

doi: 10.6046/zrzyyg.2021122

引用格式: 薛白,王懿哲,刘书含,等. 基于孪生注意力网络的高分辨率遥感影像变化检测[J]. 自然资源遥感,2022,34(1):61–66. (Xue B, Wang Y Z, Liu S H, et al. Change detection of high-resolution remote sensing images based on Siamese network[J]. Remote Sensing for Natural Resources, 2022, 34(1): 61–66.)

基于孪生注意力网络的高分辨率遥感影像变化检测

薛白, 王懿哲, 刘书含, 岳明宇, 王艺颖, 赵世湖

(自然资源部国土卫星遥感应用中心, 北京 100048)

摘要: 随着遥感影像空间分辨率的提升,地物成像特征愈加复杂,基于纹理表达和局部语义等技术的变化检测方法已很难满足需求。为提升高分辨率遥感影像的变化检测精度,构建了一套较大规模的0.8~2 m 高分辨率遥感人类活动变化检测数据集(HRHCD-1.0);同时将空间注意力和通道注意力机制引入孪生变化检测网络中,设计了具有更强上下文变化语义特征提取能力的孪生注意力变化检测网络。模型对比实验中,孪生注意力变化检测模型相比非注意力机制模型在验证集上平均交并比提升24%,检测结果更完整,有效缓解了非注意力模型边界较差、局部漏检和空洞等问题。后处理方法对检测结果的小图斑去除、填洞和图形学平滑等处理,提升了图斑图形效果。变化检测训练中样本量增加对于模型应用的精度和泛化能力有显著提升作用。

关键词: 高分辨率遥感影像;变化检测;深度学习;注意力机制

中图法分类号: TP 79 **文献标志码:** A **文章编号:** 2097-034X(2022)01-0061-06

0 引言

近年来,随着高分辨率遥感影像的普及,在土地、环保和减灾等领域的大范围高分辨率土地覆被整体变化或特定要素变化的应用需求日益旺盛。传统基于人工作业的方式分析遥感影像变化费时费力,自动实现时序影像变化检测具有重要的科学研究和应用价值,遥感界对该领域的研究已经开展了数十年^[1-3]。高分辨率影像随着空间分辨率的提升,影像的精细细节和复杂纹理特征更加丰富,同时受到成像角度、配准及辐射处理等差异、传感器差异等影响^[4],急剧增加了自动变化检测难度,不同地物光谱和纹理特征混淆性增加,可分性降低。传统基于光谱特征阈值差分图^[5]、主成分分析^[6]和典型相关分析^[7]等方法应用于高分辨率遥感影像变化检测时存在图斑较破碎等问题。基于面向对象分析的方法采用对象纹理等特征表达的方式可以有效提升检测质量,例如李亮等^[8]提出充分融合光谱特征和纹理特征的面向对象多特征融合的变化检测方法,提高了变化检测的精度,但该类方法往往先进行对象化分割,对象化分割结果对变化检测效果影响

显著,且主要适用于均质性较好的地物变化场景。

随着深度学习方法在计算机视觉领域的成功,具有更强影像语义特征提取能力的深度学习方法逐渐被引入到高分辨率遥感影像变化检测研究中,其无需先对影像对象化,通过神经网络直接提取多尺度特征实现像素级变化区域检测。引入深度学习的变化检测方法主要分为2类:仅用深度学习做特征提取的方法和端到端直接提取变化的方法。前者首先通过深度信念网络、自动编码器和卷积神经网络等深度学习技术提取像素或对象的深度特征,然后生成差分向量通过聚类或分类方法获取变化^[9];后者直接对输入的时序影像提取特征获取变化结果^[10]。由于深层神经网络具有很强的深层特征提取能力,这些方法取得了优于传统方法的性能,但 these 方法大多数是从单影像语义分割网络中修改过来,缺乏对变化检测问题的针对性设计,在较复杂场景应用中存在检测完整度低、易误检漏检等问题,与工程化生产应用需求尚存在一定差距。研究针对复杂多样化的高分辨率遥感影像典型人类活动地表变化检测应用,提出一种孪生注意力变化检测深度学习方法。

同时,深度学习方法的精度和泛化性严重依赖样本,虽然目前已有多个高分辨率遥感变化检测数

收稿日期: 2021-04-23; 修订日期: 2021-11-08

基金项目: 民用航天“十三五”预先研究项目“测绘信息产品作业测试和服务规范研究”(编号: D040401)资助。

第一作者: 薛白(1988-),女,硕士,工程师,主要从事遥感数字图像处理、土地利用与动态监测等相关研究。Email: 1069460245@qq.com。

据集公开(表 1),但大多针对建筑类变化^[11-14],仅少量覆盖多类型地表变化^[15-16],并且数据量较小且同本研究的人类活动地表变化检测业务应用类型不能匹配。针对典型人类活动地表变化检测应用,需要构建面向人类活动变化检测应用的高分辨率遥感变化检测数据集。

表 1 主要的开源高分辨率遥感变化检测数据集

Tab. 1 Main open source high-resolution remote sensing change detection data set

数据集	变化类型	分辨率/ m	数据量/ 组	数据尺寸/ 像素
LEVIR-CD ^[11]	建筑物	0.5	637	1 024 × 1 024
Google Data-set ^[12]	建筑物	0.55	20	1 006 × 1 168 ~ 4 936 × 5 224
WHUCD ^[13]	建筑物	0.2	1	15 354 × 32 507
AICD ^[14]	建筑物、临时目标	约 0.1 ~ 0.5	10 000	256 × 256
SECOND ^[15]	多类型	约 1	4 662	512 × 512
HRSCD ^[16]	多类型	0.5	2	10 000 × 10 000

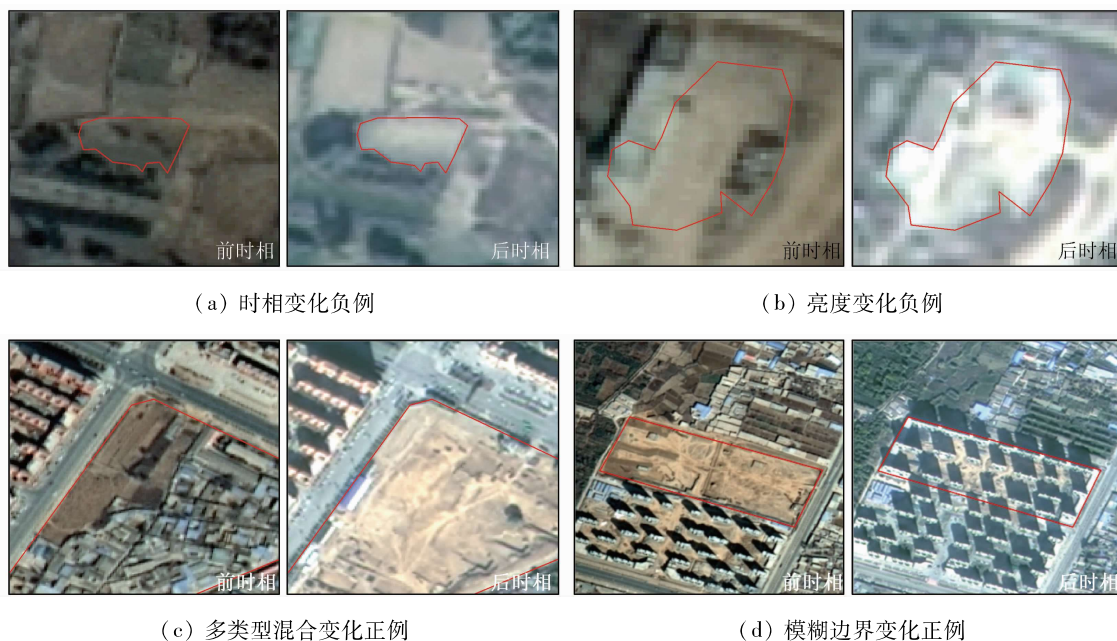


图 1 HRHCD 数据集样本正例与负例示例

Fig. 1 Examples of positive and negative examples of HRHCD data set

存在多类型混合或没有明显边界的情况,例如图 1(c)和图 1(d)。这些因素使得本数据集成为一个很有挑战性的变化检测数据集。

1.2 孪生注意力机制变化检测网络

孪生网络已经被证明在变化检测任务中有良好表现^[17],研究采用双分支孪生网络进行特征提取,2期影像分别接入 2 个结构相同权重共享的卷积分支提取特征。直接使用全卷积孪生变化检测网络存在检测完整度低、易误检漏检等问题,这主要受限于网络特征提取能力不足和空间与通道域上下文语义信息未有效利用。考虑到高分辨率遥感影像变化检测具有更丰富的特征和语义信息提取与利用能力的要

1 研究方法

1.1 NHRCD-1.0 变化检测数据集构建

研究首先构建高分辨率遥感人类活动变化检测数据集(high-resolution remote sensing human activity change detection dataset, HRHCD-1.0)。选择 GF-1、GF-2 和 ZY-3 等国产主流高分辨率卫星影像处理生成 2 m 和 0.8 m 融合数据构建成组的影像对,采集建筑物、推填土、道路、沟渠和矿区等各类目标新增和减少的变化图斑,按照 512 像素 × 512 像素数据尺寸成对裁剪得到 23 026 组样本。本数据集涉及变化类型多样,变化场景有明确的业务应用定义,考虑到存在很多同真实变化类似的影像变化,选择易混淆的非变化区域生成负例样本,例如图 1(a)中的时相变化和图 1(b)中的亮度变化均是负例。同时,变化图斑按土地利用场景采集,场景图斑

求,研究一方面引入具有注意力机制的骨干网络提取特征,另一方面引入空间注意力和通道注意力机制模块^[18]对 2 个分支提取的孪生特征进行融合,获取更好的空间与通道域语义信息,设计得到孪生注意力机制变化检测网络,简称 Siam-Atte。网络结构(图 2)的主要流程包括:①孪生特征提取。分别以变化前时相(T1)和变化后时相(T2)影像为输入,利用深度卷积神经网络进行影像特征提取。研究采用具有分隔注意力机制的 ResNeSt50 网络作为主干实现丰富的特征提取能力,为了有效联合 2 个时相的特征构建具有可比性的同一特征空间,T1 和 T2 的特征提取网络选用同样的结构并且参数共享。②特征

融合。将2个分支提取的特征作为输入,通过双注意力模块(spatial attention and channel attention, SACA)

分别实现注意力特征提取,对特征进行拼接,再通过特征金字塔卷积分类实现变化检测。

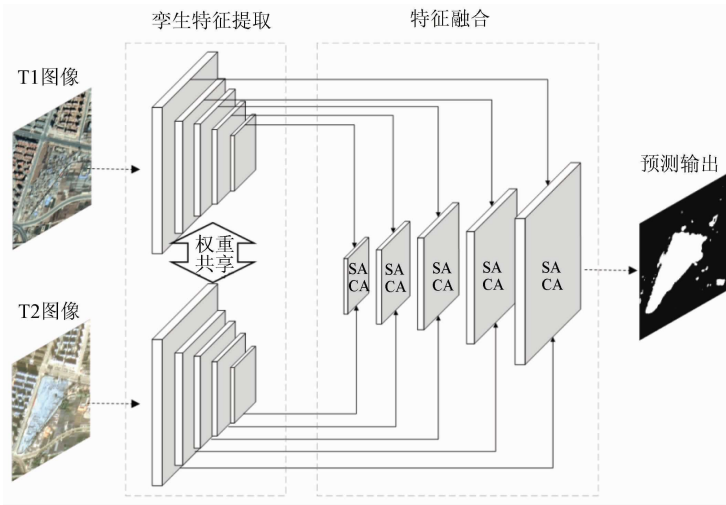


图2 孪生注意力机制变化检测网络结构

Fig. 2 Siamese attention mechanism change detection network structure

网络中对于变化与非变化区域复杂特征提取的优化主要是通过注意力机制模块实现,包括空间注意力机制和通道注意力机制。首先分别从孪生特征转换得到注意力特征,然后将2个注意力模块的输出特征进行聚合,以获得更好的特征表示。空间注意力特征的计算过程为:①利用空间注意力矩阵对孪生特征空间上的任意2个像素之间的空间关系进行建模;②在空间注意力矩阵和原始孪生特征之间进行矩阵乘法;③对第二步得到的矩阵进行逐元素求和运算,得到最终的特征矩阵。通道注意力特征的计算过程为:①计算通道维度上的通道注意力矩阵;②在通道注意力矩阵和原始孪生特征之间进行

矩阵乘法;③将第二步得到的矩阵与原始特征逐元素求和。最终,将2个注意力模块的输出进行逐元素求和,并通过卷积聚合。

1.3 模型训练

1.3.1 损失函数

变化目标呈现多样化的尺度特征,且相对背景占比较小,考虑到现实中变化区域一般较少,存在显著的变化与非变化区域样本类别不平衡的问题,研究采用对样本不均衡有效的平衡二分类交叉熵和Dice系数损失综合的损失函数,损失函数 L 为两者的加权和,公式为:

$$L = L_{bce} + \lambda L_{dice}, \quad (1)$$

$$L_{bce} = -\beta \sum_{j \in Y_+} \ln Pr(y_j = 1) - (1 - \beta) \sum_{j \in Y_-} \ln Pr(y_j = 0), \quad (2)$$

$$L_{dice} = 1 - \frac{2|Y' \cap Y|}{|Y'| + |Y|}, \quad (3)$$

式中: L_{bce} 为平衡二分类交叉熵损失; L_{dice} 为Dice系数损失; λ 为权重系数,取值0.2; $\beta = |Y_-| / (|Y_+| + |Y_-|)$ 和 $1 - \beta = |Y_+| / (|Y_+| + |Y_-|)$, $|Y_+|$ 和 $|Y_-|$ 为样本数据统计的变化和非变化像素数; $Pr(y_j)$ 为像素 j 的sigmoid输出; Y' 和 Y 分别为预测和标签中变化像素的个数。

1.3.2 数据增强

考虑到应用中变化影像对的前后期色彩、分辨率、亮度等差异性,研究采用了较多的增强处理方法以丰富数据提升模型的鲁棒性,训练中采用的动态增强方法包括:①2期影像成对的 90° 、 180° 和 270° 旋转, $0 \sim 40^\circ$ 的随机旋转, $0.8 \sim 1.25$ 倍的随机缩放,水平和垂直翻转;②2期影像独立的随机噪声,

高斯模糊,色彩扰动和对比度变化。

1.3.3 网络训练

为了提升模型训练的效率,采用ResNeSt50网络在ImageNet数据上的预训练权重,设置输入影像尺寸为 512×512 像素,采用2阶段训练方法。第一阶段训练时冻结ResNeSt50特征提取层权重设置批大小为16,迭代轮数为300,采用Adam优化器,初始学习率为 $2E-3$,余弦衰减到 $1E-6$ 。第二阶段训练时取消ResNeSt50权重冻结,冻结批归一化层设置批大小为4,迭代轮数为200,采用Adam优化器,初始学习率为 $1E-4$,分段衰减到 $1E-8$,同时为避免训练过拟合,采用早停策略。

1.4 业务应用后处理

对于人类活动变化检测业务化应用,深度学习提取的结果存在图斑空洞、小图斑和边界不平滑等问题,需要进行优化处理。对于存在空洞的图斑通

过形态学方法进行填洞处理；对于业务应用不关注的小图斑按像素阈值进行删除处理；针对边界优化进行 7 个像素的形态学膨胀与腐蚀操作去除提取结果边界的毛刺和尖锐凹槽，以获取更加平滑的图斑边界。栅格结果图形优化后，还需要将栅格结果矢量化以得到业务可直接应用的矢量成果。矢量化中采用抽稀平滑操作进一步解决矢量成果的锯齿状和节点多等问题，提升成果的美观性。

2 实验及结果分析

2.1 评价指标

研究分别对训练和应用建立评价指标，评价在不同阶段评估模型的性能。训练评价指标采用像素级的平均交并比，应用阶段采用图斑级的召回率和正确率。平均交并比是语义分割领域常用的评价指标，其表达式为：

$$IoU = \frac{P_{ii}}{P_i + P_j - P_{ii}}, \quad (4)$$

$$mIoU = \frac{\sum_{l=1}^k IoU_l}{k}, \quad (5)$$

式中： IoU 为交并比； $mIoU$ 为平均交并比； P_i 为真实值像素数； P_j 为预测值像素数； P_{ii} 为将真实值预测正确的像素数； l 为第 l 类； k 为类别数量，单类变化中仅考虑变化像素， k 为 1。

在应用阶段为了更直观地评估模型效果，采用图斑级的召回率和正确率指标。首先对预测结果进行矢量化，同标注矢量进行逐图斑对比，定义预测与标注图斑交并比大于 30% 的图斑为正确预测的正例图斑，正确率 P 和召回率 R 分别为：

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

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

式中： TP 为正确预测的正例； FP 为错误预测的正例； FN 为错误预测的错例。

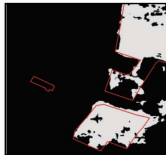
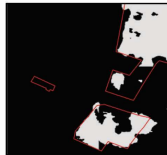
2.2 实验分析

2.2.1 模型精度分析

研究采用相同的特征提取网络 ResNeSt50 训练非注意力机制的全卷积孪生网络模型 (Siam - Diff 和 Siam - Conc) 与本文方法 (Siam - Atte) 进行训练精度对比分析。Siam - Diff, Siam - Conc 和 Siam - Atte 这 3 种网络模型在同一测试数据集上的 $mIoU$ 分别为 0.357, 0.324 和 0.443。Siam - Atte 网络具有显著的精度优势， $mIoU$ 比 Siam - Diff 网络高 0.086，相对提升 24%。建筑群、推填土和道路 3 种典型场景下 3 种网络的检测结果如表 2 所示。从表 2 中可以看出：① Siam - Diff 和 Siam - Conc 结果存在边界较差、局部漏检和较多空洞等现象，Siam - Atte 结果建筑群、道路等完整性较好，相对更准确；② 3 种模型均存在少量的误检测，表现为检测范围过大或过小。

表 2 不同网络检测结果对比

Tab. 2 Comparison of detection results of different networks

类别	前时相影像	后时相影像	Siam - Diff 结果	Siam - Conc 结果	Siam - Atte 结果
建筑群					
推填土					
道路					

2.2.2 后处理提取图斑效果分析

研究对模型检测结果进行多种图形学后处理以优化检测结果。图 3 展示了后处理优化效果，图中浅灰色区域为原始检测结果，红线和蓝线分别为标

注图斑和优化后的矢量图斑。从图 3 中可以看出，后处理方法对检测结果实现了小图斑去除、填洞和图形学平滑，提升了图斑完整度、精度和图形美观性，同时对于小图斑误提取有去伪作用。

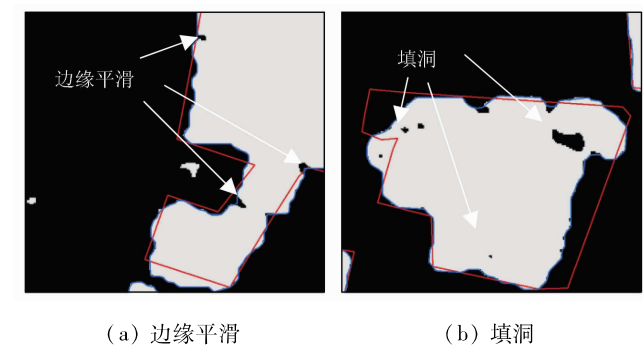


图 3 后处理优化结果

Fig. 3 Results with post – processing optimization

2.2.3 不同数据量训练模型的应用测试

研究在不同训练数据规模下训练模型并分析模型的应用效果。为了独立评估应用效果,选择独立的测试数据采用正确率和召回率对不同模型应用效果进行评价分析。训练样本量分别设置为 5 000, 10 000 和 20 000 组,表 3 给出了不同数据集训练模型的测试指标。结果表明模型应用的召回率和正确

率随着样本量的增大而提升,20 000 组样本量模型精度最佳,小样本量模型在应用中存在相对较多误检,样本量对于模型应用的精度和泛化能力有显著提升作用。

表 3 不同训练数据规模下模型应用效果				
Tab.3 Model application effects under different training data scales				
样本量/组	预测图斑/个	真值图斑/个	R/%	P/%
5 000	1 197	390	51.3	16.8
10 000	734	390	66.9	35.6
20 000	667	390	72.8	42.7

尽管数据集样本量达到 20 000 组,但是测试中发现模型正确率仍然偏低,主要的误检类型包括时相差异、高大建筑物阴影、云雾干扰等,典型的误检图斑如图 4 所示。误检测主要由于实际应用场景中影像变化类型更加多样,而模型的适配性和鲁棒性还较低,需要针对应用结果设计相应的负例样本进行优化训练。

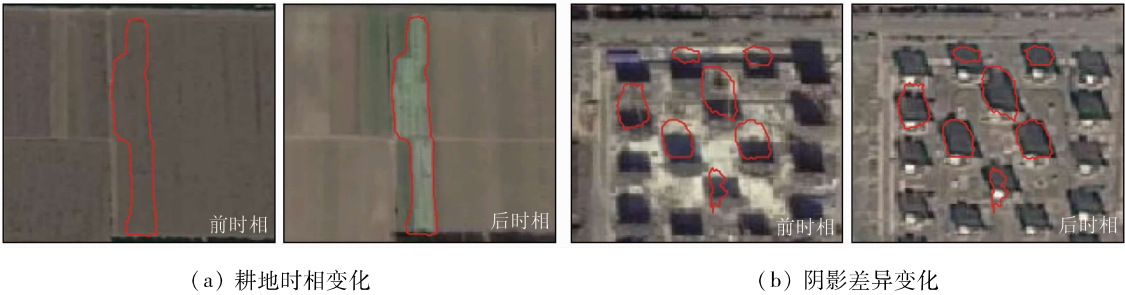


图 4 非业务应用类型的误检测图斑示例

Fig. 4 Examples of misdected patterns for non – business types

4 结论

1) 本文面向高分辨率遥感数据场景下的典型人类活动变化检测应用,一方面利用高分辨率遥感影像构建了一套面向人类活动变化检测应用的 0.8 ~ 2 m 高分辨率遥感变化检测数据集 (HRHCD – 1.0); 另一方面构建了用于 2 期影像变化检测的孪生注意力机制变化检测网络,特别地引入空间注意力模块和通道注意力模块进行注意力特征提取和融合,实现具有更强上下文语义特征提取能力的变化检测网络。

2) 本文模型相比全卷积孪生模型提取的变化结果更完整,有效缓解了后者存在的边界较差、局部漏检和较多空洞等问题; 后处理方法对检测结果的小图斑去除、填洞和图形学平滑有良好效果,提升了图斑完整度,减少了小型误提图斑; 在较复杂的变化场景应用中,小样本量模型存在相对较多的误检,样本量增加对于模型应用的精度和泛化能力有显著提升作用。

3) 然而,在复杂的人类活动变化检测应用场景中,面临着业务应用中变化场景复杂多样判定难、变化区域边界模糊、图斑边界业务规则化要求高等挑战,本文变化检测方法还存在正确率不高、边界套合精度不足等问题。同时,业务化应用中对变化检测算法的正确率和召回率均有较高的需求,进一步的研究工作主要考虑针对业务应用扩展样本、研究更好的特征提取网络、设计更具鲁棒性的网络和研究变化结果的误检图斑去除方法等。

参考文献 (References) :

[1] Tewkesbury A P,Comber A J,Tate N J,et al. A critical synthesis of remotely sensed optical image change detection techniques[J]. Remote Sensing of Environment,2015,160:1 – 14.

[2] Bruzzone L,Prieto F. Automatic analysis of the difference image for unsupervised change detection[J]. IEEE Transactions on Geoscience and Remote Sensing,2000,38(3):1171 – 1182.

[3] Celik T. Unsupervised change detection in satellite images using principal component analysis and K – means clustering[J]. IEEE Geoscience and Remote Sensing Letters,2009,6(4):772 – 776.

[4] 张良培,武 辰. 多时相遥感影像变化检测的现状与展望[J].

- 测绘学报, 2017, 46(10): 1447 – 1459.
- Zhang L P, Wu C. Advance and future development of change detection for multi – temporal remote sensing imagery[J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(10): 1447 – 1459.
- [5] Saux B L, Randrianarivo H. Urban change detection in SAR images by interactive learning[C]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2013: 3990 – 3993.
- [6] Jackson R D. Spectral indices in N – space[J]. *Remote Sensing of Environment*, 1983, 13(5): 409 – 421.
- [7] 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. doi: 10.6046/gtzyyg. 2016. 01. 04.
- Huang W, Huang J L, Wang L H, et al. Remote sensing image change detection based on change vector analysis of PCA component[J]. *Remote Sensing for Land and Resources*, 2016, 28(1): 22 – 27. doi: 10.6046/gtzyyg. 2016. 01. 04.
- [8] 李亮, 舒宁, 王凯, 等. 融合多特征的遥感影像变化检测方法[J]. *测绘学报*, 2014, 43(9): 945 – 953.
- Li L, Shu N, Wang K, et al. Change detection method for remote sensing images based on multi – features fusion[J]. *Acta Geodaetica et Cartographica Sinica*, 2014, 43(9): 945 – 953.
- [9] Zhang H, Gong M G, Zhang P Z, et al. Feature – level change detection using deep representation and feature change analysis for multispectral imagery[J]. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(11): 1666 – 1670.
- [10] Peng D F, Zhang Y J, Guan H Y. End – to – end change detection for high resolution satellite images using improved UNet + + [J]. *Remote Sensing*, 2019, 11(11): 1382.
- [11] Chen H, Shi Z. A spatial – temporal attention – based method and a new dataset for remote sensing image change detection[J]. *Remote Sensing*, 2020, 12(10): 1662.
- [12] Peng D, Bruzzone L, Zhang Y, et al. SemiCDNet: A semisupervised convolutional neural network for change detection in high resolution remote – sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(7): 5891 – 5906.
- [13] Ji S P, Wei S Q, Lu M, et al. 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.
- [14] Bourdis N, Marraud D, Sahbi H. Constrained optical flow for aerial image change detection[C]//2011 IEEE International Geoscience and Remote Sensing Symposium, 2011: 4176 – 4179.
- [15] Yang K P, Xia G S, Liu Z C, et al. Asymmetric siamese networks for semantic change detection[C/OL]. [2020 – 10 – 12]. <https://arxiv.org/abs/2010.05687v1>.
- [16] Daudt R C, Saux B L, Boulch A, et al. Multitask learning for large – scale semantic change detection[J]. *Computer Vision and Image Understanding*, 2019, 187: 102783.
- [17] 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). IEEE, 2018.
- [18] Chen J, Yuan Z, Peng J, et al. DASNet: Dual attentive fully convolutional siamese networks for change detection of high resolution satellite images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 1194 – 1206.

Change detection of high – resolution remote sensing images based on Siamese network

XUE Bai, WANG Yizhe, LIU Shuhan, YUE Mingyu, WANG Yiyang, ZHAO Shihu

(Land Satellite Remote Sensing Application Center, Ministry of Natural Resources, Beijing 100048, China)

Abstract: With the improvement of the spatial resolution of remote sensing images, the imaging features of ground objects have become increasingly complex. As a result, the change detection methods of remote sensing images based on texture expression and local semantics are difficult to meet the demand. To improve the change detection accuracy of high – resolution remote sensing images, this study constructed a large – scale remote sensing – based human activity change detection dataset (HRHCD – 1.0) with a high resolution of 0.8 ~ 2 m. Moreover, this study designed an attention – based Siamese change detection network with a strong capability to extract contextual semantic features by introducing spatial attention and channel attention mechanisms. In the model comparative experiment, the attention – based Siamese change detection network proposed in this study increased the mean intersection over union on the validation set by 24% and showed more complete detection results compared to the models using non – attention mechanisms, effectively alleviating the problems of poor boundary, local omission, and holes of models using non – attention mechanisms. The post – processing method allows for small polygon removal, hole filling, and graphic smoothing of the detection results, improving the processing graphic effects of polygons. Furthermore, the increase in the sample size in the training of change detection significantly improves the application accuracy and generalization ability of the attention – based Siamese change detection network proposed in this study.

Keywords: high – resolution remote sensing images; change detection, deep learning; attention mechanism

(责任编辑: 陈理)