

柴文光, 罗崇熙. 半监督遥感图像建筑物变化检测算法[J]. 广东工业大学学报, 2025, 42(3): 36–43. doi: 10.12052/gdutxb.240045.
Chai Wenguang, Luo Chongxi. Semisupervised remote sensing image building change detection algorithm[J]. Journal of Guangdong University of Technology, 2025, 42(3): 36–43. doi: 10.12052/gdutxb.240045.

半监督遥感图像建筑物变化检测算法

柴文光, 罗崇熙

(广东工业大学 计算机学院, 广东 广州 510006)

摘要: 建筑物的变化检测在遥感图像处理和模式识别领域具有重要意义, 但在深度学习算法的应用中, 数据标注一直是一个显著的挑战, 特别是在变化检测的场景下。因此, 针对基于深度学习的变化检测算法中数据标注的难题, 本文提出了一种半监督学习方法。该方法采用了融合双时相特征的孪生网络来进行特征提取, 并构建了一个教师-学生网络框架以实施模型的半监督训练。为了进一步提升半监督变化检测的准确度, 本文在深度特征上引入了随机扰动, 以此来实现一致性正则化。此外, 在图像深度特征的层面上, 本文还提出通过捕获双时相图像特征的差异来形成决策边界以区分双时相图像变化的方法。该方法在Levir-CD和WHU_Building两个公开数据集上分别实现了83.04%和85.57%的交并比(Intersection over Union, IoU)。实验结果表明, 本文提出的方法在使用少量标注数据的前提下, 能够达到与全监督训练相近的性能。

关键词: 遥感图像; 变化检测; 一致性正则化; 半监督学习

中图分类号: TP391.4; TP751.1

文献标志码: A

文章编号: 1007-7162(2025)03-0036-08

Semisupervised Remote Sensing Image Building Change Detection Algorithm

Chai Wenguang, Luo Chongxi

(School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: Change detection in buildings holds significant importance in the fields of remote sensing image processing and pattern recognition. However, data annotation has always been a prominent challenge in the application of deep learning algorithms, especially in change detection scenarios. To address the data annotation challenges in deep learning-based change detection algorithms, this paper proposes an innovative semi-supervised learning method. This method employs a Siamese network that fuses bi-temporal features for feature extraction and constructs a teacher-student network framework for semi-supervised model training. To further enhance the accuracy of semi-supervised change detection, this paper introduces random perturbations in deep features to achieve consistency regularization. Additionally, on the level of image deep features, the paper proposes a method for forming decision boundaries by capturing differences in bi-temporal image features to distinguish changes in bi-temporal images. This method achieved Intersection over Union (IoU) scores of 83.04% and 85.57% on the Levir-CD and WHU Building datasets, respectively. Experimental results show that the proposed method can achieve performance levels close to fully supervised training with a limited amount of labeled data.

Key words: remote sensing images; change detection; consistency regularization; semi-supervised learning

在遥感图像处理和模式识别领域, 建筑物变化检测扮演着至关重要的角色。它在城市规划、灾害评估和环境监测等多个领域中有着广泛的应用。近年来, 深度学习技术在图像分类、目标检测和语义分割

等任务上取得了显著进展, 并被引入到变化检测中。自全卷积网络的方法被提出, 编码器-解码器成为图像预测任务的常用架构, 变化检测也不例外, 但由于变化检测的核心在于识别和理解两个时刻之间发生

收稿日期: 2024-03-26 录用日期: 2024-04-19 网络首发日期: 2024-09-25

基金项目: 国家自然科学基金资助项目(61772143); 广东省重点领域研发计划项目(2021B0101220006)

作者简介: 柴文光(1969-), 男, 副教授, 博士, 主要研究方向为人工智能与应用技术、计算机系统结构及高性能计算, E-mail: chaiwg@gdut.edu.cn

的变化,传统的单编码器-单解码器并不能满足任务需求,而这时孪生网络因其处理成对数据的优势而广泛应用于图像比对任务。在变化检测中,孪生网络通过比较不同时期的双时相图像,有效识别变化与不变区域的特征差异,因此孪生网络成为了变化检测的主流方法,例如深度监督图像融合网络^[1]、全卷积孪生网络^[2]与特征约束网络^[3],这些方法均采用了共享权重的双编码器来作为孪生网络。张枫幸等^[4]提出的D-WNet,他们将输入的双时相遥感图像送入双编解码器,对每个图像中的目标物体进行分割,然后在单个解码器中融合双时特征以检测变化的对象。随着遥感技术的快速发展,获取的图像数据量呈现爆炸式增长,传统的手动变化检测方法由于耗时且劳动力密集,已经无法满足当前的需求,在变化检测领域,获取大规模的、精准的标注数据是一个重大挑战,因为标注过程不但需要大量的人力物力,同时也受到专业知识的限制。为了解决数据标注问题,半监督学习方法应运而生。这类方法只需要少量标注样本,便可以利用大量未标注数据进行有效学习。例如基于教师-学生训练框架的MeanTeacher^[5]方法,其提出平均指数移动(Exponential Moving Average, EMA)将学生网络的参数通过加权衰减的方式赋予教师网络,从而提高模型的鲁棒性和泛化能力,并通过比较这2个网络产生的预测结果之间的差异,来计算一致性损失。Fixmatch^[6]、Unimatch^[7]等半监督方法通过数据增强手段,在输入图像的图像域进行不同程度的强弱数据增强来将输入差异化,以此来计算一致性损失。以上基于一致性正则化的半监督方法在变化检测任务上并不奏效,原因在于涉及双时相图像的聚类假设并不适用于图像域。聚类假设是指,

在图像处理中,具有相似特征的像素点通常会在图像域内自然地聚集成群体或簇。然而,在变化检测的背景下,这种假设更适合应用于深度特征层面而非图像域。这一观点得到了遥感图像变化检测中RSCD^[8]方法的验证。

遵循先前工作的思路,本文采用孪生网络和教师-学生框架来实现半监督训练,保留教师网络在图像域上进行的强数据增强与弱数据增强。并通过在深度特征上引入随机扰动计算一致性损失来解决双时相图像聚类假设在图像域不成立的问题。最后,为了更好地捕获双时相图像特征的差异特征,提出了通过在特征空间中计算双时相图之间的特征距离来量化两幅图像在语义层面的差异程度的方法。这个特征距离不仅提供了一种量化两幅图像差异的手段,而且可以作为一种有效的决策边界,成功地帮助模型准确判断两幅图像之间的相似或差异。

1 半监督变化检测基本原理

1.1 变化检测网络模型结构

本文采用孪生网络来提取变化特征,其以SegFormer^[9]作为网络的骨干,共享训练参数。解码器使用了FPN^[10],一种用于特征聚合的多尺度空间层次特征的深度学习网络。本文定义输入的双时相图为 $T_1 = \{t_1(i)\}_{i=1}^D$ 和 $T_2 = \{t_2(i)\}_{i=1}^D$,其中 T_1 和 T_2 表示双时相遥感图像集, D 为样本集合。整体结构如图1所示,其中Stage_{*i*}表示多层特征编码的第*i*个阶段,其中 $i = \{1, 2, 3, 4\}$,每个阶段由PatchEmbed、SegFormer以及LayerNorm组成,2个解码器共享参数。这样做的好处是共享参数使得网络更容易泛化到未见过的数据。

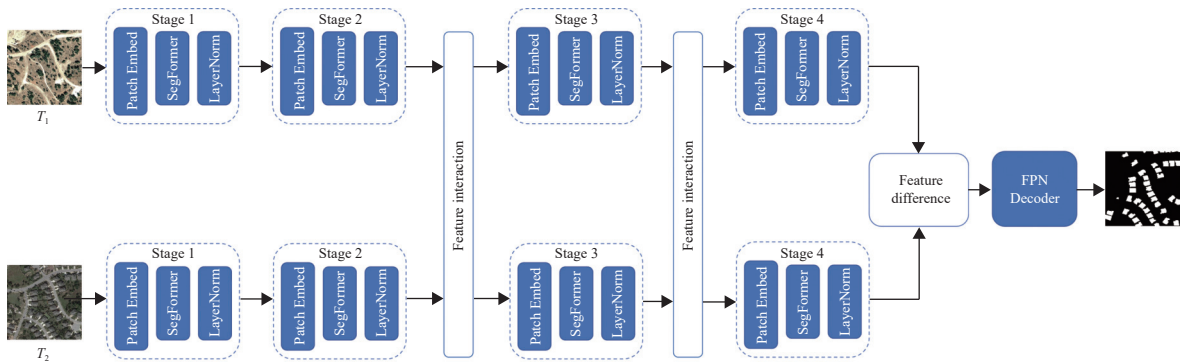


图1 整体网络框架和特征交互模块

Fig.1 Overall network framework and feature interaction module

多层编码器的不同层次可以捕捉不同的特征。通常情况下,较浅的层次更容易捕捉到输入数据的

局部结构和空间信息,而较深的层次则更倾向于捕捉更抽象的、高级的语义信息。孪生网络可以通过特

征交换来相互学习语义信息,所以本文在前两个阶段单独学习各自的样本特征,然后在中层进行特征交互(Feature Interaction),在最后一层融合特征差异(Feature Difference)。第1次特征交互是在双时相特征编码的空间维度上进行,如图2(a)所示,此时特征编码蕴含着足够多的空间信息。

第2次特征交互采取通道交换,在通道维度上进行交换,如图2(b)所示。前两次特征交换有2个好处:(1)通过特征交换可以使2个分支特征之间的分布更加相似,一定程度上实现了双时域间的域自适应;(2)这种操作可以被看作是在深度特征上进行的一种扰动,这样的扰动可以在一定程度上引入随机性,可以提高模型的泛化能力。在最后一次特征交互中,通过计算特征之间的差异(Feature difference)来确定决策边界。这里,较远的距离表示特征差异大,可能意味着数据点属于不同类别;而较近的距离表示特征差异小,可能意味着数据点属于同一类别。具体实现方式类似自注意力机制,将双时相特征都重塑成

维度为 $HW \times C$ 的语义向量 Q 和 K ,之后将2个向量点乘得到相似度矩阵(Similarity Matrix)。如图2(c)所示,最后经过sigmoid函数后与原向量 V 相乘得到聚合了双时相图相似度特征的向量Similarity,如式(1)所示。相似度矩阵取反可以得到一个蕴含了双时相图差异的差异矩阵,与原向量 V 相乘后可以得到差异向量Dissimilarity,如式(2)所示。

$$\text{Similarity} = \text{sigmoid} \left(\frac{QK^T}{\sqrt{\text{dim}}} \right) V \quad (1)$$

$$\text{Dissimilar} = \left(1 - \text{sigmoid} \left(\frac{QK^T}{\sqrt{\text{dim}}} \right) \right) V \quad (2)$$

在式(1)与式(2)中, dim代表维度,它用于对相似度矩阵进行归一化处理。Sigmoid函数能够将向量中的每个元素单独映射到(0, 1)区间。本文模型利用了这一机制引入平滑的决策边界,从而在特征域中识别出差异。对这些差异向量进行特征解码,便可获得最终的变化预测。

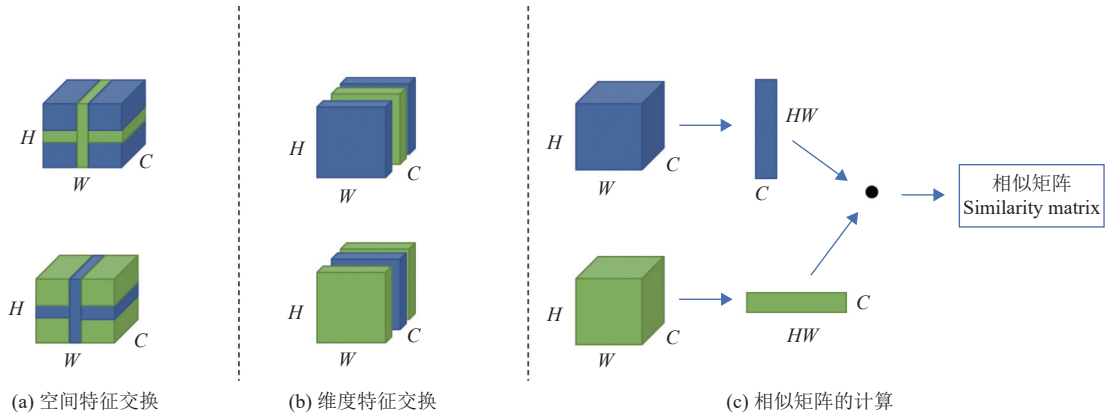


图2 3个不同的特征交互模块

Fig.2 Three different feature interaction modules

1.2 半监督学习

特征交换操作没有产生任何需要学习的参数,这样既能减少模型的参数量,又能增加泛化性,这样的网络很合适使用在半监督学习上。如图3所示,本文的半监督学习框架使用教师-学生训练架构,学生网络进行有监督学习,对双时相图进行弱数据增强后输入到模型进行训练。教师网络不进行学习,其同时进行2个预测,一个是对经过弱数据增强的双时相图进行预测,另一个为对经过了强数据增强的双时相图的预测。教师-学生两个网络并行训练,教师网络以EMA的方式更新网络的参数。学生网络的损失函数由4部分组成,分别为有标签数据的二元交叉熵损失、无标签数据强数据增强和弱数据增强计算

得到的一致性损失、在深度特征层中进行随机扰动的一致性损失和针对变化检测任务的差异正则化。总损失函数表示为

$$L = L_{ce} + L_{cr} + L_{scr} + L_{fip} \quad (3)$$

式中: L_{ce} 为学生网络部分的有监督损失; L_{cr} 为传统半监督学习中,由强弱两种不同数据增强方式得到的最终预测的一致性损失; L_{scr} 为相似一致性正则化; L_{fip} 为在深层特征中进行随机扰动的一致性损失。整个训练过程可以表示为

$$Y_1 = \sum_{i=1}^N F_s(\text{Aug}_w(t_1(i), t_2(i))) \quad (4)$$

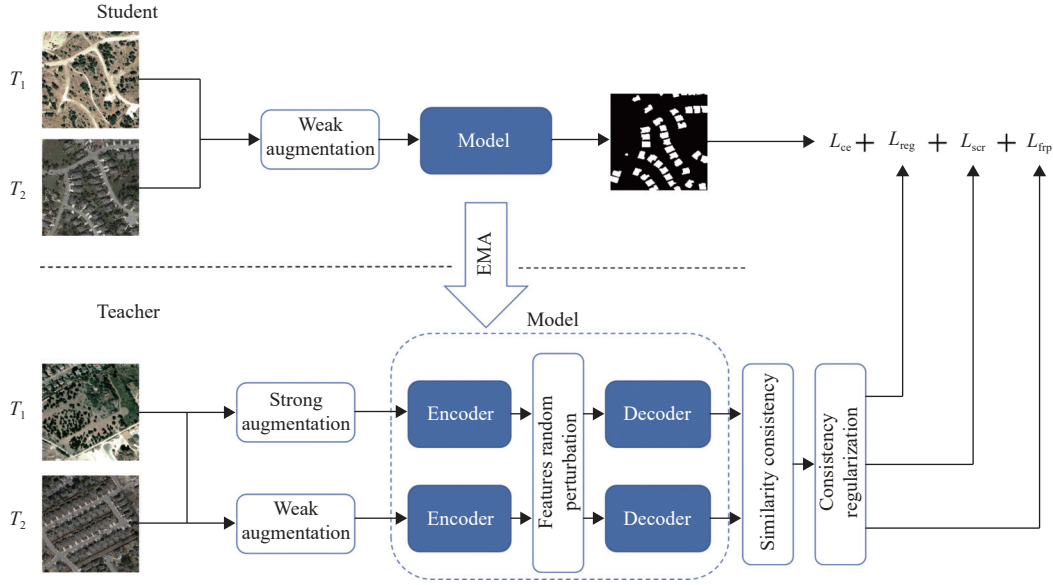


图3 教师-学生训练框架

Fig.3 The teacher-student training framework.

$$Y_u^w = \sum_{i=1}^N F_t(\text{Aug}_w(t_1(i), t_2(i))) \quad (5)$$

$$Y_u^s = \sum_{i=1}^N F_t(\text{Aug}_s(t_1(i), t_2(i))) \quad (6)$$

式中: F_s 和 F_t 为学生和教师的网络模型, N 为输入图像中的像素数量, Aug_w 和 Aug_s 表示弱数据增强和强数据增强, 弱增强是指随机裁剪、翻转和旋转等不改变像素内部信息的增强方式, 强增强是指在像素内部信息中添加干扰的增强方式, 具体增强方式如表1所示。 Y_t 为有标签数据的预测值, 教师网络同时对弱增强和强增强数据进行预测得到2个输出 Y_u^w 和 Y_u^s 。有监督损失函数 L_{ce} 如式(7)所示。

表1 强增强与弱增强

Table 1 Strong augmentation and weak augmentation

数据增强	方式
强增强	颜色空间畸变
	直方图均衡化
	图像锐度增强
	色调分离
弱增强	图像反色
	随机裁剪
	随机翻转
	随机旋转

$$L_{ce} = \frac{1}{|D_t|} \sum_{i=1}^{|D_t|} \frac{1}{H \times W} \sum_{j=1}^{H \times W} B(y_i(j), \hat{y}_i(j)) \quad (7)$$

式中: 有监督网络使用二元交叉熵 B 计算预测损失,

其中 $Y = \{y_i\}_{i=1}^{|D_t|}$ 为样本标签, $\hat{Y} = \{\hat{y}_i\}_{i=1}^{|D_t|}$ 为有监督网络的预测。 D 为总样本数据, 用 D_t 表示有标签数据, D_u 表示无标签的数据。 H, W 为预测图像的高度和宽度。一致性损失同样使用二元交叉熵计算, 以 $Y^{frp} = \{y_i^{frp}\}_{i=1}^{|D_u|}$ 表示经历了深层特征随机扰动的预测, 本文所使用的特征扰动实现方式为对经历了弱增强的特征进行参数为0.5的Dropout, 并与随机高斯噪声相加。使用弱增强的预测值作为伪标签 $P = \{p_i\}_{i=1}^{|D_u|}$, 利用置信度作为伪标签的掩码, I 为阈值函数, 仅有置信度大于 τ 的值计入损失计算, 实验中 τ 设为0.95, 由数据增强带来的一致性损失 L_{cr} 与由深层特征带来的一致性损失 L_{frp} 如式(8)与式(9)所示。

$$L_{cr} = \frac{1}{|D_u|} \sum_{i=1}^{|D_u|} \frac{1}{H \times W} \sum_{j=1}^{H \times W} I(\max(\hat{y}_i(j) > \tau)) B(p_i(j), \hat{y}_i(j)) \quad (8)$$

$$L_{frp} = \frac{1}{|D_u|} \sum_{i=1}^{|D_u|} \frac{1}{H \times W} \sum_{j=1}^{H \times W} I(\max(\hat{y}_i^{frp}(j) > \tau)) B(p_i(j), \hat{y}_i^{frp}(j)) \quad (9)$$

变化检测任务需要同时输入2张图片进行预测, 而在半监督训练的变化检测任务中, 模型内部的深层特征中双时相图之间的特征距离是相对固定的。因此, 无论双时相图在图像域经历了何种数据增强, 或者在深度特征层经历了何种随机扰动, 双时相特征向量之间的差异都应该保持一致性。过程如式(10)~(12)所示。

$$S_w = \|E(\text{Aug}_w(T_1)) - E(\text{Aug}_w(T_2))\| \quad (10)$$

$$S_s = \|E(\text{Aug}_s(T_1)) - E(\text{Aug}_s(T_2))\| \quad (11)$$

$$L_{\text{scr}} = \frac{1}{|D_u|} \sum_{i=1}^{|D_u|} \frac{1}{H \times W} \sum_{j=1}^{H \times W} (S_w(j) - S_s(j))^2 \quad (12)$$

式中： S_w 和 S_s 为分别经历了弱增强与强增强的相似度， E 表示模型的编码器，使用MSE损失计算相似度一致性损失 L_{scr} 。教师网络使用MEA策略进行参数更新，具体来说是在每次迭代完毕时，学生网络的权重进行衰减后赋予教师网络，其基本的原理是在训练过程中对学生网络的参数进行平滑处理后赋予教师网络，这样教师网络可以更好地捕捉数据分布的特征，从而拥有良好的泛化性，更好地学习没见过的数据。EMA的计算过程为

$$\theta_t = \alpha \theta_{t-1} + (1 - \alpha) \hat{\theta}_t \quad (13)$$

式中： θ_t 为 t 时刻的参数， α 为平滑因子， $\hat{\theta}_t$ 为 t 时刻学生网络的参数。

2 测量实验与结果

2.1 数据集介绍

本文实验采用2个开源数据集Levir-CD与WHU_Building。Levir-CD由637对分辨率非常高图像块组成，每个图像块的大小为1 024×1 024像素，这些图像具有5~14 a的时间跨度，并且经历了显著的土地变化，其涵盖了各种类型的建筑，如别墅住宅、高层公寓、小车库和大型仓库。本文专注于与建筑物变化相关的部分，这包括建筑增长(即从土地、草地、硬化地面或正在建设的建筑物变化到新建区域的过程)和建筑物的减少。这些双时相图像采用遥感图像，并通过二进制标签进行标注，其中变化标为1，未变化则标为0。WHU_Building变化检测数据集是武汉大学在2018年发布的开源数据集，该数据集包含了2张高分辨率的遥感图像，分别在2012年和2016年拍摄于同一地区。这些图像特别注明了不同年份的建筑物数量，其中2012年的图像中标注了12 796个不同类型的建筑物，而2016年的图像中则标注了16 077个不同类型的建筑物。在本文的实验中，将利用Levir-CD和WHU_Building这两个数据集，探讨本文模型及半监督训练方法的效果和性能。

2.2 数据增强方式

在半监督学习中，为计算图像域的一致性损失，使用了2种数据增强方法：强数据增强和弱数据增强。强数据增强通过剧烈变换生成差异大的图像，提升模型泛化能力；而弱数据增强则通过较小的变化，

如轻微旋转或亮度调整，增强模型对不同输入的适应性。具体增强方法见表1。

2.3 实验

本文网络模型使用SegFormer^[9]作为backbone，FPN作为解码器。在训练过程中，采取AdamW^[11]作为优化器，权重衰减设置为0.05，初始值设为0.001。使用NVIDIA GeForce RTX 3 090 GPU进行训练，批量大小设置为8，对Levir-CD与WHU_Building数据集均进行了40 000次迭代训练。半监督学习策略中，教师网络的参数更新使用平均指数移动EMA，其平滑因子设置为0.01，有标签数据占比分别为5%，10%，20%，40%。在该实验环境与配置中，本文模型单张图片(分辨率512×512像素)的推理速度达到了14.20 ms。

2.4 评价指标

网络模型性能评估使用语义交并比(Intersection over Union, IoU)以及 F_1 分数作为变化检测精度的指标。这2个指标是常用于语义分割和变化检测任务的度量标准，用于评估模型对图像中变化区域的准确性。IoU通过计算预测的变化区域与实际变化区域之间的交集与并之比来衡量模型性能。其计算公式为

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (14)$$

式中：TP(True Positives)为真正例，FP(False Positives)为假正例，FN(False Negatives)为假负例，TN(True Negatives)为真负例。 F_1 分数计算公式为

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (15)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (16)$$

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (17)$$

2.5 消融实验

本章节展示各个消融实验的结果以证明本文模型是可行的。表2对3种不同的一致性正则化进行消融，分别为图像域上进行强弱数据增强的一致性损失(Consistency Regularization, CR)、在深层特征上进行随机扰动得到的一致性损失(Features Random Perturbation, FRP)和利用变化检测双时相图像之间相似度一致性计算的损失(Similarity Consistency Regularization, SCR)。

从表2中的实验结果可以看出，由于聚类假设在变化检测中的图像域上不成立，网络无法从双时相

图中直接进行分类,故而仅计算CR的结果的性能指标会低于加入了FRP的性能指标。并且针对变化检测任务的相似度一致性也是成立的。

表2 在一致性正则化方法上的消融实验
Table 2 Ablation experiments on the consistent regularization method %

指标	IoU	F_1
CR	80.65	89.29
CR+FRP	82.55	90.44
CR+FRP+SCR	83.04	90.73

本文对不同阶段中进行不同方式的特征交互所得到的结果进行了消融实验。设计了3种不同的特征交互方式搭配,第1种为Changer^[12]的空间交换(Spatial Exchange, SE)与通道交换(Channel Exchange, CE)搭配,第2种将最后一阶段的特征交互换为计算特征差异(Feature Difference, FD),第3种将无可学习参数的FD替换为可学习的LFD(Learnable Feature

Difference),具体做法类似自注意力机制,对特征编码使用线性层得到可学习的权重矩阵,得到自注意力矩阵后取反得到差异注意力。其实验结果如表3所示,拥有可学习参数的LFD导致了性能下降,可见模型关注特征向量之间的距离,而非提取某种差异特征。

表3 在特征交互层上消融实验
Table 3 Ablation experiments on the differential attention method %

指标	IoU	F_1
SE+CE+CE	81.38	89.73
SE+CE+FD	83.04	90.73
SE+CE+LFD	81.67	89.30

为研究有标签数据在总样本数量上的占比对模型性能的影响,在Levir-CD和WHU Building两个不同数据集上取占比为5%,10%,20%,40%的有标签样本的训练数据进行消融实验,并加入其他方法进行对比,其结果如表4所示。

表4 在含有不同占比有标签数据的数据集上进行消融实验
Table 4 Ablation experiments were carried out on the proportion of labeled data in different datasets %

方法	Levir-CD IoU				WHU Building IoU			
	5	10	20	40	5	10	20	40
s4gan ^[13]	64.00	67.00	73.40	75.40	18.30	62.60	70.80	76.40
semiCD ^[8]	72.50	75.50	76.20	77.20	65.80	68.10	74.80	77.20
Unimatch ^[7]	75.60	79.00	79.00	78.20	77.50	78.90	82.90	82.40
ours	71.37	79.76	80.15	83.04	78.64	83.02	83.15	85.57

本文还设计了跨数据集训练的实验以证明所提出方法的高稳定性和高泛化性。如表5所示,本文使用Levir-CD作为有标签的数据,而WHU-Building作为无标签数据,和用WHU-Building作为有标签数据,而Levir-CD作为无标签数据分别进行了2次实验,从实验结果可以看出,本文模型的泛化性均高于以往方法,在一定程度上依然能有效工作。

表5 跨数据集进行半监督学习IoU性能指标
Table 5 IoU performance metrics for semi-supervised learning across datasets %

方法	有标签Levir-CD数据集训练无标签的WHU Building数据集			有标签WHU Building数据集训练无标签的Levir-CD数据集		
	5	10	20	5	10	20
s4gan ^[13]	31.20	31.70	41.90	1.60	7.80	8.70
SemiCDNet ^[14]	28.40	32.50	39.50	3.40	8.90	9.70
semiCD ^[8]	35.40	39.30	42.10	7.80	11.70	10.60
ours	45.56	49.57	46.23	28.97	25.69	27.24

为了更好地评估本文所提出方法的性能,在保证实验条件和参数相似的情况下与近年来主流的方

法进行对比,如表6所示。此次实验同时对比了有监督学习与半监督学习在2个数据集上的表现,若无特殊说明,半监督有标签数据占比均为20%,可以看出,本文所提出方法在交并比IoU与 F_1 指标上均超过所列出的其他半监督方法,并且在有监督方法中也有一定的竞争力。

2.6 结果可视化

本文利用结果可视化在Levir-CD数据集上与有监督方法Changer模型与半监督方法Unimatch进行了对比。如图4所示,第1列为双时相遥感变化图的变化前(befor)图像,第2列为变化后(after)图像,后4列分别为标签(label)、Changer模型、Unimatch方法,以及本文方法(Ours)所得到的结果可视化,可以看出对比其他半监督方法,本文方法得到的结果更接近有监督模型的预测结果。

3 结论

本文提出了一种半监督学习方法,其在建筑物

表 6 各个网络在数据集Levir-CD与WHU Building上的性能指标
Table 6 Performance metrics of each network on the datasets Levir-CD and WHU Building

		Levir-CD数据集			WHU Building数据集				
		Method	IoU/%	F_1 /%	#Param/M	Method	IoU/%	F_1 /%	#Param/M
有监督	Changer ^[12]	82.64	90.49	11.39		SNUNet ^[15]	78.83	88.16	27.06
	SNUNet-ECAM ^[15]	66.81	80.10	12.03		SNUNet-ECAM	44.67	28.75	12.03
	tinycd ^[16]	83.57	91.05	0.28		tinycd ^[16]	84.30	91.48	0.28
	changeformer ^[17]	82.48	90.40	20.75		BIT ^[18]	72.39	83.98	3.55
半监督	semiCDNet ^[14]	69.22	81.81			semiCDNet ^[14]	43.12	60.25	
	s4GAN ^[13]	67.72	80.76			s4GAN ^[13]	42.51	26.99	
	semiSANet ^[19]	76.98	86.99			semiSANet ^[19]	78.34	87.86	
	SemiBuildingChange ^[20]	71.46	83.36	17.59		SemiBuildingChange ^[20]	72.70	84.19	17.59
	ours (20%)	80.15	88.98	12.34		ours (20%)	83.15	90.20	12.34
	ours (40%)	83.04	90.73	12.34		ours (40%)	85.57	92.22	12.34

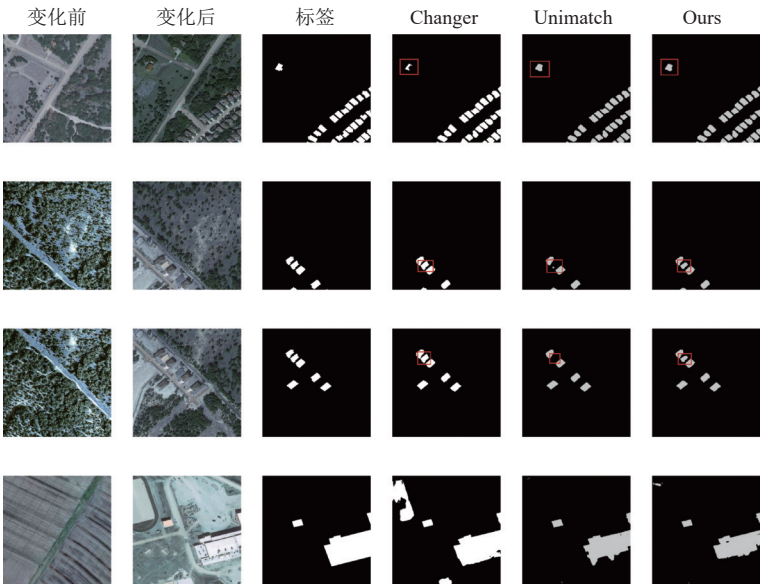


图 4 预测结果可视化

Fig.4 Visualization of prediction results.

变化检测领域得到了较好的实验结果。本文使用孪生网络构建了一个教师-学生网络框架,进行模型的半监督训练。为了提高模型的性能,提出了在编码器网络的深度特征层中引入随机扰动来计算一致性正则化的方法,并且通过捕获双时相图的语义特征差异来有效区分图像中的变化。在Levir-CD和WHU_Building数据集上获得了83.04%和85.57%的交并比(IoU),该结果证明此方法即使在标注数据有限的情况下,也能达到与有监督训练相近的性能水平。实验结果不仅超越了近年来的许多半监督方法,而且在与有监督方法的比较中也显示出较好的性能。实验验证了本文方法的有效性和鲁棒性,为解决变化检测任务中的数据标注难题提供了新的解决途径。

参考文献:

[1] 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.

[2] RODRIGO C D, BERTRAND L S, ALEXANDREI B. Fully convolutional siamese networks for change detection[C]// 2018 25th IEEE International Conference on Image Processing (ICIP). Athens: IEEE, 2018: 4063-4067.

[3] CHEN P, ZHANG B, HONG D, *et al.* FCCDN: Feature constraint network for VHR image change detection [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 187: 101-119.

[4] 张枫幸, 黄健, 李浩. 轻量级双侧输入D-WNet航空影像建筑物变化检测[J]. 激光与光电子学进展, 2024, 61(8): 3.

- ZHANG F X, HUANG J, LI H. Light-weight bilateral input D-WNet for building change detection in aerial imagery [J]. *Laser & Optoelectronics Progress*, 2024, 61(8): 3.
- [5] 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: 1195-1204.
- [6] 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.
- [7] 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*. Vancouver: IEEE, 2023: 7236-7246.
- [8] BANDARA W G C, PATEL V M. Revisiting consistency regularization for semi-supervised change detection in remote sensing images[EB/OL]. arXiv: 2204.08454 (2022-04-18) [2024-04-16]. <http://arxiv.org/abs/2204.08454>.
- [9] 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.
- [10] LIN T Y, DOLLÁR P, GIRSHICK. *et al.* Feature pyramid networks for object detection[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Hawaii: IEEE, 2017: 2117-2125.
- [11] LOSHCHILOV I, HUTTER F. Decoupled weight decay regularization[EB/OL]. arXiv: 1711.05101 (2017-11-14) [2024-04-16]. <https://arxiv.org/abs/1711.05101>.
- [12] FANG S, LI K, LI Z. Changer: feature interaction is what you need for change detection[J/OL]. arXiv: 2209.08290v1 (2022-09-17)[2024-01-20]. <https://github.com/likyoo/open-cd>.
- [13] 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.
- [14] PENG D, LORENZO B, ZHANG Y J, *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*, 2020, 59(7): 5891-5906.
- [15] 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.
- [16] CODEGONI A, LOMBARDI G, FERRARI A. TINYCD: a (not so) deep learning model for change detection [J]. *Neural Computing and Applications*, 2023, 35(11): 8471-8486.
- [17] 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*. Malaysia: IEEE, 2022: 207-210.
- [18] KOLESNIKOV A, BEYER L, ZHAI X, *et al.* Big transfer (bit): general visual representation learning[C]//*Computer Vision-ECCV*. Glasgow: ECCV, 2020: 491-507.
- [19] 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.
- [20] SUN C, CHEN H, DU C, *et al.* SemiBuildingChange: a semi-supervised high-resolution remote sensing image building change detection method with a pseudo bitemporal data generator [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1-19.

(责任编辑: 张玮欣 英文审核: 熊荣斌)