引文格式：张常旺,王雅萌,段振龙,等. 多尺度差异特征融合的高分辨率光学影像建筑变化检测[J]. 地理空间信息, 2026, 24(7): 1-5. doi:10.3969/j.issn. 1672-4623.2026.07.001

多尺度差异特征融合的高分辨率光学影像建筑变化检测

张常旺[1,2,3]，王雅萌[1,2,3]，段振龙[1,2,3]，彭壮壮[1,2,3]，崔维康[1,2,3]，张　柯[1,2,3]
（1. 河南省地质局矿产资源勘查中心，河南 郑州 450006；2. 河南省天空地遥感智能监测工程技术研究中心，河南 郑州 450006；
3. 河南省自然资源科技创新中心（天空地遥感智能监测研究），河南 郑州 450006）

摘　要：针对当前遥感建筑变化检测模型在提取目标细节与全局特征时，未充分利用各层级特征差异信息而导致的变化检测精度不理想问题，提出了一种融合多尺度全局与细节差异特征的遥感建筑变化检测模型。利用C2f-UIB结构的卷积核构建孪生骨干网络，并在各层间引入差异特征提取单元；采用聚焦线性注意力构建编码器—解码器网络；检测阶段利用密集上采样卷积单元进行特征图上采样，精确恢复图内细节差异特征。在大规模遥感建筑变化检测数据集、武汉大学变化检测数据集上的实验结果表明，所构建模型的F_1-score、总体精度、交并比均优于对照组模型；建筑变化提取结果更完整清晰，与真实掩膜更接近，在精度、泛化性方面均有良好的表现。

关键词：遥感建筑变化检测；孪生骨干网络；差异特征提取；聚焦线性注意力；密集上采样卷积
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2026）07-0001-05

Building Change Detection in High-resolution Optical Images Based on Multi-scale Differential Feature Fusion

ZHANG Changwang[1,2,3], WANG Yameng[1,2,3], DUAN Zhenlong[1,2,3], PENG Zhuangzhuang[1,2,3],
CUI Weikang[1,2,3], ZHANG Ke[1,2,3]
(1. Mineral Resources Exploration Center of Henan Geological Bureau, Zhengzhou 450006, China;
2. Henan Sky-ground Remote Sensing Intelligent Monitoring Engineering Technology Research Center, Zhengzhou 450006, China;
3. Henan Natural Resources Science and Technology Innovation Center (Sky-ground Remote Sensing Intelligent Monitoring Research), Zhengzhou 450006, China)

Abstract: To address the limited exploitation of differential information across hierarchical features in current remote sensing building change detection models leading to sub-optimal accuracy in capturing both fine-grained details and global contexts, we proposed a remote sensing building change detection model fusing multi-scale global and detail differential features. We used C2f-UIB convolutional kernels to build a twin backbone, and inserted differential feature extraction units between every extraction layer. Then, we used focused linear attention to construct an encoder-decoder network. At the detection stage, we used dense upsample convolutional units to upsample the feature maps and precisely recover intra-image detail differential features. Experimental results on the learning, vision and remote sensing laboratory-change detection dataset and Wuhan University change detection dataset show that this model surpasses state-of-the-art counterparts in F_1-score, overall accuracy and inner over union. Across diverse remote sensing scenes, the extracted building change masks are more complete and clearer than those of the compared models, and they align more closely with the ground truth masks, demonstrating excellent accuracy and generalization.

Key words: remote sensing building change detection, twin backbone network, differential feature extraction, focused linear attention, dense upsample convolution

随着航天与传感器技术的飞速发展，遥感影像已步入高分辨率、多时相时代[1-2]。基于多时相高分辨率遥感影像的建筑变化检测技术，能为城市规划、土地利用监测、灾害评估等领域提供有力的技术支持[3-4]。基于大尺度遥感影像准确提取地物变化信息，是当今遥感领域的研究热点[5]。早期遥感影像变化检测方法直接对比两幅影像间的像素、灰度、光谱特征差异或结合人工设计特征与经典机器学习算法对遥感影像开展对象级分割，再比较对象间的差异[6-7]；但受人工设计特征表达能力有限、传统机器学习方法性能较弱等因素影响，上述方法对影像间的辐射差异、相似背景地物等噪声信息更敏感，在大尺度且背景复杂的遥感影像场景中，提取精度与泛化性能表现较弱[8-9]。得益于深度学习模型在非线性特征挖掘方面的出色表现，

项目来源：河南省地质局局管地质科研资助项目（豫地科研[2024]3号）。
第一作者简介：张常旺（1991—），硕士，工程师，研究方向为摄影测量与遥感，E-mail：cwppw2084@163.com。

卷积神经网络（CNN）与视觉转换器（ViT）架构的语义分割模型，现已基本取代传统方法。采用双分支编码器的孪生架构模型[10-12]在双时相遥感影像特征挖掘方面具有独特优势，在遥感变化检测领域得到广泛应用。现有模型中仅以CNN算子构建的编码器缺乏全局上下文特征建模能力，而基于ViT算子构建的编码器则在细节特征的学习能力上有待加强。在特征解码阶段，大多数模型采用双线性插值等方法完成特征图上采样，导致采样所得特征图内缺乏目标细节特征；在差异特征图计算方面，现有模型大多在编码器或解码器末端计算单层差异特征图，未能充分利用多尺度特征图内的目标差异信息。鉴于此，本文在参考双时态图像变换器（BIT）模型架构的基础上，构建了一种融合细节与全局上下文差异特征的遥感建筑变化检测模型。

1 卷积—聚焦线性注意力孪生模型

1.1 模型总体结构

本文模型由孪生骨干网络、基于聚焦线性注意力（FLA）单元搭建的编码器—解码器以及预测输出头组成。孪生骨干网络各分支均由4组双层瓶颈卷积—通用逆向瓶颈（C2f-UIB）结构封装的卷积单元搭建，两组分支在共享权值信息的同时，采用差异特征提取（DFE）单元逐层计算各尺度下的差异特征图。在骨干网络末端，通过位置补丁与线性编码单元对底层差异特征图编码并投影至一维向量；再通过FLA单元构建的编码器—解码器网络，计算全局上下文信息；最后利用带有密集上采样卷积（DUC）单元的预测头，预测双时相遥感影像中的建筑目标变化信息。

1.2 孪生骨干网络

为捕获浅层特征图中的建筑细节特征，基于4组特征提取层搭建孪生骨干网络。提取层采用C2f结构搭建，将输入特征图沿通道切分后，采用级联多尺度UIB单元对一半特征图开展特征提取；另一半特征图则以恒等映射的方式与提取后特征图进行重新拼接。3组UIB单元封装了尺寸为 $i\times i$（$i=1,3,5$）的深度卷积核（DWConv）、1×1 逐点卷积核（PWConv）、批量归一化（BN）单元和高斯误差线性单元（GELU）激活函数，利用多尺度卷积核开展特征提取，以获取不同粒度下的目标细节特征。特征提取单元C2f-UIB结构见图1。

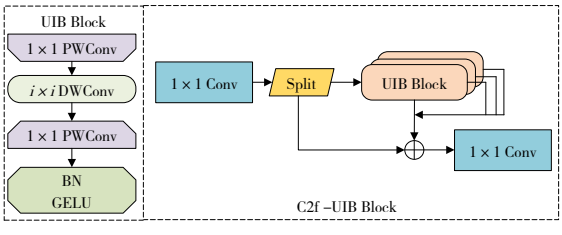


图1 骨干网络特征提取单元详细结构

在孪生骨干网络两组分支之间引入DFE单元，计算各层间双时相差异特征。DFE单元同时接收孪生网络两侧的输出特征图，通过像素级差分操作获取差异特征图 P_d，再将 P_d 与下采样后的差异图 P_{d-1} 进行融合，并通过解耦全连接注意力（DFC）层计算注意力权重，以过滤差异特征图内的负样本信息。DFC层依次采用水平全连接（HFC）、垂直全连接（VFC）沿不同方向计算注意力权重。对于尺寸为 $H\times W$ 的特征图，其上任意像素点对应注意力 a_{hw} 的计算公式为：

$$\begin{cases} a`_{hw}=\sum\limits_{h`=1}^{H}F^{H}_{h,h`w}\odot z_{h`,w} \\ a_{hw}=\mathrm{Sigmoid}\left(\sum\limits_{w`=1}^{W}F^{W}_{w,hw`}\odot a`_{hw}\right) \end{cases} \tag{1}$$

式中，F^H 与 F^W 分别为水平、垂直方向的变换权重，用于对特征图实施全连接操作；Sigmoid为乙型激活函数；$\odot$ 为逐元素相乘操作；$z_{h`,w}$ 为差异特征图上的像素点；$a`_{hw}$ 为沿水平方向的注意力。

采用逐元素相乘的方式，将 a_{hw} 赋权到差异特征图上，DFE单元完整结构见图2。

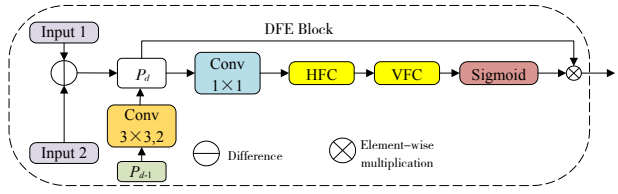


图2 DFE单元结构图

1.3 特征编码器—解码器网络

本文模型的编码器—解码器以FLA单元为核心算子（图3）。FLA单元在线性注意力的基础上引入聚焦映射函数，用于调整查询（$\boldsymbol{Q}$）、键（$\boldsymbol{K}$）特征矩阵的方向，以提高模型对目标上下文的聚焦能力。对于来自骨干网络的差异特征图，首先进行位置补丁、线性编码处理，得到 $\boldsymbol{Q}$、$\boldsymbol{K}$ 和值（$\boldsymbol{V}$）向量；再采用多头线性注意力层计算注意力权重。对于第 i 个计算头，其上注意力的计算过程为：

$$O_i=\phi_p\left(\boldsymbol{Q}_i\right)\phi_p\left(\boldsymbol{K}_i^{\mathrm{T}}\right)\boldsymbol{V}_i+\mathrm{DWC}(\boldsymbol{V}) \tag{2}$$

式中，$\boldsymbol{K}_i^{\mathrm{T}}$ 为键向量的转置；DWC() 为深度可分离卷积；$\phi_p()$ 为对 $\boldsymbol{Q}$ 与 $\boldsymbol{K}$ 编码过程中采用的聚焦映射函数。

$$\phi_p(\boldsymbol{t})=\frac{\left\|\mathrm{ReLU}(\boldsymbol{t})\right\|}{\left\|\mathrm{ReLU}(\boldsymbol{t})^{**p}\right\|}\mathrm{ReLU}(\boldsymbol{t})^{**p} \tag{3}$$

式中，$\boldsymbol{t}$ 为输入向量；ReLU() 为修正线性单元激活函数；$**p$ 为指数为 p 的逐元素幂运算。

1.4 密集连接上采样预测头

在预测阶段，首先通过DUC单元对来自解码器的特征图进行精密尺寸上采样，再利用卷积算子预测建筑变化信息。DUC算子通过可学习的卷积操作，准确

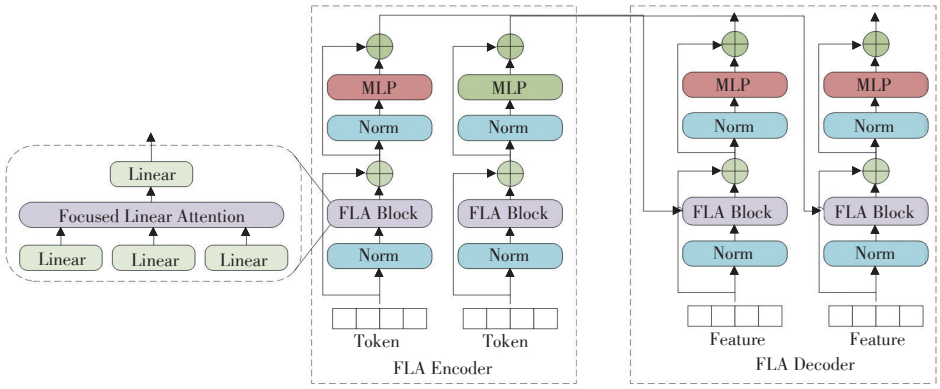


图3　FLA编码器—解码器详细结构图

还原特征图内的细节信息。以尺寸为 $(H/d,W/d,c)$ 的差异特征图为例，DUC算子首先通过尺寸为3×3，操作步长为1的卷积核对输入特征图进行卷积操作，将特征图通道数由 c 调整至 $d^2\times L$，L 为所有像素的类别数；再通过重塑操作将特征图尺寸变换为 (H,W,L)，将不同通道维度上的特征信息展开至平面，以实现小尺寸特征图的精密上采样（图4）。

$$\begin{cases} F^2_{H/d,W/d,d^2\times L}=\mathrm{Conv}_{3\times 3}\left(F^1_{H/d,W/d,c}\right) \\ F^3_{H\times W\times L}=\mathrm{Reshape}\left(F^2_{H/d,W/d,d^2\times L}\right) \end{cases} \tag{4}$$

式中，$F^1_{H/d,W/d,c}$、$F^2_{H/d,W/d,d^2\times L}$、$F^3_{H\times W\times L}$ 分别为输入特征图、卷积所得特征图和重塑所得上采样特征图，本文中 $d=4$；$\mathrm{Conv}_{3\times 3}$ 为尺寸为3×3的标准卷积核。

经过多次迭代训练，DUC算子能从小尺寸特征图中还原精细的建筑细节信息。

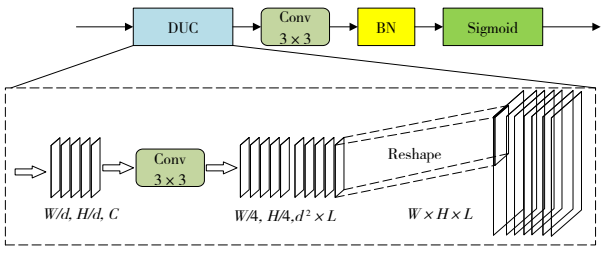


图4　预测头完整结构图

2　实验结果与分析

2.1　实验数据集

为验证本文模型在不同场景下的有效性，在公开的大规模遥感建筑变化检测数据集（LEVIR-CD）和武汉大学变化检测数据集（WHU-CD）上开展模型训练与对比实验。LEVIR-CD数据集包含637对尺寸为1 024像素×1 024像素，空间分辨率为0.5 m的遥感影像。采用重叠率为0，尺寸为512像素×512像素的滑动窗口切分原始影像及其对应的掩膜图像，得到2 548对样本，再按照8∶1∶1的比例随机划分为训练集、测试集、验证集。WHU-CD数据集内包含两幅2012年与2016年的新西兰地区的航空影像，影像尺寸为32 507像素×15 354像素，空间分辨率为0.075 m。采用重叠率为0，尺寸为512像素×512像素的滑动窗口切分两幅原始影像及其对应的掩膜图像，获得1 827张图像后，按照8∶1∶1的比例划分为训练集、测试集、验证集。

2.2　实验环境与模型参数

所有参与实验的模型均采用PyTorch框架搭建，模型的训练与测试工作均在台式主机上完成。实验设备搭载的CPU型号为Intel Xeon(R) Platinum 8352V，搭载显卡型号为Nvidia RTX 4090，运行内存大小为64 GB，硬盘容量为1 TB，操作系统为Ubuntu 22.04。训练过程中以带有权重衰减的自适应矩估计算法为参数优化器，设置动量参数为0.905，模型初始学习率为0.005，训练迭代次数为100次，单批处理样本数量为8。

为客观评估本文模型的性能，采用精确率（P）、召回率（R）、F_1-score、总体精度（OA）、平均交并比（IoU）作为评价指标。

2.3　模型对比实验

本文选择当前主流的遥感变化检测模型配对视频变化检测网络（P2V）[11]、BIT[12]、变化检测孪生嵌套UNet（SNUNet-CD）[13]、Siam-UNet++、变化转换器（ChangeFormer）[14]、移动窗口转换器—变化网络（SwinChangeNet）、时空自适应差异增强网络（STADE）[15]、多尺度特征转换器孪生网络（MFTSNet）作为对照组，与本文模型在相同配置环境下，分别在LEVIR-CD、WHU-CD数据集上开展模型训练与性能测试。对比结果表明（表1、2），在LEVIR-CD数据集上，本文模型的 F_1-score 达到92.56%，较对照组模型分别提高了11.22%、7.81%、7.33%、5.61%、5.11%、3.64%、2.74%、2.43%，且 OA、IoU均较对照组模型有不同程度的提高；在WHU-CD数据集上，本文模型的 F_1-score、OA、IoU分别为93.15%、99.43%、89.67%，其中 F_1-score 较对照组模型分别提高10.68%、7.92%、7.19%、5.83%、5.26%、3.71%、2.91%、1.61%，OA、IoU均优于对照

组模型。

表1　LEVIR-CD数据集上不同模型结果对比/%

模型	*P*	*R*	F_1-score	*OA*	IoU
P2V	79.46	83.31	81.34	97.43	74.28
BIT	82.51	87.20	84.75	97.82	79.89
SNUNet-CD	83.37	87.17	85.23	97.95	81.25
Siam-UNet++	85.64	88.31	86.95	98.04	82.53
ChangeFormer	85.45	89.55	87.45	98.39	83.76
SwinChangeNet	87.61	90.26	88.92	98.53	85.34
STADE	88.92	90.74	89.82	98.78	87.52
MFTSNet	89.74	90.53	90.13	98.95	87.93
本文模型	90.87	94.31	92.56	99.14	88.56

表2　WHU-CD数据集上不同模型结果对比/%

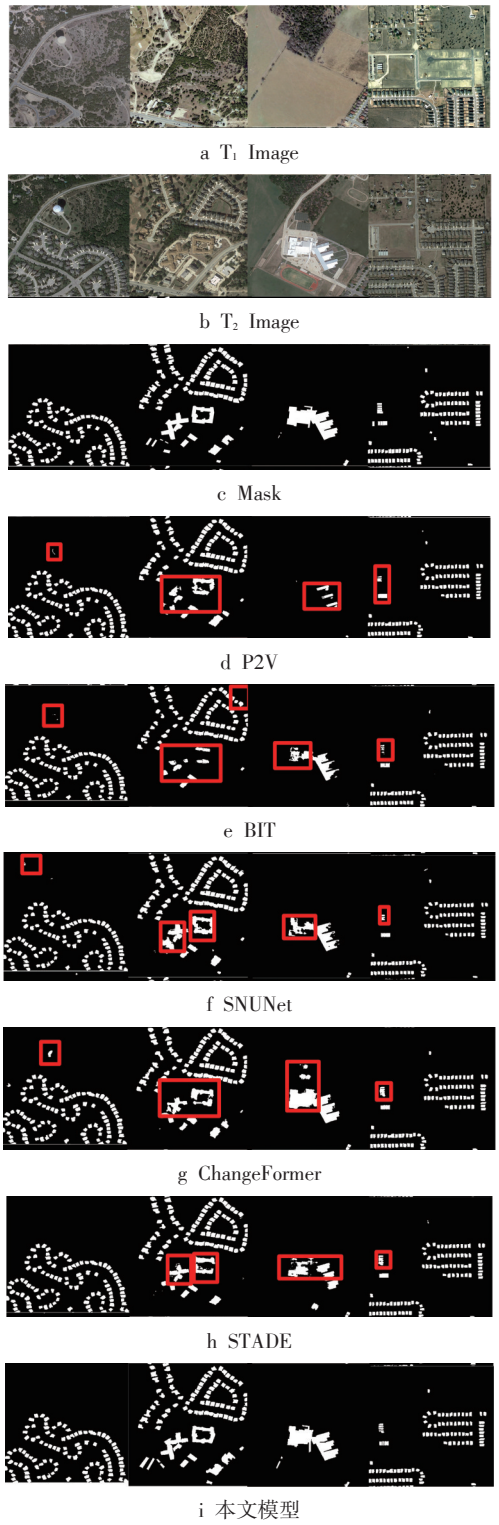
模型	*P*	*R*	F_1-score	*OA*	IoU
P2V	80.49	84.55	82.47	97.68	75.37
BIT	83.72	86.80	85.23	97.96	80.45
SNUNet-CD	84.18	87.82	85.96	98.17	83.23
Siam-UNet++	85.48	89.25	87.32	98.54	84.47
ChangeFormer	85.39	90.54	87.89	98.68	84.84
SwinChangeNet	87.59	91.37	89.44	98.79	86.43
STADE	88.47	92.08	90.24	98.92	88.52
MFTSNet	90.25	92.87	91.54	99.09	89.18
本文模型	92.54	93.77	93.15	99.43	89.67

各种模型在两组测试集上的部分提取结果见图5，可以看出，对照组模型在提取不同尺寸不同分布密度的变化建筑时均出现了不同程度的提取结果空洞残缺、背景误提取，邻近建筑混淆等问题，而本文模型提取的建筑结构更完整，在几何形态与数量方面与真实掩膜更接近，进一步证明了本文模型在遥感建筑变化检测任务上的精度优势。

上述对比实验均在同一组数据集之内开展训练拟合与性能验证。为进一步测试本文模型的泛化性能，在LEVIR-CD与WHU-CD数据集上开展跨数据集交叉应用实验，结果见表3，基于LEVIR-CD数据集训练拟合的模型，在WHU-CD测试集上的F_1-score、*OA*、IoU分别为91.27%、99.06%、88.14%，而基于WHU-CD数据集训练拟合的模型，在LEVIR-CD测试集上的F_1-score、*OA*、IoU分别为92.28%、99.15%、89.22%；说明对于不同区域场景下的变化建筑目标，本文模型均能保持较高的检测精度，具备良好的泛化性能。

2.4　消融实验

为进一步评估本文模型中各模块对于模型精度提升的具体贡献，以BIT为基线模型，将其骨干网络、编码器—解码器、预测头依次替换本文模型中的DFE骨干网络（DFE-BB）、FLA编码器—解码器（FLA-ED）、DUC预测头（DUC-Head）构建多组消融模型，并在LEVIR-CD数据集上开展对比实验，结果见表4。通过消融实验，证明了本文模型中各模块对



a T_1 Image　b T_2 Image　c Mask　d P2V　e BIT　f SNUNet　g ChangeFormer　h STADE　i 本文模型

图5　不同模型建筑变化检测结果对比

表3　交叉验证实验结果对比/%

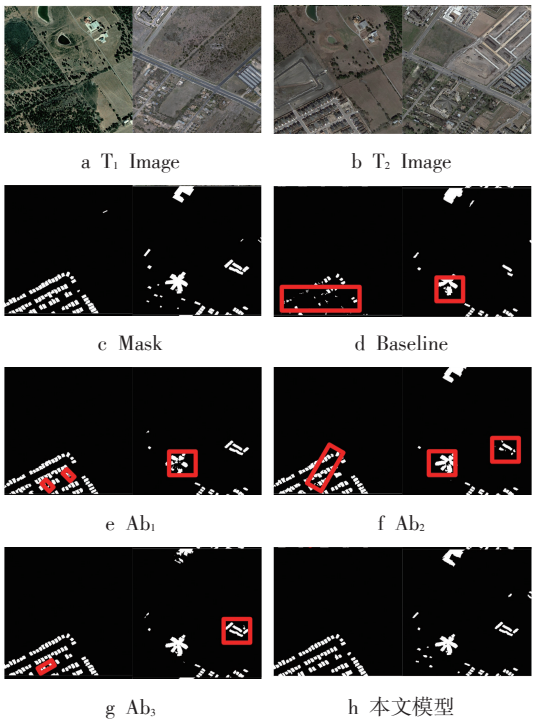
训练集	测试集	*P*	*R*	F_1-score	*OA*	IoU
LEVIR-CD	WHU-CD	89.34	93.29	91.27	99.06	88.14
WHU-CD	LEVIR-CD	91.73	92.84	92.28	99.15	89.22

提升模型精度的具体共享。各消融模型在不同场景下的检测结果见图6，在基准模型中引入DFE-BB、FLA-ED、DUC-Head后，各消融模型对于变化建筑的提取效果均得到了不同程度的优化，在建筑真实形态

和数量方面更接近掩膜中的真实建筑。

表4　消融实验结果对比/%

模型	消融模块			LEVIR-CD	
	DFE-BB	FLA-ED	DUC-Head	F_1-score	*OA*
Baseline	×	×	×	84.75	97.82
1	√	×	×	88.27	98.64
2	×	√	×	85.86	97.87
3	×	×	√	86.79	98.45
本文模型	√	√	√	92.56	99.14



a T_1 Image　b T_2 Image　c Mask　d Baseline　e Ab_1　f Ab_2　g Ab_3　h 本文模型

图6　消融模型检测结果

3　结　语

为充分提取双时相遥感影像中的建筑变化信息，构建了一种融合CNN与FLA单元的遥感建筑变化检测模型。以C2f-IUB架构下的多尺度深度卷积核搭建孪生骨干网络，提取双时相遥感影像中的建筑细节特征；在孪生网络层间通过DFE单元，捕获多尺度差异特征。基于FLA单元构建编码器—解码器网络，在差异特征图上计算全局上下文信息。在预测头中，引入DUC单元对特征图进行精密上采样，恢复特征图尺寸的同时，充分重建其内部的建筑细节特征。实验结果表明，本文模型在LEVIR-CD与WHU-CD数据集上均能准确检出各类建筑的变化情况，各项评价指标均优于对照组模型，对于数据集外的样本，仍能保持良好的泛化性能，可为违章建筑调查、应急灾害救援等应用场景提供有力的技术支持。在后续研究工作中，将进一步优化模型结构并引入更多训练样本，以提升模型的检测精度，同时将尝试通过模型蒸馏、结构剪枝等手段实现模型轻量化，扩充模型的应用场景。

参考文献

[1] 王子明. 基于注意力机制的遥感影像道路提取方法[J]. 地理空间信息,2025,23(1):13-16

[2] 刘力荣,唐新明,甘宇航,等. 新增建设用地卫星遥感智能监测技术研究[J]. 遥感学报,2024,28(11):2828-2837

[3] 范琳,张富玲,张理黎,等. 夜光视角下的西宁—海东都市圈城镇空间格局变化分析[J]. 地理空间信息,2025,23(3):31-34

[4] 戴诗涛,张晶. 基于注意力机制与可变形卷积的遥感图像变化检测方法[J]. 地理空间信息,2024,22(10):22-25

[5] Shangguan Y,Li J,Chen Z,et al. Multiscale Attention Fusion Graph Network for Remote Sensing Building Change Detection[J]. IEEE Transactions on Geoscience and Remote Sensing,2024,62:1-18

[6] Li J,Huang X,Tu L,et al. A Review of Building Detection from very High Resolution Optical Remote Sensing Images[J]. GIScience and Remote Sensing,2022,59(1):1199-1225

[7] Chen H,Li W,Shi Z. Adversarial Instance Augmentation for Building Change Detection in Remote Sensing Images[J]. IEEE Transactions on Geoscience and Remote Sensing,2021,60:14-16

[8] Bai B,Fu W,Lu T,et al. Edge-guided Recurrent Convolutional Neural Network for Multi-temporal Remote Sensing Image Building Change Detection[J]. IEEE Transactions on Geoscience and Remote Sensing,2021,60:10-13

[9] Shafique A,Cao G,Khan Z,et al. Deep Learning-based Change Detection in Remote Sensing Images: A Review[J]. Remote Sensing,2022,14(4):871

[10] 李健慷,张桂欣,祝善友,等. 融合多尺度特征Transformer的高分辨率遥感图像变化检测[J]. 遥感学报,2025,29(1):266-278

[11] Lin M,Yang G,Zhang H. Transition Is a Process: Pair-to-video Change Detection Networks for very High Resolution Remote Sensing Images[J]. IEEE Transactions on Image Processing,2022,32:57-71

[12] Zhang H,Zhang R,Ning X,et al. LIM-CD: A Large-scale Remote Sensing Change Detection Dataset for Incremental Monitoring[J]. ISPRS Annals of the Photogrammetry,Remote Sensing and Spatial Information Sciences,2023,10:903-910

[13] Fang S,Li K,Shao J,et al. SNUNet-CD: A Densely Connected Siamese Network for Change Detection of VHR Images[J]. IEEE Geoscience and Remote Sensing Letters,2021,19:1-5

[14] 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,2022:207-210

[15] Li Z,Cao S,Deng J,et al. STADE-CDNet: Spatial-temporal Attention with Difference Enhancement-based Network for Remote Sensing Image Change Detection[J]. IEEE Transactions on Geoscience and Remote Sensing,2024,62:1-17