

# 基于通道变换和Transformer的高光谱图像变化检测方法

刘文力<sup>1),2)</sup> 高峰<sup>1),2)</sup> 张浩鹏<sup>1),2)</sup> 董军宇<sup>1),2)</sup> 吴淳桐<sup>1),2)</sup>

<sup>1)</sup>(中国海洋大学计算机科学与技术学院 山东 青岛 266100)

<sup>2)</sup>(海洋信息技术教育部工程研究中心 山东 青岛 266100)

**摘 要** 当前基于Transformer的高光谱图像变化检测方法通过自注意力机制模拟长距离依赖,能够有效建模全局上下文信息。然而,现有方法仍面临着两个主要问题:一是Transformer模型计算复杂度高,导致模型在处理高维度数据时效率低下;二是现有模型对高光谱图像的波段信息利用有限,在光谱维度上缺乏特征交互。针对这些问题,本文提出了一种基于通道变换和Transformer的高光谱图像变化检测方法,以更高效地利用高光谱图像中复杂的光谱和空间信息。创新之处主要体现在两个方面:其一,采用基于通道变换和注意力机制的特征提取模块。该模块改进了传统自注意力计算方式并加入通道信息交互模块,一方面降低了传统Transformer二次方的计算复杂度,使模型更适用于处理高维度数据;另一方面实现了对高光谱图像空间和光谱信息的高效利用,增强了模型对高级语义信息的理解及对复杂变化的感知能力。其二,设计了双分支门控前馈神经网络。该网络实现了模型对特征信息的细粒度调控,提升了模型对关键地物变化和细微差异的捕捉能力。实验结果显示,本文方法在River和Hermiston数据集上的准确率分别达到了96.28%和95.97%,Kappa系数分别达到了79.44%和88.90%。相比于当前主流方法,本文模型在这两个数据集上准确率分别提升了0.60%和0.69%,Kappa系数也分别提升了10.30%和2.33%,验证了本文方法在高光谱图像变化检测任务中的有效性。

**关键词** 变化检测;高光谱图像;注意力机制;双分支门控前馈神经网络;通道变换模块

中图分类号 TP391

DOI号 10.11897/SP.J.1016.2025.00971

## Channel Shuffle and Transformer-Based Hyperspectral Images Change Detection Method

LIU Wen-Li<sup>1),2)</sup> GAO Feng<sup>1),2)</sup> ZHANG Hao-Peng<sup>1),2)</sup> DONG Jun-Yu<sup>1),2)</sup> WU Chun-Tong<sup>1),2)</sup>

<sup>1)</sup>(School of Computer and Technology, Ocean University of China, Qingdao, Shandong 266100)

<sup>2)</sup>(Engineering Research Center of the Ministry of Education for Marine Information Technology, Qingdao, Shandong 266100)

**Abstract** The existing hyperspectral image change detection methods based on Transformer simulate long-range dependencies through self-attention mechanisms, effectively modeling global contextual information. However, the current methods still face two main challenges: one is the high computational complexity of the Transformer models, leading to inefficiencies when processing high-dimensional data; another is the limited utilization of spectral information in hyperspectral images, resulting in a lack of spectral dimension feature interactions, which results in inadequate feature interaction along the spectral dimension and hinders the comprehensive understanding and

收稿日期:2024-06-25;在线发布日期:2025-02-24。本课题得到新一代人工智能国家科技重大专项(2022ZD0117202)、国家自然科学基金(42106191)资助。刘文力,硕士研究生,主要研究领域为遥感图像处理、深度学习。E-mail: 2377087649@qq.com。高峰(通信作者),博士,副教授,主要研究领域为遥感图像处理、深度学习。E-mail: gaofeng@ouc.edu.cn。张浩鹏,硕士研究生,主要研究领域为遥感图像处理、深度学习。董军宇,博士,教授,主要研究领域为海洋大数据挖掘、水下视觉信息处理。吴淳桐,硕士研究生,主要研究领域为遥感图像处理、深度学习。

exploitation of the feature information in hyperspectral images. To address these issues, this paper proposes a novel hyperspectral image change detection method based on channel shuffle and Transformer. The proposed method aims to more effectively enhance the utilization of the complex spectral and spatial information in hyperspectral images. Specifically, a feature extraction module based on channel shuffle and attention mechanism is designed, improving the traditional self-attention computation by serially integrating channel information interaction during modeling global attention. This not only facilitates efficient utilization of both spatial and spectral information in hyperspectral images but also enhances the model's understanding of high-level semantic information and perception of complex changes. Additionally, this design also improves computational efficiency by reducing the quadratic computational complexity of traditional Transformer models. Consequently, this model is more adaptable for processing high-dimensional data. Furthermore, a dual-branch gated feedforward neural network is proposed. The network, which is designed in parallel, simultaneously employs two GELU activation functions and element-wise multiplication operations to more effectively filter potential noise and capture the local information by combining different convolution operations, so as to realize fine-grained modulation of the feature information and improve the model's ability to capture the changes and subtle difference of the key features in hyperspectral images. This design not only enhances the model's ability to transform nonlinear features but also improves its capability to capture complex relationships in hyperspectral images. The proposed modules are finally incorporated into the twin network structure, which is utilized to process the dual-temporal-phase hyperspectral images. In the twin network structure, these modules can work together and complement each other, thus enabling the model to more accurately capture the change information between dual-temporal-phase images and can comprehensively learn the information between each temporal-phase image. It can also assist the model in capturing the changes between different temporal phases and realizing the sensitive detection of the changes of the features. The proposed method effectively addresses the challenges faced by existing hyperspectral image change detection methods by enhancing the utilization of spectral and spatial information and reducing computational complexity. Through the integration of channel shuffle, attention mechanism, and dual-branch gated feedforward neural network into a twin network structure, the proposed method achieves significant improvements in change detection accuracy and sensitivity. Experimental results show that the percentage correct classification (PCC) of this paper's method on the River and Hermiston datasets reaches 96.28% and 95.97%, and the kappa coefficient (KC) reaches 79.44% and 88.90%, respectively. Compared with CDFormer, the percentage correct classification of this paper's model on these two datasets is improved by 0.60% and 0.69%, respectively, and the kappa coefficient is also improved by 10.30% and 2.33%, which verifies the effectiveness of this paper's method in the hyperspectral image change detection task.

**Keywords** change detection; hyperspectral image; attention mechanism; dual-branch gated feedforward neural network; channel shuffle block

## 1 引 言

遥感图像变化检测(Change Detection, CD)是一种利用多时相遥感图像来检测地表或地物在时

间上变化的方法。这项技术已经被广泛应用于国民经济、社会生活和国家安全的各个方面,为环境监测<sup>[1]</sup>、城市规划<sup>[2]</sup>和灾害评估<sup>[3]</sup>等领域提供了数据支撑。这些影像数据通常来自普遍的光学传感器,使用这些普通传感器的局限性在于获取的遥感图

像的波段数量有限,在一定程度上限制了对地物细节和复杂变化的解析能力。而利用高光谱图像传感器获取的高光谱图像(Hyperspectral Image, HSI)具有丰富的光谱信息,在地物变化检测中具有显著的应用潜力。高光谱图像中的每个像素包含了从可见光到短波红外光的全频谱反射特性,可以为区别不同地物提供强有力的依据。用户可以根据需求选择不同的波段进行分析,从而更好地适应变化检测的实际需求。此外,高光谱图像的波段间隔小,使得不同类型的地物在不同波段下会表现出独特的反射特性,利用这些细微的反射差异,研究人员能够捕捉到地物表面更微妙的变化。因此,高光谱图像在变化检测领域展现出巨大的优势,利用这些光谱特性,可以帮助研究人员更精准地识别和分析地物的变化。

传统的变化检测研究主要致力于开发和改进基于遥感影像、计算机视觉和机器学习的方法,利用这些方法来识别和分析地理和环境变化。传统方法如影像差分法(Image Difference Method, IDM<sup>[4]</sup>)、主成分分析(Principal Component Analysis, PCA<sup>[5]</sup>)和变化矢量分析(Change Vector Analysis, CVA<sup>[6]</sup>)提供了简单而有效的检测手段。这些方法通过对图像进行简单的数学运算或统计分析来检测变化,使用这些方法进行变化检测容易忽略数据中的上下文信息,对于复杂的场景可能会产生误报或漏报。近年来,基于深度学习的方法也被广泛应用于变化检测,如卷积神经网络(Convolutional Neural Network, CNN<sup>[3]</sup>)、循环神经网络(Recurrent Neural Network, RNN<sup>[7]</sup>)、长短期记忆网络(Long Short-Term Memory Network, LSTM<sup>[8]</sup>)和Transformer<sup>[9]</sup>等模型通过学习数据中的高级特征和复杂模式,可以有效提高变化检测的准确性。卷积神经网络通过卷积、池化和全连接层等组件来提取特征,学习其中的复杂关系;Transformer模型则利用自注意力机制来捕捉序列中的依赖关系,增强模型对全局上下文的理解。这些方法能够有效处理大规模复杂数据,尤其适用于对精度和效率要求较高的变化检测任务,未来针对这一领域的研究将持续呈现迅猛发展的趋势。

尽管国内外学者已经围绕高光谱图像变化检测开展了大量研究,并在多个公开数据集上取得了较好的检测效果,但由于高光谱图像的复杂特性,现有的研究仍难以满足遥感图像应用的实际需求,主要因为如下两个技术难点:

#### (1) Transformer模型计算复杂度高

Transformer模型的计算复杂度通常为二次方级别,这意味着在处理大尺寸的高光谱数据时需要庞大的计算资源。并且高光谱数据通常具有数百甚至数千个波段,使用传统的Transformer模型来处理这些波段信息可能会面临计算资源不足。因此,需要寻找高效的计算策略和精简的模型结构来确保在保持模型性能的同时降低计算成本。

#### (2) 光谱信息利用受限,在光谱维度上缺乏特征交互

高光谱图像具有丰富的波段信息,每个像素点包含了多个波段的光谱反射率或辐射亮度值,这些波段信息能够提供丰富的光谱特征。然而,现有的模型往往只关注建模空间注意力,忽视了对图像波段信息的依赖性,导致了模型在通道维度上的交互不足。这种局限性使得模型无法充分利用高光谱图像中每个波段的信息,限制了模型对图像整体特征的理解和表达能力,从而影响变化检测任务的准确性。

为了解决以上两个问题,本文提出了一种基于通道变换和Transformer的高光谱图像变化检测方法CSTFormer(Channel Shuffle and Transformer)。具体而言,本文改进了传统自注意力的计算方式,降低了Transformer模型的计算复杂度;还提出了利用通道分割与移动来实现通道信息交互的通道变换模块(Channel Shuffle Block, CSB),利用这一模块可以有效增强模型对高光谱图像中光谱信息的利用。此外,本文还设计了一个双分支门控前馈神经网络(Dual-branch Gated Feedforward Neural Network, DGFFN)来替换Transformer模块中的FFN,使用DGFFN可以使模型更有针对性地过滤数据中的无效信息,提高变化检测的准确性。本文的主要贡献如下:

(1)设计了一种基于通道变换和注意力机制的特征提取模块,利用孪生网络结构对输入的双时相高光谱图像进行全局建模,改进了传统自注意力的计算方式,串行地结合了通道变换模块来增强通道信息交互,实现了对高光谱图像中空间和光谱信息的高效利用。

(2)基于门控前馈神经网络,本文提出了一种全新的双分支门控前馈神经网络DGFFN,使用这种网络不仅能够有效过滤高光谱图像中的无用信息,还可以很好地捕捉其中的复杂关系,提升模型对复杂信息的表达能力。



(3)针对 River 和 Hermiston 两个高光谱图像数据集,利用对比实验和消融实验从定量分析和定性分析两个角度验证了本文所提出的 CSTFormer 模型在变化检测方面的有效性。

## 2 相关工作

### 2.1 传统方法

传统方法在遥感图像变化检测领域扮演着重要角色,其主要包括基于像素的变化检测和基于对象的变化检测。基于像素的方法从单个像素或者单个像素的邻域提取深度特征,逐像素地比较两个时相图像的变化<sup>[10]</sup>,典型的方法包括基于阈值的变化检测法<sup>[11]</sup>、图像差分法<sup>[12]</sup>以及图像比值法<sup>[13]</sup>。基于阈值的方法通过将两幅图像的差异与预定义的阈值进行比较,从而确定发生变化的像素。图像差分法则是通过计算两幅图像的差异,然后根据差异值进行分类来检测变化。图像比值法则是根据两个时相图像的比率来判断变化,通常用于消除图像之间的光照和辐射条件的差异。尽管这些方法简单易行,但在处理噪声、亮度不一致等问题上仍然存在一定的局限性<sup>[14]</sup>。

基于对象的方法提取的是对应一定实体、内部相对均匀的像素组合,这些对象能够有效结合光谱和空间纹理特征<sup>[15]</sup>,提供更精确的地物信息<sup>[16]</sup>。在基于对象的方法中,对象生成是最重要的步骤之一,会直接影响网络的检测性能,通常使用目标分割技术<sup>[17]</sup>生成对象<sup>[18, 19, 20]</sup>,使用模型分析这些对象来检测变化。基于对象的比较分析利用了对观测条件不敏感的地物空间特征,如图像中物体的形状、大小、纹理以及空间关系,在高分辨率遥感图像变化检测中具有一定优势。通过比较这些特征,可以有效减弱随机噪声和季节变化对变化检测结果的影响<sup>[21]</sup>。

传统方法以其简单易行、计算效率高和硬件要求低的优点,在许多基础应用中表现良好。然而,由于图像特征提取和变化检测算法的局限性,这些方法在复杂场景中的准确性和鲁棒性可能不如深度学习方法。为了克服这些局限性,实际应用中常将传统方法与现代技术结合使用<sup>[22]</sup>,以充分发挥各自的优势。

### 2.2 深度学习方法

深度学习方法在遥感图像变化检测领域具有十分重要的地位,基于深度学习的模型具有强大的数据拟合能力,能够通过学习大规模数据集中的复杂

模式和特征来提高变化检测的准确性<sup>[23]</sup>。此外,基于深度学习的模型具有端到端学习的特点<sup>[24]</sup>,能够直接从原始数据中学习特征表示,无需手动设计特征提取器,可以有效简化算法流程。

近年来,基于深度学习的方法在高光谱图像变化检测领域得到了广泛关注,许多研究者提出了创新性的技术和模型<sup>[25, 26]</sup>。在卷积神经网络(CNN)方面,传统CNN的改进模型具有强大的上下文信息提取能力,能够接受任意大小的图像输入进行端到端的训练,这些方法将变化检测视为密集像素分类任务,Peng 等人<sup>[27]</sup>使用了一个具有编码器和解码器结构的 U-Net 网络来融合多尺度空洞卷积和深度监督策略,实现了对高光谱图像的变化检测,这种方法有效提高了检测的精确性。Liu 等人<sup>[28]</sup>着眼于提高变化检测的准确性,通过对卷积神经网络加入空间约束,使卷积仅在局部区域进行,提高了模型对局部信息的挖掘和利用。Wang 等人<sup>[29]</sup>设计了一种可变形残差卷积神经网络,根据图像的实际结构来动态调整卷积核形状,这种设计能够有效提取并利用图像中的特征信息。Ou 等人<sup>[30]</sup>利用差分矩阵来丰富变化信息的层次,通过选择慢特征带和快特征带实现了对变化特征的精准捕获,这种方法在处理复杂变化场景时表现尤为突出。

在多尺度特征提取和注意力机制方面,Tang 等人<sup>[31]</sup>设计了一种基于多尺度三维卷积的自动编码器,展示了其在变化检测领域的独特优势,这种方法通过多尺度特征提取,能够捕捉更多细节信息,从而提高变化检测性能。Gong 等人<sup>[32]</sup>使用原始的多时相高光谱图像,设计了一个端到端的网络来提高变化检测的准确性,简化了特征提取和变化检测的流程,使整个过程更加高效和精准。Song 等人<sup>[33]</sup>通过增强模型对变化信息的特征表征,并利用空间注意力和通道注意力实现了对图像特征的高效利用,解决了尺度变化和伪变化对模型性能的影响。Bidari 等人<sup>[34]</sup>在解码阶段使用注意力融合策略,增加了模型对变化像素的敏感性,通过重点关注变化区域,提高了检测精度。Wang 等人<sup>[35]</sup>提出了一种特征融合网络,增强了模型捕捉多时相高光谱图像中复杂特征的能力。Song 等人<sup>[36]</sup>引入了一种创新的跨时间交互对称注意机制,增强了模型对时序变化的敏感性,使得模型能够有效捕捉不同时刻之间的变化特征。Ding 等人<sup>[37]</sup>首次将 Transformer 技术引入高光谱图像变化检测,利用注意力机制实现了对全局信息的有效整合,提高了检测性能。Transformer 模型

利用注意力计算,使得模型能够捕捉到图像中更广泛的变化特征<sup>[38]</sup>,进一步提升了变化检测的效果。除了整合全局信息,Transformer模型还能有效处理长距离依赖关系<sup>[39]</sup>,使模型能够很好地利用远距离像素之间的关联性。这些方法不仅有效增强了模型对复杂变化的识别能力,还提高了变化检测的鲁棒性和泛化能力,为高光谱图像变化检测提供了一种全新的技术路径。

### 2.3 对比方法介绍

与上述基于深度学习的网络不同,本文创新性地提出了一种基于孪生网络结构的通道变换和 Transformer 模型。该方法将所提出的 CSB 模块和 DGFFN 模块融入孪生网络结构之中,借助孪生网络有效地提取并处理双时相高光谱图像中的特征信息。在实验部分,我们对现有变化检测领域中具有代表性的一些方法进行了对比,具体的对比方法如下:

(1)CVA<sup>[6]</sup>通过对高光谱图像进行预处理和配准,然后构建变化向量来描述同一位置的光谱差异,通过计算光谱变化向量的模来确定变化区域,这种方法比较简单,能够捕捉图像之间的显著变化。

(2)IR-MAD<sup>[40]</sup>采用迭代重新加权,利用统计方法进行多次迭代来增强变化信息,减小噪声和无关信息的干扰,以适应图像中的复杂变化。通过不断优化权重,逐步提高变化检测的精度,尤其在处理高噪声数据时表现出色。

(3)ReCNN<sup>[41]</sup>提出了一种循环卷积神经网络来捕捉高光谱图像数据随时间变化的动态信息,同时考虑了图像中的空间关系和光谱特性。这种方法既能捕捉时间序列的动态特征,又能提取空间和光谱信息,为变化检测提供了一个全面的分析框架。

(4)CDFormer<sup>[35]</sup>首次将 Transformer 技术引入到高光谱图像变化检测任务中,CDFormer 通过建模全局自注意力,捕捉图像间的复杂变化信息。这种方法通过利用注意力计算来关注全局上下文信息,能够有效整合全局特征,从而提升变化检测的性能。

(5)CSANet<sup>[34]</sup>提出了一种新的跨时间交互对称注意算法,通过这种方法,模型能够更好地捕捉和利用时间序列中的细微变化,增强模型对每个时相图像的特征表示能力。CSANet 的设计使得模型在捕捉时序变化和空间特征方面具有更高的灵活性和准确性。

(6)DA-Former<sup>[42]</sup>提出了一种用于高光谱图像

变化检测的半监督模型,它能够利用大量无标签数据通过双分支自动编码器实现域对齐,并使用少量标签数据训练分类器来分离不同特征,能够有效对齐不同时态高光谱图像的特征域,从而提高变化检测的精度。

## 3 本文方法

### 3.1 网络整体架构

本文提出的 CSTFormer 模型的总体架构如图 1 所示,该框架主要由通道变换模块和 Transformer 模块组成。本文使用孪生网络结构来处理双时相的高光谱图像,这种结构能够有效学习每个时相图像之间的信息差异,从而帮助模型捕捉不同时相间的变化。在这个网络中,首先,使用主成分分析(PCA)方法对特征在通道维度进行降维,保留高光谱图像中的重要信息,同时减少数据维度和计算复杂度。接下来,将降维处理后的信息输送到通道变换模块(CSB)进行通道信息交互。然后,将处理后的特征信息传递到 Transformer 模块,利用多头自注意力机制对特征进行建模,利用多个注意力头从不同角度关注特征之间的相互关系增强模型捕捉全局上下文信息的能力。最后,利用双分支门控前馈神经网络(DGFFN)进一步增强模型对复杂特征的表达能力。将处理后的特征输送到下一层进行迭代处理,每次迭代时,逐步提取并融合双时相高光谱图像的特征,生成更为精细的差异特征图。通过这种逐层递进的方式,逐步提升对变化信息的捕捉精度。最后,使用 Softmax 函数进行处理得到最终的二分类变化图。

### 3.2 通道变换模块

为了加强特征通道之间的信息交互,本文提出了一种通道信息交互模块。该模块能够有效地促进输入特征在不同通道间的信息流动,从而提高模型对通道间潜在关联性的捕捉与利用。模块主要由两个部分组成:通道映射模块和深度卷积模块。通道映射模块的结构如图 2 所示,在这个模块中,首先对输入特征  $X \in \mathbb{R}^{H \times W \times C}$  进行归一化处理,然后将其通道分为两部分:一部分保持不变,另一部分通过逐点的多层感知机(Multilayer Perception, MLP)进行通道混合。随后,将经过 MLP 处理的特征与未处理的部分进行拼接和重组,生成新的特征图。最后,新的特征图与原始输入特征进行残差连接,得到最终输

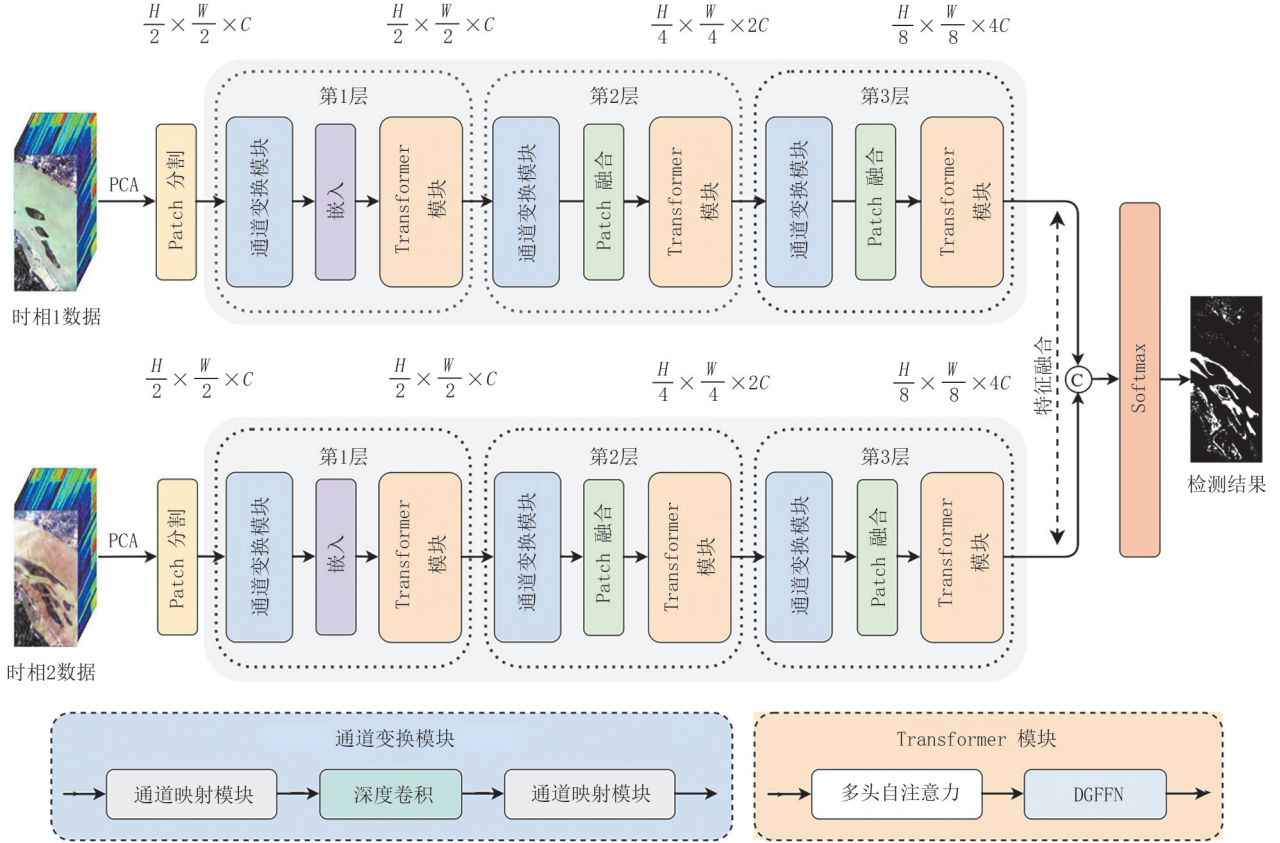


图1 基于通道变换与Transformer的高光谱图像变化检测网络结构

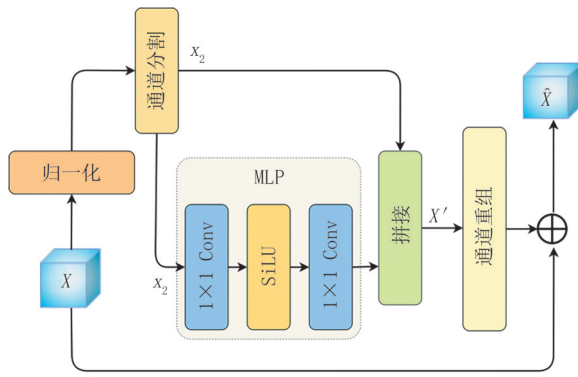


图2 通道映射模块结构

出 $\hat{X}$ 。在逐点MLP模块中,使用SiLU作为非线性激活函数,并在其前后各添加一个 $1 \times 1$ 卷积,以捕捉输入数据中的复杂关系。第一个 $1 \times 1$ 卷积将特征维度从 $C \times H \times W$ 调整为 $2C \times H \times W$ ,第二个 $1 \times 1$ 卷积将特征维度恢复为 $C \times H \times W$ ,利用这种结构能够有效增强模型的特征表达能力。通过结合逐点卷积和激活函数,可以有效促进通道间的信息交互,提升模型对全局上下文的理解。具体计算公式如下:

$$\text{LayerNorm}(X) \xrightarrow{\text{Split}} X_1 + X_2 \quad (1)$$

$$\hat{X}_2 = W_{1 \times 1}^1(\sigma(W_{1 \times 1}^0(X_2))) \quad (2)$$

$$\hat{X} = \text{Shuffle}([\hat{X}_2, X_1]) + X \quad (3)$$

这种设计通过对输入特征进行打乱重组来增强通道信息间的交互,利用逐点MLP将非线性激活函数和线性变换进行结合来提升通道变换模块的表达能力。通过引入深度卷积来增强模型对特征的学习,使模型能够同时提取高光谱图像中的空间和光谱信息,并将两者有效融合,保证模型能够有效学习高光谱图像中的边缘、纹理和形状等低级特征以及更复杂的高级特征。CSB模块在提升模型计算效率的同时,还增强了网络对高光谱图像中高级语义信息的理解以及复杂变化的感知能力。

### 3.3 优化注意力计算

传统的视觉Transformer(Vision Transformer, ViT)的注意力计算主要由 $Q$ 、 $K$ 、 $V$ 产生,其计算如公式(4),设 $Q$ 与 $K$ 的维度为 $N \times d$ ,则 $QK^T$ 的计算复杂度为 $O(N^2 \times d)$ 。而余下计算的复杂度均不超过 $O(N^2 \times d)$ ,因此,ViT自注意力机制的计算复杂度为 $O(N^2 \times d)$ 。



$$Q = XW_Q, K = XW_K, V = XW_V \quad (4)$$

Transformer模块中使用的多头注意力通过并行处理多个注意力头,每个头专注于不同的特征子空间,为模型提供了更丰富和多样化的上下文信息,使模型可以从不同的角度和抽象层次捕捉信息,增强模型对全局上下文的理解。在输入矩阵 $X$ 中随机采样 $r$ 个样本形成 $X_r$ ,然后使用 $X_r$ 计算出 $Q_r$ 和 $K_r$ ,特征维度就从原来的 $N \times d$ 降到 $r \times d$ 。然后,将计算全维度的 $Q$ 和 $K$ 与采样的 $Q_r$ 和 $K_r$ 分别相乘,得到新的 $\hat{Q} \in \mathbb{R}^{N \times r}$ 和 $\hat{K} \in \mathbb{R}^{N \times r}$ 。最后,利用生成的 $\hat{Q}$ 和 $\hat{K}$ 计算出 $N \times N$ 的相似性矩阵,该过程的计算复杂度为 $O(N^2 \times r)$ ,这种方法显著地降低了计算的复杂度,并且其中每个元素都包含了输入序列的全局上下文信息。改进的计算如公式(5),具体的计算如公式(6)和(7)。相似性矩阵的具体计算如图3所示,其中 $A$ 为相似性矩阵。实验中 $r$ 取值为10, $d$ 取值为256。

$$\text{GlobalAttention}(\hat{Q}, \hat{K}, V) = \text{Softmax}\left(\frac{\hat{Q}\hat{K}^T}{\sqrt{d'}}\right)V \quad (5)$$

$$Q_r = X_r W_Q, K_r = X_r W_K \quad (6)$$

$$\hat{Q} = QQ_r^T, \hat{K} = KK_r^T \quad (7)$$

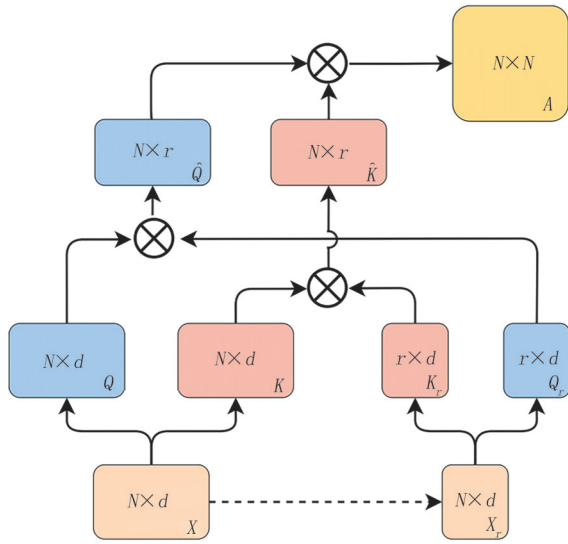
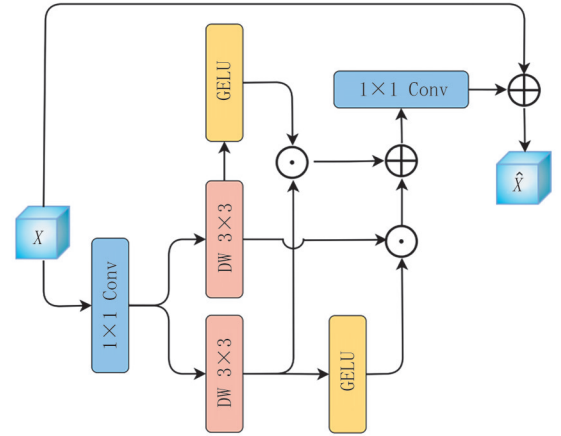


图3 相似性矩阵计算方式

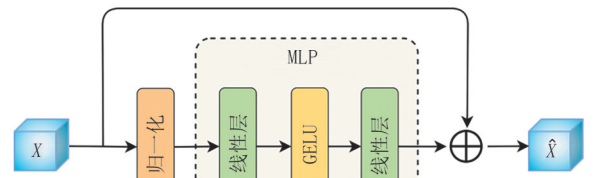
### 3.4 双分支门控前馈神经网络

考虑到高光谱图像对变化检测模型的表征能力的要求,本文设计了一种创新的双分支门控前馈神经网络,取代了传统ViT模型中的前馈神经网络。DGFFN的结构如图4(a)所示,它利用双GELU激活函数和元素乘积在两条并行路径上融

合有效信息,通过 $1 \times 1$ 卷积和 $3 \times 3$ 的深度卷积来提取局部信息。相比于图4(b)中展示的ViT中使用的前馈神经网络,本文设计的DGFFN的优势在于:



(a) 双分支门控前馈神经网络



(b) ViT中的前馈神经网络

图4 门控神经网络

(1)提升模型表征能力: 高光谱图像中的噪声会影响变化检测的效率,DGFFN采用双分支的设计,能够更有效地捕捉和整合不同层次的特征信息,而传统的FFN通常为单一路径,在信息融合方面相对有限。其次,双GELU激活函数的运用使得DGFFN能够更好地捕提高光谱图像中的非线性特征,相比传统FFN单一的激活函数,具有更强的表达能力。

(2)增强局部信息: DGFFN在每条路径上巧妙地运用了 $1 \times 1$ 卷积和 $3 \times 3$ 深度卷积,这一设计极大地丰富了局部信息,使模型对高光谱图像细节的感知更为敏感。相比之下,传统FFN在局部信息提取方面存在一定不足。此外,DGFFN在设计时充分考虑了抗噪声信息的干扰,门控机制可以有效地筛选和过滤噪声,提高特征的纯净度,使模型能够在复杂的高光谱图像环境中稳定地进行学习和信息整合。而传统的FFN在面对噪声时,往往难以有效

处理噪声干扰。

模型的输入为  $X \in \mathbb{R}^{H \times W \times C}$ , 使用 DGFFN 的完整计算如下所示:

$$\hat{X} = W_{1 \times 1}(\text{DGFFN}(X)) + X \quad (8)$$

$$\text{DGFFN}(X) = \phi_1(X) + \phi_2(X) \quad (9)$$

$$\phi_1(X) = \sigma(W_{3 \times 3}^1 W_{1 \times 1}^1(X)) \odot (W_{3 \times 3}^2 W_{1 \times 1}^2(X)) \quad (10)$$

$$\phi_2(X) = (W_{3 \times 3}^1 W_{1 \times 1}^1(X)) \odot \sigma(W_{3 \times 3}^2 W_{1 \times 1}^2(X)) \quad (11)$$

其中,  $\hat{X} \in \mathbb{R}^{H \times W \times C}$  表示经过 DGFFN 处理后的输出特征,  $W_{1 \times 1}$ 、 $W_{1 \times 1}^1$  和  $W_{1 \times 1}^2$  为  $1 \times 1$  的卷积,  $W_{3 \times 3}^1$  和  $W_{3 \times 3}^2$  代表  $3 \times 3$  的深度卷积, DGFFN 表示双分支门控机制,  $\phi_i$  为第  $i$  个分支上的输出,  $\odot$  表示逐元素乘法,  $\sigma$  表示 GELU 激活函数。

### 3.5 变换检测性能评估指标

在变化检测任务中, 单纯的可视化和定性描述容易受到主观因素的影响, 因此需要使用客观指标来量化模型的性能, 衡量算法的有效性和可靠性。主要的客观指标如下:

(1) 总误差像素 (Overall Errors, OE): OE 表示检测的总体错误数量, 即所有错误分类的像素的数量。其计算公式如下所示:

$$\text{OE} = \text{FN} + \text{FP} \quad (12)$$

其中, FN 表示被模型错误检测为未发生变化的像素的数量, FP 表示被模型错误检测为发生变化的像素的数量。

(2) 准确率 (Percentage Correct Classification, PCC): PCC 表示被正确检测的像素占总像素数的比例。PCC 的计算如式(12)所示:

$$\text{PCC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (13)$$

其中, TP 表示被正确检测为变化的像素数量, TN 表示被正确检测为不变的像素数量。

(3) Kappa 系数 (Kappa Coefficient, KC): 这个系数主要用于检验结果图与真值图的一致性, 具体计算公式为:

$$\text{KC} = \frac{\text{PCC} - \text{PRE}}{1 - \text{PRE}} \quad (14)$$

其中 PRE 的计算公式为

$$\text{PRE} = \frac{(\text{TP} + \text{FP})(\text{TP} + \text{FN}) + (\text{FN} + \text{TN})(\text{FP} + \text{TN})}{(\text{TP} + \text{FP} + \text{TN} + \text{FN})^2} \quad (15)$$

## 4 实验结果与分析

### 4.1 数据集

本文分别在 River<sup>[43]</sup> 和 Hermiston<sup>[44]</sup> 两个公开的高光谱图像数据集上进行实验, 数据集信息如表1所示。

表1 数据集信息

数据集	River	Hermiston
大小	463 × 241	307 × 241
波段数	198	154
$T_1$	May 3rd, 2013	May 1st, 2004
$T_2$	Dec 31st, 2013	May 8th, 2007

(1) River 数据集记录了江苏省内河流的变化情况, 采集时间分别为 2013 年 5 月 3 日和 2013 年 12 月 31 日, 由 Earth Observing-1 (EO-1) Hyperion 收集。该数据集的光谱范围为 0.4 至 2.5  $\mu\text{m}$ , 空间分辨率为 30 m, 光谱分辨率达到 10 nm。图像尺寸为 463 × 241 像素, 预处理后剩余 198 个可用波段。

(2) Hermiston 数据集拍摄的是美国俄勒冈州 Hermiston 市 Umatilla 县的灌溉农田, 两个时相的图像分别于 2004 年 5 月 1 日和 2007 年 5 月 8 日由 EO-1 Hyperion 传感器获取。图像大小为 307 × 241 像素, 去除噪声波段后用于变化检测任务的图像包含 154 个光谱波段。

### 4.2 实验参数分析

(1) 训练样本比例: 在高光谱图像变化检测任务中, 进行训练时所使用的训练样本占总样本的比例是一个重要的参数, 本文分别随机选取 5%、10%、15%、20%、25% 和 30% 的样本作为训练集, 对应的 Kappa 系数变化如图 5 所示。结果表明, 当训练样本为 20% 时, 模型的性能在两个数据集上均达到

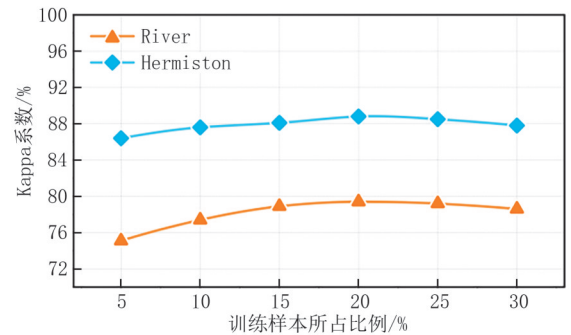
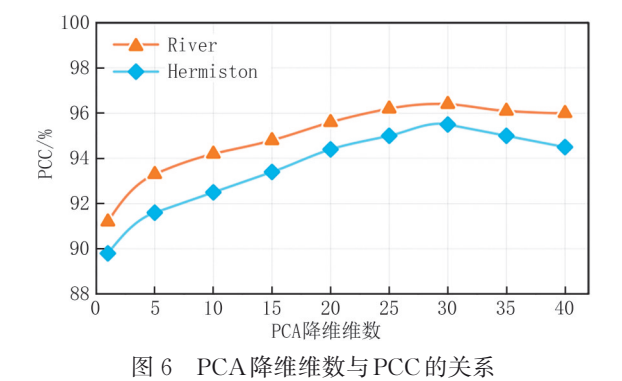


图5 训练样本所占比例及其训练效果



最佳。当训练样本为5%时,由于特征维度与模型复杂度较高,模型出现了较为明显的欠拟合现象;当训练样本超过20%时,模型出现了轻微的过拟合现象。因此,本文实验将训练样本设定为20%。

(2)PCA降维维数:PCA降维是一种能够有效降低模型训练时计算成本的方法,使用这种方法需要选择适当的维数以确保模型性能。在本节实验中,将PCA降维的维数分别设置为[1,5,10,15,20,25,30,35,40],实验结果如图6所示。结果显示,在River和Hermiston数据集上,降维维数为30时均达到了最佳效果。根据这一实验结果,本文实验将两个数据集的降维维数均设定为30。



### 4.3 对比实验分析

为了全面评估CSTFormer在高光谱图像变化检测任务上的性能,本节在两个不同数据集上对CSTFormer进行了深入分析,同时将其性能与CVA<sup>[6]</sup>、IR-MAD<sup>[42]</sup>、ReCNN<sup>[43]</sup>、CDFormer<sup>[35]</sup>、CSANet<sup>[34]</sup>以及DA-Former<sup>[44]</sup>六个经典的高光谱图像变化检测方法进行了对比。

本文在Python 3.8.1和PyTorch 1.10.0环境下编写并运行代码,在Google Colaboratory云端开展实验,利用Tesla T4 16 GB显存GPU进行模型训练和测试。为确保实验的可靠性,采用完全一致的实验环境、训练样本集和测试样本集。其中,训练样本占总样本的20%。将学习率设定为1e-4,并使用Adam优化器对模型进行优化,模型训练历经200个epoch,分别在River和Hermiston两个数据集上进行实验验证。

表2呈现了不同方法在River数据集上进行训练时模型的参数量与计算量,对比表中数据发现,CSTFormer的参数量在几种方法中优势显著。此外,我们还在图7中直观地展示了不同方法所取得的准确率。与CDFormer、CSANet以及DA-Former

这三种采用注意力机制的模型相比,CSTFormer的计算量仅为1.7 GFlops,在几种对比方法中处于最低水平,这再次验证了优化后的注意力在降低模型计算量方面的优势。同时,表2中还展示了不同方法在训练时训练一个epoch所用的时间,通过几种深度学习方法的对比,CSTFormer训练一个epoch所需的时间最短,这得益于CSTFormer中所使用的优化后的注意力计算方法。

表2 在River数据集上的各个模型的参数量和计算量			
方法	Params(K)	GFLOPs	Time(s)
CVA <sup>[6]</sup>	-	-	<b>2.35</b>
IR-MAD <sup>[42]</sup>	-	-	3.74
ReCNN <sup>[43]</sup>	<b>291.2</b>	<u>5.6</u>	3.35
CDFormer <sup>[35]</sup>	402.6	13.7	3.87
CSANet <sup>[34]</sup>	387.8	9.1	3.51
DA-Former <sup>[44]</sup>	415.3	15.3	4.05
CSTFormer	<u>333.8</u>	<b>1.7</b>	<u>3.07</u>

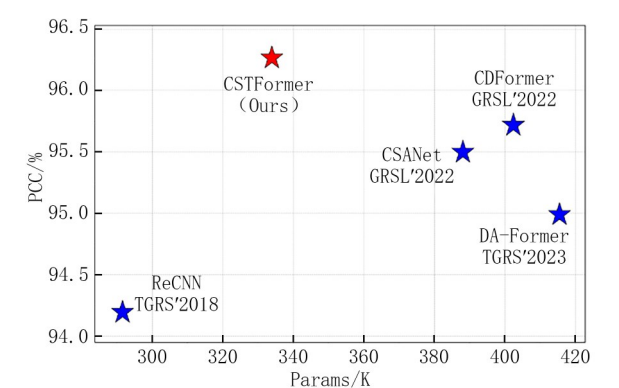


图7 不同模型训练时的参数量及其对应的准确率

我们在表3中详细地展示了几种方法在River和Hermiston两个数据集上的实验结果。在River数据集中,其发生变化的区域较为分散,这种特性使得该数据集能够很好地检验各种模型对细微变化的识别能力。通过对比图8发现,CSTFormer的可视化结果最接近于真值图,尤其是在右下角蓝色区域发生的细微变化区域,使用CSTFormer训练得到的结果图表现得最为理想。在此区域中,PCC和KC分别达到了96.28%和79.44%。与其他六种方法中表现最为出色的CDFormer相比,CSTFormer在PCC和KC上分别提升了0.60%和10.30%,这一显著的提升充分彰显了CSTFormer在处理细微变化方面的强大优势。

相较于River数据集,Hermiston数据集收集的是不同年份农业灌溉区域发生的变化,使用不同数

表3 River和Hermiston数据集对比实验结果			
方法	评价指标	数据集	
		River	Hermiston
CVA <sup>[6]</sup>	OE	8763	7257
	PCC(%)	92.15	90.18
	KC(%)	56.12	68.87
IR-MAD <sup>[42]</sup>	OE	7182	6108
	PCC(%)	93.56	91.74
	KC(%)	58.34	76.60
ReCNN <sup>[43]</sup>	OE	6476	4971
	PCC(%)	94.20	93.28
	KC(%)	58.02	80.32
CDFormer <sup>[45]</sup>	OE	<u>4787</u>	3466
	PCC(%)	<u>95.71</u>	95.31
	KC(%)	<u>72.02</u>	86.88
CSANet <sup>[34]</sup>	OE	5016	3979
	PCC(%)	95.50	94.31
	KC(%)	68.23	84.90
DA-Former <sup>[44]</sup>	OE	4893	<u>3417</u>
	PCC(%)	94.99	<u>95.53</u>
	KC(%)	72.08	<u>87.01</u>
CSTFormer	OE	<b>4150</b>	<b>2978</b>
	PCC(%)	<b>96.28</b>	<b>95.97</b>
	KC(%)	<b>79.44</b>	<b>88.90</b>

据集来验证模型在不同场景下的检测能力,能够检验模型对不同地物变化模式的泛化能力。CSTFormer在Hermiston数据集上同样也取得了最好的检测结果,尤其是图9中框选红色和蓝色区域,使用CSTFormer训练得到的结果与真值图的相似度极高。这主要是因为CSB模块有效加强了特征通道之间的信息交互,DGFFN模块剔除了噪声信息的干扰,这两个关键模块的协同作用,使得模型能够高度聚焦于对高光谱图像中高级语义的理解以及复杂变化的精准感知。相比于DA-Former,

CSTFormer在PCC和KC上分别提升了0.46%和2.17%,进一步证明了CSTFormer在不同数据集上的出色表现。

CSTFormer在两个数据集上均取得了最好的结果,其中,较小的OE值表明模型对图像中是否发生变化的识别极为准确,这不仅证明了DGFFN模块在过滤潜在噪声干扰方面的作用,还充分显示出CSTFormer能够有效控制错误分类的像素数量。由此可见,CSTFormer在变化检测任务中具有更显著的优势,利用CSTFormer提取高光谱图像中的空间和通道信息,能够更准确地识别变化区域。

4.4 消融实验

为了验证本文所提出的CSB和DGFFN模块的有效性,分别在River和Hermiston两个数据集上进行了消融实验。针对每个数据集分别设计了三种不同的消融网络,包括骨干网络、去除CSB模块的网络以及去除DGFFN模块的网络,利用消融实验来测试这两个模块对整体网络结构的影响,实验结果如表4所示。从表中可以看出,使用骨干网络进行变化检测在River和Hermiston两个数据集上得到的PCC分别为93.87%和93.15%,加入CSB模块后提升效果分别为1.93%和2.18%;加入FFN模块后提升效果分别为0.62%和0.65%,相比于使用普通的前馈神经网络,使用我们所设计的双分支门控前馈神经网络(DGFFN)在River和Hermiston两个数据集上PCC分别达到了95.73%和95.01%,相比于骨干网络,加入DGFFN模块后提升效果分别为1.98%和2.00%;同时加入CSB和DGFFN两个模块后PCC分别在River和Hermiston数据集上分别达到了96.28%和95.97%,相比于骨干网络,同时加入两个模块后效果分别提升了2.57%和3.03%。

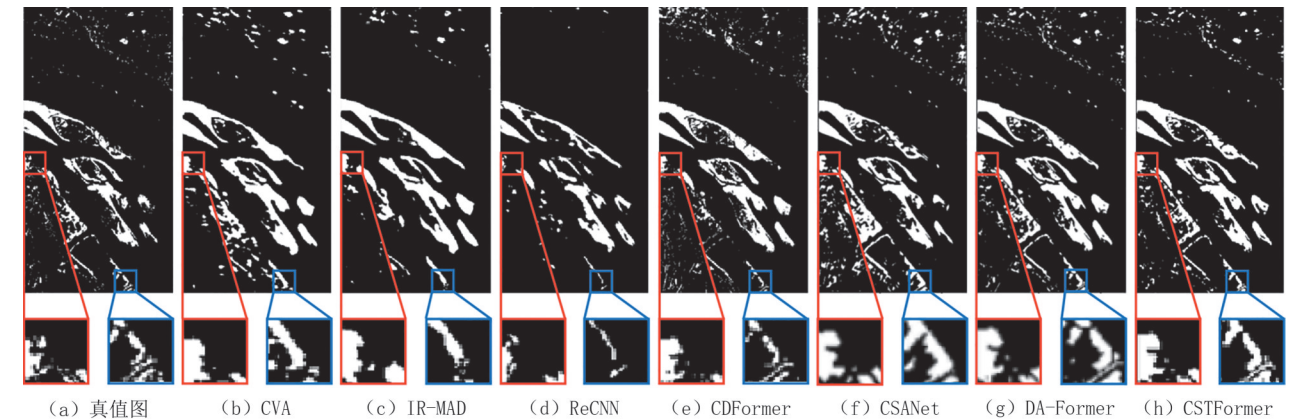


图8 River数据集上不同方法的变化检测结果

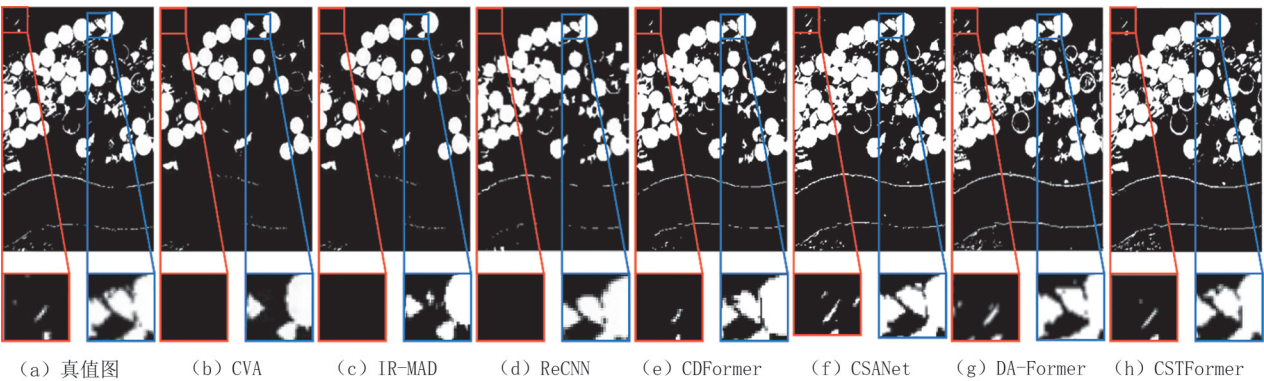


图9 Hermiston数据集上不同方法的变化检测结果

表4 基于通道变换和Transformer网络的消融实验

网络结构	PCC(%)	
	River	Hermiston
骨干网络	93.87	93.15
+CSB	95.68	95.18
+FFN	94.45	93.76
+DGFFN	95.73	95.01
+CSB+DGFFN	<b>96.28</b>	<b>95.97</b>

我们还通过消融实验进一步验证了使用优化的注意力前后所取得的效果,同时对比了模型的参数量、计算量以及处理时效,详细的对比情况如表5所示。使用优化后的注意力计算,模型参数量从382.6 K降至333.8 K,计算量也从10.2 GFLOPs大幅降低至1.7 GFLOPs,训练一个epoch所需时间也从3.54秒缩短至3.07秒。这充分表明优化后的注意力机制在降低模型参数量和计算负担、提升处理时效方面效果显著,为模型的性能提升和实际应用提供了有力支撑。

表5 优化注意力计算的消融实验

注意力	Params(K)	GFLOPs	Time(s)
未优化	382.6	10.2	3.54
优化	333.8	1.7	3.07

为了更直观地观察CSTFormer对特征图的影响,本文使用了t-SNE<sup>[45]</sup>可视化技术对Hermiston数据集输入和输出的特征图进行了处理,如图10所示。其中,绿色点表示发生了变化的像素点,红色点表示未发生变化的像素点。通过对比可视化图不难发现,经过模型训练后的特征点的区分度更加明显,进一步验证了CSB模块和DGFFN模块在变化检测任务中的有效性。

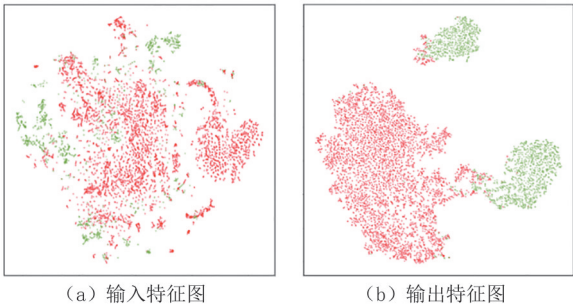


图10 t-SNE特征可视化

5 结 论

本文基于传统的Transformer模型,设计了通道变换模块来增强特征通道间信息的交互,利用逐点MLP将非线性激活函数和线性变换进行结合来提升通道变换模块的表达能力,通过引入深度卷积来增强模型对特征的学习,使模型能够同时提取高光谱图像中的空间和光谱信息,并将两者有效融合。同时,本文还利用多头注意力来增强模型对全局上下文信息的理解,通过优化传统Transformer模型中Q、K、V的计算方式,将模型的计算复杂度从 $O(N^2 \times d)$ 降低到 $O(N^2 \times r)$ 。此外,本文还设计了一个双分支门控前馈神经网络,并行使用了两个GELU激活函数和元素乘积来过滤潜在噪声,利用不同的卷积操作来捕捉局部信息,使模型能够更好地学习关键地物的变化和细微差异。通过将通道变换模块和Transformer串行应用于高光谱图像变化检测任务,使用注意力机制模拟视觉任务中的长距离依赖关系,从而实现了全局上下文的充分建模。CSTFormer不仅在关注高光谱图像空间信息的同时增强了模型对波段信息的依赖,还提升了模型在通道维度上的交互能力,从而有效提高了变化



检测任务的准确性。

在 River 数据集和 Hermiston 数据集上分别使用 PCC 和 KC 评价指标对实验结果进行分析,并结合变化检测的可视化结果进行了综合评估,实验结果表明本文所提出的 CSTFormer 模型在变化检测任务中的具有很好的性能。值得注意的是,本文仅实现了高光谱图像这一种模态数据的变化检测,未来的研究可以进一步探索将所提出的模型用于其他遥感图像(如 SAR 图像、多光谱图像、热红外图像等)的变化检测,从而拓展 CSTFormer 模型的应用领域。

### 参 考 文 献

- [1] Chen H, Shi Z. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection. *Remote Sensing*, 2020, 12(10): 1662
- [2] De Bem P P, de Carvalho Junior O A, Fontes Guimarães R, et al. Change detection of deforestation in the Brazilian Amazon using landsat data and convolutional neural networks. *Remote Sensing*, 2020, 12(6): 901
- [3] Xu J Z, Lu W, Li Z, et al. Building damage detection in satellite imagery using convolutional neural networks. *arXiv preprint arXiv:1910.06444*, 2019
- [4] Peng X, Zhong R, Li Z, et al. Optical remote sensing image change detection based on attention mechanism and image difference. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 59(9): 7296-7307
- [5] Wei H, Jinliang H, Lihui W, et al. Remote sensing image change detection based on change vector analysis of PCA component. *Remote Sensing Land Resource*, 2016, 28(1): 22-27
- [6] Chen J, Chen X, Cui X, et al. Change vector analysis in posterior probability space: A new method for land cover change detection. *IEEE Geoscience and Remote Sensing Letters*, 2010, 8(2): 317-321
- [7] Chen H, Wu C, Du B, et al. Change detection in multisource VHR images via deep siamese convolutional multiple-layers recurrent neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 58(4): 2848-2864
- [8] Shi C, Zhang Z, Zhang W, et al. Learning multiscale temporal-spatial-spectral features via a multipath convolutional LSTM neural network for change detection with hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-16
- [9] Bandara W G C, Patel V M. A transformer-based siamese network for change detection//*Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*. Malaysia, Kuala Lumpur, 2022: 207-210
- [10] Lyu H, Lu H, Mou L. Learning a transferable change rule from a recurrent neural network for land cover change detection. *Remote Sensing*, 2016, 8(6): 506
- [11] Khabou N, Rodriguez I B, Gharbi G, et al. A threshold based context change detection in pervasive environments: Application to a smart campus. *Procedia Computer Science*, 2014, 32: 461-468
- [12] Bruzzone L, Prieto D F. Automatic analysis of the difference image for unsupervised change detection. *IEEE Transactions on Geoscience and Remote sensing*, 2000, 38(3): 1171-1182
- [13] Lu D, Mausel P, Brondizio E, et al. Change detection techniques. *International journal of remote sensing*, 2004, 25(12): 2365-2401
- [14] Zhang L, Hu X, Zhang M, et al. Object-level change detection with a dual correlation attention-guided detector. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 177: 147-160
- [15] Dian Y Y, Fang S H, Yao C H. Change detection for high-resolution images using multilevel segment method. *Journal of Remote Sensing*, 2016, 20(1): 129-137
- [16] Lei Y, Liu X, Shi J, et al. Multiscale superpixel segmentation with deep features for change detection. *Ieee Access*, 2019, 7: 36600-36616
- [17] Wang M, Tan K, Jia X, et al. A deep siamese network with hybrid convolutional feature extraction module for change detection based on multi-sensor remote sensing images. *Remote Sensing*, 2020, 12(2): 205
- [18] Zhan T, Gong M, Jiang X, et al. Unsupervised scale-driven change detection with deep spatial-spectral features for VHR images. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(8): 5653-5665
- [19] Liu T, Yang L, Lunga D. Change detection using deep learning approach with object-based image analysis. *Remote Sensing of Environment*, 2021, 256: 112308
- [20] Gong M, Zhan T, Zhang P, et al. Superpixel-based difference representation learning for change detection in multispectral remote sensing images. *IEEE Transactions on Geoscience and Remote sensing*, 2017, 55(5): 2658-2673
- [21] Lu M, Chen J, Tang H, et al. Land cover change detection by integrating object-based data blending model of Landsat and MODIS. *Remote Sensing of Environment*, 2016, 184: 374-386
- [22] Xu L, Jing W, Song H, et al. High-resolution remote sensing image change detection combined with pixel-level and object-level. *IEEE Access*, 2019, 7: 78909-78918
- [23] Lin Y, Li S, Fang L, et al. Multispectral change detection with bilinear convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 2019, 17(10): 1757-1761
- [24] Peng D, Zhang Y, Guan H. End-to-end change detection for high resolution satellite images using improved UNet++ . *Remote Sensing*, 2019, 11(11): 1382
- [25] Liu S, Li H, Chen J, et al. Adaptive convolution kernel network for change detection in hyperspectral images. *Applied Optics*, 2023, 62(8): 2039-2047
- [26] Wang L, Wang L, Wang H, et al. SPCNet: A subpixel convolution-based change detection network for hyperspectral images with different spatial resolutions. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-14
- [27] Peng D, Bruzzone L, Zhang Y, et al. SCDNET: A novel convolutional network for semantic change detection in high

- resolution optical remote sensing imagery. *International Journal of Applied Earth Observation and Geoinformation*, 2021, 103: 102465
- [28] Liu Z G, Zhang Z W, Pan Q, et al. Unsupervised change detection from heterogeneous data based on image translation. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 60: 1-13
- [29] Wang J, Gao F, Dong J. Change detection from SAR images based on deformable residual convolutional neural networks// *Proceedings of the 2nd ACM International Conference on Multimedia in Asia*. Singapore, 2021: 1-7
- [30] Ou X, Liu L, Tu B, et al. A CNN framework with slow-fast band selection and feature fusion grouping for hyperspectral image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-16
- [31] Tang Y, Fan Y, Feng S, et al. Hyperspectral image change detection based on multi-scale 3D convolution autoencoder// *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*. Malaysia, Kuala Lumpur, 2022: 3211-3214
- [32] Gong M, Zhao J, Liu J, et al. Change detection in synthetic aperture radar images based on deep neural networks. *IEEE transactions on neural networks and learning systems*, 2015, 27(1): 125-138
- [33] Song K, Jiang J. AGCDetNet: An attention-guided network for building change detection in high-resolution remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 4816-4831
- [34] Bidari I, Chickerur S, Kadam S. Semantic Segmentation Using U-Net Architecture for Change Detection on Hyperspectral Imagery// *Proceedings of the International Conference on Sustainable Communication Networks and Application (ICSCNA)*. Theni, India, 2023: 932-937
- [35] Wang X, Ni W, Feng Y, et al. AgF 2 Net: Attention-guided Feature Fusion Network for Multi-temporal Hyperspectral Image Change Detection. *IEEE Geoscience and Remote Sensing Letters*, 2023, 20: 1-5
- [36] Song R, Ni W, Cheng W, et al. CSANet: Cross-temporal interaction symmetric attention network for hyperspectral image change detection. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 1-5
- [37] Ding J, Li X, Zhao L. CDFormer: A hyperspectral image change detection method based on transformer encoders. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 1-5
- [38] Zhang W, Zhang Y, Gao S, et al. Spectrum-Induced Transformer based Feature Learning for Multiple Change Detection in Hyperspectral Images. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 62: 1-12
- [39] Wang X, Li S, Zhao X, et al. Inter-Domain Collaboration between Hyperspectral and VHR Remote Sensing Images: A Cross-Scene Few-Shot Learning Framework for Change Detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 1-16
- [40] Nielsen A A. The regularized iteratively reweighted MAD method for change detection in multi-and hyperspectral data. *IEEE Transactions on Image processing*, 2007, 16(2): 463-478
- [41] Mou L, Bruzzone L, Zhu X X. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 57(2): 924-935
- [42] Wang Y, Sha J, Gao L, et al. A semi-supervised domain alignment transformer for hyperspectral images change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1-11
- [43] Wang Q, Yuan Z, Du Q, et al. GETNET: A general end-to-end 2-D CNN framework for hyperspectral image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 57(1): 3-13
- [44] Hasanlou M, Seydi S T. Hyperspectral change detection: An experimental comparative study. *International journal of remote sensing*, 2018, 39(20): 7029-7083
- [45] Van der Maaten L, Hinton G. Visualizing data using t-SNE. *Journal of machine learning research*, 2008, 9(11): 2579-2605



**LIU Wen-Li**, M. S. candidate. His research interests include remote sensing image processing and deep learning.

**GAO Feng**, Ph. D., associate professor. His research interests include remote sensing image processing and deep

learning.

**ZHANG Hao-Peng**, M. S. candidate. His research interests include remote sensing image processing and deep learning.

**DONG Jun-Yu**, Ph. D., professor. His research interests include remote sensing image processing and deep learning

**WU Chun-Tong**, M. S. candidate. His research interests include remote sensing image processing and deep learning.

## Background

In the field of remote sensing, hyperspectral images utilizing multi-temporal phases have significant application potential in the field of change detection. Hyperspectral images acquired using hyperspectral image sensors are rich in spectral information, and each pixel of hyperspectral images contains the full spectrum of reflectance properties from visible light to short-wave infrared light, which can provide a strong basis for distinguishing different features.

Traditional change detection research has focused on developing and improving methods based on remote sensing imagery, computer vision, and machine learning to identify and analyze geographic and environmental changes. Traditional methods such as IDM, PCA, and CVA provide simple and effective means of detection, and these methods are suitable for simple change detection tasks, where changes are detected by performing simple mathematical operations or statistical analysis on images, but using these methods generically ignores the contextual information, which is prone to cause a large amount of noise interference and may result in false alarms for complex scenarios. The application of deep learning techniques significantly improves the accuracy and efficiency of change detection. Models such as CNNs, RNNs, LSTMs, and Transformer can achieve more efficient and accurate change detection by learning high-level features and complex patterns in the data.

The current Transformer-based hyperspectral image change detection methods are able to effectively model global contextual information by simulating long-range dependencies through the self-attention mechanism. However, the existing methods still face two main challenges: one is the high computational complexity of the Transformer models, leading to the inefficiency

of the model in dealing with high-dimensional data; the other is the limited utilization of the spectral information of hyperspectral images, resulting in the lack of feature interaction in the spectral dimension. To address these issues, this paper proposes a change detection method for hyperspectral images based on channel shuffle and Transformer, aiming to utilize the complex spectral and spatial information in hyperspectral images more effectively. A feature extraction module based on channel shuffle and attention mechanism is used to improve the traditional self-attention computation, and this improved computation successfully reduces the computational complexity of the quadratic in the traditional Transformer model in terms of computational efficiency, which makes the model more suitable for dealing with high-dimensional data; the use of the channel shuffle module serially combines the global attention modeling when performing the channel information interaction, which realizes the efficient utilization of spatial and spectral information of hyperspectral images, and also enhances the model's understanding of high-level semantic information in hyperspectral images and its ability to perceive complex changes.

In this paper, a dual-branch gated feedforward neural network is also designed, which employs parallel processing with two GELU activation functions and element product operation at the same time. This design helps to filter potential noise more efficiently and capture local information by combining different convolutional operations, thus enabling fine-grained modulation of feature information. Through the integration of channel shuffle, attention mechanisms, and dual-branch gated feedforward neural networks into a twin network structure, the proposed method achieves significant improvements in change detection accuracy and sensitivity.