文章编号:1674-2974(2025)08-0033-11 DOI:10.16339/j.cnki.hdxbzkb.2025281

时序无关和鲁棒性增强的遥感影像变化检测方法

杨景玉[1,2†],张文驰[1],党建武[1,2],王锋[3],火久元[1]
(1. 兰州交通大学 电子与信息工程学院,甘肃 兰州 730070;
2. 甘肃省人工智能与图形图像处理工程研究中心,甘肃 兰州 730070;
3. 甘肃路桥飞宇交通设施有限公司,甘肃 兰州 730070)

摘 要:在遥感影像变化检测中,基于深度学习的方法大多采用孪生网络结构.然而,大量实验发现,此类方法会出现改变输入图像的顺序后性能严重下降的现象,其中ChangeFormer方法在LEVIR-CD数据集的交并比指标下降了79.86%,表明模型的时序鲁棒性不足,严重影响变化检测模型的实用性.对此,提出了一种结合时序对齐与跨层特征混合的变化检测方法CINet(chronologic invariant network),在特征提取时设计时序对齐模块,通过对特征图进行空间混合和时序重建,在特征层面减少双分支的时序差异.然后设计了跨层特征混合模块,使用全尺度连接和差异引导来充分利用双分支中每一层级的特征图,提高在不同时序下的检测能力.最后,在LEVIR-CD数据集的实验结果显示,CINet的召回率和交并比分别达到了90.63%、84.13%,相较于ChangeFormer分别提高了1.83个百分点、1.65个百分点.在多个数据集上的实验结果也表明,即使在改变输入顺序后,所提方法仍能取得良好的变化检测结果,显示出优于其他方法的检测性能和更强的时序鲁棒性.

关键词:遥感影像;变化检测;孪生网络;时序对齐
中图分类号:TP391 **文献标志码**:A

A Chronological Invariant and Robust Enhancement Method for Change Detection in Remote Sensing Image

YANG Jingyu[1,2†],ZHANG Wenchi[1],DANG Jianwu[1,2],WANG Feng[3],HUO Jiuyuan[1]
(1. School of Electronics and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;
2. Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphics & Image Processing, Lanzhou 730070, China;
3. Gansu Luqiao Feiyu Transportation Facilities Co., Ltd., Lanzhou 730070, China)

Abstract: Deep learning-based methods, particularly those employing Siamese network structure, are widely used in remote sensing image change detection. However, extensive experiments reveal that these methods often suffer significant performance degradation when the order of input images is altered. Notably, the ChangeFormer

* **收稿日期**:2024-09-23
基金项目:国家自然科学基金资助项目(62067006,62367005,62262038), National Natural Science Foundation of China(62067006,62367005,62262038);甘肃省自然科学基金资助项目(24JRRA253), Natural Science Foundation of Gansu Province (24JRRA253);甘肃省教育科技创新项目(2021JYJBGS-05), Gansu Province Educational Science and Technology Innovation Project (2021JYJBGS-05)
作者简介:杨景玉(1979—),男,甘肃白银人,兰州交通大学副教授,博士
† 通信联系人,E-mail:yangjy@mail.lzjtu.cn

method exhibits a 79.86% drop in the intersection over the union (IoU) metric on the LEVIR-CD dataset, indicating a lack of chronological robustness that severely impacts the practical utility of the change detection model. To address this issue, this article proposes a novel change detection method called chronologic invariant network (CINet), which integrates chronological alignment module and cross-layer feature mixing module. During the feature extraction phase, a chronological alignment module is introduced. This module employs spatially cross-mix feature maps and chronologic reconstruction to reduce temporal information discrepancies between the two branches at the feature level. Additionally, a cross-layer feature mixing module is designed to blend deep and shallow features using a full-scale connection and difference-guided approach, effectively utilizing feature information from every layer of both branches to improve change detection accuracy under varying input sequences. Experimental results on the LEVIR-CD dataset show that CINet achieves a recall of 90.63% and an IoU of 84.13%, which represent improvements of 1.83 percentage points and 1.65 percentage points over ChangeFormer, respectively. Results from multiple datasets further demonstrate that the proposed method consistently maintains high change detection performance and robust chronological invariance, even when the order of input images is altered, outperforming other methods.

Key words: remote sensing image; change detection; Siamese network; chronological alignment

遥感影像变化检测通过分析同一地理区域不同时间点的地理空间数据,识别该区域的变化范围[1],是遥感技术中的一项重要课题.作为监测地表变化的重要手段,遥感影像变化检测在环保监测[2]、灾害预警[3]、地图更新[4]等多个应用场景中发挥着关键作用.

早期的遥感影像变化检测使用差分法、分类后比较法、回归分析法等方法进行检测[5].然而,随着遥感数据量的增加和图像分辨率的提高,这些传统方法愈加难以应对如此复杂的数据和任务.相比之下,深度学习凭借其端到端的结构优势,不仅大幅降低了变化检测任务研究的复杂性,而且显著提高了检测方法的准确性[6],为复杂场景下的遥感影像变化检测带来了重大突破.

近年来,基于深度学习的变化检测模型通常采用双分支网络架构.例如,Caye等人[7]提出的全卷积孪生网络模型FC-Siam-Conc和FC-Siam-Diff.前者由于将特征图拼接,计算量较大;后者用特征图差值,虽降低了计算量,但损失了部分特征信息.Zhang等人[8]提出双流网络模型IFN,将特征图输入判别网络来提取差异特征,但该模型未能充分利用双时相遥感影像的上下文关系,在地物类型复杂时容易出现误检、漏检.注意力机制能够扩大网络的感受野,有些研究以此提高模型性能.例如,Chen等人[9]提出了多尺度时空注意力机制,使用度量的方式生成变化检测结果;Fang等人[10]提出的SNUNet采用了集成通道注意力模块,有效缓解了浅层特征融合较差的问题.虽然基于注意力机制的方法增强了卷积神经网络捕捉全局细节的能力,但在像素空间关系建模方面仍显不足,难以充分利用双时相遥感影像的远程细节信息和长距离依赖关系.随着Transformer在自然语言处理领域的成功应用,不少研究者将其应用到遥感影像变化检测.例如,BIT[11]和ICIF-Net[12]方法将卷积神经网络与Transformer方法结合.不同的是,BIT采用串行方式,将卷积神经网络提取的特征送入Transformer进一步捕捉全局信息;而ICIF-Net是并行结构,使用卷积神经网络和Transformer同时提取遥感图像的特征.ChangeFormer[13]则是一种基于Transformer的孪生网络变化检测方法.尽管这些结合Transformer的方法取得了较好的性能,但易受复杂背景干扰,未能充分地利用多层级特征图,导致精度受到一定影响.

在上述方法的模型训练中,输入通常是数据集中变化前和变化后的一组遥感图像对,模型输出的往往是精度较高的二值变化结果图.其中,数据集固定的图像输入顺序为正时序,即输入T1时相到T2时相两个时间点的遥感图像数据.与之相对,反时序是指输入T2时相到T1时相两个时间点的遥感图像数据.然而,当模型在测试阶段输入反时序的遥感图像对时,输出的结果出现了不同程度的恶化,如图1

所示.

针对上述问题,本文提出一种结合时序对齐与跨层特征混合的遥感影像变化检测方法 CINet(chronologic invariant network).本文主要工作如下:

1)首先,通过大量实验,证实了目前多数基于孪生网络的遥感影像变化检测方法存在改变双时相遥感影像的输入顺序后,变化检测模型的结果表现不佳这一问题.

2)其次,提出一种结合时序对齐与跨层特征混合的变化检测方法 CINet.通过设计的时序对齐模块和跨层特征混合模块,充分混合双时相遥感影像的时序信息和多层级特征,进一步提高特征图利用率,减少噪声和复杂背景干扰,生成对时序具有更强鲁棒性的特征图.

3)最后,在多个遥感数据集上的实验结果显示:CINet 相较于其他方法具有更优的检测性能和更强的时序鲁棒性,能够在改变输入图像顺序后仍取得良好的变化检测结果.



图 1　ChangeFormer 模型在 LEVIR-CD 数据集上的可视化结果

Fig.1　Visualization results of ChangeFormer model on the LEVIR-CD dataset

1　遥感图像输入顺序分析

为了进一步探究图像输入顺序对变化检测模型结果的影响,本文使用 8 种基于深度学习的遥感变化检测方法在不同遥感数据集上进行了实验.实验结果如图 2 和图 3 所示,令人惊讶的是,当输入反时序的双时相遥感影像时,这些模型的性能指标均出现显著下降.

实验结果清晰地表明,无论是基于卷积神经网络的还是结合 Transformer 方法的变化检测模型,在正时序场景下均能取得较高的变化检测精度,但在反时序场景下评价指标大幅下降,严重影响了变化检测模型在遥感场景中的实用性.目前对此现象的相关研究较少,文献[14]指出 FC-EF-Res 模型也存在这一问题,但并未证实此问题的普遍性.

模型输入反时序的双时相遥感影像,从实际场景的角度来看,尽管变化方向不同,但最终的变化检测结果应当一致.例如,在建筑物变化检测中,若在变化前的图像 T1 中某处为空地,而在变化后的图像 T2 中该区域出现了建筑物,即为新增建筑;反时序时,T2 中的建筑物在 T1 中变为空地,即为拆除建筑.模型理应能够识别此处为变化区域.然而,实验结果显示,多数变化检测模型在此情况下无法有效地识

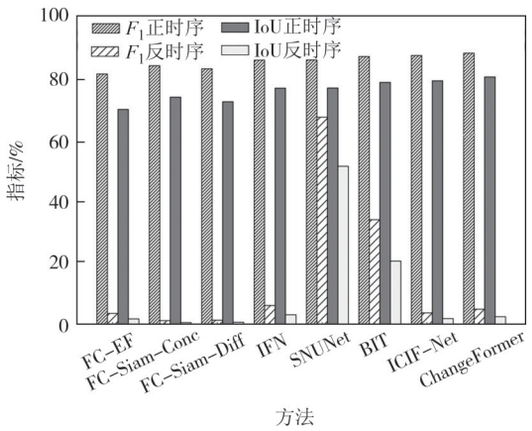


图2 在LEVIR-CD数据集上不同检测方法输出结果的F_1分数和交并比(IoU)

Fig.2 F_1 and intersection over union (IoU) of results from different methods on the LEVIR-CD dataset

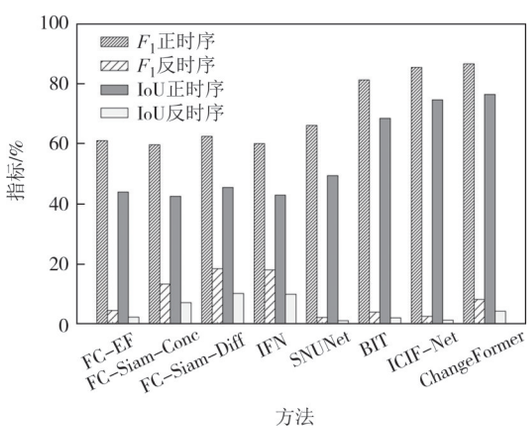


图3 在DSIFN数据集上不同检测方法输出结果的F_1分数和交并比

Fig.3 F_1 and IoU of results from different methods on the DSIFN dataset

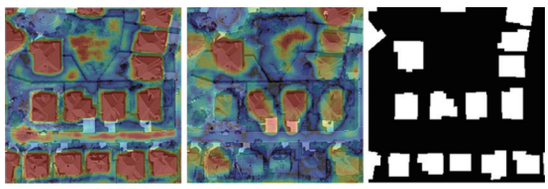
别其为变化区域,评价指标大幅下降.这表明需要进一步改进模型,提高模型对输入时序的鲁棒性.

目前多数变化检测模型以权重共享的孪生网络为编码器,即对T1图像和对T2图像进行特征提取的网络权重相同.在提取特征后,通常采用特征拼接的方式处理特征差异图,表示如下:

$$\boldsymbol{F}_{\mathrm{diff}}=\mathrm{Active}(\mathrm{Norm}(\mathrm{Conv}(\mathrm{Cat}(\boldsymbol{F}_{\mathrm{T1}},\boldsymbol{F}_{\mathrm{T2}})))) \quad (1)$$

式中:Cat表示特征拼接操作;$\boldsymbol{F}_{\mathrm{T1}}$、$\boldsymbol{F}_{\mathrm{T2}}$表示编码器对两幅图像提取的特征;Active表示激活函数;Norm表示归一化操作;Conv表示卷积.

特征拼接通常将两个时相的特征图沿通道维度进行连接,能够较好地保留特征图的完整信息.但也正因为这种方式,模型将正时序特征和反时序特征拼接后的结果视为两个截然不同的特征图,使得模型在改变输入顺序后难以精确地识别出变化区域,从而出现变化检测精度下降的现象,如图4所示.



(a)正时序 (b)反时序 (c)GT

图4 ChangeFormer在正时序和反时序的热图对比

Fig.4 Comparison of heatmaps of ChangeFormer model in normal chronology and anti-chronology

2 本文方法

2.1 整体网络结构

在遥感场景中,双时相遥感影像中的地物目标通常呈现出尺度和形态的多样性,加大了变化检测任务的难度.且受到光照、土壤类型等因素干扰,两幅遥感图像往往存在显著的图像风格差异.这导致改变图像输入顺序后,多数模型难以准确检测到变化的地物目标.

因此,本文受ChangeFormer变化检测模型结构的启发,提出了一种结合时序对齐与跨层特征混合的遥感影像变化检测方法CINet,所提方法的网络模型整体结构如图5所示.模型呈现编码-解码结构,模型工作时,首先由Transformer构成的多层级编码器,对遥感图像中地面变化目标的特征信息进行提取.其中,编码器通过提出的时序对齐模块,使双时相遥感影像在特征层面充分混合,缩小两幅遥感图像之间的数据分布距离,以缓解反时序对变化检测结果的影响.然后,通过设计的跨层特征混合模块,更有效地将编码器提取的浅层特征与深层特征进行混合,充分利用双时相遥感影像的空间几何特征和时序物候特征,从而进一步提高模型对多层级特征图的利用率.最后,通过转置卷积解码器和分类器,实现变化区域和非变化区域的区分.

2.2 编码器

CINet采用文献[15]中的Transformer Block作为主干网络(Backbone),用于提取双时相遥感影像的特征.相较于卷积神经网络,该方法得益于自注意力机制,可以计算像素块间的全局相关性,因此具有更大的有效感受野和更强的上下文建模能力[16].编码器采用分层结构,通过不同大小和步长的卷积将图像下采样为四个尺度,分别对应四个阶段,在每个阶段内通过Transformer Block提取不同大小的特征.这

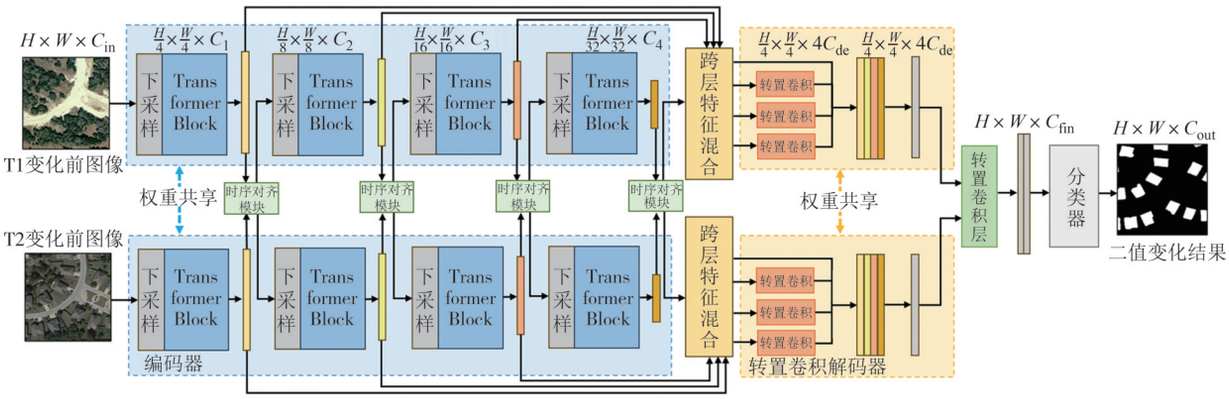


图 5　网络模型整体结构图

Fig.5　Overall architecture of the model

种层次化架构能够生成兼具高分辨率空间细节特征与低分辨率语义特征的多层级特征图,有效提升变化检测的性能.

考虑到遥感影像变化检测数据集通常规模有限,而 Transformer 方法在数据集图像数量较少时难以获得较好的特征表示能力[17],因此本文采用了有预训练权重的主干网络.尽管预训练阶段未对变化检测任务进行特定优化,但在处理颜色、纹理等视觉特征方面具有通用性[18].其中,纹理特征通过刻画像素与周围环境的空间关系,能够更好地表征地物的内部结构,这对变化检测任务具有重要意义[19].

2.3　时序对齐模块

囿于遥感影像在空间、光谱上的异质性[20],即使是同一地区的图像,由于获取时间和条件的不同,其数据分布也可能存在显著差异,具体表现为地物光谱响应曲线的偏移、图像统计特性的变化等.这种差异使得稳定的特征匹配关系难以建立.同时,反时序改变了特征差异的方向性,而模型训练时往往隐式地学习了正时序的变化模式,难以有效泛化到反时序的变化特征,导致预测结果不佳.

为了克服上述不足,本文受特征交互思想[21]的启发,提出一种时序对齐模块,如图 6 所示.该模块基于变化检测中的基本假设:任一像素或区域在两个时间点上仅存在“变化”或“未变化”两种状态.当像素值或空间特征发生显著变化时,无论时间顺序如何,均被视为变化.

该模块中,首先将第 i 阶段输入的 T1、T2 图像特征重塑(Reshape)为 $\boldsymbol{I}^i_{\text{T1}}$、$\boldsymbol{I}^i_{\text{T2}}\in H^i\times W^i\times C^i$,其中特征图的高度为 H^i、宽度为 W^i、通道数为 C^i.随后,按照图中所示方式在空间维度上进行交叉混合,交换特征映射,从而得到混合后的特征图 $\boldsymbol{M}^i_{\text{T1}}$、$\boldsymbol{M}^i_{\text{T2}}\in H^i\times W^i\times C^i$.

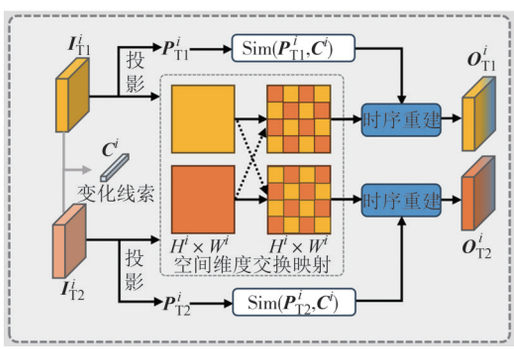


图 6　时序对齐模块

Fig.6　Chronological alignment module

然而,这种时序间的特征混合可能引入额外的噪声.为减轻这种噪声和复杂背景对 Transformer 序列化操作的影响,本模块对输入的特征图生成聚合全局变化信息的变化线索 $\boldsymbol{C}^i$.该变化线索具有时序不变性,能够有效地对双时相遥感影像中的固有变化特征进行建模.简单起见,使用绝对差值计算变化线索:

$$\boldsymbol{C}^i=(\text{GAP}(\text{Conv}(|\boldsymbol{I}^i_{\text{T1}}-\boldsymbol{I}^i_{\text{T2}}|))) \tag{2}$$

式中:GAP 为全局平均池化(global average pooling)操作.GAP 可以将 C 个形状为 $H^i\times W^i$ 的变化特征图的全局信息压缩为 $1\times1\times C$ 形状的数列,其中的每个实数计算方式如下:

$$\boldsymbol{C}^i_c=\frac{1}{H^i\times W^i}\sum_{m=1}^{H^i}\sum_{n=1}^{W^i}x_c(m,n) \tag{3}$$

式中:$x_c(m,n)$表示第 c 个变化特征图对应位置的值.

得到的变化线索用于后续的时序重建.为了将特征图与变化线索的特征表示对齐在同一空间,使用可学习参数的投影函数对特征图进行变换,这里使用简单的 1×1 卷积层实现,表示如下:

$$\boldsymbol{P}^i_{\text{T1}}=\text{ReLU}(\text{BN}(\text{Conv}(\boldsymbol{I}^i_{\text{T1}}))) \tag{4}$$

式中:ReLU 为激活函数;BN 为批归一化(batch nor-

malization, BN)操作.$\boldsymbol{I}_{\mathrm{T1}}^{i}$通过投影函数,得到特征图投影后的$\boldsymbol{P}_{\mathrm{T1}}^{i}$,用同样的方式计算$\boldsymbol{P}_{\mathrm{T2}}^{i}$.

投影后得到的两个特征$\boldsymbol{P}_{\mathrm{T1}}^{i}$与$\boldsymbol{P}_{\mathrm{T2}}^{i}$分别与变化线索$\boldsymbol{C}^{i}$进行相似度计算,得到关系$r^{i}=\mathrm{Sim}(\boldsymbol{P}^{i},\boldsymbol{C}^{i})$,此处相似度估计函数$\mathrm{Sim}(\cdot,\cdot)$采用点积实现,以提高计算效率.

最后,通过关系r^{i}来对混合后的特征图$\boldsymbol{M}^{i}$进行时序重建,重建后的特征映射$\boldsymbol{O}^{i}$计算方式如下:

$$\boldsymbol{O}^{i}=\frac{\boldsymbol{W}_{\mathrm{re}}(\boldsymbol{M}^{i})}{1+\mathrm{e}^{-r^{i}}} \tag{5}$$

式中:$\boldsymbol{W}_{\mathrm{re}}$为可学习的权重参数.

时序对齐模块与Transformer Block在编码阶段交替应用,能更好地减小双时相特征的时序差异.这有助于网络从复杂的数据中学习到更丰富的特征表示,同时通过时序重建降低了模型对于输入时序的依赖,增强时序鲁棒性,使其对反时序有更好的泛化能力.

2.4 跨层特征混合模块

受光照条件和季节变化的影响,不同时相中的地面目标常呈现弱边缘特征,这使得简单的特征融合策略难以充分利用多个层级特征图.文献[22]在目标检测任务中舍弃了浅层特征,认为深层特征具有更高级、更丰富的语义信息.而文献[23]指出,在遥感变化检测中,深层特征有助于变化目标的定位,而浅层特征则包含丰富的细节信息,对区分不同地物类型的边界至关重要.传统的特征融合方法主要局限于单一时相图像的特征融合,未能充分利用双时相遥感影像的空间几何特征和时序物候特征,导致出现时序差异.

为了克服上述不足,本文提出一种跨层特征混合模块.如图7所示,该模块通过全尺度连接结构和差异信息引导,实现在多个层级上充分混合T1、T2双分支所提取的浅层和深层特征.这种设计能够有效利用不同层级中与变化目标相关的上下文信息,提升多尺度特征图的利用效率,从而减少时序差异,提高时序鲁棒性.

特征金字塔结构能够使不同层级的特征图识别不同尺度的目标[23].因此,该模块在不同阶段的多个层级间进行特征混合,使得每个阶段的特征图能够有效聚合来自其他层级的、与目标相关的上下文特征,从而缓解特征在逐层传递过程中的信息衰减,生成更丰富、更精细的特征表示.此外,通过在各阶段引入另一时相的特征信息,利用差异信息引导网络捕获时相间的空间变化模式.

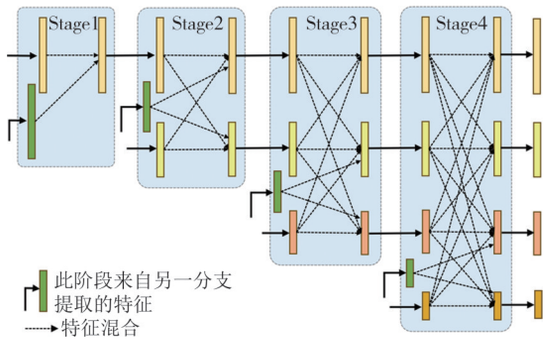


图7 跨层特征混合模块图

Fig.7 Cross-layer feature mixing module

图8详细展示了该模块在第二阶段的特征混合流程:它利用差异信息引导网络提取各层级中与变化目标相关的上下文特征,并自适应地聚合两幅遥感图像的特征表示,进一步增强了模型对于不同时序的鲁棒性.

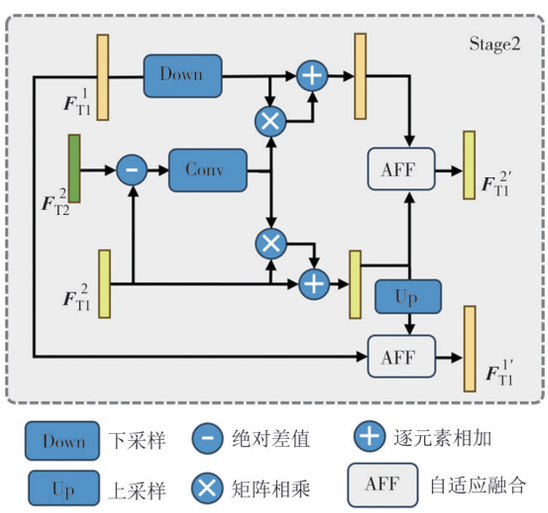


图8 跨层特征混合模块在Stage2的特征混合

Fig.8 Process of feature mixing in stage 2 of cross-layer feature mixing module

具体而言,模块输入包括T1图像在第一阶段的特征$\boldsymbol{F}_{\mathrm{T1}}^{1}$、第二阶段特征$\boldsymbol{F}_{\mathrm{T1}}^{2}$以及T2图像的第二阶段特征$\boldsymbol{F}_{\mathrm{T2}}^{2}$.首先,为了保持特征形状一致并减少计算量,对于重塑后的特征$\boldsymbol{F}_{\mathrm{T1}}^{1}$进行下采样和通道压缩.这里使用1×1的卷积调整通道数,然后使用大小为3、步长为2的卷积对特征图进行下采样,调整至合适的尺寸.对于第二阶段特征$\boldsymbol{F}_{\mathrm{T1}}^{2}$则进行上采样处理.随后,使用3×3的卷积层处理两幅图像特征图的差异信息$\boldsymbol{D}^{2}=\mathrm{Conv}(|\boldsymbol{F}_{\mathrm{T1}}^{2}-\boldsymbol{F}_{\mathrm{T2}}^{2}|)$,通过特征叠加突显变化区域,从而引导网络更聚焦于变化目标的特征表示.

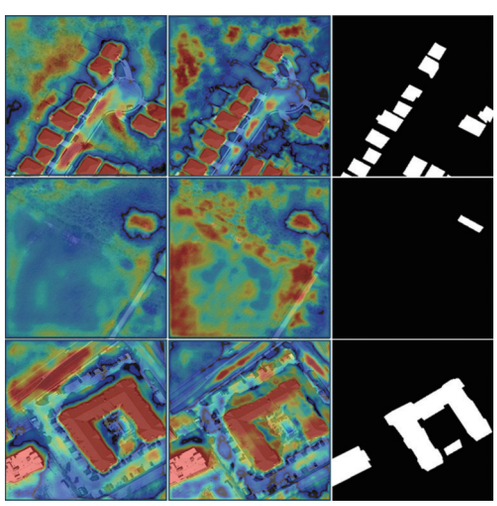
自适应空间上下文特征能够有效处理遥感图像

中形状和尺寸多样的地面目标边缘[24].因此,在融合的最后阶段,为了充分利用不同层级中的上下文信息来准确定位地面变化目标,将上采样后的$\boldsymbol{F}_{\mathrm{T1}}^{2}$与原$\boldsymbol{F}_{\mathrm{T1}}^{1}$进行自适应融合(adaptive feature fusion, AFF),生成混合后的$\boldsymbol{F}_{\mathrm{T1}}^{1'}$;将下采样后的$\boldsymbol{F}_{\mathrm{T1}}^{1}$和$\boldsymbol{F}_{\mathrm{T1}}^{2}$进行AFF得到$\boldsymbol{F}_{\mathrm{T1}}^{2'}$.具体而言,AFF首先将输入特征在通道维度上拼接,通过1×1卷积层学习各特征图的贡献度,然后使用Softmax函数归一化后赋予每层适当的权重,从而自适应聚合来自不同层级特征的丰富细节和语义信息.第二阶段的自适应融合AFF表示如下:

$$\mathrm{AFF}^{2}(\boldsymbol{x}_1,\boldsymbol{x}_2)=\alpha^{2}\times\boldsymbol{x}_1+\beta^{2}\times\boldsymbol{x}_2+\gamma^{2}\boldsymbol{D}^{2} \tag{6}$$

式中:$\alpha^{2},\beta^{2},\gamma^{2}$表示第二阶段的权重参数;$\boldsymbol{x}_1,\boldsymbol{x}_2$表示输入的特征图;$\boldsymbol{D}^{2}$表示第二阶段的差异信息.

此模块通过跨层级的特征混合,有效传递特征信息,使网络既能利用高层次特征校正地面变化目标的位置,又能以低层次特征的细节区分不同变化目标的边界.图9展示了特征可视化热图的结果对比,可以看出:与未加入该模块的结果相比,加入模块后的建筑物轮廓更加完整、边界更加清晰,对非变化区域的关注度也更低.



(a)有模块　(b)未加模块　(c)GT

图9　特征可视化热图

Fig.9　Feature visualization heatmaps

2.5　转置卷积解码器

在解码阶段,与其他多数模型的做法不同,为了更好地保留关键特征信息,缓解正时序特征与反时序特征在拼接后不一致的问题,CINet并未在早期阶段将双时相遥感影像的特征图进行融合.具体而言,经过跨层特征混合模块处理后的特征图分别送入图5所示的转置卷积解码器处理,随后进行最终分类.

为了克服传统上采样方式中重复像素或插值法导致的模糊、锯齿问题,解码器中采用了转置卷积.该方法有效利用像素的空间邻域关系,将低分辨率的特征图更有效地重建为高分辨率的特征图,以提高模型二分类的最终准确性.

3　实验结果与分析

3.1　实验环境设置

实验基于PyTorch框架,使用NVIDIA RTX 4080Super显卡加速运算.模型训练时,使用在ImageNet数据集上的预训练权重来初始化网络,采用二元交叉熵损失训练模型,使用AdamW作为优化器,学习率设置为0.004,衰减系数为0.01,批数量大小为16,训练周期为200轮.

3.2　数据集

为了测试CINet的性能,本文采用了不同场景的公开遥感数据集:LEVIR-CD和DSIFN数据集.

LEVIR-CD数据集[9]是Google Earth收集的美国得克萨斯州几个城市的637对高分辨率(0.5米/像素)图像对,尺寸为1 024×1 024像素,涵盖了不同季节、光照条件的多种类型的建筑.该数据集主要用于建筑物变化检测任务,单个变化实例的变化区域一般较小.本实验中将LEVIR-CD数据集中的图像裁剪为不重叠的256×256像素的图像,将其按原比例分为:7 120张图像的训练集、1 024张的验证集和2 048张的测试集.

DSIFN数据集[8]同样是Google Earth收集的遥感变化检测数据集,包括中国北京、深圳、成都、重庆、武汉、西安六个城市的394个双时相图像对,尺寸大小为512×512像素.该数据集经过数据增强扩充数量,包含大小各异的变化实例,变化区域平均发生在大尺度范围.本实验中将DSIFN数据集中的图像分为:3 600张的训练集、340张的验证集和48张的测试集.

3.3　评价指标

采用遥感影像变化检测中常用的四项指标来定量评估本次的实验结果,分别是精确率P(precision)、召回率R(recall)、F_1分数和交并比(IoU).计算公式如下:

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{7}$$

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{8}$$

$$F_1=\frac{2\times P\times R}{P+R} \tag{9}$$

$$\mathrm{IoU}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} \tag{10}$$

式中:TP表示图像中实际发生变化并且被正确检测出的像素数量;FP表示图像中实际未发生变化但被错误地检测为变化的像素数量;FN表示图像中实际发生了变化但被错误地检测为未变化的像素数量.

3.4 实验结果

为了验证CINet(本文方法)的性能,本实验将其与多个具有代表性的基于深度学习的遥感影像变化检测方法进行了对比实验.

表1呈现了在LEVIR-CD数据集上不同变化检测方法的评价指标比较,其中最优指标以加粗突出显示.实验结果表明,CINet在召回率、F_1分数和交并比指标上的表现均优于其他变化检测方法,证明了本方法的有效性.IFN方法在精确率上表现最佳,达到94.02%,但其他指标稍逊一筹.时序对齐模块和跨层特征混合模块的加入提升了多项评价指标.在正时序的场景下,相较于ChangeFormer方法,CINet的精确率、召回率、F_1分数和交并比分别提高了0.09个百分点、1.83个百分点、0.98个百分点、1.65个百分点,改进效果在图10的变化检测可视化结果中得到了直观体现.在反时序场景中,本方法的各项评价指标大幅领先于其他方法,能够保持良好的变化检测性能,其效果也可以在图11中观察到.

表1 在LEVIR-CD数据集上不同方法的变化检测结果的评价

Tab.1 Evaluation of change detection results of different methods on LEVIR-CD dataset %

方法	正时序测试集				反时序测试集			
	P	R	F_1	IoU	P	R	F_1	IoU
FC-EF[7]	85.92	81.22	83.50	71.68	4.10	3.36	3.69	1.88
FC-Siam-Conc[7]	88.53	84.03	86.22	75.78	8.88	0.68	1.26	0.63
FC-Siam-Diff[7]	90.70	80.35	85.21	74.23	33.82	0.73	1.43	0.72
IFN[8]	**94.02**	82.93	88.13	78.77	71.37	3.32	6.34	3.27
SNUNet[10]	89.18	87.17	88.16	78.83	90.53	55.80	69.05	52.73
BIT[11]	89.24	89.37	89.31	80.68	91.58	21.57	34.92	21.15
ICIF-Net[12]	91.06	88.24	89.63	81.20	46.65	2.02	3.87	1.97
ChangeFormer[13]	92.05	88.80	90.40	82.48	84.76	2.63	5.10	2.62
CINet(本文方法)	92.14	**90.63**	**91.38**	**84.13**	**91.67**	**91.09**	**91.38**	**84.12**

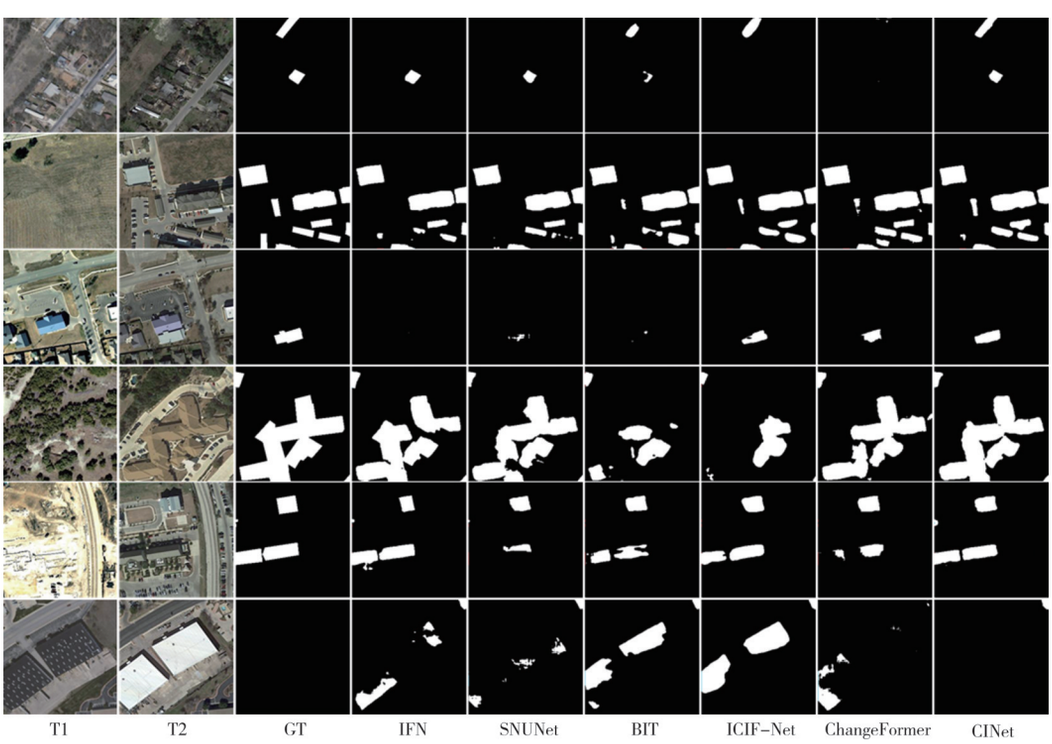


图10 不同方法在LEVIR-CD数据集上正时序的可视化结果

Fig.10 Visualization results of different methods in normal chronology on LEVIR-CD dataset

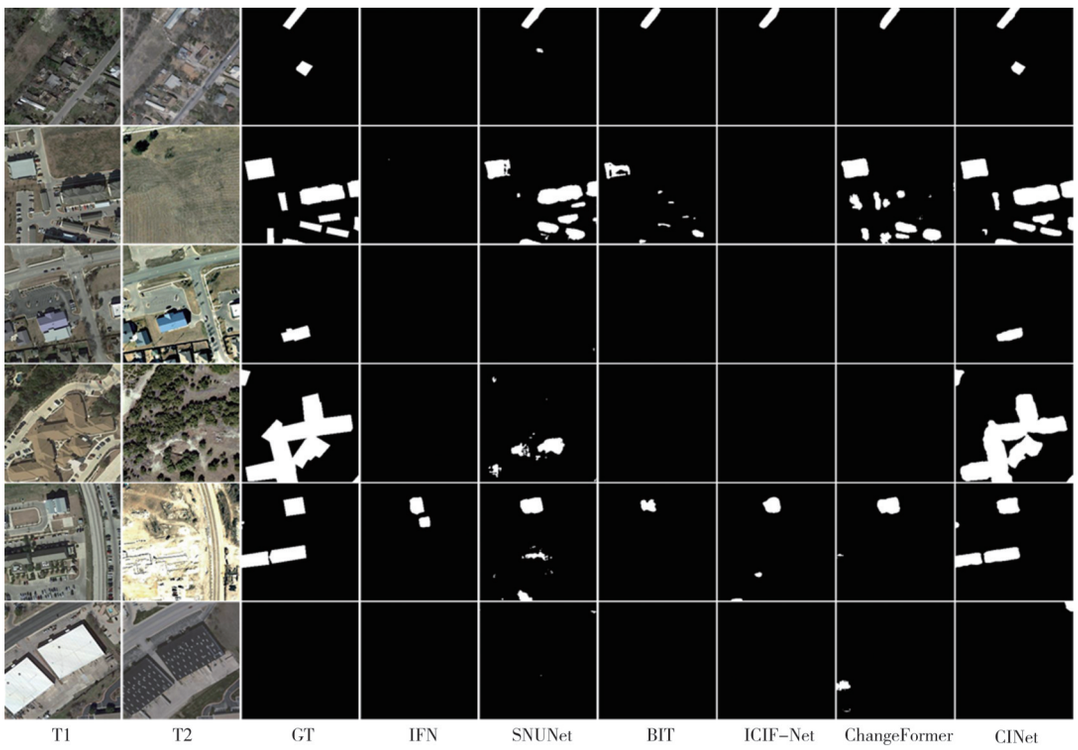


图 11　不同方法在 LEVIR-CD 数据集上反时序的可视化结果

Fig.11　Visualization results of different methods in anti-chronology on LEVIR-CD dataset

值得注意的是,图 12 展示了图 10 中第一行样本的详细分析.样本中,上方圈出区域为拆除建筑物(T1 时相存在,T2 中被拆除),下方圈出区域为新增建筑.在正时序情况下,大多数模型无法完整识别出上方的拆除建筑,但能识别下方的新增建筑;而在反时序时,多数模型却能较好地识别上方的新增建筑.这一现象表明,多数模型仅能识别单一方向的变化,对于反时序场景的鲁棒性较差.相比之下,CINet 在正时序、反时序场景下均能准确识别图中变化区域.

对于图 10 中第二行样本,多数方法能识别出较大的建筑物变化区域,但由于 T2 建筑物大小不一且存在阴影干扰,BIT 和 ChangeFormer 的结果出现了误检,其他方法也因建筑物较小出现不同程度的漏检,而本文方法漏检情况相对较少.第三行样本中,建筑物间隔较小,其他方法均出现严重漏检,而本方法对边界分割更加清晰.第四行样本中,建筑物轮廓较为复杂,BIT 和 ICIF-Net 出现大面积漏检,仅本方法能较为完整地分割建筑物.第五行样本中,SNUNet 出现漏检,IFN 的建筑物轮廓不够完整,其他方法因图像明显差异而导致结果粗糙,ICIF-Net 和本方法的结果较为完整.最后一行样本中,由于建筑物颜色不同且房顶有较多纹理,其他方法均误判为变化建筑物,仅有本方法正确识别为未变化区域.

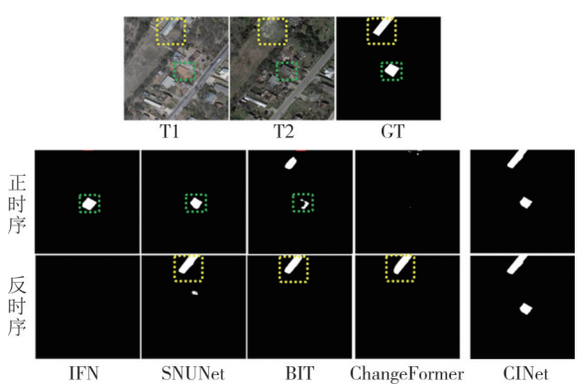


图 12　实验中第一组结果对比

Fig.12　The first set of results comparison

DSIFN 数据集的实验结果如表 2 所示,本文方法在精确率、召回率、F_1 分数和交并比四项指标上均表现最佳,在正时序场景下,相较于 ChangeFormer 方法,分别提高了 1.06 个百分点、1.97 个百分点、1.54 个百分点和 2.42 个百分点.得益于跨层特征混合模块,CINet 能够充分利用多层级特征图,可有效捕获跨尺度的变化信息,从而提高了模型识别 DSIFN 数据集中各种尺度变化实例的准确性.

为了进一步探究时序对齐模块在特征提取不同阶段的作用,本文在 CINet 主干网络的不同阶段加入时序对齐模块,实验结果的评价指标对比如表 3 所示.

表 2 在DSIFN数据集上不同方法的变化检测结果的评价

Tab.2 Evaluation of change detection results of different methods on DSIFN dataset %

方法	正时序测试集				反时序测试集			
	P	*R*	F_1	IoU	*P*	*R*	F_1	IoU
FC-EF[7]	72.61	52.73	61.09	43.98	6.74	3.41	4.53	2.31
FC-Siam-Conc[7]	66.45	54.21	59.71	42.56	19.01	10.27	13.34	7.14
FC-Siam-Diff[7]	59.67	65.71	62.54	45.50	33.91	12.69	18.47	10.17
IFN[8]	67.86	53.94	60.10	42.96	20.26	16.27	18.04	9.92
SNUNet[10]	60.60	72.89	66.18	49.45	7.03	1.33	2.24	1.13
BIT[11]	87.05	75.30	81.32	68.52	24.76	2.18	4.01	2.05
ICIF-Net[12]	85.19	85.84	85.51	74.69	6.03	1.63	2.57	1.30
ChangeFormer[13]	88.48	84.94	86.67	76.48	14.20	5.80	8.24	4.30
CINet	**89.54**	**86.91**	**88.21**	**78.90**	**89.52**	**86.93**	**88.21**	**78.91**

表 3 在LEVIR-CD数据集上时序对齐模块不同位置的结果对比

Tab.3 Comparison of chronological alignment module at different positions on LEVIR-CD dataset

Stage1	Stage2	Stage3	Stage4	*P*/%	*R*/%	F_1/%	IoU/%
▮				90.71	84.72	87.61	77.95
	▮			92.15	88.95	90.52	82.68
		▮		**93.23**	89.41	91.28	83.96
			▮	93.20	89.41	91.27	83.94
▮	▮			89.40	68.40	77.50	63.27
		▮	▮	92.14	**90.63**	**91.38**	**84.12**
	▮	▮	▮	93.21	87.36	90.19	82.13
▮	▮	▮	▮	91.43	76.09	83.06	71.02

实验结果表明,在CINet主干网络的第三、四阶段集成时序对齐模块能够获得最佳效果.具体而言,在第三阶段引入时序对齐模块时,精确率达到最高(93.23%);而在第三、第四阶段同时引入该模块时,召回率(90.63%)、F_1分数(91.38%)和IoU(84.12%)均达到最优.

然而,在网络更早期阶段引入时序对齐模块会导致变化检测结果指标下降,特别是当模块应用于主干网络的第一阶段时,评价指标显著降低.深入分析发现,这主要是由于第一阶段提取的浅层特征容易受到图像噪声和无关细节的影响,且语义信息较少,不足以有效引导网络模型准确识别地面变化目标.这一发现凸显了跨层特征混合模块的重要性,该模块能够有效利用早期阶段的低级特征与后续阶段的丰富语义特征,从而提高变化检测的整体性能.

3.5 方法效率

表4展示了不同方法的浮点运算次数(floating point operations, FLOPs)和模型参数量(parameters),测试图像的尺寸为256×256×3.从计算效率的角度来看,本文方法的FLOPs为29.46 G,相比IFN(82.26 G)和SNUNet(50.82 G)有显著降低,分别减少了64.2%和42.0%.在参数量方面,CINet的参数量为45.23 M,相较于IFN减少了10.8%,但明显高于SNUNet(12.04 M)和BIT(3.55 M).

总体而言,本文的CINet在获得更好检测性能的同时也带来了一定的资源开销.未来研究可以考虑进一步优化结构,以降低参数量,提高其效率.

表 4 方法效率对比

Tab. 4 Comparison of method efficiency

方法	FLOPs/G	参数量/M
IFN[8]	82.26	50.71
SNUNet[10]	50.82	12.04
BIT[11]	10.63	3.55
ICIF-Net[12]	25.36	23.82
ChangeFormer[13]	21.18	29.75
CINet	29.46	45.23

4 结 论

1) 揭示了现有遥感影像变化检测方法在时序鲁棒性方面存在的不足:多数基于孪生网络的遥感影像变化检测方法,在改变双时相遥感影像的输入顺序时,结果表现不佳.

2) 提出了一种结合时序对齐与跨层特征混合的遥感影像变化检测方法CINet,实现了在反时序场景下对地面变化目标的有效检测,为提高变化检测模型的时序鲁棒性提供了新的方案.

3) CINet因其独特的时序对齐模块和跨层特征混合模块,能够有效利用时序信息和多层级特征,相比其他方法展现出更优的检测性能和更强的时序鲁

棒性.

尽管本文所提出的方法在实验中表现出较好的性能,但仍存在一定的局限性.由于采用的Transformer方法自身计算量较大,所提方法中又存在多处卷积运算,因此模型参数量较大,解决此问题需要对该方法进行更深入的研究.

参考文献

[1] WOODCOCK C E, LOVELAND T R, HEROLD M, et al. Transitioning from change detection to monitoring with remote sensing: a paradigm shift[J]. Remote Sensing of Environment, 2020,238:111558.

[2] CHEN C F, SON N T, CHANG N B, et al. Multi-decadal mangrove forest change detection and prediction in Honduras, central America, with landsat imagery and a Markov chain model [J]. Remote Sensing,2013,5(12):6408-6426.

[3] MAHDAVI S,SALEHI B,HUANG W M,et al. A PolSAR change detection index based on neighborhood information for flood mapping[J]. Remote Sensing, 2019, 11(16): 1854.

[4] WU T J, LUO J C, ZHOU Y N, et al. Geo-object-based land cover map update for high-spatial-resolution remote sensing images *via* change detection and label transfer [J]. Remote Sensing, 2020, 12(1): 174.

[5] 眭海刚, 冯文卿, 李文卓, 等. 多时相遥感影像变化检测方法综述[J]. 武汉大学学报(信息科学版), 2018, 43(12): 1885-1898.
SUI H G,FENG W Q,LI W Z,et al. Review of change detection methods for multi-temporal remote sensing imagery [J]. Geomatics and Information Science of Wuhan University, 2018, 43(12): 1885-1898.(in Chinese)

[6] BAI T, WANG L, YIN D M, et al. Deep learning for change detection in remote sensing: a review[J]. Geo-spatial Information Science,2023,26(3):262-288.

[7] DAUDT R C, SAUX B L, BOULCH A. Fully convolutional Siamese networks for change detection [C]//2018 25th IEEE International Conference on Image Processing (ICIP). Athens, Greece. IEEE,2018:4063-4067.

[8] ZHANG C X, YUE P, TAPETE D, et al. A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 166:183-200.

[9] CHEN H, SHI Z W. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection[J]. Remote Sensing, 2020, 12(10): 1662.

[10] FANG S, LI K Y, SHAO J Y, et al. SNUNet-CD: a densely connected Siamese network for change detection of VHR images [J]. IEEE Geoscience and Remote Sensing Letters, 2021, 19: 8007805.

[11] CHEN H, QI Z P, SHI Z W. Remote sensing image change detection with transformers[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 5607514.

[12] FENG Y C, XU H H, JIANG J W, et al. ICIF-net: intra-scale cross-interaction and inter-scale feature fusion network for bitemporal remote sensing images change detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 4410213.

[13] BANDARA W G C, PATEL V M. A transformer-based Siamese network for change detection [C]//IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium. Kuala Lumpur, Malaysia. IEEE, 2022: 207-210.

[14] 郑思宇,胡华浪,黄进,等. 双通道孪生网络在遥感变化检测中的应用[J]. 计算机系统应用,2022,31(3):56-64.
ZHENG S Y, HU H L, HUANG J, et al. 2-channel Siamese network for remote sensing change detection [J]. Computer Systems and Applications, 2022, 31(3): 56-64.(in Chinese)

[15] XIE E Z, WANG W H, YU Z D, et al. SegFormer: simple and efficient design for semantic segmentation with transformers [J]. Advances in Neural Information Processing Systems, 2021, 34: 12077-12090.

[16] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [M]//Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 213-229.

[17] DING X H, ZHANG X Y, HAN J G, et al. Scaling up your kernels to 31 × 31: revisiting large kernel design in CNNs [C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA. IEEE, 2022: 11953-11965.

[18] SOWRIRAJAN H, YANG J, NG A Y, et al. Moco pretraining improves representation and transferability of chest X-ray models [C]//Medical Imaging with Deep Learning. Lubeck, Germany, 2021: 728-744.

[19] 王建华,秦其明,高中灵,等. 加入空间纹理信息的遥感图像道路提取[J]. 湖南大学学报(自然科学版),2016,43(4):153-156.
WANG J H, QIN Q M, GAO Z L, et al. Road extraction from high-resolution remote sensing imagery by including spatial texture feature [J]. Journal of Hunan University (Natural Sciences), 2016, 43(4): 153-156.(in Chinese)

[20] 柳钦火,曹彪,曾也鲁,等. 植被遥感辐射传输建模中的异质性研究进展[J]. 遥感学报,2016,20(5):933-945.
LIU Q H, CAO B, ZENG Y L, et al. Recent progresses on the remote sensing radiative transfer modeling over heterogeneous vegetation canopy[J]. Journal of Remote Sensing, 2016, 20(5): 933-945.(in Chinese)

[21] FANG S, LI K Y, LI Z. Changer: feature interaction is what you need for change detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5610111.

[22] SUN X D, WU P C, HOI S C H. Face detection using deep learning: an improved faster RCNN approach[J].Neurocomputing, 2018, 299: 42-50.

[23] LI Z L, TANG C, WANG L Z, et al. Remote sensing change detection *via* temporal feature interaction and guided refinement [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022,60:5628711.

[24] ZHANG D, ZHANG H W, TANG J H, et al. Feature pyramid transformer [M]//Computer vision-ECCV 2020. Cham: Springer International Publishing, 2020: 323-339.

[25] LV Z Y, ZHANG P F, SUN W W, et al. Novel land-cover classification approach with nonparametric sample augmentation for hyperspectral remote-sensing images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 4407613.