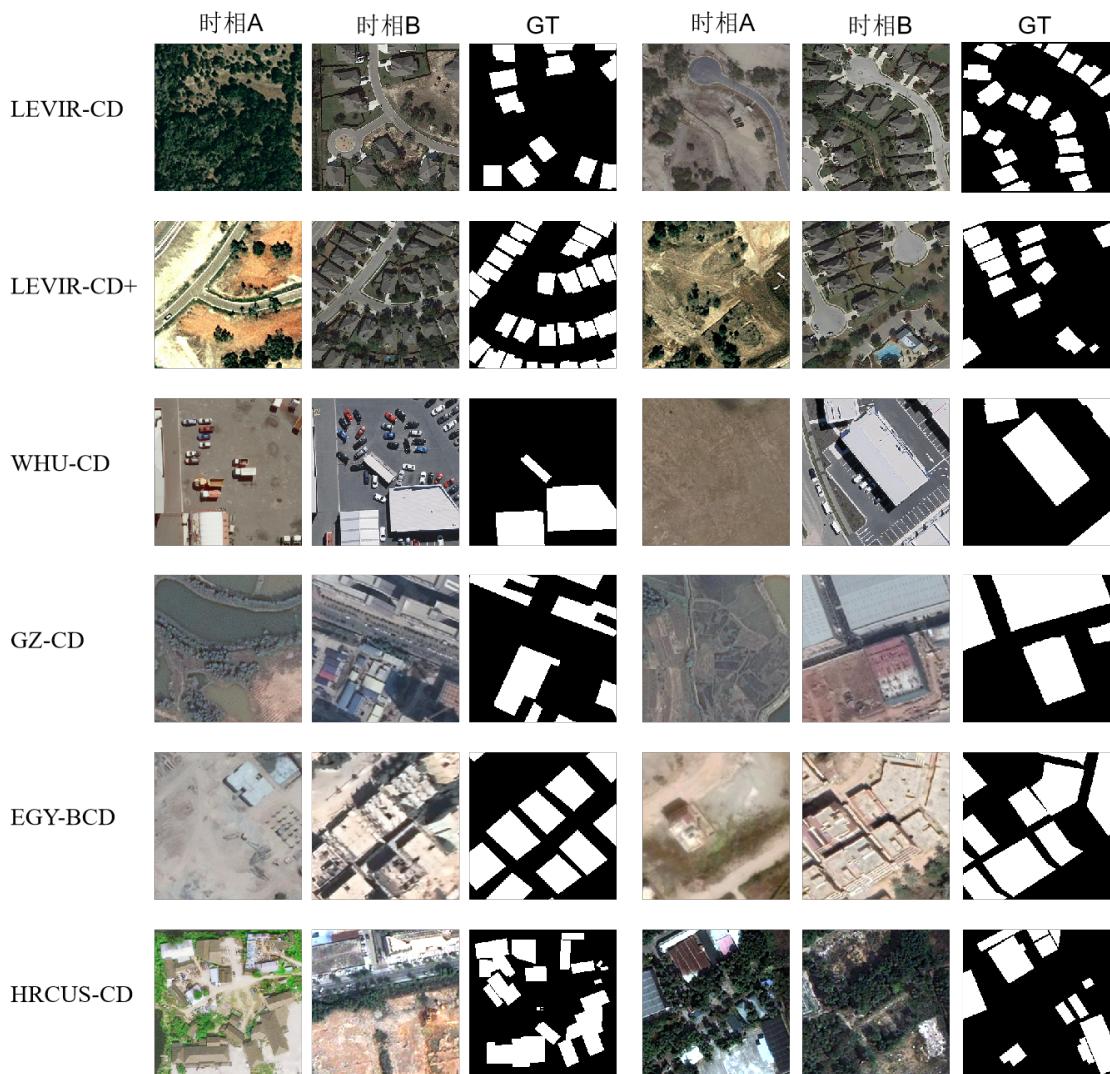


图 2-10 建筑物变化检测数据集部分样例图

广州郊区。共收集了 19 幅不同尺寸的红、绿、蓝波段空间分辨率为 0.55m 的超分辨率图像，注释集中在建筑物变化上。经过切割成统一尺寸之后，共包含 3603 对样本，其中训练样本 2882 对，验证样本 360 对，测试样本 361 对。

EGY-BCD 数据集： EGY-CD 数据集包含 6091 张 2015 - 2022 年拍摄的 256×256 像素的图像对，空间分辨率为 0.25m/像素，主要标注了埃及 4 个城市和沿海地区的建筑变化区域。不用额外处理，我们将其中 70% 用于训练，20% 用于验证，剩下的 10% 用于测试。

HRCUS-CD 数据集： 该数据集由 11388 对高分辨率遥感图像组成，裁剪为 256×256 像素，空间分辨率为 0.5m。主要分为两个征集区，一个是城市建成区，时间跨度从 2019 年到 2022 年，建筑变化面积较少；另一个是正在建设的新城区，从 2010 年到 2018 年，包含农田、山地等多种地貌，建筑物变化面积较大。

CDD 数据集： CDD 数据集包含 16000 对 256×256 像素的双时相图像对，像素分辨率从 0.03 到 1 米不等。所有这些双时间图像都是从谷歌地球收集的 7 对 4725×2700 像

素的季节性变化图像对中裁剪出来的。分别有 10000 对、3000 对和 3,000 对用于训练、验证和测试。

DSIFN-CD 数据集: DSIFN-CD 是在谷歌地球上人工采集的，它由覆盖中国 6 个城市（北京、成都、深圳、重庆、武汉、西安）的 6 幅大型双时相超高分辨率图像组成。其中五个大图像对（北京、成都、深圳、重庆、武汉）被裁剪成 394 个子图像对，大小为 512×512 像素。经过数据增强后，得到 3940 对双相图像。将西安图像对裁剪为 48 个子图像对进行测试。最终切割为统一的 256×256 像素之后，训练数据集中有 14400 对图像，验证数据集中有 1360 对图像，测试数据集中有 192 对图像。

SYSU-CD 数据集: 该数据集由 2007 年至 2014 年在香港拍摄的 20000 对 256×256 像素、空间分辨率为 0.5 米航拍图像组成。数据集中的主要变化类型包括：(a) 新建城市建筑；(b) 郊区扩张；(c) 施工前基础工作；(d) 植被变化；(e) 扩大道路；(f) 近海建筑。我们将其中 70% 的样本用于训练，20% 的样本用于验证，剩下的 10% 的样本集用于测试。

CL-CD 数据集: CL-CD 数据集由 600 对农田变化样本图像组成，其中 320 对用于训练，120 对用于验证，120 对用于测试。CL-CD 的双时相影像是 2017 年和 2019 年由中国广东省高分二号卫星采集的，空间分辨率范围为 0.5-2 m。每组样本由两个 512×512 像素图像和对应的字段变化的二进制标签组成。CL-CD 中指出的主要变化类型包括建筑物、道路、湖泊和裸露的土地。我们将训练、验证、测试集都切割为 256×256 大小之后进行使用。

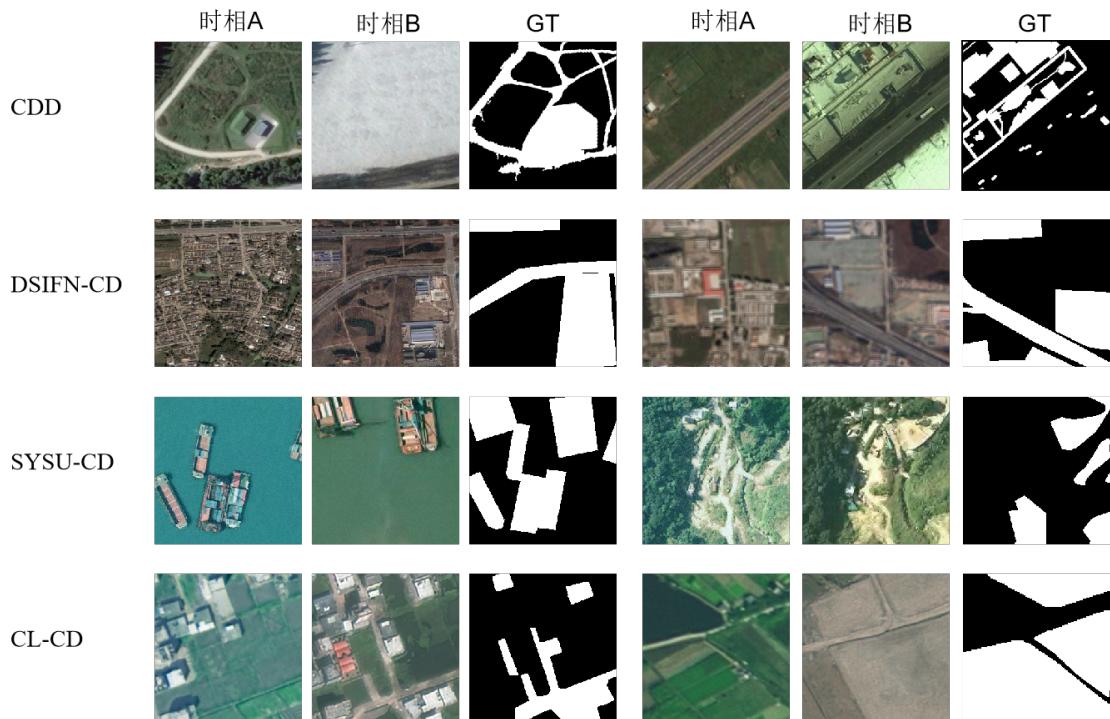


图 2-11 多类变化检测数据集部分样例图

2.5 本章小结

本章主要介绍了本文研究内容的相关技术，首先简要介绍了本研究用到的 ResNet 卷积神经网络，随后阐述了与之相关的 Vison Transformer 模型的架构和原理。接着大致介绍了本文研究中进行对比实验的几种半监督变化检测 SOTA 方法。以及实验中使用的变化检测的两种最常用的评价指标。最后介绍了本文使用的十个公开数据集，所有实验均在这十个数据集上进行训练和验证。

第3章 基于自适应动态阈值的半监督变化检测算法

3.1 引言

半监督变化检测方法中基于伪标记和一致性正则化的方法取得了巨大的成功，其关键思想在于模型应该根据半监督学习中的平滑假设和低密度假设，在不同的扰动下对相同的未标记数据产生类似的预测或相同的伪标签。然而，这种方法在利用未标记数据时存在明显局限性。由于它们要么采用固定阈值，要么依赖临时性的阈值调整方案，仅能筛选出部分可信的未标记样本用于训练。更重要的是，这些方法缺乏针对不同数据分布的自适应调整机制，导致模型性能欠佳且收敛速度缓慢。

通常情况下，为保证伪标签的质量，需要设定一个高阈值。然而，固定的高阈值可能导致早期训练阶段的数据利用率低，并且忽略了不同类别的学习难度差异。*Dash*^[107] 和 *admatch*^[108] 等半监督学习方法提出随着训练的进行，逐渐增加固定的全局阈值（特定于数据集）。虽然这些方法提高了对未标记数据的利用率，但这种特设阈值调整方案是由超参数进行控制，因此与模型的学习过程脱节。*FlexMatch*^[109] 表明，不同的类应该有不同的局部阈值（特定于类别）。虽然局部阈值考虑了不同类别的学习差异，但它们仍然是从预定义的固定全局阈值映射出来的。*Adsh*^[110] 通过优化每个类的伪标签数量，从预定义的非平衡半监督学习阈值中获得自适应阈值。总而言之，这些方法在根据模型的学习进度调整阈值方面均存在一些不足之处，从而阻碍了训练过程，特别是当标记数据过于稀缺而无法提供足够的监督时。

此外，此前的基于伪标签的半监督变化检测方法中，对于那些被筛选掉的低置信度预测都是直接丢弃，而这样可能造成的问题就是，训练后的模型可能会过度拟合到容易学习的样本上，而忽略了难学习的样本，从而产生“马太效应”，即本来就强的变得更强，本来就弱的变得更弱。

因此，我们认为需要根据模型的学习状态来自适应地确定阈值。具体来说，使用一个较低的全局阈值来利用更多的未标记数据，并在模型早期训练阶段加快收敛速度。随着训练过程的进行，预测置信度不断增加，此时使用更高的全局阈值来过滤掉错误的伪标签来减轻确认偏差。此外，由于变化检测中前景和背景类别的极度不平衡，还应该根据模型对变化和不变类预测的置信度，分别为前景、背景定义一个局部阈值，以减轻类别不平衡带来的预测偏差。并且我们增加了一个低置信度学习模块，旨在最大限度地利用整个未标记数据集，以提高训练过程中的泛化程度。具体来说，我们将来自低置信度预测的软伪标签中可能存在的正确信息通过一致性约束利用起来，以提高模型的性能和表征能力，最终提出了 *AdTSemiCD* (*Semi-supervised change detection with an adaptive dynamic thresholding*) 半监督变化检测框架。

本章介绍了基于自适应动态阈值的半监督变化检测算法——*AdTSemICD*，首先整

体性的阐述了算法的框架和训练流程，随后详细介绍了基于预测概率分布的全局动态阈值和类别动态阈值，以及设计的低置信度学习模块。最后在实验部分报告了本章方法在十个公开数据集上取得的实验结果，从定性和定量两个角度进行了全方位的对比，以及全面的消融实验分析，验证了 AdTSemiCD 的有效性。

3.2 基于自适应动态阈值的半监督变化检测框架

3.2.1 整体框架

图3-1展示了 AdTSemiCD 的网络框架。半监督变化检测的任务描述如下：给定一个标记数据集 $D_l = \{\{x_{a,i}^l, x_{b,i}^l\}, y_i^l\}_{i=1}^m$ ，一个无标注数据集 $D_u = \{x_{a,j}^u, x_{b,j}^u\}_{j=1}^n$ ，其中 $\{\{x_{a,i}^l, x_{b,i}^l\}, y_i^l\}$ 代表第 i 个双时相图像对和真实标签，以及 $\{x_{a,j}^u, x_{b,j}^u\}$ 代表第 j 对未标记图像对，下标 a 和 b 分别是用于标识双时相图像对的前、后时相的图像，标注数据集和无标注数据集的样本数量分别是 n 和 m ，并且 $n \gg m$ ，模型不仅可以从 D_l 中提取有意义的信息，还可以通过利用 D_u 中大量未标记的训练样本来捕获更广泛的特征，以提高模型的泛化能力。

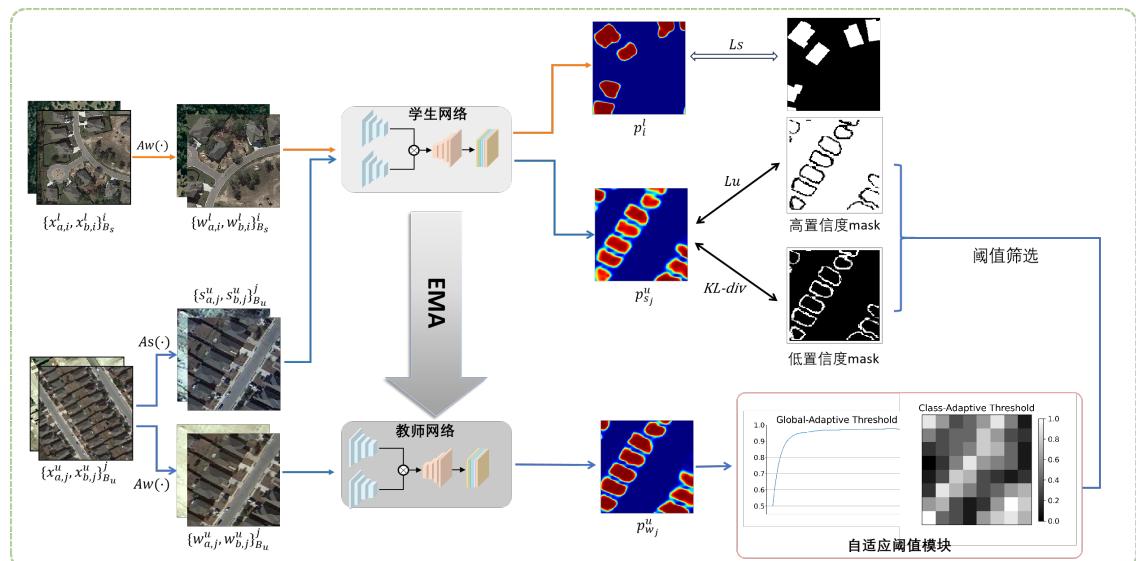


图 3-1 AdTSemiCD 算法的整体框架

模型架构：在本研究中，我们采用广泛应用的平均师生框架来完成半监督变化检测任务。该网络由两个组件组成，学生模型 M_{stu} 和教师模型 M_{tea} 具有相同的网络结构，分别由 θ_s 和 θ_t 参数化。其权重通过梯度下降方法进行优化，训练学生模型从少量标记样本和大量未标记样本中提取变化特征。相反，而教师模型 M_{tea} 用于生成伪标签来指导学生学习未标记的数据，教师模型通过 EMA 方法更新参数。

无标记样本在输入到教师和学生网络之前会分别经历弱增强 $A_w(\cdot)$ 以及进一步的强增强 $A_s(\cdot)$ 过程，以提供输出一致性正则化项。

训练目标：目标是最小化模型在 D_l 上训练的监督损失 L_s 和在 D_u 上训练的无监督损失 L_u 。在训练过程中，样本以随机打乱的小批次（标记样本批次 B_l 和无标记样本批

次 B_u ）的形式输入网络。一个小批次内在标记样本上的 L_s 损失计算为模型对其输出的预测概率 P_i^l 与真实标签 y_i^l （Ground Truth, GT）之间的交叉熵（Cross Entropy, CE）：

$$\mathcal{L}_s = \frac{1}{|\mathcal{B}_l|} \sum_{i=1}^{|\mathcal{B}_l|} \text{CE}(p_i^l, y_i^l) \quad (3-1)$$

其中 $|\mathcal{B}_l|$ 表示标记样本批次的大小， $p_i^l = M_{stu}^\theta(x_{a,i}^l, x_{b,i}^l)$ 表示学生模型对第 i 对标注图像对的预测概率。

一个批次内在未标记样本上的无监督损失 L_u 也类似，这里我们使用来自 M_{tea} 的伪标签作为监督，因此 L_u 计算为：

$$\mathcal{L}_u = \frac{1}{|\mathcal{B}_u|} \sum_{j=1}^{|\mathcal{B}_u|} \text{CE}(p_{s,j}^u, \hat{y}_j^u), \quad (3-2)$$

$$\hat{y}_j^u = \arg \max(p_{w,j}^u) > \tau \quad (3-3)$$

其中 $|\mathcal{B}_u|$ 表示未标记样本小批次的大小， $p_{w,j}^u = M_{tea}^\theta(w_{a,j}^u, w_{b,j}^u)$ 表示教师模型对第 i 对弱增强未标记图像对的变化的预测概率结果，进一步利用 $\arg \max$ 算子对其进行取最大运算得到伪标签 \hat{y}_j^u ，如公式3-3所示， $p_{s,j}^u = M_{stu}^\theta(s_{a,j}^u, s_{b,j}^u)$ 表示教师模型对第 i 对强增强未标记图像对的变化检测预测概率， τ 为固定的筛选阈值，用于剔除低置信的预测。

综上所述，AdaSemiCD 训练过程总的损失为：

$$\mathcal{L} = \mathcal{L}_s + \lambda(\cdot)\mathcal{L}_u \quad (3-4)$$

在本研究中，为了加快模型收敛速度，与以往的研究^{[50][51]}保持一致， $\lambda(\cdot)$ 采用固定超参数表示。

训练策略：学生网络的参数 θ 通过随机梯度下降（SGD）技术最小化总训练损失来优化，以最小化总体损失 \mathcal{L} 。而教师网络的参数 θ' 通过学生模型参数 θ 在训练时间序列上的指数移动平均来更新，如3-5所示。超参数 β 作为动量参数， β 值越大，移动指数平均的窗口越宽。通常， β 选择接近 1.0 的数字，例如在本研究中，选择 0.996。

$$\theta' = \beta\theta' + (1 - \beta)\theta \quad (3-5)$$

问题定义：监督学习过程即通过最小化学生模型预测概率与真实样本标签之间的交叉熵损失来优化模型的参数；无监督学习过程通常情况下将教师模型的输出作为训练指导，对于教师模型在无标记样本上的输出概率，通过一个阈值来筛选出可信的预测作为伪标签，然后最小化学生模型输出概率和伪标签之间的交叉熵损失来优化学生模型的参数。教师模型的参数更新则由学生模型参数的移动指数平均累计得到。此外通常情况下，在此框架中还引入了一致性正则化，即对无标记样本分别进行弱增强以及在此之上的进一步的强增强，分别作为教师模型和学生模型的输入，基于低密度假设和平滑假设，

二者的输出结果应该保持一致。通过一致性正则化，能够进一步提高模型的泛化能力。

很明显，生成伪标签过程中的阈值筛选是一个非常重要的过程，阈值过高会导致对无标记数据的利用率不足，而阈值过低又会导致引入错误标签。并且对于不同数据分布，最佳阈值的选择往往不相同，需要通过实验来验证。因此，我们设计了一种自适应的动态阈值调整机制，根据模型在不同数据上的预测概率分布，动态更新模型对于前景和背景的概率筛选阈值，即公式??中的 τ 为自适应动态改变的。此外，还设计了一个额外的低置信度学习模块，进一步提高对无标记数据的利用率。我们将在下面的小节中分别对这些设计进行详细阐述。

3.2.2 全局自适应阈值

在我们的自适应阈值机制中，首先估计一个全局阈值作为模型置信度的均线。当训练开始时，阈值较低，此时接受更多可能正确的样本进入训练以加快模型的收敛速度。随着模型预测变得更加自信，全局阈值自适应地增加，以过滤掉可能不正确的样本，减少确认偏差。具体说来，我们基于以下两个原则设计全局阈值。首先，全局阈值应该与

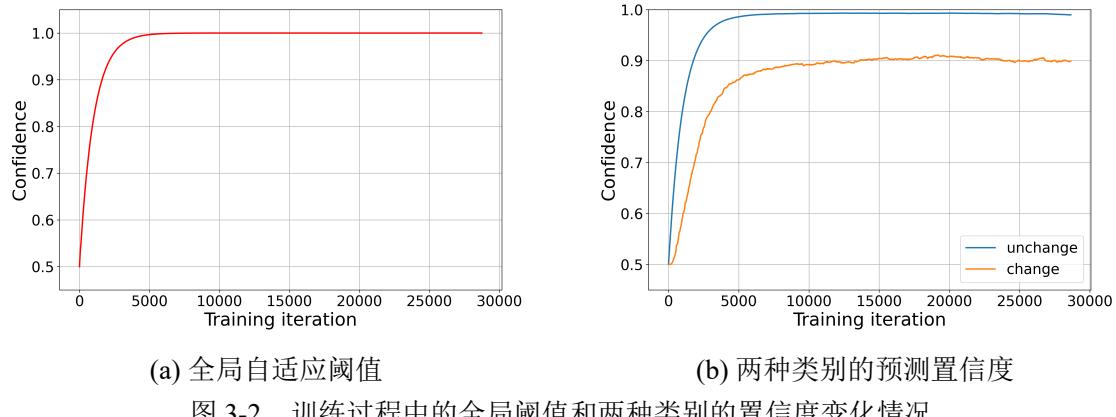


图 3-2 训练过程中的全局阈值和两种类别的置信度变化情况

模型对未标记数据预测的概率分布相关，要尽可能地反映模型的整体学习状况。此外，全局阈值应该在训练过程中稳定地增加，以确保丢弃错误的伪标签。考虑到以上两点，我们将全局阈值 τ_t 设置为模型对未标记数据的平均置信度，其中 t 表示第 t 个迭代轮次。为了使得阈值保持平滑，我们将全局置信度估计为每个训练时间步的置信度的指数移动平均（EMA），并且在训练初始时间将全局阈值设置为每种类别的平均预测概率，即 0.5。综上，全局动态阈值的计算公式如3-6所示：

$$\tau_t = \begin{cases} 0.5, & \text{if } t = 0 \\ \beta\tau_{t-1} + (1 - \beta)\frac{1}{|\mathcal{B}_u|} \sum_{b=1}^{|\mathcal{B}_u|} \max(p_{w,j}^u), & \text{otherwise} \end{cases} \quad (3-6)$$

其中， $p_{w,j}^u$ 为教师模型在弱增强输入图像上的预测概率输出， β 与公式3-5中一致，为 EMA 的衰减动量。

3.2.3 类别自适应阈值

如图3-3所示，变化检测任务存在严重的类别失衡问题，模型在背景类别上的预测必然更加自信。类别动态阈值旨在以特定于前景变化类别和背景不变类的方式调节各自的阈值，以考虑类内多样性和两种类别的差异性。我们分别计算模型对两种类别的预测概率分布的期望，以此来估计模型特定于类别的学习状态，具体地说，就是基于模型特定于类别的平均置信度，动态调整全局置信度在每个类别上的表示。同样地，我们也是将类别平均置信度估计为在每个训练时间步的置信度的指数移动平均（EMA），并且在训练初始时间将类别置信度设置为一个均匀分布，即（0.5, 0.5）。计算公式如下：

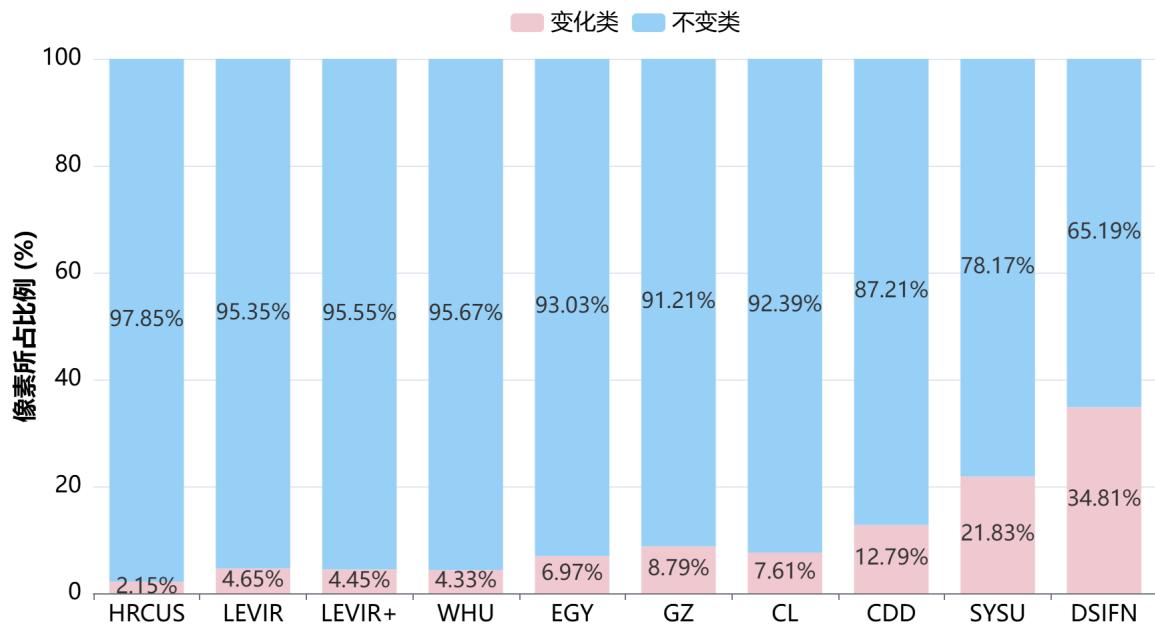


图 3-3 公开变化检测数据集的类别不平衡统计

$$\tilde{p}_t(c) = \begin{cases} 0.5, & \text{if } t = 0 \\ \lambda\tilde{p}_{t-1}(c) + (1 - \lambda)\frac{1}{|\mathcal{B}_u|} \sum_{b=1}^{|\mathcal{B}_u|} p_{w,j}^u(c), & \text{otherwise} \end{cases} \quad (3-7)$$

其中 $\tilde{p}_t = [\tilde{p}_t[0], \tilde{p}_t[1]]$ 是包含前景和背景类别的预测概率分布的期望。然后对其进行一个最大归一化，将归一化后的值作为类别阈值权重，最终特定于类别的自适应动态阈值结合了全局动态阈值和类别置信度期望，公式表示为：

$$\mathcal{W}(c) = \frac{\tilde{p}_t(c)}{\max \{\tilde{p}_t(c) : c \in [0, 1]\}}, \quad (3-8)$$

$$\tau_t(c) = \mathcal{W}(c) \times \tau_t \quad (3-9)$$

则式3-2中的伪标签 \hat{y}_i^u 的筛选机制被优化为：

$$\hat{y}_i^u = p_{w,j}^u(\arg \max (p_{w,j}^u)) \geq \tau_t (\arg \max (p_{w,j}^u)), \quad (3-10)$$

3.2.4 低置信度学习

在上一步的筛选过程中，剔除掉了许多低置信度的标签，仅使用高置信度的预测作为硬伪标签。这一做法基于这样一种假设：高置信度的标签都是正确的，而低置信度的标签都是错误的。但是我们在研究中发现，被筛选掉的一部分低置信度标签中也包含着正确的标签信息，但是在模型学习过程中没有得到充分的利用。因此我们将这部分低置信的标签并不完全丢弃，而是将其作为软伪标签，利用强弱增强图像对之间的输出一致性正则化对其进行学习。具体地，我们使用 KL 散度作为低置信度学习的损失函数，计算公式如下：

$$\mathcal{L}_{KL} = \frac{1}{|\mathcal{B}_u|} \sum_{b=1}^{|\mathcal{B}_u|} D_{KL}(p_{s,j}^u, p_{w,j}^u) \cdot (p_{w,j}^u(\arg \max(p_{w,j}^u)) < \tau_t(\arg \max(p_{w,j}^u))) \quad (3-11)$$

综上，我们的 AdTSemiCD 的总体损失被重新定义为：

$$\mathcal{L}_{total} = \mathcal{L}_s + \lambda_u \mathcal{L}_u + \lambda_k \mathcal{L}_{KL} \quad (3-12)$$

λ_u 和 λ_k 分别是硬标签交叉熵损失和软标签 KL 散度损失的权重超参数，我们会在消融实验部分对其取值选择进行探究。

3.3 实验结果及分析

数据划分：我们对所有数据集遵循相同的划分规则，定义了四种半监督设置，我们分别选择训练集中 5%，10%，20% 和 40% 作为标记样本训练集，训练集中其他样本均视作无标记样本，在训练过程中不使用任何真实标注信息。其中在 LEVIR-CD 和 WHU-CD 两个数据集上我们遵循^{[6][7]} 中现有的半监督划分，在其他数据集上我们进行了随机划分，但在所有方法的对比实验上我们都是使用此相同的训练样本集，以进行公平和科学的实验对比。

实现细节：为了防止网络结构带来的性能差异影响我们的结论，我们所有对比方法均使用 ResNet50+PPM 网络作为我们的变化检测网络。我们将初始学习率设置为 0.01，它以 0.9 的动量随训练进程逐渐线性降低到 1e-4，并使用 SGD 优化器训练所有方法。所有的模型都进行了 80 个 epoch 的训练，每个批次的标记和未标记每次迭代的小批次大小均设置为 8。此外，所采用的弱增强包括随机翻转、随机缩放、随机裁剪，强增强包括^[11] 中引用的 9 种强增强方式。此外，对于全局动态阈值，我们设定了与对比试验中固定阈值（0.95）一样的上限，到达上限之后就保持此阈值不再变化。所有的实验均利用 PyTorch 框架实现并且在 4 块 Tesla V100-SXM2-32GB GPU 上进行模型训练和验证。

3.3.1 对比试验

我们与 RCR 等几种半监督变化检测算法和 s4GAN 等半监督语义分割方法进行了全方位的对比实验，包括以 IoU^c 和 OA 为指标的定量实验结果对比，以及在每个数据

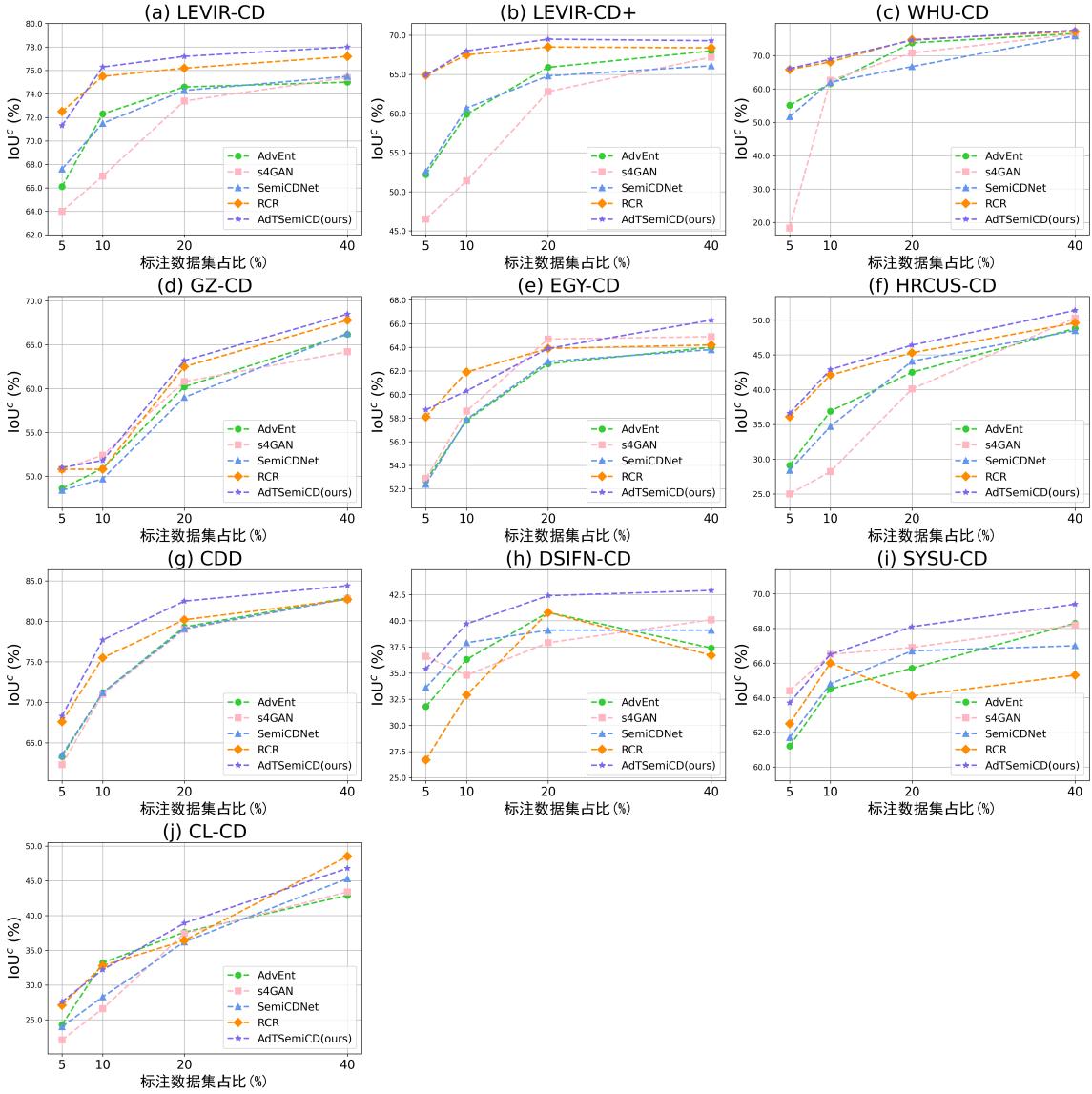


图 3-4 AdTSemiCD 与 SOTA 方法在 10 个数据集上的性能对比折线图

集的测试集上的推理结果可视化对比，并且 AdTSemiCD 都取得了优异的结果。

(1) 定量对比

表3-1和表3-2分别展示了我们在六个二值建筑物变化检测数据集和四个多类别变化检测数据集上的定量实验结果，其中蓝色高亮标注并且加粗显示的数值为最优结果，下划线标注的数值为次优结果。此外，Oracle 表示在整个训练数据集上以 100% 标注率进行训练的结果，为我们对比的上界。

我们的 AdTSemiCD 在几乎所有二值变化检测数据集上都取得了最好的结果，如表3-1所示。尤其是在 LEVIR-CD+ 和 HRCUS-CD 两个数据集上在所有实验设置下都达到了最佳，在 IoU^c 上分别获得了 0.6 和 0.9 和百分点的提升。在其余数据集上除了个别半监督实验设置下，整体上相比于所有对比方法也都取得了一定的提升效果。但是，需要说明的情况是，我们的方法在标注率极低（例如 5% 时），我们设计的自适应动态阈

表 3-1 不同 CD 方法在二值建筑物变化检测数据集上的平均定量性能对比

数据集	方法	5%		10%		20%		40%	
		IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
LEVIR-CD	仅监督	61.0	97.60	66.8	98.13	72.3	98.44	74.9	98.60
	AdvEnt ^[112]	66.1	98.08	72.3	98.45	74.6	98.58	75.0	98.60
	s4GAN ^[113]	64.0	97.89	67.0	98.11	73.4	98.51	75.4	98.62
	SemiCDNet ^[5]	67.6	98.17	71.5	98.42	74.3	98.58	75.5	98.63
	RCR ^[6]	72.5	98.47	<u>75.5</u>	<u>98.63</u>	<u>76.2</u>	<u>98.68</u>	<u>77.2</u>	<u>98.72</u>
	AdTSemiCD	<u>71.3</u>	<u>98.40</u>	76.3	98.66	77.2	98.75	78.0	98.79
Oracle		$IoU^c=77.9$ and OA=98.77							
LEVIR-CD+	仅监督	52.0	97.72	58.4	98.06	66.1	98.31	66.2	98.42
	AdvEnt ^[112]	52.2	97.68	59.9	98.11	65.9	98.37	68.0	98.51
	s4GAN ^[113]	46.5	97.25	51.4	97.66	62.8	98.18	67.2	98.46
	SemiCDNet ^[5]	52.6	97.66	60.7	98.24	64.8	98.37	66.1	98.38
	RCR ^[6]	<u>64.9</u>	<u>98.25</u>	<u>67.5</u>	<u>98.45</u>	<u>68.5</u>	<u>98.52</u>	<u>68.4</u>	<u>98.51</u>
	AdTSemiCD	64.9	98.26	68.0	98.48	69.5	98.59	69.3	98.60
Oracle		$IoU^c=70.5$ and OA=98.63							
WHU-CD	仅监督	50.0	97.48	55.7	97.53	65.4	98.20	76.1	98.94
	AdvEnt ^[112]	55.1	97.90	61.6	98.11	73.8	98.80	76.6	98.94
	s4GAN ^[113]	18.3	96.69	62.6	98.15	70.8	98.60	76.4	98.96
	SemiCDNet ^[5]	51.7	97.71	62.0	98.16	66.7	98.28	75.9	98.93
	RCR ^[6]	<u>65.8</u>	<u>98.37</u>	<u>68.1</u>	<u>98.47</u>	74.8	98.84	<u>77.2</u>	<u>98.96</u>
	AdTSemiCD	66.1	98.41	68.9	98.54	<u>74.6</u>	<u>98.83</u>	77.6	99.03
Oracle		$IoU^c=85.5$ and OA=99.38							
GZ-CD	仅监督	47.5	93.56	51.4	94.26	58.0	95.65	66.3	96.62
	AdvEnt ^[112]	48.6	<u>94.39</u>	50.9	94.89	60.2	95.79	66.2	96.58
	s4GAN ^[113]	<u>50.8</u>	94.38	52.4	94.98	60.8	95.94	64.2	96.39
	SemiCDNet ^[5]	48.4	93.58	49.7	94.79	59.0	95.66	66.3	96.57
	RCR ^[6]	<u>50.8</u>	93.82	50.8	94.69	<u>62.5</u>	<u>96.07</u>	67.8	96.61
	AdTSemiCD	51.0	94.47	<u>51.8</u>	<u>94.93</u>	63.2	96.26	68.5	96.81
Oracle		$IoU^c=69.0$ and OA=96.93							
EGY-CD	仅监督	49.8	95.73	54.6	96.38	61.4	96.83	<u>65.1</u>	97.25
	AdvEnt ^[112]	52.7	96.01	57.8	96.58	62.6	96.86	64.0	97.19
	s4GAN ^[113]	52.9	95.94	58.6	96.50	64.7	97.09	64.9	97.27
	SemiCDNet ^[5]	52.4	96.00	57.9	96.31	62.8	96.95	63.8	97.19
	RCR ^[6]	<u>58.1</u>	<u>96.50</u>	<u>59.9</u>	<u>96.77</u>	<u>63.9</u>	<u>97.08</u>	64.2	97.18
	AdTSemiCD	58.7	96.55	60.3	96.80	<u>63.9</u>	<u>97.05</u>	66.3	97.41
Oracle		$IoU^c=67.6$ and OA=97.54							
HRCUS-CD	仅监督	29.5	98.11	36.0	98.45	43.4	98.68	48.9	98.84
	AdvEnt ^[112]	29.1	98.11	36.9	98.40	42.5	98.61	48.8	98.71
	s4GAN ^[113]	25.0	97.86	28.2	98.24	40.1	98.62	<u>50.3</u>	<u>98.85</u>
	SemiCDNet ^[5]	28.4	98.00	34.7	98.44	44.1	98.68	48.5	98.74
	RCR ^[6]	<u>36.1</u>	<u>98.36</u>	<u>42.1</u>	<u>98.69</u>	<u>45.3</u>	<u>98.76</u>	49.6	98.66
	AdTSemiCD	36.6	98.43	42.9	98.72	46.4	98.81	51.4	98.85
Oracle		$IoU^c=59.0$ and OA=99.06							

表3-2 不同CD方法在多类变化检测数据集上的平均定量性能对比

数据集	方法	5%		10%		20%		40%	
		IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
CDD-CD	仅监督	60.4	94.25	67.9	95.46	75.6	96.59	82.3	97.56
	AdvEnt ^[112]	63.3	94.65	71.2	96.01	79.3	97.14	82.9	97.66
	s4GAN ^[113]	62.3	94.69	71.0	95.94	79.0	97.10	82.8	97.63
	SemiCDNet ^[5]	63.5	94.68	71.2	95.99	79.1	97.13	82.8	97.63
	RCR ^[6]	67.6	95.40	75.5	96.57	80.2	97.26	82.7	97.61
	AdTSemiCD	68.3	95.52	77.7	96.92	82.5	97.54	84.4	97.77
Oracle		$IoU^c=87.8$ and OA=98.10							
DSIFN-CD	仅监督	34.8	78.34	<u>38.9</u>	83.41	40.2	87.00	39.6	87.00
	AdvEnt ^[112]	31.8	77.83	36.3	83.86	<u>40.8</u>	85.92	37.4	86.31
	s4GAN ^[113]	36.6	84.10	34.8	86.87	37.9	<u>87.69</u>	<u>40.1</u>	86.52
	SemiCDNet ^[5]	33.6	78.60	37.9	84.18	39.1	86.77	39.1	<u>87.05</u>
	RCR ^[6]	26.7	83.78	32.9	86.05	40.8	86.70	36.7	86.08
	AdTSemiCD	<u>35.4</u>	79.83	39.7	84.58	42.4	87.74	42.9	87.90
Oracle		$IoU^c=58.1$ and OA=90.82							
SYSU-CD	仅监督	62.9	89.57	64.4	90.18	66.0	90.82	66.4	90.93
	AdvEnt ^[112]	61.2	89.36	64.5	90.18	65.7	90.35	<u>68.3</u>	91.24
	s4GAN ^[113]	64.4	90.02	<u>66.5</u>	90.48	<u>66.9</u>	90.26	68.2	<u>91.51</u>
	SemiCDNet ^[5]	61.7	89.32	64.8	90.25	66.7	<u>90.97</u>	67.0	91.08
	RCR ^[6]	62.5	89.76	66.0	<u>90.75</u>	64.1	90.22	65.3	90.56
	AdTSemiCD	<u>63.7</u>	<u>89.92</u>	66.5	90.89	68.1	91.26	69.4	91.96
Oracle		$IoU^c=68.2$ and OA=91.64							
CL-CD	仅监督	18.1	91.90	31.4	92.42	37.2	<u>93.32</u>	45.9	94.98
	AdvEnt ^[112]	24.3	92.13	33.2	93.01	<u>37.6</u>	<u>93.59</u>	42.9	94.06
	s4GAN ^[113]	22.1	92.00	26.6	93.09	37.4	<u>93.59</u>	43.4	93.87
	SemiCDNet ^[5]	24.0	<u>92.20</u>	28.3	93.42	36.2	92.41	45.3	94.22
	RCR ^[6]	<u>27.1</u>	91.63	<u>32.8</u>	92.99	36.4	93.07	48.5	94.94
	AdTSemiCD	27.6	92.32	32.2	<u>93.13</u>	38.9	93.75	46.8	95.09
Oracle		$IoU^c=50.1$ and OA=95.66							

值带来的改进效果比较微弱，这可能是由于标记数据的极端稀少导致模型的检测分割质量比较差，而由此导致的结果就是整体预测置信度较低，从而自适应动态阈值在训练初期长时间保持较低的值，因此虽然提高了标签利用率但是引入了更多的噪声。而当标记数据量增加时，模型能够生成更加可信的预测概率，在提高标签利用率的同时，能够保证其正确性。因此表现在实验结果就是，在更高的标记率下，AdTSemiCD 可以更加稳定和高效地提升变化检测性能。

在多类别变化检测数据集上，我们的方法是能够在有限的标记数据上，能够通过额外的无标记训练样本稳定地提升变化检测的性能，如表3-2所示。而其余的对比方法在某些场景下会失效，性能甚至低于仅在有限的标记样本上进行监督训练的基线，例如表中 DSIFN-CD 和 SYSU-CD 数据集上 5% 标记率下，除 s4GAN 和我们的 AdTSemiCD 之外，其余方法都比基线更差。这是由于多类别变化检测任务其本身比单类变化检测更加复杂模型在其中某一类上的性能表现好坏会对整体性能产生影响，而我们的方法在阈值调整时，考虑了类别的预测置信度，对于欠学习的变化特征，会通过调整阈值来强化对该类别的学习。总体来看，AdTSemiCD 虽然没能够在所有实验设置下都达到最佳，但是总体上还是达到了性能优化的目的。

(2) 定性对比

我们对所有数据集上的对比实验在相应的测试集上所取得的变化检测结果都进行了可视化，以更加直观地对比不同方法的变化检测效果。表3-6、表3-5、表3-8、表3-9分别展示了在 5% 标记样本训练下，不同半监督方法在 LEVIR-CD、WHU-CD、CDD、CL-CD 四个数据集上的预测可视化结果，每组结果中的第一行为预测概率热力图表示是，第二行为预测结果，红色表示变化区域，其余为背景区域。

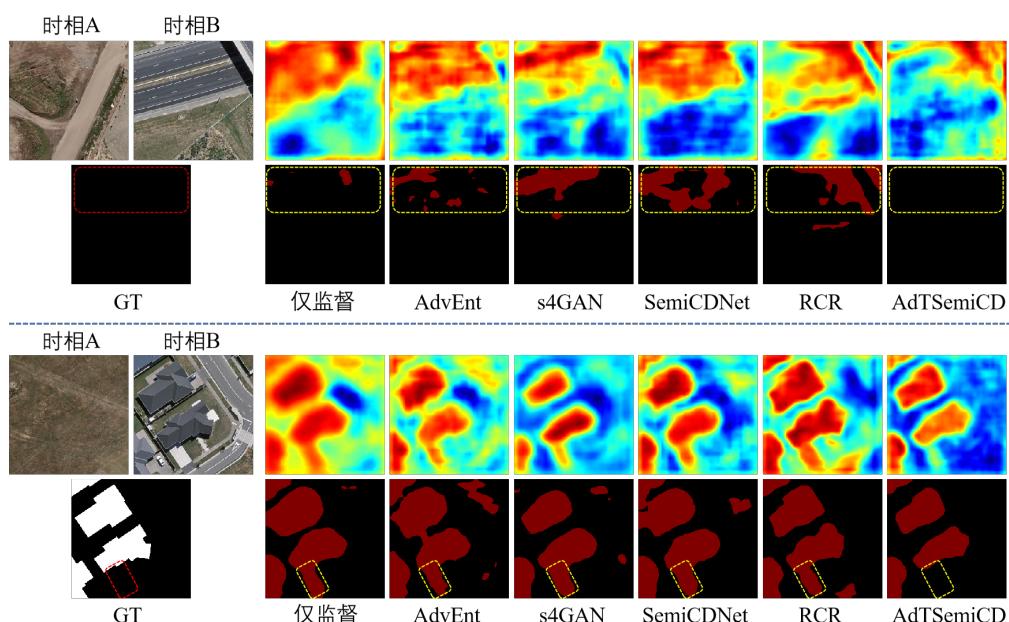


图 3-5 5% 标记训练样本下 AdTSemiCD 与 SOTA 方法在 WHU-CD 测试集上的可视化对比图

显然在 LEVIR-CD 数据集上，我们的 AdTSemiCD 能够能够更好地识别边界，并且从热力图可以看出我们的方法在变化类别和不变类别上的预测置信度都比较高，区分度比较明显，这是因为我们提出的类别自适应阈值能够保证模型对两种类别的学习比较平衡并且充分学习到更多的上下文知识。在 WHU-CD 数据集上，我们的 AdTSemiCD 更好地排除了建筑物变化目标之外的干扰项，例如图3-5中的高速公路和水泥地，这些类型与建筑物比较相似，并且由于训练的标记样本稀少，进一步提高了准确识别建筑物变化检测的挑战性。AdTSemiCD 更加充分地挖掘了样本中的隐藏信息，进而提升了在这些区域的识别能力。在 CDD 数据集上，本章方法的一个特点就是在细节上表现更佳，比如第一组可视化结果中草地间的土路。而 CL-CD 数据集上，所有方法比较容易出现的问题就是大面积的错检、漏检，这对定量结果中 IoU 指标的影响比较剧烈，问题原因是标记样本的稀少，CL-CD 数据集是十个数据集中数据规模最小的，在半监督设置下就更加稀少了。我们的方法在这些情况下表现良好，虽然边界识别可能并非完全准确，存在一些误差，但是总体上表现尚可，从热力图上可以看出，相比其他半监督方法，AdTSemiCD 能够正确的判定出变化目标。

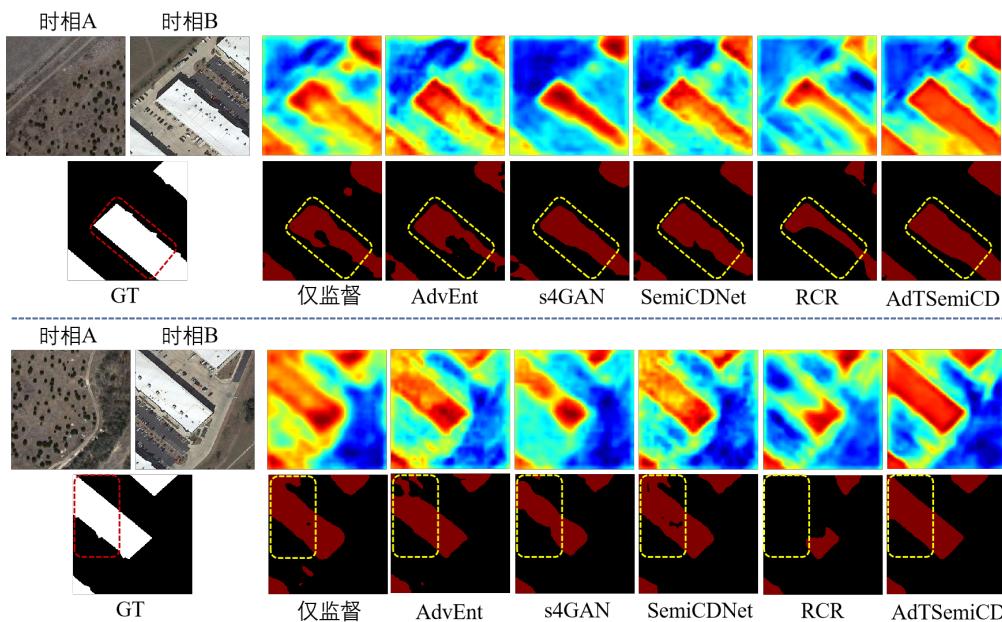


图 3-6 5% 标记训练样本下 AdTSemiCD 与 SOTA 方法在 LEVIR-CD 测试集上的可视化对比图

图3-7展示了不同标记率下训练的模型在相同测试样本上的变化检测可视化结果。可以看到随着标记训练样本的增多，模型的检测效果越好，尤其是在一些细节处的表现，并且我们的半监督变化检测算法仅在 40% 甚至更少的标记需求下就取得了与 Oracle 相当的性能。除给出的可视化示例之外，在其余数据集以及其余样本上，我们的方法也表现出了与定性实验结果一致的优异性能。总而言之，自适应阈值使得我们的方法能够更加充分和平衡地学习到无标注样本的特征分布，低置信度学习进一步减小了错误伪标签带来的影响，因此在一些复杂区域、细节区域以及稀少样本类型上，我们的 AdTSemiCD

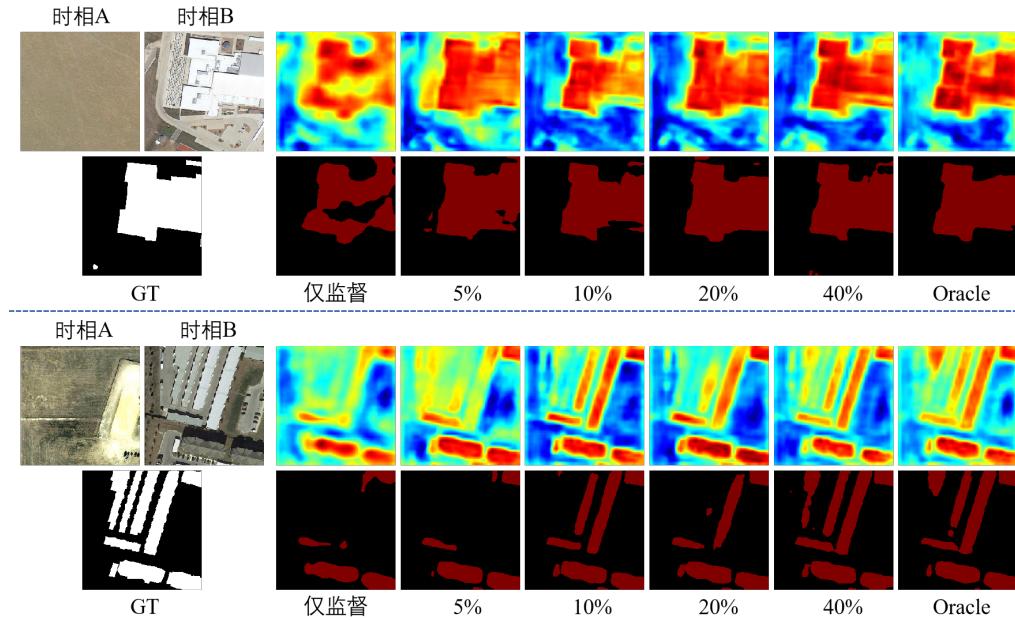


图 3-7 不同标记率下 AdTSemiCD 在 LEVIR-CD 测试集上的可视化对比图

展现了其强大的判别能力。

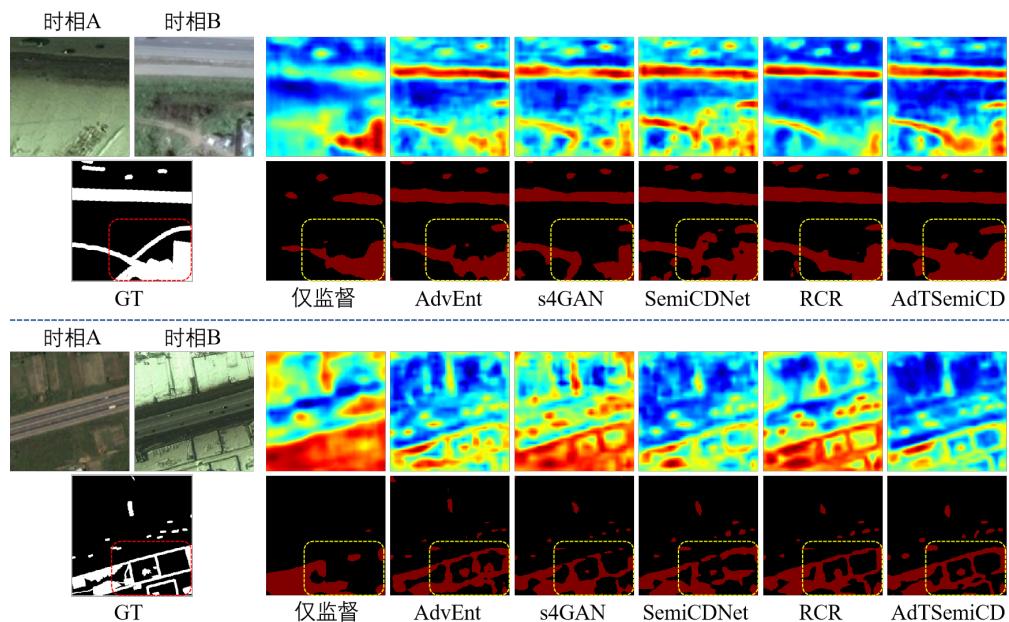


图 3-8 5% 标记训练样本下 AdTSemiCD 与 SOTA 方法在 CDD 测试集上的可视化对比图

3.3.2 消融实验

表 3-3 在 LEVIR-CD 数据集上的自适应动态阈值消融研究

Method	5%		10%		20%		40%	
	IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
仅监督	61.0	97.60	66.8	98.13	72.3	98.44	74.9	98.60
MT+NT	61.8 (+1.8)	98.66 (+0.06)	67.9 (+1.1)	98.30 (+0.17)	74.6 (+2.3)	98.59 (+0.15)	76.4 (+1.5)	98.70 (+0.10)
MT+FT	67.1 (+6.1)	98.14 (+0.54)	75.0 (+8.2)	98.63 (+0.50)	76.6 (+4.3)	98.71 (+0.27)	77.0 (+2.1)	98.73 (+0.13)
MT+AT	69.9 (+8.9)	98.38 (+0.78)	75.3 (+8.5)	98.67 (+0.54)	76.9 (+4.6)	98.74 (+0.30)	77.8 (+2.9)	98.79 (+0.19)

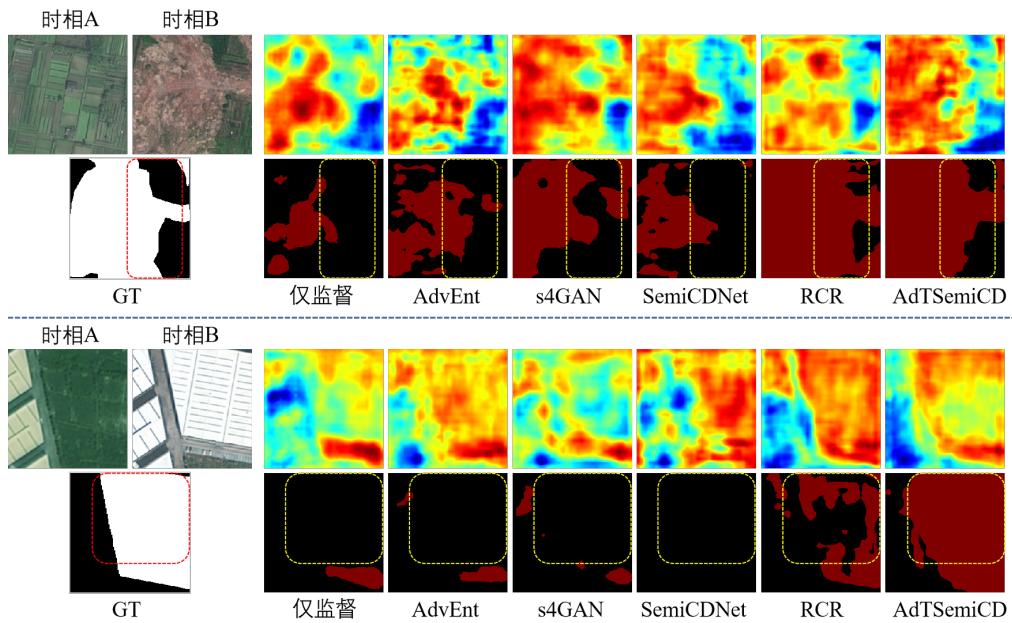


图 3-9 5% 标记训练样本下 AdTSemiCD 与 SOTA 方法在 CL-CD 测试集上的可视化对比图

为了验证本章所提出的自适应动态阈值和低置信度学习的有效性，我们在 LEVIR-CD 数据集所有标记率下进行了消融实验，实验结果如表3-3和表3-4所示。其中，表中的缩写及其含义解释如下：

MT: Mean-Teacher, 平均教师半监督学习框架。

NT: No Threshold, 无阈值筛选，即将所有标签均视为可信标签。

FT: Fixed Threshold, 固定阈值筛选，固定阈值设置为 0.95。

AT: Adaptive Threshold, 自适应阈值筛选。

LL: Low confidence Learning, 低置信度学习。

在关于自适应动态阈值筛选的消融实验中，当对教师模型生成的伪标签不进行阈值筛选而直接指导学生模型训练时，相比基线，大量的无标记样本带来的提升却是微乎其微，这也说明了阈值筛选的重要性。使用 0.95 的固定高阈值能够带来巨大的改进，这是因为过滤掉了那些不可信的标签，仅使用高置信度的可靠伪标签指导模型的训练，这种自训练能够极大地扩充模型学习到的样本特征空间。最后，使用我们的基于模型学习状态的自适应动态阈值筛选机制，在固定阈值筛选的基础上进一步提升了模型的性能。得益于动态调整过程，模型在不同训练阶段使用了不同数量和质量的伪标签，不仅加快了模型的收敛速度，还提高了模型的泛化性。

表 3-4 AdTSemiCD 在 LEVIR-CD 数据集上的消融实验

Method	5%		10%		20%		40%	
	IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
仅监督	61.0	97.60	66.8	98.13	72.3	98.44	74.9	98.60
MT+FT	67.1 (+6.1)	98.14 (+0.54)	75.0 (+8.2)	98.63 (+0.50)	76.6 (+4.3)	98.71 (+0.27)	77.0 (+2.1)	98.73 (+0.13)
MT+AT	69.9 (+8.9)	98.38 (+0.78)	75.5 (+8.7)	98.64 (+0.50)	76.8 (+4.5)	98.72 (+0.28)	77.7 (+2.8)	98.75 (+0.15)
MT+FT+LL	68.3 (+7.3)	98.21 (+0.61)	75.7 (+8.9)	98.64 (+0.51)	77.0 (+4.7)	98.72 (+0.28)	77.3 (+2.4)	98.75 (+0.15)
AdTSemiCD	71.3 (+10.3)	98.40 (+0.80)	76.3 (+9.5)	98.66 (+0.53)	77.2 (+4.9)	98.75 (+0.31)	78.0 (+3.1)	98.79 (+0.19)

在表3-4的 AdTSemiCD 的整体消融实验中，我们一方面在经典的平均教师半监督学习框架基线上，逐步替换和增加我们所设计的自适应动态阈值和低置信度学习模块，其中每个部分都取得了预期的提升效果，见表中第三行、第五行与仅监督基线的指标提升对比。另一方面，我们通过在两种不同的阈值方案上添加我们的低置信度学习模块，独立地证明了这一模块设计的有效性，见表中第二行与第四行以及表中第三行与第五行的指标提升对比，低置信模块分别在固定阈值筛选和自适应动态阈值筛选上取得了 0.65 和 0.73 个百分点的 IoU^c 提升，并且值得注意的是，低置信度学习在标记率越少的情况下效果越好，很显然这是因为标记数据越少，模型的预测更加不可信，因此那些被筛选掉的伪标签包含真实语义的概率也就越大，通过低置信度学习这种软伪标签可以将过滤掉的语义信息加以利用。

表 3-5 损失权重消融实验结果

λ_u	λ_w	LEVIR-CD		WHU-CD		CDD		CL-CD	
		IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
0	0 (Sup.Only)	61.0	97.60	50.0	97.48	60.4	94.25	18.1	91.90
0.5	0	68.1	98.20	62.9	98.17	64.3	94.76	23.5	92.08
1.0	0	69.9	98.38	65.0	98.32	65.7	94.96	25.9	92.14
2.0	0	67.2	98.15	63.8	98.24	64.6	94.87	25.8	92.15
1.0	0.5	70.6	98.35	66.1	98.41	67.9	95.37	27.6	91.90
1.0	1.0	71.3	98.40	66.0	98.41	68.3	95.52	26.5	92.27
1.0	2.0	71.0	98.39	65.2	98.34	67.7	95.42	27.6	92.32

自适应动态阈值伪标签筛选的设计初衷在于提高无标记样本的利用率，因此我们进行了消融实验，在同样的半监督学习框架之上，以及固定阈值和自适应动态阈值两种方案下对无标记样本的利用率进行了对比，如图3-10所示，我们的自适应动态阈值筛选能够更大程度地挖掘无标记样本富含的信息，利用率从训练初期到训练完成均高于固定阈值筛选方法。此外还可以观察到的一点就是，在这种筛选机制下，能够更好地遏制由于少量标记样本造成的过拟合现象，在模型中后期仍然能够在无标记样本上生成可靠的伪标签。

此外，在总体损失公式3-12中，无监督损失 \mathcal{L}_u 和低置信度学习损失 \mathcal{L}_{KL} 的权重分别为超参数 λ_u 和 λ_k ，我们对其取值进行了消融实验。表3-5展示了我们在四个数据集上以 5% 标记率进行训练时的消融实验结果，其中在 LEVIR-CD 和 CDD 数据集上，当 λ_u 和 λ_k 分别取值为 1.0、1.0 时取得了最优结果；而在 WHU-CD 上，当 λ_u 和 λ_k 分别取值为 1.0、0.5 时取得了最优结果；以及在 CL-CD 上，当 λ_u 和 λ_k 分别取值为 1.0、2.0 时取得了最优结果。并且 λ_k 在 CDD、CL-CD 数据集上选择更加敏感。但整体都围绕 (1.0, 1.0) 的参数组合上下略微波动，因此我们在其余数据集上均选择这两个数值作为两部分损失的权重。

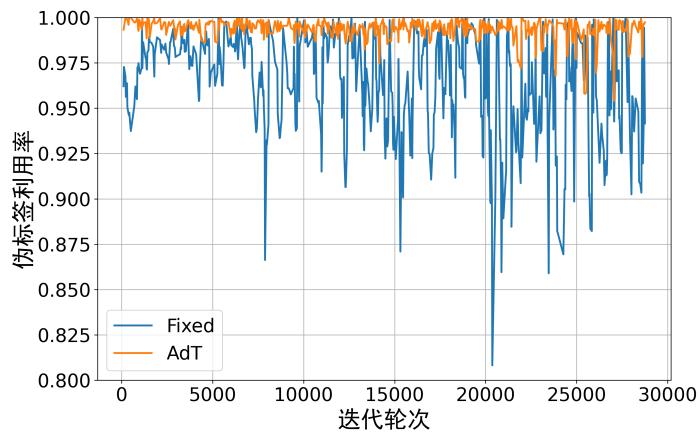


图 3-10 训练过程中的伪标签利用率

3.4 本章小结

本章提出了基于自适应动态阈值的半监督变化检测算法，首先在引言部分分析了此前的阈值方法存在的局限性，针对模型对无标记样本学习不够充分的问题，提出了设计自适应阈值的必要性。之后详细介绍了本章的方法，从整体框架入手，再分别介绍了自适应动态阈值调整方法和低置信度学习模块。最后在所有十个数据集上进行了实验验证，从定性和定量两个角度验证了本章方法与其他几种对比半监督方法的性能，以及进行了消融实验证明本章方法中每个部分设计的合理性和有效性。

第 4 章 基于伪标签评估的自适应半监督变化检测算法

4.1 引言

变化检测是遥感的一个重要领域，其主要重点是确定卫星在同一区域以不同间隔拍摄的双时相图像对中的变化区域。考虑到为变化检测任务准确标注掩码的过程是劳动密集型的，半监督变化检测是更有前景的方法。半监督变化检测的范式通常是通过利用有限的可用标签和大量未标记的样本来提高变化检测性能。对未标记的数据生成伪标签，这些伪标签通常是具有较高预测概率的临时预测掩码，将其作为训练过程中的指导信号，通过混合训练具有实际标签的有限数据和具有伪标签的丰富数据，学生模型可以学习到更多重要的特征，从而显著提高性能。最流行的方法是平均教师^[4] 框架，它使用一个教师模型来生成伪标签，在训练过程中为学生模型提供指导。随后使用学生模型的指数移动平均（EMA）更新教师模型。

虽然这些方法已经获得了一定的成功，但仍然存在重大问题：即模型以同样的方式处理所有样本，而不考虑不同样本之间必然存在的差异性，并且训练过程缺乏灵活性。首先，很明显，未标记的样本可能并不总是能够胜任“教师”角色。模型在为复杂的样本生成可靠的高质量伪标签时经常遇到困难，这反过来又引入了额外的噪声，可能会误导模型的训练。

此外，EMA 更新过程也没有考虑这种噪声干扰。考虑到训练批次中可能存在偏差或包含噪声，动态确定训练更新有助于训练过程的稳定性。这些因素强调需要更精确的监督方法，否则可能会对模型的训练产生负面影响。在本研究中，我们引入了一种自适应动态学习策略 AdaSemiCD，旨在提高伪标签的准确性并简化训练过程。我们的框架结合了传统的半监督训练方法，并辅以两个创新的功能模块 AdaFusion 和 AdaEMA。首先，我们利用 AdaFusion 在单个样本水平上抑制噪声，从而提高伪标签的准确性。与之前依赖于完全随机融合区域的 Augseg^[48] 或 CutMix^[114] 等方法相反，我们的 AdaFusion 模块主动识别可能最不确定的区域，并将其替换为来自高质量标记数据集或未标记数据集的可靠内容。在此之后，我们通过 AdaEMA 模块动态调整师生模型的参数更新选择批次，以确保提高稳定性。虽然传统的 EMA 有效地减轻了模型参数的波动从而提高了稳定性，但它在每次训练迭代后都统一的更新参数，在处理一系列训练样本时忽略了模型在不同迭代中的不同学习效果。如果未标记的样本包含大量错误信息，它可能会误导模型的训练。因此，我们的 AdaEMA 引入了模型级参数更新的自适应选择过程，使模型能够充分集成优越的参数。

本章阐述了基于伪标签评估的自适应半监督变化检测算法——AdaSemiCD，首先介绍了模型的整体框架以及训练流程，随后详细介绍了设计的伪标签评估指标，以及基于此指标设计的两个自适应模块。最后在实验部分报告了该研究方法在十个公开数据集上

取得的实验结果，并与其他经典半监督变化检测算法在定性和定量两个维度进行了公平的对比，证明了我们的方法优越性。最后通过消融实验证明了每个模块的有效性。

4.2 基于伪标签评估的自适应半监督变化检测框架

4.2.1 整体框架

AdaSemiCD的整体框架如图4-1所示，与第三章 AdTSemiCD类似，也才应平均教师模型作为主体框架。不同之处在于，AdaSemiCD基于教师模型对无标记样本的输出结果，在无标记样本对上进行了进一步的样本融合操作，以及在教师模型权重更新时采用了AdaEMA方式。

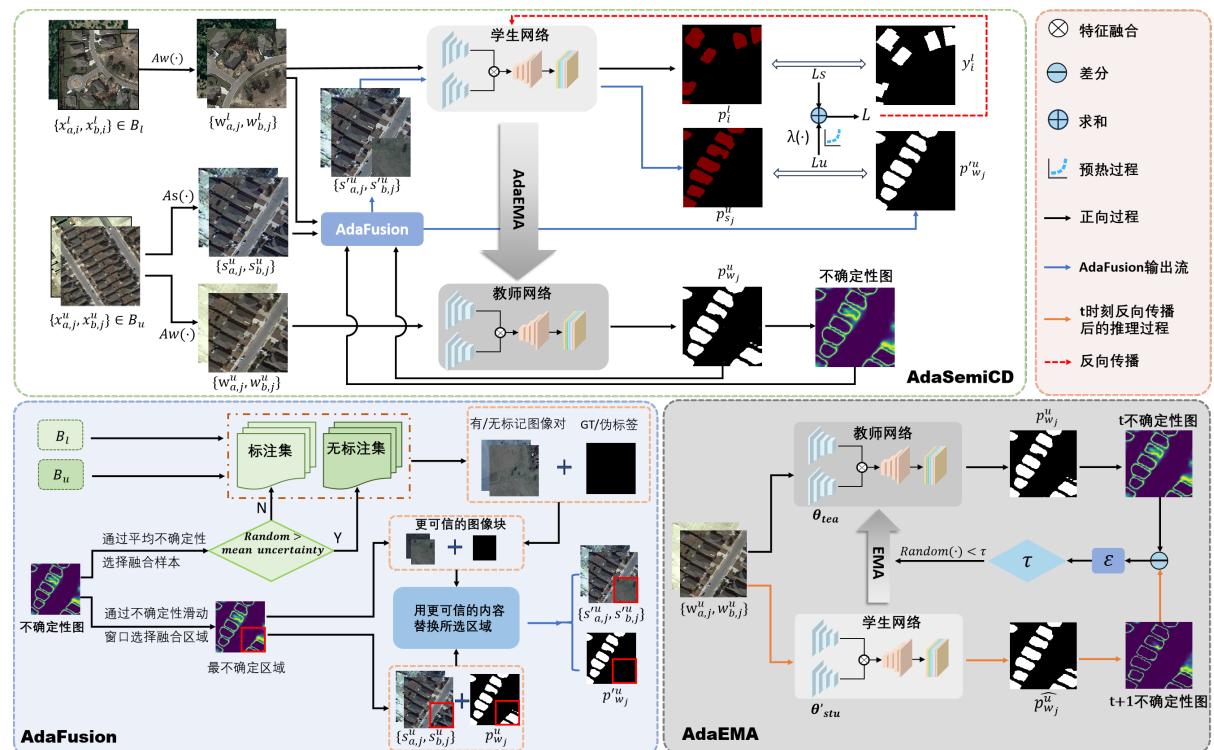


图 4-1 AdasemiCD 的整体框架图

此外，在初始训练阶段，我们的模型对未标记的样本生成的伪标签是高度不可靠的；在这个阶段过度依赖无监督训练可能会引入严重的噪声。相反，随着训练进入中后期阶段，模型越来越多地从有限的标记数据中学习，伪标签的质量得到提高，这时减少监督训练相对于非监督训练的比例是合理的。为了减轻过拟合现象和增强特征空间，必须实现一个受控过程来动态调整这两个分量在损失函数中的贡献。因此，在 AdaSemiCD 中我们使用一个预热过程来控制未标记部分的训练速度。 $\lambda(\cdot)$ 为无监督损失随训练步长变化的权值，用于动态调整不同阶段有监督和无监督训练的比例，表达式如4-1所示。

$$\lambda(\cdot) = w_{\max} \times e^{-\phi \times (1 - \text{iter cur} / \text{iter max})^2}, \quad (4-1)$$

$$\text{iter}_{\max} = \gamma \times \text{iter}_{\sum} \quad (4-2)$$

其中 w_{\max} 和 ϕ 是超参数， w_{\max} 表示无监督损失的最大可达权重， ϕ 用于控制上升的剧烈程度； iter_{cur} 表示当前迭代周期； iter_{\max} 为预热过程的总步长，由4-2计算得到， iter_{sum} 为整个训练过程的总迭代次数，且 $0 < \gamma < 1.0$ ；在上升过程之后，无监督损失的权重保持为 w_{\max} 不变。在训练的早期，该权值相对较低，无监督训练的作用可以忽略不计，但在训练的中后期，该权值逐渐增大，加权后的无监督损失超过了有监督损失，因此无监督训练占主导。

自适应模块：半监督学习的本质在于伪标签的质量。然而，很明显，前面提到的过程没有考虑到不同样本个体对模型训练的不同影响。本章研究主要关注与伪标签生成直接相关的两个元素：未标记图像对，以及伪标签生成网络（即教师模型）在识别变化方面的性能。为了给未标记信息提供更可靠的监督指导，减少训练的不确定性，我们设计了一种自适应训练策略来解决这两个关键问题。我们首先设计了一个度量标准，用于量化伪标签的不确定性，作为自适应调整的基础。接下来，我们提出图像层面对未标记的训练样本进行自适应改造。此外，还提出在训练阶段对教师网络应用自适应选择性的EMA更新，以集成更加正确的参数，从而产生更一致和更高质量的伪标签。这些努力将在下面的小节中进行详细阐述。

4.2.2 伪标签评估指标设计

为了通过准确测量伪标签的质量来提高其有效性，关键的一步就是评估指标的设计。这个度量将有助于识别可靠的标签和确定每个训练样本的效力。与带有标记图像对的场景不同，可以将真实标签作为基准，通过预测与真是标签的平均交并比、F1分数等来评估模型的推理质量，而伪标签缺乏任何这样的参考，其只能与自己进行比较。

因此，我们专门为伪标签设计了一个可量化的评估指标，旨在总体信息熵的基础之上，同时考虑到类别不平衡和混淆区域等因素。信息熵是一种衡量模型性能的常用指标，计算公式如下：

$$E(x_i) = -P(x_i) \log_2 P(x_i) \quad (4-3)$$

其中 $P(x_i)$ 为模型对于样本 x_i 的输出概率。我们认为，预测值中的信息熵越低，预测结果的可信度越高。相反，信息熵越高，预测结果的变化越大，同一像素上的预测概率分布越均匀，类别之间的差异越小。

在变化检测任务中，由于类别不平衡带来的巨大挑战，直接应用信息熵通常不会产生良好的结果。如图3-3所示，我们在 10 个公开的变化检测数据集上进行了统计分析，变化和不变类别的比例是极度失衡的。这种类别不平衡现象会导致模型在训练过程中学习目标类别不够充分，而过度地拟合不变类别的特征分布。因此，在推理过程中，模型更倾向于将像素分类为背景类。因此，尽管该模型的总体精度（OA）看起来很高，但在目标类别的交并比（ IoU^c ）上的性能仍然令人不满意。为了在评估伪标签质量时最大限度地减少这种类别不平衡的影响，我们在计算信息熵时为两个类别分配了不同的权重，

如式(4-4)所示。

$$E'(x_i) = w_1 \times E(x_i)[0] + w_0 \times E(x_i)[1] \quad (4-4)$$

其中 w_0 和 w_1 分别表示当前小批次中属于不变类和变化类别的像素的比例。其计算公式如式(4-5)和式(4-6)所示。

$$w_0 = \frac{\sum_{i=1}^{|Bu|} \sum_{k=1}^{H \times W} (P(x_i) == 0)}{|Bu| \times H \times W}, \quad (4-5)$$

$$w_1 = \frac{\sum_{i=0}^{|Bu|} \sum_{k=0}^{H \times W} (P(x_i) == 0)}{|Bu| \times H \times W} \quad (4-6)$$

此外，来自边界、目标与背景相似的区域的像素，模型的预测通常具有很大的不确定性，应该得到更多的重视。首先，我们通过计算两类预测概率之差的绝对值来增强这些易混区域的影响：

$$D(x_i) = \text{abs}(P(x_i)[1] - P(x_i)[0]) \quad (4-7)$$

这里， abs 表示取绝对值运算，这是一种用来防止前后时相变化检测不幂问题的措施。然后利用信息熵进行逐像素相乘运算，最终得到图像的不确定性映射图 $U(x_i)$ ：

$$U(x_i) = 1 - D(x_i) \cdot E'(x_i) \quad (4-8)$$

显然，在低信息熵的像素位置，该过程不会导致显著的变化，而在高信息熵的像素位置，其值将显著减小，直接使用该指标与所表达的意义相反。因此，我们采用校正信息熵的逆作为其描述。

综上所述，为了评估变化检测伪标签的质量，我们提出了一个可量化的计算度量 U 来衡量预测不确定性，该度量考虑了类别不平衡等因素并增加了对混淆区域的关注，从而在总体信息熵的基础上表达了更多有价值的信息。

4.2.3 自适应样本融合机制

图像融合经常被用来增强或者改造训练样本以提高样本多样性，CutMix^[114] 和 MixUp^[115] 是典型的融合方法。在本研究中，我们的目标是利用图像融合技术来排除训练样本中的不可靠区域。这个过程包括两个步骤：1) 融合区域选择；2) 融合图像选择。

融合区域的自适应选择。与传统的随机选择混合区域的 CutMix 技术不同，我们的方法更加精细。首先我们将初始化一个随机大小的边界框，然后沿着图像坐标滑动窗口，维护一个总体最大不确定性框，最终选择总体不确定性最高的区域作为融合区域。

这些区域通常是边界或复杂的区域，模型在这些区域的检测性能受限，很难准确识别变化目标，因此如果不做处理地直接用于训练，就会引入大量的噪声干扰，这对模型训练有很大的影响。

融合内容的自适应选择。对于最大不确定性的区域，我们可以选择从样本集中选择一个其余更可信的样本替代这部分，可选样本来自标记数据集 B_l 或 y 有着更高可靠性

的未标记数据集 B_u 。这种策略可以防止过度使用仅有的少量标记样本，进一步降低过拟合的风险。这里融合内容使用图像块而不使用实例对象，这是因为我们认为过多的人为干预可能会破坏模型的泛化能力，并且仅替换实例对象可能造成周围像素的失真。

融合内容的选择主要通过自适应调整决策上界来确定融合内容。我们直接使用同一个小批次内计算的不确定性均值作为决策上界，对每个小批次随机生成一个概率值（0 到 1 之间），如果超过不确定性均值，代表当前样本包含的噪声过多，则随机从标记图像批次中选择一对图像作为融合内容。我们会维护根据不确定性排序维护一个批次不确定队列集合，如果小于此上界，则从此队列集合中均匀选择 TopK 的未标记的图像对。很容易理解，无标注样本的伪标签质量越高。可以被认为足够可靠。而不确定性越高的样本噪声越大，融合标记样本可以显著降低噪声密度，融合过程如图??左下角的 AdaFusion 子图。

算法 4.1 AdaEMA 算法

Input:

- 1: 学生模型 M_{stu}^θ , 教师模型 M_{tea}^θ
- 2: 当前批次的训练样本集 $\mathcal{B} = \{\mathcal{B}_l, \mathcal{B}_u\}$

Output:

- 3: 更新后的教师模型 $M_{tea}^{\theta'}$
 - 4:
 - 5: 在标记样本 \mathcal{B}_l 上根据??计算监督损失 \mathcal{L}_s ;
 - 6: 在无标记样本 \mathcal{B}_u 上根据??计算无监督损失 \mathcal{L}_u ;
 - 7: 根据 SGD 最小化总训练损失来更新学生网络参数 M_{stu}^θ 为 $M_{stu}^{\theta'}$;
 - 8: 根据4-8计算教师模型在无标记样本上生成的伪标签的不确定性 U_{tea} ;
 - 9: 根据4-8计算更新后的学生模型在无标记样本上的预测结果的不确定性 U_{stu} ;
 - 10: 根据公式4-9和4-10计算教师模型此次迭代更新概率的上界 τ ;
 - 11: **if** $random(\cdot) < \tau$ **then**
 - 12: 根据指数移动平均公式3-5更新教师模型参数为 $M_{tea}^{\theta'}$;
 - 13: **else**
 - 14: $M_{tea}^{\theta'} = M_{tea}$;
 - 15: **end if**
 - 16: **Return:** $M_{tea}^{\theta'}$;
-

4.2.4 自适应师生模型参数更新机制

在平均教师框架中，通过对学生模型参数在时间序列上的移动指数平均得到教师模型的参数，这种方法与共享权重的孪生网络相比，更加稳定可靠。然而，我们对平均教师模型的期望是达到共同进化的最优状态，在这种状态下，教师模型是不断进化的学生模型的累积表示。实现这种共同进化的关键因素是每次教师模型更新时，学生模型是处于更优状态还是处于波动状态。那么，我们如何评估学生模型的状态呢？这个问题本质上与训练过程的验证阶段类似。通常，在训练几个 epoch 后，我们评估当前模型，将验证集中的样本输入到模型中进行推理，并将结果与实际标签进行比较以确定模型的表现。然而，如果我们在每次迭代之后都进行这样的验证过程，那样是非常耗费计算资源

的，导致训练时间的显著增加，这是一种不现实的方案。降低验证集中的样本数量（仅为几对）可能会解决可行性问题，但样本量太小无法全面准确地评估模型的性能，这会带来新的挑战。那么，有没有一种方法可以将这两个概念融合在一起呢？

在每个训练阶段，我们首先根据第3.3.1小节中描述的训练策略更新学生模型的参数 M_{stu}^θ ，从而得到 $M_{stu}^{\theta'}$ 。接下来，我们在当前的未标记训练样本批次 B_u 上评估更新后的学生模型 $M_{stu}^{\theta'}$ 和教师模型 M_{tea}^θ 。我们根据4-8计算不确定性图，并将其分别表示为 U_{tea} 和 U_{stu} 。进一步根据4-9得到不确定性变化值 ε 。

$$\varepsilon = \frac{\sum U_{stu} - \sum U_{tea}}{|B_u|} \quad (4-9)$$

如果 $\varepsilon < 0$ ，则代表模型性能发生退化或处于振荡。相反，当 $\varepsilon \geq 0$ ，模型得到正确的训练。最终，只有进化的学生模型被选择参与到教师模型参数的更新中，为了加入一定的动态性，我们据此得到一个更新概率上界，如式(4-10)所示，再根据此上界决策此迭代轮次是否更新。

$$\tau = \begin{cases} \frac{1}{\text{iter}^2 + \epsilon} & , \quad \varepsilon \leq 0 \\ 1.0 & , \quad \varepsilon > 0 \end{cases} \quad (4-10)$$

这里， $\epsilon = 1e - 5$ 用来防止除数为零， iter 表示当前的迭代计数。如果模型不确定性变化值 $\varepsilon \leq 0$ ，也不会完全丢弃它们。而是赋予一个较低的概率上界去选择更新。有了这个上限规则（范围从0到1），引入了一些随机性来确定教师网络参数更新，进一步的细节可以在算法4.1中找到。

4.3 实验结果及分析

4.3.1 实验设置

本章所有实验设置均与第三章相同，因此部分实验数据沿用了第三章取得实验结果。

4.3.2 对比试验

我们与几种最先进的半监督变化检测方法进行了比较，其中SemiCDNet^[5]、RCR^[6]和FPA^[7]是过去几年出现的获得最优性能的半监督变化检测方法。此外，我们还比较了两种半监督语义分割方法，s4GAN^[113]和AdvEnt^[112]。

(1) 定量对比

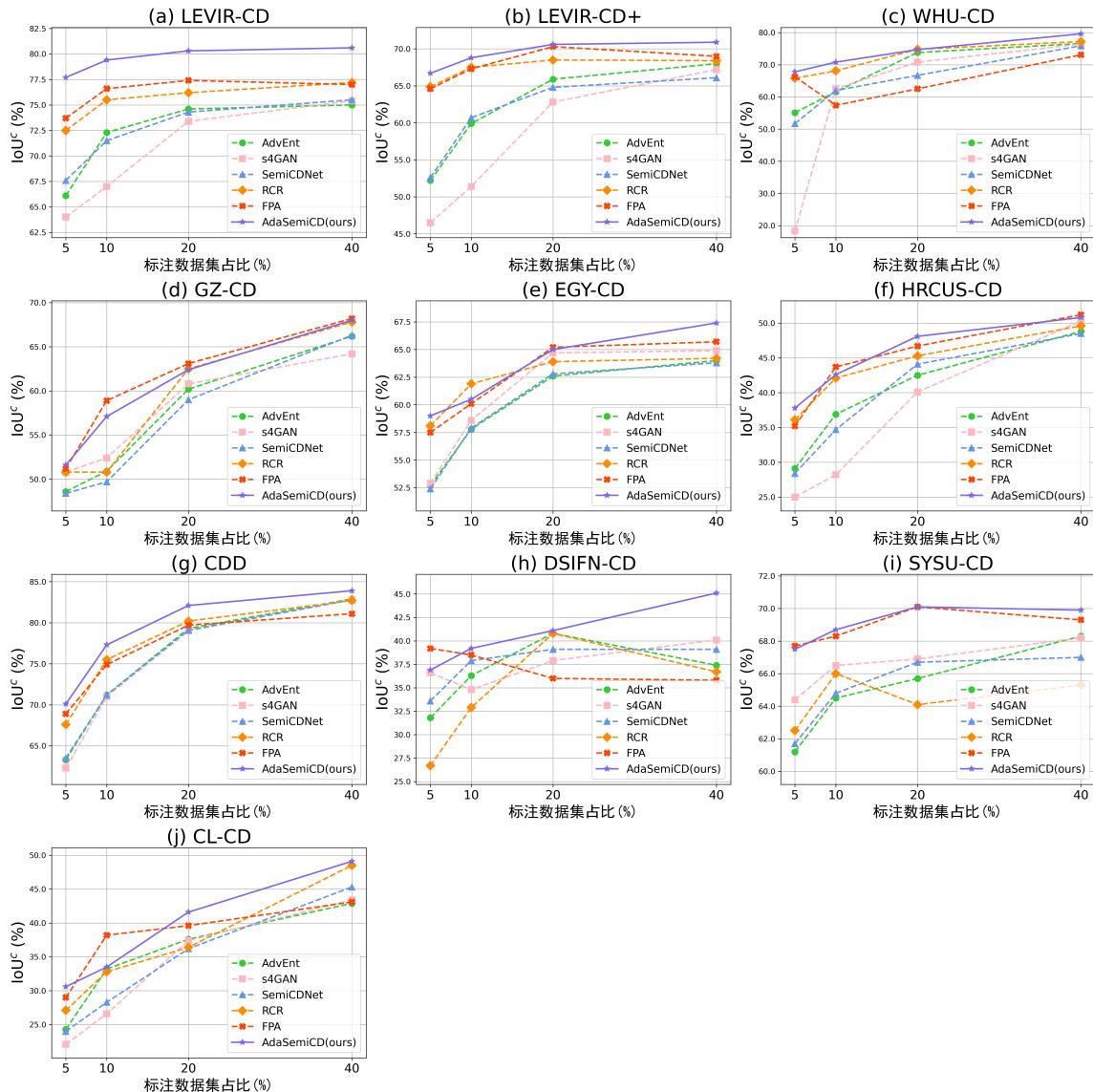


图 4-2 AdaSemiCD 与 SOTA 方法在 10 个数据集上的性能对比折线图

我们在所有 10 个数据集上进行了实验，表4-1和表4-2分别报告了我们在建筑物变化检测数据集和多类变化检测数据集上的实验结果。其中，表中的 Sup.Only 是指仅在有限的标记数据集上进行监督训练，而 Oracle 是在整个全部训练数据集上进行全监督训练的结果。另外图4-2更直观地展示了我们的所有实验定量结果。值得注意的是，我们的方法在所有数据集的几乎所有半监督划分设置中都实现了最先进的性能。可以肯定的是，在大多数情况下，所有的半监督变化检测方法在相应的划分设置下都比监督方法表现得更好，这证明了半监督方法在从大量未标记的训练样本中学习到了更多的知识，同时，我们的方法取得的领先证明了我们对于如何高效、正确地从未标记样本集中学习的自适应机制是具有实际意义的。

建筑物变化检测: 如表4-1所示，其中蓝色高亮部分代表着最优结果，下划线标注

表4-1 不同CD方法在二值建筑物变化检测数据集上的平均定量性能对比

Dataset	Method	5%		10%		20%		40%	
		IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
LEVIR-CD	Sup. only	61.0	97.60	66.8	98.13	72.3	98.44	74.9	98.60
	AdvEnt ^[112]	66.1	98.08	72.3	98.45	74.6	98.58	75.0	98.60
	s4GAN ^[113]	64.0	97.89	67.0	98.11	73.4	98.51	75.4	98.62
	SemiCDNet ^[5]	67.6	98.17	71.5	98.42	74.3	98.58	75.5	98.63
	RCR ^[6]	72.5	98.47	75.5	98.63	76.2	98.68	77.2	98.72
	FPA ^[7]	73.7	98.57	76.6	98.72	77.4	98.75	77.0	98.74
	AdaSemiCD	77.7	98.78	79.4	98.87	80.3	98.92	80.6	98.93
Oracle		$IoU^c=77.9$ and OA=98.77							
LEVIR-CD+	Sup. only	52.0	97.72	58.4	98.06	66.1	98.31	66.2	98.42
	AdvEnt ^[112]	52.2	97.68	59.9	98.11	65.9	98.37	68.0	98.51
	s4GAN ^[113]	46.5	97.25	51.4	97.66	62.8	98.18	67.2	98.46
	SemiCDNet ^[5]	52.6	97.66	60.7	98.24	64.8	98.37	66.1	98.38
	RCR ^[6]	64.9	98.25	67.5	98.45	68.5	98.52	68.4	98.51
	FPA ^[7]	64.6	98.30	67.3	98.40	70.3	98.64	69.0	98.59
	AdaSemiCD	66.7	98.49	68.8	98.51	70.6	98.63	70.9	98.64
Oracle		$IoU^c=70.5$ and OA=98.63							
WHU-CD	Sup. only	50.0	97.48	55.7	97.53	65.4	98.20	76.1	98.94
	AdvEnt ^[112]	55.1	97.90	61.6	98.11	73.8	98.80	76.6	98.94
	s4GAN ^[113]	18.3	96.69	62.6	98.15	70.8	98.60	76.4	98.96
	SemiCDNet ^[5]	51.7	97.71	62.0	98.16	66.7	98.28	75.9	98.93
	RCR ^[6]	65.8	98.37	68.1	98.47	74.8	98.84	77.2	98.96
	FPA ^[7]	66.3	98.45	57.4	97.69	62.5	98.48	73.1	98.69
	AdaSemiCD	67.8	98.62	70.8	98.70	74.7	98.86	79.6	99.13
Oracle		$IoU^c=85.5$ and OA=99.38							
GZ-CD	Sup. only	47.5	93.56	51.4	94.26	58.0	95.65	66.3	96.62
	AdvEnt ^[112]	48.6	94.39	50.9	94.89	60.2	95.79	66.2	96.58
	s4GAN ^[113]	50.8	94.38	52.4	94.98	60.8	95.94	64.2	96.39
	SemiCDNet ^[5]	48.4	93.58	49.7	94.79	59.0	95.66	66.3	96.57
	RCR ^[6]	50.8	93.82	50.8	94.69	62.5	96.07	67.8	96.61
	FPA ^[7]	51.2	93.92	58.9	95.78	63.1	96.26	68.2	96.82
	AdaSemiCD	51.6	94.56	57.1	95.57	62.4	96.21	68.0	96.75
Oracle		$IoU^c=69.0$ and OA=96.93							
EGY-CD	Sup. only	49.8	95.73	54.6	96.38	61.4	96.83	65.1	97.25
	AdvEnt ^[112]	52.7	96.01	57.8	96.58	62.6	96.86	64.0	97.19
	s4GAN ^[113]	52.9	95.94	58.6	96.50	64.7	97.09	64.9	97.27
	SemiCDNet ^[5]	52.4	96.00	57.9	96.31	62.8	96.95	63.8	97.19
	RCR ^[6]	58.1	96.50	61.9	96.77	63.9	97.08	64.2	97.18
	FPA ^[7]	57.5	96.52	60.1	96.86	65.2	97.25	65.7	97.34
	AdaSemiCD	59.0	96.55	60.5	96.80	65.0	97.20	67.4	97.39
Oracle		$IoU^c=67.6$ and OA=97.54							
HRCUS-CD	Sup. only	29.5	98.11	36.0	98.45	43.4	98.68	48.9	98.84
	AdvEnt ^[112]	29.1	98.11	36.9	98.40	42.5	98.61	48.8	98.71
	s4GAN ^[113]	25.0	97.86	28.2	98.24	40.1	98.62	50.3	98.85
	SemiCDNet ^[5]	28.4	98.00	34.7	98.44	44.1	98.68	48.5	98.74
	RCR ^[6]	36.1	98.36	42.1	98.69	45.3	98.76	49.6	98.66
	FPA ^[7]	35.2	98.37	43.7	98.65	46.7	98.82	51.2	98.81
	AdaSemiCD	37.8	98.59	42.6	98.70	48.1	98.84	50.8	98.87
Oracle		$IoU^c=59.0$ and OA=99.06							

的是次优结果。我们所提出的 AdaSemiCD 框架在几乎所有建筑物变化检测数据集上都表现出了出色的性能。值得注意的是，在 LEVIR-CD、LEVIR-CD+ 和 WHU-CD 数据集上，AdaSemiCD 在 IoU^c 上实现了全面的、显著的改进，分别提高了 3.1、1.3 和 1.7 个百分点。在 EGY-CD 和 HRCUS-CD 数据集上，AdaSemiCD 总体上相较于之前的最佳方法有所提升，分别平均提高了 0.75 和 0.4 个百分点。然而，在某些实验设置下，其性能略低于之前表现最好的方法。这可以归因于这些数据集带来的独特挑战：EGY-CD 数据集存在大量的过度曝光问题，这经常导致建筑物与背景混合在一起，增加了准确区分的难度。同样，HRCUS-CD 数据集也受到了某些建筑物变化区域存在的植被遮挡的影响。这些挑战阻碍了模型在用有限的真实标签进行训练时生成可靠的伪标签的能力。

显然，AdaSemiCD 在 GZ-CD 数据集上遇到了严峻的挑战，除了在 5% 标记率的半监督设置下保持领先之外，其余大多数实验设置下的性能表现都逊于 FPA。我们将其归因于该数据集的固有特性和我们方法的固有局限性。首先，在所有进行实验的数据集中，GZ-CD 数据集的空间分辨率最低。这种限制，再加上小建筑物的普遍存在，导致在这种分辨率下很难准确辨别出变化目标，即使是对人类观察者用肉眼来检测也是如此。其次，该数据集的人工注释非常粗糙，当然这也与我们前边的推测相同。如图2-10所示，对于具有多个目标的区域，Ground Truth 中通常是直接将一整块直接标注为变化区域，而这其中包含大部分背景区域。这会在标记的数据中引入大量噪声，特别是对于建筑物变化检测任务。最后，数据集采集的背景也带来了额外的挑战，该数据集采集于城市郊区，而在超过五年的时间跨度内，城市的快速发展和扩张导致了前后两幅时相的图像之间存在巨大差异，除了带注释的建筑变化区域之外。背景的变化也非常显著，例如整片区域的植被变为裸土，以至于很难确定两幅图像之间的关系，这引入了大量非目标类别变化，严重干扰了模型关注建筑物变化区域的能力。然而，与 AdaSemiCD 不同，FPA 和 RCR 并不完全依赖单一硬标签。相反，它们利用特征一致性约束指导模型训练，这很大程度上有效地缓解了这个问题。

此外，从实验结果来看，由于背景类占多数，所有方法在这些二值建筑物变化检测数据集上的总体准确率（OA）都很高。OA 往往直接映射着平均交并比 IoU^c ，即便是小幅度的 OA 提升， IoU^c 就会有比较明显的改善。我们的 AdaSemiCD 在几乎所有的数据集上都取得了提高，因为它有效地降低了噪声，最大限度地减少了错误信号的影响，从而在像素级实现了更准确的预测。

多类变化检测：值得注意的是，在 CDD 数据集的所有实验设置上都取得了最优的结果， IoU_c 和 OA 分别平均增加了 1.8 个百分点和 0.25 个百分点，这可以归因于其异常准确的标注，如图 3 所示。这种高质量的标注对于基于伪标签的半监督方法特别有利。此外，AdaSemiCD 在 SYSU-CD 和 CL-CD 数据集上的整体性能有所提高。然而，在某些实验配置中，它的性能略低于 SOTA 结果。这可以归因于在识别这些数据集中存在的一些山脉和土地变化方面的固有挑战。这些变化通常是大规模的，这些地区的错误分类

可能会导致性能指标的波动。

表4-2 不同CD方法在多类变化检测数据集上的平均定量性能对比

Dataset	Method	5%		10%		20%		40%	
		IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
CDD-CD	Sup. only	60.4	94.25	67.9	95.46	75.6	96.59	82.3	97.56
	AdvEnt ^[112]	63.3	94.65	71.2	96.01	79.3	97.14	<u>82.9</u>	<u>97.66</u>
	s4GAN ^[113]	62.3	94.69	71.0	95.94	79.0	97.10	82.8	97.63
	SemiCDNet ^[5]	63.5	94.68	71.2	95.99	79.1	97.13	82.8	97.63
	RCR ^[6]	67.6	95.40	<u>75.5</u>	<u>96.57</u>	<u>80.2</u>	<u>97.26</u>	82.7	97.61
	FPA ^[7]	<u>68.9</u>	<u>95.66</u>	74.9	96.55	79.7	97.20	81.1	97.37
	AdaSemiCD	70.1	95.89	77.3	96.89	82.1	97.56	83.9	97.80
Oracle		$IoU^c=87.8$ and OA=98.10							
DSIFN-CD	Sup. only	34.8	78.34	<u>38.9</u>	83.41	40.2	87.00	39.6	87.00
	AdvEnt ^[112]	31.8	77.83	36.3	83.86	40.8	85.92	37.4	86.31
	s4GAN ^[113]	36.6	<u>84.10</u>	34.8	<u>86.87</u>	37.9	87.69	40.1	86.52
	SemiCDNet ^[5]	33.6	78.60	37.9	84.18	39.1	86.77	39.1	<u>87.05</u>
	RCR ^[6]	26.7	83.78	32.9	86.05	40.8	86.70	36.7	86.08
	FPA ^[7]	39.2	84.27	38.5	87.12	36.0	<u>87.41</u>	35.8	86.50
	AdaSemiCD	<u>36.9</u>	80.46	39.2	82.94	41.1	85.45	45.1	87.12
Oracle		$IoU^c=58.1$ and OA=90.82							
SYSU-CD	Sup. only	62.9	89.57	64.4	90.18	66.0	90.82	66.4	90.93
	AdvEnt ^[112]	61.2	89.36	<u>64.5</u>	90.18	65.7	90.35	68.3	91.24
	s4GAN ^[113]	64.4	90.02	66.5	90.48	66.9	90.26	68.2	91.51
	SemiCDNet ^[5]	61.7	89.32	64.8	90.25	66.7	90.97	67.0	91.08
	RCR ^[6]	62.5	89.76	66.0	90.75	64.1	90.22	65.3	90.56
	FPA ^[7]	67.7	90.95	<u>68.3</u>	<u>91.09</u>	<u>70.1</u>	<u>92.01</u>	<u>69.3</u>	91.97
	AdaSemiCD	<u>67.5</u>	91.16	68.7	91.59	70.1	92.03	69.9	91.90
Oracle		$IoU^c=68.2$ and OA=91.64							
CL-CD	Sup. only	18.1	91.90	31.4	92.42	37.2	93.32	45.9	<u>94.98</u>
	AdvEnt ^[112]	24.3	92.13	33.2	93.01	37.6	93.59	42.9	94.06
	s4GAN ^[113]	22.1	92.00	26.6	93.09	37.4	93.59	43.4	93.87
	SemiCDNet ^[5]	24.0	<u>92.20</u>	28.3	<u>93.42</u>	36.2	92.41	45.3	94.22
	RCR ^[6]	27.1	91.63	32.8	92.99	36.4	93.07	<u>48.5</u>	94.94
	FPA ^[7]	<u>29.0</u>	91.00	38.2	93.37	<u>39.6</u>	<u>93.88</u>	43.1	94.15
	AdaSemiCD	30.6	92.52	<u>33.5</u>	92.40	41.6	94.21	49.1	95.85
Oracle		$IoU^c=50.1$ and OA=95.55							

如表4-2所示，我们的AdaSemiCD同样在几乎所有多类数据集和所有划分设置上都实现了最佳性能，只是提升程度低于在简单的建筑物变化检测数据集上所取得的，这是由于多类别变化检测任务本身的复杂性，最明显的一点是，在某些情况下，之前的一些半监督的变化检测方法完全失效了，比监督基线的检测性能更差。然而，我们的半监督变化检测方法仍然可以利用额外的未标记样本来提高性能。值得注意的是，在CDD数据集的所有实验设置下AdaSemiCD都取得了最优的结果， IoU^c 和OA分别平均提高了1.8和0.25个百分点，这可以归因于其准确的手工标注，如图2-11所示。这种高质量的标注对基于伪标签自训练的半监督方法极为有利。此外，AdaSemiCD在SYSU-CD和CL-CD数据集上的整体性能有所提高。然而，在某些实验配置中，它的性能略低于最优

结果。这可以归因于在识别这些数据集中存在的一些山脉或土地变化方面的固有挑战。这些变化通常是大规模的，这些地区的错误分类可能会导致性能指标的波动。

另一点值得注意的是，在这些多类变化检测数据集中，所有方法的整体精度都降低了， OA 和 IoU^c 不再直接对应。其中一个现象是 IoU^c 达到了最高，而 OA 没有，这是由于其侧重点不同。 OA 更注重全局精度，在背景类占主导的数据集下就是优先背景类的检测，而 IoU^c 更关注目标类别的检测，在变化检测中我们更加关注于目标对象的检测，因此 IoU^c 更加重要。从 DSIFN-CD 数据集的实验结果中可以看到，我们的方法在一些设置下尽管在 OA 上表现不佳，但是在 IoU^c 上取得了最好的结果。

在所有此类数据集中，我们的方法在 DSIFN-CD 上表现出最差的性能。虽然它在关键指标 IoU^c 上表现良好，但在 OA 上的表现并不是最优的。这可以归因于 DSIFN-CD 和 GZ-CD 的相似之处，因为这两个数据集都具有人工标注粗糙和低空间分辨率这两个特点，如图2-11所示。此外，DSIFN-CD 类别不平衡现象基本消失，如图3-3所示，其变化类别的比例高达 34.81%。因此，我们设计的类别重平衡策略的增益有限，并且可能导致预测向前景倾斜，从而对背景像素预测错误，这或许就是 OA 表现不佳的直接原因。

4.3.3 定性对比

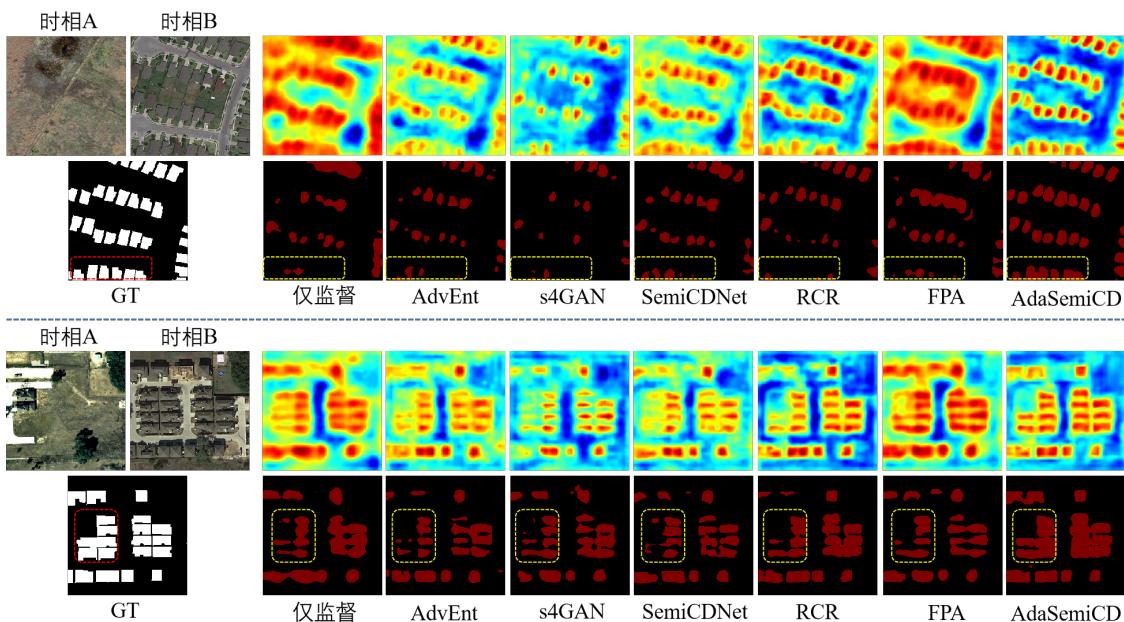


图 4-3 AdaSemiCD 与 SOTA 方法在 LEVIR-CD 测试集上的变化检测结果可视化对比

为了更清楚地说明我们提出的方法的变化检测性能提升，我们在所有十个不同的数据集的测试集上都进行了可视化，我们展示了部分可视化样例，其中图4-3、图4-4、图4-5以及图4-6分别展示了在 LEVIR-CD、WHU-CD、CDD、CL-CD 数据集的测试集上的预测结果可视化示例。图中用方框标注的区域是容易出错的区域。显然，在这些数据集上，我们的方法显著减少了遗漏和错误检测的常见问题。在具有挑战性的场景中，我们的方法仍然可以有效地识别出感兴趣的变化目标。

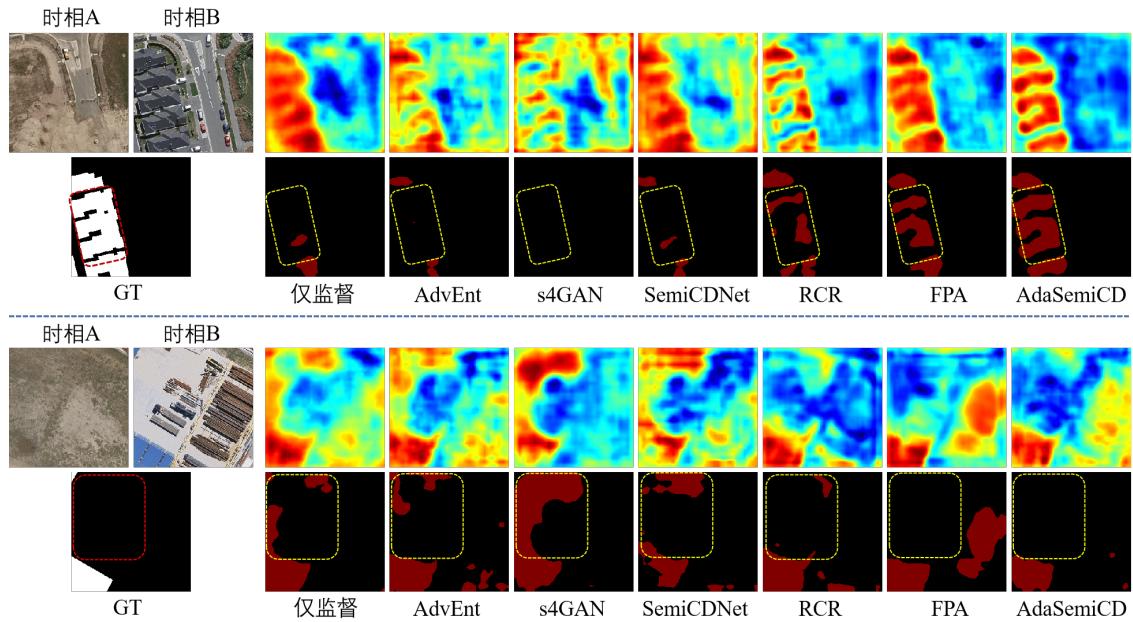


图 4-4 AdaSemiCD 与 SOTA 方法在 WHU-CD 测试集上的变化检测结果可视化对比

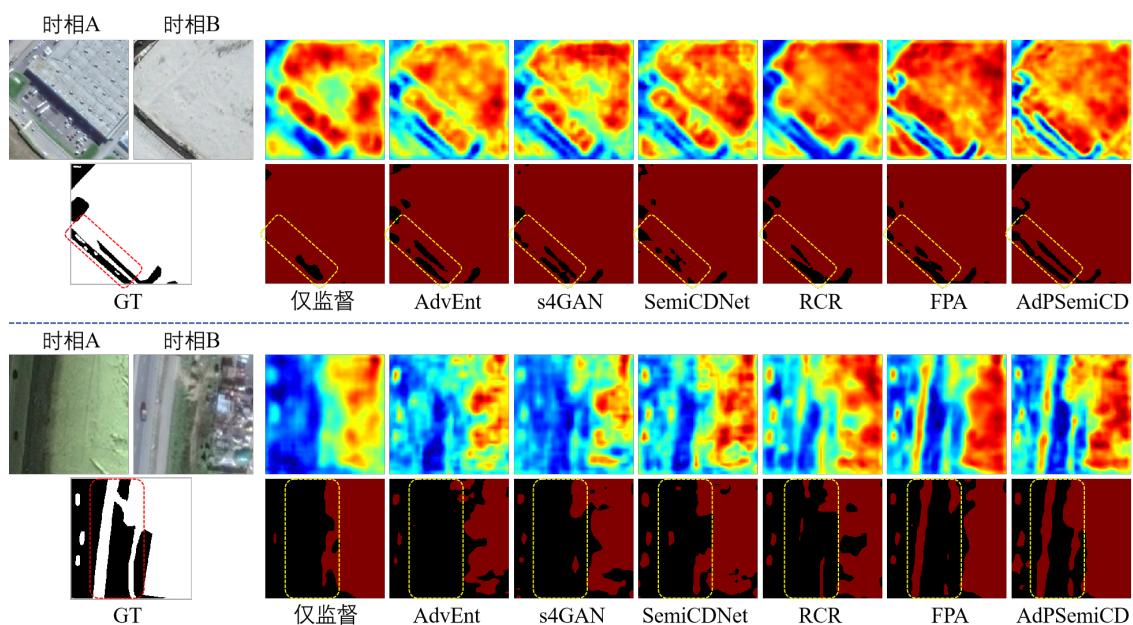


图 4-5 AdaSemiCD 与 SOTA 方法在 CDD 测试集上的变化检测结果可视化对比

在具有挑战性的场景下，半监督变化检测方法可能会遇到错误引导的问题，从而导致这种偏差在训练过程中被进一步放大。因此，在这种情况下，从可视化图中可以看出，一些半监督方法可能比仅监督训练的基线方法表现得更差。然而，我们的模型在训练阶段通过自适应融合，一定程度上排除了部分噪声，并在整个训练过程中自适应地集成完成进化的模型参数，在这些复杂的情况下，伪标签的质量会逐渐提高，直到它们达到可靠的标准，从而为模型训练提供准确的信号。因此，我们的方法在这些具有挑战性的领域表现出了优越的性能，如图4-4和图4-5中强调的区域。

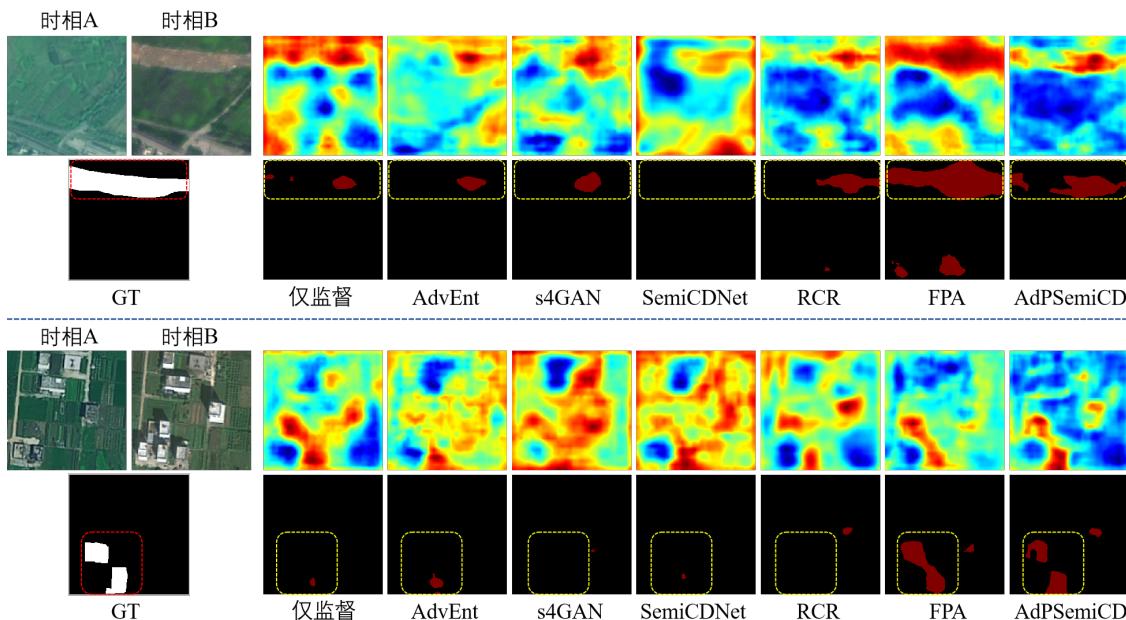


图 4-6 AdaSemiCD 与 SOTA 方法在 CL-CD 测试集上的变化检测结果可视化对比

4.3.4 消融实验

在本小节中，我们进行了消融研究，主要是为了验证每个模块在我们提出的方法中的作用，并探索我们在实验中的一些超参数选择。

(1) 自适应机制有效性

由于对比的半监督方法与本文模型结构的差异，我们一开始并没有急于验证我们所提出的自适应机制的优越性。相反，我们首先使用经典的平均教师框架作为基线进行了实验，除了为其设置的固定超参数来控制无监督损失的权重外，其余的数据增强和变化检测网络与所有方法保持相同。与单模型和共享权值的双分支网络相比，平均教师框架的优势非常明显。这也证明了扰动一致性和移动指数平均的有效性，在此基础上我们探讨了所提出的自适应机制的效力，此实验在 LEVIR-CD 数据集上进行。

我们将所提出的 AdaEMA 模块和 AdaFusion 模块分别添加到平均教师框架中，在 IoU^c 上分别取得了 1.2 和 5.1 个百分点的平均改进，如表4-3所示。表中的 * 表示，在 AdaFusion 中，采用自适应机制判断是否进行融合，而在选择融合区域时采用随机选择策略。这与 AdaFusion 自适应融合之间在性能上存在明显的差距。最后，完整的方法在

表 4-3 AdaSemiCD 在 LEVIR-CD 数据集上的消融研究

Method	5%		10%		20%		40%	
	IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
仅监督	61.0	97.60	66.8	98.13	72.3	98.44	74.9	98.60
MT-EMA	67.1 (+6.1)	98.14 (+0.54)	75.0 (+8.2)	98.63 (+0.50)	76.6 (+4.3)	98.71 (+0.27)	77.0 (+2.1)	98.73 (+0.13)
MT-AEMA	68.9 (+7.9)	98.23 (+0.63)	76.1 (+9.3)	98.66 (+0.53)	77.7 (+5.4)	98.78 (+0.34)	77.8 (+2.9)	98.78 (+0.18)
(MT-EMA)+AF*	72.0 (+11.0)	98.43 (+0.83)	76.8 (+10.0)	98.72 (+0.59)	77.5 (+5.2)	98.74 (+0.30)	78.5 (+3.6)	98.80 (+0.20)
(MT-EMA)+AF	77.0 (+16.0)	98.72 (+1.12)	78.8 (+12.0)	98.83 (+0.70)	80.4 (+8.1)	98.91 (+0.47)	80.0 (+5.1)	98.90 (+0.30)
AdaSemiCD	77.7 (+16.7)	98.78 (+1.18)	79.4 (+12.6)	98.87 (+0.74)	80.3 (+8.0)	98.92 (+0.48)	80.6 (+5.7)	98.93 (+0.33)

任意单个模块的性能上去的了进一步的提升，这表明我们提出的两个模块是解耦的，模型体系结构是合理的。

表 4-4 在 LEVIR-CD 数据集上伪标签量化指标的消融研究

Method	5%		10%		20%		40%	
	IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
仅监督	61.0	97.60	66.8	98.13	72.3	98.44	74.9	98.60
Entropy	73.2 (+12.2)	98.50 (+0.90)	77.4 (+10.6)	98.75 (+0.62)	78.6 (+6.3)	98.81 (+0.37)	79.2 (+4.3)	98.89 (+0.29)
Entropy+rebalance	76.0 (+15.0)	98.63 (+1.03)	78.7 (+11.9)	98.82 (+0.69)	79.6 (+7.3)	98.85 (+0.41)	79.6 (+4.7)	98.86 (+0.26)
Entropy+confusion	74.6 (+13.6)	98.61 (+1.01)	78.0 (+11.2)	98.79 (+0.66)	79.2 (+6.9)	98.82 (+0.38)	79.8 (+4.9)	98.87 (+0.27)
Uncertainty	77.7 (+16.7)	98.78 (+1.18)	79.4 (+12.6)	98.87 (+0.74)	80.3 (+8.0)	98.92 (+0.48)	80.6 (+5.7)	98.93 (+0.33)

(2) 伪标签评估指标

如前面小节所述，我们提出的伪标签评估度量基于信息熵，并通过类重新平衡和混淆区域放大来增强其表示能力。为了验证该方设计的有效性，我们首先仅使用信息熵对伪标签进行评估，然后逐步地加入类别重平衡和易混淆区域放大，最后使用完整的不确定性指标进行评估，通过不同评估指标下对模型最终性能的影响来验证我们的设计。

实验结果如表4-4所示。当仅使用信息熵作为实现自适应训练机制的评估指标时，性能已经得到了显著的提高，这很大程度上归功于我们设计的 AdaFusion 和 AdaEMA 模块。结合类别重平衡和混淆区域放大之后都进一步提高了性能，表明这些改进能够在训练过程中更准确地对伪标签作出评估，从而促进更精确的自适应操作。值得注意的是，类别重平衡带来的改进超过了易混淆区域放大所带来的改进。这主要是因为在类别重平衡的影响下，伪标签的评估更加强调前景，在一定程度上忽略了背景的识别，从而允许 AdaFusion 更多地关注前景中的不可靠区域。此外，随着标记数据量的增加，易混淆区域放大产生了更大的正向影响。最后，使用包含以上所有部分的不确定性作为评估指标，进一步的在各自的基础上提升了性能，这表明这两个改进是相互独立的，可以有效地结合在一起以进行更准确的评估。

此外，我们对训练过程中生成的伪标签进行了存储，图4-7展示了在整个训练过程中，伪标签评估指标中是否进行类别重平衡的情况下，伪标签和对应的 Ground Truth 之间计算的 IoU^c (注意：这里使用的 Ground Truth 仅用于计算 IoU^c ，而不涉及训练过程中的任何其他方面)。显然，尽管类别重平衡不直接作用于改进伪标签，但它能够更准确地评估伪标签，从而在后续的自适应训练机制中间接地提高了伪标签质量，使它们更接近真正的标签。

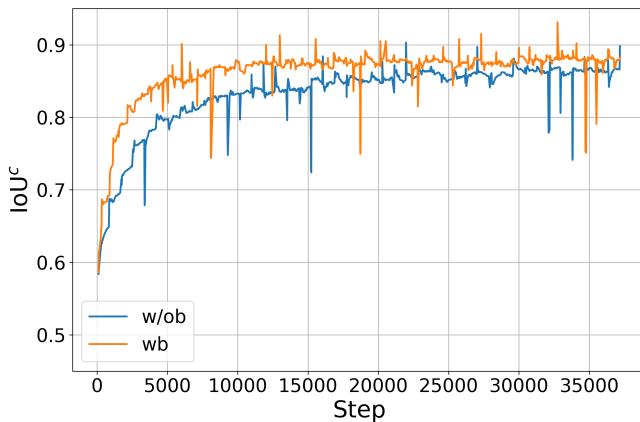


图 4-7 类别重平衡对训练中未标记样本生成的伪标签质量的影响

(3) 超参数选择

无监督损失权重的预热过程对 AdaSemiCD 在半监督变化检测上的性能有重大影响。因此，我们对控制预热过程的两个超参数 (γ 和 w_{max}) 的选择进行了消融实验。如表4-5所示，在参数 (0.1, 10)、(0.1, 0.1) 和 (0.1, 1.0) 的组合下，我们的方法在 LEVIR-CD、WHU-CD、CDD 三个数据集上的性能最佳。而且，我们的方法对该超参数敏感，参数选择不当会造成较大的性能差异。这是因为我们的方法同时对标记样本进行监督训练和对未标记样本进行无监督训练。如果不能很好地平衡两者之间的关系，就会导致标记样本的过拟合或未标记样本的过多噪声干扰。我们剩下的所有实验都是在这个超参数设置下进行的，比较方法的超参数与原论文中的最佳设置一致。但总的来说，通过经验积累发现， γ 我们总是设置为 0.1， w_{max} 可以在 0.1, 1.0, 10.0, 30.0 中进行尝试，由于深度学习的黑盒性，单阶段半监督学习中监督和无监督的互相影响因素较多，我们暂时无法得出一组在所有情况下都适用的最优参数。在其余数据集上我们仅进行了轻微的超参调整就取得了可观的实验结果，我们将 w_{max} 分别设置为：LEVIR-CD+: 10.0; GZ-CD: 5.0; EGY-CD: 5.0; HRCUS-CD: 1.0; DSIFN-CD: 10.0; SYSU-CD: 1.0; CL-CD: 1.0。

(4) 计算资源

虽然引入了额外的自适应机制有效地提高了检测性能，但是一个不可避免的问题就是增加了额外的计算资源消耗，如果耗费了大量的资源而仅仅带来了有限的提升，则得不偿失，因此我们进行了计算资源耗费的实验验证。

实验结果如表4-6所示，由于我们对所有实验方法都公平地采用了相同的变化检测网络，因此训练参数数量 (46.85M) 和计算量 (585.85 GFLOPs) 都保持相同，我们的 AdaSemiCD 推理时间方面与其他方法相当。训练时间较长的主要原因是每次训练迭代过程中需要生成和评估两次伪标签，经过统计这一步大约需要 0.3s 左右，而融合和 EMA 参数更新分别只需要 0.006s 和 0.03s 左右。我们的模型仅仅是前期训练时间增加

表 4-5 超参数敏感性消融实验结果

γ	w_{max}	LEVIR-CD		WHU-CD		CDD	
		IoU^c	OA	IoU^c	OA	IoU^c	OA
0	0 (Sup.Only)	66.8	98.13	55.7	97.53	67.9	95.46
0.05	1.0	67.2	98.17	53.8	97.02	74.4	96.35
0.1	1.0	71.8	98.32	61.0	98.10	77.3	96.89
0.3	1.0	69.9	98.26	59.4	98.03	76.2	96.56
0.5	1.0	68.7	98.15	60.9	98.25	76.3	96.58
1.0	1.0	67.3	98.13	60.5	98.18	72.3	95.98
0.1	0.1	65.2	97.60	70.8	98.70	71.6	95.77
0.1	0.5	68.3	98.14	66.9	98.54	75.8	96.67
0.1	5.0	73.9	98.75	60.1	98.00	69.1	95.51
0.1	10.0	79.4	98.87	52.4	97.40	68.2	95.50
0.1	30.0	71.9	98.42	50.34	97.12	65.4	95.20

了 25%，但是取得了显著的性能优势，很好地平衡了性能和计算耗费，这对其实际应用具有重要的意义。

表 4-6 LEVIR-CD 数据集上不同方法的参数、计算复杂度和训练时间的比较

Method	Params(M)	FLOPs(G)	Training Time(s)	Inference Time(ms)	IoU^c
Sup.Only	46.85	585.85	77	56	61.0
AdvEnt ^[112]	46.85	585.85	405	63	66.1
s4GAN ^[113]	46.85	585.85	585	58	64.0
SemiCDNet ^[5]	46.85	585.85	408	75	67.6
RCR ^[6]	46.85	585.85	742	59	72.5
FPA ^[7]	46.85	585.85	727	68	73.7
AdaSemiCD	46.85	585.85	915	67	77.7
Oracle	46.85	585.85	293	55	77.9

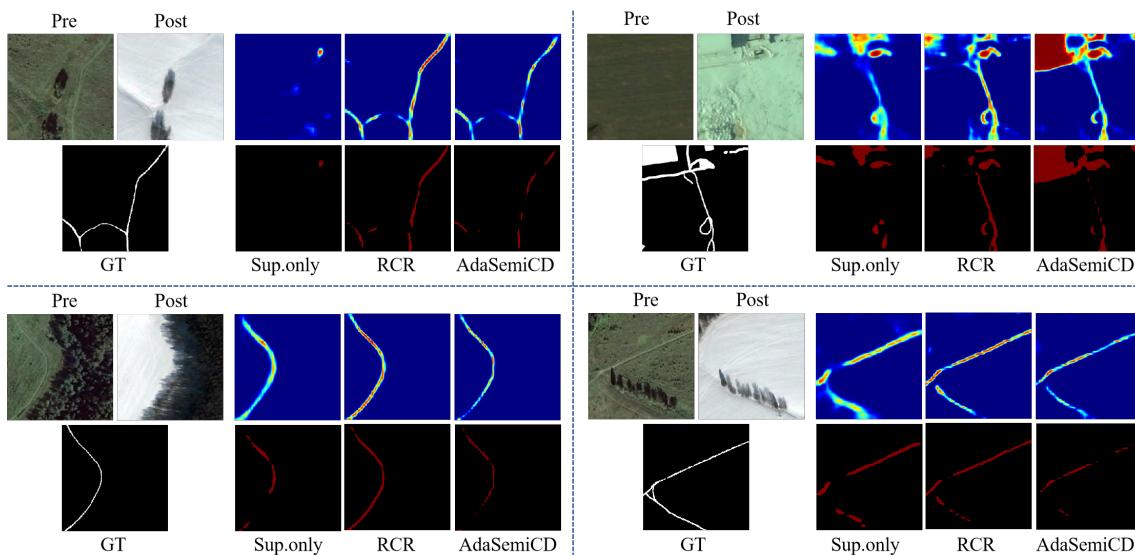


图 4-8 AdaSemiCD 的一些失效场景

4.3.5 局限性分析

虽然我们的方法在大多数情况下都是非常有效的，但我们在实验中也发现在一些特殊场景下依然存在一定的局限性。如图4-8所示，在检测一些细长状变化区域时，我们的方法表现比较挣扎，甚至相比监督基线表现更差。而在识别和检测块状变化区域方面具有很大的优势，如图中右上角的例子所示。这是因为当我们的 AdaFusion 进行样本融合时，它是 patch 到 patch 的融合，很容易截断图像中这种细长状变化的区域，而这种变化类型带有标注的真实训练数据太少，导致没有有效的知识可供学习。

因此，或许更有针对性的融合方法可以避免这种信息丢失。例如，将变化区域视为一个实例，在实例级别进行样本融合。这样，在去除噪声、引入额外的监督信息、扩大样本多样性的同时，可以在很大程度上保留不同变化类型的特征。

4.4 本章小结

本章提出了基于伪标签评估的自适应半监督变化检测算法，首先介绍了基于平均教师模型的整体半监督变化检测框架，包括监督训练、基于一致性正则化的无监督训练两个部分，介绍了总体的损失函数。具体来说，我们对教师模型对未标记训练样本生成的伪标签质量进行评估，并根据评估结果对样本进行自适应的融合，提出了 AdaFusion 模块，以及对模型参数更新进行自适应的选择，提出了 AdaEMA 模块，最终提出了 AdaSemiCD 自适应半监督框架。随后我们介绍了统一的实验设置，与其他几种方法进行了对比实验，从定量和定性两个维度证明了 AdaSemiCD 的先进性，并设计消融实验验证了各个自适应模块的有效性以及我们的方法中超参数的最优选择。最后从性能和计算开销方面进行了实验探究，我们的方法仅增加了微弱的训练时间开销，而取得了性能上的巨大提升，证明我们的方法是具有实际应用意义的。

第5章 基于自适应扰动的半监督变化检测算法

5.1 引言

在此前的研究中，半监督变化检测领域的最佳方法是基于一致性正则化 + 自训练的整体方法，它假设对未标记的图像进行准确的预测，使得扰动不会将图像特征推到真实(隐藏)分类决策边界的错误侧。但是其中存在的两个关键问题就是：(1) 模型对未标记图像的预测并不总是足够准确的；(2) 对无标记图像施加的扰动并不总是对模型训练有效的。

针对问题一，我们在第四章的工作从样本融合方面入手，一定程度地排除了错误伪标签。但这种基于图像块融合的方法还是产生了一些不可避免的问题，即丢失掉了一些变化类型（例如河流、公路）的特征，使得模型在这些类型上的检测效果不佳。我们考虑采用更加准确的融合方式，例如实例级融合，然而这还涉及额外的双时相图像实例分割以及相应的边缘平滑处理运算，这会导致计算资源的耗费成倍数增加。于是我们从 RCR^[6] 以及经典的半监督分类算法 FixMatch^[50] 中受到启发，采用模型在图像及其扰动版本下的预测概率分布差异作为一致性约束，而不使用单一硬标签作为训练指导，从而减轻错误预测的影响。

半监督学习中的数据扰动主要作用于使得模型在相同的输入上产生预测不一致，从而扩充模型学习从一张样本上能够学到的特征空间，但是这种扰动没有恒定和具体的目标，能确定的仅有预先定义的离散搜索空间，造成的影响就是有可能造成过大的确认偏差，即在扰动前后模型对其预测概率分布差异较大时进行强行拉近，这反而会破坏模型的泛化性，进而学到无关或错误的信息，特别是在模型的预测置信度较低的情况下。因此针对问题二，对特征扰动进行精确的调整以更好地适应半监督训练是有必要的。最终我们提出了 AdPSemiCD (Semi-supervised change detection algorithm based on sample level adaptive feature perturbation)，我们的基本思想是：对模型在无标记样本上的预测进行评估，基于评估结果为每个样本生成一个受限制的特定离散扰动范围，包括特征扰动的数量以及每种扰动的扰动强度。从而在实现一定程度的可控随机性扰动，减少模型的确认偏差，并且提供更加有效的一致性约束，提高模型的泛化性能。

本章介绍了基于自适应特征扰动的半监督变化检测算法——AdPSemiCD，首先从全局角度阐述了算法的框架和流程，随后具体地介绍了我们设计的基于无标记样本预测评估的自适应特征扰动方法。最后在实验部分报告了我们在十个公开数据集上对本章方法进行实验验证取得的实验结果，包括定量实验指标和定性可视分析，此外还对本章提出的自适应扰动模块进行了消融分析，证明了本章研究的有效性。

5.2 基于自适应扰动的半监督变化检测框架

5.2.1 整体框架

与第三章和第四章的研究相同，本章的研究仍然建立在平均教师模型这种简单的半监督学习框架之上，虽然简单但是有效，整体框架如图5-1所示。

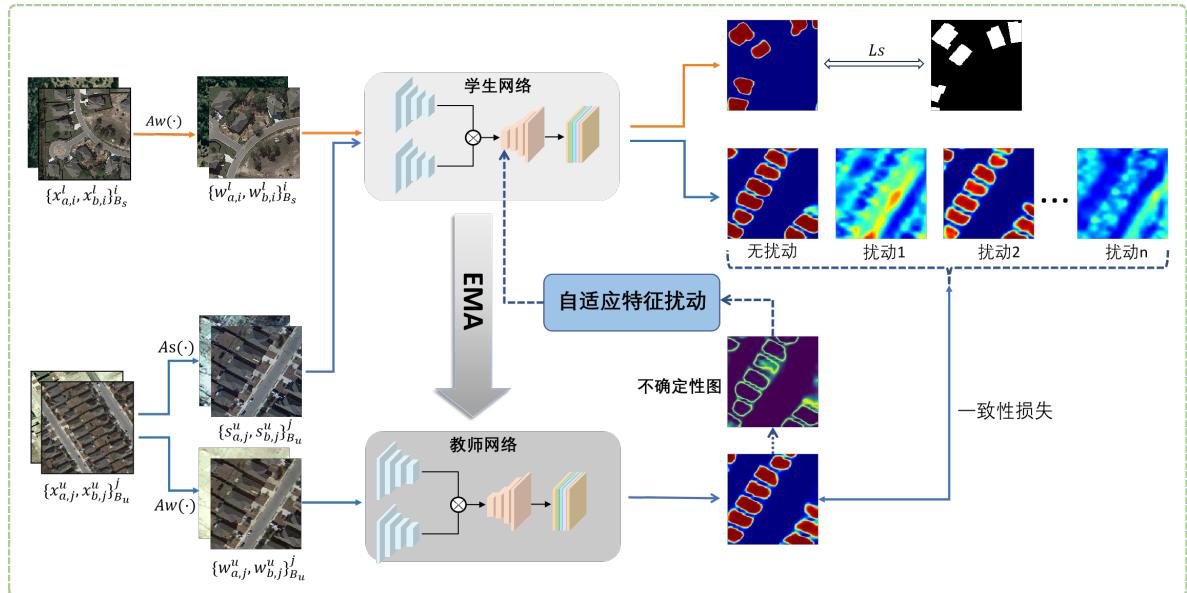


图 5-1 AdPSemiCD 算法的整体框架

该框架由学生模型和教师模型组成，分别由 θ_s 和 θ_t 进行参数化。具体地说，教师模型能够对无标记图像对进行预测并对其进行评估，教师模型的参数权重通过学生模型权重的指数移动平均来平滑地进行更新，公式表示如下：

$$\theta_t = \beta\theta_t + (1 - \beta)\theta_s \quad (5-1)$$

其中， β 是常用的动量参数，在本章方法中默认设置为 0.996。

另外，在每次训练迭代中，给定一批标记样本 $B_l = \{\{x_{a,i}^l, x_{b,i}^l\}, y_i^l\}_{i=1}^{|B_l|}$ 和一批未标记样本 $B_u = \{x_{a,i}^u, x_{b,i}^u\}_{i=1}^{|B_u|}$ ， $|B_l|$ 和 $|B_u|$ 分别是标记样本批次和无标记样本批次的样本数量。我们的训练目标是同时最小化有监督损失 \mathcal{L}_s 和无监督一致性损失 \mathcal{L}_u 来训练学生模型。因此，学生模型的总训练损失是：

$$\mathcal{L} = \mathcal{L}_s + \lambda_u \mathcal{L}_u \quad (5-2)$$

其中 λ_u 为无监督损失的权重，在本章中采取与第四章相同的预热函数作为动态调整方案。与大多数图像分割算法相同，我们选择标准的交叉熵损失作为监督训练损失，其计算如下：

$$\mathcal{L}_s = \frac{1}{|\mathcal{B}_l|} \sum_{i=1}^{|\mathcal{B}_l|} \mathcal{L}_{ce}(M(\theta_s, (x_{a,i}^l, x_{b,i}^l)), y_i^l) \quad (5-3)$$

这里的 y_i 和 \hat{y}_i 分别表示真实二值标签和模型预测概率。

对于无监督损失，我们需要的是衡量模型在特征扰动前后预测的相似性。因此选择均方误差 (Mean Squared Error, MSE)，它的核心思想是计算预测值与真实值之间的平方差的平均值，其计算如下：

$$\mathcal{L}_{MSE} = \frac{1}{|\mathcal{B}_u|} \sum_{i=1}^{|\mathcal{B}_u|} (y_i - \hat{y}_i)^2 \quad (5-4)$$

其中 y_i 和 \hat{y}_i 分别表示无特征扰动和施加特征扰动之后的模型预测概率，其具体定义我们将在下一小节给出。

最后，对于正则化常数 λ_u ，我们采用和第四章相同的预热函数，该权重随着训练进程从初始权值 0 不断增大直至达到预定的最大权重。

5.2.2 自适应特征扰动

在现有的大多数半监督学习研究中，无论是在分类任务中还是在分割任务中，各种扰动技术得到了广泛应用。在半监督学习中，数据扰动的目标是从同一幅图像生成两个不同的视图，其中不需要特定的最优增强策略。此外，正如文献^[116] 中所讨论的，过度失真的扰动会损害数据分布，降低半监督学习的效果。因此对于每个训练样本，我们应该限定一个扰动空间的范围，以达到最佳的效果。为此，我们设计了一种基于评估无标记样本预测结果的自适应特征扰动方法。

首先我们继续沿用在第四章中设计的伪标签评估指标，对模型在无标记样本上的预测结果进行评估，但不同之处在于我们为了方便利用此评估计算自适应扰动的扰动数量和扰动强度，我们对第四章中的评估指标进行了逆变换，即计算所得从不确定性图变为了确定性图，计算公式如下：

$$C(y_i) = D(y_i) \cdot E'(y_i) \quad (5-5)$$

其中 y_i 表示教师模型在无标记样本上的预测概率，此外 $D(\cdot)$ 和 $E'(\cdot)$ 计算公式如4-3、4-4以及4-7。

接下来对整个无标记样本批次内的样本确定性图 $\{C_j\}_{j=1}^{|\mathcal{B}_u|}$ 进行归一化得到 $\{C'_j\}_{j=1}^{|\mathcal{B}_u|}$ ，在此我们使用最大归一化，公式如下：

$$C'_j = \frac{C_j - C_{\min}}{C_{\max} - C_{\min}} \quad (5-6)$$

其中 $C'_j \in [0, 1]$ ，接下来将其作为一个缩放因子，来计算得到每对样本的特征扰动数量，如公式5-7，其中 Num 表示扰动池中的扰动方法总数。最后从扰动池中为该样本随机抽取相应数量的扰动方法作用于其在学生模型上融合后的高维特征。

$$num_j = C'_j \times Num \quad (5-7)$$

其中扰动池包含以下七种特征扰动，扰动名称、扰动方式以及主要参数和其预定义取值

范围见表5-1。

表 5-1 AdPSemiCD 使用的特征扰动方法

扰动名称	作用描述	主要参数
无扰动	保留原始特征	—
特征加噪声	添加噪声，干扰特征表示	$uniform_{range} \in [0, 1.0]$
特征丢弃	丢弃部分特征	$threshold \in [0, 1.0]$
特征裁剪	仅保留一定范围内的特征	$threshold \in [0, 1.0]$
引导特征遮挡	遮挡部分背景特征区域	$erase \in [0, 1.0]$
特征模糊	通过平滑操作使特征值分布更加均匀	$\sigma \in [0.5, 2.0]$
通道扰动	不同通道的特征进行互换	$num \in [10, 50]$
对抗扰动	对部分区域特征施加对抗扰动	$\epsilon \in [1.0, 5.0]$

对于选中的扰动，我们同样使用归一化后的确定性作为缩放因子，为每对样本计算所选取扰动的参数，来自适应地控制扰动的强度。计算公式如下：

$$\theta_j^i = C'_j \times \theta^i \quad (5-8)$$

$\{\theta^i\}_{i=1}^{Num}$ 即所有扰动的预定义参数集合，如表5-1中所列举。

最后，将每对样本自适应选择的扰动施加到相应样本经过学生模型编码器和特征差异模块之后的融合特征值上，扰动后的特征再经过解码器进行上采样从而输出为扰动版本的预测概率。假设经过编码器和特征差异模块之后的图像特征为 $\mathcal{F}_{s,j}$ ，则扰动过程可表示为：

$$\hat{\mathcal{F}}_{s,j} = \mathcal{F}_{s,j} \oplus \mathcal{P}_i(\theta^i) \quad (5-9)$$

得到 num_j 个扰动特征 $\{(\hat{\mathcal{F}}_{s,j})_n\}_{n=1}^{num_j}$ ，再分别输入解码器中得到 num_i 个扰动预测概率 $\{(\hat{p}_{s,j})_n\}_{n=1}^{num_j}$ ，

$$\hat{p}_{s,j} = f_d(\hat{\mathcal{F}}_{s,j}) \quad (5-10)$$

当然此外还有一个无扰动版本的学生模型输出，即 $p_{s,j} = f_d(\mathcal{F}_{s,j})$ 。于是，在一致性约束之下，我们在无标记样本上的无监督训练目标就是拉近这些扰动版本输出和无扰动版本输出以及教师模型在弱增强样本上的输出之间的距离，最小化输出概率分布之间的差异。无监督损失计算如下：

$$\mathcal{L}_u = \frac{1}{|\mathcal{B}_u|} \sum_{j=1}^{|\mathcal{B}_u|} \left(\sum_{i=1}^{num_j} \mathcal{L}_{MSE} ((\hat{p}_{s,j})_i, p_{w,j}) + \mathcal{L}_{MSE} ((p_{s,j})_i, p_{w,j}) \right) \quad (5-11)$$

5.3 实验结果及分析

5.3.1 实验设置

本章所有实验数据集划分以及共同的实现细节部分均与第四章保持一致。

5.3.2 对比实验

为了验证本章提出的 AdPSemiCD 的有效性，与第四章类似，我们在十个数据集上与三种半监督变化检测算法与两种半监督语义分割算法进行了全方位的对比实验，包括实验指标对比和可视化结果对比。

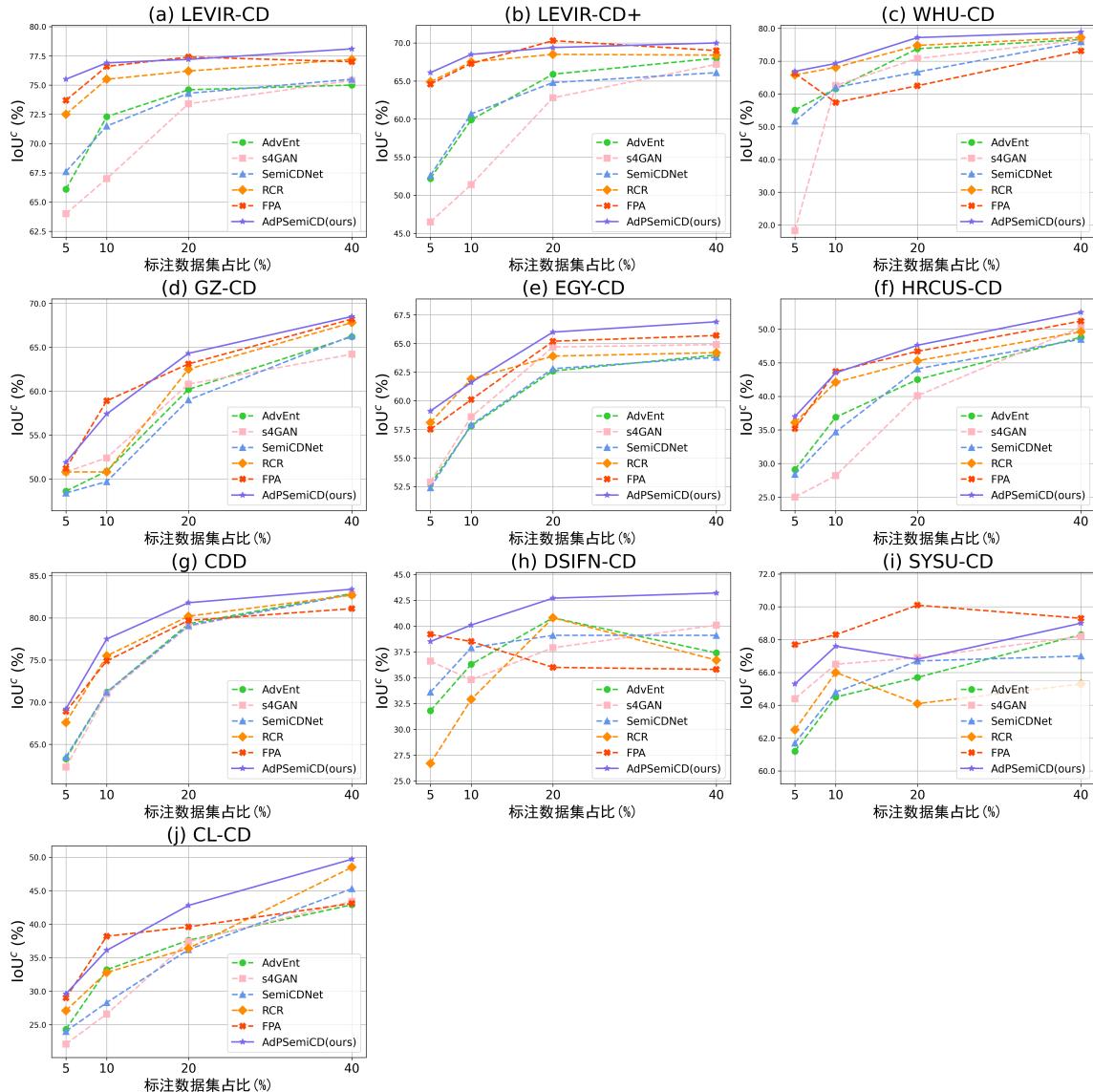


图 5-2 AdPSemiCD 与 SOTA 方法在 10 个数据集上的性能对比折线图

(1) 定量对比

我们在所有 10 个数据集上进行了实验，表5-2和表5-3分别报告了我们在建筑物变化检测数据集和多类变化检测数据集上的实验结果。其中，Oracle 是在整个训练数据集上进行全监督训练的结果。另外图5-2更直观地展示了我们的所有实验定量结果。值得注意的是，我们的方法在所有数据集的几乎所有半监督划分设置中都实现了最先进的性能。在大多数情况下，所有的半监督变化检测方法在相应的划分设置下都比仅监督的基线方法表现得更好，这证明了半监督方法在从大量未标记的训练样本中学习到了更多的

知识。

如表5-2所示，其中蓝色高亮部分代表着最优结果，下划线标注的是次优结果。我们所提出的 AdPSemiCD 框架在几乎所有建筑物变化检测数据集上都达到了最佳的效果，并且在所有数据集上表现都比 RCR 更好，这说明我们的自适应特征扰动一致性比 RCR 中采用的随机特征扰动一致性更加适合半监督变化检测。

建筑物变化检测:

具体来说，我们的 AdPSemiCD 在 LEVIR-CD、LEVIR-CD+、WHU-CD、GZ-CD、EGY-CD、HRCUS-CD 上分别在四种半监督比例设置下平均提升了 0.70、0.58、1.48、0.28、1.1、0.65 个百分点的 IoU^c 。此外，值得说明的是，在第四章 AdaSemiCD 表现不佳的数据集——GZ-CD 上，本章提出的 AdPSemiCD 取得了更好的效果，在三种半监督实验设置下都超越了 FPA 方法。正如第四章所分析的，这是由于 GZ-CD 的分辨率最低并且手工标注比较粗糙，对于 AdaSemiCD 这样基于伪标签自训练的方法提出了巨大的挑战，而本章的方法并不依赖于二值硬伪标签，而是利用一个可控搜索空间内的扰动一致性作为训练约束，克服了伪标签不准确这一困难。但是在其他相对简单的数据集上，本章 AdPSemiCD 带来的性能提升相比 AdaSemiCD 就比较有限了，这是由于在伪标签足够可靠的情况下，一致性正则化 + 自训练的半监督方法普遍比单纯的一致性正则化方法表现更好。

多类变化检测: 如表5-3所示，与建筑物变化检测数据集上的趋势和规律相同，我们的 AdPSemiCD 虽然同样在几乎所有多类别变化检测数据集和所有半监督比例设置下都实现了最佳性能，并且在所有实验中都超越了 RCR，但是总体上带来的提升相比 AdaSemiCD 更有限。

具体来说，在 CDD 数据集上，AdPSemiCD 实现了平均 1.1 个百分点的 IoU^c 提高。在与 GZ-CD 类似，标注质量低、空间分辨率低的 DSIFN-CD 上，本章的 AdPSemiCD 要比 AdaSemiCD 表现更好，实现了大幅度的性能提升， IoU^c 平均提高了 1.48 个百分点，弥补了 AdaSemiCD 存在的局限性。不难注意到，AdPSemiCD 在 SYSU-CD 数据集上几乎全方位落后于 FPA 方法，这是由于 FPA 同时使用了特征对齐和预测对齐，在 SYSU-CD 这种高质量标注的数据集上能够很好地利用伪标签指导模型学习不同类别的特征。而对于 AdPSemiCD 和 RCR，存在的挑战就是缺乏语义信息，而 SYSU-CD 中又存在一些非常难以识别的变化类型，例如稀疏的草地变化为裸土，两种类型特征极为相似，导致难以区分。因此 FPA 在该数据集上领先 RCR 以及在其基础上改进的本章算法。在 CL-CD 数据集上，AdPSemiCD 平均提高了 0.73 个百分点的 IoU^c 。

表5-2 不同CD方法在建筑物变化检测数据集上的定量性能对比

数据集	方法	5%		10%		20%		40%	
		IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
LEVIR-CD	仅监督	61.0	97.60	66.8	98.13	72.3	98.44	74.9	98.60
	AdvEnt ^[112]	66.1	98.08	72.3	98.45	74.6	98.58	75.0	98.60
	s4GAN ^[113]	64.0	97.89	67.0	98.11	73.4	98.51	75.4	98.62
	SemiCDNet ^[5]	67.6	98.17	71.5	98.42	74.3	98.58	75.5	98.63
	RCR ^[6]	72.5	98.47	75.5	98.63	76.2	98.68	77.2	98.72
	FPA ^[7]	73.7	98.57	76.6	98.72	77.4	98.75	77.0	98.74
	AdPSemiCD	75.5	98.63	76.9	98.70	77.2	98.74	78.1	98.80
Oracle		$IoU^c=77.9$ and OA=98.77							
LEVIR-CD+	仅监督	52.0	97.72	58.4	98.06	66.1	98.31	66.2	98.42
	AdvEnt ^[112]	52.2	97.68	59.9	98.11	65.9	98.37	68.0	98.51
	s4GAN ^[113]	46.5	97.25	51.4	97.66	62.8	98.18	67.2	98.46
	SemiCDNet ^[5]	52.6	97.66	60.7	98.24	64.8	98.37	66.1	98.38
	RCR ^[6]	64.9	98.25	67.5	98.45	68.5	98.52	68.4	98.51
	FPA ^[7]	64.6	98.30	67.3	98.40	70.3	98.64	69.0	98.59
	AdPSemiCD	66.1	98.40	68.5	98.49	69.4	98.58	70.0	98.62
Oracle		$IoU^c=70.5$ and OA=98.63							
WHU-CD	仅监督	50.0	97.48	55.7	97.53	65.4	98.20	76.1	98.94
	AdvEnt ^[112]	55.1	97.90	61.6	98.11	73.8	98.80	76.6	98.94
	s4GAN ^[113]	18.3	96.69	62.6	98.15	70.8	98.60	76.4	98.96
	SemiCDNet ^[5]	51.7	97.71	62.0	98.16	66.7	98.28	75.9	98.93
	RCR ^[6]	65.8	98.37	68.1	98.47	74.8	98.84	77.2	98.96
	FPA ^[7]	66.3	98.45	57.4	97.69	62.5	98.48	73.1	98.69
	AdPSemiCD	66.9	98.54	69.3	98.65	77.2	98.94	78.9	99.09
Oracle		$IoU^c=85.5$ and OA=99.38							
GZ-CD	仅监督	47.5	93.56	51.4	94.26	58.0	95.65	66.3	96.62
	AdvEnt ^[112]	48.6	94.39	50.9	94.89	60.2	95.79	66.2	96.58
	s4GAN ^[113]	50.8	94.38	52.4	94.98	60.8	95.94	64.2	96.39
	SemiCDNet ^[5]	48.4	93.58	49.7	94.79	59.0	95.66	66.3	96.57
	RCR ^[6]	50.8	93.82	50.8	94.69	62.5	96.07	67.8	96.61
	FPA ^[7]	51.2	93.92	58.9	95.78	63.1	96.26	68.2	96.82
	AdPSemiCD	51.9	94.58	57.4	95.59	64.3	96.30	68.5	96.88
Oracle		$IoU^c=69.0$ and OA=96.93							
EGY-CD	仅监督	49.8	95.73	54.6	96.38	61.4	96.83	65.1	97.25
	AdvEnt ^[112]	52.7	96.01	57.8	96.58	62.6	96.86	64.0	97.19
	s4GAN ^[113]	52.9	95.94	58.6	96.50	64.7	97.09	64.9	97.27
	SemiCDNet ^[5]	52.4	96.00	57.9	96.31	62.8	96.95	63.8	97.19
	RCR ^[6]	58.1	96.50	59.9	96.77	63.9	97.08	64.2	97.18
	FPA ^[7]	57.5	96.52	60.1	96.86	65.2	97.25	65.7	97.34
	AdPSemiCD	59.1	96.58	61.6	96.90	66.0	97.31	66.9	97.37
Oracle		$IoU^c=67.6$ and OA=97.54							
HRCUS-CD	仅监督	29.5	98.11	36.0	98.45	43.4	98.68	48.9	98.84
	AdvEnt ^[112]	29.1	98.11	36.9	98.40	42.5	98.61	48.8	98.71
	s4GAN ^[113]	25.0	97.86	28.2	98.24	40.1	98.62	50.3	98.85
	SemiCDNet ^[5]	28.4	98.00	34.7	98.44	44.1	98.68	48.5	98.74
	RCR ^[6]	36.1	98.36	42.1	98.69	45.3	98.76	49.6	98.66
	FPA ^[7]	35.2	98.37	43.7	98.65	46.7	98.82	51.2	98.81
	AdPSemiCD	37.0	98.49	43.5	98.72	47.6	98.85	52.5	98.89
Oracle		$IoU^c=59.0$ and OA=99.06							

表 5-3 不同 CD 方法在多类变化检测数据集上的平均定量性能对比

数据集	方法	5%		10%		20%		40%	
		IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
CDD-CD	仅监督	60.4	94.25	67.9	95.46	75.6	96.59	82.3	97.56
	AdvEnt ^[112]	63.3	94.65	71.2	96.01	79.3	97.14	82.9	97.66
	s4GAN ^[113]	62.3	94.69	71.0	95.94	79.0	97.10	82.8	97.63
	SemiCDNet ^[5]	63.5	94.68	71.2	95.99	79.1	97.13	82.8	97.63
	RCR ^[6]	67.6	95.40	75.5	96.57	80.2	97.26	82.7	97.61
	FPA ^[7]	68.9	95.66	74.9	96.55	79.7	97.20	81.1	97.37
	AdPSemiCD	69.2	95.74	77.5	96.90	81.8	97.53	83.4	97.72
Oracle		$IoU^c=87.8$ and OA=98.10							
DSIFN-CD	仅监督	34.8	78.34	38.9	83.41	40.2	87.00	39.6	87.00
	AdvEnt ^[112]	31.8	77.83	36.3	83.86	40.8	85.92	37.4	86.31
	s4GAN ^[113]	36.6	84.10	34.8	86.87	37.9	87.69	40.1	86.52
	SemiCDNet ^[5]	33.6	78.60	37.9	84.18	39.1	86.77	39.1	87.05
	RCR ^[6]	26.7	83.78	32.9	86.05	40.8	86.70	36.7	86.08
	FPA ^[7]	39.2	84.27	38.5	87.12	36.0	87.41	35.8	86.50
	AdPSemiCD	38.5	84.65	40.1	87.22	42.7	87.98	43.2	88.13
Oracle		$IoU^c=58.1$ and OA=90.82							
SYSU-CD	仅监督	62.9	89.57	64.4	90.18	66.0	90.82	66.4	90.93
	AdvEnt ^[112]	61.2	89.36	64.5	90.18	65.7	90.35	68.3	91.24
	s4GAN ^[113]	64.4	90.02	66.5	90.48	66.9	90.26	68.2	91.51
	SemiCDNet ^[5]	61.7	89.32	64.8	90.25	66.7	90.97	67.0	91.08
	RCR ^[6]	62.5	89.76	66.0	90.75	64.1	90.22	65.3	90.56
	FPA ^[7]	67.7	90.95	68.3	91.09	70.1	92.01	69.3	91.97
	AdPSemiCD	65.3	90.78	67.6	91.35	66.8	91.54	69.0	91.76
Oracle		$IoU^c=68.2$ and OA=91.64							
CL-CD	仅监督	18.1	91.90	31.4	92.42	37.2	93.32	45.9	94.98
	AdvEnt ^[112]	24.3	92.13	33.2	93.01	37.6	93.59	42.9	94.06
	s4GAN ^[113]	22.1	92.00	26.6	93.09	37.4	93.59	43.4	93.87
	SemiCDNet ^[5]	24.0	92.20	28.3	93.42	36.2	92.41	45.3	94.22
	RCR ^[6]	27.1	91.63	32.8	92.99	36.4	93.07	48.5	94.94
	FPA ^[7]	29.0	91.00	38.2	93.37	39.6	93.88	43.1	94.15
	AdPSemiCD	29.6	92.39	36.1	93.55	42.8	94.04	49.7	95.39
Oracle		$IoU^c=50.1$ and OA=95.55							

在多类别变化检测数据集上，可以观察到的一个现象就是，半监督方法的性能似乎存在较大的波动，在某些半监督比例设置下，相比仅监督训练的基线可以大幅度提高性能，而在同样的数据集上另外的半监督比例设置下，可能比基线效果更差，甚至增加了标记训练数据量，反而降低了性能。这都是由于监督训练和无监督训练之间的平衡被破坏。而本章方法总体上表现比较稳定，主要得益于对训练样本在一个有限的可控搜索空间中进行扰动，并且基于样本难易程度自适应调整搜索空间的范围，从而在扩充模型学习的特征空间的同时，能够保证针对性扰动以及整体扰动空间可控，确保了其稳定性。

(2) 定性对比

我们还对所有对比实验中的检测结果进行了可视化，图5-3、5-5、5-6、5-7分别展示了在 LEVIR-CD、WHU-CD、CDD、CL-CD 四个数据集的测试集上的一些变化检测结果

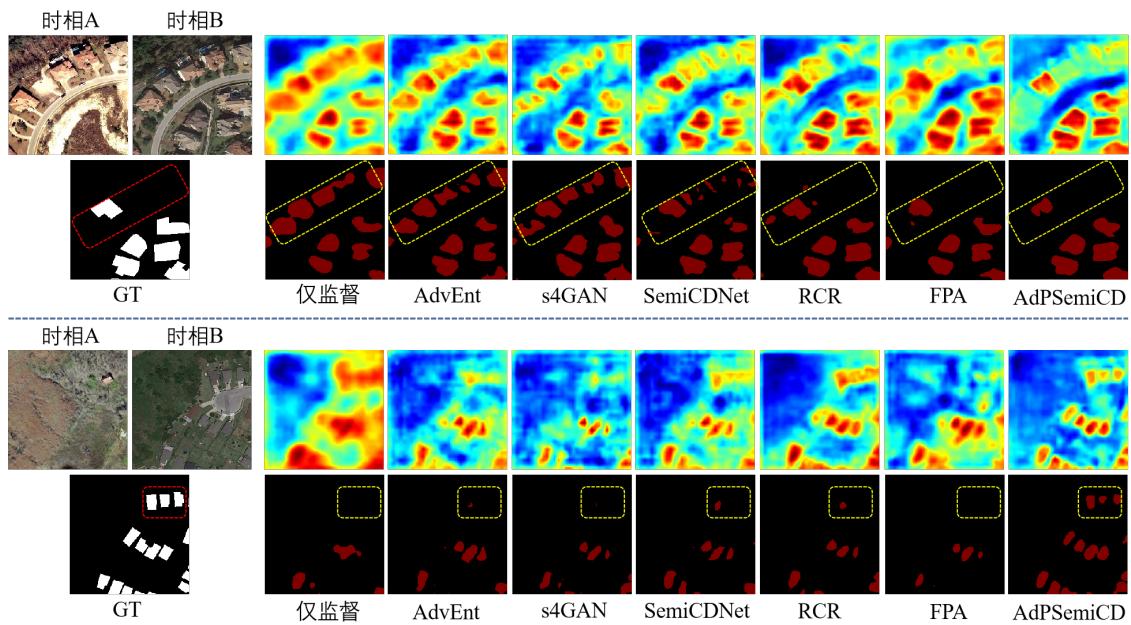


图 5-3 5% 标记训练样本下 AdPSemiCD 与 SOTA 方法在 LEVIR-CD 测试集上的可视化对比图

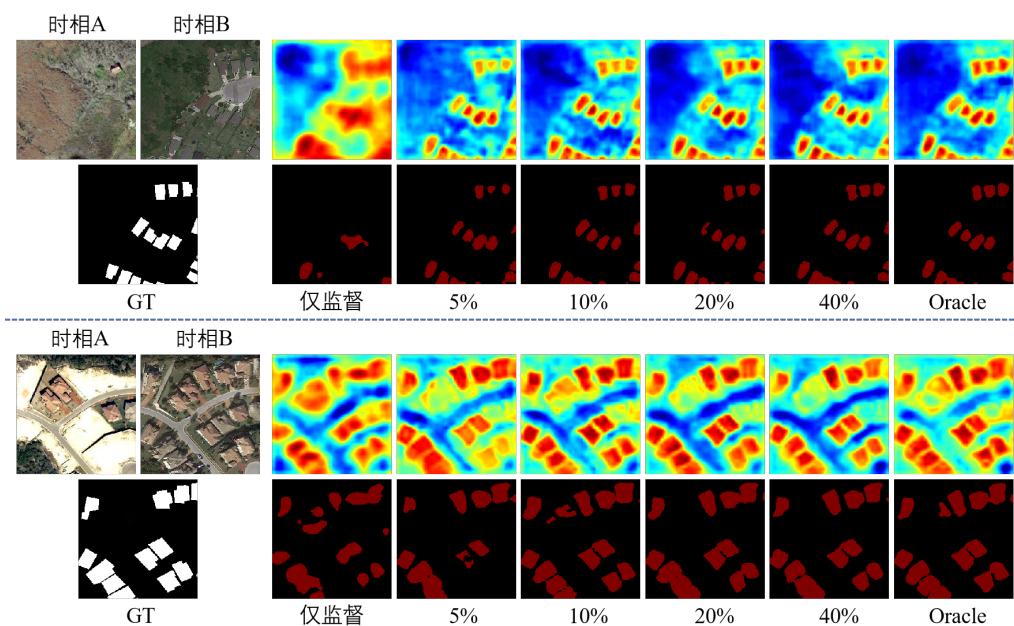


图 5-4 不同标记率下 AdPSemiCD 在 LEVIR-CD 测试集上的可视化对比图

样例，同时以热力图和二值标签的形式对模型预测结果进行了可视化表示。整体看来，本章提出的 AdPSemiCD 相比其余半监督变化检测方法在检测效果上都表现更佳，尤其是与使用随机特征扰动的 RCR 方法相比。此外，从热力图分布可以看出，在仅监督训练模式下，模型的预测概率比较分散，进一步可以推断出预测信息熵较高。半监督方法一定程度上可以更加自信地判别出变化和不变区域，表现在热力图上两类区别更加明显，高热区域更加集中，表明了半监督学习利用大量无标记样本的有效性，但是仍然存在一些判别错误的情况，我们的 AdPSemiCD 有效减轻了这一现象，取得了更好的表现。

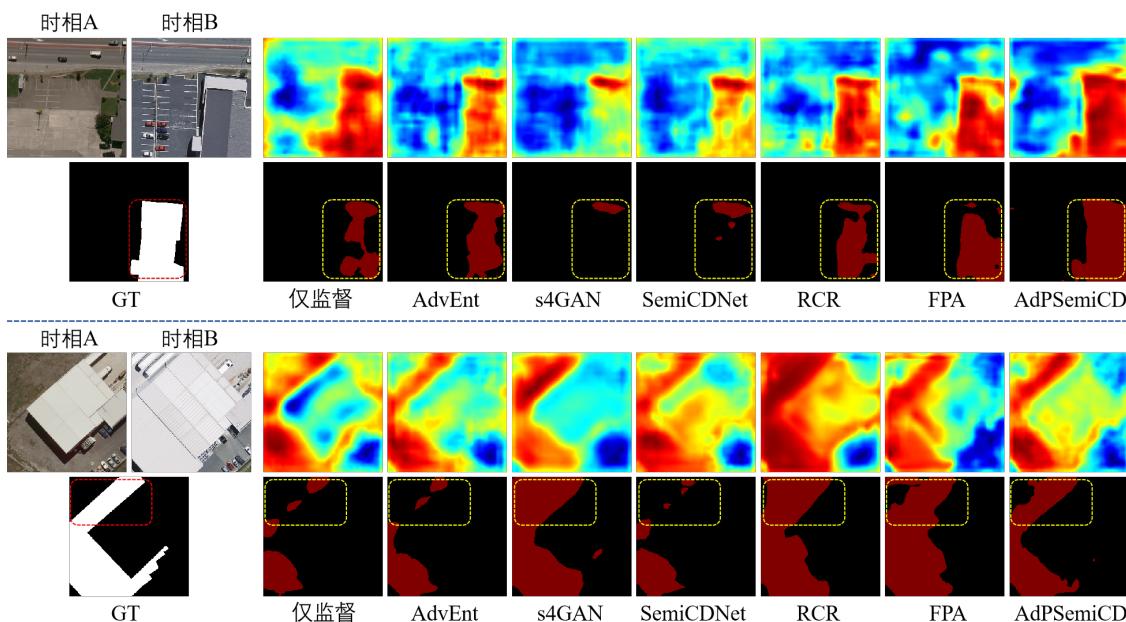


图 5-5 5% 标记训练样本下 AdPSemiCD 与 SOTA 方法在 WHU-CD 测试集上的可视化对比图

具体地说，LEVIR-CD 数据集上的可视化结果展示了本章方法在减少漏检以及小目标变化检测上的优异表现，相比其他方法，RCR 已经取得了一定的进步，而在其扰动方式上的针对性改进使得我们的方法进一步提高了性能。在 WHU-CD 数据集上，所选样例与背景的相似度极高，在第一组结果中大部分方法漏检严重，仅识别出了与背景差异较大的小部分变化区域，而第二组的结果中则是错误地将左上角的其他类别检测为建筑物变化区域。从热力图和标签图上都可以看出，本章方法能够以更高的置信度检测出来边界更加清晰的变化区域。在多类别的数据集上，本章方法同样表现很好，例如在 CDD 数据集的可视化结果中，AdPSemiCD 不仅在小目标（车辆）的变化检测上继承了 RCR 不俗的性能，在容易混淆的变化目标上同样准确地进行了检测识别。值得一提的是，从可视化结果上看，本章方法很好地弥补了第三章由于样本改造产生类别学习样本缺失进而造成细长目标识别失效的问题，如图 5-6 中第二组结果所示，最为直观地证明了本章方法达成了研究目的。对于 CI-CD 数据集，由于样本量太少，导致所有方法整体上检测效果都比较差，但正是在这种极端情况下，在图 5-7 所示的两组样例中，我们的方法在这类小目标变化上仍然表现出了强大的检测能力。

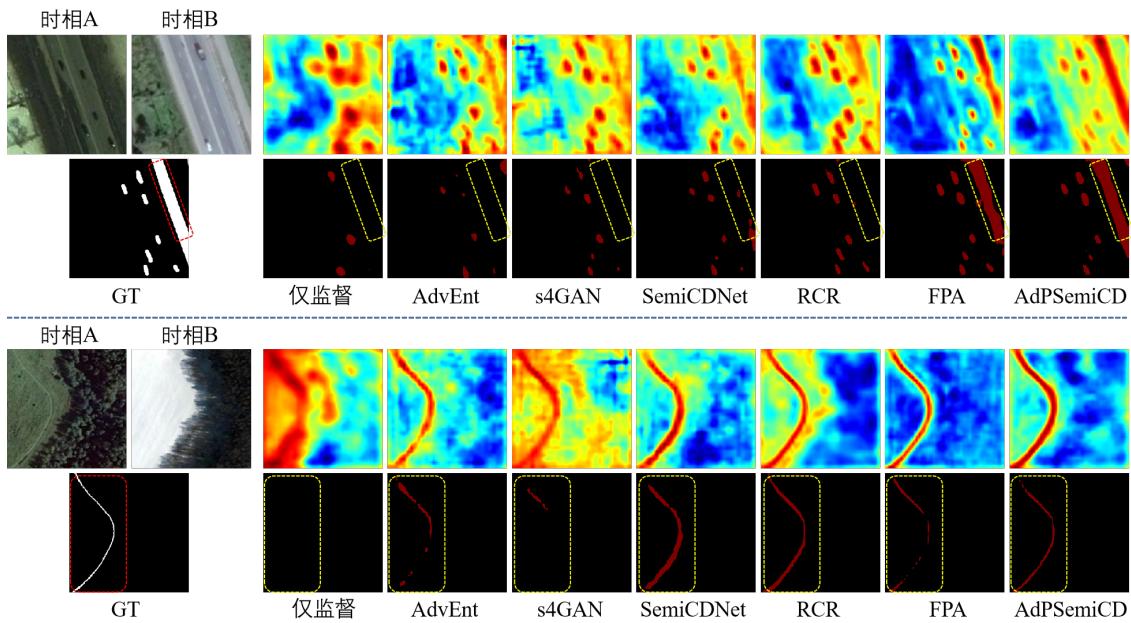


图 5-6 5% 标记训练样本下 AdPSemiCD 与 SOTA 方法在 CDD 测试集上的可视化对比图

表 5-4 AdPSemiCD 在 WHU-CD 数据集上的消融实验

Method	5%		10%		20%		40%	
	IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
仅监督	50.0	97.48	55.7	97.53	65.4	98.20	76.1	98.94
MT+NP	61.0 (+11.0)	98.08 (+0.60)	62.4 (+6.7)	98.16 (+0.63)	70.7 (+5.3)	98.69 (+0.49)	77.5 (+1.4)	98.99 (+0.05)
MT+FP	61.4 (+11.4)	98.15 (+0.67)	63.7 (+8.0)	98.28 (+0.75)	73.2 (+7.8)	98.80 (+0.60)	77.1 (+1.0)	98.96 (+0.02)
MT+RP	66.1 (+16.1)	98.40 (+0.92)	68.7 (+13.0)	98.53 (+1.00)	75.6 (+10.2)	98.88 (+0.68)	78.0 (+1.9)	98.98 (+0.04)
AdPSemiCD	66.9 (+16.7)	98.54 (+1.06)	69.3 (+13.6)	98.65 (+1.12)	77.2 (+11.8)	98.94 (+0.74)	78.9 (+2.8)	99.09 (+0.15)

我们还对 LEVIR-CD 数据集上，在不同半监督设置下的检测结果进行了可视化，如图5-4所示。首先最为显而易见的就是，我们的半监督变化检测方法所取得的检测效果相比仅监督基线提升了很多，无论是在整体识别准确率还是边界处理上都要好。随着提供的标记数据越多，变化检测性能也更好，这也从侧面反映了变化检测任务对人工标注的依赖，可以看到 Oracle（即使用整个标注训练集进行训练）和 5% 仅监督的基线之间存在非常明显的差距，因此就凸显了半监督变化检测的意义。我们的 AdPSemiCD 在 5% 标记数据上的训练效果就已经基本能够胜任变化检测任务，在 10% 至 10% 标记数据下就已经取得了和 Oracle 相当的性能，这也证明了我们的方法具有很大的应用价值。

总的来说，这些可视样例代表性地展示了本章方法与其他方法的变化检测效果对比，与定性结果相对应，AdPSemiCD 在所有数据集上的检测效果从可视化结果来看，都比 RCR 更好，这得益于其不仅保留了 RCR 中探索出的广阔的特征扰动空间对半监督变化检测更加有利这一成功经验，同时还借鉴了前文中对不同无标记样本进行自适应训练的设计，为无标记样本创造了一个自适应的可控特征扰动空间，不仅提高了模型的泛化性，还减少了确认偏差带来的负面影响，从而稳定地提升了半监督变化检测性能。

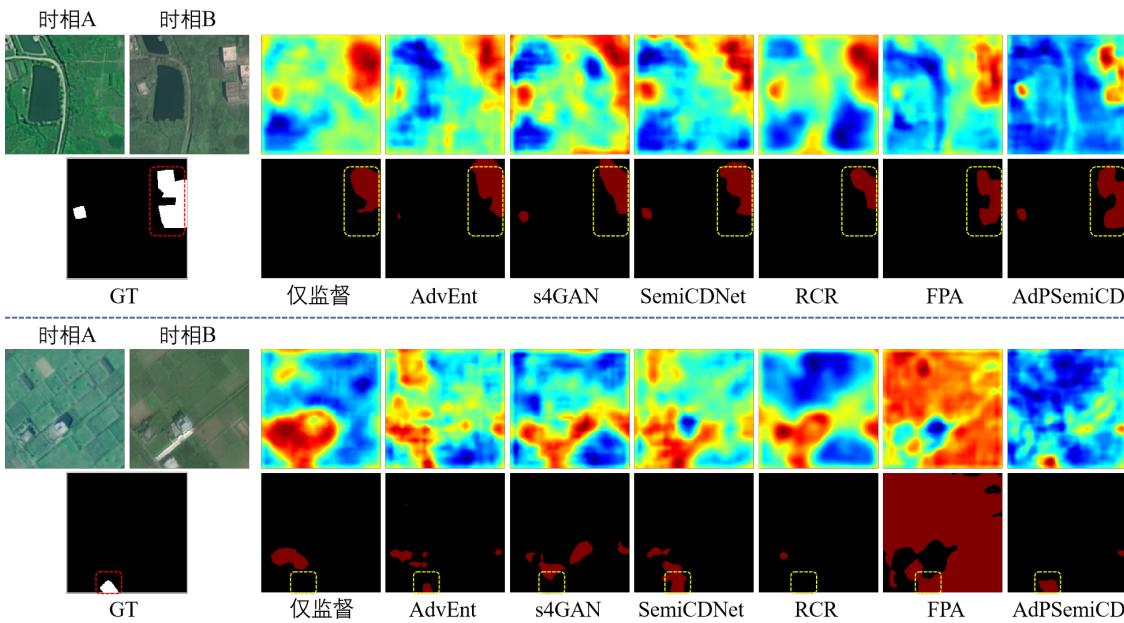


图 5-7 5% 标记训练样本下 AdPSemiCD 与 SOTA 方法在 CL-CD 测试集上的可视化对比图

5.3.3 消融实验

为了证明本章提出的自适应特征扰动有效，我们同样从定量和定性两个角度设计了消融实验。首先是在平均教师半监督学习框架之上，分别使用无特征扰动、RCR 中的随机特征扰动以及本章提出的自适应特征扰动，将前两组实验作为对照组。首先对图表中使用的缩写做出说明：

NP: 无特征扰动，即直接将教师模型、学生模型预测概率分布的均方误差作为训练损失；

FP: Fixed Perturbation，固定特征扰动，即根据经验预定义扰动强度和扰动数量；

RP: Random Perturbation，随机特征扰动，即 RCR 中使用的随机选择扰动强度和扰动数量；

最终在 WHU-CD 上取得的实验指标结果如表5-4所示，其中红色标注部分为该方法相比基线的提升值。在不进行特征扰动时，模型退化为平均教师模型实现的 FixMatch 方法，可以在基线之上大幅度地提高模型性能。使用固定特征扰动并不能进一步带来更大的改进，一方面是因为单一的固定扰动参数值很难确定最优选择，另一方面，对困难样本进一步继续施加扰动，会扩大确认偏差，所以尤其是在标记训练数据更少时，固定特征扰动带来的收益更加微乎其微。RCR 中使用的随机特征扰动具有一定的前景，相比固定扰动方案，它纳入了更多的随机性，对不同样本施加不同程度的扰动，很大程度上提高了模型的泛化性能。最后本章提出的自适应特征扰动方案进一步细化了扰动的过程，对样本进行评估，进而基于此评估自适应地调整扰动参数和数量，因此继续提升了模型的性能并且提高了稳定性。

此外，我们还对无监督损失的权重进行了消融实验，本章采取的权重与第三章

相同，由公式4-1计算，其计算主要由 γ 和 w_{max} 两个参数决定。因此我们对这两个参数的选择组合进行了消融研究，如表5-5所示，展示了我们在 LEVIR-CD、WHU-CD、CDD、CL-CD 四个数据集上 10% 标记率下取得的实验结果。我们先固定最大权重值 w_{max} 为 1.0，将 γ 作为变量控制预热进程的快慢，得到每个数据集上 γ 的最佳参数，在 LEVIR-CD、WHU-CD、CDD 和 CL-CD 上其最佳取值分别为：0.1, 0.1, 0.5, 0.1。在确定 γ 取值的基础之上，再将 w_{max} 作为唯一变量，探索预热完成后最大无监督权重的整个模型的影响，确定其最佳参数选择，实验结果表明在 LEVIR-CD、WHU-CD、CDD 和 CL-CD 上其最佳取值分别为：30.0, 30.0, 30.0, 10.0。由此，可以得出基本结论，即训练数据量越小， γ 取值应该越小，否则过拟合的可能性越高；多类别数据集上 γ 取值应该比单一类别（建筑物）变化检测数据集上更大，模型应该充分学习标记样本中各个类别的语义特征； w_{max} 取值应该比较大，因为同一样本不同扰动版本之间的均方误差值数量级较小，给予一个较大的权重才能够使其对模型训练发挥作用。因此，在其余数据集上我们基于此经验，选择的取值组合为分别为：LEVIR-CD+ (0.1, 30.0)；GZ-CD: (0.05, 10.0)；EGY-CD: (0.1, 10.0)；HRCUS-CD: (0.1, 30.0)；DSIFN-CD: (0.5, 30.0)；SYSU-CD: (0.5, 30.0)，对比实验中的所有实验结果即在此参数组合下所取得，实验证明我们的取值经验是有效的，不过在逐个数据集上都进行更细致的调参过程或许会取得更好的实验结果。

表 5-5 超参数敏感性消融实验结果

γ	w_{max}	LEVIR-CD		WHU-CD		CDD		CL-CD	
		IoU^c	OA	IoU^c	OA	IoU^c	OA	IoU^c	OA
0	0 (Sup.Only)	66.8	98.13	55.7	97.53	67.9	95.46	31.4	92.42
0.1	1.0	69.5	98.37	60.1	97.96	69.9	95.85	33.7	93.15
0.5	1.0	68.0	98.24	59.9	97.92	71.4	96.05	31.8	92.68
1.0	1.0	67.0	98.16	57.8	97.84	70.2	95.93	30.6	92.28
0.1	0.5	66.5	98.13	58.4	97.90	-	-	31.7	92.75
0.1	5.0	74.4	98.52	65.6	98.23	-	-	34.9	93.29
0.1	10.0	75.6	98.65	68.1	98.47	-	-	36.1	93.55
0.1	30.0	76.9	98.70	69.3	98.65	-	-	33.5	93.07
0.5	0.5	-	-	-	-	70.1	95.79	-	-
0.5	5.0	-	-	-	-	74.8	96.57	-	-
0.5	10.0	-	-	-	-	77.5	96.90	-	-
0.5	30.0	-	-	-	-	66.2	96.73	-	-

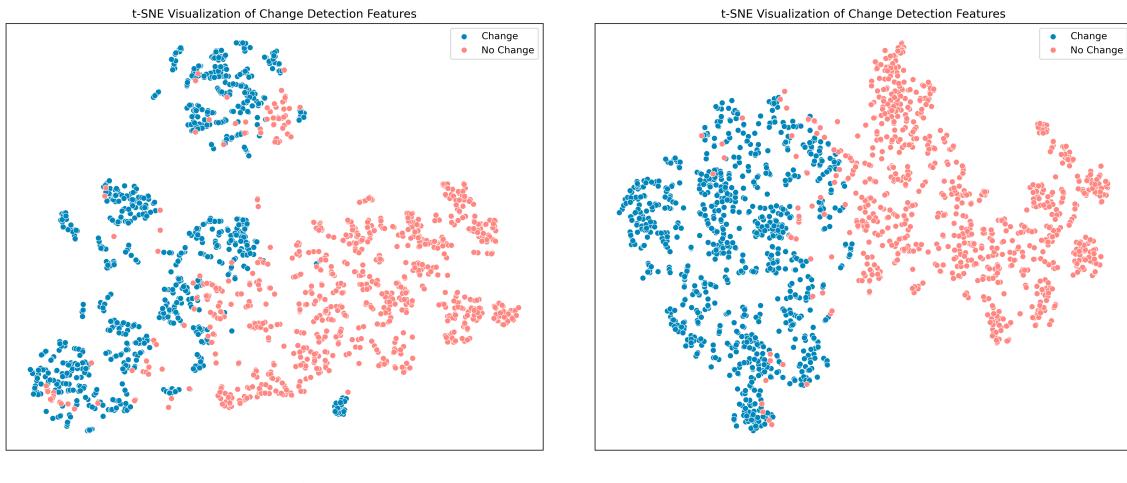


图 5-8 LEVIR-CD 测试集上的 t-SNE 特征可视化对比

最后，我们还使用 t-SNE 技术对分别使用随机特征扰动和自适应特征扰动时测试集上的特征进行了降维并投影到二维坐标进行了可视化，可视化结果如图5-8、图5-9、图5-10以及图5-11分别为 LEVIR-CD、WHU-CD、CDD 和 CL-CD 数据集上的 t-SNE 特征可视化结果。

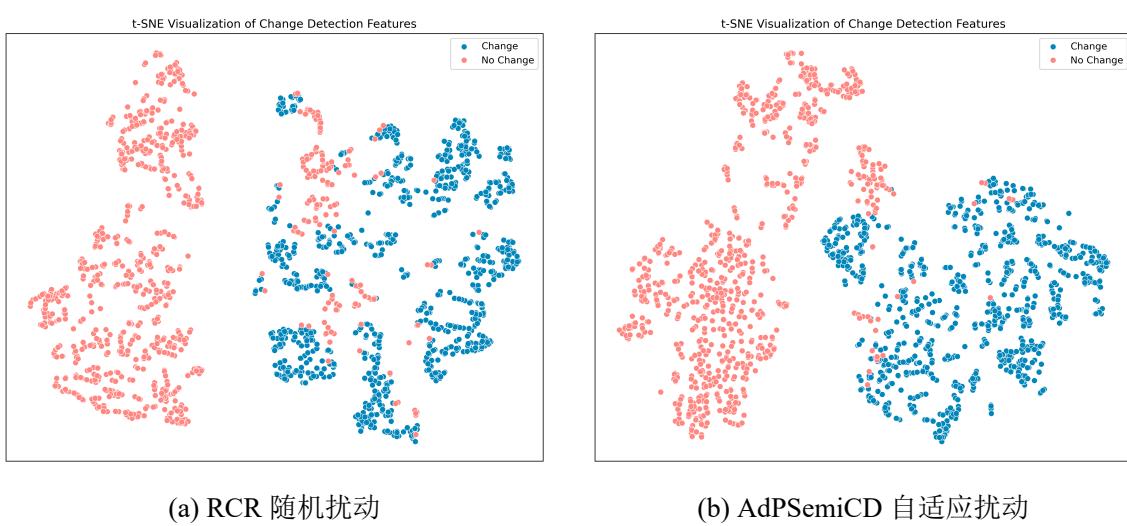


图 5-9 WHU-CD 测试集上的 t-SNE 特征可视化对比

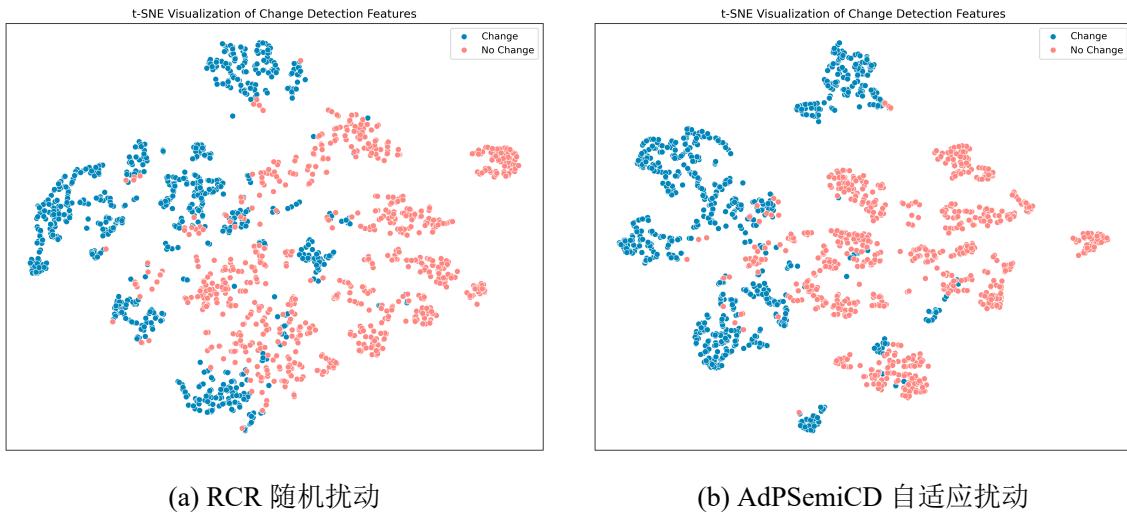


图 5-10 CDD 测试集上的 t-SNE 特征可视化对比

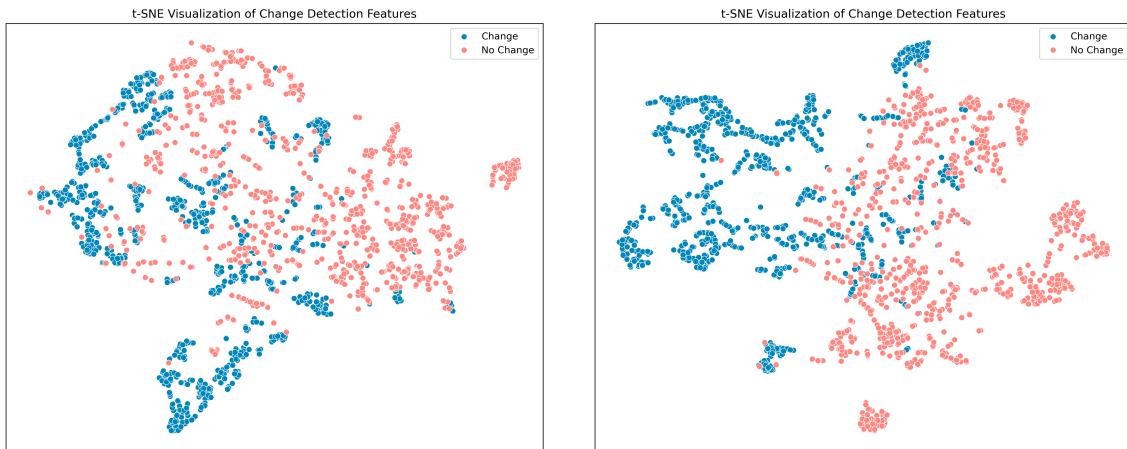


图 5-11 CL-CD 测试集上的 t-SNE 特征可视化对比

相比 RCR 中使用的随机扰动，使用自适应特征扰动的半监督方法，提取的特征空间中决策边界更加明显，变化前景和背景类别的特征在类内更加紧凑，类间更加分散，直观地验证了自适应扰动方法的优越性。

5.4 本章小结

本章提出了基于自适应特征扰动的半监督变化检测算法，首先在引言分析了此前基于一致性正则化和自训练的半监督方法存在的局限，针对伪标签不够准确和随机扰动方法可能加大确认偏差的问题，我们设计了自适应特征扰动的一致性方法。基于对模型在样本上的预测评估来把握模型的学习状态和样本的难易程度，从而对样本进行自适应的特征扰动，包括自适应选择扰动数量和扰动方式的参数自适应缩放。最后展示了我们在十个数据集上的对比实验和消融实验结果，从定量的实验指标和定性的可视化结果，全方位地证明了本章所提出的 AdPSemiCD 的有效性。

第6章 总结与展望

6.1 本文工作总结

变化检测是遥感影像自动解译领域的一个重要研究方向，在许多领域发挥着重要作用，例如，城市建设规划、森林环境保护、农村土地管理、自然灾害评估等。针对变化检测任务的数据标注非常复杂，耗时耗力。而在如今非常成熟的对地观测技术之下，无标注的数据相对来说容易获得。因此本文通过半监督学习的方法将大量的无标注样本利用起来，使得模型仅需少量的人工标注即可达到可观的检测效果，解决了基于深度学习的变化检测方法对标注训练数据的依赖问题，提高了变化检测在实际应用场景中的可行性。

本文首先综述了变化检测任务的发展历程，梳理了国内外关于半监督变化检测的相关研究现状以及半监督学习的主要途径和方法，并对卷积神经网络和此前的最佳半监督变化检测算法进行了详细介绍。最后本文着力研究了基于伪标签自训练和一致性正则化的半监督变化检测，主要从伪标签阈值、伪标签改造以及扰动一致性三个角度设计了半监督变化检测算法。本文的主要研究内容如下：

1) 针对当前基于伪标签和一致性正则化的方法中，采用固定阈值或特定阈值调整策略可能无法充分利用未标记数据的局限性，我们提出了一种基于自适应动态阈值的半监督变化检测算法——AdTSemiCD。该算法通过监测模型预测的置信度动态把握其学习状态，并据此实时调整伪标签的阈值。此外，为了充分挖掘被筛选掉的低置信度标签所包含的潜在信息，我们设计了一个低置信度学习模块，从而进一步提升训练效率。实验结果表明，所提方法减少了在某些稀少样本类别上的漏检问题，表现除了优越的性能。

2) 针对模型生成的伪标签在不同难度无标记样本上存在噪声差异，同时对模型训练的贡献度也不均衡的问题，我们提出了一种基于伪标签评估的自适应半监督变化检测算法——AdaSemiCD。首先，我们设计了 AdaFusion 模块，用于在单样本层面改造高不确定性区域，显著提高伪标签的准确性。其次，我们提出了 AdaEMA 模块，在模型参数更新过程中引入自适应选择机制，使模型能够更有效地整合优化参数。大量实验证明了该方法显著改善了伪标签质量，从而有效提升了变化检测的整体性能。

3) 针对现有一致性正则化方法在对所有样本采取相同扰动策略时可能扩大模型确认偏差的问题，以及基于硬伪标签的自训练方法中不可避免的噪声问题，我们提出了一种基于自适应特征扰动的半监督变化检测算法——AdPSemiCD。该方法通过对无标记样本的预测结果进行评估，针对难易程度不同的样本特征自适应地施加扰动，其中包括对扰动数量和扰动强度的动态选择。通过这种样本级定制的扰动策略，算法在扩展特征空间和减少确认偏差之间达到了平衡，从而为模型训练提供了更有效的一致性约束。最终，实验结果充分验证了该自适应特征扰动策略的有效性和实用性。

6.2 未来研究展望

本文对半监督遥感影像变化检测算法进行了全面、系统的调研，并分别从伪标签筛选、伪标签改造以及特征空间优化三个层面进行了改进。然而，目前还有一些工作值得继续深入探讨研究：

1) 当前基于伪标签自训练的半监督变化检测算法，都依赖于为无标记样本生成的高可信伪标签，伪标签质量直接影响模型的学习效果。但是仅仅依靠仅有少量真实标签训练的模型自身来生成伪标签是存在瓶颈的，除了在伪标签筛选、伪标签改造两个阶段进行一些处理之外，还可以探索在伪标签生成网络层面进行探索，例如引入在大量遥感数据上进行过预训练的大模型来辅助生成伪标签，或许有望突破伪标签质量瓶颈，甚至完全解决半监督变化检测这个问题。

2) 积极探索半监督变化检测与多模态学习的融合。半监督学习可以利用大量的无标记样本训练，天然适合将大量的不同模态（例如合成孔径雷达图像、红外图像等）的无标记样本利用起来，互补性地提供信息，以更好地提高模型泛化性，乃至将光学图像变化检测、合成孔径雷达图像变化检测等细分任务统一起来。

参考文献

- [1] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016: 770-778.
- [2] Dosovitskiy A. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.
- [3] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2881-2890.
- [4] Tarvainen A, Valpola H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results[J]. Advances in neural information processing systems, 2017, 30.
- [5] Peng D, Bruzzone L, Zhang Y, et al. Semicdnet: A semisupervised convolutional neural network for change detection in high resolution remote-sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(7): 5891-5906.
- [6] Bandara W G C, Patel V M. Revisiting consistency regularization for semi-supervised change detection in remote sensing images[J]. arXiv preprint arXiv:2204.08454, 2022.
- [7] Zhang X, Huang X, Li J. Semisupervised change detection with feature-prediction alignment[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-16.
- [8] Bouziani M, Goïta K, He D C. Automatic change detection of buildings in urban environment from very high spatial resolution images using existing geodatabase and prior knowledge[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2010, 65(1): 143-153.
- [9] Zhang X, Yang Y, Ran L, et al. Remote sensing image semantic change detection boosted by semi-supervised contrastive learning of semantic segmentation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-13.
- [10] K S B, N S, Kevala V D, et al. Recent advances in urban expansion monitoring through deep learning-based semantic change detection techniques from satellite imagery[C]. 2024 IEEE Space, Aerospace and Defence Conference (SPACE). 2024: 169-173.
- [11] Hao R, Chen E, Li Z. Forest cover change detection method using bi-temporal gf-1 multi-spectral data[C]. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2016: 3762-3765.
- [12] Xing Y, Zhang Q, Ran L, et al. Progressive modality-alignment for unsupervised heterogeneous change detection[J]. IEEE Transactions on Geoscience and Remote Sensing,

2023, 61: 1-12.

- [13] Xing Y, Zhang Q, Ran L, et al. Improving reliability of heterogeneous change detection by sample synthesis and knowledge transfer[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-11.
- [14] Liang Z, Xia X, Zhou B, et al. Construction and implementation of a data-oriented platform for rural ecological entrepreneurship based on rural remote sensing image change detection algorithm[C]. 2022 International Conference on Inventive Computation Technologies (ICICT). 2022: 1011-1015.
- [15] Lv Z Y, Shi W, Zhang X, et al. Landslide inventory mapping from bitemporal high-resolution remote sensing images using change detection and multiscale segmentation [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(5): 1520-1532.
- [16] Anniballe R, Noto F, Scalia T, et al. Earthquake damage mapping: An overall assessment of ground surveys and vhr image change detection after l'aquila 2009 earthquake [J]. Remote sensing of environment, 2018, 210: 166-178.
- [17] Mondini A, Guzzetti F, Reichenbach P, et al. Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images[J]. Remote sensing of environment, 2011, 115(7): 1743-1757.
- [18] Ludeke A K, Maggio R C, Reid L M. An analysis of anthropogenic deforestation using logistic regression and gis[J]. Journal of Environmental Management, 1990, 31(3): 247-259.
- [19] Du P, Wang X, Chen D, et al. An improved change detection approach using tri-temporal logic-verified change vector analysis[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 161: 278-293.
- [20] Celik T. Unsupervised change detection in satellite images using principal component analysis and k -means clustering[J]. IEEE geoscience and remote sensing letters, 2009, 6 (4): 772-776.
- [21] Marchesi S, Bruzzone L. Ica and kernel ica for change detection in multispectral remote sensing images[C]. 2009 IEEE International Geoscience and Remote Sensing Symposium: volume 2. IEEE, 2009: II-980.
- [22] Nielsen A A, Conradsen K, Simpson J J. Multivariate alteration detection (mad) and maf postprocessing in multispectral, bitemporal image data: New approaches to change detection studies[J]. Remote Sensing of Environment, 1998, 64(1): 1-19.

- [23] Wu C, Du B, Zhang L. Slow feature analysis for change detection in multispectral imagery [J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 52(5): 2858-2874.
- [24] Caye Daudt R, Le Saux B, Boulch A. Fully Convolutional Siamese Networks for Change Detection[C]. 2018 25th IEEE International Conference on Image Processing (ICIP). Athens: IEEE, 2018: 4063-4067.
- [25] Zhang H, Chen H, Zhou C, et al. BiFA: Remote sensing image change detection with bitemporal feature alignment[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-17.
- [26] Chen H, Zhang H, Chen K, et al. Continuous cross-resolution remote sensing image change detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-20.
- [27] Chen H, Qi Z, Shi Z. Remote sensing image change detection with transformers[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-14.
- [28] Bandara W G C, Patel V M. A transformer-based siamese network for change detection [C]. IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium. 2022: 207-210.
- [29] Wu C, Du B, Zhang L. Fully convolutional change detection framework with generative adversarial network for unsupervised, weakly supervised and regional supervised change detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45 (8): 9774-9788.
- [30] Li Z, Tang C, Liu X, et al. MS-Former: Memory-supported transformer for weakly supervised change detection with patch-level annotations[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-13.
- [31] Ma L, Huang Y, Shi W, et al. Weakly supervised building change detection based on deepcut and temporal invariant[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-18.
- [32] Bandara W G C, Patel V M. Deep metric learning for unsupervised remote sensing change detection[J]. arXiv preprint arXiv:2303.09536, 2023.
- [33] Yan L, Yang J, Wang J. Domain knowledge-guided self-supervised change detection for remote sensing images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16: 4167-4179.
- [34] Tang X, Zhang H, Mou L, et al. An unsupervised remote sensing change detection method based on multiscale graph convolutional network and metric learning[J]. IEEE Transac-

- tions on Geoscience and Remote Sensing, 2022, 60: 1-15.
- [35] Chen H, Li W, Shi Z. Adversarial instance augmentation for building change detection in remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-16.
- [36] Wang Z, Liu D, Wang Z, et al. A new remote sensing change detection data augmentation method based on mosaic simulation and haze image simulation[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16: 4579-4590.
- [37] Ren C, Wang X, Gao J, et al. Unsupervised change detection in satellite images with generative adversarial network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(12): 10047-10061.
- [38] Bandara W G C, Nair N G, Patel V M. DDPM-CD: Denoising diffusion probabilistic models as feature extractors for change detection[C]. IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). 2025.
- [39] Ran L, Zhan W, Li Y, et al. DTFSeg: A dynamic threshold filtering method for semi-supervised semantic segmentation[C]. 2023 China Automation Congress (CAC). Chongqing, China: IEEE, 2023: 7571-7576.
- [40] Ran L, Li Y, Liang G, et al. Pseudo labeling methods for semi-supervised semantic segmentation: A review and future perspectives[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2024.
- [41] Cai Z, Ravichandran A, Maji S, et al. Exponential moving average normalization for self-supervised and semi-supervised learning[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 194-203.
- [42] Van Engelen J E, Hoos H H. A survey on semi-supervised learning[J]. Machine learning, 2020, 109(2): 373-440.
- [43] Lee D H, et al. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks[C]. Workshop on challenges in representation learning, ICML: volume 3. Atlanta, 2013: 896.
- [44] Laine S, Aila T. Temporal ensembling for semi-supervised learning[J]. arXiv preprint arXiv:1610.02242, 2016.
- [45] Mi P, Lin J, Zhou Y, et al. Active teacher for semi-supervised object detection[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 14482-14491.
- [46] Mittal S, Tatarchenko M, Brox T. Semi-supervised semantic segmentation with high-and

- low-level consistency[J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 43(4): 1369-1379.
- [47] Yang L, Qi L, Feng L, et al. Revisiting weak-to-strong consistency in semi-supervised semantic segmentation[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 7236-7246.
- [48] Zhao Z, Yang L, Long S, et al. Augmentation matters: A simple-yet-effective approach to semi-supervised semantic segmentation[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 11350-11359.
- [49] Zhao Z, Long S, Pi J, et al. Instance-specific and model-adaptive supervision for semi-supervised semantic segmentation[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 23705-23714.
- [50] Sohn K, Berthelot D, Carlini N, et al. Fixmatch: Simplifying semi-supervised learning with consistency and confidence[J]. Advances in neural information processing systems, 2020, 33: 596-608.
- [51] Li X, Yu L, Chen H, et al. Transformation-consistent self-ensembling model for semisupervised medical image segmentation[J]. IEEE transactions on neural networks and learning systems, 2020, 32(2): 523-534.
- [52] Zhang Y, Cheng Y, Qi Y. SemiSAM: Exploring sam for enhancing semi-supervised medical image segmentation with extremely limited annotations[C]. IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2024.
- [53] Xie Q, Dai Z, Hovy E, et al. Unsupervised data augmentation for consistency training[J]. Advances in neural information processing systems, 2020, 33: 6256-6268.
- [54] Lai X, Tian Z, Jiang L, et al. Semi-supervised semantic segmentation with directional context-aware consistency[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 1205-1214.
- [55] Yang L, Zhuo W, Qi L, et al. St++: Make self-training work better for semi-supervised semantic segmentation[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 4268-4277.
- [56] Wang Y, Wang H, Shen Y, et al. Semi-supervised semantic segmentation using unreliable pseudo-labels[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 4248-4257.
- [57] Blum A, Mitchell T. Combining labeled and unlabeled data with co-training[C]. Proceedings of the eleventh annual conference on Computational learning theory. 1998:

92-100.

- [58] Kingma D P, Mohamed S, Jimenez Rezende D, et al. Semi-supervised learning with deep generative models[J]. Advances in neural information processing systems, 2014, 27.
- [59] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [60] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25.
- [61] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [62] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]. Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. Springer, 2015: 234-241.
- [63] Shi Q, Liu M, Li S, et al. A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection[J]. IEEE transactions on geoscience and remote sensing, 2021, 60: 1-16.
- [64] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]. Proceedings of the European conference on computer vision (ECCV). 2018: 3-19.
- [65] Zhang Q, Lu Y, Shao S, et al. Mfnet: Mutual feature-aware networks for remote sensing change detection[J]. Remote Sensing, 2023, 15(8): 2145.
- [66] Fang S, Li K, Shao J, et al. Snunet-cd: A densely connected siamese network for change detection of vhr images[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 19: 1-5.
- [67] Zhou Z, Rahman Siddiquee M M, Tajbakhsh N, et al. Unet++: A nested u-net architecture for medical image segmentation[C]. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. Springer, 2018: 3-11.
- [68] Zheng Z, Ma A, Zhang L, et al. Change is everywhere: Single-temporal supervised object change detection in remote sensing imagery[C]. Proceedings of the IEEE/CVF international conference on computer vision. 2021: 15193-15202.
- [69] Vaswani A. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017.
- [70] Zaremba W. Recurrent neural network regularization[J]. arXiv preprint arXiv:1409.2329,

2014.

- [71] Li Q, Zhong R, Du X, et al. Transunetcd: A hybrid transformer network for change detection in optical remote-sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-19.
- [72] Xie E, Wang W, Yu Z, et al. Segformer: Simple and efficient design for semantic segmentation with transformers[J]. Advances in neural information processing systems, 2021, 34: 12077-12090.
- [73] Bandara W G C, Patel V M. A transformer-based siamese network for change detection [C]. IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2022: 207-210.
- [74] Jiang B, Wang Z, Wang X, et al. Vct: Visual change transformer for remote sensing image change detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023.
- [75] Jiang M, Chen Y, Dong Z, et al. Multiscale fusion cnn-transformer network for high-resolution remote sensing image change detection[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, 17: 5280-5293.
- [76] Li W, Xue L, Wang X, et al. Mcnet: A multi-scale cnn-transformer network for change detection in optical remote sensing images[C]. 2023 26th International Conference on Information Fusion (FUSION). 2023: 1-5.
- [77] Li W, Xue L, Wang X, et al. Convtransnet: A cnn-transformer network for change detection with multiscale global-local representations[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-15.
- [78] Kirillov A, Mintun E, Ravi N, et al. Segment anything[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 4015-4026.
- [79] Li K, Cao X, Meng D. A new learning paradigm for foundation model-based remote-sensing change detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-12.
- [80] Radford A, Kim J W, Hallacy C, et al. Learning transferable visual models from natural language supervision[C]. International conference on machine learning. PMLR, 2021: 8748-8763.
- [81] Ding L, Zhu K, Peng D, et al. Adapting segment anything model for change detection in vhr remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-11.
- [82] Liu C, Chen K, Zhang H, et al. Change-agent: Towards interactive comprehensive

- change interpretation and analysis from change detection and change captioning[J]. arXiv preprint arXiv:2403.19646, 2024.
- [83] Dong S, Wang L, Du B, et al. Changeclip: Remote sensing change detection with multimodal vision-language representation learning[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2024, 208: 53-69.
- [84] Zheng Z, Zhong Y, Zhang L, et al. Segment any change[J]. arXiv preprint arXiv:2402.01188, 2024.
- [85] Bousias Alexakis E, Armenakis C. Evaluation of semi-supervised learning for cnn-based change detection[J]. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2021, 43: 829-836.
- [86] Mao Z, Tong X, Luo Z. Semi-supervised remote sensing image change detection using mean teacher model for constructing pseudo-labels[C]. ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2023: 1-5.
- [87] Sun C, Wu J, Chen H, et al. Semisanet: A semi-supervised high-resolution remote sensing image change detection model using siamese networks with graph attention[J]. Remote Sensing, 2022, 14(12): 2801.
- [88] Hafner S, Ban Y, Nascetti A. Urban change detection using a dual-task siamese network and semi-supervised learning[C]. IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium. 2022: 1071-1074.
- [89] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[J]. Advances in neural information processing systems, 2014, 27.
- [90] Liu J, Chen K, Xu G, et al. Semi-supervised change detection based on graphs with generative adversarial networks[C]. IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2019: 74-77.
- [91] Nie W, Gou P, Liu Y, et al. Semi supervised change detection method of remote sensing image[C]. 2022 IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). IEEE, 2022: 1013-1019.
- [92] Yang S, Hou S, Zhang Y, et al. Change detection of high-resolution remote sensing image based on semi-supervised segmentation and adversarial learning[C]. IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2022: 1055-1058.
- [93] Li Y, Chen H, Dong S, et al. Multi-temporal samplepair generation for building change

- detection promotion in optical remote sensing domain based on generative adversarial network[J]. *Remote Sensing*, 2023, 15(9): 2470.
- [94] He T, Zhang Z, Zhang H, et al. Bag of tricks for image classification with convolutional neural networks[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019: 558-567.
- [95] Xie S, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1492-1500.
- [96] Shi W, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1874-1883.
- [97] Murphy K P. Machine learning: a probabilistic perspective[M]. MIT press, 2012.
- [98] Miyato T, Maeda S I, Koyama M, et al. Virtual adversarial training: A regularization method for supervised and semi-supervised learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(8): 1979-1993.
- [99] Chen H, Shi Z. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection[J]. *Remote Sensing*, 2020, 12(10): 1662.
- [100] Ji S, Wei S, Lu M. Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set[J]. *IEEE Transactions on geoscience and remote sensing*, 2018, 57(1): 574-586.
- [101] Holail S, Saleh T, Xiao X, et al. Afde-net: Building change detection using attention-based feature differential enhancement for satellite imagery[J]. *IEEE Geoscience and Remote Sensing Letters*, 2023, 20: 1-5.
- [102] Zhang J, Shao Z, Ding Q, et al. Aernet: An attention-guided edge refinement network and a dataset for remote sensing building change detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1-16.
- [103] Lebedev M, Vizilter Y V, Vygolov O, et al. Change detection in remote sensing images using conditional adversarial networks[J]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2018, 42: 565-571.
- [104] Zhang C, Yue P, Tapete D, et al. A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 166: 183-200.
- [105] Shi Q, Liu M, Li S, et al. A deeply supervised attention metric-based network and an

open aerial image dataset for remote sensing change detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-16.

- [106] Liu M, Chai Z, Deng H, et al. A cnn-transformer network with multiscale context aggregation for fine-grained cropland change detection[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 4297-4306.
- [107] Xu Y, Shang L, Ye J, et al. Dash: Semi-supervised learning with dynamic thresholding [C]. International conference on machine learning. PMLR, 2021: 11525-11536.
- [108] Berthelot D, Roelofs R, Sohn K, et al. Adamatch: A unified approach to semi-supervised learning and domain adaptation[J]. arXiv preprint arXiv:2106.04732, 2021.
- [109] Zhang B, Wang Y, Hou W, et al. Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling[J]. Advances in Neural Information Processing Systems, 2021, 34: 18408-18419.
- [110] Guo L Z, Li Y F. Class-imbalanced semi-supervised learning with adaptive thresholding [C]. International conference on machine learning. PMLR, 2022: 8082-8094.
- [111] Cubuk E D, Zoph B, Shlens J, et al. Randaugment: Practical automated data augmentation with a reduced search space[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020: 702-703.
- [112] Vu T H, Jain H, Bucher M, et al. Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 2517-2526.
- [113] Mittal S, Tatarchenko M, Brox T. Semi-supervised semantic segmentation with high-and low-level consistency[J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 43(4): 1369-1379.
- [114] Yun S, Han D, Oh S J, et al. Cutmix: Regularization strategy to train strong classifiers with localizable features[C]. Proceedings of the IEEE/CVF international conference on computer vision. 2019: 6023-6032.
- [115] Zhang H, Cisse M, Dauphin Y N, et al. mixup: Beyond empirical risk minimization[J]. arXiv preprint arXiv:1710.09412, 2017.
- [116] Yuan J, Liu Y, Shen C, et al. A simple baseline for semi-supervised semantic segmentation with strong data augmentation[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 8229-8238.

致 谢

伴随着此文落笔，我的西工大研究生生活即将画上句号。曾经无比期待这一天的到来，但当它真的来临时，心中却充满了复杂的情感。都说读研只有录取和毕业是最开心的时刻，但当我回首这几年，发现一路走来，我收获的不仅是知识，更是无数值得铭记的瞬间。还记得，犹记得刚来到这里的第一天，无比迷茫的我在夜晚骑车一圈圈地环绕校园，借着路灯的微光，仿佛找到了前行的方向。但其实，那灯光不过是心灵的寄托，帮助我走出自己的误区，与自己和解；真正指引我不断前行的，是我的导师、同门、朋友和家人。他们的陪伴和支持让我克服困惑，跨越难关，沿着正确的道路坚定向前。对此，我怀着深深的感激之情，铭记于心。

我要感谢我的导师冉令燕老师。冉老师和蔼可亲，虽为师生，却更像朋友，彼此尊重、互相关怀。冉老师时常关心我们，流行病期间提醒大家注意健康，逢年过节组织组内聚餐，营造了温馨和谐的实验室氛围。科研上，冉老师治学严谨，责任感强，耐心指导我们读文献、理思路、找创新、做实验和写文章，可以说是我心中理想型导师的典范。感谢您三年来的指导，助我顺利完成学业，也让我度过了一段多年后回想起来仍觉美好的时光。

我要感谢我的同门、师兄师姐、师弟师妹们，在读研期间两点一线的生活中，和大家在实验室的相处时光最多，大家一起在科研上互帮互助、共同进步，以及闲暇时光一起约饭、打羽毛球的这些经历非常值得留恋，感恩遇见你们，为我本来平常的生活增添了色彩，也真诚地祝愿大家都将拥有锦绣前程。

我要感谢我的室友和好朋友们，感谢你们的陪伴，一个人固然可以埋头走得很快，但一群人才能够相互扶持走得更远，感谢你们在我生病期间的关心和照顾，以及在很多精神内耗的时间段，愿意成为我倾诉的对象，帮助我走出阴霾。也在此祝福大家此去一路顺利，多年之后归来仍是少年。

最后，我还要感谢我的家人，感谢你们无论是经济上还是情感上的坚定支持，读研这一路走来没有你们的默默支持我是无法心无旁骛、集中精力完成学业的，是你们给了我前进的底气和勇气。我会继续努力，在未来的工作中脚踏实地，用自己的能力回报你们。

同时，我也要感谢在繁忙的科研工作中抽出宝贵时间对我的论文进行评审的专家教授和参加论文答辩的各位老师，衷心地感谢你们在评审和答辩过程中提出的宝贵意见和建议！

“道路是曲折的，前途是光明的”。行文至此，我想以这一句作为结尾，也作为未来生活的勉励。梦虽遥，追则能达；愿虽艰，持则可圆；不怕困难，迎难而上。也祝福我们的党和国家繁荣昌盛，河山添锦绣，星光映万家，早日实现中华民族的伟大复兴。

在学期间发表的学术成果和参加科研情况

I 攻读硕士学位期间发表论文情况

- [1] Ran, Lingyan, Dongcheng Wen, Zhuo Tao, Shizhou Zhang, Xiuwei Zhang and Yanning Zhang. "AdaSemiCD: An Adaptive Semi-Supervised Change Detection Method Based on Pseudo-Label Evaluation." IEEE Transactions on Geoscience and Remote Sensing (TGRS), minor revision.

II 攻读硕士学位期间参与项目情况

- [1] 参与科研项目. 数据判读只能支撑平台算法库开发

III 攻读硕士学位期间获奖情况

[1] 获 2022-2023 学年、2023-2024 学年、2024-2025 学年西北工业大学二等奖学金

[2] 首届重庆市人工智能创新大赛遥感赛道——二等奖

[3] 陕西省语音与图像处理重点实验室 2023 年度——服务贡献奖

西北工业大学
学位论文知识产权声明书

本人完全了解学校有关保护知识产权的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属于西北工业大学。学校有权保留并向国家有关部门或机构送交论文的复印件和电子版。本人允许论文被查阅和借阅。学校可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。同时本人保证，毕业后结合学位论文研究课题再撰写的文章一律注明作者单位为西北工业大学。

本学位论文属于（在下方框内打“√”）：

- 保密论文，保密期（ 年 月 日至 年 月 日）。
- 公开论文。

学位论文作者签名：_____

指导教师签名：_____

年 月 日

年 月 日

西北工业大学

学位论文原创性声明

秉承学校严谨的学风和优良的科学道德，本人郑重声明：所呈交的学位论文，是本人在导师的指导下进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容和致谢的地方外，本论文不包含任何其他个人或集体已经公开发表或撰写过的研究成果，不包含本人或其他已申请学位或其他用途使用过的成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式表明。

本人学位论文与资料若有不实，愿意承担一切相关的法律责任。

学位论文作者签名：_____

年 月 日