引文格式：王航,李炎寅. 场景分类驱动的遥感影像建筑物变化检测[J]. 地理空间信息, 2026, 24(5): 103-106. doi:10. 3969/j.issn. 1672-4623.2026.05.020

场景分类驱动的遥感影像建筑物变化检测

王　航，李炎寅*
（园测信息科技股份有限公司，江苏 苏州 215027）

摘　要：在复杂的城市环境中，现有建筑物变化检测方法存在误检和漏检问题，特别是在小目标变化检测方面。为解决这些问题，提出了一种场景分类驱动的建筑物变化检测网络。首先利用残差块分别从前后时相的遥感影像中提取多尺度深度特征；然后通过构建的差分模块实现前后时相特征的交互融合；最后构建检测头和分类头，分别用于检测精细化的变化区域和判断变化存在性。实验结果表明，该网络显著提升了建筑物变化检测的准确性和鲁棒性，为高精度建筑物变化检测提供了一种有效的解决方案。

关键词：遥感影像；深度学习；变化检测；变化场景分类；特征差分
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2026）05-0103-04

Building Change Detection from Remote Sensing Images Driven by Scene Classification

WANG Hang, LI Yanyin
(Yuance Information Technology Co., Ltd., Suzhou 215027, China)

Abstract: In the complex urban environments, the existing building change detection methods still face issues of false positives and missed detections, especially when detecting changes in small targets. To address these problems, we proposed a building change detection network driven by scene classification. Firstly, we utilized residual blocks to extract multi-scale deep features separately from bi-temporal remote sensing images. Then, we constructed a feature difference module to achieve interactive fusion of the bi-temporal features. Finally, we built detection and classification heads to perform fine-grained detection of change regions and to determine the existence of changes, respectively. Experimental results show that this network significantly enhances the accuracy and robustness of building change detection, providing an effective solution for high-precision building change detection.

Key words: remote sensing image, deep learning, change detection, change scene classification, feature difference

建筑物变化检测广泛应用于城市规划、土地利用监测和灾后评估等方面。随着城市扩张和自然灾害的频繁发生，对高精度建筑物变化检测方法的需求日益迫切[1-2]。尽管遥感技术的发展使得在大范围内有效检测和表征建筑物变化成为可能，但由于场景的复杂性，从双时相影像中准确提取和分析建筑物的变化信息仍然面临挑战[3]。常用的变化检测方法包括代数法[4]、分类法[5]、变换法[6]以及近年兴起的深度学习法[7-9]。代数法、分类法、变换法的操作简单，但在处理复杂的建筑环境和影像质量差异时效果有限；且在多时相影像间的特征提取上存在局限，易忽略细微的变化信息或受环境噪声干扰。神经网络凭借其独特的架构设计和强大的特征提取能力，被广泛应用于遥感变化检测。尽管上述研究取得了良好成果，但在大范围场景中进行建筑物变化检测时，仅依赖像素级的变化检测仍存在较多的误检和漏检问题，特别是在检测小目标的变化方面。因此，本文提出了一种创新的场景分类驱动的变化检测网络（SC-CDNet）。首先利用孪生特征提取网络和基于Transformer的交叉融合模块，分别从前后时相影像中提取特征，并通过交叉融合实现信息交互；再构建检测头和分类头，分别用于变化检测和判断变化存在性，以实现对建筑物变化区域的精细识别。

1　研究方法

SC-CDNet首先利用残差块分别从前后时相的遥感影像中提取多尺度深度特征，标记为 $\{F_1, F_2, F_3, F_4\}$；然后通过构建的差分模块对各尺度对应的前后时相特征进行处理，获得语义变化特征，标记为 $\{D_1, D_2, D_3, D_4\}$；最后构建分类头和检测头，分别用于变化场景分类和变化检测。

第一作者简介：王航（1987—），本科，工程师，从事遥感影像解译工作，E-mail：project1024@126.com。
通信作者：李炎寅（1984—），硕士研究生，高级工程师，从事遥感影像解译工作，E-mail：652412359@qq.com。

1.1 残差块提取特征

本文采用残差块进行特征提取，以提升模型的深度和特征表示能力[10]。残差块通过引入跳跃连接，使网络能在更深层次上保持有效的梯度传播，缓解了深层网络中的梯度消失问题。这一设计能在提取多时相影像的细微特征时实现更高的稳定性和准确性，确保网络在深层结构中对建筑物变化的细节保持敏感。此外，残差块的加法跳跃连接结构提高了特征的复用性，有助于模型捕获建筑物的复杂结构特征，从而提升整体变化检测性能。本文采用下采样残差块和标准残差块两种结构。下采样残差块通过3×3卷积核、步长为2的卷积层降低空间分辨率并增加通道数，以逐步提取更高级的语义信息；通过包含1×1卷积核的卷积层进行通道压缩与扩展，利用跳跃连接中的1×1卷积进行维度匹配，从而实现输入输出的一致性并与主路径进行逐元素相加。标准残差块则保持输入输出维度不变，确保信息流的无损传递。在网络结构中，每个残差块中使用一个下采样残差块和3个标准残差块。在输入尺寸为 $(3, H, W)$ 的遥感图像情况下，经过4个残差阶段的处理后，输出的多尺度特征图分别为 $(32, H/2, W/2)$、$(64, H/4, W/4)$、$(128, H/8, W/8)$ 和 $(256, H/16, W/16)$。

1.2 差分模块

为充分利用多时相影像中的变化信息，增强模型对建筑物细微变化的检测能力，本文设计了一个基于单层多头交叉注意力结构的前后时相特征融合模块。首先通过感知器提取输入的前后时相特征的初始特征，再分别进行维度转换，生成查询矩阵（$\boldsymbol{Q}$）、键矩阵（$\boldsymbol{K}$）和值矩阵（$\boldsymbol{V}$）。特征融合过程中，通过矩阵相乘操作生成注意力权重，以捕捉各自的特征差异。需要注意的是，该步骤使用的是对方的 $\boldsymbol{Q}$，以实现前后时相特征的关联计算[11]。最终对前后时相注意力矩阵相乘后的结果取绝对差分，获得包含细微变化信息的融合特征。绝对差分操作确保前后时相之间的特征差异得到有效表达，增强了模型在多时相变化检测任务中的稳定性和准确性。实验中注意力头数设置为特征通道数 $C/16$，每个子头的维度为 C/Head。最终输出的4个语义差分特征图 $\{D_1, D_2, D_3, D_4\}$ 的尺寸分别为 $(32, H/2, W/2)$、$(64, H/4, W/4)$、$(128, H/8, W/8)$ 和 $(256, H/16, W/16)$。

1.3 检测头和分类头

本文构建的检测头和分类头分别承担空间定位与语义判别任务，协同完成建筑物变化检测。检测头主要用于精确捕捉变化区域在图像中的空间位置，输出像素级的变化结果；分类头则用于从全局尺度判断目标区域是否存在变化，提供场景级的辅助监督信号。这种双分支结构有效平衡了局部定位与全局语义的建模能力，从而提升了模型对复杂建筑变化模式的适应性。具体而言，检测头通过卷积和上采样操作将 $\{D_2, D_3, D_4\}$ 统一为与 D_1 相同的空间尺寸，再经过特征拼接和一个3×3卷积层融合语义信息，最终通过一个1×1卷积和上采样操作得到变化检测结果。分类头通过卷积和下采样操作将 $\{D_1, D_2, D_3\}$ 统一为与 D_4 相同的空间尺寸，再经过特征拼接融合语义信息，并利用过池化层聚合空间信息，最后通过感知器实现二分类输出，即判断当前输入区域是否存在建筑物变化。

2 实验与结果分析

2.1 实验设置

本文采用武汉大学公开发布的建筑变化检测数据集。该数据集专用于遥感图像变化检测[12]，包含新西兰基督城地区在2012年和2016年拍摄的成对图像，影像分辨率为0.2 m，影像尺寸为32 507像素×15 354像素，包含红、绿、蓝3个波段。本文首先将原始图像裁剪为不重叠的256×256影像块，用于模型训练与评估。整个数据集包含4 461个训练影像块、1 487个验证影像块和1 486个测试影像块。

实验基于Windows10操作系统，使用Pytorch1.10深度学习框架进行模型开发与训练。硬件配置为NVIDIA RTX 2080 Ti显卡，32 GB内存，确保计算任务的高效处理与加速。CUDA版本为11.3，用于实现与GPU的并行计算。编程语言为Python3.8，代码编辑器为PyCharm Community 2020。初始学习率设定为0.000 5，采用余弦退火策略进行动态调整，以加快模型收敛并防止陷入局部最优。批大小设置为6，训练共进行300轮。采用交叉熵损失评估预测结果与真实标签的差异。采用随机梯度下降优化器，通过逐步调整模型参数来最小化损失，提高模型的预测精度。

为验证SC-CDNet在建筑物变化场景分类与检测中的有效性，对比分析各种模型的效果。分类任务中的对比模型包括ResNet50[13]、VGG16[14]、MobileNetV2[15]、DenseNet[16]、Vision Transformer[17]和Swin Transformer[18]；检测任务中的对比模型包括FC-EF[19]、SEIFNet[12]、BIT[20]、SNUNet[21]、ChangeFormer[22]和A2Net[23]。实验从定性和定量两个方面评估了SC-CDNet的性能。定性评估通过可视化提取结果，直观对比不同模型的性能。定量评估采用精确度（P）、召回率（R）和F_1-score指标全面衡量SC-CDNet的效果[21]。

2.2　对比实验

1）定量评价。各模型在变化场景分类任务中的性能评估结果见表1，SC-CDNet在所有指标上均取得了最佳表现。具体而言，Vision Transformer在 P 和 F_1-score上的表现次优，分别为96.57%和95.93%，比SC-CDNet低1.57、2.3个百分点；MobileNetV2在 R 上的表现次优，为95.94%，比SC-CDNet低2.39个百分点，但在 P 上的表现最差，仅为92.17%，比SC-CDNet低5.97个百分点；DenseNet在 R 和 F_1-score上表现最差，分别为91.11%和92.06%，分别比SC-CDNet低7.22、6.17个百分点。

表1　变化场景分类任务中的性能评估结果/%

模型	P	R	F_1-score
ResNet50[13]	96.45	95.09	95.77
VGG16[14]	93.85	92.43	93.13
MobileNetV2[15]	92.17	95.94	94.01
DenseNet[16]	93.03	91.11	92.06
Vision Transformer[17]	96.57	95.30	95.93
Swin Transformer[18]	92.37	94.95	93.64
SC-CDNet	98.14	98.33	98.23

各模型在变化检测任务中的定量评估结果见表2，SC-CDNet在所有指标上均取得了领先。具体而言，SC-CDNet与性能较优的BIT模型相比，尽管二者在 P 指标上仅相差0.78个百分点，但在 R 和 F_1-score上分别提升了2.20、1.49个百分点。值得注意的是，F_1-score作为综合考虑精度与召回能力的核心评价指标，提升尤为重要，表明SC-CDNet在保证准确识别变化区域的同时，有效减少了误检和漏检现象。BIT在 P 和 F_1-score上的表现次优，分别为93.24%和92.20%，比SC-CDNet低0.78、1.49个百分点；SEIFNet在 R 上的表现次优，为91.38%，比SC-CDNet低1.99个百分点；而FC-EF在所有指标上的表现均最差。综上所述，SC-CDNet在两个任务中均展现了显著优势，验证了其在变化场景分类与检测任务中的卓越性能和可靠性。

表2　变化检测任务中的性能评估结果/%

模型	P	R	F_1-score
FC-EF[19]	90.97	89.89	90.43
SEIFNet[12]	91.15	91.38	91.27
BIT[20]	93.24	91.17	92.20
SNUNet[21]	92.14	91.18	91.65
ChangeFormer[22]	93.23	89.92	91.55
A2Net[23]	91.91	91.26	91.58
SC-CDNet	94.02	93.37	93.69

2）定性评价。各模型在多个典型场景中的可视化对比结果见图1，第一行为小目标建筑物变化场景，前后时相均存在建筑，但发生细微变更；第二行为大目标建筑物变化场景，变化区域范围较大且结构清晰；第三行为无变化场景，用于测试模型在无变化场景中的鲁棒性；第四行为新增小型密集建筑场景；第五行为新增大型建筑场景；第六行为存在树木遮挡的复杂背景场景。结果表明，尽管所有模型在不同场景中均存在一定程度的误检或漏检，但SC-CDNet在整体检测准确性和边界保持能力上表现最优。无论是在小目标还是大目标建筑物变化场景中，SC-CDNet都能准确捕捉变化区域，显著减少错误响应。在无变化场景中，SC-CDNet未产生误检，说明其在场景分类与变化判别的协同机制下，具有更强的非变化区域识别能力。在复杂遮挡场景中，SC-CDNet的误检现象最少，说明其在应对背景干扰、边界模糊等复杂情况时具有更强的鲁棒性和泛化能力。

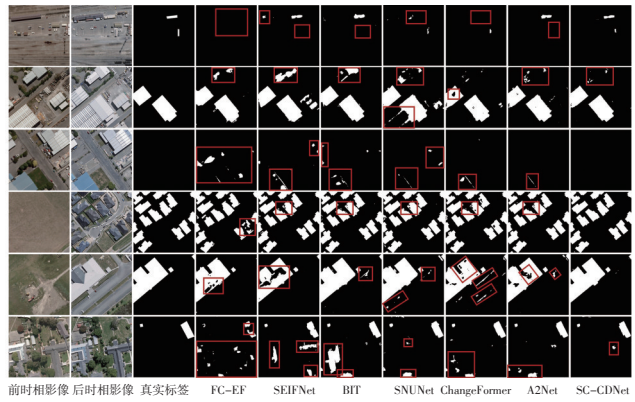


图1　变化检测的可视化结果

2.3　多任务方法的有效性

SC-CDNet能同时完成变化场景分类和变化检测任务，可视为一个多任务的网络模型。为验证这两个任务是否对彼此有益，本文进行单任务与多任务的对比实验。其中，单任务方法中的变化场景分类是通过去掉检测头实现的；类似地，变化检测任务则是通过去掉分类头进行的。结果表明（表3），多任务方法在变化场景分类和变化检测两个任务上均比单任务有所提升。具体来说，多任务方法在变化场景分类任务上的 P、R 和 F_1-score分别提高1.79、1.28、1.53个百分点；在变化检测任务上的 P、R 和 F_1-score分别提升0.69、2.31、1.51个百分点。多任务方法能有效利用任务间的关联性，从而在两个任务上都取得更优的性能。

表3　变化检测任务中的性能评估结果/%

方法		P	R	F_1-score
单任务方法	变化场景分类	96.35	97.05	96.70
	变化检测	93.33	91.06	92.18
多任务方法	变化场景分类	98.14	98.33	98.23
	变化检测	94.02	93.37	93.69

3 结 论

本文提出了一种场景分类驱动的建筑物变化检测网络SC-CDNet，旨在解决复杂城市环境中建筑物变化检测的误检和漏检问题，特别是针对小目标的变化检测。利用残差块从前后时相遥感影像中提取多尺度深度特征，再结合基于Transformer的差分模块，实现了前后时相特征的交互融合，有效捕捉了细微的变化信息。此外，构建检测头和分类头分别用于精细化的变化区域检测和变化存在性判断，进一步提升了模型的准确性和鲁棒性。实验结果表明，SC-CDNet在公开建筑物变化检测数据集上取得了优异的性能，定性和定量分析均证明了本文方法在建筑物变化检测任务中的有效性和可靠性。综上所述，本文方法为高精度建筑物变化检测提供了一种有效的解决方案，具有重要的实际应用价值。未来的工作将致力于进一步优化模型结构，提升对更复杂场景和多源遥感数据的适应性，以及探索在其他变化检测领域的应用潜力。

参考文献

[1] 鄢薪,慎利,潘俊杰,等. 多尺度特征融合与空间优化的弱监督高分遥感建筑变化检测[J]. 测绘学报,2024,53(8):1586-1597

[2] 厉芳婷,张过,周方栋,等. 基于双流并行全向扫描Mamba的遥感影像建筑物变化检测[J]. 武汉大学学报(信息科学版),2024,49(9):1574-1585

[3] 林娜,孙鹏林,王玉莹,等. 多尺度特征孪生神经网络的建筑物变化检测方法[J]. 测绘科学,2022,47(5):185-192

[4] 张晗,倪维平,严卫东,等. 利用分形和多尺度分析的中低分辨率SAR图像变化检测[J]. 武汉大学学报(信息科学版),2016,41(5): 642-648

[5] 唐朴谦,杨建宇,张超,等. 基于像素比值的面向对象分类后遥感变化检测方法[J]. 遥感信息,2010,25(1):69-72

[6] 黄维,黄进良,王立辉,等. 基于PCA的变化向量分析法遥感影像变化检测[J]. 国土资源遥感,2016,28(1):22-27

[7] 孙剑明,赵梦鑫,郝旭耀. 遥感图像变化检测方法研究综述[J]. 计算机工程与应用,2024,60(20):30-48

[8] 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

[9] 祁鑫博. 遥感影像面向对象变化检测综述[J]. 北京测绘,2021,35(4):427-431

[10] 孔文学,罗亦泳,陈心龙,等. 高分辨率遥感影像建筑物提取卷积神经网络[J]. 北京测绘,2024,38(1):44-49

[11] Yang C K,Chen M H,Chuang Y Y,et al. 2D-3D Interlaced Transformer for Point Cloud Segmentation with Scene-level Supervision[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV),Paris,2023

[12] Huang Y Y,Li X H,Du Z S,et al. Spatiotemporal Enhancement and Interlevel Fusion Network for Remote Sensing Images Change Detection[J]. IEEE Transactions on Geoscience and Remote Sensing,2024,62:1-14

[13] He K M,Zhang X Y,Ren S Q,et al. Deep Residual Learning for Image Recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),Seattle,WA,2016

[14] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-scale Image Recognition[J]. arXiv preprint arXiv,2014,1:16-20

[15] Sandler M,Howard A,Zhu M L,et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks[C]//Proceedings of the 31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),Salt Lake City,UT,2018

[16] Huang G,Liu Z,Van Der Maaten L,et al. Densely Connected Convolutional Networks [C]//Proceedings of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),Honolulu,2017

[17] Dosovitskiy A. An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale[J]. arXiv preprint arXiv,2020,16:22-35

[18] Liu Z,Lin Y T,Cao Y,et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows[C]//Proceedings of the 18th IEEE/CVF International Conference on Computer Vision (ICCV),Electr Network,2021

[19] Daudt R C,Le Saux B,Boulch A. Fully Convolutional Siamese Networks for Change Detection[C]//Proceedings of the 25th IEEE International Conference on Image Processing (ICIP),Athens,Greece,2018

[20] Chen H, Qi Z P, Shi Z W. Remote Sensing Image Change Detection with Transformers[J]. IEEE Transactions on Geoscience and Remote Sensing,2022,60:14-23

[21] 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,2022,19:5-16

[22] Bandara W G C, Patel V M. A Transformer-based Siamese Network for Change Detection[C]//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS),Kuala Lumpur,Malaysia,2022

[23] Li Z L,Tang C,Liu X W,et al. Lightweight Remote Sensing Change Detection with Progressive Feature Aggregation and Supervised Attention[J]. IEEE Transactions on Geoscience and Remote Sensing,2023,61:1-12