

基于改进 BIT 的耕地变化检测算法*

徐世亮¹, 赖民权², 刘继忠^{2†}

(1. 江西省自然资源事业发展中心, 江西 南昌 330025; 2. 南昌大学 先进制造学院, 江西 南昌 330031)

摘要: 耕地“非农化”严重威胁全球粮食安全与生态稳定。遥感变化检测技术因具有大范围动态监测优势, 已成为识别耕地“非农化”过程的核心手段。然而, 现有方法在处理复杂耕地场景时, 受限于多尺度特征表征能力, 难以在提取破碎耕地边缘细节的同时兼顾大范围地块的全局语义一致性, 从而导致边缘模糊与局部特征丢失。针对上述问题, 本文提出一种基于改进双时相图像 Transformer(Bitemporal Image Transformer, BIT)的耕地变化检测算法 Far-CDNet。首先, 引入并联普通卷积及多种差分卷积的细节增强卷积模块, 并通过动态加权和残差连接强化特征提取网络的边缘细节表征能力。其次, 将 BIT 模块中语义标记器的普通卷积替换为深度可分离卷积, 以增强局部特征捕获能力并生成具有更高层语义的输出特征。最后, 增加一条残差分支, 进一步融合 Transformer 前后的局部及全局信息。实验结果表明, 改进后的模型 $F1$ 分数为 79.18%, IoU 为 69.32%, 相较于 BIT 模型, $F1$ 分数提升 4.17%, IoU 提升 4.24%。

关键词: 变化检测; 非农化; 双时相图像 Transformer; 细节特征增强卷积模块; Transformer

DOI: 10.13568/j.cnki.651094.651316.2025.07.07.0004

中图分类号: TP391 **文献标识码:** A **文章编号:** 2096-7675(2026)02-0144-012

引文格式: 徐世亮, 赖民权, 刘继忠. 基于改进 BIT 的耕地变化检测算法[J]. 新疆大学学报(自然科学版中英文), 2026, 43(2): 144-155.

英文引文格式: Xu Shiliang, Lai Minquan, Liu Jizhong. Farmland change detection algorithm based on improved BIT[J]. Journal of Xinjiang University(Natural Science Edition in Chinese and English), 2026, 43(2): 144-155.

Farmland Change Detection Algorithm Based on Improved BIT

Xu Shiliang¹, Lai Minquan², Liu Jizhong²

(1. Jiangxi Natural Resources Development Center, Nanchang Jiangxi 330025, China;

2. School of Advanced Manufacturing, Nanchang University, Nanchang Jiangxi 330031, China)

Abstract: Farmland non-agriculturalization is a serious threat to global food security and ecological stability. Remote sensing change detection technology has become a core tool for identifying the process of farmland non-agriculturalization by virtue of its advantage of large-scale dynamic monitoring. However, existing methods face challenges in balancing the extraction of fine edge details in fragmented farmland with the maintenance of global semantic consistency in large-scale fields, often resulting in blurred edges and lost local features. To address these issues, a farmland change detection algorithm based on improved BIT, named Far-CDNet, is proposed. Firstly, a detail enhancement convolution module that connects ordinary convolution and multiple differential convolutions in parallel is introduced, and the edge detail representation capability of the feature extraction network is enhanced through dynamic weighting and residual connection. Secondly, the ordinary convolution of the semantic tokenizer in the BIT module is replaced by a deep separable convolution to enhance the local feature capture ability and generate output features with higher-level semantics, so as to improve the overall feature expression ability of the model. Finally, a residual branch is added to further integrate the local and global information before and after the Transformer. The experimental results show that the improved model $F1$ score is 79.18%, and IoU is 69.32%. Compared with the BIT model,

* 收稿日期: 2025-07-07; 修回日期: 2026-01-09; 录用日期: 2026-01-14.

基金项目: 江西省高层次高技能领军人才培养工程项目“基于 GIS 与视频融合的自然资源监管关键技术研究”(2022233); 江西省自然资源厅科技创新项目“基于智能监控与地理信息融合技术在自然资源监管中的应用研究”(202317).

作者简介: 徐世亮(1981—), 男, 硕士, 正高级工程师, 从事计算机软件技术、网络与信息安全的研, E-mail: vigor2005@163.com.

† 通信作者: 刘继忠(1974—), 男, 博士, 教授, 博士生导师, 主要从事智能机电系统与机器人、机器视觉与图像处理的研究, E-mail: liujizhong@ncu.edu.cn.

the *F1* score is increased by 4.17%, and *IoU* is increased by 4.24%.

Key words: change detection; non-agriculturalization; bitemporal image Transformer; detail feature enhancement convolution module; Transformer

0 引言

农业生产是全球粮食安全的保障。然而,受人口快速增长和气候变化的影响,耕地作为农业活动的基本要素,存在一些“非农化”活动,包括不合理造林、挖掘湖泊、扩大储备和非法建筑等^[1],这不仅扰乱了当地的农业生态系统,还会威胁粮食供应。为了及时获取耕地信息,确保耕地生产和粮食安全,耕地的快速动态变化检测(Change Detection, CD)极为重要。

现有耕地分布范围广泛,通过人工实地调查获取耕地动态变化既费力又耗时。随着卫星影像的广泛应用,遥感技术已成为一种有效且实用的方法,可用于地形分类、建筑物覆盖区提取和土地覆盖CD等方面。相较于城市扩张或森林砍伐等变化,耕地变化检测面临更为严峻的挑战。首先,耕地、裸地和部分建设用地在光谱特征上存在高度相似性^[2],尤其在作物生长周期的不同阶段易产生混淆。其次,耕地内部结构复杂,田埂、垄沟等细微纹理丰富,且“非农化”活动(如非法建筑、挖塘)往往呈现不规则、破碎的边缘,对算法的边缘细节感知能力提出了极高要求^[3]。最后,耕地变化情况多样,从点状的零星占用到面状的大规模开发,变化信号在初期往往十分微弱,且易受邻近地物(如道路、水体)变化的干扰^[4]。

传统CD方法主要基于图像变换,如主成分分析(Principal Component Analysis, PCA)^[5]、变化向量分析(Change Vector Analysis, CVA)^[6]和多变量变化检测^[7]等,这些方法存在光谱-空间特征耦合性不足、边界模糊及噪声敏感等缺陷,在处理耕地场景时易存在边缘模糊、细节丢失、误检漏检率高等问题。随机器学习的持续发展,基于分类处理方法在遥感土地覆盖变化检测中被广泛应用。Bai等^[8]通过对比随机森林、支持向量机、深度信念网络和AdaBoost在高分辨率卫星影像变化检测中的性能,发现支持向量机在小样本场景下的检测精度最高,而随机森林随样本量增加展现更优的总体准确率。Devries等^[9]基于随机森林的后分类和传统CVA更新年度耕地变化地图。这些分类方法通过时序影像分类比较实现变化检测,但其非线性建模能力受限,且依赖高质量的分类模型与大规模训练样本,导致计算成本较高。

随着深度学习被广泛应用于遥感领域,以深度学习为代表的人工智能技术逐渐被引入目标变化检测任务中。基于深度学习的变化检测方法可以有效捕捉更多的复杂特征,进一步提取不同目标的空间特征,使检测结果更加准确。循环神经网络(Recurrent Neural Network, RNN)在捕获序列关系方面效果较好,并已广泛应用于与序列数据相关的领域。Lyu等^[10]设计了一种具有可转移性的有效变化规则,以检测土地覆盖二元和多元类别的变化。其所设计的模型采用改进长短期记忆(Long Short-Term Memory, LSTM)模型记录核心记忆单元中的差异信息或多类变化来学习稳定且可迁移的变化规则。相较于卷积神经网络(Convolutional Neural Network, CNN),该方法通过对图像的时间序列信息进行建模取得了更为出色的结果,然而其并未充分考虑图像在空间中的特征信息。

在训练样本充足的情况下,CNN强大的学习能力使其在高维信息研究中获得更准确丰富的特征。CNN及其架构不断改进对遥感影像场景分类、目标检测和变化检测等应用作出了重大贡献。Wang等^[11]首次将Faster R-CNN迁移至高分辨率遥感变化检测,构建了端到端检测框架,但对复杂耕地边界变化敏感度不足。詹必伟等^[12]提出了多时相光谱融合策略,通过深度学习实现了耕地变化信息的自动化精准提取。然而,这些基于传统CNN的耕地CD模型普遍采用深层卷积堆叠和大尺度卷积核维持全局特征,导致模型复杂度显著增加,存在较高的计算资源依赖问题。

起源于自然语言处理领域的Transformer架构,近年来在计算机视觉领域展现出显著的跨领域适应性。该架构在图像分类、语义分割、目标检测等核心视觉任务中均得到广泛应用,其突出的全局依赖建模能力有效缓解了传统CNN存在的长程语义关联缺失问题。受这些工作的启发,Chen等^[13]将CNN和Transformer同时引入变化检测领域,创新性地构建双时相图像Transformer(BIT)模型,在语义标记器中将输入图像编码为上下文丰富的语义标记。尽管BIT模型在常规遥感CD方面具有较好性能,但在处理耕地变化场景时受限于多尺度特征表征能力,难以在提取耕地破碎边缘细节的同时兼顾大范围地块的全局语义一致性。这种多尺度表征能力的缺失,直接导

致了边缘细节丢失与全局信息融合不充分的连锁反应. 针对上述问题,本文以BIT模型为基础,提出了一种改进的耕地变化检测网络模型(Farmland Change Detection Network, Far-CDNet).

本文的贡献如下:

1)通过引入并联的标准卷积(Vanilla Convolution, VC)、中心差分卷积(Center Difference Convolution, CDC)、角度差分卷积(Angle Difference Convolution, ADC)、水平差分卷积(Horizontal Difference Convolution, HDC)和垂直差分卷积(Vertical Difference Convolution, VDC)的细节增强卷积模块,并通过动态加权和残差连接以强化特征提取网络ResNet18的边缘细节表征能力.

2)通过将BIT模块中语义标记器的普通卷积替换为深度可分离卷积,以增强局部特征捕获能力并生成具有更高层语义的输出特征.

3)通过增加一条残差分支,进一步融合Transformer前后的局部及全局信息,以较小的成本融合和增强不同尺度的特征,提高网络对耕地变化区域的整体检测性能.

1 Far-CDNet模型结构

BIT模型是一种端到端的孪生网络,本文基于BIT模型改进的Far-CDNet模型结构如图1所示. 其由5个模块组成,分别为卷积特征提取模块、改进的语义标记器、Transformer模块、残差分支模块与预测模块. Far-CDNet模型采用两支并行的特征处理通路,最后通过残差分支将二者融合输出:1)细节保持通路(卷积流). 利用ResNet18_S4F及特征提取增强模块(Feature Extraction Enhancement Module, FEEM)专注于提取田埂、田块边界等高频几何特征,确保耕地轮廓的完整性,具体结构见图2(a);2)全局语义通路(Transformer流). 利用Transformer捕捉长程上下文依赖,以克服作物生长周期不同带来的光谱异质性问题,保证大面积耕地内部预测的一致性. 最后,通过残差分支将两者融合,实现对耕地“非农化”区域的精准定位.

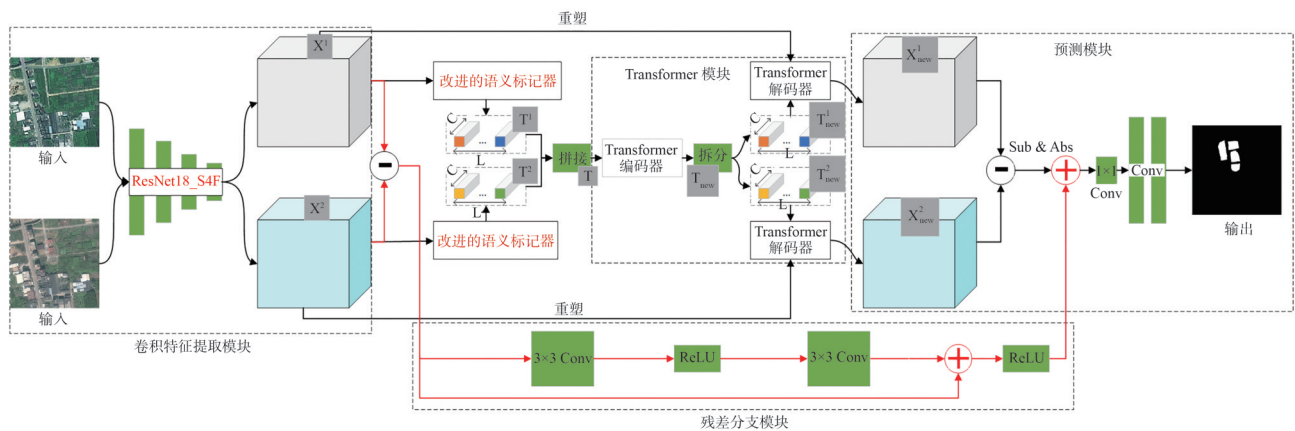


图 1 Far-CDNet模型结构

Figure 1 Far-CDNet model structure

1.1 卷积特征提取模块

Far-CDNet特征提取模块网络如图2所示,以ResNet18_S4F作为基础框架,通过移除初始全连接层构建特征提取器,具体由 7×7 卷积核构成的首层卷积模块、3个级联的残差块和使用双线性插值法(Bilinear)进行2倍的上采样组成. 首层卷积采用步幅2进行下采样,生成半分辨率浅层特征. 继而通过 3×3 最大池化层(步幅2)实施二次下采样,获得输入图像 $1/4$ 尺度的中级特征.

各残差块包含双 3×3 卷积层的级联结构,其间集成批量归一化层与ReLU激活函数. 特征映射通过跳跃连接实现元素级加和操作,该残差连接机制在ReLU激活前完成特征融合. 经FEEM增强的特征输入至ResBlock2,其首卷积层配置步幅2,实现特征图尺寸的逐级减半;而ResBlock3的首卷积层采用步幅1,经3阶段残差处理,最终输出特征图的空间分辨率降为原始输入的 $1/8$.

耕地场景中往往包含大量细微的线性特征,普通卷积容易对其进行平滑处理从而导致边界模糊. 为此,本文在首个残差块后嵌入FEEM模块,如图2所示. 其中原细节增强卷积模块(Detail-Enhanced Convolution, DE-

Conv)采用异构卷积并行架构,集成VC、CDC、ADC、HDC与VDC 5类特征提取通路,如图3(a)所示. 然而,部署5个并行卷积层进行特征提取会导致参数和推理时间增加. DEConv利用卷积层的可加性,将并行部署的卷积简化为单一的标准卷积,如图3(b)所示. 即给定输入特征 F_{in} ,DEConv可以利用重参数化技术将具有相同计算代价和推理时间的 F_{out} 输出到普通卷积层. 表达式为(为了简化,省略偏差项):

$$\begin{aligned}
 F_{out} &= \text{DEConv}(F_{in}) = \sum_{i=1}^5 F_{in} * K_i \\
 &= F_{in} * \left(\sum_{i=1}^5 K_i \right) = F_{in} * K_{cvt}
 \end{aligned} \tag{1}$$

式中:DEConv()为本文所提DEConv操作, $K_i = 1:5$ 分别为VC、CDC、ADC、HDC和VDC的核,*为卷积操作, K_{cvt} 为转换后的核,将并行卷积结合.

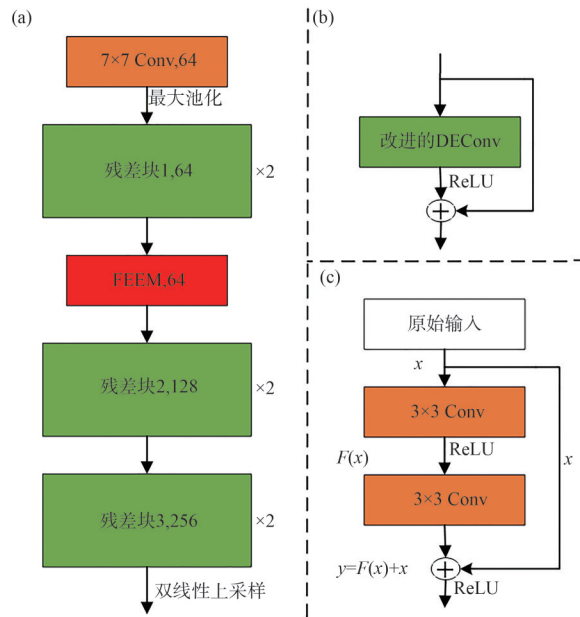


图2 (a)ResNet18_S4F,(b)FEEM,(c)残差块网络结构

Figure 2 (a) ResNet18_S4F, (b) FEEM, (c) Residual block network structure

然而,DEConv的核心局限在于其采用简单的权重线性叠加策略,将5种不同方向差分卷积的权重直接相加,这种各向同性融合机制忽略了不同方向卷积核在细节捕捉中的异质性贡献,导致高频特征(如边缘、纹理)与低频背景信息在融合过程中产生混淆,尤其在复杂场景下易引入噪声或弱化细微结构. 另外,DEConv单一的输出路径缺乏对多分支特征的动态调控能力,难以适应图像局部区域的频域特性差异. 为克服上述问题,本研究首先利用通道拼接与 1×1 卷积实现多方向权重的非线性交互,生成兼具全局感知与局部敏感性的融合权重,如图3(c)所示. 在此基础上,通过双路卷积增强与Sigmoid门控函数构建自适应加权机制,使网络能够依据区域内容动态调节不同增强路径的贡献比例,在保留原始信息(残差连接)的同时,优先强化有效细节并抑制冗余响应. 门控加权模块的权重表达式为:

$$a = \text{sigmoid}(res1 + res2). \tag{2}$$

1.2 改进的语义标记器

在图像处理领域,Transformer架构常采用非重叠分块的输入策略,即将输入图像分割为固定尺寸的补丁后展开其像素灰度值构成二维矩阵,其中每个补丁对应一个Token. 与这种直接空间离散化的方法不同,BIT网络创新性地通过卷积操作构建结构化特征矩阵作为Transformer输入. 针对耕地“非农化”过程中常见的零星违建或小规模非法挖塘现象,单一的全局建模可能忽略这些局部异常信号. 故将改进的语义标记器引入深度卷积,在生成Transformer的Token前先进行局部的细粒度特征聚合,防止微小变化目标在全局投影过程中丢失.

图4为改进的语义标记器,设前一卷积模块输出特征张量 $X \in R^{H \times W \times C}$,其处理流程包含以下关键步骤:首先,采用深度卷积提取输入特征空间维度的细粒度特征,通过独立处理各输入通道的空间相关性,有效捕获局部上下文信息;其次,通过 1×1 卷积将通道维度投影至 L 维,获得 $X' \in R^{H \times W \times L}$;随后沿通道维度展平得到 $X'' \in R^{HW \times L}$,并在 HW 维度施加Softmax生成注意力图 $p \in R^{HW \times L}$;同时将原始特征 X 按通道展平为 $R \in R^{HW \times C}$;最终通过矩阵相乘得到Transformer输入 $T \in R^{L \times C}$. 表达式为:

$$T = (p)^T R \left(\sigma \left(\left(\phi_2 \left(\phi_1 \left(X; w_d \right) \right); w_p \right) \right)^{reshape} \right)^T X^{reshape}, \quad (3)$$

式中: σ 为Softmax函数, ϕ_2 为采用 $w_p \in R^{L \times C \times 1 \times 1}$ 的逐点卷积, ϕ_1 为采用卷积核 $w_d \in R^{C \times k \times k \times k}$ 的深度卷积操作(k 为空间卷积核尺寸).

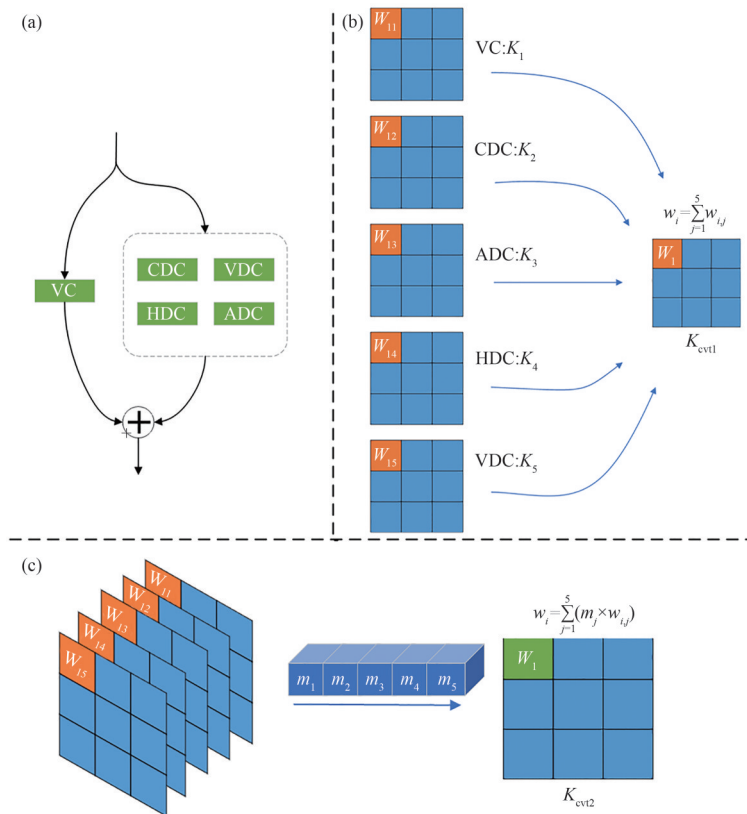


图3 改进的(a)DEConv, (b)重参数化, (c)通道拼接再卷积压缩

Figure 3 Improved (a) DEConv, (b) Reparametrisation, (c) Channel concatenation followed by re-convolution compression

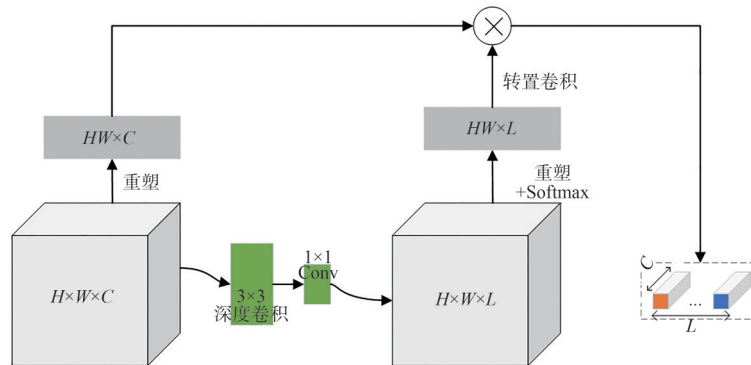


图4 改进的语义标记器

Figure 4 Improved semantic labeller

1.3 Transformer 模块

上述语义标记器的作用是将卷积神经网络提取的高维特征图通过空间注意力机制下采样为一组紧凑的语义标记。这些标记能够有效聚合影像中的语义信息,从而形成一种高度抽象的视觉词汇表。Transformer 编码器则负责在这些标记之间建立全局的语义关系,通过多头自注意力机制和多层感知机模块,对时空上下文信息进行建模。这一过程能够捕捉标记之间的长距离依赖关系,并生成富含上下文信息的标记表示。最后,Transformer 解码器将这些上下文丰富的标记重新投影回像素空间,通过多头交叉注意力机制与原始像素特征进行融合,从而增强原始特征的语义表达能力。

1.4 残差分支模块

针对直接串联 Transformer 导致的局部特征退化现象,在卷积特征层与 Transformer 模块间构建残差连接机制,将卷积层提取的高频空间细节(耕地边界)直接注入 Transformer 输出的低频语义特征(变化区域主体)中,这种跨域融合机制可有效避免深层网络对耕地边缘信息的平滑效应。具体而言,通过引入 ResNet 残差块作为跨模态特征桥接器,建立局部细节特征与全局语义特征的并行传输路径。表达式为:

$$F_{\text{enhanced}} = \mathcal{R}(F_{\text{cnn}}) + \mathcal{T}(F_{\text{cnn}}), \quad (4)$$

式中: $\mathcal{R}(\cdot)$ 为残差块的非线性映射, $\mathcal{T}(\cdot)$ 为 Transformer 的全局注意力运算。该设计通过特征加性融合策略,使网络在保留卷积层高频细节信息的同时,充分融合 Transformer 提取的长程依赖关系。

1.5 预测模块

在特征解码阶段,计算 Transformer 解码器输出的双时相特征图间的差分张量以捕获时序变化特征,随后将该差分特征与残差连接分支提供的空间细节特征进行通道维度拼接,形成具备时空联合表征能力的融合特征。

基于上述融合特征,预测模块采用极简架构设计的全卷积网络(Fully Convolutional Network, FCN)实现端到端变化判别。该架构首先通过 1×1 卷积层实现通道压缩,构建可学习参数空间约简;随后采用双线性插值进行空间分辨率重建,并通过级联的 3×3 卷积核实现非线性映射与通道数降维;最终采用 Sigmoid 激活函数实施逐像素逻辑回归,输出二值变化图。表达式为:

$$BC = \text{sigmoid}(f_{3 \times 3}^{(2)} \circ f_{3 \times 3}^{(1)} \circ \uparrow \circ f_{1 \times 1}(X)), \quad (5)$$

式中: $\text{sigmoid}(\cdot)$ 为 Sigmoid 函数, $f_{3 \times 3}^{(1)}$ 与 $f_{3 \times 3}^{(2)}$ 均为 3×3 卷积, \uparrow 为双线性插值, $f_{1 \times 1}(\cdot)$ 为 1×1 卷积, X 为输入。

2 实验与分析

为验证本文所提 Far-CDNet 算法在耕地变化检测任务中的性能和改进模块的有效性,设计了对比实验与消融实验。

2.1 开发环境及数据集构建

2.1.1 开发环境

本文在 Ubuntu 操作系统下基于 Pytorch 框架进行开发训练,具体硬件配置及软件系统配置分别如下。

操作系统:Ubuntu 操作系统;处理器(CPU)Intel(R) Xeon(R) Gold 6152;内存:30 G;显卡(GPU):RTX 3090(24 GB);Pytorch:2.1.2;Python:3.8;CUDA:12.1。

2.1.2 数据集构建

本文采用公开的耕地变化检测基准数据集 CLCD 作为基础实验数据。该数据集包含 600 组双时相遥感影像样本(2017 vs 2019),按 6:2:2 比例划分为训练集(360 组)、验证集(120 组)和测试集(120 组)。影像数据来源于中国高分二号(GF-2)卫星采集的广东省农业区域多光谱数据,空间分辨率具有 0.5~2 m 的多尺度特征,每组样本包含:1)512×512 像素的双时相配准影像;2)耕地“非农化”二进制标注图。针对原始数据集样本多样性不足的问题,本文通过江西省地理信息公共服务平台采集九江地区 2020—2023 年间 GF-2 卫星多光谱影像,经几何配准、辐射校正等预处理后进行耕地二值化标注,再通过遍历对比双时相图片标签的像素形成变化标签,以此补充数据集。考虑收集及制作的数据集仍不足,易产生过拟合现象,故通过旋转、平移、镜像等操作方式对数据集进行增强扩充,使数据集扩充至 3 250 组。变化标签的制作过程如图 5 所示。

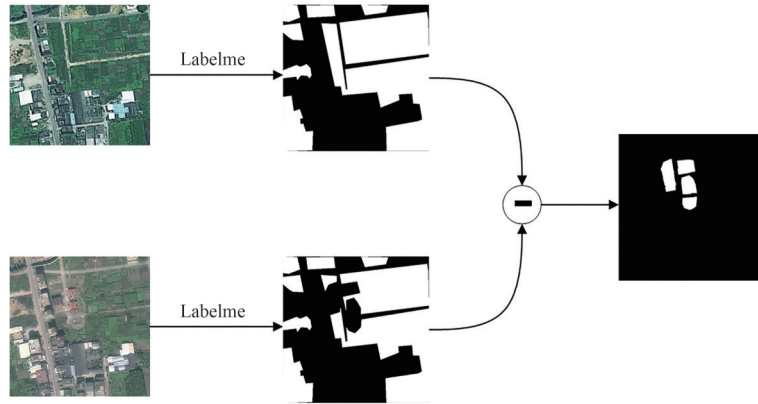


图 5 变化标签制作过程

Figure 5 Variable label production process

2.2 实验设置与评价指标

2.2.1 实验设置

将数据增强后的 3 250 组数据集按 CLCD 相同的 6:2:2 比例划分为训练集(1 950 组)、验证集(650 组)和测试集(650 组)。算法运行参数:epochs 设置为 200;图片输入尺寸为 512×512;Batchsize 为 8;动量设置为 0.99;权重衰减系数为 0.000 5;学习速率初始设置为 0.01,线性衰减系数为 0。

2.2.2 评价指标

为结合遥感图像的语义分割和变化检测,本文采用了经典和常用的评价标准,即精确率 P (Precision)、召回率 R (Recall)、 $F1$ 分数和 IoU 。表达式为:

$$P = \frac{TP}{TP + FP} \times 100\%, \quad (6)$$

$$R = \frac{TP}{TP + FN} \times 100\%, \quad (7)$$

$$F1 = \frac{P_{re} \times R_e \times 2}{P_{re} + R_e} \times 100\%, \quad (8)$$

$$IoU = \frac{TP}{TP + FN + FP} \times 100\%, \quad (9)$$

式中: TP 和 TN 分别为成功预测为耕地信息的像素数和其他非耕地信息的像素数, FP 和 FN 分别为未检测到的耕地和其他非耕地的像素数。每个评价指标的具体含义为: P_{re} 精确率是衡量模型在所有正例中真正检测到多少正例的指标,受误判为正例的样本数量影响,当模型对负例做出错误预测时,精确率会降低; R_e 召回率衡量模型对正例的检测能力,受未检测到的正例数量影响,漏检时召回率会降低; $F1$ 分数是考虑了精确率和召回率的加权数字评估; IoU 是衡量预测区域与真实区域间的交集比例,受每个类别预测准确性影响。为消除类别分布偏差对模型评估的影响,本实验采用类别间(变化类与非变化类)的宏平均指标,以均衡度量模型在全域场景下的综合分类性能。

2.3 对比实验

将 Far-CDNet 网络与 FC-EF^[14]、FC-Siam-Diff^[14]、FC-Siam-Conc^[14]、Siam-NestedUNet^[15]、SwinSUNet^[16]、ChangeFormer^[17]、BIT 及 CGNet(Change Guiding Network)^[18]等主流变化检测模型进行比较。各对比模型简介如下。

1)FC-EF(全卷积早期融合):基于图像级融合策略,通过通道维度拼接将双时相影像整合为单一输入,并输入全卷积网络进行特征提取与分类。

2)FC-Siam-Diff(全卷积孪生差异网络):采用特征级差异融合机制,通过共享参数的 Siamese FCN 提取多级特征,利用特征张量差分运算实现双时相信息融合。

3)FC-Siam-Conc(全卷积孪生连接网络):基于特征级连接融合范式,通过双分支 Siamese FCN 并行提取多

尺度特征,采用通道拼接方式融合时空特征.

4) Siam-NestedUNet: 结合 Siamese 网络与 NestedUNet 的多尺度特征连接架构,通过密集跳跃连接保留高分辨率语义特征. 该模型在解码器各层级引入通道注意力机制,并采用深度监督策略增强中间特征判别性.

5) SwinSUNet: 基于纯 Transformer 架构的对称 U 型网络,其编码器、特征融合模块和解码器均采用 Swin Transformer 构建,通过窗口多头自注意力机制捕获全局上下文信息.

6) ChangeFormer: 集成层次化 Transformer 编码器与 MLP 解码器的孪生网络,通过多尺度自注意力机制建模长程依赖关系,有效提取变化检测所需的跨时空细节特征.

7) CGNet: 基于深度学习的变化检测方法,提出一种创新性的多尺度特征融合策略,结合变化引导模块,以解决传统 U-Net 在变化特征表达上的不足. 该网络通过生成深度变化图作为先验信息,增强对变化区域的定位能力,同时通过自注意力机制捕获长距离依赖关系,显著提升边缘检测的精度和内部区域的完整性.

对比实验结果如表 1 所示, Far-CDNet 的精确率(81.19%)、召回率(77.45%)、F1 分数(79.18%)和 IoU (69.32%) 4 项指标均超越主流方法: 相较于基于传统卷积网络的方法,如 FC-EF (其 F1 分数为 71.38%), Far-CDNet 通过集成 Transformer 模块和语义标记器以增强全局信息融合,实现了 F1 分数 7.80% 的提升. 与同样采用 Transformer 架构但专注于多层次自注意力机制建模的 ChangeFormer (IoU 为 61.80%) 相比, Far-CDNet 利用增加的残差分支和优化后的语义标记器来改进特征融合策略,使得 IoU 提升 7.52%. 面对结合孪生网络结构并结合多尺度特征融合的 CGNet (IoU 为 66.06%), Far-CDNet 凭借其细节增强卷积模块 (FEEM) 进一步强化边缘细节捕捉能力,实现 IoU 额外增长 3.26%. 即便与基础 BIT 模型相比,经过模块优化后的 Far-CDNet 仍能在精确率上提高 2.37%. 实验验证了 Far-CDNet 在全局建模、局部特征融合及边缘精度上的综合优势.

表 1 对比实验

Table 1 Comparative experiments

模型	P/%	R/%	F1/%	IoU /%
FC-EF	73.06	69.98	71.38	61.43
FC-Siam-Diff	72.95	69.90	71.29	61.35
FC-Siam-Conc	68.02	72.04	69.75	59.51
SwinSUNet	78.37	74.38	75.44	65.53
ChangeFormer	77.75	67.71	72.43	61.80
Siam-NestedUNet	76.87	72.16	74.19	63.16
CGNet	79.32	73.61	75.94	66.06
BIT	78.82	72.24	75.01	65.08
Far-CDNet	81.19	77.45	79.18	69.32

图 6 的可视化结果直观反映出不同模型在测试集上的变化检测结果,其中: Image1 和 Image2 分别是同一区域两个不同时相的遥感影像, Label 是人工标注的真实耕地变化区域. 传统卷积架构模型(如 FC-EF 系列)虽能实现基础变化边界检测,但存在显著的漏检与误检现象,同时伴随检测区域的空间不连续性缺陷. 在纯 Transformer 架构中, ChangeFormer 表现出严重的特征遗漏问题,而 SwinSUNet 虽可有效捕捉宏观变化区域,但在处理边缘细节时表现出明显的特征退化问题. 值得注意的是,基于 Siamese 网络与 NestedUNet 构建的多尺度特征融合架构 Siam-NestedUNet,在特定测试样本中产生了明显的假阳性区域. 相较之下,结合变化引导模块和自注意力机制的 CGNet 模型的检测精度指标较 SwinSUNet 显著提升,但仍存在局部误判问题. 作为 CNN-Transformer 混合架构的基准模型, BIT 在测试中暴露出边界模糊及局部漏检缺陷,这表明其未能有效实现卷积操作的局部敏感性与 Transformer 全局上下文建模的优势融合. 而通过引入细节增强卷积模块、改进语义标记器及增加残差融合信息分支的 Far-CDNet 模型能较精确地识别出大面积的变化区域,并且在边缘区域仍能保持较高的检测精度.

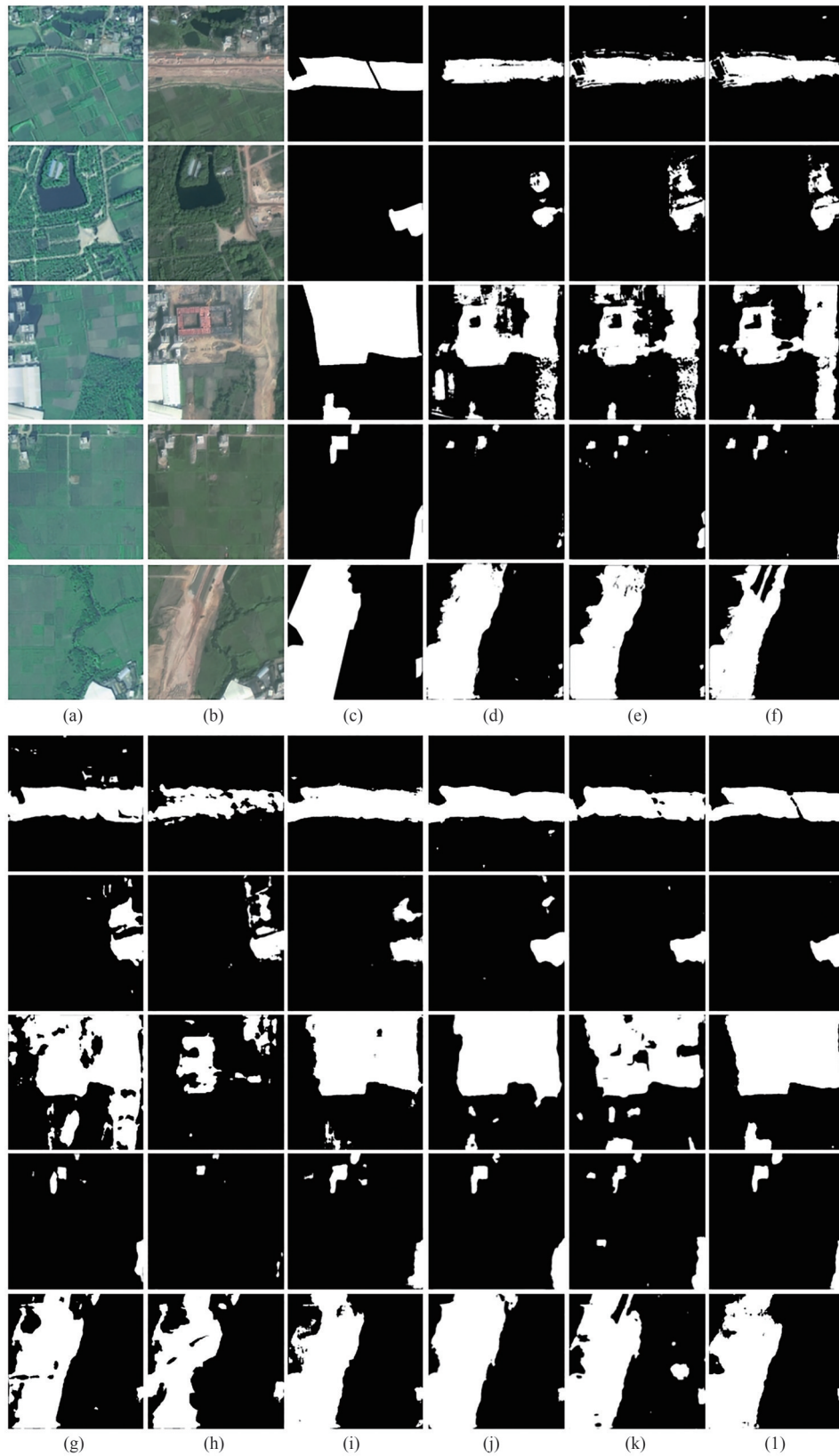


图 6 不同方法在测试集上的可视化结果:(a)Image1,(b)Image2,(c)Label,(d)FC-EF,(e)FC-Siam-Diff,(f)FC-Siam-Conc,(g)Siam-NestedUNet,(h)ChangeFormer,(i)SwinSUNet,(j)CGNet,(k)BIT,(l)Far-CDNet

Figure 6 Visualisation results of different methods on the test set: (a) Image1, (b) Image2, (c) Label, (d) FC-EF, (e) FC-Siam-Diff, (f) FC-Siam-Conc, (g) Siam-NestedUNet, (h) ChangeFormer, (i) SwinSUNet, (j) CGNet, (k) BIT, (l) Far-CDNet

2.4 消融实验

为验证改进模型各模块性能,设计了基于基础BIT模型($F1$ 分数为75.01%, IoU 为65.08%)的消融实验,逐步引入改进模块进行性能验证,实验结果如表2所示.首先,集成改进的FEEM后,模型在保持精确率稳定和轻量化的同时,召回率显著提升至74.92%,增幅达2.68%.相较于仅通过增加特征提取网络深度(BIT+ResNet18_S5)而参数量及计算量显著增加的方法,FEEM使 $F1$ 分数提升0.62%,这进一步证实了其多方向卷积核非线性融合机制在强化边缘细节表达方面的有效性.其次,在引入改进的语义标记器后, $F1$ 分数进一步提升至77.66%,相对提升2.65%.该语义标记器采用深度卷积与逐点卷积的级联结构,增强了局部特征的细粒度表达能力,从而提高了对复杂耕地变化区域的识别精度.最后,Far-CDNet通过增加残差分支实现动态信息融合,达到79.18%的 $F1$ 分数和69.32%的 IoU ,分别较基础BIT模型提升4.17%和4.24%,且参数量及计算量仅增加了1.40%和1.09%.模块协同分析表明,FEEM提供了高频特征信息,改进后的语义标记器增强了局部上下文建模能力,残差分支则通过全局与局部特征的动态融合,显著提升了模型的鲁棒性和整体表现.实验结果充分验证了各改进模块在优化耕地变化检测中的协同效应及其重要性.

表 2 消融实验
Table 2 Ablation experiments

模型	$P/\%$	$R/\%$	$F1/\%$	$IoU/\%$	Params/M	FLOPs/G
BIT	78.82	72.24	75.01	65.08	3.006	34.01
BIT+FEEM	78.52	74.92	76.57	66.56	3.027	34.03
BIT+ResNet18_S5	80.30	72.87	75.95	66.03	11.474	105.24
BIT+FEEM+改进语义标记器	79.02	76.44	77.66	67.67	3.027	34.04
Far-CDNet	81.19	77.45	79.18	69.32	3.048	34.38

消融实验可视化对比结果如图7所示. BIT模型的预测结果整体上能够识别出主要的变化区域,但在细节层面存在明显缺陷:变化区域边界呈现锯齿状、变化区域中部出现小部分漏检现象.引入FEEM后,模型对变化区域边界的刻画能力及中部漏检现象得到显著改善.在此基础上,进一步集成改进的语义标记器,模型不仅保持了良好的边缘细节,而且减少了部分区域的假阳性误判.最后,通过残差分支有效融合Transformer前后的局部细节与全局语义信息,其预测结果在整体上最为接近真实标签(Label).该模型精准地勾勒出大范围变化区域的宏观形态,同时在微观尺度上也保留了丰富的细节结构.

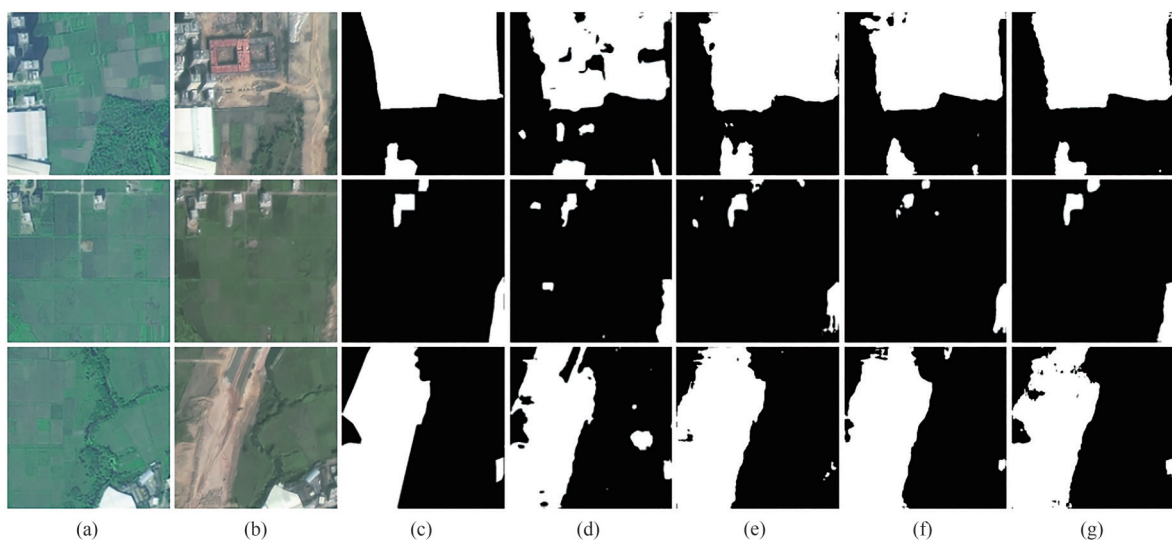


图 7 消融实验可视化对比结果:(a)Image1,(b)Image2,(c)Label,(d)BIT,(e)BIT+FEEM,
(f)BIT+FEEM+改进语义标记器,(g)Far-CDNet

Figure 7 Visualisation comparison of ablation experiments: (a) Image1, (b) Image2, (c) Label, (d) BIT, (e) BIT+FEEM, (f) BIT+FEEM+improved semantic labeller, (g) Far-CDNet

3 结论

本文围绕耕地“非农化变化中难以兼顾细碎物理边界提取与大范围田块语义一致性建模这一关键难点,构建了Far-CDNet变化检测模型.对比实验表明,Far-CDNet在测试集上取得 P 为81.19%、 R 为77.45%、 $F1$ 分数为79.18%、 IoU 为69.32%,整体优于多种主流方法,较BIT分别提升4.17%($F1$ 分数)和4.24%(IoU).消融结果进一步说明,特征提取增强模块、改进的语义标记器与残差分支融合局部与全局信息三者带来稳定增益,且模型参数量与计算量仅小幅增加,具备较好的精度-效率权衡与工程可部署性,为有效监测和管理农业用地“非农化”提供了可靠的技术手段.

参考文献:

- [1] 关洁琳. 国内外耕地非农化研究进展及展望[J]. 南方农机,2023,54(22):42-45.
Guan J L. Research progress and prospect of non-agriculturalization of arable land at home and abroad[J]. China Southern Agricultural Machinery,2023,54(22):42-45. (in Chinese)
- [2] 汪杰,孙孝林,伍丹华,等. 高密度植被覆盖区基于无人机高光谱数据的农田土壤有机质反演[J]. 光谱学与光谱分析,2025,45(10):2885-2896.
Wang J,Sun X L,Wu D H,et al. Prediction of soil organic matter for farmlands covered with high density of vegetation based on UAV hyperspectral data[J]. Spectroscopy and Spectral Analysis,2025,45(10):2885-2896. (in Chinese)
- [3] 李林源,黄华国,穆西哈,等. 低空无人机植被定量遥感:进展、挑战与展望[J]. 遥感学报,2025,29(6):2083-2113.
Li L Y,Huang H G,Mu X H,et al. Low-altitude UAV-based quantitative remote sensing of vegetation:Advances, challenges, and prospects[J]. National Remote Sensing Bulletin,2025,29(6):2083-2113. (in Chinese)
- [4] 习文强,杜世宏,杜守基. 多时相耕地覆盖提取和变化分析:一种结合遥感和空间统计的时空上下文方法[J]. 地球信息科学学报,2022,24(2):310-325.
Xi W Q,Du S H,Du S J. Multi-temporal cropland cover extraction and change analysis: A spatio-temporal contextual approach combining remote sensing and spatial statistics[J]. Journal of Geo-information Science,2022,24(2):310-325. (in Chinese)
- [5] Marpu P,Gamba P,Benediktsson J A. Hyperspectral change detection using IR-MAD and feature reduction[C]//2011 IEEE International Geoscience and Remote Sensing Symposium,July 24-29,2011,Vancouver,BC,Canada. New York:IEEE,2011:98-101.
- [6] Singh S,Talwar R. Review on different change vector analysis algorithms based change detection techniques[C]//2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013),December 9-11,2013,Shimla,India. New York:IEEE,2013:136-141.
- [7] Li P J,Cheng T,Moser G,et al. Multitemporal change detection by spectral and multivariate texture information[C]//2007 IEEE International Geoscience and Remote Sensing Symposium,July 23-28,2007,Barcelona,Spain. New York:IEEE,2007:1922-1925.
- [8] Bai T,Sun K M,Deng S Q,et al. Comparison of four machine learning methods for object-oriented change detection in high-resolution satellite imagery[C]//MIPPR 2017:Remote Sensing Image Processing,Geographic Information Systems,and Other Applications,October 28-29,2017,Xiangyang,Hubei,China. New York:SPIE,2018:99-106.
- [9] Devries B,Decuyper M,Verbesselt J,et al. Tracking disturbance-regrowth dynamics in tropical forests using structural change detection and Landsat time series[J]. Remote Sensing of Environment,2015,169:320-334.
- [10] Lyu H B,Lu H,Mou L C. Learning a transferable change rule from a recurrent neural network for land cover change detection[J]. Remote Sensing,2016,8(6):506-528.
- [11] Wang Q,Zhang X D,Chen G Z,et al. Change detection based on Faster R-CNN for high-resolution remote sensing images[J]. Remote Sensing Letters,2018,9(10):923-932.
- [12] 詹必伟,唐小丽,何奕萱,等. 基于改进U-Net网络的耕地变化检测方法研究[J]. 测绘,2023,46(2):76-79.
Zhan B W,Tang X L,He Y X,et al. Research of farmland change detection technology based on improvement U-Net[J]. Surveying and Mapping,2023,46(2):76-79. (in Chinese)
- [13] 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: 1-14.
- [14] Daudt R C, LE Saus B, Boulch A. Fully convolutional Siamese networks for change detection [C]//2018 25th International Conference on Image Processing, October 7-10, 2018, Athens, Greece. New York: IEEE, 2018: 4063-4067.
- [15] Fang S, Li K Y, Shao J Y, et al. SNUNet-CD: A densely connected Siamese network for change detection of VHR images [J]. Geoscience and Remote Sensing Letters, 2021, 19: 1-5.
- [16] Zhang C, Wang L J, Cheng S L, et al. SwinSUNet: Pure transformer network for remote sensing image change detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-13.
- [17] 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, July 17-22, 2022, Kuala Lumpur, Malaysia. New York: IEEE, 2022: 207-210.
- [18] Han C X, Wu C, Guo H N, et al. Change guiding network: Incorporating change prior to guide change detection in remote sensing imagery [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16: 8395-8407.

责任编辑: 刘 敏