E-mail: jig@aircas.ac.cn
Website: www.cjig.cn
Tel: 010-58887035
中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS
©中国图象图形学报版权所有

中图法分类号:TP751　　**文献标识码**: A　　**文章编号**: 1006-8961(2026)09-3266-12

论文引用格式:Ding C, Zhang J B, Hao X F, Zheng S R and Yan S. 2026. Time-difference-guided network for hyperspectral image change detection. Journal of Image and Graphics, 31(9):3266-3277(丁晨, 张景博, 郝小锋, 郑思睿, 闫松. 2026. 时间差引导的高光谱图像变化检测网络. 中国图象图形学报, 31(9):3266-3277)[DOI:10.11834/jig.260214]

时间差引导的高光谱图像变化检测网络

丁晨[1*],张景博[1],郝小锋[1],郑思睿[1],闫松[2]

1. 西安邮电大学计算机学院,西安　710121; 2. 西安航天动力研究所,西安　710100

摘　要:**目的**　高光谱图像变化检测在环境监测、土地利用分析及灾害评估等领域具有重要应用价值。然而,现有深度学习方法多在高层或决策阶段进行特征比较,难以充分利用多层级时间差异信息,导致细微变化难以准确捕获,限制检测精度。针对上述问题,本文提出时间差引导网络(time-difference-guided network, TDG-Net)。**方法**　以Vision Mamba为主干构建孪生架构,设计时间存储器模块(time storage module,TSM),基于简化长短期记忆网络(simplified long short-term memory,sLSTM)对双时相高光谱图像多层特征进行时序建模,提取显式时间差特征;进一步通过时域差引导策略将时间差信息转化为空间注意力权重,并以残差方式反馈至多层特征提取网络,实现对变化区域的强化与未变化区域的抑制。同时引入Dice损失函数缓解类别不平衡问题。与近年Mamba-based方法相比,本文核心区别在于通过轻量LSTM在多层级显式建模并引导时间差,而非仅在高层隐式融合。**结果**　在River、Farmland和Hermiston数据集上,所提方法在总体精度(overall accuracy,OA)和Kappa系数上均优于CVA(change vector analysis)及自监督高光谱空间—光谱特征方法。其中,在River数据集上OA为96.54%、Kappa为78.51%;在Farmland数据集上分别为95.87%、90.11%;在Hermiston数据集上分别为98.21%、91.79%。消融实验表明TSM与时域差引导策略均带来性能的显著提升。**结论**　本文通过显式多层级时间差建模与层次化引导,有效提升高光谱变化检测精度与鲁棒性,在复杂场景下具备良好泛化能力。

关键词:高光谱图像(HSI);变化检测(CD);Vision Mamba;时间差建模;深度学习;遥感

Time-difference-guided network for hyperspectral image change detection

Ding Chen[1*], Zhang Jingbo[1], Hao Xiaofeng[1], Zheng Sirui[1], Yan Song[2]

1. School of Computer Science and Technology, Xi'an University of Posts and Telecommunications, Xi'an　710121, China;

2. Xi'an Aerospace Propulsion Institute, Xi'an　710100, China

Abstract: **Objective** Hyperspectral images (HSIs) capture reflectance values across hundreds of contiguous spectral bands at each spatial location, thus facilitating the precise discrimination of materials with subtle spectral differences by generating rich spectral-spatial information. This high spectral resolution makes HSIs particularly valuable for many application, such as environmental monitoring, land cover mapping, precision agriculture, and change detection (CD) applications. CD in multitemporal HSIs aims to identify meaningful surface alterations by comparing images of the same scene acquired at different times. Nowadays, accurate HSI-CD plays a critical role in ecological monitoring, urban expansion analysis, and disaster assessment. However, this task faces significant challenges, including high dimensionality, spectral redundancy, spatial heterogeneity, and the difficulty of effectively integrating spatial, spectral, and temporal features. Tra-

收稿日期:2026-04-16;**修回日期**:2026-05-19;**预印本日期**:2026-05-26

*通信作者:丁晨　dingchen@xupt.edu.cn

基金项目:国家自然科学基金项目(62472350)

Supported by: National Natural Science Foundation of China(62472350)

ditional methods like change vector analysis (CVA), principal component analysis (PCA), and multivariate alteration detection (MAD), rely on hand-crafted features and empirical thresholding. Despite being computationally efficient, these approaches are sensitive to noise, illumination variations, and atmospheric effects. Deep learning methods, particularly convolutional neural networks (CNNs), have demonstrated superior performance by automatically learning hierarchical features. While most CNN-based frameworks adopt dual-branch Siamese architectures, they still struggle to capture long-range dependencies and explicit temporal dynamics. **Method** To address the abovementioned limitations, a novel time-difference-guided network (TDG-Net) is introduced in this paper. The proposed method employs a Siamese architecture with Vision Mamba as the feature extraction backbone to efficiently capture long-range spatial dependencies with linear computational complexity. Two core components are integrated: the time storage module (TSM) and the temporal difference guidance strategy. Bitemporal HSIs are first fed into the dual-branch Vision Mamba backbone to extract multilevel spatial-spectral features. Then, the TSM performs sequential temporal modeling on these features, using a simplified long short-term memory (sLSTM). To reduce computational cost, spectral compression via 1×1 convolution and spatial down-sampling via max pooling are applied before feeding features into the sLSTM. The sLSTM generates explicit temporal difference representations. The temporal difference guidance strategy then converts these low-resolution difference features into spatial attention weights through bilinear up-sampling and ReLU activation. These weights are then fed back to the dual-branch network via residual connections, adaptively emphasizing change-relevant regions and suppressing unchanged areas at each hierarchical level. Finally, to produce the binary change detection map, the enhanced multilevel features are fused and passed to the classification head. To mitigate severe class imbalance, Dice loss is employed instead of conventional cross-entropy loss, directly optimizing the overlap between predicted and ground-truth change regions. Compared with recent Mamba-based methods, the core distinction of TDG-Net lies in its explicit modeling and hierarchical guidance of temporal differences across multiple feature levels via a lightweight TSM and residual feedback, rather than relying primarily on high-level implicit fusion. **Result** Comprehensive experiments are conducted on three widely used benchmark hyperspectral datasets: River, Farmland, and Hermiston. All experiments are implemented on an NVIDIA RTX 3090 GPU using TensorFlow-GPU 2.5.0. Performance is then evaluated using overall accuracy (OA) and Kappa coefficient, with results averaged over 10 independent runs. Ablation studies confirm the significant contributions of the TSM and temporal difference guidance strategy. Comparative experiments against state-of-the-art methods demonstrate that TDG-Net consistently outperforms all competitors. On the River dataset, TDG-Net achieves an OA of 96.54% and a Kappa of 78.51%. On the Hermiston dataset, it reaches an OA of 98.21% and a Kappa of 91.79%. On the Farmland dataset, it attains an OA of 95.87% and a Kappa of 90.11%. Additional analyses on model complexity show that TDG-Net maintains competitive parameter count, floating point operations (FLOPs), and inference time while achieving superior accuracy. **Conclusion** The proposed TDG-Net effectively addresses the key limitations of existing HIS CD methods by explicitly modeling temporal differences at multiple hierarchical levels and guiding feature learning with adaptive attention. Furthermore, by integrating a lightweight TSM based on simplified LSTM and a temporal difference guidance strategy within a Vision Mamba backbone, the method captures rich spatial-spectral-temporal dynamics while effectively suppressing pseudo-changes. The introduction of Dice loss further alleviates the severe class imbalance problem. Extensive experiments on three benchmark datasets demonstrate that TDG-Net achieves state-of-the-art performance in terms of quantitative metrics and visual quality, exhibiting strong robustness and generalization in complex scenarios. Additionally, compared with existing Mamba-based methods, such as SAVDGN and CDMamba, which primarily rely on implicit high-level fusion, TDG-Net explicitly models and hierarchically guides temporal differences across multiple feature levels through a lightweight TSM and residual feedback mechanism, thus achieving more fine-grained change representation.

Key words: hyperspectral image (HSI); change detection (CD); Vision Mamba; temporal difference modeling; deep learning; remote sensing

0 引 言

高光谱图像(hyperspectral image, HSI)在每个空间位置包含数百个连续光谱波段,具有丰富光谱—空间信息,可实现对细微材料差异的精确识别(Bhargava等,2024;Fauvel等,2013),因此在环境监测、土地覆盖制图、精准农业及变化检测等领域具有重要应用价值(Lu和Weng, 2007; Thenkabail等, 2004)。高光谱变化检测通过对比不同时相影像识别地表变化,在生态监测、城市扩展及灾害评估中具有重要意义(Radke等, 2005; Bruzzone和Carlin, 2006; Singh, 1989),但仍面临高维冗余、空间异质性及光谱—空间—时间融合困难等挑战(Plaza等, 2009; Zhao和Du, 2016)。

传统方法如变化向量分析(change vector analysis, CVA)、主成分分析、多元变化检测以及支持向量机和K近邻等,通常基于差异图像或统计指标进行变化检测(Malila, 1980; Nielsen, 2007)。尽管计算高效,但依赖人工特征与经验阈值,对噪声、光照及大气变化敏感,难以建模复杂光谱—空间关系(Bai等, 2023)。周承乐等人(2024)提出了一种光谱—频域融合的高光谱变化检测方法,结合梯度相关性光谱距离与傅里叶显著性增强,有效提升了变化像元的可分性与检测鲁棒性。

随着深度学习的发展,卷积神经网络(convolutional neural network, CNN)广泛应用于高光谱图像变化检测,典型方法多采用孪生网络结构对双时相影像进行特征提取,并通过差异或相似度建模变化信息(Zhu等, 2017; Wang等, 2019),但在长程依赖建模及显式时间建模方面仍存在不足。

为增强时间建模能力,研究者提出循环神经网络及卷积长短期记忆网络以联合建模空间—光谱—时间特征(Mou等, 2019),但其计算开销较高。同时,自监督学习在标注稀缺场景中展现出良好潜力,例如自监督高光谱空间—光谱特征理解网络(self-supervised hyperspectral spatial-spectral feature understanding network, HyperNet)通过自监督学习获得鲁棒表示(Hu等, 2022)、高效多时相自注意力网络利用多时相自注意力机制抑制冗余信息并强化时间差异(Hu等, 2023)。此外,李湘潭等人(2025)提出基于混合注意力与双向门控网络的变化检测方法,通过混合注意力机制抑制噪声并融合全局与局部特征,在复杂噪声环境下显著提升了检测精度。

近年来,Transformer通过自注意力建模全局依赖被引入HSI变化检测。光谱—时间Transformer及其扩展可有效捕捉跨波段与跨时间关系(Li和Ding, 2023; Zou等, 2024)。刘文力等人(2025)提出基于通道变换和Transformer的高光谱图像变化检测方法,通过通道洗牌提升效率并强化全局上下文建模,进一步缓解了长程依赖问题。与此同时,相关研究从空间增强与跨域建模展开探索,如通过空间信息增强提升特征表达能力(肖岱峰 等, 2026)、轻量型Res-3D-CNN结合3D卷积与Transformer以增强跨域特征表示能力(杨丽霞 等, 2026)。在此基础上,Gao等人(2025)提出光谱注意力与视觉Mamba融合的变化检测网络(spectral attention and vision Mamba-based detection graph network, SAVDGN);Qu等人(2025)提出变化检测Transformer(change detection transformer, CDAFormer)实现无监督域自适应;Zhang等人(2025)提出变化检测Mamba网络(change detection Mamba, CDMamba),通过引入状态空间模型增强长程依赖建模能力。

尽管上述方法取得了一定进展,但其时间信息建模方式仍存在一定局限:一方面,这些方法多在高层特征空间中对双时相信息进行隐式融合,缺乏对多层级时间差异的显式刻画,导致不同层级特征之间的时间信息难以协同利用,从而限制了时间差异特征的充分表达。时间差异是区分变化区域与未变化区域的核心判别依据,其能够刻画同一场景在不同时刻的细微变化,若未能在多层级上有效建模,将降低模型对细粒度变化的感知能力与检测精度。另一方面,这类方法通常将时间建模过程与空间特征提取耦合在统一框架中,缺乏针对“差异敏感性”的专门设计,使得模型对变化区域与非变化区域的区分能力仍然有限。

针对上述问题,本文提出时间差引导网络(time-difference-guided network, TDG-Net),通过轻量级序列记忆模块显式提取双时相时间差信息,并在多层特征学习中进行引导,增强了时空特征的融合和时间信息的建模能力。具体而言,时间存储器模块(time storage module, TSM)基于简化长短期记忆网络(simplified long short-term memory, sLSTM)生成显式时间差特征,并通过多层次时间信息协同建模提升

不同层级特征之间的时间关联能力，从而克服现有Mamba-based方法主要依赖高层隐式融合的问题。而时域差引导策略进一步利用差异特征生成空间注意力权重，并通过残差连接反馈至双分支网络，引导模型更加关注变化区域、抑制不变区域的冗余响应，从而减少伪变化干扰并显著提升检测精度与鲁棒性。

1 方　法

1.1 框架概述

本文方法整体结构如图 1 所示，旨在充分挖掘双时相高光谱图像的时间差异信息以提升细微变化检测能力。该网络基于孪生框架构建，以 Vision Mamba 作为特征提取主干，包含时间存储器模块 TSM 与时域差引导策略两个核心组件，二者在多层级上协同增强时空变化特征。首先，Vision Mamba 主干对双时相输入进行多层空间—光谱特征提取；随后，TSM 通过谱维压缩与空间下采样降低计算开销，并结合 sLSTM 建模时间依赖，生成显式时间差特征并抑制冗余响应；进一步，时域差引导策略将低分辨率时间差特征上采样并映射为注意力权重，通过残差方式反馈至双分支网络，引导模型聚焦真实变化区域。最终，多层特征融合后经分类层输出变化检测结果。

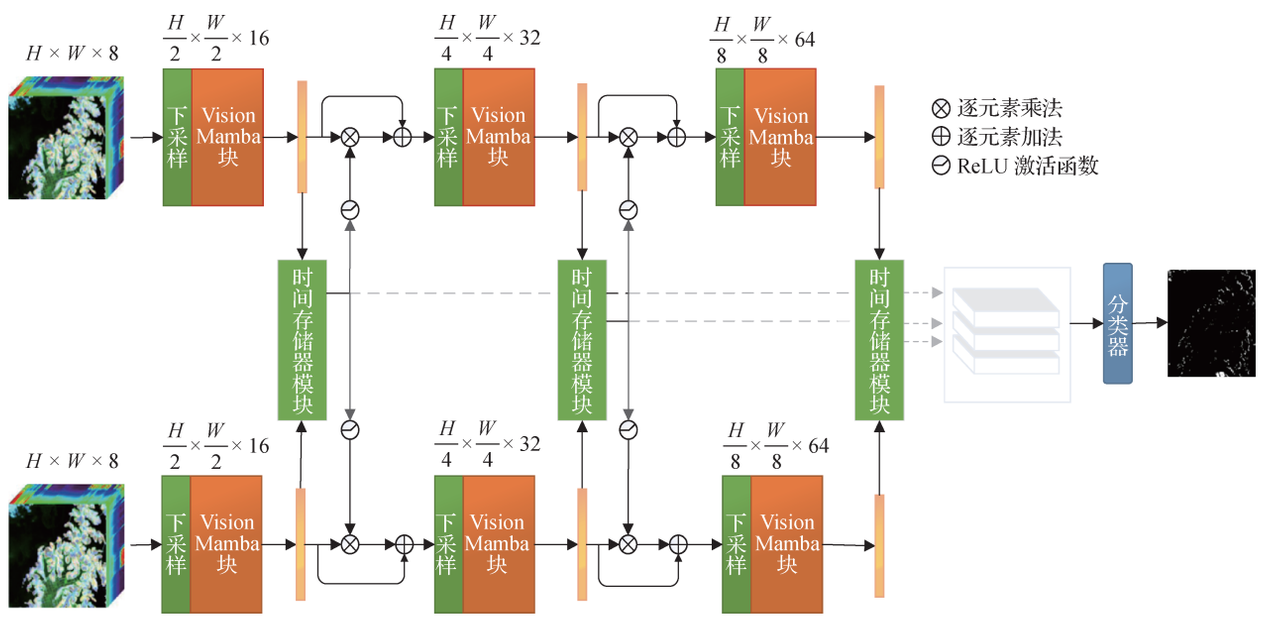


图 1　时间差异引导网络整体框架图

Fig. 1　Overall framework of the proposed time-difference-guided network (TDG-Net)

1.2 时间存储器模块

本文方法基于简化长短期记忆网络 sLSTM 设计了时间存储器模块 TSM。该模块在继承标准长短期记忆网络（long short-term memory，LSTM）门控机制优势的基础上，通过简化网络结构显著减少参数数量，既保留了对时序信息的强建模能力，又有效降低了高维高光谱数据下的训练复杂度，提高了训练稳定性与效率。

具体而言，sLSTM 相较于标准 LSTM、卷积长短期记忆网络（convolutional LSTM，ConvLSTM）和卷积门控循环单元（convolutional gated recurrent unit，ConvGRU）的主要简化策略包括：1）门控参数共享机制：将部分卷积核参数在不同门控单元之间共享，减少冗余权重，降低模型复杂度；2）状态降维与谱维压缩协同：在输入 sLSTM 之前通过 1 × 1 卷积对通道维度进行压缩，并结合最大池化进行空间下采样，有效降低输入特征维度和计算负担；3）优化门控逻辑：精简部分非线性变换操作，同时在卷积操作中融入空间关联信息，使时间建模自然融合空间特征。

这些简化措施使 sLSTM 在保持门控记忆能力的同时，参数量较标准 ConvLSTM 减少约 30%～45%，浮点运算次数（floating point operations，FLOPs）显著下降，有效缓解了高光谱数据高维特性导致的梯度消失与训练低效问题。此外，在时序特征提取过程中，sLSTM 通过卷积门控机制将空间关联信息与时间依赖自然融合，避免单独处理时间维度造成的信

息割裂，从而实现时间—空间特征的协同捕获，为多层级显式时间差特征的提取提供了可靠基础。时间存储器模块结构如图2所示。

具体而言，设网络在第i个特征层级提取到的双时相特征分别为

$$\boldsymbol{F}_i^{(t)} \in \mathbf{R}^{H_i \times W_i \times C_i} \tag{1}$$

式中，$\boldsymbol{F}_i^{(t)}$为第i层在时刻t的特征图（图像特征），H_i、W_i分别表示空间高度与宽度，C_i表示通道数，$i \in \mathbf{N}$表示网络层索引，t表示时间索引，考虑高光谱数据在光谱维存在冗余信息，首先采用逐点卷积进行谱维压缩，具体为

$$\hat{\boldsymbol{F}}_i^{(t)} = Conv_{1\times1}\left(\boldsymbol{F}_i^{(t)}\right),\ \ t \in \{1,2\} \tag{2}$$

式中，$Conv_{1\times1}(\cdot)$表示1×1卷积操作，用于将通道维从C_i压缩至较低维度，同时保留判别性光谱信息。

在完成谱维压缩后，为进一步降低时间建模阶段的计算成本，本文在sLSTM之前引入最大池化（max pooling，MP）层，对空间维进行下采样，具体为

$$\boldsymbol{X}_i^{(t)} = MP\left(\hat{\boldsymbol{F}}_i^{(t)}\right) \tag{3}$$

式中，$MP(\cdot)$表示最大池化算子，用于减小特征图空间分辨率；$\boldsymbol{X}_i^{(t)}$表示输入至sLSTM的特征表示。该操作在保留局部显著响应的同时，有效减少了sLSTM中卷积门控的参数数量和计算复杂度，从而降低训练成本。随后，处理后的双时相特征按照时间顺序输入sLSTM单元进行时间建模。首先，将时刻$T1$的特征$\boldsymbol{X}_i^{(1)}$输入sLSTM，并在初始隐藏状态与记忆状态均为零的条件下更新状态，具体为

$$\left(\boldsymbol{H}_i^{(1)} \middle| \boldsymbol{C}_i^{(1)}\right) = sLSTM\left(\boldsymbol{X}_i^{(1)}\right) \tag{4}$$

式中，$\boldsymbol{H}_i^{(1)}$和$\boldsymbol{C}_i^{(1)}$分别表示sLSTM在第1个时间步输出的隐藏状态与记忆状态，用于刻画时刻$T1$的时间特征表示。随后，在处理时刻$T2$的特征$\boldsymbol{X}_i^{(2)}$时，将前一时刻的隐藏状态与记忆状态作为初始条件传递给sLSTM，具体为

$$\left(\boldsymbol{H}_i^{(2)} \middle| \boldsymbol{C}_i^{(2)}\right) = sLSTM\left(\boldsymbol{X}_i^{(2)} \middle| \boldsymbol{H}_i^{(1)}, \boldsymbol{C}_i^{(1)}\right) \tag{5}$$

式中，$\left(\boldsymbol{H}_i^{(1)}, \boldsymbol{C}_i^{(1)}\right)$为从时刻$T1$继承的时间上下文信息，使得sLSTM在更新当前状态时能够显式考虑两时相之间的关联性。

在sLSTM内部，时间状态更新由卷积门控机制完成。以时间步t为例，各门控计算为

$$\begin{cases} \boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_f * \boldsymbol{X}_t + \boldsymbol{U}_f * \boldsymbol{H}_{t-1} + \boldsymbol{b}_f\right) \\ \boldsymbol{i}_t = \sigma\left(\boldsymbol{W}_i * \boldsymbol{X}_t + \boldsymbol{U}_i * \boldsymbol{H}_{t-1} + \boldsymbol{b}_i\right) \\ \tilde{\boldsymbol{C}}_t = \tanh\left(\boldsymbol{W}_c * \boldsymbol{X}_t + \boldsymbol{U}_c * \boldsymbol{H}_{t-1} + \boldsymbol{b}_c\right) \\ \boldsymbol{C}_t = \boldsymbol{f}_t \odot \boldsymbol{C}_{t-1} + \boldsymbol{i}_t \odot \tilde{\boldsymbol{C}}_t \\ \boldsymbol{o}_t = \sigma\left(\boldsymbol{W}_o * \boldsymbol{X}_t + \boldsymbol{U}_o * \boldsymbol{H}_{t-1} + \boldsymbol{b}_o\right) \\ \boldsymbol{H}_t = \boldsymbol{o}_t \odot \tanh\left(\boldsymbol{C}_t\right) \end{cases} \tag{6}$$

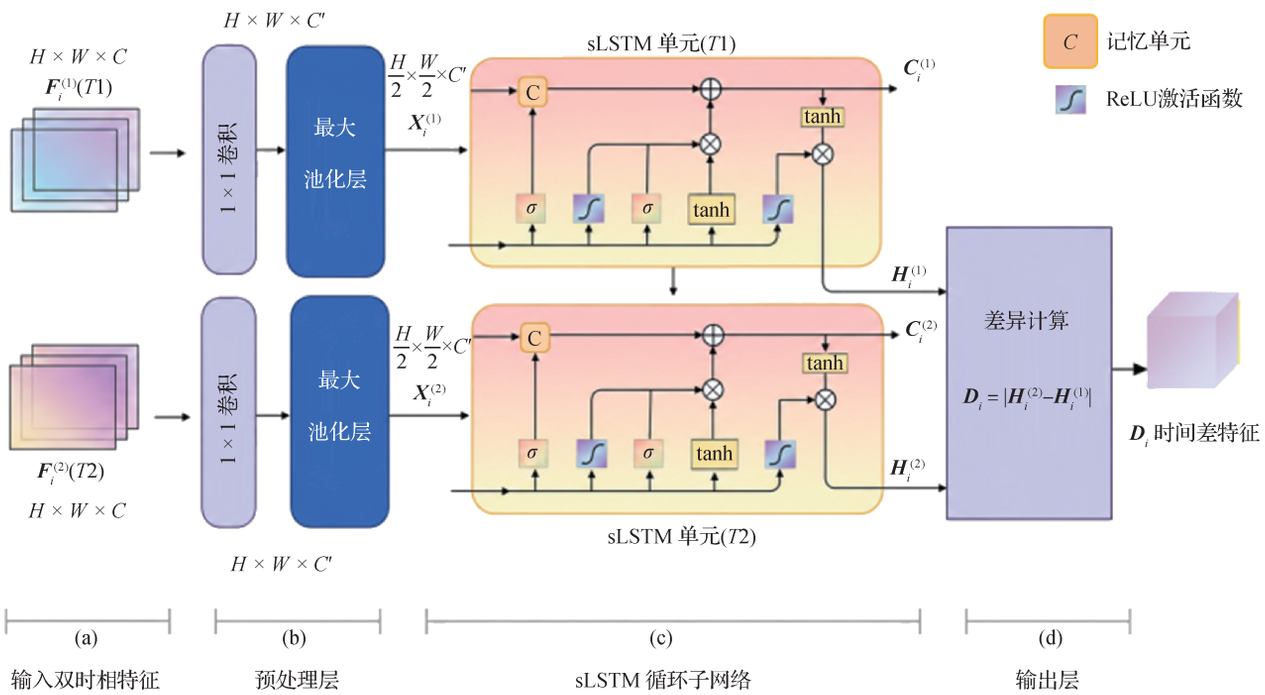


图2　时间存储器模块示意图

Fig. 2　Schematic diagram of the time storage module((a)input bitemporal features; (b)preprocessing layer; (c)sLSTM subnetwork; (d)output layer)

式中，$\boldsymbol{f}_t$、$\boldsymbol{i}_t$、$\boldsymbol{o}_t$分别表示遗忘门、输入门和输出门的输出；$\sigma(\cdot)$为 sigmoid 激活函数；$\boldsymbol{C}_t$和$\tilde{\boldsymbol{C}}_t$分别表示记忆状态和候选记忆状态；$\boldsymbol{H}_t$表示隐藏状态；$\boldsymbol{W}_f$、$\boldsymbol{U}_f$为对应门控的卷积核参数；$\boldsymbol{b}_f$为偏置项；⊙表示逐元素乘法；*表示二维卷积运算。

完成双时相时间建模后，将两次时间步的隐藏状态差异作为显式时间差特征，具体为

$$\boldsymbol{D}_i = \left| \boldsymbol{H}_i^{(2)} - \boldsymbol{H}_i^{(1)} \right| \tag{7}$$

式中，$\boldsymbol{D}_i$表示第 i 层时间差特征，反映双时相在该特征层级上的变化响应。

1.3 时域差引导策略

本文提出时域差引导特征提取策略，利用时间存储器模块输出的时空差异特征$\boldsymbol{D}_i$显式引导特征学习，使网络逐步聚焦真实变化区域。具体而言，将$\boldsymbol{D}_i$转换为空间注意力权重并反馈至双支路光谱特征提取网络。由于时间存储器模块采用最大池化下采样以降低计算复杂度，所得时间差特征$\boldsymbol{D}_i$的空间分辨率低于当前层级输入特征$\boldsymbol{F}_i^{(1)}$和$\boldsymbol{F}_i^{(2)}$。

为了实现像素级的特征对齐与引导，首先对$\boldsymbol{D}_i$进行上采样操作，使其空间维度恢复至与$\boldsymbol{F}_i^{(t)}$一致。随后，利用非线性激活函数将对齐后的强度映射为注意力权重矩阵，计算为

$$\boldsymbol{W}_i^{\text{guide}} = \delta\left(Up\left(\boldsymbol{D}_i\right)\right) \tag{8}$$

式中，$Up(\cdot)$表示双线性插值上采样操作（bilinear interpolation upsampling），用于恢复特征的空间分辨率；$\delta(\cdot)$表示激活函数 ReLU（rectified linear unit），用于将差异值映射为非负权重系数；$\boldsymbol{W}_i^{\text{guide}}$表示引导权重矩阵，其直观反映了当前层级的变化概率分布。

在获得引导权重后，将其分别作用于双时相光谱特征图$\boldsymbol{F}_i^{(1)}$和$\boldsymbol{F}_i^{(2)}$，通过逐元素相乘抑制背景噪声并增强变化区域响应。为避免深层网络中权重连乘导致的特征退化与梯度消失，引入残差连接，将增强特征与原始特征相加，作为下一阶段输入。修正后的双时相特征表示为

$$\hat{\boldsymbol{F}}_i^{(1)} = \boldsymbol{F}_i^{(1)} \odot \boldsymbol{W}_i^{\text{guide}} + \boldsymbol{F}_i^{(1)} \tag{9}$$

$$\hat{\boldsymbol{F}}_i^{(2)} = \boldsymbol{F}_i^{(2)} \odot \boldsymbol{W}_i^{\text{guide}} + \boldsymbol{F}_i^{(2)} \tag{10}$$

通过层级递进的反馈机制，时间存储器模块提取的时域差异不仅作为判别依据，还以先验形式嵌入特征提取过程。随着网络加深，该策略能够有效抑制光照、阴影及配准误差引起的伪变化干扰，引导模型聚焦真实变化区域，从而提升变化检测精度与鲁棒性。

1.4 损失函数

本文实验所采用的损失函数为 Dice loss，这是一种基于集合相似度度量的损失函数，计算为

$$L_{\text{Dice}} = 1 - \frac{2\sum_{i=1}^{N} p_i y_i + e}{\sum_{i=1}^{N} p_i + \sum_{i=1}^{N} y_i + e} \tag{11}$$

式中，N表示输入图像块中的像素总数；$p_i \in [0,1]$表示模型预测第 i 个像素属于变化类的概率值；$y_i \in \{0,1\}$表示第 i 个像素的真实标签（1 表示变化，0 表示未变化）；e为极小的平滑项，用于防止分母为零并增强数值稳定性。

相比交叉熵损失，Dice loss 直接基于预测概率与真实标签的交并比进行优化，而非逐像素独立计算分类误差。交叉熵损失对大量背景样本更为敏感，容易导致模型偏向多数类，从而忽略稀疏的变化区域；而 Dice loss 的核心思想是直接优化预测结果与真实变化区域之间的重叠程度，能够有效抑制背景样本的干扰，引导模型更加关注少量但关键的变化区域。因此，在变化检测这类正负样本极度不平衡的任务中，Dice loss 能更有效地提升模型的检测精度与鲁棒性。

2 实验结果和分析

2.1 实验设置

实验在配备 NVIDIA RTX 3090（24 GB）的工作站上完成，基于 TensorFlow-GPU 2.5.0 实现。采用 Adam 优化器（$\beta_1 = 0.9$，$\beta_2 = 0.999$），学习率设为 1×10^{-4}，训练 200 个 epoch。网络通道数按 8、16、32、64 逐级递增，以增强多层特征表达能力。为保证结果稳定性，所有实验基于随机采样重复 10 次，并统计平均值与标准差。性能评估采用总体精度（overall accuracy，OA）和 Kappa 系数，并结合定量与定性结果进行综合分析。

2.2 实验数据集

River、Hermiston 和 Farmland 数据集是高光谱变化检测中常用的公开数据集，均由 EO-1 卫星搭载的 Hyperion 传感器获取的多时相高光谱影像构成。River 数据集位于中国江苏省某河流流域，主要反映清淤等人为活动引起的河道变化，包含 2013 年 5 月

3日与12月31日两期影像。Hermiston数据集位于美国俄勒冈州，包含2004年5月1日与2007年5月8日影像，空间分辨率为30 m，尺寸为307×241像素，涵盖农田、水体及裸土等地物类型，并提供像素级标注。Farmland数据集位于江苏盐城地区，反映作物轮作引起的农田变化，包含2006年5月3日与2007年4月23日影像，尺寸为450×140像素，分辨率为30 m，同样保留154个有效波段，并提供标注。

2.3 实验结果

2.3.1 对比实验

为验证所提方法的有效性与优越性，选取多种代表性方法进行对比，具体包括：经典无监督方法CVA、光谱—空间多信息融合方法（spectral-spatial multi-information fusion，SSMIF）（Wu等，2024）、基础CNN方法2DCNN、HyperNet、门控增强Transformer网络（gated enhanced transformer network，GETNET）、通道—空间注意力网络（channel-spatial attention network，CSANet）（Song等，2022）、多层级编码—解码注意力网络（multi-level encoder-decoder attention network，ML-EDAN）（Qu等，2022）、光谱—空间—时间Transformer（spectral-spatial-temporal Transformer，SST-Former）（Wang等，2022）。分别从无监督、空间特征提取、注意力建模、时序建模及全局依赖建模等角度进行对比。实验结果如表1所示。

整体来看，深度学习方法均显著优于传统CVA方法，说明基于深度特征学习的模型能够更有效地挖掘双时相影像中的变化信息。相比之下，本文方法在River、Hermiston和Farmland 3个数据集上均取得了最优性能，充分验证了模型的有效性、稳定性以及良好的跨场景泛化能力。在River数据集上，OA和Kappa分别为96.54%和78.51%，较SST-Former提升0.40%和1.01%，表现出更强的判别能力。在Hermiston数据集上，OA和Kappa分别为98.21%和91.79%，较SST-Former提升0.80%和0.03%，仍保持优势。在Farmland数据集上，OA和Kappa分别为95.87%和90.11%，较SSMIF提升1.29%和3.10%，体现出更强的鲁棒性。

表1 不同方法在3个数据集上的检测精度对比

Table 1 Quantitative comparison of different methods on three datasets

/%

方法	River		Hermiston		Farmland	
	OA	Kappa	OA	Kappa	OA	Kappa
CVA	85.73	42.72	94.54	83.86	75.23	45.62
SSMIF	95.83	75.68	96.32	86.89	<u>94.58</u>	<u>87.01</u>
HyperNet	94.97	70.95	93.96	82.21	90.70	76.83
2DCNN	95.17	73.17	96.00	88.27	92.54	83.20
GETNET	95.34	74.28	96.56	90.18	93.97	86.37
ML-EDAN	95.55	73.51	96.03	90.21	93.37	85.06
CSANet	95.76	74.49	95.92	88.03	93.82	83.25
SST-Former	<u>96.14</u>	<u>77.50</u>	<u>97.41</u>	<u>91.76</u>	94.02	85.20
本文	**96.54**	**78.51**	**98.21**	**91.79**	**95.87**	**90.11**

注：加粗、下划线字体分别表示各列最优、次优结果。

此外，从可视化结果（图3—图5）可见，相比其他对比方法，本文方法在细节保留与噪声抑制方面表现最佳，预测更接近真实标签。在Farmland数据集上，本文方法能够有效抑制椒盐噪声，并保持小目标变化区域的完整性；在River数据集上，模型能够准确保留河流边缘及细长变化区域，获得更好的边界连续性；在Hermiston数据集上，模型对大范围变化区域具有更强的一致性识别能力，减少了区域内部的断裂和空洞现象。上述结果表明，本文方法不仅能够充分融合双时相特征信息，而且能够同时兼顾全局语义建模与局部细节表达，从而在不同场景下均获得更加准确和鲁棒的变化检测结果。

2.3.2 消融实验

为验证各关键模块对模型性能的贡献，在River、Farmland和Hermiston数据集上开展消融实验。以仅包含Vision Mamba主干的模型为基线，逐步引入时间存储器模块（TSM）和时域差引导策略，分析其对变化检测性能的影响，结果如表2所示。

实验结果表明，移除时间存储器模块（TSM）或时域差引导策略后，模型在3个数据集上的性能均出现不同程度下降，说明二者在时间差异建模与空间特征增强中均具有重要作用。其中，仅引入TSM时，River、Farmland和Hermiston数据集上的OA分别提升2.18%、3.53%和2.20%，明显优于仅使用时域差引导策略的情况，表明TSM基于sLSTM的时间建模机制能够有效捕获双时相依赖关系并提取细粒度时间差异特征，从而增强变化感知能力。当两者联合使用时，模型达到最佳性能，OA分别提升3.41%、4.73%和2.89%，Kappa系数同步提升，说明时域差

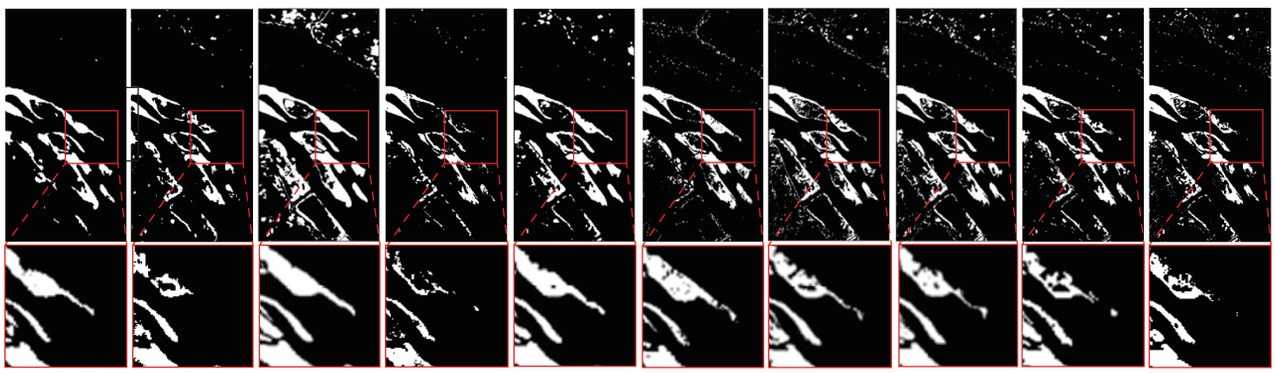


图 3　在 River 数据集上不同方法的二分类图

Fig. 3　Binary classification maps of different methods on the River dataset ((a) CVA; (b) SSMIF; (c) HyperNet; (d) 2DCNN; (e) GETNET; (f) ML-EDAN; (g) CSANet; (h) SST-Former; (i) TDG-Net (ours); (j) ground truth)

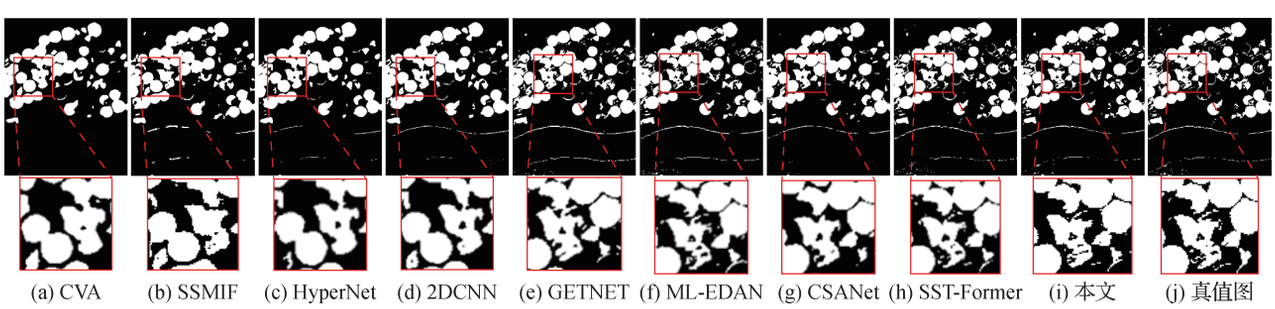


图 4　在 Hermiston 数据集上不同方法的二分类图

Fig. 4　Binary classification maps of different methods on the Hermiston dataset ((a) CVA; (b) SSMIF; (c) HyperNet; (d) 2DCNN; (e) GETNET; (f) ML-EDAN; (g) CSANet; (h) SST-Former; (i) TDG-Net (ours); (j) ground truth)

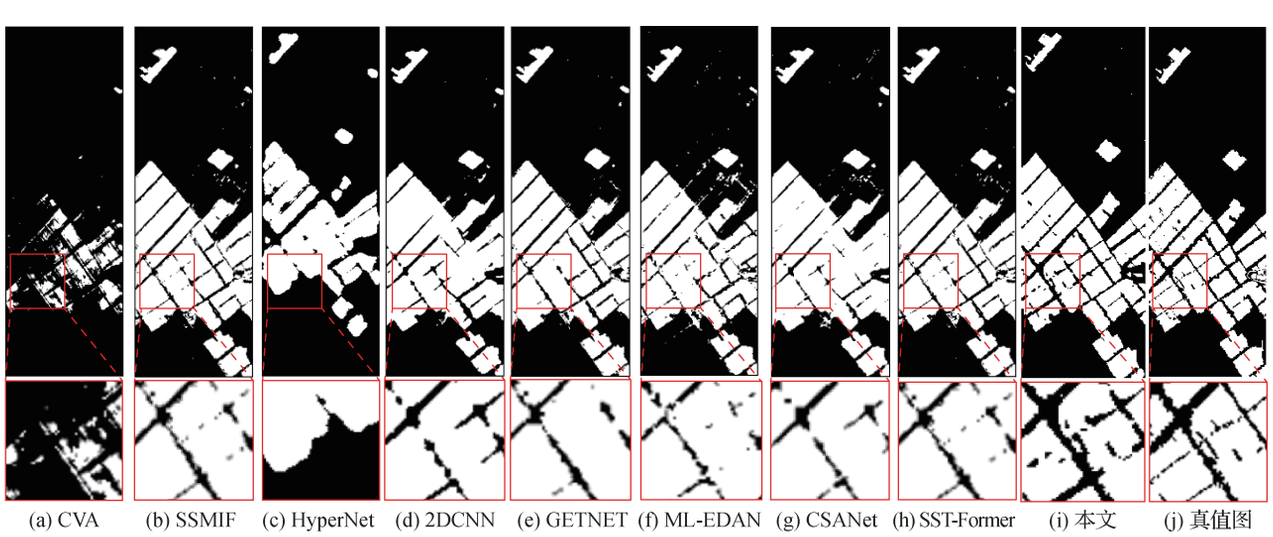


图 5　在 Farmland 数据集上不同方法的二分类图

Fig. 5　Binary classification maps of different methods on the Farmland dataset ((a) CVA; (b) SSMIF; (c) HyperNet; (d) 2DCNN; (e) GETNET; (f) ML-EDAN; (g) CSANet; (h) SST-Former; (i) TDG-Net (ours); (j) ground truth)

引导策略可在 TSM 生成的时间差异特征基础上进一步强化空间注意力响应，抑制光照、阴影及噪声引起的伪变化，从而提升对真实变化区域的关注。总体来看，两模块在不同数据集上均表现出良好稳定性与互补性，验证了所提出方法在时间建模与空间增强协同作用上的有效性与必要性。

表2　3个数据集上消融实验结果

Table 2　Results of ablation experiments on three datasets

/%

方法	时间存储器模块	时域差引导策略	River		Farmland		Hermiston	
			OA	Kappa	OA	Kappa	OA	Kappa
1	×	×	93.13	50.21	91.14	81.66	95.32	78.21
2	√	×	95.31	70.92	94.67	86.55	97.52	90.42
3	×	√	94.87	65.35	93.21	84.29	96.32	85.31
4	√	√	**96.54**	**78.51**	**95.87**	**90.11**	**98.21**	**91.79**

注：加粗字体表示各列最优结果。“×”表示未使用该模块，“√”表示使用该模块。

3　讨　论

为进一步验证所提方法在模型复杂度、参数设置及复杂场景下的实际性能表现，本文从模型计算效率、超参数敏感性以及鲁棒性3个方面展开分析。考虑到Hermiston数据集具有较复杂的地物分布与细粒度变化特征，能够较好地反映模型对时空变化信息的建模能力，因此本节讨论均以Hermiston数据集实验结果为例进行分析。

3.1　模型复杂度与推理效率分析

为进一步验证所提方法在模型规模与计算效率方面的优势，统计了不同方法的参数量（Params）、FLOPs以及单幅图像平均推理时间，结果如表3所示。相比SST-Former、CSANet等Transformer类方法，本文方法在保持较低模型复杂度的同时取得了更优的检测精度，表明所设计的时间存储器模块与时域差引导策略能够有效提升时间差异建模能力，并在精度与效率之间实现较好的平衡。此外，传统方法虽然具有较快推理速度，但由于缺乏深层特征表达能力，其检测性能明显低于深度学习方法。综合来看，本文方法在检测精度、模型规模与推理效率之间取得了较优折中，具备较好的实际应用潜力。

表3　不同方法的模型复杂度与推理效率对比

Table 3　Comparison of model complexity and inference efficiency of different methods

方法	Params/M	FLOPs/G	推理时间/ms
CVA	–	–	**3.2**
SSMIF	2.15	5.84	14.7
HyperNet	3.42	8.63	19.5
2DCNN	**1.87**	**4.25**	11.8
GETNET	2.96	7.41	16.2
ML-EDAN	5.28	13.76	27.4
CSANet	6.14	15.83	29.1
SST-Former	8.37	22.45	36.8
本文	4.12	10.37	18.6

注：加粗字体表示各列最优结果。“–”表示传统方法无可学习参数或难以统一统计其计算复杂度。

3.2　超参数敏感性分析

为分析关键超参数对模型性能的影响，本文分别对学习率、池化大小以及通道数配置进行了敏感性实验，结果如表4—表6所示。学习率实验表明，当学习率设置为1×10^{-4}时，模型取得最佳检测性能，OA与Kappa分别达到98.21%和91.79%。较大的学习率虽然能够加快收敛速度，但容易导致训练振荡；而较小的学习率则会降低模型收敛效率。池化大小实验结果表明，2×2池化在保留空间结构信息与降低特征冗余之间取得了较好平衡，检测性能优于1×1、3×3及4×4设置。对于网络通道数配置，8/16/32/64在模型复杂度与检测精度之间表现出较优折中。虽然16/32/64/128配置能够进一步提升少量性能，但参数量与计算开销显著增加，因此本文最终采用8/16/32/64作为网络通道配置。

3.3　鲁棒性分析

为验证本文方法在复杂场景下的稳定性与泛化能力，进一步开展了不同噪声水平下的鲁棒性实验，结果如表7所示。在噪声实验中，随着高斯噪声强度增加，各方法检测性能均出现下降，但本文方法仍保持较高的OA与Kappa系数。当噪声方差增大至0.1时，OA仍达到95.88%，表明模型具有较

表 4　不同学习率下的敏感性分析

Table 4　Sensitivity analysis under different learning rates

学习率	OA/%	Kappa/%
1×10^{-3}	96.84	88.21
5×10^{-4}	97.56	89.74
1×10^{-4}	**98.21**	**91.79**
1×10^{-5}	97.31	89.18

注：加粗字体表示各列最优结果。

表 5　不同池化大小下的敏感性分析

Table 5　Sensitivity analysis under different pooling sizes

池化大小	OA/%	Kappa/%
1 × 1	97.42	89.63
2 × 2	**98.21**	**91.79**
3 × 3	97.68	90.12
4 × 4	96.95	88.57

注：加粗字体表示各列最优结果。

表 6　不同通道数配置下的敏感性分析

Table 6　Sensitivity analysis under different channel configurations

通道数配置	Params/M	FLOPs/G	OA/%	Kappa/%
4/8/16/32	2.31	6.42	97.26	89.04
8/16/32/64	4.12	10.37	98.21	91.79
16/32/64/128	8.35	21.64	98.34	91.96

好的抗噪能力。这主要得益于时间存储器模块对时间差异特征的显式建模以及时域差引导策略对伪变化区域的有效抑制。实验结果表明，本文方法在复杂噪声环境下具有较好的鲁棒性与实用价值。

表 7　不同噪声水平下的鲁棒性分析

Table 7　Robustness analysis under different noise levels

高斯噪声方差	OA/%	Kappa/%
0	**98.21**	**91.79**
0.01	97.84	91.02
0.03	97.26	89.85
0.05	96.71	88.63
0.1	95.88	86.94

注：加粗字体表示各列最优结果。

4　结　论

针对现有方法仅在决策层进行特征对比、难以充分利用多层信息且忽略双时相时间差异的问题，本文提出一种时间差引导的高光谱图像变化检测方法。该方法以 Vision Mamba 为主干提取双时相空间特征，并设计时间存储器模块 TSM，在谱维压缩与空间下采样基础上引入 sLSTM 进行时序建模，通过门控与记忆机制提取显式时间差特征并抑制不变区域响应。在此基础上，构建时域差引导策略，将时间差特征映射为空间注意力权重，并通过残差连接反馈至双分支网络，引导多层特征学习聚焦真实变化区域。此外，引入 Dice loss 以缓解类别不平衡问题。

在 River、Hermiston 和 Farmland 数据集上的实验结果表明，所提方法能够有效利用多层级时间差异信息，在总体精度（OA）和 Kappa 系数上均优于多种先进方法，同时在模型规模、计算复杂度与推理效率之间取得了良好平衡，验证了模型的有效性与鲁棒性。未来工作将进一步面向标注样本稀缺场景，引入自监督学习策略以提升模型泛化能力；同时探索多模态数据（如光学与合成孔径雷达（synthetic aperture radar，SAR））融合下的变化检测方法，并研究端到端联合优化框架，以进一步提升模型性能与实际应用价值。

参考文献（References）

Bai T, Wang L, Yin D M, Sun K M, Chen Y P, Li W Z, et al. 2023. Deep learning for change detection in remote sensing: a review. Geo-Spatial Information Science, 26(3): 262-288 [DOI: 10.1080/10095020.2022.2085633]

Bhargava A, Sachdeva A, Sharma K, Alsharif M H, Uthansakul P and Uthansakul M. 2024. Hyperspectral imaging and its applications: a review. Heliyon, 10(12): #33208 [DOI: 10.1016/j.heliyon.2024.e33208]

Bruzzone L and Carlin L. 2006. A multilevel context-based system for classification of very high spatial resolution images. IEEE Transactions on Geoscience and Remote Sensing, 44(9): 2587-2600 [DOI: 10.1109/TGRS.2006.875360]

Fauvel M, Tarabalka Y, Benediktsson J A, Chanussot J and Tilton J C. 2013. Advances in spectral-spatial classification of hyperspectral images. Proceedings of the IEEE, 101(3): 652-675 [DOI: 10.1109/JPROC.2012.2197589]

Gao H M, Li J Y, Chen Z H and Xu S F. 2025. Spectral attention and visionmamba difference guided network for hyperspectral image change detection. Optics and Laser Technology, 192: #113869 [DOI: 10.1016/j.optlastec.2025.113869]

Hu M Q, Wu C and Du B. 2023. EMS-Net: efficient multi-temporal self-attention for hyperspectral change detection//2023 IEEE International Geoscience and Remote Sensing Symposium. Pasadena, USA: IEEE: 6664-6667 [DOI: 10.1109/IGARSS52108.2023.10282722]

Hu M Q, Wu C and Zhang L P. 2022. HyperNet: self-supervised hyperspectral spatial-spectral feature understanding network for hyperspectral change detection. IEEE Transactions on Geoscience and Remote Sensing, 60: #5543017 [DOI: 10.1109/TGRS.2022.3218795]

Li X R and Ding J G. 2023. Spectral-temporal transformer for hyperspectral image change detection. Remote Sensing, 15(14): #3561 [DOI: 10.3390/rs15143561]

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(4): 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(4): 1017-1026) [DOI: 10.11834/jig.240360]

Liu W L, Gao F, Zhang H P, Dong J Y and Wu C T. 2025. Channel shuffle and Transformer-based hyperspectral images change detection method. Chinese Journal of Computers, 48(4): 971-984 (刘文力, 高峰, 张浩鹏, 董军宇, 吴淳桐. 2025. 基于通道变换和Transformer的高光谱图像变化检测方法. 计算机学报, 48(4): 971-984) [DOI: 10.11897/SP.J.1016.2025.00971]

Lu D and Weng Q. 2007. A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing, 28(5): 823-870 [DOI: 10.1080/01431160600746456]

Malila W A. 1980. Change vector analysis: an approach for detecting forest changes with Landsat//Proceedings of the 6th Annual Symposium on Machine Processing of Remotely Sensed Data. West Lafayette, USA: Purdue University: 326-335

Mou L C, Bruzzone L and Zhu X X. 2019. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. IEEE Transactions on Geoscience and Remote Sensing, 57(2): 924-935 [DOI: 10.1109/TGRS.2018.2863224]

Nielsen A A. 2007. The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data. IEEE Transactions on Image Processing, 16(2): 463-478 [DOI: 10.1109/TIP.2006.888195]

Plaza A, Benediktsson J A, Boardman J W, Brazile J, Bruzzone L, Camps-Valls G, et al. 2009. Recent advances in techniques for hyperspectral image processing. Remote Sensing of Environment, 113(S1): S110-S122 [DOI: 10.1016/j.rse.2007.07.028]

Qu J H, Hou S X, Dong W Q, Li Y S and Xie W Y. 2022. A multilevel encoder-decoder attention network for change detection in hyperspectral images. IEEE Transactions on Geoscience and Remote Sensing, 60: #5518113 [DOI: 10.1109/TGRS.2021.3130122]

Qu J H, Zhao J Y, Dong W Q, He J, Li Z and Li Y S. 2025. CDAFormer: hybrid transformer-based contrastive domain adaptation framework for unsupervised hyperspectral change detection. Neural Networks, 190: #107633 [DOI: 10.1016/j.neunet.2025.107633]

Radke R J, Andra S, Al-Kofahi O and Roysam B. 2005. Image change detection algorithms: a systematic survey. IEEE Transactions on Image Processing, 14(3): 294-307 [DOI: 10.1109/TIP.2004.838698]

Singh A. 1989. Review article digital change detection techniques using remotely-sensed data. International Journal of Remote Sensing, 10(6): 989-1003 [DOI: 10.1080/01431168908903939]

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: #6010105 [DOI: 10.1109/LGRS.2022.3179134]

Thenkabail P S, Enclona E A, Ashton M S and Van Der Meer B. 2004. Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. Remote Sensing of Environment, 91(3/4): 354-376 [DOI: 10.1016/j.rse.2004.03.013]

Wang Q, Yuan Z H, Du Q and Li X L. 2019. GETNET: a general end-to-end 2-D CNN framework for hyperspectral image change detection. IEEE Transactions on Geoscience and Remote Sensing, 57(1): 3-13 [DOI: 10.1109/TGRS.2018.2849692]

Wang Y H, Hong D F, Sha J J, Gao L R, Liu L, Zhang Y G, et al. 2022. Spectral-spatial-temporal transformers for hyperspectral image change detection. IEEE Transactions on Geoscience and Remote Sensing, 60: #5536814 [DOI: 10.1109/TGRS.2022.3203075]

Wu L X, Peng J T, Yang B, Sun W W and Ye Z J. 2024. SSMIF: enhanced spatial-spectral Mamba interactive fusion network for hyperspectral change detection//Proceedings of 2024 IEEE International Conference on Signal, Information and Data Processing. Zhuhai, China: IEEE: 1-5 [DOI: 10.1109/ICSIDP62679.2024.10868119]

Xiao D F, Yu D B, Wang Y and Xiao J. 2026. SE-Mamba: hyperspectral image classification method based on spatial enhancement [J/OL]. Journal of Image and Graphics, 1-19 (肖岱峰, 于东伯, 王颖, 肖俊. 2026. SE-Mamba: 空间信息增强的高光谱图像分类方法[J/OL]. 中国图象图形学报, 1-19) [DOI: 10.11834/jig.260119]

Yang L X, Bao Y J, Zhang R and Yang S Y. 2026. Lightweight Res-3D-CNN with Transformer embedding for cross-domain hyperspectral image classification. Journal of Image and Graphics, 31(3): 927-943 (杨丽霞, 鲍雅君, 张瑞, 杨淑媛. 2026. 面向跨域高光谱图像分类的嵌入Transformer层的轻量型Res-3D-CNN. 中国图象图

形学报，31(3)：927-943)［DOI：10.11834/jig.250020］

Zhang H T，Chen K Y，Liu C Y，Chen H，Zou Z X and Shi Z W. 2025. CDMamba：incorporating local clues into Mamba for remote sensing image binary change detection. IEEE Transactions on Geoscience and Remote Sensing，63：#4405016［DOI：10.1109/TGRS.2025.3545012］

Zhao W Z and Du S H. 2016. Spectral-spatial feature extraction for hyperspectral image classification：a dimension reduction and deep learning approach. IEEE Transactions on Geoscience and Remote Sensing，54(8)：4544-4554［DOI：10.1109/TGRS.2016.2543748］

Zhou C L，Shi Q，Li J and Zhang X C. 2024. Spectral-frequency domain attribute pattern fusion for hyperspectral image change detection. National Remote Sensing Bulletin，28(1)：105-120（周承乐，石茜，李军，张新长. 2024. 光谱—频域属性模式融合的高光谱遥感图像变化检测. 遥感学报，28(1)：105-120）［DOI：10.11834/jrs.20232600］

Zhu X X，Tuia D，Mou L C，Xia G S，Zhang L P，Xu F，et al. 2017. Deep learning in remote sensing：a comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine，5(4)：8-36［DOI：10.1109/MGRS.2017.2762307］

Zou C Z，Liang W F，Liu L and Zou C W. 2024. Hyperspectral image change detection based on an improved multi-scale and spectral-wise transformer. International Journal of Remote Sensing，45(6)：1904-1925［DOI：10.1080/01431161.2024.2321469］

作者简介

丁晨，男，副教授，主要研究方向为图像处理、机器学习和模式识别。E-mail：dingchen@xupt.edu.cn

张景博，男，硕士研究生，主要研究方向为遥感图像处理。E-mail：zjb@stu.xupt.edu.cn

郝小锋，男，硕士研究生，主要研究方向为遥感图像处理。E-mail：hxf@stu.xupt.edu.cn

郑思睿，男，硕士研究生，主要研究方向为遥感图像处理。E-mail：zhengsirui@stu.xupt.edu.cn

闫松，男，高级工程师，主要研究方向为计算机视觉与光学故障诊断研究。E-mail：ys060599@126.com