

中图法分类号: TP751 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-16

论文引用格式: Hou Wenhui, Peng Jiangtao, Yang Bing, Deng Ziyang, Sun Weiwei. Bidirectional contrastive and spatial-frequency synergetic Transformer network for cross-scene hyperspectral multi-class change detection [J/OL]. Journal of Image and Graphics, XXXX: 1-16. DOI: 10.11834/jig.260365. (候文慧, 彭江涛, 杨冰, 邓紫莹, 孙伟伟. 跨场景高光谱多类变化检测的双向对比与空频协同Transformer网络[J/OL]. 中国图象图形学报, XXXX: 1-16. DOI: 10.11834/jig.260365.) [DOI: 10.11834/jig.260365]

跨场景高光谱多类变化检测的双向对比与空频协同Transformer网络

候文慧^{1,2}, 彭江涛^{1,2*}, 杨冰³, 邓紫莹^{1,2}, 孙伟伟^{4,5}

1. 湖北大学数学与统计学学院, 武汉 430062; 2. 应用数学湖北省重点实验室, 武汉 430062; 3. 中国计量大学理学院, 杭州 310018; 4. 宁波大学地理科学与遥感技术学院, 宁波 315211; 5. 宁波市海岸带遥感与生态安全重点实验室, 宁波 315211

摘要: 目的 高光谱多类变化检测旨在识别同一区域不同时间地表覆盖的具体变化类型, 在生态监测、资源调查等领域具有重要应用价值。然而, 双时相高光谱图像之间由于光照条件、气候变化等因素产生伪变化现象, 容易对检测结果造成负面干扰。同时获取高光谱图像标注数据的成本较高, 导致现有深度学习方法在标记样本有限时难以获得理想的检测效果。针对上述问题, 本文提出了一种面向小样本高光谱多类变化检测的双向对比与空频协同Transformer网络 (bidirectional contrastive and spatial-frequency synergetic Transformer network, BCSFST)。**方法** 首先设计频域感知增强模块, 通过傅里叶变换提取频域信息, 利用幅度与相位特征构建动态校正机制自适应优化原始空间特征, 并生成频域先验与空间特征共同构建空频协同Transformer, 从而增强变化特征的判别能力并抑制伪变化干扰。其次, 提出自适应跨场景关联对比学习策略, 通过双向匹配概率构建软正样本对, 挖掘源域和目标域之间的潜在关联, 实现跨场景特征对齐。**结果** 在3个高光谱多类变化检测数据集上进行实验验证, 在每类仅使用5个标记样本参与训练的条件下, BCSFST在Irrigated Area、HangZhouWan和China数据集上分别取得了90.80%、83.67%和85.78%的总体准确率, 均优于现有先进算法。同时, 消融实验进一步验证了本文设计的频域感知增强模块和自适应跨场景关联对比学习策略对模型性能提升的有效性。**结论** 本文所提出的BCSFST模型通过频谱变换与对比学习, 能够充分结合Transformer网络与少样本学习的优势, 有效提升了少样本情形下高光谱多类变化检测的准确性。

关键词: 高光谱图像; 变化检测; Transformer网络; 傅里叶变换; 对比学习; 少样本学习

Bidirectional contrastive and spatial-frequency synergetic Transformer network for cross-scene hyperspectral multi-class change detection

Hou Wenhui^{1,2}, Peng Jiangtao^{1,2*}, Yang Bing³, Deng Ziyang^{1,2}, Sun Weiwei^{4,5}

1. Faculty of Mathematics and Statistics, Hubei University, Wuhan 430062, China; 2. Hubei Key Laboratory of Applied Mathematics, Wuhan 430062, China; 3. College of Sciences, China Jiliang University, Hangzhou 310018, China; 4. School of Geographical Science and Remote Sensing Technology, Ningbo University, Ningbo 315211, China; 5. Ningbo Key Laboratory of Remote Sensing and Ecological Security of Coastal Zone, Ningbo 315211, China

Abstract: **Objective** Hyperspectral multi-class change detection aims to identify specific types of changes in land cover at

收稿日期: 2026-06-25; 修回日期: 2026-09-29

* 通信作者: 彭江涛 pengjt1982@126.com

基金项目: 国家自然科学基金 (项目编号: 42471417; 42471404; 62501557)

Supported by: National Natural Science Foundation of China under Grant Nos. 42471417, 42471404 and 62501557.

© 中国图象图形学报版权所有

different times in the same area, which has important application value in fields such as ecological monitoring, land use analysis, and resource surveys. In recent years, deep learning-based methods have been widely used for change detection in remote sensing images due to their powerful feature extraction capabilities, and have achieved good detection results when labels are sufficient. However, the hyperspectral image annotation process relies heavily on manual intervention, and the cost of obtaining annotations is high. This makes it difficult for existing deep learning methods to achieve ideal detection results when the labels are limited, which is one of the important challenges currently facing hyperspectral change detection. Meanwhile, dual-temporal hyperspectral images are easily affected by factors such as lighting conditions, climate change, and differences in imaging environment during the imaging process, resulting in pseudo changes, which makes it difficult for the model to distinguish between real changed and unchanged areas accurately. Furthermore, for cross-scene change detection tasks, the common differences in domain distribution across different scenes further increase the difficulty of hyperspectral change detection under small sample conditions. How to make full use of the knowledge in existing scenes and reduce the distribution differences between different domains is one of the key problems that need to be solved in cross-scene small sample hyperspectral image multi-class change detection. To address the aforementioned issues, this paper proposes a bidirectional contrastive and spatial-frequency synergetic Transformer network (BCSFST) for cross-scene hyperspectral multi-class change detection in small samples. The BCSFST based on a Transformer network, effectively integrates few-shot learning, frequency-domain prior modeling and contrastive learning, constructing a unified learning mechanism from dual-temporal feature enhancement to cross-scene knowledge transfer. **Method** Firstly, a frequency aware enhancement module is designed to map the features of the dual-temporal image to the frequency domain space through the Fourier transform, and extract the amplitude and phase information respectively. Subsequently, a dynamic correction mechanism is constructed using the amplitude and phase features to adaptively optimize the original spatial features, and the correction weights are adaptively adjusted according to the feature differences at different positions. Based on this, the corrected frequency domain information is mapped back to the spatial domain to further construct the frequency domain prior information. Subsequently, the frequency domain prior and spatial features are input into the Transformer, and the correlation between frequency information and spatial semantic features is established through a spatial-frequency domain collaborative attention mechanism. This enables collaborative modeling of the two types of complementary information, thereby enhancing the ability to discriminate changed features and suppressing pseudo change interference. Secondly, to address the significant distributional differences between the source and target domains in cross-scene few-shot learning, this paper proposes an adaptive cross-scene association contrastive learning strategy. This method generates adaptive neighbor weights by calculating the bidirectional consistency probability between samples from two domains, and mines potential sample matching relationships between different domains under the premise of limited target domain labeled information. Then, soft positive samples are constructed using the generated weights to establish cross-scene contrastive relationships, enhancing the feature compactness among similar samples. By using the bidirectional consistency constraints, the reliability of the feature association between the source and target domains is further improved. The feature distribution differences between different domains are reduced, and the cross-scene knowledge transfer capability and generalization performance of the model under the finite labeling conditions are enhanced. The synergistic effect of the two modules mentioned above enables the model to simultaneously mitigate the problems of pseudo change interference and domain shift from two aspects: frequency domain feature enhancement and cross-scene feature alignment, providing a more stable feature representation for the hyperspectral multi-class change detection in small labeled samples. **Result** Experiments are conducted on three hyperspectral multi-class change detection datasets. Under the condition of using only five labeled samples for training in each class, the BCSFST proposed in this paper achieved overall accuracies of 90.80%, 83.67%, and 85.78% on the Irrigated Area, HangZhouWan, and China datasets, respectively, all of which are superior to existing state-of-the-art algorithms. This demonstrates that the proposed method can still uncover discriminative features across different classes using a very small number of labeled samples, exhibiting good cross-scene generalization ability. Based on the quantitative results, the visualization of experimental results can further demonstrate the effectiveness of the proposed method. Compared with other comparison algorithms, the proposed BCSFST can more accurately identify the boundaries of changed regions and effectively reduce the false detection rate of background regions. Meanwhile, ablation

experiments further validated the effectiveness of the frequency-aware enhancement module and the adaptive cross-scene association contrastive learning strategy designed in this paper in improving model performance. After removing either of the two modules or the adaptive weights, the detection performance of the model decreased to varying degrees, indicating that the design of the two modules played a positive role in the overall model. In terms of computational time and the number of parameters, the BCSFST achieves both high detection accuracy and computational efficiency, demonstrating excellent overall performance. Finally, experiments with different numbers of labeled samples show that the proposed BCSFST achieves optimal performance in all cases, demonstrating that the proposed method has good adaptability and robustness for hyperspectral multi-class change detection tasks. **Conclusion** The BCSFST model proposed in this paper effectively mitigates pseudo change interference and cross-scene domain shift problems through spectral transformation and contrastive learning. Experimental results show that the proposed method can fully combine the advantages of Transformer network and few-shot learning, and significantly improve the accuracy of hyperspectral multi-class change detection in the case of few labeled samples. In the future, this paper will further explore hyperspectral multi-class change detection methods in unsupervised scenes, reduce the model's dependence on labeled samples and improve its generalization ability and practical application value in complex remote sensing scenes.

Key words: Hyperspectral image; change detection; Transformer network; Fourier transform; contrastive learning; few-shot learning

论文引用格式: [DOI: 10.11834/jig.260365]

0 引言

高光谱图像 (hyperspectral image, HSI) 相较于普通 RGB 图像具有更加丰富的空间-光谱信息, 能够更好地描述地物的精细特征, 从而有效提升不同地物之间的可区分性 (Plaza 等, 2009; Zhou 等, 2026; Ding 等, 2024)。得益于这一优势, 高光谱变化检测 (change detection, CD) 逐渐成为当前的研究热点之一, 其旨在从双时相图像中提取可靠的变化信息, 已被广泛应用于土地覆盖、灾害监测、城市扩张等领域 (Zhang 等, 2022; Zheng 等, 2021; Lv 等, 2023; Tu 等, 2024; Wang 等, 2024; Zhang 等, 2026)。

根据检测任务的不同, 高光谱变化检测通常可以分为二分类变化检测和多分类变化检测。多分类变化检测在判断是否存在变化的基础上需要进一步确定变化的具体类型, 因此检测难度更高, 对特征判别需求也更大。本文主要关注多分类变化检测。早期, 传统的高光谱变化检测方法往往通过人工设计特征并结合阈值分割、聚类分析等方法完成变化检测任务, 比如基于代数运算的变化向量分析方法 (Malila, 1980)、基于变换的主成分分析方法 (Rodarmel 等, 2002; Licciardi 等, 2011) 和基于直接分类的支持向量机 (support vector machine, SVM) 方法 (Bovolo 等, 2008; Habib 等, 2009) 等。尽管这

些方法在一定程度上能够识别变化区域, 但其特征提取过程高度依赖人工经验, 在复杂场景下存在特征表达能力不足、鲁棒性较弱等问题, 限制了变化检测性能的进一步提升。

随着技术的快速发展, 深度学习 (deep learning, DL) 逐渐被应用于高光谱变化检测问题中并取得了显著的性能提升 (Qu 等, 2022; Luo 等, 2023; Yang 等, 2024)。例如, Song 等人 (2022) 提出了一种跨时空交互对称注意力网络 (cross-temporal interaction symmetric attention network, CSANet) 用于二值变化检测, 通过跨时空注意交互和空间-光谱-时间特征提取来增强变化特征的判别能力。Li 等人 (2025) 提出了一种基于混合注意力和双向门控网络的高光谱变化检测方法, 通过逐步提取图像的全局-局部特征实现了多时相高光谱图像特征的深度融合。Ding 等人 (2024) 提出了一种多级特征融合与变化信息增强神经网络 (multi-level features fused and change information enhanced neural network, MFCEN) 同时用于二分类和多分类变化检测。该方法通过孪生神经网络提取双时相图像的多层特征, 并结合变化信息增强与多级特征融合策略提升高光谱变化检测的整体性能。Wu 等人 (2025) 提出了一种用于二值变化检测的自适应信息加权和同步增强网络 (adaptive information weighting and synchronized enhancement network, AIWSEN), 该网络利用最大离散熵和交互式聚合来增强差异信息, 并通过

门控机制筛选重要特征。

相较于传统的卷积神经网络 (convolutional neural networks, CNN) 而言, Transformer 凭借自注意力机制能够从全局范围捕获特征之间的依赖关系, 并实现空间、光谱以及时间信息的协同建模, 从而展现出更强的特征表达能力 (Huang 等, 2024)。近年来, 越来越多的研究将 Transformer 引入遥感图像变化检测任务。例如, Sun 等人 (2026) 提出了一个频谱引导边缘增强的 Transformer 网络用于高光谱变化检测, 通过光谱频率与空间融合模块挖掘频域与空间域互补信息, 并结合拉普拉斯引导边缘增强机制实现精细边界提取。Tang 等人 (2025) 提出了一种基于边界检测和任务交互的语义变化检测网络 (semantic change detection network based on boundary detection and task interaction, BT-SCD)。该网络以 Transformer 为基线, 通过引入边界检测任务以增强语义分割任务和变化检测任务之间的相关性。Wang 等人 (2025) 提出了一种用于高光谱二值变化检测的全局-局部注意 Transformer 网络, 通过交叉门控机制来突出显著特征并抑制噪声干扰。尽管基于深度学习的方法取得了良好的检测性能, 但其优异表现往往依赖于大量标记样本。然而在实际中, 高光谱图像标签的获取需要耗费大量人力物力。因此, 当可用标签数量有限时, 现有深度学习方法容易出现过拟合, 模型的泛化能力和检测性能会受到一定程度的影响。

为了解决标记样本稀缺带来的挑战, 学者们提出了一系列解决方法, 如数据增强、域适应、少样本学习等 (Li 等, 2019; Zhang 等, 2026)。例如, Wang 等人 (2021) 通过生成不同窗口大小的相似样本来扩充训练集, 以缓解标记样本不足的问题。Zhang 等人 (2025) 则采用基于域适应的三支 Transformer 实现标签有限条件下的高光谱图像分类。近年来, 少样本学习 (few-shot learning, FSL) 也逐渐成为遥感领域的重要研究方向之一。FSL 能够从有限标注样本中快速学习到类别特征, 并将已有信息迁移到未知类别的识别任务中 (Vinyals 等, 2016; Sun 等, 2019; Yang 等, 2020)。然而, 由于遥感数据通常具有明显的场景差异和分布偏移, 传统的 FSL 方法难以直接适用于不同区域或不同传感器获取的数据。因此, 跨场景 FSL 已被逐渐应用到高光谱图像处理中。现有基于跨域 FSL 的方法主要围绕特征

分布对齐、关系建模、知识迁移等方向展开研究。在高光谱图像分类任务中, Li 等人 (2021) 将 FSL 与域适应问题相结合, 提出了一种深度跨域 FSL 方法 (deep cross-domain FSL, DCFSL) 用于高光谱图像分类。Zhang 等人 (2024) 提出了一种图信息聚合跨域 FSL 方法 (graph information aggregation cross-domain FSL, Gia-CFSL) 用于高光谱图像分类。通过域内分布提取模块与跨域相似性感知模块来建模域内非局部关系并捕获域间分布的相似性。Zhang 等人 (2025) 为了突破全局表征对地物类别表征能力有限的问题, 在跨场景 FSL 的基础上构建局部表征机制。随着相关研究的不断深入, 跨域 FSL 逐渐被引入高光谱变化检测任务中, 为缓解标记样本不足和场景差异带来的挑战提供了新的解决思路。Wang 等人 (2023) 提出了一种基于双向生成的跨异构域 FSL 框架 (cross-heterogeneous domain FSL framework based on bidirectional generation, BiG-FSLF) 用于二分类高光谱变化检测, 该方法通过基于生成对抗策略的双向生成域适应方法实现源域及目标域特征分布的自适应对齐, 减轻域偏移问题对 FSL 的影响。在此基础上, Li 等人 (2026) 进一步建模样本间的结构关系, 提出了一种异构图感知聚合驱动的 FSL 网络, 通过构建多级图结构聚合局部与远距离特征信息, 并利用跨域知识迁移提升少样本条件下的检测性能。在多类变化检测任务中, Feng 等人 (2025) 提出了一种基于分数域信息的跨域 FSL 多类高光谱变化检测方法, 通过构建光谱-空间-分数域联合特征表征并引入拓扑结构对齐机制, 实现了跨域知识迁移与域间分布对齐。尽管上述方法在跨域 FSL 任务中取得了一定进展, 但是现有基于跨域 FSL 的高光谱变化检测方法大多面向二值任务展开研究, 对于具有多类别的变化检测问题关注不足, 难以充分利用不同变化类别之间的语义差异。另外, 现有方法大多主要关注源域与目标域之间的整体分布对齐, 对成像条件变化、光谱波动等因素引起的伪变化干扰考虑不足, 如何增强模型对于真实变化特征与伪变化特征之间的判别能力仍有待进一步研究。

综上所述, 针对跨场景高光谱变化检测中标记样本匮乏、域间分布差异及伪变化干扰等问题, 本文提出了一种双向对比与空频协同 Transformer 网络 (bidirectional contrastive and spatial-frequency syner-

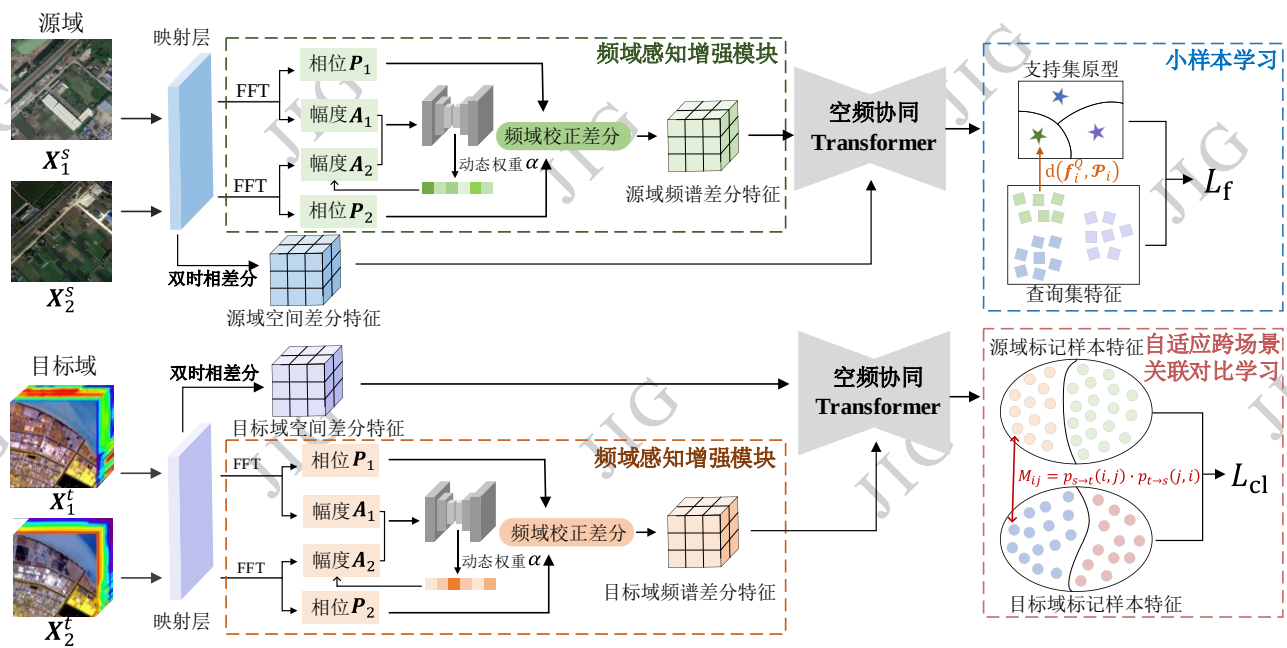


图1 BCSFST网络框架示意图

Fig. 1 The framework of BCSFST

getic Transformer network, BCSFST) 用于高光谱多类变化检测。该网络以 Transformer 为主体, 将 FSL、频域先验建模和对比学习有效融合, 构建了从双时相特征增强到跨场景知识迁移的统一学习机制。具体而言, 首先, 本文设计了一个基于傅里叶变换的频域感知增强模块, 通过对双时相特征的频域响应进行自适应建模与优化, 并作为先验知识构建空频协同 Transformer 框架, 缓解伪变化问题带来的干扰。同时, 为了缓解域偏移问题, 本文设计了一种基于双向对比学习的跨域特征对齐策略, 挖掘源域与目标域之间潜在对应关系。所提出的 BCSFST 算法在 3 个高光谱变化检测数据集上均取得了较为优异的检测结果。本文主要贡献如下:

1) 针对变化检测问题中由于外界因素引起的伪变化干扰问题, 设计了一种空频协同 Transformer 网络。通过频域感知增强模块对双时相特征进行频域建模与自适应优化, 并作为先验知识输入 Transformer 网络, 实现频域信息与空间语义信息的协同建模。

2) 针对源域和目标域之间存在的域间分布差异, 提出了一个自适应跨场景关联对比学习策略。通过挖掘跨域样本之间的潜在关联, 利用双向匹配概率构建自适应互近邻权重, 引导模型实现跨场景特征对齐, 提升模型的泛化能力。

1 提出方法

1.1 整体框架

本文所提出的 BCSFST 网络整体框架如图 1 所示。以双时相高分辨率遥感影像 (very high-resolution images, VHRI) X_1^s 和 X_2^s 作为源域, 双时相高光谱图像的 X_1^t 和 X_2^t 作为目标域, 通过跨域协同学习实现模型的迭代优化。以源域为例, 首先对输入影像进行统一维度映射以缓解异构数据间的特征差异, 统一维度设置为 $b = 100$ 。随后采用频域感知增强模块对 T_1 和 T_2 的特征进行傅里叶变换, 并对二者幅度值进行自适应加权, 构建频域引导的差分特征 D_i^* 。与此同时, 对原始特征直接差分得到空间差分特征 D_i^s 。将 D_i^* 输入 Transformer 网络以捕获全局空间依赖关系, 并利用 D_i^s 对空间语义信息进行协同校正, 提升真实变化区域的表征能力同时抑制伪变化干扰。最后, 设计了自适应跨场景关联对比学习模块, 通过自适应权重实现源域和目标域特征空间的有效对齐。在测试阶段, 本文采用最近邻 (nearest neighbor, NN; Samaniego 等, 2008) 分类器对目标域剩余无标签样本进行预测, 完成多类变化检测任务。

1.2 少样本学习

少样本学习 (few-shot learning, FSL) 旨在利用

具有充足标记样本的源域来引导仅有极少量标记样本的目标域完成任务。在跨场景 FSL 问题中,源域和目标域的数据类型及类别数均不一定相同。相较于目标域类别数 C_t ,源域类别数 C_s 满足 $C_s \geq C_t$ 。源域和目标域的训练方法保持一致,以源域数据为例,首先对选取的极少量标记样本进行数据增强,并将增强后的数据划分为支持集 $S^s = \{(x_i^s, y_i^s) | i=1\}$ 和查询集 $Q^s = \{(x_j^s, y_j^s) | j=1\}$,其中 n_s 和 n_q 分别代表了每类中选取的支持集和查询集样本数量,在本文中,将二者分别设置为 1 和 19 (Liu 等, 2020)。对于每一次元训练过程,采用基于原型网络的策略来优化支持集和查询集的特征。在利用空频协同 Transformer 进行特征提取后,源域中支持集 S^s 的特征空间可以表示为 $F_s^s = \{f_1^s, \dots, f_{n_s}^s\}$,则第 c 类的原型可表示为:

$$\mathcal{P}_c^s = \frac{1}{n_s} \sum_{f_m^s \in F_s^s} f_m^s \quad (1)$$

式中, f_m^s 为样本 x_m^s 对应的特征, $F_{s,c}^s$ 为源域支持集中第 c 类样本特征的集合。则样本 $x_i^s \in Q^s$ 对于第 c 类的概率分布可表示为:

$$p(y = cx_i^s) = \frac{\exp(-d(f_i^s, \mathcal{P}_c^s))}{\sum_{c_i=1}^{C_s} \exp(-d(f_i^s, \mathcal{P}_{c_i}^s))} \quad (2)$$

式中, $d(\cdot)$ 表示度量方式,本文选择欧氏距离进行计算; $\mathcal{P}_{c_i}^s$ 为第 c_i 类的原型。最终,通过最小化查询集样本的类概率分布的负值得到源域的损失函数 L_t^s :

$$L_t^s = \mathbb{E}_{s, Q^s} \left[- \sum_{(x^s, y^s) \in Q^s} \log p(y^s = cx^s) \right] \quad (3)$$

同理可得目标域 FSL 损失 L_t^t 。最终模型整体的 FSL 损失为:

$$L_t = L_t^s + L_t^t \quad (4)$$

1.3 空频协同 Transformer 网络

相较于传统 CNN 而言,Transformer 网络能够通过自注意力建立空间位置之间的关联,从而突破卷积操作局部感受野的限制,实现全局上下文信息建模。因此,本文选择基于 Transformer 的网络作为特征提取器,在此基础上针对变化检测任务特性引入频域感知增强模块与空间-频域协同注意力,实现空间-频域的协同感知。

1.3.1 频域感知增强模块

在高光谱变化检测任务中,由于成像条件变化、

光照差异、传感器噪声等影响因素,双时相影像之间可能存在伪变化区域,从而影响模型对变化特征的判别能力。为了减轻伪变化问题对变化检测性能的影响,本文设计了一个针对双时相影像的频域感知增强模块 (frequency-aware enhancement module, FAEM),利用频域信息对双时相特征进行自适应优化,抑制伪变化干扰。

以源域为例,对于给定的双时相影像经过统一映射处理的特征 D_1^s 和 D_2^s ,沿空间维度对二者分别执行二维快速傅里叶变换 (fast Fourier transform, FFT):

$$\mathcal{F}(D_k^s)(m, n) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} D_k^s(h, w) e^{-j2\pi(\frac{h}{H}m + \frac{w}{W}n)} \quad (5)$$

式中, $k \in \{1, 2\}$, $j^2 = -1$, $D_k^s \in \mathbb{R}^{b \times H \times W}$ (H, W 分别为输入图像的高度和宽度), (h, w) 为空间坐标, (m, n) 为频域坐标。经过 FFT,对于 T_1 和 T_2 的频域特征分解为幅度谱与相位谱:

$$\mathcal{F}(D_k^s) = A_k e^{jP_k} \quad (6)$$

式中, $A_k \in \mathbb{R}^{b \times H \times W}$ 和 $P_k \in \mathbb{R}^{b \times H \times W}$ 分别代表了第 k 时相的幅度和相位。根据已有研究,幅度更容易受到成像条件变化、噪声干扰等外界因素的影响,而相位信息则相对稳定 (Yang 等, 2020)。根据频域特征这一特性,本文重点针对双时相特征的幅度信息进行自适应优化校准。

为了同时考虑双时相特征的幅度信息,FAEM 首先对 2 个幅度信息进行拼接,得到综合信息 A_{mix} :

$$A_{\text{mix}} = \text{Concat}(A_1, A_2) \quad (7)$$

针对 A_{mix} ,本文设计了一个轻量级注意力网络,旨在充分挖掘双时相特征在频域中的关联关系,自适应评估不同频率成分的重要性,并动态调节其在特征重构过程中的贡献。其动态权重可以表示为:

$$\alpha = \sigma(G(A_{\text{mix}})) \quad (8)$$

式中, $G(\cdot)$ 表示由卷积层、归一化层、ReLU 激活函数构成的映射函数, σ 代表 Sigmoid 激活函数。随后利用学习得到的动态权重 α 对 T_2 时相幅度谱进行自适应优化校准,得到校正后的相位 A_2^{corr} :

$$A_2^{\text{corr}} = \alpha \odot A_1 + (1 - \alpha) \odot A_2 \quad (9)$$

上述过程构建了一种频域一致性约束机制。当双时相频域响应具有较高一致性时,FAEM 认为其为伪变化区域,因此倾向于增强共享频率成分,从而削弱由于外界环境带来的伪变化响应;而对真实变

化区域, FAEM 能够保留频域间的差异信息, 突出表达变化区域。

获得 T_2 优化后的幅度谱后, 将其与原始相位谱重新组合并通过 FFT 逆变换将其转换回空间域:

$$\hat{D}_2^s = \mathcal{F}^{-1}(A_2^{\text{corr}} e^{iP_2}) \quad (10)$$

式中, $\hat{D}_2^s$ 为优化后的 T_2 时相特征, 其特征维度与原始保持一致。

最后得到结合频域信息的差分特征 D_i^s :

$$D_i^s = \hat{D}_2^s - D_1^s \quad (11)$$

FAEM 能够抑制由成像条件差异导致的异常频域响应, 同时保留与真实地物变化相关的判别性频率信息, 从而增强变化特征的可靠性。基于此, 本文将经过频域优化增强后的 D_i^s 视为一种可靠变化先验, 在保留原始空间特征中丰富结构信息的基础上, 将 D_i^s 作为引导信息嵌入 Transformer 网络, 实现频域信息与空间语义信息的协同建模。

1.3.2 空频协同 Transformer 网络

本文构建了一个基于 Transformer 网络的全局特征提取器, 其具体结构如图 2 所示。设计的空频协同 Transformer 网络通过空间-频域协同注意力机制在全局范围内建立像素点间的依赖关系, 增强特征的判别能力。以针对源域的特征提取过程为例。首先, 对于输入的双时相影像做差分, 得到空间差分特征 D_s^s :

$$D_s^s = D_2^s - D_1^s \quad (12)$$

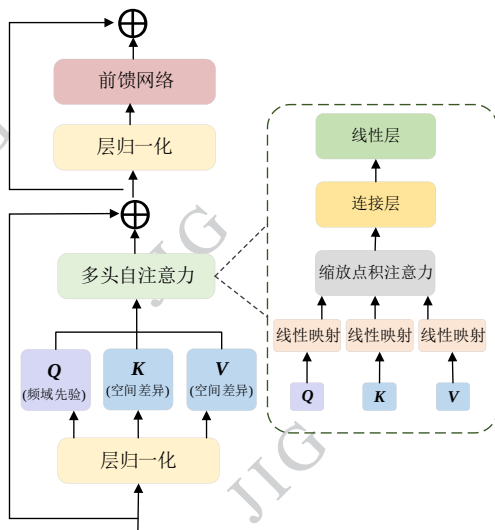


图 2 空频协同 Transformer 网络结构示意图

Fig. 2 Schematic diagram of spatial-frequency synergetic transformer network

在 Transformer 网络中, 往往通过缩放点积注意力机制计算空间位置之间的相关性, 即:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (13)$$

式中, Q, K, V 分别代表了查询 (Query)、键 (Key) 和价值 (Value)。为了充分利用 FAEM 获得的先验信息, 使 Transformer 在保留空间语义信息的同时增强对真实变化区域的响应能力, 本文设计了一个频域引导交叉注意力机制。对于 Q, K 和 V , 本文采用频域先验差分特征 D_i^s 构建 Q , 而利用空间特征 D_i^s 构建 K 和 V , 即:

$$\begin{cases} Q_f = w_q D_i^s \\ K_s = w_k D_i^s \\ V_s = w_v D_i^s \end{cases} \quad (14)$$

式中, w_q, w_k 和 w_v 为投影矩阵参数; Q_f, K_s 和 V_s 分别代表更新后的 Q, K, V 值。通过将频域先验与原始空间信息分别用于 Q 与 $K-V$ 的构建, 使频域信息参与全局关系建模过程, 进而引导 Transformer 优先关注真实变化的位置区域。因此, 空间-频域协同注意力得到的特征可表示为 $\text{Cr_Att}(Q, K, V)$:

$$\text{Cr_Att}(Q_f, K_s, V_s) = \text{softmax}\left(\frac{Q_f K_s^T}{\sqrt{d_k}}\right)V_s \quad (15)$$

将多头注意力得到的特征进行拼接, 并采用残差连接与前馈网络对输出特征进行优化, 得到空频协同 Transformer 网络最终的输出特征 F^s 。

通过上述过程, 频域信息被显式引入 Transformer 自注意力的计算过程, 最终获得兼具全局语义信息和频域变化先验的高层变化表征, 为后续跨域特征对齐和多类变化检测任务提供了更加可靠的特征基础。

1.4 自适应跨场景关联对比学习

由于源域与目标域之间存在显著的分布差异, 且二者的类别空间不是完全一一对应的。若直接通过语义标签实现监督对比学习, 容易产生类别偏移, 导致学习到的信息缺乏跨域泛化能力。因此, 本文设计了一种基于软互近邻的自适应跨场景关联对比学习 (adaptive cross-scene association contrastive learning, ACACL) 策略, 通过挖掘源域和目标域样本之间的潜在对应关系, 构建双向概率一致性约束, 实现无监督跨域特征对齐。

具体而言, 经过空频协同 Transformer 网络后, 源

域和目标域中参与训练的标记样本均可得到一组维度为 128 的特征,记为 $\mathbf{F}^s = \{f_1^s, \dots, f_{n_s+n_q}^s\}$ 和 $\mathbf{F}^t = \{f_1^t, \dots, f_{n_s+n_q}^t\}$ 。对上述特征进行归一化处理,构建样本间的相似性矩阵。样本 x_j^t 相较于 x_i^s 的相似度可以表示为:

$$s_{i \rightarrow j} = \frac{\phi(f_i^s f_j^t)}{\tau} \quad (16)$$

式中, $f_i^s, f_j^t \in \mathbf{R}^{128}$, ϕ 表示相似度度量方法,本文通过余弦相似度计算特征之间的相似性; τ 表示温度参数,用于调节相似度分布的平滑程度,本文中选取 $\tau = 0.1$ 。同理可定义反向相似度 $s_{j \rightarrow i}$ 。

为了刻画跨域样本之间的双向关联关系,本文分别从源域和目标域两个方向建立概率映射。源域样本指向目标域的匹配概率为:

$$p_{s \rightarrow t}(i, j) = \frac{\exp s_{i \rightarrow j}}{\sum_{k=1}^{n_s+n_q} \exp s_{i \rightarrow k}} \quad (17)$$

类似地,可以计算得到目标域样本指向源域的匹配概率 $p_{t \rightarrow s}(j, i)$ 。

由于单向最近邻易受到跨域噪声和特征偏移的影响,本文认为只有同时满足双向高置信匹配的样本对才更有可能属于同一潜在空间中。因此,本文通过双向概率的一致性构建自适应互近邻权重 M_{ij} :

$$M_{ij} = p_{s \rightarrow t}(i, j) \cdot p_{t \rightarrow s}(j, i) \quad (18)$$

当且仅当样本对 (x_i^s, x_j^t) 在两个方向上均具有较高匹配概率时,其对应权重才会被显著增强,从而提高跨域样本关联的可靠性。自适应互近邻权重 M_{ij} 可以视作软正样本分布,其不再局限于单一最近邻样本,而是为每个源域样本动态分配多个潜在正样本,充分利用了跨域特征空间中的局部结构信息。基于 M_{ij} ,将样本自适应划分为正样本对和负样本对:

$$\begin{cases} x_{i,+}^s = \sum_{j=1}^{n_s+n_q} M_{ij} \cdot \exp s_{i \rightarrow j} \\ x_{i,-}^s = \sum_{j=1}^{n_s+n_q} (1 - M_{ij}) \cdot \exp s_{i \rightarrow j} \end{cases} \quad (19)$$

最终,构建基于双向一致性的跨域对比损失:

$$L_{cl}^{s \rightarrow t} = -\frac{1}{n_s + n_q} \sum_{i=1}^{n_s+n_q} \log \left(\frac{x_{i,+}^s}{x_{i,+}^s + x_{i,-}^s} \right) \quad (20)$$

同理可得目标域对比损失 $L_{cl}^{t \rightarrow s}$,二者相加即为

整体的跨域对比损失 L_{cl} 。该损失通过最大化正样本相似度占比,引导源域与目标域中具有相同潜在语义的样本在特征空间内逐渐聚集,有效降低错误匹配带来的噪声传播问题,提高跨域对齐过程的稳定性与鲁棒性。

1.5 整体损失函数

本文所提出的 BCSFST 网络在 FSL 损失的基础上,针对跨场景多类变化检测中存在的伪变化问题及跨域特征分布问题引入了基于频域信息的 FAEM 及基于对比学习的 ACACL。最终,BCSFST 的损失函数 L 由两部分共同组成:

$$L = L_f + L_{cl} \quad (21)$$

2 实验结果

2.1 实验数据集

本文共使用了 4 个数据集,包括源域的 1 个 VHRI 数据集和目标域的 3 个 HSI 数据集,其类别信息及样本数量如表 1 所示,其中 C0 表示未变化, C1-C6 代表了不同类型的变化。4 个数据集的相关描述如下:

表 1 源域及目标域数据集的类别信息

Table 1 Category information of the datasets of the source and target domains

		源域		目标域	
颜色	类别	Second	Irrigated Area	HangZhouWan	China
	C0	188354	30579	84891	40407
	C1	3750	1034	1846	6019
	C2	3537	1048	5539	12229
	C3	6458	5111	10032	145
	C4	19649	1261	3330	-
	C5	4615	479	-	-
	C6	35781	988	-	-
	未知类别	-	-	54362	-

注:-表示无该类别样本。

1) **Second 数据集**: Second 是一个面向语义变化检测任务的 VHRI 数据集,包含来自中国多个城市的多组 512×512 双时相遥感影像。本文随机选取其中一组作为源域数据参与训练,其双时相伪彩色图

像及地物如图3所示。

2) Irrigated Area 数据集: Irrigated Area (IA) 数据集采集于美国俄勒冈州本顿县的一处灌溉农田区域, 包含了两幅分别获取于2004年5月1日和2007年5月8日的高光谱图像。该数据集大小为 180×225 , 包含159个光谱波段。其双时相伪彩色图像及地物如图4所示。



(a) T_1 时刻图像 (b) T_2 时刻图像 (c) 地物真值

((a) image on T_1 ; (b) image on T_2 ; (c) Ground truth)

图3 Second数据集的双时相伪彩色图像及地物真值

Fig. 3 Dual-temporal pseudo-color images and ground truth of the Second dataset



(a) T_1 时刻图像 (b) T_2 时刻图像 (c) 地物真值

((a) image on T_1 ; (b) image on T_2 ; (c) Ground truth)

图4 IA数据集的双时相伪彩色图像及地物真值

Fig. 4 Dual-temporal pseudo-color images and ground truth of the IA dataset

3) HangZhouWan 数据集: HangZhouWan (HZW) 数据集是一个新的高光谱变化检测数据集, 其由ZY 1-02D卫星分别于2021年1月29日和2022年2月27日采集于中国杭州湾沿海区域, 包含 400×400 个像素及166个光谱波段, 其中有54362个未知类别像素不参与训练及测试。其双时相伪彩色图像及地物如图5所示。

4) China 数据集: China 数据集采集于中国江苏省盐城市的一个农田区域, 获取时间分别为2006年5月3日和2007年4月23日。数据集大小为 450×140 像素, 包含154个光谱波段。根据地物变化情况, 共标注了3类变化类别及1类未变化类别, 其双



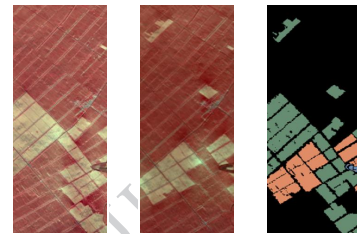
(a) T_1 时刻图像 (b) T_2 时刻图像 (c) 地物真值

((a) image on T_1 ; (b) image on T_2 ; (c) Ground truth)

图5 HZW数据集的双时相伪彩色图像及地物真值

Fig. 5 Dual-temporal pseudo-color images and ground truth of the HZW dataset

时相伪彩色图像及地物如图6所示。



(a) T_1 时刻图像 (b) T_2 时刻图像 (c) 地物真值

((a) image on T_1 ; (b) image on T_2 ; (c) Ground truth)

图6 China数据集的双时相伪彩色图像及地物真值

Fig. 6 Dual-temporal pseudo-color images and ground truth of the China dataset

2.2 实验设置及对比算法

为验证所提出BCSFST网络的有效性, 本文选取了8种对比方法, 包含传统算法SVM (Bovolo 等, 2008); 基于深度学习的4种方法: BT-SCD (Tang 等, 2025)、CSANet (Song 等, 2022)、MFCEN (Ding 等, 2024) 和 AIWSEN (Wu 等, 2025); 以及基于FSL的算法 BiG-FSLF (Wang 等, 2023)、DCFSL (Li 等, 2021) 和 Gia-CFSL (Zhang 等, 2024)。为确保实验的公平性, 所有对比方法均采用原文推荐的参数设置, 并统一选取每类5个标记样本参与训练。实验结果采用总体准确率 (overall accuracy, OA)、类别准确率 (class accuracy, CA) 和 kappa 系数 (κ) 作为评价指标, 综合评价各方法的检测性能。

所有实验均在配备 NVIDIA RTX 2080Ti 显卡的设备上完成训练, 实验环境为 Python 3.7, 深度学习框架采用 PyTorch 实现。对于本文提出的 BCSFST 网络, 学习率设置为 0.001, 训练迭代次数为 800。

表2 9种算法在IA数据集上的检测结果

	SVM (Bovolo 等, 2008)	BT-SCD (Tang 等, 2025)	CSANet (Song 等, 2022)	AIWSEN (Wu 等, 2025)	MFCEN (Ding 等, 2024)	DCFSL (Li 等, 2021)	BiG-FSLF (Wang 等, 2023)	Gia-CFSL (Zhang 等, 2024)	BCSFST (本文)
C0	75.09	72.64	74.79	69.10	80.91	87.59	90.38	85.56	93.35
C1	92.09	98.05	85.34	89.64	90.59	88.11	91.27	89.88	88.73
C2	75.90	75.92	76.75	82.76	67.61	72.18	77.70	65.95	73.64
C3	94.12	86.57	87.96	91.99	80.14	83.83	89.21	85.93	86.18
C4	58.38	72.75	70.78	83.01	76.81	69.88	73.96	66.60	74.21
C5	18.83	89.67	93.80	96.47	93.19	93.53	94.26	90.08	85.15
C6	68.36	77.09	82.88	82.50	74.74	85.23	85.98	77.04	79.88
OA	76.60	75.45	77.07	73.96	80.58	86.20	89.36	84.47	90.80
κ	57.90	56.29	58.78	55.41	61.72	70.68	76.54	67.84	78.84

注:粗体表示最佳结果。

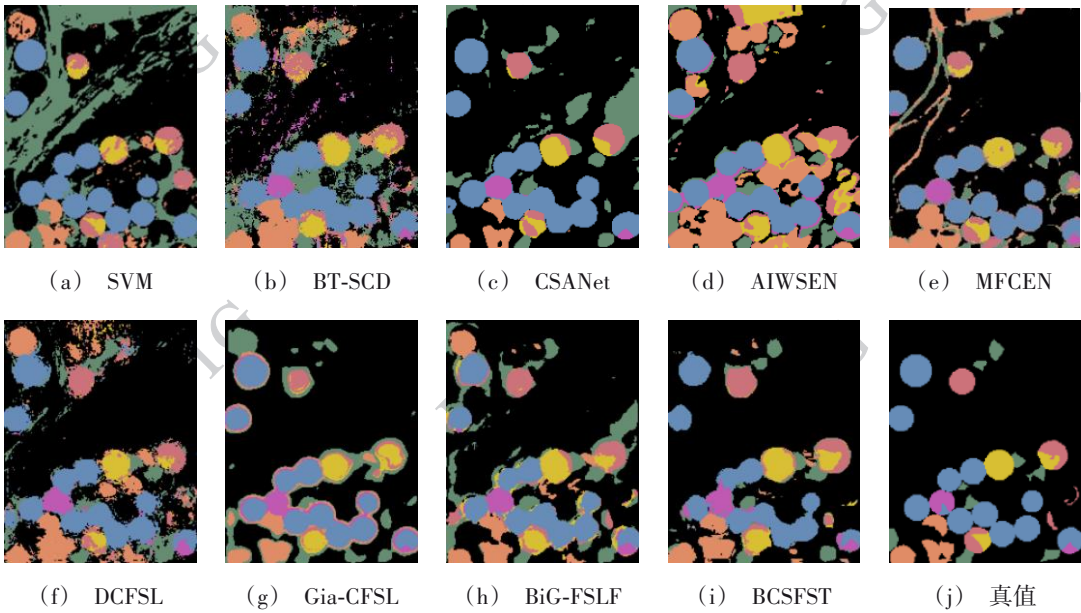


图7 IA数据集检测结果示意图

Fig. 7 The detection map on IA dataset. (a) SVM; (b) BT-SCD; (c) CSANet; (d) AIWSEN; (e) MFCEN; (f) DCFSL; (g) Gia-CFSL; (h) BiG-FSLF; (i) BCSFST; (j) Ground truth.

2.3 检测结果及分析

2.3.1 IA数据集上的检测结果

IA数据集上的实验结果如表2所示。从表2中可以看出,在每类仅使用5个标记样本(即一共35个)的情况下,传统算法SVM和基于深度学习的4种方法整体表现均欠佳,这主要是由于深度学习模型通常依赖大量标签学习具有良好泛化能力的特征表示,而在少样本场景下容易出现特征学习不充分

和过拟合现象,限制了模型对复杂变化模式的刻画能力。但是,一些方法在部分变化类别上取得了较好的检测结果,如BT-SCD算法在C1中取得了最好的CA,表明其对特定类别具有较强的拟合能力,但整体泛化能力仍有不足。相较而言,基于FSL的算法能够充分利用有限标注样本挖掘类别间的判别信息,取得了良好的检测结果。本文所提出的BCSFST网络在OA和 κ 上均取得最优结果,OA相较次优算

法 BiG-FSLF 有 1.44% 的提升。此外,BCSFST 在未变化类别 C0 上取得了最佳检测结果,说明本文提出的空频协同Transformer网络中的FAEM能够有效缓解伪变化带来的负面干扰。

图7展示了9种算法在IA数据集的检测结果示意,其中图7(j)为真实地物图。从图7可见,由于标记样本较少,传统算法和基于深度学习的算法难以准确学习变化类别的判别特征,检测结果中存在

较为明显的漏检、误检以及边界错分现象,与真实结果差距较大。本文所提出的BCSFST(即图7(i))与真实地物图更加接近,在复杂边界区域和未变化区域上存在明显的检测优势。图7也说明了本文设计的空频协同Transformer网络及ACACL策略能够有效缓解伪变化的干扰,同时增强类间可分性和类内聚合性,从而获得更可靠的检测结果。

表3 9种算法在 HZW 数据集上的检测结果
Table 3 Detection results of nine algorithms on the HZW dataset

	SVM (Bovolo 等, 2008)	BT-SCD (Tang 等, 2025)	CSANet (Song 等, 2022)	AIWSEN (Wu 等, 2025)	MFCEN (Ding 等, 2024)	DCFSL (Li 等, 2021)	BiG-FSLF (Wang 等, 2023)	Gia-CFSL (Zhang 等, 2024)	BCSFST (本文)
C0	52.03	52.65	62.09	68.45	54.72	77.25	84.36	80.17	87.74
C1	71.04	84.23	66.53	75.34	33.37	71.97	60.66	74.53	75.48
C2	53.71	70.86	66.00	75.57	62.52	70.51	57.44	69.63	61.76
C3	64.59	84.01	69.41	78.99	64.51	71.07	69.04	74.05	72.79
C4	40.96	63.05	66.89	58.31	41.38	62.96	58.33	61.27	53.54
OA	53.29	57.46	63.22	69.63	55.97	75.77	80.26	78.34	83.67
κ	31.77	31.64	34.98	42.29	26.58	48.06	51.52	50.69	57.34

注:粗体表示最佳结果。

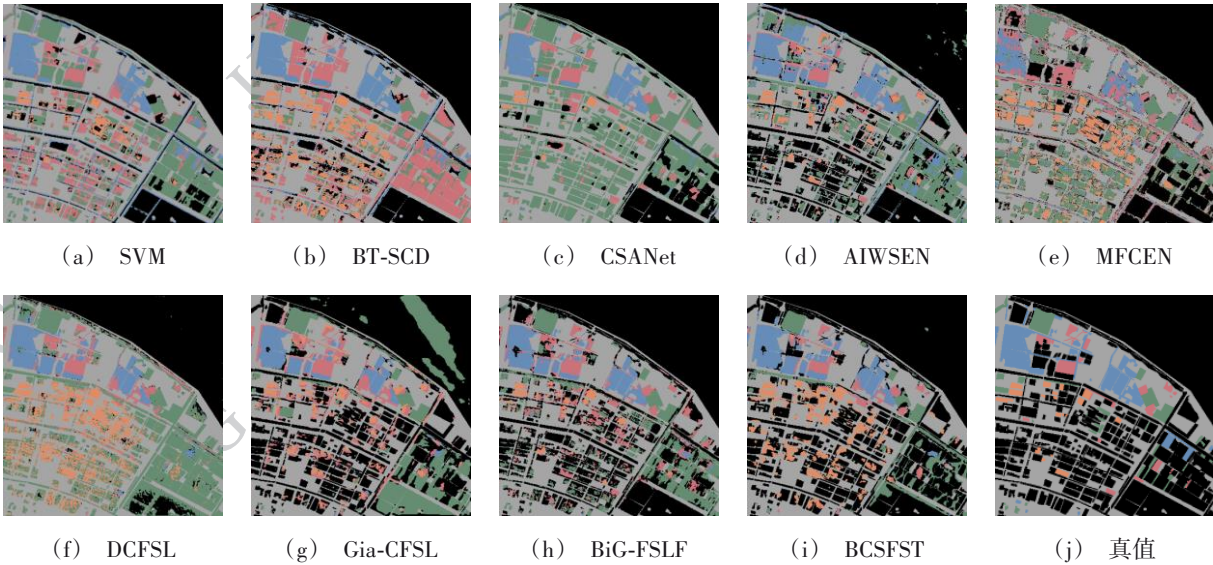


图8 HZW数据集检测结果示意图

Fig. 8 The detection map on HZW dataset. (a) SVM; (b) BT-SCD; (c) CSANet; (d) AIWSEN; (e) MFCEN; (f) DCFSL; (g) Gia-CFSL; (h) BiG-FSLF; (i) BCSFST; (j) Ground truth.

2.3.2 HZW数据集上的检测结果

表3和图8分别展示了所提出BCSFST算法和8种对比方法的定量评价结果及相应可视化。由于

HZW数据集中的变化类别空间分布较为离散,因此该数据集具有较高的检测难度。如表3所示,相较于传统算法,4种深度学习方法有一定性能提升,

AIWSEN 依托自适应信息加权机制获得了最佳结果。然而,受限于标记样本数量,深度学习方法仍难以充分学习稳定的判别特征,其性能提升受到明显限制。相比之下,FSL方法通过引入跨域知识迁移缓解了标记样本不足的影响。从整体指标来看,本文提出的BCSFST取得了最高的OA和 κ ,表明其具有更优的整体检测能力。同时,BCSFST在C0上取得了84.36%的CA,且在变化类别上也展现出稳定的性能,进一步验证了本文方法在复杂场景下的鲁棒性和适用性。

结合图8中的可视化结果可以发现,大部分方法在未变化区域周围产生了较多离散噪声点,导致检测结果与真实标签图存在明显差异。相较而言,本文提出的BCSFST通过引入跨域对比学习与频域空间校正,能够更准确地刻画变化区域的空间分布,并有效减少了背景区域的误检现象。

2.3.3 China数据集上的检测结果

China数据集的定量结果如表4所示。从表4来

看,由于China数据集类别较少,各方法均取得了相对较高的整体检测精度。由于China数据集样本分布不均匀,不同方法在CA上的表现仍存在较大差异。BT-SCD、CSANet和AIWSEN在C1上均取得了优异的检测结果,但是在某些类别有明显性能下降,说明这些方法在训练过程中更容易受到类别分布不均衡的影响,过拟合于部分类别。同时,基于FSL的算法整体性能优于传统和深度学习算法,而BCSFST在取得最高OA的同时在各类别之间表现出更加均衡的检测能力,这说明了本文的方法能够充分挖掘有限监督信息中的潜在结构关系,从而获得更加稳定的检测结果。

图9展示了不同方法的检测结果。由于China数据集的区域边界较为复杂,存在明显的交错和过渡区域,因此大部分方法均出现了不同程度的检测错误。相比之下,本文提出的BCSFST能够在保证各类别识别能力的同时兼顾整体场景的变化分布,与真实地物图更接近,展现出更好的整体一致性。

表4 9种算法在China数据集上的检测结果
Table 4 Detection results of nine algorithms on the China dataset

	SVM (Bovolo 等, 2008)	BT-SCD (Tang 等, 2025)	CSANet (Song 等, 2022)	AIWSEN (Wu 等, 2025)	MFCEN (Ding 等, 2024)	DCFSL (Li 等, 2021)	BiG-FSLF (Wang 等, 2023)	Gia-CFSL (Zhang 等, 2024)	BCSFST (本文)
C0	83.52	86.18	86.94	71.39	87.78	86.92	85.84	86.11	88.29
C1	86.62	91.69	92.22	92.47	78.28	87.16	86.68	89.67	88.90
C2	88.51	65.77	74.09	84.89	61.88	77.82	82.24	75.85	76.04
C3	0.00	98.28	63.45	64.97	91.86	75.29	77.29	76.90	78.36
OA	84.67	82.53	84.75	76.34	81.43	85.02	85.16	84.32	85.78
κ	69.80	65.69	69.31	59.29	63.29	70.45	71.13	69.61	71.54

注:粗体表示最佳结果。

3 讨论

3.1 消融实验

本文主要提出了两个模块,分别是FAEM和ACACL。为了进一步评估本文提出的各个模块对整体算法的贡献,本文在三个数据集上分别展开消融实验,实验结果如表5所示。其中,基线为仅使用Transformer提取特征并计算FSL损失。在此基础上逐渐加入FAEM和ACACL两个模块,每个数据集的第五行为完整模型的结果。

表5结果表明,仅使用FSL框架时模型整体性能仍较为有限,难以有效应对复杂地物变化。在此基础上加入FAEM后,三个数据集的OA和 κ 都有了一定程度的提升,尤其是China数据集,其OA提升了1.66%。该结果表明通过频域先验对双时相特征进行自适应校准,可以有效抑制由光照变化、气候扰动等因素引起的伪变化,从而获得更加可靠的变化特征表示。单独引入ACACL后,模型性能较基线也有一定提升,其中IA数据集上OA提升了1.32%,说明所提出的ACACL能够有效挖掘源域与目标域之间的潜在关联,缓解不同分布差异带来的域偏移。

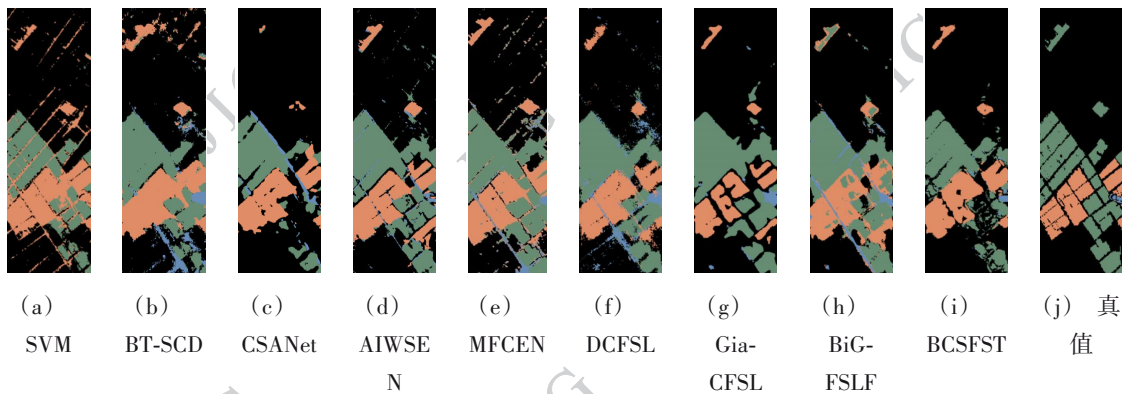


图9 China数据集检测结果示意图

Fig. 9 The detection map on China dataset. (a) SVM; (b) BT-SCD; (c) CSANet; (d) AIWSEN; (e) MFCEN; (f) DCFSL; (g) Gia-CFSL; (h) BiG-FSLF; (i) BCSFST; (j) Ground truth.

表5 3个数据集上的消融实验

Table 5 The ablation experiments on three datasets

数据集	FSL	FAEM	ACACL	OA	κ
IA	√			88.91	76.06
	√	√		89.81	77.51
	√		√	90.23	78.11
	√	√(固定 α)	√	90.43	78.55
	√	√	√	90.80	78.84
HZW	√			80.26	53.93
	√	√		81.57	55.65
	√		√	82.38	57.10
	√	√(固定 α)	√	81.45	54.84
	√	√	√	83.67	57.34
China	√			83.50	68.29
	√	√		85.16	70.68
	√		√	84.75	70.20
	√	√(固定 α)	√	84.28	69.14
	√	√	√	85.78	71.53

注:√表示使用该模块,粗体表示最佳结果。

进一步将 FAEM 和 ACACL 同时加入基线,模型在三个数据集上均取得了最优值,说明两者能够从频域特征增强和跨场景知识迁移两方面协同提升模型的检测能力。

此外,为了验证 FAEM 中的动态权重 α 的有效性,本文将 α 固定为 0.5 进行实验。从表 5 可以看出,固定权重后,模型在三个数据集上的性能均有一定下降,这是因为 FAEM 无法根据不同空间位置的

特征差异自适应调整幅度校正程度,可能在抑制伪变化的同时削弱有效变化信息,从而对最终性能产生一定影响。

综上所述,本文提出的两个模块均对模型性能有积极作用。在二者协同作用下,模型能够同时从频域特征优化和跨域特征对齐两个层面提升特征表达能力,进一步增强类内聚合和类间可分性,从而获得更优的检测性能。

3.2 计算时间和参数量

本节对 8 种方法的参数量及在 3 个数据集上的训练时间进行统计分析,其结果如表 6 所示。为保证实验结果的公平性和可比性,所有方法均在相同硬件环境下完成训练。

从表 6 可以看出,基于深度学习的方法整体具有较高的训练效率,尤其是 BT-SCD,在训练样本较少的情况下仅需几十秒即可完成训练。与此同时,部分方法由于模型设计融入注意力机制或特征融合结构,在提升性能的同时带来了较大的参数开销。而基于 FSL 的方法通常引入了元学习、跨域迁移等额外策略,需要计算额外的优化损失,导致训练时间普遍偏长。相较而言,本文提出的 BCSFST 方法计算时间较短,与部分结构复杂的深度学习方法(如 CSANet 和 AIWSEN)相比,其在计算开销相近甚至更低的情况下取得了更优的检测性能,在计算时间、参数量及检测精度上实现了较好的平衡。

3.3 不同样本数量的影响

为了进一步验证各类方法在不同少样本条件下的适应能力,本节分别选取目标域标记样本为每类 1-5 个开展实验,并对不同方法在不同规模下的检测

性能进行对比分析,实验结果如图 10 所示。

如图 10 所示,随着标记样本数量逐渐增加,大部分方法的检测性能呈现逐步提升的趋势,且传统

方法和基于深度学习的方法性能普遍低于基于 FSL 方法,说明 FSL 能够更有效地利用有限信息,更适用于小样本条件下的跨场景变化检测任务。由于目

表 6 不同算法的计算时间和参数量分析
Table 6 The analysis of computational time and parameter of different methods

数据集	BT-SCD (Tang 等, 2025)	CSANet (Song 等, 2022)	AIWSEN (Wu 等, 2025)	MFCEN (Ding 等, 2024)	DCFSL (Li 等, 2021)	BiG-FSLF (Wang 等, 2023)	Gia-CFSL (Zhang 等, 2024)	BCSFST (本文)
参数量(M)	25.44	40.42	12.23	153.78	4.25	0.65	0.85	0.42
训练时间 (秒)	IA	48.83	367.70	578.59	126.06	8289.45	828.03	8727.68
	HZW	56.48	520.01	530.27	313.28	7171.98	1269.72	6944.87
	China	47.95	507.61	492.85	167.18	5737.52	557.75	5203.85

标域标记样本数量过少,且 HZW 数据集场景较为复杂,少数方法在不同 shot 设置下存在一定波动。相比之下,本文所提出的方法在所有情形下均取得最优性能,在不同少样本条件下均保持了较好的检测性能和适应能力。

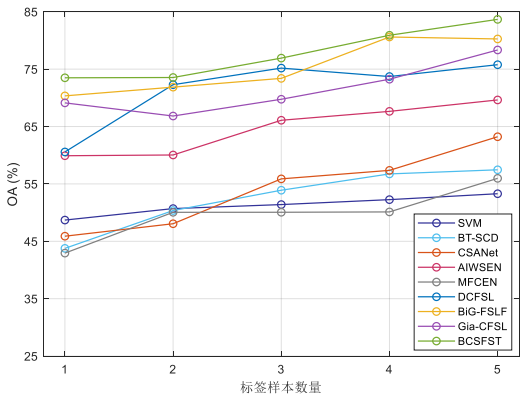


图 10 HZW 数据集上不同标记样本数量的 OA
Fig. 10 The OA of different labeled samples on the HZW dataset

4 结 论

针对跨场景高光谱变化检测中存在的伪变化干扰和域分布偏移问题,本文提出了一种面向小样本高光谱多类变化检测的双向对比与空频协同 Transformer 网络 (BCSFST),将小样本学习、频域先验建模及对比学习有机融入到 Transformer 框架中。对于双时相特征,本文引入频域信息,通过幅度和相位特

征对双时相图像进行自适应校正并构建频域先验,引导 Transformer 网络学习到更具判别性的特征。在此基础上,通过双向一致性对比学习来挖掘源域和目标域特征空间中的潜在对应关系,实现跨场景特征对齐,有效缓解了域偏移带来的影响。在 IA、HZW 和 China 三个高光谱多分类变化检测数据集的实验结果表明,本文提出的 BCSFST 在标记样本有限的条件下取得了优于现有先进方法的检测性能。未来,将考虑探索无监督情形下的高光谱变化检测方法,进一步降低模型对标记样本的依赖,提高模型在复杂遥感场景中的泛化能力和实际应用价值。

参考文献 (References)

Bovolo F, Bruzzone L and Marconcini M. 2008. A novel approach to unsupervised change detection based on a semisupervised SVM and a similarity measure. IEEE Transactions on Geoscience and Remote Sensing, 46(7): 2070-2082 [DOI: 10.1109/TGRS.2008.916643]

Ding C, Chen J Y, Zheng M M, Zhang L, Wei W and Zhang Y N. 2024. Survey of hyperspectral image change detection method. Journal of Image and Graphics, 29(06): 1714-1729 (丁晨, 陈静怡, 郑萌萌, 张磊, 魏巍, 张艳宁. 2024. 高光谱图像变化检测技术研究进展. 中国图象图形学报, 29(06): 1714-1729) [DOI: 10.11834/jig.240031]

Ding J G, Li X R, Xiang S and Chen S H. 2024. Multilevel features fused and change information enhanced neural network for hyperspectral image change detection. IEEE Transactions on Geoscience and Remote Sensing, 62, Art. No. 5502413 [DOI: 10.1109/TGRS.2023.3344784]

- Feng S, Zhang J H, Fan Y Z, Liu X Y, Zhao C H and Li W. 2025. Cross-domain few-shot learning method based on fractional domain information for hyperspectral image multi-class change detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(12): 12680-12691 [DOI: 10.1109/TCSVT.2025.3570466]
- Habib T, Inglada J, Mercier G and Chanussot J. 2009. Support vector reduction in SVM algorithm for abrupt change detection in remote sensing. *IEEE Geoscience and Remote Sensing Letters*, 6(3): 606-610 [DOI: 10.1109/LGRS.2009.2020306]
- Huang Y X, Zhang L F, Qi W C, Song R X, Huang C P and Cen Y. 2024. Semi-supervised hypermatch-driven cross temporal and spatial interaction transformer for hyperspectral change detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 6426-6443 [DOI: 10.1109/JSTARS.2024.3370151]
- Li S Y, Xie X H, Zhao X Y and Wang X H. 2026. HGA2-FSL: A heterogeneous graph aware-aggregation driven few-shot learning network for hyperspectral image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 64, Art. No. 5508314 [DOI: 10.1109/TGRS.2026.3673671]
- Li W, Chen C, Zhang M M, Li H C and Du Q. 2018. Data augmentation for hyperspectral image classification with deep CNN. *IEEE Geoscience and Remote Sensing Letters*, 16(4): 593-597 [DOI: 10.1109/LGRS.2018.2878773]
- Li X T, Gao F, Sun Y and Dong J Y. 2025. Hybrid attention and bidirectional-gated network for hyperspectral image change detection. *Journal of Image and Graphics*, 30(04): 1017-1026 (李相潭, 高峰, 孙悦, 董军宇. 2025. 针对高光谱遥感图像变化检测的混合注意力和双向门控网络. *中国图象图形学报*, 30(04): 1017-1026) [DOI: 10.11834/jig.240360]
- Li Z K, Liu M, Chen Y S, Xu Y M, Li W and Du Q. 2021. Deep cross-domain few-shot learning for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60, Art. No. 5501618 [DOI: 10.1109/TGRS.2021.3057066]
- Licciardi G, Marpu P R, Chanussot J and Benediktsson J A. 2011. Linear versus nonlinear PCA for the classification of hyperspectral data based on the extended morphological profiles. *IEEE Geoscience and Remote Sensing Letters*, 9(3): 447-451 [DOI: 10.1109/LGRS.2011.2172185]
- Liu B, Yu A Z, Yu X C, Wang R R, Gao K L and Guo W Y. 2020. Deep multiview learning for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(9): 7758-7772 [DOI: 10.1109/TGRS.2020.3034133]
- Luo F L, Zhou T Y, Liu J M, Guo T, Gong X W and Ren X W. 2023. Multiscale diff-changed feature fusion network for hyperspectral image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61, Art. No. 5502713 [DOI: 10.1109/TGRS.2023.3241097]
- Lv Z Y, Zhang M, Sun W W, Benediktsson J, Lei T and Falco N. 2023. Spatial-contextual information utilization framework for land cover change detection with hyperspectral remote sensed images. *IEEE Transactions on Geoscience and Remote Sensing*, 61, Art. No. 4411911 [DOI: 10.1109/TGRS.2023.3336791]
- Malila W A. 1980. Change vector analysis: An approach for detecting forest changes with Landsat/LARS symposia. 385.
- Plaza A, Benediktsson J A, Boardman J W, Brazile J, Bruzzone L and Camps-Valls G. 2009. Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment*, 113: S110-S122 [DOI: 10.1016/j.rse.2007.07.028]
- Qu J H, Xu Y S, Dong W Q, Li Y S and Du Q. 2022. Dual-branch difference amplification graph convolutional network for hyperspectral image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60, Art. No. 5519912 [DOI: 10.1109/TGRS.2021.3135567]
- Rodarmel C and Shan J. 2002. Principal component analysis for hyperspectral image classification. *Surveying and Land Information Science*, 62(2): 115-122.
- Samaniego L, Bárdossy A, Schulz K. 2008. Supervised classification of remotely sensed imagery using a modified k-nn technique. *IEEE Transactions on Geoscience and Remote Sensing*, 46(7): 2112-2125 [DOI: 10.1109/TGRS.2008.916629]
- Song R X, Ni W H, Cheng W and Wang X H. 2022. CSANet: Cross-temporal interaction symmetric attention network for hyperspectral image change detection. *IEEE Geoscience and Remote Sensing Letters*, 19, Art. No. 6010105 [DOI: 10.1109/LGRS.2022.3179134]
- Sun L, Shi B C, Zheng Y H, Xiao L and Cao Y L. 2026. SFIEET: spectral frequency-induced edge enhancement transformer for hyperspectral change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 64, Art. No. 5501515 [DOI: 10.1109/TGRS.2026.3651827]
- Sun Q R, Liu Y Y, Chua T-S and Schiele B. 2019. Meta-transfer learning for few-shot learning//*Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*. Long Beach: 403-412 [DOI: 10.1109/CVPR.2019.00049]
- Tang Y J, Feng S, Zhao C H, Chen Y Q, Lv Z Y and Sun W W. 2025. A semantic change detection network based on boundary detection and task interaction for high-resolution remote sensing images. *IEEE Transactions on Neural Networks and Learning Systems*, 36(9): 17184-17198 [DOI: 10.1109/TNNLS.2025.3570425]
- Tu L L, Huang X, Li J, Yang J Y and Gong J Y. 2024. A multi-task learning method for extraction of newly constructed areas based on bi-temporal hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 208: 308-323 [DOI: https://doi.org/10.1016/j.isprsjprs.2024.01.016]
- Vinyals O, Blundell C, Lillicrap T, Kavukcuoglu K and Wierstra D. 2026. Matching networks for one shot learning// *Proceedings of the Advances in Neural Information Processing Systems*, 29: 3637 - 3645.

- Wang W N, Liu X B and Mou X Q. 2021. Data augmentation and spectral structure features for limited samples hyperspectral classification. *Remote Sensing*, 13(4): 547 [DOI: <https://doi.org/10.3390/rs13040547>]
- Wang X H, Li S Y, Zhao X Y and Zhao K Y. 2023. BiG-FSLF: A cross heterogeneous domain few-shot learning framework based on bidirectional generation for hyperspectral image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61, Art. No. 5516213 [DOI: 10.1109/TGRS.2023.3292249]
- Wang X, Li Y Y and Zhang X L. 2024. Ghost-guided UNet++ for high-resolution remote sensing image change detection. *Journal of Image and Graphics*, 29(05): 1460-1478 (王鑫, 李莹莹, 张香梁. 2024. Ghost引导UNet++的高分遥感影像变化检测. *中国图象图形学报*, 29(05): 1460-1478) [DOI: 10.11834/jig.230212]
- Wang Z Y, Gao F, Dong J Y and Du Q. 2025. Global and local attention-based transformer for hyperspectral image change detection. *IEEE Geoscience and Remote Sensing Letters*, 22, Art.No. 5500405 [DOI: 10.1109/LGRS.2024.3505294]
- Wu L X, Peng J T, Yang B, Sun W W and Ye Z J. 2025. AIWSEN: Adaptive information weighting and synchronized enhancement network for hyperspectral change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 63, Art. No. 5503412 [DOI: 10.1109/TGRS.2025.3531478]
- Yang B, Sun W W and Peng J T. 2024. SAGN: Sharpening-aware graph network for hyperspectral image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 62, Art. No. 5518812 [DOI: 10.1109/TGRS.2024.3403971]
- Yang R, Xu X, Li X R, Wang L and Pu F L. 2020. Learning relation by graph neural network for SAR image few-shot learning// *Proceeding of the IEEE International Geoscience and Remote Sensing Symposium*. Waikoloa: 1743-1746 [DOI: 10.1109/IGARSS39084.2020.9323139]
- Yang Y and Soatto S. 2020. FDA: Fourier domain adaptation for semantic segmentation//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: 4085-4095 [DOI: 10.1109/CVPR42600.2020.00414]
- Zhang L F and Zhang L P. 2022. Artificial intelligence for remote sensing data analysis: A review of challenges and opportunities. *IEEE Geoscience and Remote Sensing Magazine*, 10(2): 270-294 [DOI: 10.1109/MGRS.2022.3145854]
- Zhang Y X, Li W, Jia W, Zhang M M, Tao R and Liang S L. 2025. Cross-domain hyperspectral image classification based on bi-directional domain adaptation. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(12): 12038-12051 [DOI: 10.1109/TCSVT.2025.3586282]
- Zhang Y X, Li W, Zhang M M and Tao R. 2025. Local Representation Few-Shot Learning for Hyperspectral Image Cross-Scene Classification. *Acta Electronica Sinica*, 53(1): 248-258 (张宇翔, 李伟, 张蒙蒙, 陶然. 2025. 基于局部表征少样本学习的高光谱图像跨场景分类. *电子学报*, 53(1): 248-258) [DOI: 10.12263/DZXB.20230937]
- Zhang Y X, Li W, Zhang M M, Han J W, Tao R and Liang S L. 2026. SpectralX: Parameter-efficient domain generalization for spectral remote sensing foundation models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 239: 774-792 [DOI: <https://doi.org/10.1016/j.isprsjprs.2026.06.008>]
- Zhang Y X, Li W, Zhang M M, Wang S, Tao R and Du Q. 2024. Graph information aggregation cross-domain few-shot learning for hyperspectral image classification. *IEEE Transactions on Neural Networks and Learning Systems*, 35(2): 1912-1925 [DOI: 10.1109/TNNLS.2022.3185795]
- Zheng Z, Zhong Y F, Wang J J, Ma A L and Zhang L P. 2021. Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters. *Remote Sensing of Environment*, 265: 112636 [DOI: 10.1016/j.rse.2021.112636]
- Zhou T Y, Liu J M, Guo T, Fu C and Luo F L. 2026. Joint central difference feature and spatial-spectral attention method for hyperspectral image change detection. *National Remote Sensing Bulletin*, 30(2): 358-370 (周天媛, 刘嘉敏, 郭坦, 付川, 罗甫林. 2026. 联合中心差异特征和空谱注意力的高光谱图像变化检测. *遥感学报*, 30(2): 358-370) [DOI: 10.11834/jrs.20255012]

作者简介

候文慧,女,博士研究生,主要研究方向为高光谱图像变化检测。E-mail: 320645980@qq.com.

彭江涛,通信作者,男,教授,主要研究方向为高光谱图像处理。E-mail: pengjt1982@126.com.

杨冰,男,讲师,主要研究方向为高光谱图像处理。E-mail: bingyang0517@163.com.

邓紫莹,女,硕士研究生,主要研究方向为高光谱图像变化检测。E-mail: 1459966243@qq.com.

孙伟伟,男,教授,主要研究方向为高光谱图像处理。E-mail: sunweiwei@nbu.edu.cn.