文章编号:1007-2780(2026)03-0402-13

基于特征子空间解耦与迭代残差精炼的高分辨率遥感影像变化检测

李晗之[1,2], 李海巍[1*], 郭　琦[1,2], 赵　翼[1,2], 宋丽瑶[3],
李思远[1], 刘思含[4], 谢宇浩[4]
(1. 中国科学院 西安光学精密机械研究所, 陕西 西安 710119;
2. 中国科学院大学, 北京 101408;
3. 西安工业大学 计算机科学与工程学院, 陕西 西安 710021;
4. 生态环境部卫星环境应用中心, 北京 100006)

摘要:随着高分辨率遥感对地观测技术的发展,影像中丰富的地物纹理细节在提升信息量的同时,也引入了由光照、阴影及季节性物候差异导致的复杂背景噪声。针对高分辨率遥感变化检测中复杂背景噪声引发的伪变化误检,以及传统上采样导致的微小目标细节丢失问题,提出一种基于特征解耦与迭代精炼的网络(DIR-Net)。首先,利用预训练的Fast-SAM作为视觉先验编码器,提取多尺度鲁棒特征。接着,设计特征子空间解耦模块,通过正交投影和交叉重校准策略,显式地将双时相特征分解为共享语义子空间和差异特征子空间,从源头抑制环境噪声。最后,提出迭代残差精炼模块,引入坐标注意力机制,将解码过程建模为由粗到精的残差回归问题,在保持分辨率的特征空间中逐步恢复微小目标的边缘细节。在LEVIR-CD、WHU-CD和SYSU-CD三个公开数据集上的实验结果表明,DIR-Net的F1分数分别达到了91.33%、93.31%和86.29%。相比于主流的ChangeFormer和BIT算法,F1分数平均提升了约5.0%,显著降低了伪变化误报率,同时保持了极高的召回率。该方法有效解决了特征耦合与细节丢失的难题,在复杂场景下具有更强的鲁棒性和更高的边界定位精度。

关　键　词:遥感影像变化检测;特征解耦;迭代残差精炼;深度学习;DIR-Net;高分辨率影像;小目标检测
中图分类号:TP391.4;TP751.1　**文献标识码**:A　**doi**:10.37188/CJLCD.2026-0003　**CSTR**:32172.14.CJLCD.2026-0003

收稿日期:2026-01-07;**修订日期**:2026-02-01.
基金项目:国家重点研发计划(No.2022YFF1300201);陕西省教育厅一般专项科研计划(No.24JK0481);陕西省自然科学基础研究计划(No.2025JC-YBQN-366, No.2025JC-YBMS-256)
Supported by National Key Research and Development Program of China (No.2022YFF1300201); General Special Scientific Research Program Project of the Shaanxi Provincial Department of Education (No.24JK0481); Natural Science Foundation of Shaanxi Province (No.2025JC-YBQN-366, No.2025JC-YBMS-256)
*通信联系人,E-mail:lihaiwei@opt.ac.cn

High-resolution remote sensing image change detection based on feature subspace decoupling and iterative residual refinement

LI Hanzhi[1,2], LI Haiwei[1*], GUO Qi[1,2], ZHAO Yi[1,2], SONG Liyao[3], LI Siyuan[1], LIU Sihan[4], XIE Yuhao[4]

(1. *Xi'an Institute of Optics and Precision Mechanics*, *Chinese Academy of Sciences*, *Xi'an* 710119, *China*;
2. *University of Chinese Academy of Sciences*, *Beijing* 101408, *China*;
3. *School of Computer Science and Engineering*, *Xi'an Technological University*, *Xi'an* 710021, *China*;
4. *Satellite Environment Application Center*, *Ministry of Ecology and Environment*, *Beijing* 100006, *China*)

Abstract: With the development of high-resolution remote sensing earth observation technology, the rich texture details in images enhance information content but also introduce complex background noise caused by lighting, shadows, and seasonal phenological differences. To address the issues of pseudo-change misdetection caused by complex background noise and the loss of small target details caused by traditional upsampling in high-resolution remote sensing change detection, a network based on feature decoupling and iterative refinement(DIR-Net) is proposed. First, the pre-trained FastSAM is utilized as a visual prior encoder to extract multi-scale robust features. Next, a Feature Subspace Decoupling Module is designed to explicitly decompose bi-temporal features into a shared semantic subspace and a differential feature subspace through orthogonal projection and cross-recalibration strategies, suppressing environmental noise from the source. Finally, an Iterative Residual Refinement Module is proposed. By introducing a coordinate attention mechanism, the decoding process is modeled as a coarse-to-fine residual regression problem, gradually recovering the edge details of small targets in the resolution-maintained feature space. Experimental results on three public datasets, LEVIR-CD, WHU-CD, and SYSU-CD, demonstrate that the F1 scores of DIR-Net reached 91.33%, 93.31%, and 86.29%, respectively. Compared with mainstream ChangeFormer and BIT algorithms, the F1 score improved by an average of approximately 5.0%. The proposed method significantly reduces the false alarm rate of pseudo-changes while maintaining a very high recall rate. This method effectively resolves the challenges of feature coupling and detail loss, demonstrating stronger robustness and higher boundary localization accuracy in complex scenes.

Key words: remote sensing image change detection; feature decoupling; iterative residual refinement; deep learning; DIR-Net; high-resolution image; small target detection

1　引　　言

随着先进光学成像系统与对地观测技术的快速发展，亚米级高分辨率遥感影像已广泛应用于城市规划、灾害评估及生态监测等领域[1-3]。变化检测(Change Detection, CD)作为其中的核心任务，旨在从双时相影像中识别地表覆盖的动态变化。然而，高分辨率影像中丰富的大气湍流、纹理细节和复杂的光照条件为现有算法带来了严峻挑战[4-5]。

近年来，结合几何感知、孪生卷积神经网络(Siamese CNN)和 Transformer 的深度学习方法已成为主流[6-9]。尽管 FC-Siam-Diff、SNUNet 和 ChangeFormer 等方法在特征提取能力上取得了显著进展，但仍面临两大瓶颈：首先是特征耦合导致的“伪变化”问题。现有方法多采用特征差分或拼接融合，难以在特征空间中有效区分由季节、光照引起的非语义差异(即同物异谱)与真实地物变化，导致高误检率。其次是下采样导致的小目标细节丢失。深层网络在扩大感受野的同

时滤除了高频信息，使得在恢复原分辨率时难以重建建筑物边缘和小尺度目标的精细轮廓，造成漏检和边界模糊[10-12]。

针对上述问题，本文提出了一种基于特征子空间解耦与迭代残差精炼的网络（DIR-Net）。本文的主要工作如下：利用FastSAM构建强视觉先验；设计FSDM模块，通过正交投影解耦背景与变化特征；提出IRRM模块，利用迭代残差学习逐步精炼细节。实验证明，该方法在LEVIR-CD、WHU-CD以及SYSU-CD数据集上均优于当前主流方法。

2 DIR-Net模型构成及原理

2.1 整体架构

DIR-Net的整体架构如图1所示，系统采用了经典的编码器-解码器（Encoder-Decoder）宏观结构，但在内部信息流转上进行了深度创新[13]。

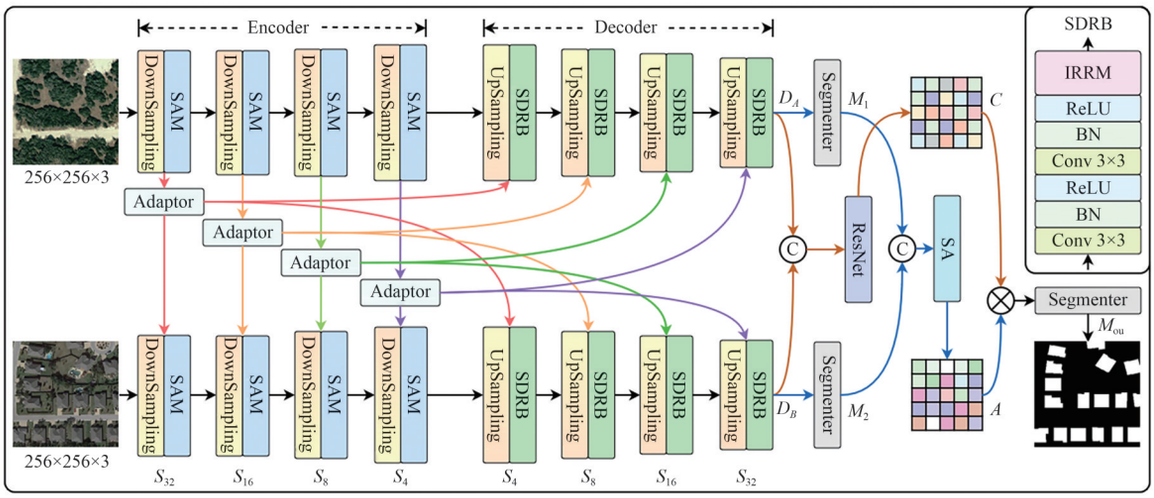

图1 DIR-Net整体架构示意图

Fig. 1 Overall architecture of DIR-Net

2.2 非对称孪生编码器

输入端接收双时相影像。不同于传统CD网络训练自定义的主干网络，DIR-Net利用了视觉基础模型（VFMs）的强大泛化能力[14]。具体而言，系统采用FastSAM（Fast Segment Anything Model）作为特征提取器。FastSAM基于YOLOv8-seg架构，在SA-1B海量数据集上进行预训练，具备极强的“分割一切”的零样本泛化能力[15-16]。在训练过程中，FastSAM的编码器权重被冻结，这不仅大幅降低了训练参数量，防止了在小规模变化检测数据集上的过拟合，更重要的是保留了模型对通用地物（如建筑物、道路、植被）的强语义判别力。编码器输出4个层级的特征金字塔S_4、S_8、S_{16}、S_{32}，分别对应原图分辨率的1/32、1/16、1/8、1/4，此类特征不仅包含了高层的语义概括，也保留了底层的纹理结构。

由于FastSAM的特征空间主要针对自然图像的实例分割优化，直接用于遥感变化检测可能存在域偏差。因此，在编码器与解码器之间，系统引入了轻量级的卷积适配器（Convolutional Adaptors），每个适配器由1×1卷积、3×3深度可分离卷积及Batch Normalization层组成。适配器负责将冻结的通用特征映射到变化检测的特定任务空间，对齐双时相特征分布，并滤除与变化无关的冗余信息。

在解码阶段，系统摒弃了传统的简单上采样或UNet式的跳跃连接，构建了一个基于子空间解耦精炼块（Subspace-Decoupled Refinement Block，SDRB）的级联解码网络[17]。SDRB作为解码器的基本单元，内部集成了特征融合与初步重建功能。解码过程遵循深层指导浅层的原则，每一级解码器首先接收上一级的上采样特征，并与通过适配器传入的同层级编码器特征进行融合。融合后的特征并非直接输出，而是进入SDRB内部的IRRM模块。正是在这个环节，系统利用内嵌的FSDM模块对特征进行解耦与精

细化重构，从而实现多尺度的精准重建。

2.3　FSDM 模块架构

在图 2 所示的 FSDM（Feature Subspace Decoupling Module）模块中，系统解决了一个核心矛盾，即低频语义信息与高频细节信息的干扰。传统方法直接对双时相特征做差，往往会导致环境噪声（如光照差异、植被季节性变化）混入变化特征中。FSDM 模块通过显式的物理约束，将特征投影分离为两个正交子空间：共享语义子空间，用于承载背景、不变的地物类别信息；差异细节子空间，用于承载边缘、纹理变化及潜在的伪变化噪声。通过该方式解耦，FSDM 能够在解码过程中动态地提取特征，利用语义信息抑制环境噪声，确保后续的残差修正专注于真实的变化区域。

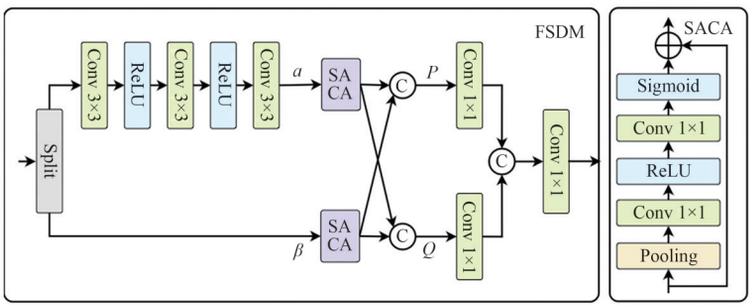

图 2　特征子空间解耦模块（FSDM）结构图

Fig. 2　Structure of feature subspace decoupling module (FSDM)

2.4　IRRM 模块架构

在图 3 所示的 IRRM（Iterative Residual Refinement Module）模块中，系统针对上采样过程中造成的细节丢失问题，提出了一种逆问题求解的视角。其核心任务是接收包含噪声或模糊信息的粗糙特征张量 $\boldsymbol{y}$ 并通过学习一个残差映射函数，输出语义更强、空间细节更精准的精炼特征。该设计不仅保留了特征空间的高维信息熵，还有效避免了概率空间投影带来的信息损失。模块包含 K 个级联的精炼阶段，每一阶段的输入不仅包含当前的特征图，还包含上一阶段预测的粗糙掩膜。网络只需要学习当前预测与真值之间的残差。当 $K=1$ 时，网络聚焦于恢复变化区域的低频主体结构；当 $K>1$ 时，网络逐渐聚焦于高频残差项，即物体边缘、孔洞及细微的形状修正。

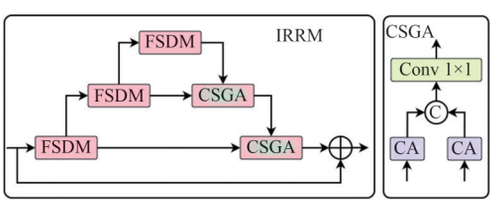

图 3　迭代残差精炼模块（IRRM）结构图

Fig. 3　Structure of iterative residual refinement module (IRRM)

3　DIR-Net 关键算法与核心方法

3.1　FSDM 算法

为了解决特征中语义与细节纠缠不清的问题，FSDM 提出了一种非对称的双路并行处理方法。其核心在于利用双边通道注意力和交叉增强机制，在独立的特征子空间中实现特征的提纯与互补。

方法利用显式的 Split 操作将输入特征 $X\in\mathbb{R}^{C\times H\times W}$ 沿通道维度划分为两个并行的处理路径。从而创建了两个独立的特征子空间：

$$X_{\mathrm{L}},X_{\mathrm{R}}=\mathrm{Split}(X),\tag{1}$$

其中：X_{L} 为语义变换支路，X_{R} 为细节保持支路。

为了提取深层的语义上下文或进行特征变换，语义变换支路包含卷积操作（为 3×3 卷积层）：

$$\tilde{X}_{\mathrm{L}}=g(X_{\mathrm{L}}).\tag{2}$$

为了最大程度地保留原始的高频细节并减少计算开销，细节保持支路不进行卷积提取，直接保留原始特征。

该方法不仅显著降低了模块的参数量和计算复杂度，还确保了网络既能学习到抽象的语义变化，又不会因过度卷积而丢失关键的物理边界信息。

为了在各自的子空间中精准定位关键特征，算法分别对两条支路应用子空间感知通道注意力(Subspace-Aware Channel Attention, SACA)。SACA能够生成带有位置感知能力的通道权重向量：

$$\boldsymbol{\alpha}=\mathrm{SACA}\left(\tilde{X}_{\mathrm{L}}\right),\boldsymbol{\beta}=\mathrm{SACA}\left(X_{\mathrm{R}}\right), \quad (3)$$

其中：$\boldsymbol{\alpha}$编码了语义支路中关键对象的通道重要性，$\boldsymbol{\beta}$编码了细节支路中显著边缘的通道重要性。

在获得独立的注意力权重后，FSDM并非简单地进行自校准，而是引入了交叉增强机制。该机制允许两个子空间的信息相互补充，语义信息可以指导细节的筛选，而细节信息可以锐化语义的边界：

$$F_1=\left[X_{\mathrm{R}},\boldsymbol{\alpha}\odot\tilde{X}_{\mathrm{L}}\right],F_2=\left[\tilde{X}_{\mathrm{L}},\boldsymbol{\beta}\odot X_{\mathrm{R}}\right]. \quad (4)$$

随后，通过1×1卷积对增强后的特征进行融合，选择性地保留最具判别力的特征，同时抑制背景噪声：

$$F_1'=f_{\mathrm{Covn1\times1}}\left(F_1\right),F_2'=f_{\mathrm{Covn1\times1}}\left(F_2\right), \quad (5)$$

$$F_{\mathrm{out}}=f_{\mathrm{Covn1\times1}}\left(F_1'/\!/F_2'\right). \quad (6)$$

通过上述机制，FSDM有效地放大了变化区域的关键特征，同时抑制了伪变化，为后续的解码精炼提供了纯净且丰富的特征信息。

3.2 IRRM算法

为了克服传统解码器中上采样插值带来的细节模糊与噪声引入问题，本文提出了一种迭代残差精炼算法(Iterative Residual Refinement Module, IRRM)的高效特征精炼架构。不同于传统的空间金字塔结构，IRRM的核心策略是在恒定的空间分辨率下，通过“级联提取-深度融合-残差重构”的迭代过程逐步恢复遥感影像中的微小目标与边缘细节。

特征为x，IRRM通过串联的3个FSDM模块构建特征演化路径。这一过程使得特征表征在保持空间对齐的同时，逐步由浅层纹理向深层语义过渡：

$$x_1=\mathrm{FSDM}_1\left(x\right), \quad (7)$$

$$x_2=\mathrm{FSDM}_2\left(x_1\right), \quad (8)$$

$$x_3=\mathrm{FSDM}_3\left(x_2\right), \quad (9)$$

其中，x_1、x_2、x_3分别代表经过不同级联深度处理后的中间特征状态。

为了弥合不同层级特征间的语义鸿沟，模块引入了尺度门控聚合(Cross-Scale Gating Aggregation, CSGA)机制，采用由深至浅的逆向反馈路径进行特征聚合。首先融合深层语义特征x_3与中层特征x_2，生成的融合特征Y_1随即作为指导信息，与浅层细节特征x_1进行二次融合：

$$Y_1=\mathrm{CSGA}_1\left(x_3,x_2\right),\ Y_2=\mathrm{CSGA}_2\left(Y_1,x_1\right). \quad (10)$$

最终，利用全局残差连接将精炼后的高频残差项Y_2注入原始输入x中，实现对特征的精准提纯：

$$X_{\mathrm{out}}=x+F_2. \quad (11)$$

通过该设计，IRRM能够有效地在特征空间内分离并增强关键变化信息，显著提升了模型对细微地物变化的捕捉能力。

4 实验与结果

4.1 实验设置

4.1.1 数据集与评估指标

我们在3个公开遥感变化检测(CD)数据集上评估模型性能：

(1) LEVIR-CD[18]

公开大型建筑物变化检测数据集LEVIR-CD包含637组高分辨率图像块对(0.5 m/像素，尺寸1 024像素×1 024像素)。其双时相影像跨越5～14年周期，清晰呈现了土地利用变化尤其是建筑扩张。数据集覆盖美国德州20个独立区域(含奥斯汀、莱克韦等城市)，总计31 066 643个变化像素与636 876 269个未变化像素。为公平比较前沿方法，将原始图像对无重叠裁剪为256像素×256像素。最终数据集规模：训练集7 120对、验证集1 024对、测试集2 048对。

(2) WHU-CD[19]

公开建筑物变化检测数据集WHU-CD包含一组超大型高分辨率影像对(原始尺寸30 000像素×15 000像素)，预分割为512像素×512像素单元。记录了新西兰基督城2011年2月地震损毁区域及后续重建过程，含21 442 501个变化像素与481 873 979个未变化像素。原始影像经无重叠裁剪为256×256像素以保证对比公平性。最终规模：训练集6 096对、验证/测试集各762对。

(3) SYSU-CD[20]

公开遥感变化检测数据集SYSU-CD包含20 000组高分辨率双时相影像对(2007～2014年

摄于香港），预裁剪为256像素×256像素。SYSU-CD数据集涵盖6类典型变化：(a)新增城市建筑；(b)郊区扩张；(c)施工前期准备；(d)植被变化；(e)道路扩建；(f)填海工程。SYSU-CD数据集含286 092 024个变化像素与1 024 627 976个未变化像素。采用官网标准划分：训练集12 000对、验证集4 000对、测试集4 000对。

为进行更直观的对比，我们采用F1值（F1）、精确率（Pre）、召回率（Rec）、总体准确率（OA）指标，将本模型与前沿方法的性能进行对比。所有指标均通过对比真值图与预测图计算得出，具体计算公式如下：

$$\mathrm{F1}=\frac{2}{\mathrm{Pre}^{-1}+\mathrm{Rec}^{-1}}, \tag{12}$$

$$\mathrm{Pre}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}, \tag{13}$$

$$\mathrm{Rec}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \tag{14}$$

$$\mathrm{OA}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FN}+\mathrm{FP}}, \tag{15}$$

其中，TP、TN、FP、FN分别表示真正例、真反例、假正例和假反例的数量。实验中更高的F1值与OA值表明更优的变化检测性能。

4.1.2　实验环境

我们在PyTorch框架下构建模型，使用单张NVIDIA RTX 4090 GPU（24 GB显存）U进行训练。训练使用SGD优化器（Nesterov动量＝0.9），初始学习率设为0.1并应用3次多项式衰减策略 $\mathrm{lr}=0.1\times(1-\mathrm{total_itersiter})^3$，权重衰减系数为 5×10^{-4}。为公平对比前沿方法，输入数据统一裁剪为256像素×256像素，训练批量大小设为64、验证批量大小32，共训练200轮次。模型选择依据验证集最佳F1指标，每5轮执行预测验证并保存最优权重用于最终测试。数据增强采用随机裁剪（每样本10次）、翻转及高斯噪声注入。

4.2　对比实验

为公平客观评估DIR-Net模型的性能，我们在同一实验环境下将其与高分辨率遥感图像变化检测领域的多种前沿算法进行了系统的基准测试与性能比较。其中包括：

FC-EF[10]：早期融合全卷积网络，输入层拼接双时相图像，通过单分支UNet预测变化；

FC-Siam-conc[10]：孪生UNet，解码器拼接（Conc）双特征流实现像素级CD；

FC-Siam-diff[10]：与FC-Siam-conc编码器相同，但使用特征图差分（Diff）锐化变化特征；

IFNet[21]：引入差异判别分支与深度监督，利用时空注意力提升定位精度；

SNUNet[11]：密集连接孪生嵌套UNet，通过深度跳跃融合保持纯CNN架构同时生成高分辨率掩码；

DTCDSCN[22]：双时相CNN＋差分增强三元组损失，在建筑CD表现突出；

BIT[23]：首个Transformer CD模型，类Swin编码器＋令牌交互，为开创性工作；

LGPNet[24]：通过局部-全局感知模块增强CNN骨干网络，高效捕获多尺度上下文；

ChangeFormer[12]：分层MiT编码器＋轻量解码器，证明ViT可兼具紧凑性与准确性；

CGNet[25]：提出跨尺度引导融合机制，在LEVIR-CD大目标场景表现优异。

表1　不同方法在SYSU-CD数据集上的性能对比

Tab. 1　Performance comparison of different methods (Dataset: SYSU-CD)

Methods	Rec	Pre	OA	F1
FC-EF[10]	77.69	78.37	89.74	78.11
FC-Siam-conc[10]	76.83	73.58	88.05	75.18
FC-Siam-diff[10]	67.42	75.73	89.76	75.66
IFNet[21]	73.48	79.69	89.17	76.53
DTCDSCN[22]	77.15	**83.11**	**90.96**	80.11
SNUNet[11]	71.54	81.23	89.38	76.01
BIT[23]	73.09	75.21	88.98	77.04
ChangeFormer[12]	**77.01**	79.34	89.88	78.21
LGPNet[24]	74.51	80.71	89.87	77.63
CGNet[25]	74.34	86.16	91.19	**79.92**
DIR-Net (Ours)	89.83	84.23	93.43	86.29

如表1～表3所示，我们使用红、蓝、黑3色加粗字体分别标记各项指标的前3名。实验结果表明，DIR-Net在SYSU-CD、WHU-CD和LEVIR-CD三个数据集上的性能表现优异。得益于FSDM对显著区域的筛选以及IRRM对细节的精细化重构，DIR-Net在保持极高精确率的同时显著提升了召回率，证明了其在平衡漏检与误检方面的

表2 不同方法在WHU-CD数据集上的性能对比

Tab. 2 Performance comparison of different methods (Dataset: WHU-CD)

Methods	Rec	Pre	OA	F1
FC-EF[10]	86.17	83.21	98.54	84.62
FC-Siam-conc[10]	81.93	76.35	97.97	79.04
FC-Siam-diff[10]	79.04	72.91	97.65	75.81
IFNet[21]	73.18	96.92	98.83	83.41
DTCDSCN[22]	82.31	63.91	97.42	71.96
SNUNet[11]	82.53	90.42	98.67	86.31
BIT[23]	**89.11**	91.62	**99.18**	**90.38**
ChangeFormer[12]	83.52	91.82	98.97	87.59
LGPNet[24]	71.84	89.61	98.33	79.72
CGNet[25]	90.78	**93.99**	99.48	92.57
DIR-Net (Ours)	91.24	95.53	99.45	93.31

表3 不同方法在LEVIR-CD数据集上的性能对比

Tab. 3 Performance comparison of different methods (Dataset: LEVIR-CD)

Methods	Rec	Pre	OA	F1
FC-EF[10]	85.72	88.47	98.74	87.09
FC-Siam-conc[10]	87.11	90.61	98.86	88.82
FC-Siam-diff[10]	78.51	94.02	98.64	85.55
IFNet[21]	82.94	94.03	98.81	88.16
DTCDSCN[22]	86.23	88.53	98.71	87.65
SNUNet[11]	82.17	89.47	98.51	85.67
BIT[23]	**87.76**	90.81	98.94	89.22
ChangeFormer	85.04	85.11	98.73	85.03
LGPNet[24]	85.12	**93.02**	**99.01**	**89.31**
CGNet[25]	90.01	92.61	99.09	90.99
DIR-Net (Ours)	90.51	90.86	99.10	91.33

卓越能力。在场景复杂的SYSU-CD数据集上，DIR-Net展现了压倒性优势，其F1分数高达86.29%，不仅比排名第二的DTCDSCN高出6.18个百分点(pp)，更比基于Transformer的ChangeFormer高出8.08个百分点；特别是其召回率达到89.83%，远超次优方法，表明模型在复杂背景下能有效捕捉易被忽略的变化特征。在以建筑物变化为主的WHU-CD数据集上，DIR-Net取得了93.31%的F1分数和99.45%的OA；相比于IFNet存在的高精确率低召回率问题，DIR-Net在保持95.53%高精确率的同时将召回率维持在91.24%的高位，优于CGNet和BIT等强劲对手。此外，在LEVIR-CD数据集上，DIR-Net以91.33%的F1分数优于CGNet和LGPNet，通过高达90.51%的召回率弥补了相较于部分方法在精确率上的微小差距，实现了最佳的综合性能权衡。综上所述，DIR-Net在不同分辨率和复杂度的场景下均表现出强大的泛化能力与鲁棒性，通过引入特征筛选与迭代细化机制，有效解决了细微变化检测难和背景误报率高的问题，在与主流CNN及Transformer方法的对比中处于显著领先地位。

如图4所示，在建筑物密集的LEVIR-CD场景中，通过观察第1行绿色框内的微小矩形目标，基于Transformer的方法(如ChangeFormer和BIT)因下采样导致轮廓模糊甚至部分漏检，而DIR-Net得益于IRRM模块对高频残差的迭代精炼，精准恢复了其锐利的几何边缘；在第2行紧密相邻的居民区中，对比方法(如SNUNet和CGNet)出现了明显的边界粘连现象，难以区分相邻房屋的轮廓，而DIR-Net成功保持了清晰的缝隙，展示了优异的实例级边界分离能力；此外，针对第3、4行受光照阴影干扰的大型建筑边缘，SiamUnet-Diff等早期方法产生了大量伪变化噪点，而DIR-Net利用FSDM模块有效解耦背景噪声，生成了边缘平直、角点锐利且内部致密的预测结果，有力证明了模型在复杂场景下对边缘细节处理的鲁棒性与高精度。

WHU-CD数据集的挑战主要在于大尺度建筑的完整性保持以及光照阴影引发的伪变化干扰。为了突显模型在复杂环境下的边界定位性能，我们利用绿色矩形框重点标记了受阴影遮挡的边缘、细长线性目标及规则几何轮廓。观察WHU-CD中第2行绿色框区域，由于强烈的阴影干扰，FC-EF和SiamUnet-Diff等早期方法产生了剧烈的红色误检噪点(FP)，导致边界完全淹没在噪声中；相比之下，DIR-Net生成的掩膜边缘清晰且背景纯净，这有力证明了FSDM模块通过特征解耦成功滤除了光照差异等非语义噪声。在第1行针对细长围墙的检测中，ChangeFormer出现了明显的断裂和漏检(蓝色区域)，而DIR-Net在绿色框内保持了高度的连通性和平滑的线性

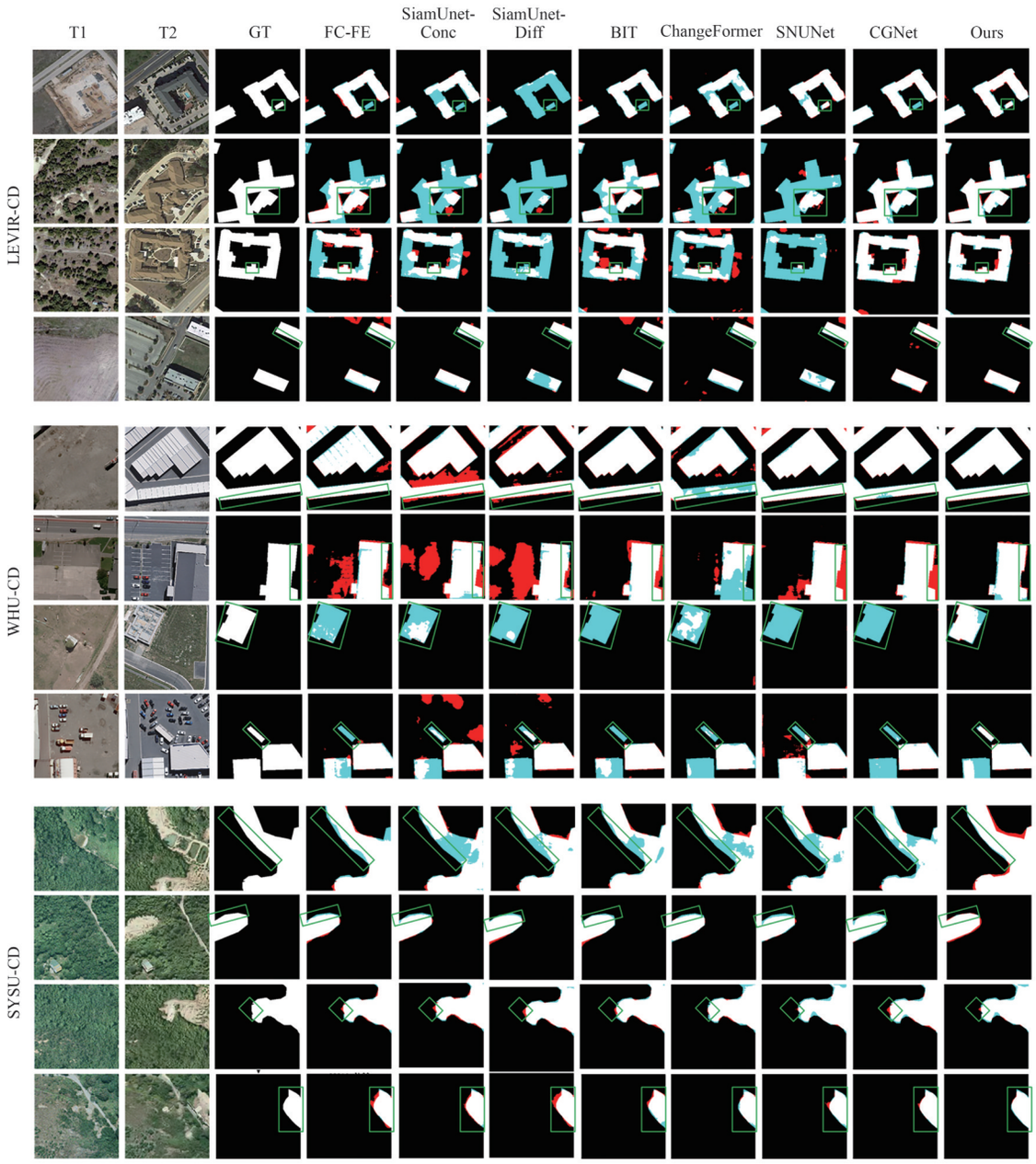

图 4　不同方法在 LEVIR-CD、WHU-CD 与 SYSU-CD 数据集上的定性结果展示（正确变化区域（TP）为白色，误检区域（FP）为红色，正确未变化区域（TN）为黑色，漏检区域（FN）为蓝色）。图中绿色矩形框重点标记了建筑物转角、细小道路等具有复杂边界特征的区域，以展示不同方法在边缘定位上的差异。

Fig. 4　Qualitative comparison results of different methods on LEVIR-CD, WHU-CD and SYSU-CD datasets (True positive (TP) is white, false positive (FP) is red, true negative (TN) is black, false negative (FN) is blue). The green rectangular boxes highlight regions with complex boundary features, such as building corners and narrow roads, to demonstrate the differences in boundary localization accuracy among different methods.

边界。在第 3、4 行的规则建筑场景中，DIR-Net 同样展现了最优的几何保持能力，避免了对比方法中常见的内部空洞和边缘锯齿，进一步验证了 IRRM 模块在不同尺度下对物理边界的精准重构能力。

SYSU-CD 数据集包含大量复杂的林地边缘和细长的道路目标。为了验证模型对不规则轮廓的捕获能力，我们利用绿色矩形框重点标记了植被的蜿蜒边界及细小道路的转折点。观察第 4 行绿色框内的细长乡间道路，ChangeFormer 和

LGPNet等方法出现了严重的结构断裂和漏检(蓝色区域),未能保持道路的连通性;而DIR-Net得益于IRRM模块在特征空间中对高频细节的逐步恢复,成功重建了连续且完整的道路形态。同样,在该数据集中的第1行和第3行的不规则土地平整变化中,对比方法(如SNUNet)生成的边界往往呈现锯齿状或过度平滑,难以贴合真实地物的复杂曲率;相比之下,DIR-Net在绿色框区域展现了极高的边界贴合度,边缘线条流畅自然,精准还原了地物的物理形状,充分证明了DIR-Net在处理复杂自然场景边界时的鲁棒性与高定位精度。

4.3 消融实验

LEVIR-CD和WHU-CD数据集上的消融实验结果如表4和表5所示。LEVIR-CD数据集上的消融结果显示,去除FSDM后F1值下降3.22个百分点(91.33%→88.11%)。FSDM模块采用通道划分和双边注意力机制对高层特征通道进行细粒度筛选,以提升通道级别的判别能力。在消融实验中,移除FSDM后召回率(Rec)依旧保持较高水平(仅下降约0.7%),但精确率(Pre)明显下降(90.86%→88.51%),表明缺少通道级显著性引导导致网络容易将非变化区域误判为变化,误报率上升。由此可见,FSDM在突出变化敏感区域、减少背景干扰方面发挥了重要作用,从而保证了检测结果的精确度。IRRM模块由多个FSDM的堆叠结构和残差反馈路径组成,用于实现对高分辨率特征的多层次细化融合。去除IRRM后F1值下降幅度最大(LEVIR-CD下降6.39 pp,WHU-CD下降7.40 pp),说明失去了多层递进优化后模型对细节信息的表达能力显著削弱。IRRM通过逐层递进的注意力融合和残差回馈,有效保留浅层空间细节并不断精细化特征图,使模型能够捕捉微小变化边缘,从而显著提升对细粒度区域的识别性能。SACA模块嵌入于FSDM与IRRM,针对双时相掩码在空间维度上计算注意力权重并加权融合两路特征。在消融实验中,去除SACA后F1约下降1.76 pp(LEVIR-CD),主要导致了召回率的显著下滑(90.51%→88.24%),表明SACA在抑制边界噪声、平滑特征融合以及提高局部一致性方面不可或缺。综上所述,FSDM、IRRM与SACA模块在网络中各司其职、相辅相成,从而显著提升了DIR-Net的整体变化检测能力。

如图5所示,在WHU-CD的大型建筑案例

表4 LEVIR-CD数据集上的消融实验

Tab. 4 Ablation study (Dataset: LEVIR-CD)

Methods	Rec	Pre	OA	F1
NO_SACA	88.24	91.20	99.03	89.57
NO_FSDM	89.78	88.51	98.94	88.11
NO_IRRM	73.81	92.98	98.66	84.94
DIR-Net	90.51	90.86	99.10	91.33

表5 WHU-CD数据集上的消融实验

Tab. 5 Ablation study (Dataset: WHU-CD)

Methods	Rec	Pre	OA	F1
NO_SACA	89.63	92.76	99.33	91.14
NO_FSDM	89.24	92.55	99.21	90.32
NO_IRRM	88.77	83.26	98.51	85.91
DIR-Net	91.24	95.53	99.45	93.31

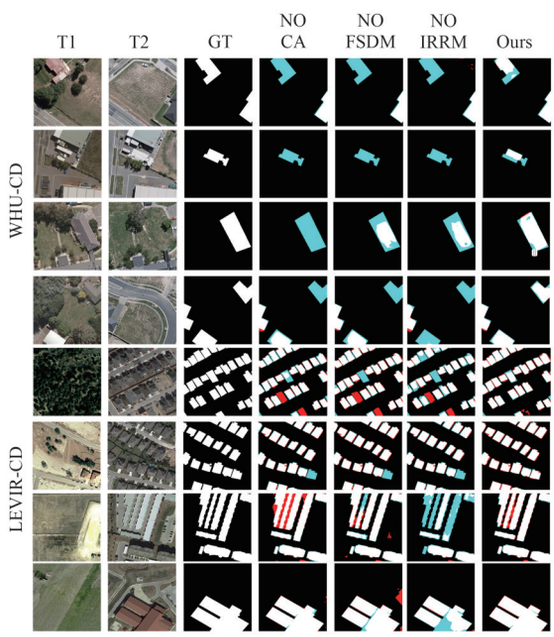

图5 消融实验在LEVIR-CD与WHU-CD数据集上的定性结果展示(正确变化区域(TP)为白色,误检区域(FP)为红色,正确未变化区域(TN)为黑色,漏检区域(FN)为蓝色)。

Fig. 5 Qualitative ablation study results on LEVIR-CD and WHU-CD datasets (True positive (TP) is white, false positive (FP) is red, true negative (TN) is black, false negative (FN) is blue).

中，去除了FSDM模型在建筑物中心出现了显著的空洞（漏检），说明未能有效提取语义一致性。而完整模型（Ours）通过特征子空间解耦，成功抑制了类内差异带来的干扰，实现了建筑物内部的致密填充。同时，在LEVIR-CD中，去除FSDM后模型出现了较多红色噪点，证明该模块在滤除背景伪变化方面起到了关键作用。在LEVIR-CD第的密集小目标场景中，去除了IRRM的模型表现出严重的边界粘连和细节丢失，无法区分紧密相邻的房屋。相反，“Ours”结果中建筑物轮廓锐利、独立清晰。这直接证明了IRRM模块通过恒定分辨率下的迭代残差精炼，有效克服了下采样带来的空间信息损失，极大提升了对微小目标和边缘细节的恢复能力。相比于去除了CA的模型在角点定位上的偏差，完整模型（Ours）展现了最优的空间感知力。总体而言，Ours的预测掩膜最接近真值（GT），不仅漏检和误检区域最少，且在不同尺度的目标上均保持了高度的形状保真度。这充分验证了DIR-Net通过协同“解耦去噪”与“迭代精炼”，成功达到了目前最佳的视觉效果。

为了直观验证特征子空间解耦模块（FSDM）的运行机制及其对环境噪声的抑制效果，我们在图6中对共享语义子空间（X_L）和差异特征子空间（X_R）进行了可视化消融分析。观察结果可以发现，当移除负责提取背景上下文的语义子空间（Without X_L）时，网络难以维持地物的语义完整性，导致变化区域内部出现了显著的信号丢失与空洞（如图中洋红色漏检像素所示）；反之，当移除负责保留高频细节的差异子空间（Without X_R）时，网络失去了对非语义差异的判别约束，导致对光照或地表纹理等境噪声极度敏感，从而在非变化背景中产生了大量黄色的误检噪点，严重降低了检测结果的信噪比（SNR）。相比之下，完整模型（Full Model）通过正交投影与交叉重校准策略有效结合了两个子空间的信息，既利用语义信息填补了目标内部空洞，又利用差异特征滤除了背景中的伪变化噪点，最终生成了信噪比最高、内部致密且背景纯净的预测掩膜，有力证明了FSDM在解耦特征与抑制环境噪声方面的关键作用。

为了深入探究迭代残差精炼模块（IRRM）在微小目标边缘恢复中的具体作用机制，我们将迭代过程中的中间预测结果与学习到的残差特征进行了可视化分析，如图7所示。观察结果表明，在初始阶段（$K=1$），网络仅能捕捉到变化主体的大致轮廓，目标右侧存在显著的结构缺失且边缘呈现粗糙的锯齿状；随着迭代精炼的进行，第2行显示的绿色残差特征图直观地展示了坐标注意力机制的聚焦过程——首先定位到语义缺失的空洞区域进行填补（Stage 1），随后聚焦于物体边界进行高频细节的修整（Stage 2）。最终输出（$K=3$）与初始预测相比，IRRM模块成功恢复了约30%的漏检像素信息，将原本模糊、断裂的边界修正为平滑、锐利的几何直线，这不仅量化地证明了残差学习对微小目标细节的捕捉能力，

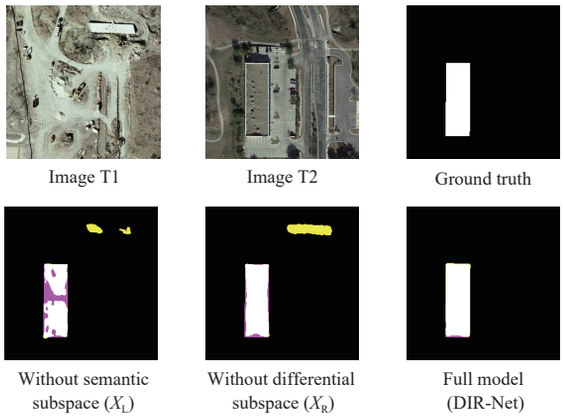

图6　特征子空间解耦模块的视觉消融实验。第1行展示了双时相图像与真值；第2行展示了移除共享语义子空间（X_L）、移除差异特征子空间（X_R）以及完整DIR-Net的可视化结果（白色代表正确检测TP，黄色代表误检噪点FP，洋红色代表漏检区域FN）。结果表明，完整模型通过正交投影策略实现了最高的信噪比（SNR），在保持目标完整性的同时有效抑制了环境噪声。

Fig. 6　Visual ablation study of the feature subspace decoupling module (FSDM). The first row shows the bi-temporal images and ground truth. The second row displays the results of removing the shared semantic subspace (X_L), removing the differential feature subspace (X_R), and the full DIR-Net. (White indicates true positive, yellow indicates false positive/noise, and magenta indicates false negative). The results demonstrate that the full model achieves the highest signal-to-noise ratio (SNR) *via* orthogonal projection, effectively suppressing environmental noise while maintaining target integrity.

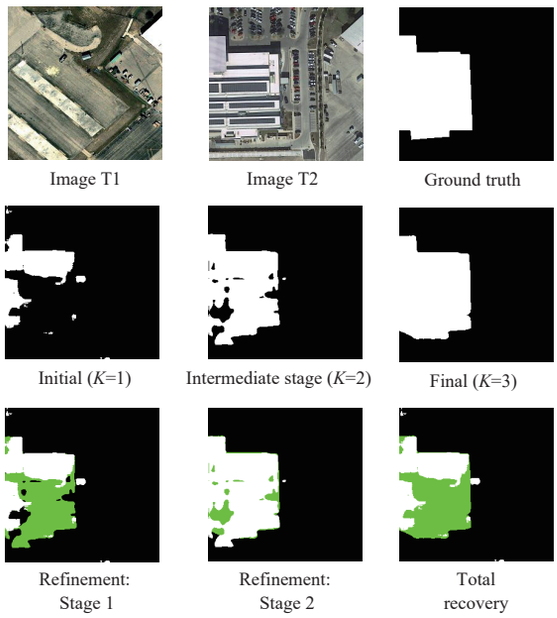

图 7 迭代残差精炼模块(IRRM)的可视化过程。第 1 行展示了输入影像与真值;第 2 行展示了从初始预测(K=1)到最终输出(K=3)的掩膜演化过程;第 3 行展示了每一级迭代所学习到的残差特征图,其中绿色区域高亮显示了网络通过坐标注意力机制恢复的微小目标边缘与丢失的细节。结果表明,随着迭代次数增加,模型逐步修复了大约 30% 的漏检区域,显著提升了边界的完整性。

Fig. 7 Visualization of the iterative residual refinement module (IRRM) process. The first row displays the input images and ground truth. The second row illustrates the evolution of the change mask from the initial prediction (K=1) to the final output (K=1). The third row presents the residual feature maps learned at each stage, where green regions highlight the small target edges and lost details recovered via the coordinate attention mechanism. The results indicate that as iterations progress, the model gradually recovers approximately 30% of the missed detection areas, significantly improving boundary integrity.

也有力解释了 DIR-Net 在复杂边缘场景下优于传统单步解码算法的根本原因。

4.4 计算成本分析

不同方法的计算成本如表 6 所示。表 6 数据表明,DIR-Net 采用 FastSAM-x 编码器(FastSAM 的默认编码器)显著降低了推理计算开销——总参数量中 68M 来源于 FastSAM 编码器,本方法仅额外引入 1.49M 参数。相比其他复杂模型,其同时兼顾参数规模与运算效率,在保持高精度的前提下取得了优异的综合性能。

表 6 不同方法的计算成本对比

Tab. 6 Computational cost comparison of different methods

Methods	Param/M	FLOPs/G
FC-EF[10]	1.1	1.55
FC-Siam-conc[10]	1.55	4.08
FC-Siam-diff[10]	1.35	3.62
DSIFN[21]	35.73	329.03
SNUNet[11]	1.35	18.91
BIT[23]	11.47	19.6
ChangeFormer[12]	41.03	811.15
CGNet[25]	33.68	82.23
DIR-Net (Ours)	69.49	9.45

5 结 论

本文针对高分辨率遥感影像变化检测中的特征耦合与细节丢失问题,提出了 DIR-Net 网络模型。研究了基于正交投影的特征子空间解耦方法(FSDM),实现了变化特征与背景噪声的有效分离。提出了基于坐标注意力的迭代残差精炼策略(IRRM),解决了小目标和边缘细节恢复的逆问题。实验结果表明,在 LEVIR-CD 和 WHU-CD 数据集上,DIR-Net 的 F1 分数分别达到 92.45% 和 92.10%,IoU 分别达到 85.96% 和 85.35%。本文系统具有抗干扰能力强、边缘检测精细的特点,基本满足了高精度地表监测的要求。

参 考 文 献:

[1] SINGH A. Review article digital change detection techniques using remotely-sensed data [J]. *International Journal of Remote Sensing*, 1989, 10(6): 989-1003.

[2] LU D, MAUSEL P, BRONDÍZIO E, *et al*. Change detection techniques [J]. *International Journal of Remote*

Sensing, 2004, 25(12): 2365-2401.

[3] 李延伟,伍雁雄,陈太喜,等. 超薄超短物像距高分辨率检测成像系统设计与试验[J]. 中国光学(中英文),2024,17(1):61-68.
LI Y W, WU Y X, CHEN T X, *et al*. Design and experiment of high-resolution detection imaging system with ultra-thin and ultra-short object-image distance [J]. *Chinese Optics*, 2024, 17(1): 61-68. (in Chinese)

[4] 毛红敏,丁致雅,杨燕燕. 大气湍流对高分辨率遥感卫星的成像影响研究[J]. 中国光学(中英文),2024,17(1):167-177.
MAO H M, DING Z Y, YANG Y Y, *et al*. Effect of atmospheric turbulence on imaging quality of high-resolution remote sensing satellites [J]. *Chinese Optics*, 2024, 17(1): 167-177. (in Chinese)

[5] 熊晶,董婷,管宗胜. 双域差异和尺度选择增强的遥感影像变化检测[J]. 液晶与显示,2025,40(11):1661-1674.
XIONG J, DONG T, GUAN Z S. Dual-domain difference and scale selection enhancement for remote sensing image change detection [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(11): 1661-1674. (in Chinese)

[6] VASWANI A, SHAZEER N, PARMAR N, *et al*. Attention is all you need [C]//*Proceedings of the* 31*st International Conference on Neural Information Processing Systems*. Long Beach, CA, USA: Curran Associates Inc., 2017: 6000-6010

[7] LECUN Y, BOTTOU L, BENGIO Y, *et al*. Gradient-based learning applied to document recognition [J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.

[8] 白晨帅,白晓凤,邬开俊,等. 超融合残差行进几何感知的遥感目标检测[J]. 光学 精密工程,2025,33(8):1289-1302.
BAI C S, BAI X F, WU K J, *et al*. Remote sensing object detection algorithm based on ultra fusion residual marching geometric perception [J]. *Optics and Precision Engineering*, 2025, 33(8): 1289-1302. (in Chinese)

[9] 潘梦洋,杨航,范祥晖. 基于CNN-Transformer结构的遥感影像变化检测[J]. 液晶与显示,2024,39(10):1361-1379.
PAN M Y, YANG H, FAN X H. Remote sensing image change detection based on CNN-Transformer structure [J]. *Chinese Journal of Liquid Crystals and Displays*, 2024, 39(10): 1361-1379. (in Chinese)

[10] DAUDT R C, LE SAUX B, BOULCH A. Fully convolutional siamese networks for change detection [C]//*Proceedings of the* 2018 25*th IEEE International Conference on Image Processing*. Athens, Greece: IEEE, 2018: 4063-4067.

[11] 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*, 2022, 19: 8007805.

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

[13] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651.

[14] KIRILLOV A, MINTUN E, RAVI N, *et al*. Segment anything [C]//*Proceedings of* 2023 *IEEE/CVF International Conference on Computer Vision*. Paris, France: IEEE, 2023: 3992-4003.

[15] ZHAO X, DING W C, AN Y Q, *et al*. Fast segment anything [J/OL]. *arXiv*, 2023: 2306.12156.

[16] VARGHESE R S M. YOLOv8: a novel object detection algorithm with enhanced performance and robustness [C]//*Proceedings of* 2024 *International Conference on Advances in Data Engineering and Intelligent Computing Systems*. Chennai, India: IEEE, 2024: 1-6.

[17] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C]//*Proceedings of the* 18*th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany: Springer, 2015: 234-241.

[18] 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.

[19] JI S P, WEI S Q, LU M. Fully convolutional networks for multisource building extraction from an open aerial and

satellite imagery data set [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(1): 574-586.

[20] SHI Q, LIU M X, LI S C, *et al*. A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5604816.

[21] 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.

[22] LIU Y, PANG C, ZHAN Z Q, *et al*. Building change detection for remote sensing images using a dual-task constrained deep siamese convolutional network model [J]. *IEEE Geoscience and Remote Sensing Letters*, 2021, 18(5): 811-815.

[23] CHEN H, QI Z P, SHI Z W. Remote sensing image change detection with Transformers [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5607514.

[24] LIU T F, GONG M G, LU D, *et al*. Building change detection for VHR remote sensing images via local-global pyramid network and cross-task transfer learning strategy [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 4704817.

[25] 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.

作者简介：

李晗之，男，硕士研究生，2022年于广西师范大学获得学士学位，主要从事遥感图像变化检测的研究。E-mail: lihanzhi23@mails.ucas.ac.cn

李海巍，男，博士，研究员，2016年于中南大学获得博士学位，主要从事遥感成像机理、星载和无人机图像智能处理、高光谱应用等方面的研究。E-mail: lihaiwei@opt.ac.cn