引文格式：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. doi:10. 3969/j. issn. 1672-4623.2024.10.006

基于注意力机制与可变形卷积的遥感图像变化检测方法

戴诗涛[1]，张　晶[1*]
（1. 广东省国土资源测绘院，广东 广州 510700）

摘　要：针对传统遥感图像变化检测语义分割方法分割精度不高、边缘分割不准确的问题，提出一种基于可变形卷积与注意力机制的遥感图像变化检测方法。该方法以PSPNet为基础模型，首先，主干网络选择轻量化的MobileNet，可以减少模型的参数量；其次，在金字塔池化模块中引入可变形卷积，能够通过可变形的接收场自适应获取遥感图像前后变化特征信息；最后，引入注意力机制模块，能够忽略无关的背景信息。该方法在CDD、LEVIR-CD遥感图像变化检测数据集上进行验证，交并比指标较基础模型分别提高了1.45%和2.01%。结果表明，相较于经典的语义分割方法，该方法融合了可变形卷积、双重注意力机制的优点，能够更加准确地提取遥感图像中变化检测的信息。

关键词：遥感图像；变化检测；可变形卷积；注意力机制
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2024）10-0022-04

Remote Sensing Image Change Detection Method Based on Attention Mechanism and Deformable Convolution

DAI Shitao[1], ZHANG Jing[1]
(1. Surveying and Mapping Institute Lands and Resource Department of Guangdong Province, Guangzhou 510700, China)

Abstract: Aiming at the problems of poor segmentation accuracy and inaccurate edge segmentation in traditional semantic segmentation methods for remote sensing image transformation detection, we proposed a remote sensing image change detection method based on deformable convolution and attention mechanism. Taking PSPNet as the base model, the backbone network chooses the lightweight MobileNet, which can reduce the number of parameters of the model at first. And then, deformable convolution is introduced into the pyramid pooling module, which is able to adaptively acquire the before and after change feature information of remote sensing images by deformable receiving field. Finally, the introduction of Attention Mechanism module enables irrelevant background information to be greatly ignored. The method was validated on CDD, LEVIR-CD remote sensing image change detection dataset. The cross-merger ratio was improved by 1.45% and 2.01%, respectively, over the base model. The results show that compared with the classical semantic segmentation methods, this method incorporates the advantages of deformable convolution and dual attention mechanism, and can extract the information of change detection in remote sensing images more accurately.

Key words: remote sensing image, change detection, deformable convolution, attention mechanism

遥感影像变化检测是利用遥感技术获得的多时相影像数据，通过比较不同时间点的影像信息，识别、定量分析地物或场景变化的情况[1-3]。随着遥感技术的不断发展和遥感影像数据的不断增多，遥感图像变化检测被广泛应用于城市规划、土地利用、灾害预测、环境保护等领域[4-7]。传统的遥感影像变化检测方法被提出和应用，如基于差值图像、基于阈值分割、基于纹理特征等方法，这些方法忽视了遥感影像内部的复杂信息和空间上下文信息，导致检测结果不准确或缺乏鲁棒性[8-10]。近年来，深度学习技术的迅速发展，为遥感影像变化检测带来了新的思路与方法[11]。宋文宣[12]等提出了一种遥感图像变化检测的新方法，该方法一定程度上提高了变化检测的精度，但并没有考虑遥感图像中的空间维度的信息，无法满足高精度的遥感图像变化检测。常振量[13]等在DeepLabV3+网络中，引入多尺度特征张量，有效的提高了遥感图像的准确率，但该方法没有考虑到空间维度与通道维度的特征信息。申传庆[14]等构建了遥感图像自动分类模型（ResNet-UNet），实验结果表明，该方法有效提高了地表覆盖的分类精度与效率，但该方法没有考虑到遥

收稿日期：2023-07-15。
项目来源：广东省科技计划项目（2021B1212100003）。
第一作者简介：戴诗涛（1990—），工程师，主要从事航空摄影测量及地理信息数据分析工作，E-mail：262834299@qq.com。
通信作者：张晶（1990—），工程师，主要从事航空摄影测量及地理信息数据分析工作，E-mail：zhangjing-1990@qq.com。

感图像的细小的特征。杨景玉[15]等针对遥感变化检测算法对光谱变化过于敏感的问题，提出了一种基于条件随机场的卷积神经网络方法，实验结果表明，该方法有效提高了变化检测的精度，但没有考虑到遥感图像复杂的特征信息。

综上所述，本文提出一种结合注意力机制模块和可变形卷积的遥感图像变化检测方法(attentional mechanisms and deformable convolutional networks，AMDC-Net），该方法以PSPNet为基础模型，首先，以MobileNet为主干特征提取网络；其次，在金字塔池化模块中引入可变形卷积，能够通过可变形的接收场自适应获取遥感图像前后变化特征信息；最后，引入注意力机制模块，能够有效的忽略无关的背景信息，使模型更加关注遥感图像中发生变化的关键特征信息。

1　研究方法

1.1　网络结构

该模型主要由深度特征提取网络、可变形卷积模块和双重注意力机制模块三部分组成。

1.1.1　主干特征提取网络

PSPNet模型中的主干特征网络采用的是传统的卷积和池化操作，它从低到高依次提取出不同维度的特征层，随着网络层数的加深，从而导致模型的参数量剧增、增加计算的复杂度。因此，本文选择轻量级的MobileNet网络作为AMDC-Net网络模型的主干特征提取网络。MobileNet网络通过深度分离卷积结构来减少模型参数量以及计算的复杂度，有效的保持了模型较高的准确性。相比于传统的卷积操作，在深度分离卷积中输入和输出通道之间进行逐点卷积，然后将每个通道应用于深度卷积中提取特征，有效的减少模型参数数量，从而提高模型的准确性。

1.1.2　可变形卷积

普通卷积是通过规则的网格 R，对每一个采样点的采样值进行权重求和，而可变形行卷积是在标准的卷积的基础上，在采样的过程中增加了一个偏移量。对于图像中心采样点位置 p_0，普通卷积的输出的具体如下：

$$y(p_0)=\sum_{p_n\in R}\omega(p_n)\times\left(p_0+p_n\right) \tag{1}$$

式中，$\omega(*)$ 为采样点的权重；p_n 为格网 R 中所列位置的枚举；x、y 分别为输入、输出特征图。

可变形卷积通过增加偏移量 $(\Delta p_n|n=1,2,3\cdots N)$ 来扩充标准卷积的格网 R，其中 N 为采样点的数量，公式（1）可变为：

$$y(p_0)=\sum_{p\in R}\omega(p_n)\times x(p_0+p_n+\Delta p_n) \tag{2}$$

式中，Δp_n 为偏置量，且偏移量通常为小数。因此，公式（2）通过双线性插值进行求解后变为：

$$x(p)=\sum_q G(q,p)\times x(q) \tag{3}$$

可变形卷积可以将遥感图像中变形最显著的地方通过卷积操作操作来获取偏移量，这种方法可以有效捕获变形前后的变化区域。在PSPNet模型中金字塔池化模块中引入可变形卷积，输出特征图的卷积核与偏移量同时学习遥感图像变化前后的特征信息，而偏移域是由可变性卷积层学习得到偏移量，其中通道维度 $2N$ 对应了 N 个偏移量。为了更有效地学习获得偏移量，采用双线性插值法快速的获取偏移后采样点值，通过改变采样点位置以及感受野，使模型可以对不同大小、形状的遥感图像前后变化的地物进行自适应采样。

1.1.3　注意力机制模块

在PSPNet模型的末端引入了注意力机制，有效忽略无关的特征背景信息，通过反向传播和向前传播来学习获得注意力的权重。注意力机制由两部分组成：通道注意力机制、空间注意力机制。

1）通道注意力机制，具体表达式如下：

$$\begin{aligned}&\boldsymbol{F}_a(F)=S\left\{M\left[\mathrm{MaxPool}(F)\right]+M\left[\mathrm{AvgPool}(F)\right]\right\}=\\&S\left\{\omega_1\left[\omega_0\left(F^c_{\max}\right)\right]+\omega_1\left[\omega_0\left(F^c_{\mathrm{avg}}\right)\right]\right\}\end{aligned} \tag{4}$$

式中，S 为Sigmoid激活函数；F 为输入特征图；ω_0、ω_1 为共享参数；$\boldsymbol{F}_a(F)$ 为通道注意力模块参数矩阵。

2）空间注意力机制，具体表达式如下：

$$\begin{aligned}&\boldsymbol{F}_b(F)=S\left\{f^{7\times7}\left[\mathrm{AvgPool}(F);\mathrm{MaxPool}(F)\right]\right\}=\\&S\left\{f^{7\times7}\left(F^s_{\mathrm{avg}};F^s_{\max}\right)\right\},\end{aligned} \tag{5}$$

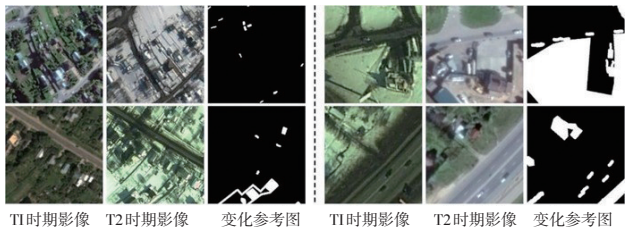
式中，F 为输入的特征图；S 为Sigmoid激活函数；$f^{7\times7}$ 为7×7卷积层；F^s_{avg} 为平均池化后的特征图、$F^s_{\max}$ 为最大池化后的特征图；$\boldsymbol{F}_b(F)$ 为空间注意力模块参数矩阵。

2　实验结果与分析

2.1　实验数据

LEVIR-CD数据集由637个高分辨率卫星影像组成，图像的大小为1 024×1 024像素，该数据集包含了不同城市的大量建筑物，如农村住宅、城市住宅、小

型车库以及高大公寓等。CDD数据集总共包括了3种类型的遥感图像，图像的大小为1 900×1 000像素。受计算机计算性能的限制，将LEVIR-CD数据集的验证集、测试集、验证集比例随机划为7∶2∶1；将CDD数据集切割剪以及数据增强后生成了10 000张图像集，其中验证集、测试集、验证集的比例随机划为8∶1∶1，2个数据集部分样本数据集，如图1所示。



a CCD数据集　　b LEVIR-CD数据集

图1 CDD、LEVIR-CD数据集部分

2.2 评价指标

本文采用召回率、*F*1分数、总体精度、交并比4个指标评价模型的可行性，采用分数对模型的精度进行评估，如果数值越高，说明遥感图像的变化图的精准度越好，4个评价指标的具体表达式如下：

$$R=\frac{TP}{TP+FN} \tag{6}$$

$$F_1=2\times\frac{R\times P}{R+P} \tag{7}$$

$$R_{OA}=\frac{TP+TN}{TP+FP+TN+FN} \tag{8}$$

$$R_{\mathrm{IoU}}=\frac{TP}{FP+TP+FN} \tag{9}$$

式中，*TP*、*FP*、*TN*和*FN*分别为真阳性、假阳性、真阴性和假阴性；*R*为预测像素正确所占真实像素的比例；F_1为召回率、精确度综合指标；R_{OA}为遥感图像中所有的像素被正确分类的比例；R_{IoU}为遥感图像中真实像素和预测像素集合的交集与并集之比。

2.3 实验设置

本研究所有的实验均基于Pytorch1.3.1深度学习框架实现，操作系统为64位的Window10系统，配置NVIDIA GeForce RTX 20808 GB显存，CUDA10.2用于加速计算，迭代次数为50次。

2.4 实验结果与分析

为了验证本文方法在遥感影像变化检测的有效性。在CDD、LEVIR-CD遥感变化检测数据集进行了实验，将本文的方法与FCN、U-Net和HRNet、PSPNet进行对比，结果如图如2、3所示。

2.4.1 LEVIR-CD数据集的结果对比

表1为不同模型在LEVIR-CD数据集的对比结果。通过对比结果可以发现，AMDC-Net模型在总体精度、召回率、交并比以及F1-score比其他4种模型均有所提高。FCN模型的交并比为75.68%，在对比实验中处于最低值，总体的分类精度仅为80.56%；U-Net模型的4个指标数值比FCN模型略有提升；HRNet模型的总体分类精度优于FCN、U-Net模型，交并比指标到达了80.13%；PSPNet模型作为本文方法的基础模型，*F*1分数为90.12%、总体分类精度达到了90.12%、召回率达到了89.67%、IoU为83.35%，四项评价指标都优于前面的对比模型；本文方法在LEVIR-CD数据集表现出较好的性能，与PSPNet模型进行对比交并比分别提高了2.01%，*F*1分数提高了1.12%，总体分类精度提高了3.42%，结果表明，本文方法能够有效的提高遥感图像的变化检测的性能。

表1 LEVIR-CD数据集在不同模型对比结果

模型	总体精度/%	召回率/%	F1-score/%	IoU/%
FCN	80.56	81.74	83.40	75.68
U-Net	85.32	86.59	87.61	79.20
HRNet	89.40	88.54	89.00	80.13
PSPNet	90.12	89.67	90.12	83.35
本文方法	93.56	90.23	91.24	85.36

为了更加直观地对比5种模型在遥感图像变化检测的提取效果，分别选取了4幅不同细节的遥感图像进行了对比，结果如图2所示。FCN模型变化检测有明显的错误，缺失了较多的变化的特征信息；U-Net模型遥感图像的变化检测效果稍微优于FCN模型，但图像的边缘分割较为粗糙，较多的细节特征没有提取处理；HRNet模型也出现较多的漏检的现象，对遥感图像变化检测前后的边缘信息划分还不够精确；PSPNet模型变化检测的精度略有提升，仍存在部分分类错误的现象，对图像的轮廓信息提取精度不够准确；本文方法在LEVIR-CD数据集表现出较好的性能，明显改善了遥感图像变化漏检、错检的问题，同时能够较好地保留遥感图像变化前后的细节特征。

2.4.2 CDD数据集的结果对比

为了进一步验证本文方法的有效性与可行性，本文方法在CDD数据集上进行验证，结果如图3所示。本文方法的遥感图像变化检测的结果明显优于FCN、PSPNet、U-Net、HRNet模型，尤其在小细节特征方面表现出较好的性能。FCN模型变化检测的结果中有着明显的空洞问题，且存在错检、漏检的现象；U-Net模型和HRNet模型遥感图像变化检测效果优于FCN模型，语义分割整体分割效果较好；PSPNet模型的检测的效果优于FCN、U-Net、HRNet模型，能够较好地分割出遥感图像变化检测后特征信息，但还是

存在部分细节特征漏检的问题。本文方法充分利用了可变形卷积、双重注意力机制的优点，使得模型的提取遥感图像前后变化的结果更加接近真实值。

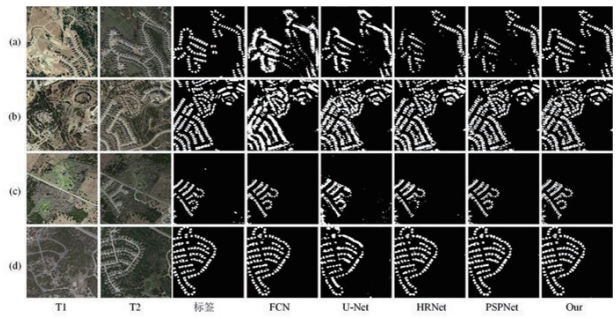


图2　LEVIR-CD数据集在不同模型的对比结果

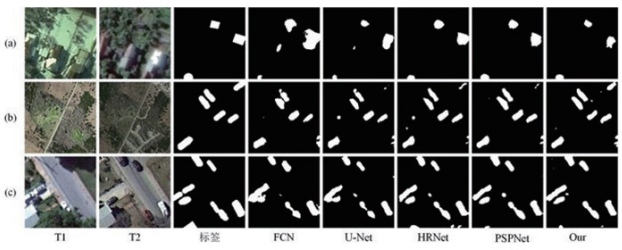


图3　CDD数据集不同模型的对比结果

表2为本文方法与U-Net、FCN、PSPNet和HRNet模型在CDD遥感图像变化检测数据集的对比结果，通过结果可以发现，对比模型中FCN模型的总体分类精度明显低于U-Net、PSPNet、HRNet以及本方法，为81.36%，反映出FCN模型不适用于遥感图像的变化检测，U-Net模型的交并比优于FCN模型；HRNet模型的总体精度为90.21%、召回率为89.36%、IoU为81.94%，*F*1分数为90.98%，4个评价指标达到较好的水平；PSPNet模型的总体精度、召回率、IoU、*F*1分数仅次于本文方法。本方法的总体精度为92.06%，较PSPNet模型提高了1.19%，IoU为83.39%，较PSPNet模型提高了1.45%。通过这些评价指标表明，本文方法在总体精度、召回率、*F*1分数、IoU均有一定的提升，所提的方法可适用于遥感图像的变化检测。

表2　CDD数据集在不同模型对比结果

模型	总体精度/%	召回率/%	F1-score/%	IoU/%
FCN	81.36	81.32	83.96	76.32
U-Net	85.35	86.97	88.33	80.24
HRNet	90.21	89.36	90.65	80.22
PSPNet	90.87	90.25	90.98	81.94
本文方法	92.06	91.38	91.68	83.39

3　结　语

本文提出一种基于注意力机制与可变形卷积的遥感图像变化检测方法。该方法首先选择轻量级的MobileNet作为主干特征提取网络，其次引入可变卷积模块和注意力机制模块构建了神经网络模型（AM-DC-Net）。该方法在CDD、LEVIR-CD数据集上进行验证，结果表明，交并比指标较PSPNet模型分别提高了1.45%和2.01%，相比传统的语义分割方法，2个数据集的遥感图像变化检测结果均证明了本文方法的有效性。下一步研究将结合多尺度机制和Transformer机制进一步提高遥感图像变化检测的精度。

参考文献

[1] 李滔,王海瑞. 基于孪生注意力网络的建设用地遥感影像变化检测[J]. 计算机科学,2023,50(增刊):288-292

[2] 沈鑫甦,嵇灵. 面向遥感地类变化检测的U型深度学习神经网络改进方法[J]. 测绘通报,2023(6):93-97

[3] 陈婕,刘纪平,徐胜华. 增强边缘信息的全卷积神经网络遥感影像建筑物变化检测[J]. 测绘通报,2023(6):61-67

[4] 张祖勋,姜慧伟,庞世燕,等. 多时相遥感影像的变化检测研究现状与展望[J]. 测绘学报,2022,51(7):1 091-1 107

[5] 王超,王帅,陈晓,等. 联合UNet++和多级差分模块的多源光学遥感影像对象级变化检测[J]. 测绘学报,2023,52(2):283-296

[6] 赵祥,王涛,张艳,等. 基于改进DeepLabv3+孪生网络的遥感影像变化检测方法[J]. 地球信息科学学报,2022,24(8):1 604-1 616

[7] 王鑫,张香梁,吕国芳. 基于融合边缘变化信息全卷积神经网络的遥感图像变化检测[J]. 电子与信息学报,2022,44(5):1 694-1 703

[8] 任秋如,杨文忠,汪传建,等. 遥感影像变化检测综述[J]. 计算机应用,2021,41(8):2 294-2 305

[9] 冯权泷,陈泊安,李国庆,等. 遥感影像样本数据集研究综述[J]. 遥感学报,2022,26(4):589-605

[10] 杜培军,林聪,陈宇,等. 多时相遥感影像样本迁移模型与地表覆盖智能分类[J]. 同济大学学报(自然科学版),2022,50(7): 955-966

[11] 王昶,张永生,王旭,等. 基于深度学习的遥感影像变化检测方法[J]. 浙江大学学报(工学版),2020,54(11): 2 138-2 148

[12] 宋文宣,彭代锋. 一种改进全卷积网络的遥感影像变化检测[J]. 遥感信息,2022,37(6):130-136

[13] 常振良,杨小冈,卢瑞涛,等. 基于改进DeepLabv3+的高分辨率遥感影像变化检测研究[J]. 激光与光电子学进展,2022,59(12):493-504

[14] 申传庆,王凯,王文杰. 基于ResNet-UNet的地表覆盖自动分类技术研究[J]. 地理空间信息,2023,21(6):21-23

[15] 杨景玉,贾多科,王阳萍. 基于超像素分割与CRF的高分辨率遥感影像变化检测[J]. 测绘通报,2022(2):25-30