

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: Chen Yongqi, Feng Shou, Ren Chunying, Tang Yingjie, Ren Huixin, Zhao Chunhui. Frequency-enhanced hierarchical interaction network for forest change detection in multispectral remote sensing images [J/OL]. Journal of Image and Graphics, XXXX: 1-16. DOI: 10.11834/jig.260051. (陈勇奇, 冯收, 任春颖, 唐英杰, 任慧鑫, 赵春晖. 面向多光谱遥感图像林地变化检测的频域增强分层交互网络[J/OL]. 中国图象图形学报, XXXX: 1-16. DOI: 10.11834/jig.260051.) [DOI: 10.11834/jig.260051]

面向多光谱遥感图像林地变化检测的频域增强分层交互网络

陈勇奇¹, 冯收¹, 任春颖², 唐英杰¹, 任慧鑫², 赵春晖¹

1. 哈尔滨工程大学信息与通信工程学院, 哈尔滨 150001; 2. 中国科学院东北地理与农业生态研究所, 长春 130102

摘要: 目的 多光谱遥感图像林地变化检测是生态监测领域的重要研究方向之一。针对双时相图像中成像条件差异引起的伪变化, 以及复杂地物变化在尺度与形态上的多样性问题, 设计了一种频域增强分层交互网络, 以提升复杂场景下林地变化检测的准确性与鲁棒性。方法 首先, 构建频域特征增强模块, 在频域对双时相特征进行增强, 以抑制由成像条件不一致引起的风格差异等伪变化干扰。其次, 设计跨时相相关引导交互模块, 通过对双时相特征的充分交互与相关性引导, 增强网络对局部细微变化区域的识别能力。最后, 通过空-谱耦合交互模块, 对不同网络深度的变化特征进行整合, 融合浅层细粒度结构信息与深层语义信息, 实现对多尺度、多形态林地变化对象的联合感知。结果 实验在3个林地变化检测数据集上与 Spectral-Former、CSANet (cross-temporal interaction symmetric attention network)、TriTF (triplet Transformer framework)、DIEFEN (differential information-enhanced feature exchange network)、FrFTML (fractional Fourier transform domain decoupling-based adaptive weighted metric learning network) 和 AIWSEN (Adaptive information weighting and synchronized enhancement network) 等先进方法进行了对比分析。在 R1 数据集中, 所提出方法的 Kappa 系数达到 0.8102, 相比性能次优方法提升 1.09%; 在 R2 数据集中, Kappa 系数提升 2.62%; 在 R3 数据集中, Kappa 系数提升 1.74%。同时, 在三个数据集上的消融实验进一步验证了所提模块的有效性。结论 本文所提方法能有效缓解双时相风格差异引起的伪变化干扰, 并提升对细微变化与多尺度变化对象的检测能力, 从而获得更为准确、稳定的林地变化检测结果。

关键词: 多光谱遥感图像; 林地变化检测; 频域特征; 特征交互; 多尺度变化感知

Frequency-enhanced hierarchical interaction network for forest change detection in multispectral remote sensing images

Chen Yongqi¹, Feng Shou¹, Ren Chunying², Tang Yingjie¹, Ren Huixin², Zhao Chunhui¹

1. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China; 2. Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China

Abstract: **Objective** Multispectral remote sensing forest image change detection aims to identify temporal variations in land-cover types or surface conditions by comparing multispectral images acquired over the same area at different times.

收稿日期: 2026-01-23; 修回日期: 2026-06-25

基金项目: 国家自然科学基金项目(62471155); 国家自然科学基金项目(62371153); 海南省自然科学基金项目(626YXQN0951); 自然资源部国土卫星遥感应用重点实验室开放基金(KLSMNR-G202503)

Supported by: National Natural Science Foundation of China (62471155); National Natural Science Foundation of China (62371153); Hainan Provincial Natural Science Foundation of China (626YXQN0951); Open Fund of the Key Laboratory of Land Satellite Remote Sensing Application, Ministry of Natural Resources (KLSMNR-G202503)

This technique has been widely applied in ecological monitoring and natural resource management, among which forest change detection is of particular importance for tracking forest cover dynamics, assessing ecosystem function changes, and supporting forest resource conservation. Multispectral remote sensing images, acquired through observations across multiple discrete spectral bands, can effectively characterize spectral response differences associated with changes in land-cover types and surface conditions. Meanwhile, they provide large-area coverage and periodic observations with controllable data acquisition costs, thereby offering stable and discriminative data support for long-term and continuous forest change monitoring. Consequently, multispectral remote sensing image-based change detection has become an essential technical approach for forest resource monitoring. In early studies, change detection was mainly conducted using traditional image processing and pattern recognition techniques, where changes were modeled based on manually designed features and predefined thresholds. Although these methods are relatively simple and interpretable, they are highly sensitive to noise, illumination variations, and imaging condition discrepancies, which limits their reliability in complex scenarios. With the rapid advancement of deep learning, convolutional neural network (CNN)-based methods have become the dominant paradigm for remote sensing change detection due to their strong capability in automatic feature representation and end-to-end learning. Siamese architectures further improve detection performance by explicitly modeling feature differences between bi-temporal images. However, the fixed receptive fields of CNNs restrict their ability to capture long-range dependencies and global contextual information. To alleviate this limitation, attention mechanisms have been introduced to emphasize change-relevant features across spatial and channel dimensions. More recently, Transformer-based models have been applied to change detection, benefiting from self-attention mechanisms to model global dependencies and complex change patterns. Despite these advances, deep learning-based methods still face significant challenges in forest change detection. Bi-temporal remote sensing images often exhibit pronounced style discrepancies caused by seasonal variations, illumination conditions, and sensor differences, which introduce substantial pseudo-change interference. Moreover, forest changes are characterized by high spatial heterogeneity, ranging from subtle local variations to large-scale degradation or recovery processes, resulting in complex multi-scale and multi-pattern change characteristics. These factors make it particularly challenging to suppress style-induced pseudo changes while accurately perceiving and localizing change targets at different spatial scales. To address these challenges, this paper proposes a frequency-enhanced hierarchical interaction network for accurate forest change detection. **Method** This paper proposes a frequency-enhanced hierarchical interaction network for forest change detection in remote sensing imagery. Bi-temporal multispectral images are taken as input and first processed by shared-weight spatial-spectral feature extraction modules to obtain multi-level spatial-spectral representations. To optimize the global representation of bi-temporal features, a frequency feature enhancement module (FFEM) is designed to enhance bi-temporal feature representations. This module performs adaptive modulation in the frequency domain to suppress low-frequency style interference and emphasize mid- and high-frequency structural information, thereby alleviating pseudo-change artifacts caused by imaging condition discrepancies. Subsequently, a cross-temporal correlation-guided interaction module (CCIM) is introduced to explicitly model temporal correlations between bi-temporal images and extract discriminative change features. By leveraging feature interaction and correlation guidance, this module captures temporal dependencies, enabling the extraction of stable change representations and improving the network's sensitivity to subtle forest changes. In addition, a spatial-spectral coupled interaction module (SCIM) is employed to fully exploit the complementary information among change features at different hierarchical levels. This module jointly models spatial and channel-wise dependencies by integrating fine-grained structural details with high-level semantic representations, thereby enhancing the perception of multi-scale forest change regions. Finally, the fused change features are fed into a classification head to generate the change detection results. **Result** Comparison experiments were conducted on three forest change detection datasets to compare the proposed method with six advanced approaches, including Spectral-Former, CSANet, TriTF, DIEFEN, FrFTML, and AIWSEN. For each dataset, 10% of the samples were randomly selected for training, while the remaining samples were used for testing. Performance was evaluated using five widely adopted metrics for change detection tasks, including overall accuracy (OA), Kappa coefficient (KC), precision (Pre), recall (Rec), and F1-score (F1). The proposed method achieves Kappa coefficients of 0.8102, 0.7355, and 0.7635 on the three datasets, respectively. Compared with the second-best method, it yields performance gains of 1.09%, 2.62%, and 1.74%. In addition,

ablation studies conducted on all three datasets further verify the effectiveness of the proposed modules. The ablation experiments indicate that the frequency feature enhancement module, the cross-temporal correlation-guided interaction module, and the spatial-spectral coupled interaction module play key roles in improving change detection accuracy by effectively suppressing pseudo-changes and enhancing the perception of subtle and multi-scale forest changes. The proposed method consistently outperforms mainstream approaches across all datasets. These results highlight the advantage of the proposed design in achieving accurate and stable forest change detection. **Conclusion** This paper proposes a frequency-enhanced hierarchical interaction network for forest change detection in remote sensing imagery. By integrating bi-temporal frequency-domain feature enhancement, cross-temporal change extraction, and collaborative interaction among multi-level change features, the proposed method effectively improves forest change detection performance. Experimental results show that the proposed approach consistently outperforms existing mainstream methods across multiple forest change detection datasets, demonstrating strong consistency and stability. Benefiting from the synergistic design of frequency enhancement and hierarchical interaction, the method effectively suppresses style interference caused by bi-temporal imaging condition discrepancies, leading to a significant reduction in pseudo-change false alarms. Meanwhile, in complex forest scenarios, the proposed approach exhibits enhanced capability in perceiving and localizing subtle change regions as well as change objects at different spatial scales, enabling accurate identification of both local structural variations and large-scale evolutionary processes. These results indicate that the proposed method achieves strong robustness and applicability under complex remote sensing imaging conditions, providing a reliable solution for forest change monitoring and related applications.

Key words: multispectral remote sensing images; forest change detection; frequency features; feature interaction; multi-scale change perception

论文引用格式: Chen Yongqi, Feng Shou, Ren Chunying, Tang Yingjie, Ren Huixin, Zhao Chunhui. Frequency-enhanced Hierarchical Interaction Network for Forest Change Detection in Multispectral Remote Sensing Images — SCID[J/OL]. Journal of Image and Graphics. DOI: 10.11834/jig.260051. (陈勇奇, 冯收, 任春颖, 唐英杰, 任慧鑫, 赵春晖. 面向多光谱遥感图像林地变化检测的频域增强分层交互网络—SCID[J/OL]. 中国图象图形学报. DOI: 10.11834/jig.260051.)

0 引言

多光谱遥感图像通过在多个离散光谱波段对地表进行观测,能够有效刻画不同地表覆盖类型及其状态变化在光谱响应上的差异,并在保证数据获取成本可控的同时,具备大范围、周期性观测能力,为森林变化的长期、连续监测提供了稳定且具有判别力的数据基础。多光谱遥感图像变化检测旨在通过对同一区域在不同时期获取的多光谱遥感图像进行对比分析,以识别地表覆盖类型或地物状态随时间发生的变化(Wu等,2024;贾萌等,2024;王鑫等,2023)。基于多光谱遥感图像的变化检测已成为森

林资源监测的重要手段,对于掌握森林覆盖动态、评估生态系统功能变化以及支撑林业资源保护具有重要意义(Liu等,2026)。

早期的变化检测主要依赖于人工目视解译或基于人工设计的特征,通过变化向量分析(change vector analysis, CVA),图像差分、主成分分析(principal component analysis, PCA)等方式对时相间图像的差异进行建模(Miao等,2024)。这类方法虽然实现简单、可解释性强,但对噪声、光照变化及成像条件差异高度敏感,难以在复杂场景下稳定识别真实变化(王仁芳等,2024;王艳恒等,2020)。

近年来,深度学习方法凭借其强大的特征表征能力,已成为遥感变化检测的主流研究方向。现有方法主要包括单流与双流两类框架。其中,全卷积早期融合(fully convolutional early fusion, FC-EF)通过将双时相图像在通道维度直接拼接,实现单流端到端变化检测(Daudt等,2018)。然而,该方法忽略了时相间的相关性建模,导致变化区域定位能力受限。相比之下,双流框架基于孪生网络对双时相图像进行特征提取,更有利于建模其相似性与差异(Miao等,2024),因而被广泛采用。例如,Zhao等人(2023)通过语义引导缓解特征融合干扰;Liu等人(2024)结合空间细节与光谱差异信息提升检测性

能;Miao等人(2024)利用全尺度跳跃连接融合多层特征;Wang等人(2024)通过跨层特征交互在降低计算复杂度的同时提高检测精度。然而,受限于卷积神经网络(convolutional neural networks, CNN)的局部感受野,这类方法在建模长程依赖和全局变化信息方面仍存在不足。

为克服上述挑战,研究人员通过引入注意力机制,引导模型关注与变化相关的关键信息,并有效捕捉特征之间的长程依赖关系,从而提升对复杂变化模式的判别能力。Lv等人(2022)通过空间-光谱注意力增强特征表达;Guo等人(2024)利用相似性引导注意力提升变化特征的判别能力;Xu等人(2023)结合注意力金字塔与通道交叉机制挖掘多维变化信息;Li等人(2022)通过混合注意力模块整合上下文信息。在上述研究基础上,研究人员进一步引入具备全局依赖建模能力的自注意力结构Transformer,将其应用于变化检测任务,并在复杂场景中取得了良好的检测效果。Chang等人(2024)通过设计基于Transformer轻量化变化检测方法,实现了对长程依赖和关键变化语义的感知。Chen等人(2022)利用Transformer建模长距离依赖以改善边界与小目标检测;Jiang等人(2023)利用Transformer学习背景一致性表示增强变化识别能力。

虽然上述方法在遥感图像变化检测领域已取得了一定进展,但在面向实际复杂场景的林地变化检测任务时仍面临诸多挑战。具体而言,受成像季节、光照条件以及传感器成像条件差异等因素影响,不同时相图像在亮度、颜色等方面往往存在显著差异,从而导致整体图像风格呈现出明显的不一致性。这类由成像条件不一致引入的伪变化并不对应真实地物变化,却容易在变化检测过程中被误识别为变化区域。另一方面,林地变化形式既包括小尺度的局部变化,也包括大范围的退化或恢复等整体演化过程,呈现出多尺度、多形态并存的空间异质性特点,导致方法难以对不同尺度变化目标的有效感知与精准定位。因此,在有效抑制由成像条件差异引入的伪变化干扰的同时,实现对不同尺度林地变化目标的准确感知与定位,是一项具有挑战性的任务。

针对上述问题,本文提出了一种多光谱遥感图像频域增强分层交互网络(frequency-enhanced hierarchical interaction network, FHI-Net),该方法从频域增强与分层特征交互两个角度对双时相变化特征

进行协同建模。首先,该方法在频域特征中抑制由成像条件差异引入的风格干扰以获得更加稳定的特征表示。随后,通过跨时相相关引导交互突出细微真实的变化区域,得到多层级变化特征。最后,通过空-谱耦合交互融合所提取到的不同层级变化特征,从而实现对多尺度、多形态林地变化对象的准确感知。主要贡献为:1)首次将频域自适应调制引入林地变化检测任务,设计了频域特征增强模块(frequency feature enhancement module, FFEM)。通过抑制低频风格干扰并强化中高频结构信息,为后续提取变化信息提供更加稳定的特征表示。2)为了充分建模双时相图像的时相相关性并获得变化特征,提出了跨时相相关引导交互模块(cross-temporal correlation-guided interaction module, CCIM)。该模块通过建模双时相特征之间的相关关系,增强网络对细粒度变化信息的感知能力。3)为了聚合不同层次变化特征之间丰富的上下文信息,构建了空-谱耦合交互模块(spatial-spectral coupled interaction module, SCIM)。该模块通过对拼接后的多层变化特征进行空间与通道维度的交互融合,实现浅层结构信息与深层语义信息的有效整合,从而提升对多尺度、多形态林地变化对象的整体感知能力。

1 研究方法

本节首先介绍所提FHI-Net的整体算法框架,而后分别介绍频域特征增强模块FFEM、跨时相相关引导交互模块CCIM、空-谱耦合交互模块SCIM以及所使用的损失函数。

1.1 算法框架

如图1所示,FHI-Net以双时相多光谱图像块为输入,通过频域特征增强、跨时相相关引导交互、空-谱耦合注意力交互,实现对变化信息的精准提取与判别。此外,FHI-Net还包含3个共享权重的空-谱特征提取块,用于初步提取多光谱图像的空-谱特征。

具体而言,定义双时相多光谱图像分别为 $I^1 \in \mathbf{R}^{H \times W \times C_1}$ 和 $I^2 \in \mathbf{R}^{H \times W \times C_2}$ 。首先,FHI-Net以两个时相中空间位置相同的像素为中心提取输入图像块,分别记为 $X^1 \in \mathbf{R}^{p \times p \times C_1}$ 和 $X^2 \in \mathbf{R}^{p \times p \times C_2}$,其中 H, W, C 分别表示多光谱图像的高度、宽度和波段数, p 表示所提取图像块的尺寸。将双时相图像块 X^1 和 X^2 输入空-谱特征提取块中,以获得输出特征

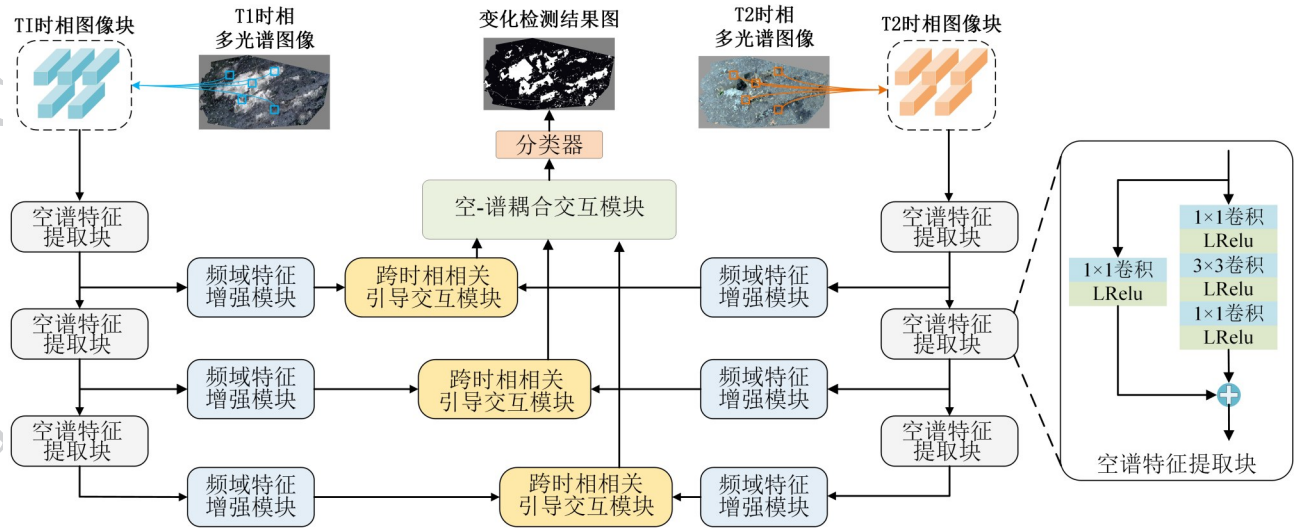


图1 频域增强分层交互网络框架图

Fig. 1 Architecture of the frequency-domain enhanced hierarchical interaction network

F_i^1 和 F_i^2 , 其中 $i \in \{1, 2, 3\}$ 代表第 i 层特征提取块的

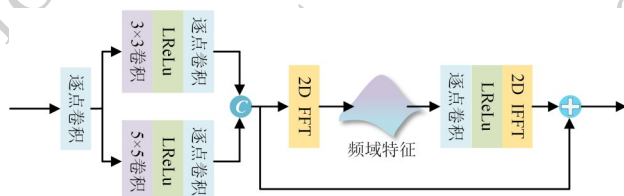


图2 频域特征增强模块

Fig. 2 Frequency feature enhancement module

输出。在各层特征编码后利用频域特征增强模块 FFEM 对单时相特征进行频域增强以减弱时相间的风格差异与噪声干扰。随后, 通过 CCIM 对增强后的双时相特征进行交互, 以捕获鲁棒的变化特征。为充分挖掘不同网络深度变化特征之间的上下文信息, 通过空-谱耦合交互模块 SCIM 在空间维度与通道维度上对多层变化特征进行联合交互与融合表达, 使浅层变化特征中的细粒度结构信息与深层变化特征中的高层语义信息得以充分融合, 从而提升网络对多尺度、多形态变化对象的整体感知能力。最终, 融合后的变化特征通过分类器得到最终的变化检测结果。

1.2 频域特征增强模块 FFEM

针对双时相特征中由成像条件差异引入的非判别性干扰, 本文设计了频域特征增强模块 FFEM。

该模块结合多尺度特征提取与频域调制, 在频率域对双时相特征进行自适应调制, 以抑制风格相关的冗余响应并增强结构性判别信息, 从而获得更

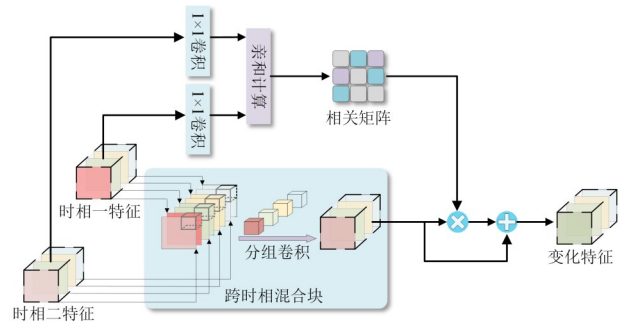


图3 跨时相相关引导交互模块

Fig. 3 Cross-temporal correlation-guided interaction module

加稳定的特征表示。

如图2所示, 以经过空-谱特征提取块所提取的一时相特征 F_i^1 为例, 首先通过点卷积对通道维度进行线性变换, 并将特征划分为两个分支以提取不同尺度的局部结构信息。具体为:

$$\hat{F}_i^1 = \text{PConv}(\sigma(\text{DConv}(\text{Split}(\text{PConv}(F_i^1)))))) \quad (1)$$

式中, $\text{DConv}(\cdot)$ 表示多尺度深度可分离卷积, $\text{PConv}(\cdot)$ 为点卷积, $\text{Split}(\cdot)$ 表示特征划分操作, σ 表示 GeLU 激活函数。

受成像条件变化的影响, 双时相图像往往在整体亮度与色调分布上表现出一定差异, 这类全局风格差异主要集中于特征的低频成分 (Huang 等, 2021; Yang 等, 2020)。相比之下, 反映地物几何结构变化的成分 (如边缘、轮廓) 主要分布于中高频频段, 具有更强的判别性 (Gao 等, 2025; Liu 等, 2022)。基于该特性, 本文引入离散傅里叶变换 (discrete

Fourier transform, DFT), 将空间域特征映射至频域, 并对不同频率分量进行自适应调制, 以削弱由成像条件变化引入的低频风格偏差, 同时强化中高频结构信息, 从而提升模型对真实地物变化的判别能力。相比其它频域变换方式, DFT 具有形式简单、计算效率高且易于在深度网络中实现等优势, 因此采用 DFT 完成特征的频域映射。具体为

$$\mathbf{F}_{\text{DFT},i}^1(u,v) = \frac{1}{\sqrt{HW}} \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} \hat{\mathbf{F}}_i^1(h,w) e^{-j2\pi\left(\frac{hu}{H} + \frac{vw}{W}\right)} \quad (2)$$

式中, $\mathbf{F}_{\text{DFT},i}^1$ 表示频域特征, (u,v) 表示频域坐标。

为进一步抑制低频风格扰动并增强与真实变化区域相关的判别性频率成分, 本文对频域特征的不同频率分量进行自适应调制, 具体为

$$\hat{\mathbf{F}}_{\text{DFT},i}^1 = \sigma\left(\text{BN}\left(\text{PConv}\left(\mathbf{F}_{\text{DFT},i}^1\right)\right)\right) \quad (3)$$

式中, $\hat{\mathbf{F}}_{\text{DFT},i}^1$ 表示更新过后的频域特征, BN 为批归一化层。该过程通过在频域引入可学习的参数, 不同频率分量的响应幅值可根据其对变化判别的贡献程度进行自适应调节, 从而在训练过程中抑制对与风格差异相关的低频成分, 并强化与真实变化相关的中高频结构信息。最后, 通过逆傅里叶变换将频域特征映射回空间域, 获得频域增强后的时相特征 $\tilde{\mathbf{F}}_i^1$, 具体为

$$\tilde{\mathbf{F}}_i^1 = \mathcal{F}^{-1}\left(\hat{\mathbf{F}}_{\text{DFT},i}^1\right) \quad (4)$$

如图 1 所示, FFEM 应用于网络的三层特征提取阶段, 对不同深度的双时相特征分别进行频域增强, 为后续提供更加稳定且判别性更强的特征表示。

1.3 跨时相相关引导交互模块 CCIM

在变化检测任务中, 充分挖掘双时相特征之间的时空依赖关系对于准确刻画细微变化具有重要意义。现有方法通常通过特征拼接或差分的方式来提取双时相差异, 从而实现变化区域的识别。然而, 当面对复杂的地物分布时, 上述方式难以充分建模不同时间相间的语义对应关系, 其对细微变化的刻画能力和特征表达的稳定性仍然受限。为此, 本文设计了一种跨时相相关引导交互模块 (CCIM), 通过建模双时相特征之间的相关关系, 并引导特征交互过程聚焦于具有一致语义对应的区域, 从而增强对细微变化的感知能力并提升变化特征的鲁棒性。

如图 3 所示, 首先通过将双时相特征的通道维度混合, 来建模时相间的时空依赖关系从而获得粗略的变化信息。具体而言, 对于经过频域增强后的

双时相特征 $\tilde{\mathbf{F}}_i^1$ 和 $\tilde{\mathbf{F}}_i^2$, 沿着通道维度进行交错得到混合特征, 具体为

$$\mathbf{F}_{\text{mix},i}(k) = \begin{cases} \tilde{\mathbf{F}}_i^1[k/2], & k \text{ 为偶数} \\ \tilde{\mathbf{F}}_i^2[k/2], & k \text{ 为奇数} \end{cases} \quad (5)$$

式中, $\mathbf{F}_{\text{mix},i}(k)$ 表示第 i 层的跨时相混合特征, k 为其通道维度的索引, $[\cdot]$ 表示向下取整操作。

$\mathbf{F}_{\text{mix},i}$ 中每一对相邻通道分别代表双时态图像的语义信息。基于此, 采用分组卷积实现双时相特征间的交互融合。具体为

$$\hat{\mathbf{F}}_{\text{mix},i} = \sigma\left(\text{BN}\left(\text{GConv}\left(\mathbf{F}_{\text{mix},i}\right)\right)\right) \quad (6)$$

式中, $\text{GConv}(\cdot)$ 表示分组卷积, 对于混合后的特征图, 分组卷积在每一对相邻的双时相通道上独立执行卷积计算, 从而在保持通道对应关系的同时, 提取双时相图像间的时空依赖关系。

为了进一步增强 $\hat{\mathbf{F}}_{\text{mix},i}$ 中包含的变化信息, CCIM 通过相关矩阵, 对双时相影像中不同空间位置之间的语义关联进行建模。具体而言, 首先分别对原始双时相特征进行线性映射。具体为

$$\mathbf{E}_i^1 = \phi(\mathbf{F}_i^1), \quad \mathbf{E}_i^2 = \varphi(\mathbf{F}_i^2) \quad (7)$$

式中, $\phi(\cdot)$ 和 $\varphi(\cdot)$ 表示线性映射操作, $\mathbf{E}_i^1$ 和 $\mathbf{E}_i^2$ 分别表示双时相的嵌入特征。

随后对 $\mathbf{E}_i^1$ 和 $\mathbf{E}_i^2$ 进行亲和计算得到相关矩阵, 具体为

$$\mathbf{A} = \text{Softmax}\left(\left(\mathbf{E}_i^2\right)^T \cdot \mathbf{E}_i^1\right) \quad (8)$$

式中, $\text{Softmax}(\cdot)$ 表示 Softmax 函数, 相关矩阵 $\mathbf{A}$ 表示不同时相、不同空间位置之间的相关性强度。

最终基于相关矩阵对混合特征 $\hat{\mathbf{F}}_{\text{mix},i}$ 进行相关引导的交互增强, 得到变化特征 $\mathbf{D}_i$, 具体为

$$\mathbf{D}_i = \hat{\mathbf{F}}_{\text{mix},i} + \hat{\mathbf{F}}_{\text{mix},i} \cdot \mathbf{A} \quad (9)$$

在该过程中, 相关矩阵 $\mathbf{A}$ 作为权重系数, 对混合特征进行加权聚合。相关矩阵 $\mathbf{A}$ 刻画了双时相特征之间在不同空间位置上的语义相似性。对于未变化区域, 由于双时相地物结构与光谱特征保持一致, 其对应特征通常具有较高的相似度, 因此在相关矩阵中表现为较大的响应权重。而在真实变化区域, 由于地物结构或光谱特征发生变化, 其特征相似度相对较低。通过利用相关矩阵对混合特征进行加权, 可以在语义一致区域内稳定地传播特征信息, 从而增强特征表示的稳定性并提升变化特征的判别能力。

1.4 空-谱耦合交互模块 SCIM

对于不同层级的变化特征,浅层特征具有丰富的边缘、纹理等细粒度结构信息,而深层特征则包含更强的语义信息与上下文信息(Xie 等, 2025)。仅依赖单一层级的变化特征往往难以同时兼顾小尺度细节变化与大尺度语义变化,进而影响对多尺度、多形态变化对象的准确感知。为此,本文构建空-谱耦合交互模块 SCIM,在融合多层变化特征的基础上,从空间维度与通道维度引导不同层级变化信息的交互与互补融合,弥合细粒度结构信息与深层语义信息间的语义鸿沟,从而增强网络对复杂变化对象的整体表征能力(Yu 等, 2025)。

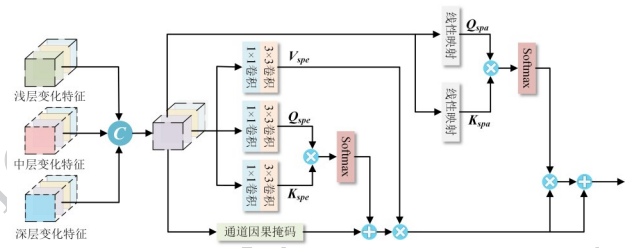


图4 空-谱耦合交互模块

Fig. 4 Spatial-spectral coupled interaction module

如图4所示,SCIM首先将多层级变化特征沿通道维度进行拼接得到 F_c ,并将其作为空-谱耦合交互模块的输入。为实现空-谱维度的协同交互,SCIM首先通过自适应投影层对输入特征进行线性映射,生成谱分支的查询(Query)、键(Key)和值(Value)特征,具体为

$$Q_{spe} = W_Q \cdot F_c \quad K_{spe} = W_K \cdot F_c \quad V_{spe} = W_V \cdot F_c \quad (10)$$

式中, W_Q , W_K 和 W_V 表示卷积映射层。

为了在挖掘不同层级变化特征间的谱间(通道)关系的同时,减少冗余干扰,设计通道因果掩码约束通道依赖范围并强化通道间的关联,具体为

$$M = \begin{cases} \gamma^{|m-n|}, & |m-n| \leq T \\ \varepsilon, & |m-n| > T \end{cases} \quad (11)$$

式中, m 和 n 为通道索引, T 为控制有效关联通道

范围的因果关联阈值, $\gamma \in (0, 1)$ 为关联衰减因子, ε 为非零的极小值。由于该掩码作用于特征通道之间的依赖建模,相邻通道通常具有更强的语义相关性,而远距离通道的相关性相对较弱。因此,采用局部通道依赖建模策略,将因果关联阈值 T 设置为4,以保留局部通道之间的有效交互。同时设置衰减因子 γ 为0.8,使通道权重随通道距离指数衰

减,从而在强化相关通道关系的同时抑制冗余特征交互(Yu 等, 2025)。最终,通道交互的输出 F_{spe} 可表示为

$$A_{spe} = \text{Softmax} \left(\frac{Q_{spe} K_{spe}^T}{\varepsilon} \odot M \right) \quad (12)$$

$$F_{spe} = A_{spe} \cdot V_{spe} \quad (13)$$

式中, $\odot$ 表示逐元素乘法。

为了挖掘变化区域在空间维度上的关联关系,如图4所示,对输入特征进行线性映射生成空间分支的 Q_{spa} 和 K_{spa} ,并将计算得到空间注意力矩阵,具体为

$$A_{spa} = \text{Softmax} \left(\frac{Q_{spa} K_{spa}^T}{\sqrt{C}} \right) \quad (14)$$

式中, $\sqrt{C}$ 为缩放因子。

为实现空间结构信息与通道间关联的交互融合,利用空间注意力矩阵对通道特征进行加权,以引导谱间变化信息在空间维度上的分布,从而强化变化区域的一致性。具体为

$$\hat{F}_c = F_{spe} \cdot A_{spa} + F_{spe} \quad (15)$$

最终,将得到的 $\hat{F}_c$ 送入分类器得到最终的预测结果。

1.5 损失函数

本文采用交叉熵损失函数作为模型训练的损失函数,引导模型有效区分变化与非变化区域,其可表示为

$$\mathcal{L}_{BCE} = -\frac{1}{m} \sum_{i=1}^m [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (16)$$

式中, m 为样本数量, y_i 为第 i 个样本的真实变化标签, $\hat{y}_i$ 为第 i 个样本的预测结果。

2 实验设置及结果分析

本节首先介绍用于实验验证的三个多光谱林地变化检测数据集,随后详细说明实验设置与评价指标。最后,展示并分析了所提FHI-Net的消融实验结果以及与现有先进方法的对比实验结果。

2.1 数据集介绍

本文选取三组林地多光谱变化检测数据集进行实验验证。图5给出了这3组数据集的双时相图 and 变化检测真值图。三组数据集的变化标签由具备专业知识的遥感解译人员基于双时相影像进行人工标

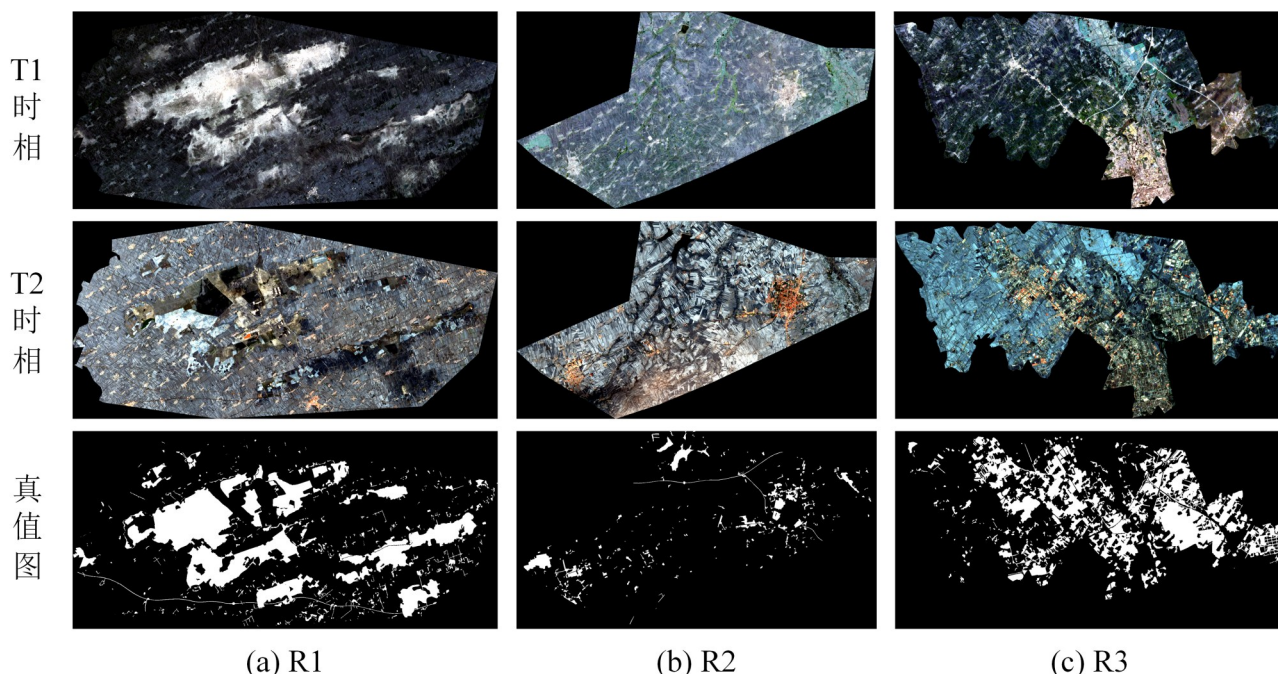


图5 三组林地多光谱变化检测数据集

Fig. 5 Three sets of forest land multi-spectral change detection datasets ((a) R1; (b) R2; (c) R3)

注,以保证变化区域标注的准确性。在真值图中变化区域被记为255,未变化区域被标记为0。考虑到森林光谱特征存在年内变化,本文在数据构建过程中对影像获取时间进行了严格控制,选取9月至10月期间的影像数据,以保证双时相影像处于相同季节与生长周期,从而减少由物候变化引起的光谱差异干扰。

1)R1:由Landsat卫星对扶余市西部采集获得,影像尺寸为1850×861像素,空间分辨率为30m。其中,T1时相影像获取于2000年9月26日,为7波段多光谱影像;T2时相影像获取于2020年10月19日,为4波段多光谱影像。该区域包含变化类样本273,342个,未变化类样本967,661个,记录了盐碱地向林地、旱地向林地、草地向林地转化的变化过程,变化面积较大且分布相对集中,变化区域具有较强的空间连通性。

2)R2:由Landsat卫星对榆树市中西部采集获得,影像尺寸为1641×901像素,空间分辨率为30m。其中,T1时相影像获取于2000年9月26日,为7波段多光谱影像;T2时相影像获取于2020年10月19日,为4波段多光谱影像。该区域包含变化类样本44,511个,未变化类样本387,376个,记录了水田向林地、旱地向林地转化的变化过程,其变化区域

整体规模较小,空间分布呈现出明显的离散化与碎片化特征。3)R3:由Landsat卫星对长春市宽城区南部采集获得,影像尺寸为1435×729像素,空间分辨率为30m。其中,T1时相影像获取于2000年9月26日,为7波段多光谱影像;T2时相影像获取于2020年10月19日,为4波段多光谱影像。该区域包含变化类样本168,166个,未变化类样本387,376个,记录了旱地向林地、林地向建设用地等人为地类转化的变化过程,其变化区域在空间上呈现出多尺度共存、形态多样的分布特征。

此外,为保证后续网络训练过程中在通道维度上的一致性,同时避免对原始光谱信息进行人为重构,对T2时相影像的最后一个波段进行通道复制扩展,使其在输入阶段与T1时相影像在通道数上保持一致。该策略在不引入额外光谱信息的前提下,有效支持了双时相异构波段条件下的变化检测建模。

2.2 实验设置及评价指标

所提网络基于Pytorch框架实现,通过单张NVIDIA GeForce RTX 3090显卡进行训练和测试。在训练阶段,采用Adam优化器,初始学习率为0.001,共训练200轮,在训练的第100、130和160轮时将学习率减半。此外,引入了L2正则化,系数为0.0005,以防止过拟合。批处理大小设置3为64。

表 1 FHI-Net在三个数据集上的消融实验结果
Table 1 Ablation experimental results of FHI-Net on three datasets

							/%
数据集	网络	OA	KC	F1	Pre	Rec	
R1	基线	90.94	0.7577	81.58	74.85	89.65	
	基线+FFEM	91.73	0.7782	83.07	77.18	89.94	
	基线+CCIM	91.97	0.7829	83.50	76.61	91.75	
	基线+SCIM	91.22	0.7623	81.96	74.81	90.63	
	FFEM+CCIM	92.73	0.7961	84.48	77.85	92.36	
	FFEM+CCIM+SCIM	93.04	0.8102	85.56	78.81	93.57	
R2	基线	94.37	0.6255	65.36	49.21	97.27	
	基线+FFEM	95.23	0.6668	69.12	53.28	98.38	
	基线+CCIM	95.47	0.6789	70.22	54.66	98.16	
	基线+SCIM	95.51	0.6892	71.16	55.72	98.44	
	FFEM+CCIM	96.01	0.7072	72.81	57.50	99.23	
	FFEM+CCIM+SCIM	96.53	0.7355	75.41	60.71	99.50	
R3	基线	83.71	0.6458	76.56	68.46	86.85	
	基线+FFEM	85.33	0.6636	78.08	70.48	87.53	
	基线+CCIM	86.72	0.6988	79.87	72.87	88.36	
	基线+SCIM	86.11	0.6886	79.15	72.51	87.12	
	FFEM+CCIM	88.19	0.7334	82.06	75.97	89.21	
	FFEM+CCIM+SCIM	89.66	0.7635	83.93	79.21	89.26	

注:加粗字体表示各数据集上的最优结果。

对于三个实验数据集,以像素为中心从双时相影像中提取大小为7×7的图像块作为基本训练样本,每个图像块中心像素对应的变化标签作为该样本的类别标注。在此基础上,从变化与未变化样本中随机抽取 10% 的样本作为训练集,其余样本作为测试集。为了定量评估所提方法的有效性,采用了变化检测任务五项常用的评估指标,即总体准确率(OA)、Kappa 系数(KC)、精确率(Pre)、召回率(Rec)和 F1 分数(F1)(Wu 等,2025)。

为进一步验证所提方法的有效性,共选取了 6 种先进的对比方法进行对比实验,包括:SpectralFormer (Hong 等, 2022)、CSANet (Song 等, 2022)、TriTF (Wang 等, 2023)、DIEFEN (Wu 等, 2024)、FrFTML (Feng 等, 2025)和 AIWSEN (Wu 等, 2025)。所有对比方法均使用各自的开源代码实现,并根据原始文章中提供的参数设置进行实验。

2.3 消融实验

为了验证 FHI-Net 中各个模块的有效性,分别在三个实验数据集上进行了消融实验,以验证频域特征增强模块 FFEM、跨时相相关引导交互模块 CCIM、空-谱耦合交互模块 SCIM 对网络的贡献。消融实验的结果如表 1 所示。

表 1 中的基线为仅具备空-谱特征提取块与分类器的基础网络。可以观察到,在基线模型中分别引入 FFEM、CCIM 和 SCIM 后,模型在三个数据集上的各项指标均获得不同程度的提升,表明各模块均具有独立有效性。其中,FFEM 通过频域信息增强提升了时相特征的判别性与稳定性,从而为后续跨时相交互过程提供更可靠的特征基础,进而带来整体检测性能的提升。CCIM 通过跨时相相关引导交互提取到了更具判别力的特征,SCIM 通过空-谱耦合交互提升了对不同尺度变化区域的感知能力。

在上述基础上进一步加入 CCIM 后,网络在变
© 中国图象图形学报版权所有

化目标呈现碎片化、小尺度且分布离散在R2中表现出更明显的性能提升;同时,在地物类型复杂、变化形态不规则的R3中,检测性能也得到了进一步增强。以R3为例,其Kappa系数由0.6636显著提升至0.7334,提升幅度达到6.98%;R2的Kappa系数也由0.6668提升至0.7072。这说明CCIM通过充分建模双时相特征之间的时空依赖关系,有效增强了对小尺度、细微变化目标的识别能力,使细碎变化在复杂背景中更加稳定地被识别。

当进一步引入空-谱耦合交互模块SCIM后,网络在三个数据集上均取得最优性能,且在复杂场景中的提升尤为明显,其中R3的OA和Kappa系数分别达到89.66%和0.7635,相较基线模型提升了4.85%和11.77%。该结果表明,SCIM在多层变化特征融合过程中有效协调浅层细节与深层语义信息,增强了网络在复杂场景下对形态多样变化区域的感知能力,从而获得更加精确的变化检测结果。

2.4 对比实验

2.4.1 定量结果

不同方法在3个数据集上的定量结果如表2—

表4所示。综合三个数据集的实验结果可以看出,尽管

表2 不同方法在R1数据集上的实验结果

Table 2 Experimental results of different methods on the R1 dataset

	/%				
方法	OA	KC	F1	Pre	Rec
Spectral-Former	93.33	0.7993	84.14	88.33	80.33
CSANet	91.17	0.7285	78.36	85.10	72.60
TriTF	89.81	0.7039	76.94	76.68	77.19
DIEFEN	91.22	0.7401	79.59	81.56	77.72
FrFTML	92.87	0.7834	82.90	86.66	79.46
AIWSEN	91.32	0.7418	79.69	82.19	77.35
FHI-Net	93.04	0.8102	85.56	78.81	93.57

注:加粗字体表示各列最优结果。

FHI-Net在部分场景中的精确率(Pre)指标并非最高,但其在召回率(Rec)、F1以及KC等综合指标上都优于6种对比方法,体现出更加均衡且实用的检测

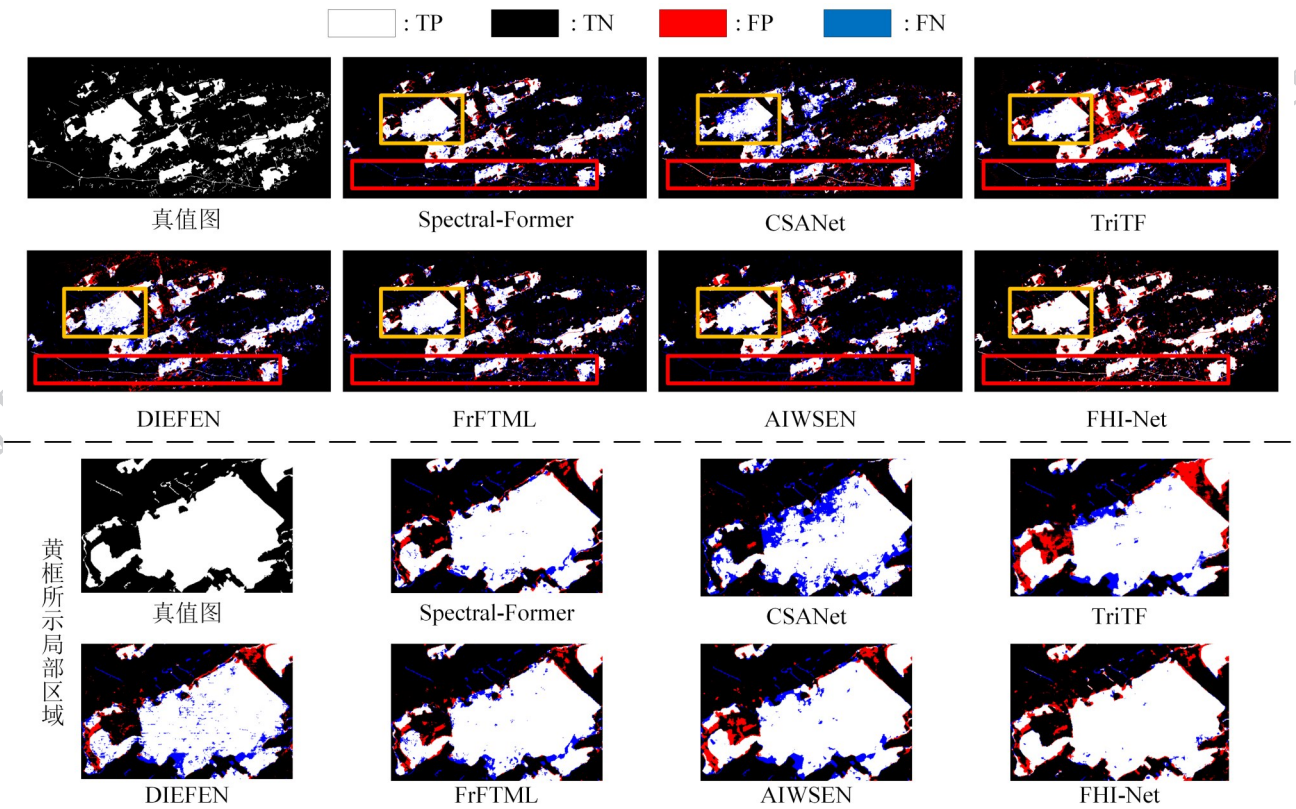


图6 不同方法在R1数据集上的变化检测结果可视化

Fig. 6 Visualization of change detection results using different methods on the R1 dataset

表 3 不同方法在 R2 数据集上的实验结果
Table 3 Experimental results of different methods on the R2 dataset

方法	OA	KC	F1	Pre	Rec
Spectral-Former	96.87	0.6566	67.25	82.33	56.84
CSANet	97.14	0.7093	72.43	79.52	66.49
TriTF	96.71	0.6035	61.89	89.94	47.18
DIEFEN	96.33	0.6172	63.62	72.43	56.72
FrFTML	96.62	0.6308	64.80	78.88	54.98
AIWSEN	96.63	0.6252	64.21	80.47	53.41
FHI-Net	96.53	0.7355	75.41	60.71	99.50

注:加粗字体表示各列最优结果。

性能。具体而言,在 R2 数据集中,变化样本极为稀少,一些对比方法如 TriTF 和 Spectral-Former 虽然在 Pre 指标上取得较高数值,分别达到 89.94% 和 82.33% 但其 Rec 分别仅为 47.18% 和 56.84%, 导致大量真实变化区域被漏检,整体 F1 指标也较低。类似现象在 DIEFEN 和 FrFTML 等方法中同样存在,其

表 4 不同方法在 R3 数据集上的实验结果
Table 4 Experimental results of different methods on the R3 dataset

方法	OA	KC	F1	Pre	Rec
Spectral-Former	88.58	0.7272	80.86	82.07	79.67
CSANet	89.31	0.7461	82.26	82.67	81.86
TriTF	79.99	0.5788	72.94	61.74	89.13
DIEFEN	82.55	0.5841	70.87	71.62	70.14
FrFTML	86.35	0.6747	77.20	78.07	76.36
AIWSEN	81.51	0.5352	65.99	74.45	59.26
FHI-Net	89.66	0.7635	83.93	79.21	89.26

注:加粗字体表示各列最优结果。

回调率 Rec 在三个数据集上都较低,限制了变化检测的完整性。相比之下,所提方法 FHI-Net 在该数据集上取得了 99.50% 的 Rec,几乎覆盖全部真实变化区域,同时在 F1 和 KC 等综合指标上均显著优于对比方法。

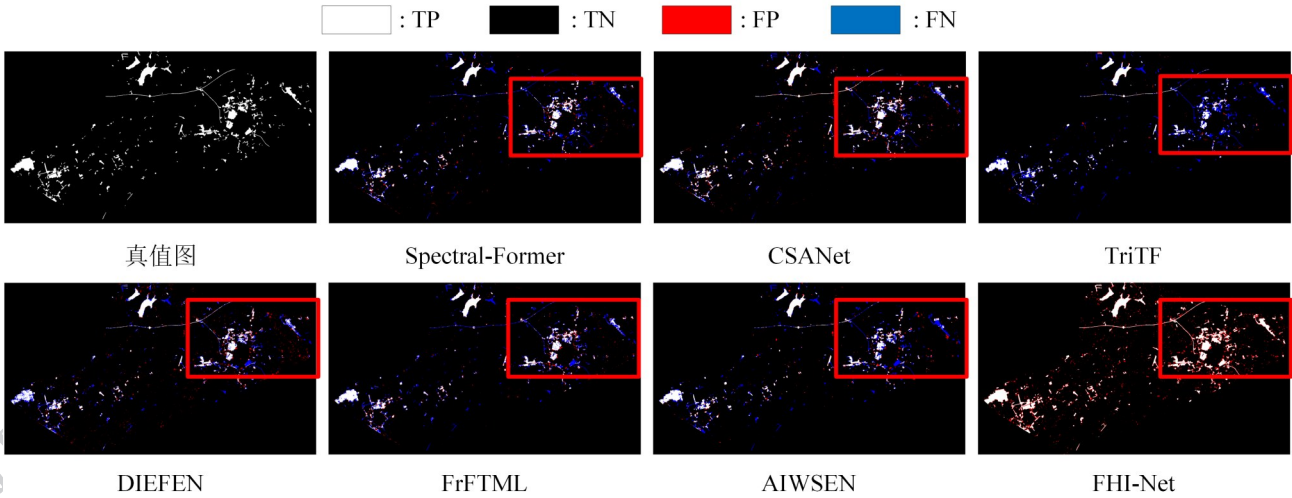


图 7 不同方法在 R2 数据集上的变化检测结果可视化

Fig. 7 Visualization of change detection results using different methods on the R2 dataset

在三个数据集上 FHI-Net 的 Pre 指标低于 Spectral-Former, CSANet 等方法。这主要与本文方法的设计目标有关。所提出的 FFEM 对特征进行频域自适应调制,使模型对变化响应更加敏感,而 CCIM 通过相关引导交互进一步放大潜在变化区域的特征表达,使模型对细微变化更加敏感。在此基础上,SCIM 通过融合多层变化特征提升了变化区域

的整体连贯性。上述设计在提高变化区域检测完整性的同时,可能引入少量误检,从而在一定程度上影响 Pre 指标。然而,这种设计能够显著提高 Rec 并减少真实变化区域的漏检。在实际变化检测应用中,漏检往往会带来更严重的影响,因此在保持整体性能平衡的情况下,提高变化区域的覆盖能力具有更高的实际意义。

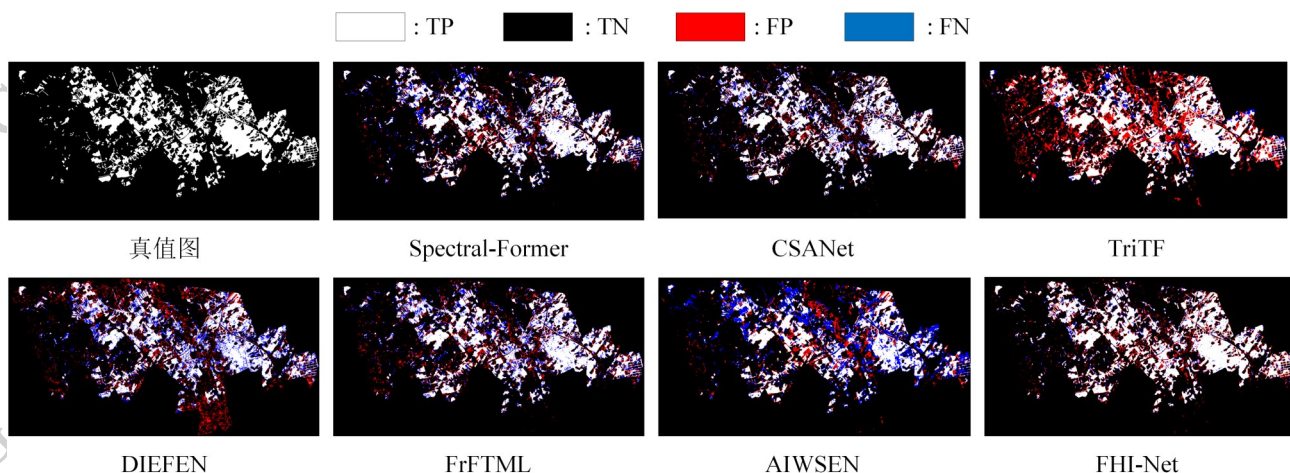


图8 不同方法在R3数据集上的变化检测结果可视化

Fig. 8 Visualization of change detection results using different methods on the R3 dataset

值得注意的是,在变化样本极为稀少的R2数据集中,FHI-Net取得了较高的召回率,但精确率相对较低。这可能是与变化样本极为稀少的分布特性有关,在高度不平衡场景下,少量误检即可在数量上超过真实变化样本,从而导致精确率下降,而只要大部分真实变化被检测到,召回率即可保持较高水平。尽管如此,所提方法在该数据集上的Kappa系数和F1值仍为所有对比方法中最高,体现了整体性能优势。同时,FHI-Net在R1和R3数据集中同样在F1和KC指标上取得最优结果,表明其较高召回率并未以牺牲整体判别性能为代价,而是在Pre与Rec之间实现了有效权衡,验证了方法在不同场景下的稳定性与鲁棒性。

2.4.2 定性结果

不同方法在3个数据集上的可视化变化检测结果如图6—图8所示。在检测结果图中,白色代表真阳性(TP),黑色代表真阴性(TN),红色代表假阳性(FP),蓝色代表假阴性(FN)。对于R1数据集,其左上角存在一处大尺度、连续分布的变化区域,而图像下部则包含一条细长、尺度较小的变化结构,二者分别代表了连续变化与细粒度变化两种典型模式(黄框与红框区域)。对于左上角的大块变化区域,部分对比方法(如Spectral-Former和CSANet)虽然能够检测到大部分变化区域,但在区域内部存在明显的破碎和断裂现象。多数对比方法难以保持细长变化结构的连续性,检测结果往往呈现出断续或局部缺失的情况,反映其对小尺度变化的敏感性不足。

相比之下,FHI-Net在上述两类区域中均表现出

更稳定的检测效果:在大尺度连续变化区域内,变化响应更加完整、连贯;在细线状变化结构处,能够较好地保持其空间连续性。这是由于FHI-Net通过FFEM与CCIM交互,有效抑制了由成像条件差异引入的伪变化干扰,并突出真实细微的变化区域的结构响应;在此基础上,SCIM进一步整合浅层与深层变化特征中所包含的互补信息,使模型能够同

时感知细粒度结构变化与高层语义变化,从而更全面地表征多尺度、多形态的变化对象。

R2与R3数据集在变化分布和场景复杂度上具有明显差异。R2数据集中变化样本极为稀少,变化区域主要集中于局部小范围内,而R3中变化区域分布更为广泛,变化形态不规则,地物类型复杂,背景干扰显著,对变化检测方法的稳定性提出了更高要求。

如图7所示,在R2数据集中,多数对比方法在红框标注区域内存在不同程度的漏检现象,尤其是AIWSEN和TriTF,其检测结果较为破碎,部分真实变化区域未被完整识别,反映出在变化样本稀少条件下无法准确识别细小变化。相比之下,FHI-Net通过频域特征增强、跨时相变化提取以及多层变化特征协同交互,充分聚合不同层级变化特征之间的上下文信息,能够在该区域内较为完整地识别变化区域,显著减少漏检现象。

在R3数据集中,由于变化区域形态复杂且边界不规则,多数方法在检测过程中容易引入大量误检。例如,部分方法在变化区域边缘产生明显的噪声响应,使变化结果呈现出碎片化特征。相

表 5 FFEM 对伪变化样本识别性能的影响
Table 5 Effect of FFEM on pseudo change sample recognition performance

数据集	未变化样本间的余弦相似度统计				伪变化样本预测准确度/%		
	范围	均值	标准差	阈值	伪变化样本个数	基线	基线+FFEM
R1	[0.3648,0.9996]	0.9574	0.0282	0.8726	12417	69.71	75.64
R2	[0.4314,0.9997]	0.9715	0.0191	0.9141	2833	95.55	96.47
R3	[0.3633,0.9998]	0.9604	0.0362	0.8517	14456	94.28	94.89

较而言,FHI-Net在保持变化区域整体结构完整性的同时,对背景干扰具有更好的抑制能力,能够较好地刻画复杂变化区域的空间形态,使变化区域更加连贯完整。

综上,在变化样本稀少和变化形态复杂的两类典型场景中,FHI-Net均表现出更稳定的检测效果,能够在减少漏检的同时保持较好的检测结果,进一步验证了其在林地复杂场景中的适应性和鲁棒性。

2.5 伪变化样本抑制实验

为定量验证FFEM在抑制伪变化方面的有效性,本文设计了基于样本相似性的伪变化验证实验。具体地,在未变化样本区域中,计算双时相对应样本之间的余弦相似度分布。考虑到双时相影像在波段数量上的差异,相似度计算仅基于共有的四个光谱波段进行,以保证计算的一致性与可比性。随后,对未变化样本的相似度分布进行统计分析,计算其均值与标准差,并采用三倍标准差准则确定伪变化判定阈值。最终,将低于该阈值的未变化样本视为潜在伪变化样本。

如表5所示,与基线模型相比,引入FFEM后模型在伪变化样本上的识别准确率在各数据集上均有不同程度提升。这表明所提出的FFEM模块能够有效抑制由成像条件差异引起的伪变化响应,从而增强模型对真实变化与伪变化的判别能力。

2.6 波段设置与通道对齐策略分析

针对T1时相为7波段而T2时相为4波段的情况,本文采用对T2时相影像的最后一个波段进行通道复制的方式以实现输入通道数对齐。该操作并未引入额外光谱信息,仅用于保持网络输入维度一致,从而支持双时相特征的联合建模。相比直接选取双时相共有的4个波段,该策略能够保留T1时相中更丰富的光谱信息,有助于提升变化检测的判别能力。

为验证该策略的合理性,本文进一步补充了统

表 6 不同波段设置下各数据集的变化检测结果对比
Table 6 Comparison of change detection performance on different datasets under different spectral band settings
/%

	波段数	OA	KC	F1	Pre	Rec
R1	4波段	91.91	0.7805	83.33	76.32	91.76
	7波段	93.04	0.8102	85.56	78.81	93.57
R2	4波段	95.99	0.7165	73.66	58.62	99.10
	7波段	96.53	0.7355	75.41	60.71	99.50
R3	4波段	88.03	0.7305	81.90	75.49	89.51
	7波段	89.66	0.7635	83.93	79.21	89.26

注:加粗字体表示各个数据集上的最优结果。

表 7 在R3数据集上的数据分离设置下的性能比较
Table 7 Performance comparison under the data separation setting on the R3 dataset
/%

方法	OA	KC	F1	Pre	Rec
CSANet	89.34	0.7439	81.94	84.39	79.63
FHI-Net	89.69	0.7647	84.04	79.26	89.44

注:加粗字体表示各列最优结果。

一采用4波段输入的对比实验,结果如表6所示。从结果可以看出,在三个数据集上,基于7波段输入的模型在OA、KC及F1等指标上均优于4波段设置,表明通道复制策略未引入显著噪声干扰,且能够在保留更多光谱信息的同时提升模型性能。

2.7 数据分离条件下的性能验证

本文在统一的数据预处理与数据划分框架下开展实验,并采用固定随机种子进行样本划分,以保证实验设置的一致性与可复现性。在此基础上,进一步设计了空间严格分离实验,即将训练样本所在的空间区域从测试集中完全剔除,其余区域作为测试集,从而实现训练集与测试集在空间上的完全隔离。

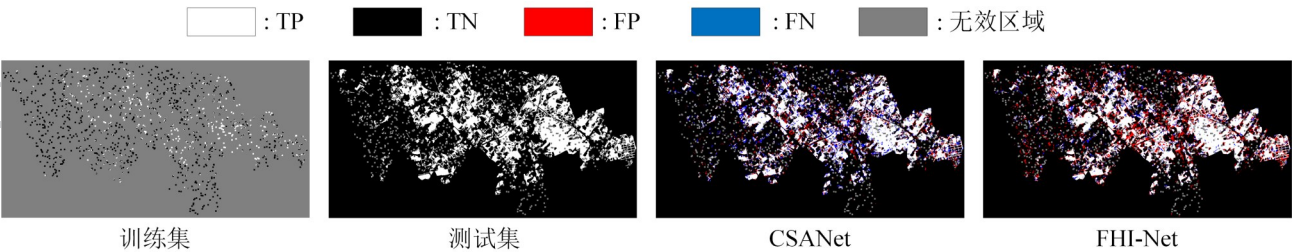


图9 数据分离条件下的变化检测结果可视化

Fig. 9 Visualization of change detection results under the data separation setting

表8 不同方法的参数量与每秒浮点运算次数对比

Table 8 Comparison of parameter size and FLOPs of different methods

方法	参数量/M	FLOPs/G
Spectral-Former	0.09	3.08
CSANet	40.33	0.16
TriTF	0.16	0.04
DIEFEN	7.62	0.04
FrFTML	26.01	1.66
AIWSEN	11.41	0.02
FHI-Net	26.78	2.23

分离数据集的训练集和测试集如图9所示,灰色表示为无效区域。

在该严格分离条件下的实验结果如表7所示,所提方法仍取得稳定的检测性能。并在整体指标上优于次优的对比方法CSANet,进一步验证了方法的有效性与实验结果的可靠性。

2.8 计算开销分析

为进一步分析所提方法的计算开销,对各方法的参数量和浮点运算量进行了统计,结果如表8所示。从表中可以看出,FHI-Net的参数量为26.78M,FLOPs为2.23G。尽管所提方法的参数量相较部分轻量级模型有所增加,但其仍低于CSANet(40.33M)等方法,同时在FLOPs方面保持在中等水平,显著低于Spectral-Former(3.08G FLOPs)。这表明FHI-Net在保证特征表达能力的同时,并未显著增加计算复杂度。相比之下,虽然部分方法(如TriTF和AIWSEN)具有较低的计算开销,但其检测性能相对有限。综合来看,FHI-Net在计算复杂度与检测性能之间实现了更优的权衡,在可接受的计算成本下获得了更显著的性能提升。

3 结 论

本文针对林地变化检测任务中双时相图像风格差异及空间异质性强等挑战,提出了一种多光谱遥感图像频域增强分层交互网络(FHI-Net)。该方法通过频域特征增强模块(FFEM)抑制风格差异引入的伪变化干扰,并结合跨时相相关引导交互模块(CCIM)提取更具判别力的变化特征,同时利用空-谱耦合交互模块(SCIM)实现多层次变化特征融合,从而实现对细微变化与多尺度变化区域的有效感知与准确定位。在三个林地变化检测数据集上的实验结果表明,所提出的FHI-Net优于六种现有先进方法,验证了其在复杂场景下的有效性与鲁棒性。未来研究将进一步扩展研究区域与数据规模,以系统评估方法在大范围场景下的泛化能力。同时,探索基于自监督学习的频域特征建模方法,以降低对高质量标注数据的依赖,并进一步拓展至多时相变化检测框架,实现林地变化过程的动态建模,从而提升模型在复杂时序场景中的鲁棒性与泛化能力。

参考文献(References)

Chang H, Wang P J, Diao W H, Xu G L and Sun X. 2024. Remote sensing change detection with bitemporal and differential feature interactive perception. IEEE Transactions on Image Processing, 33: 4543 - 4555 [DOI: 10.1109/TIP.2024.3424335]
Chen H, Qi Z P and Shi Z W. 2022. Remote sensing image change detection with transformers. IEEE Transactions on Geoscience and Remote Sensing, 60: 1 - 14 [DOI: 10.1109/TGRS.2021.3095166]
Daudt R C, Le Saux B and Boulch A. 2018. Fully convolutional Siamese networks for change detection // Proceedings of the 25th IEEE International Conference on Image Processing (ICIP). Athens, Greece: IEEE: 4063 - 4067 [DOI: 10.1109/ICIP.2018.8451652]
Feng S, Lan T Y, Fan Y Z, Zhang M M, Zhao C H and Li W et al.

2025. An adaptive weighted metric learning network based on fractional domain decoupling for hyperspectral change detection. *IEEE Transactions on Neural Networks and Learning Systems*, 36(10): 17856 – 17870 [DOI: 10.1109/TNNLS.2025.3586714]
- Gao N, Jiang X Y, Zhang X H and Deng Y. 2025. Efficient frequency-domain image deraining with contrastive regularization // *Proceedings of the European Conference on Computer Vision (ECCV). Lecture Notes in Computer Science*, vol. 15099. Berlin: Springer: 240 – 257 [DOI: 10.1007/978-3-031-72940-9_14]
- Gao H N, Su X, Wu C, Du B and Zhang L P. 2024. SAAN: similarity-aware attention flow network for change detection with VHR remote sensing images. *IEEE Transactions on Image Processing*, 33: 2599 – 2613 [DOI: 10.1109/TIP.2024.3349868]
- Hong D F, Han Z, Yao J, Gao L R, Zhang B and Plaza A et al. 2022. SpectralFormer: rethinking hyperspectral image classification with transformers. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1 – 15 [DOI: 10.1109/TGRS.2021.3130716]
- Huang J X, Guan D Y, Xiao A R and Lu S J. 2021. FSDR: frequency space domain randomization for domain generalization // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville: IEEE/CVF: 6887 – 6898 [DOI: 10.1109/CVPR46437.2021.00682]
- Jia M, Zhao Q and Lu X F. 2024. Bipartite adversarial autoencoder network for unsupervised detection of changes in heterogeneous remote sensing images. *Journal of Image and Graphics*, 29(08): 2188 – 2204 (贾萌, 赵秦, 鲁晓锋. 2024. 双边对抗自编码网络的无监督异质遥感图像变化检测. *中国图象图形学报*, 29(08): 2188 – 2204) [DOI: 10.11834/jig.230497]
- Jiang B, Wang Z T, Wang X X, Zhang Z Y, Chen L, Wang X and Luo B. 2023. VcT: visual change transformer for remote sensing image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1 – 14 [DOI: 10.1109/TGRS.2023.3327139]
- Li Z M, Yan C X, Sun Y and Xin Q C. 2022. A densely attentive refinement network for change detection based on very-high-resolution bitemporal remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1 – 18 [DOI: 10.1109/TGRS.2022.3159544]
- Liu P, Liu Y X, Guo X X, Liu Y C, Zhao W J and Xu W X. 2026. Faster, better, and more accurate mapping of burned areas using Sentinel-2 multispectral images. *Remote Sensing of Environment*, 333: 115137 [DOI: 10.1016/j.rse.2025.115137]
- Liu Y, Yu R, Wang J H, Zhao X Y, Wang Y T and Tang Y Set al. 2022. Global spectral filter memory network for video object segmentation // *Proceedings of the European Conference on Computer Vision (ECCV). Lecture Notes in Computer Science*, vol. 13689. Berlin: Springer: 648 – 665 [DOI: 10.1007/978-3-031-19818-2_37]
- Liu Y L, Zhang F, Zhang S X, Zhang K, Sun J D and Bruzzone L. 2024. Content-guided spatial – spectral integration network for change detection in HR remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1 – 16 [DOI: 10.1109/TGRS.2024.3352050]
- Lv Z Y, Wang F J, Cui G Q, Benediktsson J A, Lei T and Sun W W. 2022. Spatial – spectral attention network guided with change magnitude image for land cover change detection using remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1 – 12 [DOI: 10.1109/TGRS.2022.3197901]
- Miao L Z, Li X T, Zhou X X, Yao L, Deng Y M and Hang Tet al. 2024. SNUNet3+: a full-scale connected Siamese network and a dataset for cultivated land change detection in high-resolution remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1 – 18 [DOI: 10.1109/TGRS.2023.3344284]
- 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: 1 – 5 [DOI: 10.1109/LGRS.2022.3179134]
- Wang X H, Zhao K Y, Zhao X Y and Li S Y. 2023. TriTF: a triplet transformer framework based on parents and brother attention for hyperspectral image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1 – 13 [DOI: 10.1109/TGRS.2023.3260969]
- Wang G X, Cheng G, Zhou P C and Han J W. 2024. Cross-level attentive feature aggregation for change detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(7): 6051 – 6062 [DOI: 10.1109/TCSVT.2023.3344092]
- Wang R F, Yang Z J, Qiu H, Wang F, Gao G and Wu D. 2025. Lightweight change detection network integrating feature interaction and fusion. *Journal of Image and Graphics*, 30(3): 0870 – 0882 (王仁芳, 杨梓健, 邱虹, 王峰, 高广, 吴敦. 2025. 结合特征交互和融合的轻量级变化检测网络. *中国图象图形学报*, 30(3): 0870 – 0882) [DOI: 10.11834/jig.240280]
- 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(5): 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(5): 1460 – 1478) [DOI: 10.11834/jig.230212]
- Wang Y H, Gao L R, Chen Z C and Zhang B. 2020. Deep learning and superpixel-based method for high-resolution remote sensing image change detection. *Journal of Image and Graphics*, 25(6): 1271 – 1282 (王艳恒, 高连如, 陈正超, 张兵. 2020. 结合深度学习和超像素的高分遥感影像变化检测. *中国图象图形学报*, 25(6): 1271 – 1282) [DOI: 10.11834/jig.190319]
- Wu C, Zhang L P, Du B, Chen H R X, Wang J X and Zhong H. 2024. UNet-like remote sensing change detection: A review of current models and research directions. *IEEE Geoscience and Remote Sensing Magazine*, 12(4): 305 – 334 [DOI: 10.1109/MGRS.2024.3412770]
- Wu L X, Ma L, Peng J T, Yang B, Sun W W and Wang J Het al. 2024.

- DIEFEN: differential information-enhanced feature exchange network for hyperspectral change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1 – 12 [DOI: 10.1109/TGRS.2024.3469930]
- 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: 1 – 12 [DOI: 10.1109/TGRS.2025.3531478]
- Xie X S, Zhou Z H, Niu C and Li Z L. 2025. Revisiting semantic gap in u-shaped networks from a performance oriented perspective: Quantification, visualization, and optimization. *IEEE Transactions on Circuits and Systems for Video Technology* [DOI: 10.1109/TCSVT.2025.3626818]
- Xu X T, Yang Z and Li J J. 2023. AMCA: Attention-guided multiscale context aggregation network for remote sensing image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1 – 19 [DOI: 10.1109/TGRS.2023.3272006]
- Yang Y C and Soatto S. 2020. FDA: Fourier domain adaptation for semantic segmentation // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE/CVF: 4084 – 4094 [DOI: 10.1109/CVPR42600.2020.00414]
- Yu D B and Gao Z. 2025. VolFormer: explore more comprehensive cube

interaction for hyperspectral image restoration and beyond // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville: IEEE Computer Society: 28091 – 28101 [DOI: 10.1109/CVPR52734.2025.02616]

- Zhao S J, Zhang X L, Xiao P F and He G J. 2023. Exchanging dual-encoder – decoder: A new strategy for change detection with semantic guidance and spatial localization. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1 – 16 [DOI: 10.1109/TGRS.2023.3327780]

作者简介

陈勇奇,男,博士研究生,主要研究方向为遥感图像变化检测。E-mail:chenyq@hrbeu.edu.cn

冯收,通信作者,男,副教授,主要研究方向为遥感图像处理。E-mail:fengshou@hrbeu.edu.cn

任春颖,女,研究员,主要研究方向为森林遥感。E-mail:renchy@iga.ac.cn

唐英杰,男,博士研究生,主要研究方向为遥感图像变化检测。E-mail:tangyj@hrbeu.edu.cn

任慧鑫,女,博士研究生,主要研究方向为森林遥感。E-mail:renhuixin@iga.ac.cn

赵春晖,男,教授,主要研究方向为数字信号与图像处理、高光谱遥感图像处理。E-mail:zhaochunhui@hrbeu.edu.cn