

基于 DeepLabV3+孪生网络的遥感建筑物变化检测^{*}

郭江^{1,2} 辛月兰^{1,2} 王庆庆^{1,2} 王浩臣² 盛月²

(1. 青海师范大学计算机学院 西宁 810001; 2. 藏语智能信息处理及应用国家重点实验室 西宁 810001)

摘要:针对多尺度下小建筑易漏检及建筑物轮廓边界检测精度不足的情况,提出一种基于 DeepLabV3+的双通道孪生网络。首先,为提高分割结果的精确度,同时避免网络层数加深带来的模型过拟合问题,采用改进后的 ResNeXt50(32×4d)作为主干网络来提取特征;其次,针对孪生网络特征融合不充分的问题,设计了基于注意力的双通道融合模块;此外,为提高模型整体信息感知能力,对空洞空间卷积金字塔池化做增强处理;最后,在特征恢复阶段引入特征对齐模块和全连接 CRF 进一步补充和细化分割结果。在 LEVIR-CD 数据集上精确率(precision)、召回率(recall)和 F1 指数分别达到了 0.923 3、0.899 4 和 0.911 2。

关键词:遥感图像变化检测;双通道融合;空洞空间卷积;特征对齐;全连接 CRF

中图分类号: TP751.1 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Remote sensing building change detection based on DeepLabV3+ twin network

Guo Jiang^{1,2} Xin Yuelan^{1,2} Wang Qingqing^{1,2} Wang Haochen² Sheng Yue²

(1. School of Computer Science, Qinghai Normal University, Xining 810001, China;

2. State Key Laboratory of Tibetan Intelligent Information Processing and Application, Xining 810001, China)

Abstract: To address the situation that small buildings are easily missed and the accuracy of building contour boundary detection is insufficient at multiple scales, a dual-channel twin network based on DeepLabV3+ is proposed in this paper. First, in order to improve the accuracy of segmentation results and avoid the overfitting problem of the model caused by the deepening of network layers, the improved ResNeXt50(32×4d) is used as the backbone network to extract features and optimize the computational efficiency of the model; second, to address the problem of inadequate feature fusion in the twin network, an attention-based dual-channel fusion module is designed; in addition, to improve the overall model. Finally, feature alignment module and fully connected CRF are introduced in the feature recovery stage to further complement and refine the segmentation results. The precision, recall and F1 indices reach 0.923 3, 0.899 4 and 0.911 2, respectively, on the LEVIR-CD dataset.

Keywords: remote sensing image change detection; dual channel fusion; null space convolution; feature alignment; fully connected CRF

0 引言

城市化和人口增长的快速发展使得城市建筑物的监测和管理在城市规划、环境保护和灾害响应等领域变得重要,而建筑物变化检测已成为热门的遥感研究方向^[1-3]。近年来,随着深度学习的不断发展,更多的研究者将深度学习方法应用到建筑物变化检测中。吴海平等^[4]利用深

度学习和卷积神经网络(convolutional neural networks, CNN)提升了分类精度并自动提取新建建筑信息。王民水等^[5]基于 DeepLabV3+网络,用随机补片的优化训练策略,将两幅不同时相的遥感影像叠加后作为网络输入,相较 UNet 有更高的精度且内存消耗小。朱节中等^[6]针对小目标容易被漏检的问题,提出基于 U-net++的端到端变化检测模型,将双时序的遥感图像通道融合后输入到深

收稿日期:2023-09-16

^{*} 基金项目:国家自然科学基金(61662062)、青海省自然科学基金面上项目(2022-ZJ-929)资助

度可分离卷积,减少了小目标建筑漏检情况。上述研究方法,将不同时相数据融合后输入网络,不同图像特征之间彼此影响,破坏了原始图像的高维特征,导致检测精度不高。因此,陈良轩等^[7]提出孪生 Unet 网络模型,通过在 Unet 中引入残差结构和注意力机制,提高网络对多尺度建筑物变化的提取能力和泛化能力。高建文等^[8]提出局部—全局语义特征增强的编码—解码网络变化检测模型,将 CNN^[9]与 Transformer^[10]结合,增强信息交流共享,有效提升变化检测精度。上述研究虽然提高了检测精度,但是在复杂情形下还存在小尺度目标漏检,部分建筑物边缘轮廓粗糙等问题。

针对上述问题,本文提出一种基于 DeepLabV3+ 编解码结构的孪生网络遥感建筑物变化检测方法。首先,在编码部分使用 ResNeXt50(32×4d)^[11]作为主干网络进行特征提取,利用 Ghost Module^[12]替换普通卷积以减少计算量提高计算效率;其次,在主干网络尾部加入双通道注意力特征融合模块精炼目标建筑变化信息;同时,设计了增强型空洞空间金字塔池化(enhanced atrous spatial pyramid pooling,EASPP)以加强上下文语义感知;在解码部分引入特征对齐模块(appearance-based feature alignment module,A-FAM)^[13]进行多次高低维特征融合,充分利用浅层特征还原图像边缘和纹理信息,提升模型的细节表达能力;最后使用全连接条件随机场(fully connected conditional random field,CRF)^[14]做后处理,强化特征学习,增强模型泛化能力和整体准确性。

1 本文方法

本文改进的网络基于 DeepLabV3+ 编解码结构组成,整体网络结构如图 1 所示。

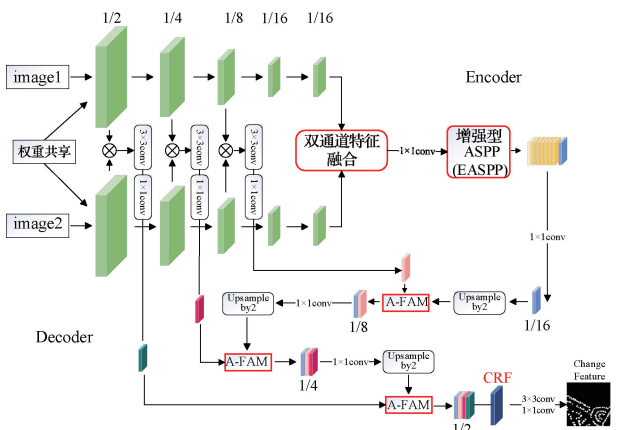


图 1 整体网络结构

编码阶段,通过改进 ResNeXt50 的孪生特征提取网络对不同时相的遥感影像进行特征提取;将提取到的特征输入到双通道特征融合模块,精炼图像变化区域信息;最后,通过 EASPP 加强模型的上下文感知能力,提高分割性能。

解码阶段,在高低维特征跨层融合前引入 A-FAM 模块,减少特征恢复过程中的信息损失,对编码层输出先进行 2 倍上采样,然后与孪生特征提取网络对应层级的特征图融合,经 1×1 卷积降维后,再进行 2 倍上采样与孪生特征提取网络中尺寸为 1/4 的特征图拼接融合,同样的操作,将 1/4 融合后的特征再与孪生特征提取网络中尺寸为 1/2 的特征图拼接融合,通过全连接 CRF 做后处理细化调整。最后使用 3×3 和 1×1 卷积细化特征输出检测结果。三次空间维度的特征融合,可极大提高变化建筑物边缘细节特征恢复。

1.1 基于 ResNeXt50(32×4d)的孪生特征提取网络

为了能够同时处理两种时序影像,提高特征表达能力,避免层数加深带来的模型退化和梯度爆炸的问题,本文采用孪生 ResNeXt50(32×4d)构成 DeepLabV3+ 的编码层进行特征提取,ResNeXt50(32×4d)网络与 ResNet101 参数几乎相同,但效果相当。该模型通过平行堆叠具有相同拓扑结构的 blocks 替代了 ResNet 三层卷积结构,在不增加参数量的情况下提升模型的准确性。ResNeXt50(32×4d)网络参数如表 1 所示。

表 1 ResNeXt50(32×4d)网络参数

Stage	Output	ResNeXt50(32×4d)
Conv1	112×112	7×7,64, stride=2 3×3max pool, stride=2
Conv2	56×56	$\begin{bmatrix} 1\times1,128 \\ 3\times3,128,C=32 \\ 1\times1,128 \end{bmatrix} \times 3$
Conv3	28×28	$\begin{bmatrix} 1\times1,256 \\ 3\times3,256,C=32 \\ 1\times1,512 \end{bmatrix} \times 4$
Conv4	14×14	$\begin{bmatrix} 1\times1,512 \\ 3\times3,512,C=32 \\ 1\times1,1024 \end{bmatrix} \times 6$
Conv5	7×7	$\begin{bmatrix} 1\times1,1024 \\ 3\times3,1024,C=32 \\ 1\times1,2048 \end{bmatrix} \times 3$
	1×1	Global average pool 100-d fc,softmax

1.2 ResNeXt50(32×4d)引入 Ghost module

参考文献[12,15]可知,更多的相似特征会提高神经网络的特征提取能力。因此,本文将 ResNeXt50(32×4d)中的常规卷积替换为 Ghost Module 以增强网络的特征提取效率,相比于常规卷积,Ghost Moudle 可以用更少的参数生成更多的特征图,即在保证网络精度的同时减少网络参数和计算量。常规卷积和 Ghost Module 如图 2 所示。

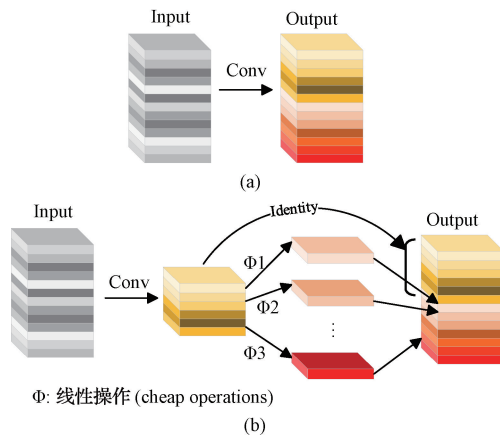


图2 (a)常规卷积层;(b) Ghost Module

Ghost Module 逐个特征图进行深度可分离卷积,其生成的特征图可以通过恒等运算和简单的线性运算获得更多的相似特征,以增强模型的特征提取能力,减少参数量^[16]。使用 Ghost Module 扩展 ResNeXt50 (32×4d) 如表 2 所示。

表 2 Ghost Module 扩展 ResNeXt50 (32×4d)

Layer name
Input
7×7,64, stride=2
MaxPool 3×3
Block1: Ghost Module(32,4d)×3
Block2: Ghost Module(32,4d, stride=2)×4
Block3: Ghost Module(32,4d, stride=2)×6
Block4: Ghost Module(32,4d, stride=2)×3
Average Pooling
Fully Connected

表 2 中,Blocks(1、2、3、4)由 GhostBottleneck 模块堆叠而成。每个 GhostBottleneck 块包括一个 1×1 卷积,以及一个轻量级操作序列,其中包括一个 3×3 卷积和一个 1×1 卷积。分组卷积 32 组,每组通道数相同。

1.3 双通道特征融合模块

深层下采样后的卷积神经网络专注于归纳概括全局图像信息,会忽视对细节影像信息的提取,导致小尺度目标建筑物检测效果差。为此,本文采用了注意力门的机制主动抑制全局影像中无关的背景信息,在改进的主干特征提取网络后部署双通道特征融合模块,旨在增强不同时间变化像素的权重,提升变化特征的表达能力,更好地保留图像中的边界和细节信息,提高特征提取结果的准确性和可靠性,结构如图 3 所示。

首先,由主干网络提取的两组特征时相 f_1 和 f_2 分别通过卷积注意力(convolutional block attention module,

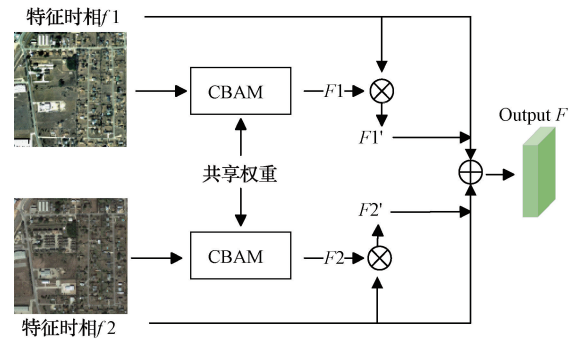


图3 双通道特征融合模块

CBAM)^[17]进行相关性特征计算:

$$F_1 = CBAM(f_1) \quad (1)$$

$$F_2 = CBAM(f_2) \quad (2)$$

其次,相关性特征对特征 f_1 和 f_2 赋以新的权重,捕捉存在差异的建筑物变化信息,约束变化建筑物高度相关的背景信息。

$$F_1' = f_1 \otimes F_1 \quad (3)$$

$$F_2' = f_2 \otimes F_2 \quad (4)$$

最后,为防止赋以新权重的特征 f_1 、 f_2 过拟合,将 f_1 、 f_2 与 F_1' 、 F_2' 相加得到编码部分的输出结果 output F :

$$output F = f_1 + f_2 + F_1' + F_2' \quad (5)$$

本文提出的双通道特征融合模块可以进一步精炼目标变化信息并抑制背景信息,在实验有效性分析可视化结果中取得了不错的结果。

1.4 增强型 ASPP(EASPP)

由于两时相特征图进行差分突出变化的特征会丢失部分细节信息。为此,本文设计了 EASPP,通过在 ASPP^[18]中引入残差连接和金字塔空间池化层,扩大感受野,增强模型的上下文感知能力,从而提高模型的准确性。EASPP 结构如图 4 所示。

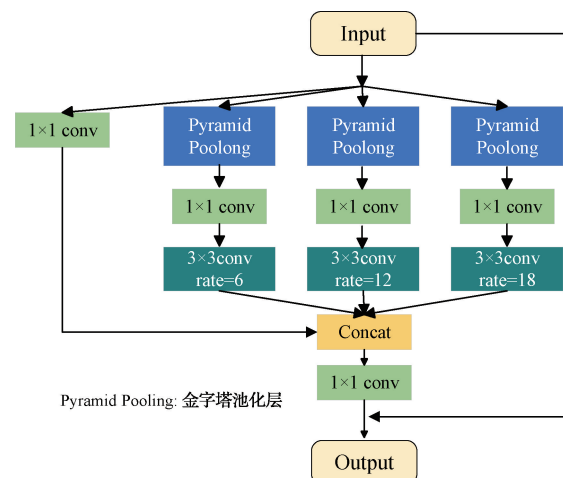


图4 EASPP 结构

改进结构中每个分支首先对输入特征图进行金字塔空间池化操作,金字塔空间池化层根据不同的感受野大小,同时进行多个池化操作,得到不同尺度的池化特征图,更全面地描述输入图像中的对象和上下文信息;之后使用 1×1 的卷积对得到的池化特征图进行通道降维;最后, 3×3 卷积层对降维后的特征进行卷积,以获取更高级别的特征信息。引入残差连接和金字塔空间池化层的EASPP模块结构不仅能够获得多尺度特征信息和语义信息,还能够有效地避免信息丢失和梯度消失的问题,提高分割结果的准确性。

1.5 特征对齐模块

DeepLabV3+网络在解码器部分在预测分割结果时,

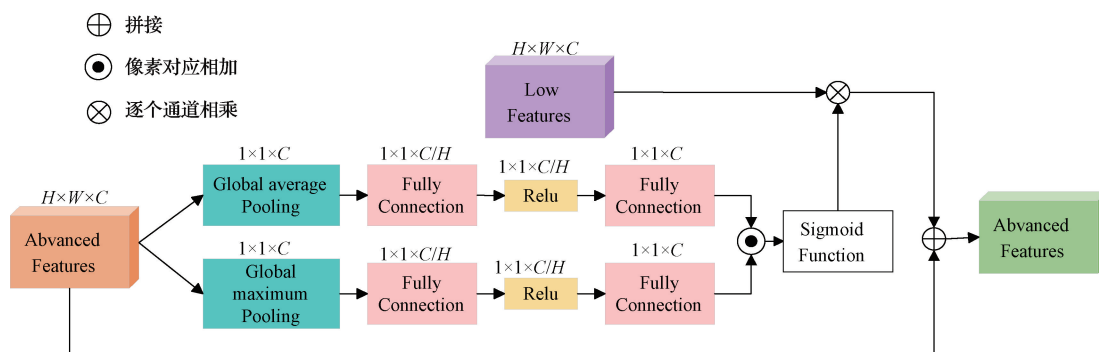


图5 特征对齐模块

该模块高维特征的两条并行注意力通道包含压缩、激励两部分。压缩部分由全局平均池化和全局最大值池化统计特征图通道信息,并获得 $1 \times 1 \times C$ 的向量 $\mathbf{A}$;激励部分由全连接层获得通道间的关系描述,公式表达为:

$$\mathbf{S} = \sigma[\omega_2 \delta(\omega_1 \mathbf{A})] \quad (6)$$

首先,全连接层降低通道数为 C/H ,权重为 ω_1 ,然后通过激活函数 δ 后由权重为 ω_2 的第2个全连接层恢复通道数;其次,两路特征向量合并相加经过sigmoid函数 σ 生成权重向量 $\mathbf{S}$,并且与低维特征 f 通道相乘,得到通道权重特征 $\mathbf{F}$;最后,权重特征 $\mathbf{F}$ 与高维特征 f' 合并输出对齐后的特征。

$$\mathbf{F} = f \cdot \sigma(\omega_2 \delta(\omega_1 [\text{Avgpool}(f') + \omega_2 \delta(\omega_1 [\text{Maxpool}(f')])]) \quad (7)$$

1.6 全连接 CRF 后处理

为了进一步提高模型的精度和平滑提取到的变化区域边界,参考文献[19-20]的图像优化方法,引入全连接CRF作为后处理阶段,将解码器部分的输出作为CRF的输入,对模型最后分割结果进行更加精细的调整。

在后处理阶段中,全连接CRF使用二元势函数来解释像素间的关系,给像素关系紧密的两个像素赋予相同的类别标签,而关系相差很大的两个像素会赋予不同的类别标签,然后通过标签判断像素间关系。其能量函数有一元势能和二元势能组成:

直接将4倍上采样的低维特征和高维特征进行特征融合会增加部分无关空间信息,并且没有考虑不同维度特征对齐程度,导致分割精度的缺失。在变化检测任务中,由于网络层数的增加,特征图涵盖更多的抽象语义信息,二者直接拼接会引入噪声,影响分割结果。

因此,本文在编码器部分引入A-FAM,使得高低维度特征能够更好地相互补充。A-FAM通过全局平均池化、全局最大值池化进行自适应学习通道权重和减少无关噪声。其中,变化程度不同的建筑物得到的特征空间和感受野的大小不同,且处于不同层次或不同分支的网络中,因此进行特征对齐可以满足本文变化检测需求。A-FAM结构如图5所示。

$$E(x | I) = \sum_i \phi_u(x_i) + \sum_{i,j} \phi_p(x_i, y_j) \quad (8)$$

其中,前半部分为一元势函数用于衡量当像素点 i 的观测值为 y_i 时,该像素点属于类别标签 x_i 的概率,来自卷积神经网络的后端输出;后半部分为二元势函数用于衡量两事件同时发生的概率 $p(x_i, y_j)$ 。

全连接CRF利用变化建筑物中相邻像素之间的相互关系进行统一优化,抑制模糊边缘和加强辨别区域轮廓,使得分割结果更加平滑和准确。

1.7 损失函数

本文采用裁剪过的遥感影像进行训练,由于变化区域的面积过小会导致正负样本不均衡,因此采用加权交叉熵损失函数来缓解这个问题。每层网络检测到的变化像素点形成变化图像,表达式如下:

$$\tilde{\mathbf{Y}} = [\tilde{y}_k, k = 1, 2, 3, \dots, H \times W] \quad (9)$$

式中: $\tilde{y}_k$ 表示每层变化图像 $\tilde{\mathbf{Y}}$ 中的像素点; H 和 W 分别表示每层变化图像的高和宽。

然后通过权重均衡正负样本进行加权交叉熵损失函数(LossWCE)的计算:

$$\text{Loss}_{\text{wce}} = \frac{1}{H \times W} \sum_{k=1}^{H \times W} w(c) \cdot \left(\log \left(\frac{\exp(\tilde{y}(k)(c))}{\sum_l \exp(\tilde{y}(k)(l))} \right) \right) \quad (10)$$

式中: $w(c)$ 代表权重均衡正负样本,且 c 的值为1或0表

示检测图像中改变的像素和没改变的像素; $\exp()$ 为指数损失。

在高分辨率遥感图像中,可能存在少量的变化区域被检测出,导致训练时损失函数出现局部值偏小。为了解决这个问题,可以增加正样本的权重。本文采用 Dice 损失函数来衡量预测正值区域与标签正值区域的相似程度。Dice 损失定义为:

$$Loss_{dice} = 1 - \frac{2 \cdot Y \cdot \tilde{Y}'}{Y + \tilde{Y}'} \quad (11)$$

式中: Y 为标签图; $\tilde{Y}'$ 为 $\tilde{Y}$ 经过 softmax 处理。

综上,通过结合加权交叉熵损失函数和 Dice 损失函数,整体损失函数如下:

$$LOSS_{\text{总}} = \sum_{i=1}^4 w_i \sum_{j=1}^4 (LOSS_{wce}^j + LOSS_{dice}^j) \quad (12)$$

这样可以考虑到不同语义层的权重和不同损失函数的贡献,使得训练更加全面。

2 实验过程及分析

2.1 数据集

实验使用 LEVIR-CD 数据集,包括各种类型的建筑,如别墅住宅、高层公寓、小车库和大型仓库。包括 637 个 $1\,024 \times 1\,024$ 的高分辨率的谷歌地球图像补丁,时间跨度在 5~16 年,涵盖了显著的土地利用变化,按照 7:1:2 的比例设置训练集、验证集、测试集。由于实验显存限制,将每个样本裁剪为 8 个 512×512 图像,获得 3 568、512、1 016 的 512×512 图像分别用于训练、验证、测试。

2.2 实验环境

实验平台软件配置为 Windows10 系统、深度学习框

架为 Pytorch,内存 43 GB;GPU 为 GeForce RTX 1080ti; batchsize 设置为 8,学习率初始化为 0.000 1,Epoch 为 120,优化器算法选择为 Adam。

2.3 评价指标

使用精确率(precision)、召回率(recall)和 F1 指数来评估分析本文改进模型分割结果。计算公式如下:

$$Precision = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (13)$$

$$Recall = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (14)$$

$$F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (15)$$

式中: N_{TP} 为预测为建筑的像素数; N_{FN} 为预测为非建筑的像素数; N_{FP} 为将非建筑预测为建筑目标像素数。

2.4 数据集风格同化

由于不同时相间的时间相隔过大,前后差异较大会导致变化建筑物的差异减弱,造成误检和分割错误。为突出变化区域的建筑物差异,尽可能保持两幅不同时期遥感图像的风格一致性,利用直方图匹配实现图像风格的统一,将后时相像图像作为参考图像,对后时相图像做直方图匹配,使得两幅不同时期高分辨率遥感图像具有类似的色彩,从而降低变化区域对准确率的影响。如图 6 所示,待匹配的原始遥感影像的整体色彩偏亮,匹配过后,两幅图像的整体色调差距变小。在双时相影像风格差异比较大的时候,实验结果会有提升。

2.5 实验结果及可视化

1) 消融实验

为验证本文选取主干网络的有效性,在 LEVIR-CD

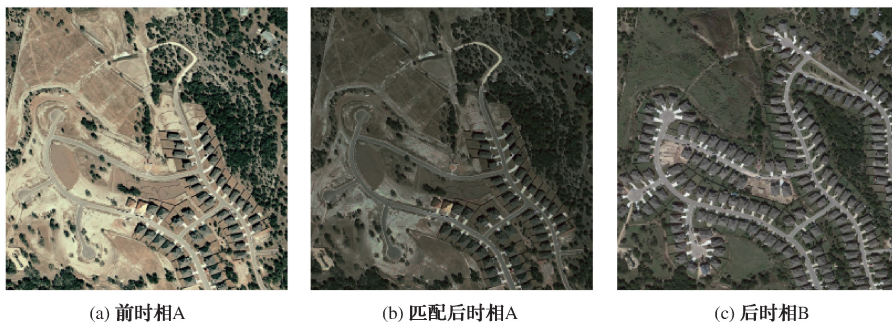


图 6 数据匹配结果对比

数据集上进行消融实验。表 3 为不同孪生主干特征提取网络对应的各项指标和参数。表 4 为数据集在直方图同化前后各模块消融实验结果。

可以看出 ResNeXt50(32×4d)各项指标均超过 Res-Net50,ResNet101 的情况下,参数量只有 ResNet101 的 1/2 左右;而将普通卷积替换为 Ghost module 的 ResNeXt50(32×4d)在参数量小幅度增加的情况下检测指标最好。

表 3 不同孪生主干特征提取网络性能和参数量对比

Backbone	F1 指数	Precision	Recall	Params
ResNet50 ^[21]	0.801 0	0.845 8	0.759 9	23.4×10^6
ResNeXt50 (32×4d) ^[11]	0.829 9	0.871 0	0.792 6	22.0×10^6
ResNet101 ^[21]	0.815 7	0.855 7	0.779 2	42.0×10^6
ResNeXt50(32×4d) with Ghost module	0.849 4	0.897 5	0.806 3	25.5×10^6

表 4 消融实验评价指标对比

	双通道特征融合	EASPP	特征对齐	全连接 CRF	F1 指数	Precision	Recall
直方图匹配	—	—	—	—	0.849 4	0.897 5	0.806 3
	✓	—	—	—	0.878 0	0.907 7	0.850 2
	✓	✓	—	—	0.885 8	0.908 1	0.864 7
	✓	✓	✓	—	0.901 0	0.920 1	0.882 4
	✓	✓	✓	✓	0.911 2	0.923 3	0.899 4
未直方图匹配	—	—	—	—	0.848 3	0.865 3	0.832 1
	✓	—	—	—	0.876 1	0.892 7	0.860 2
	✓	✓	—	—	0.882 0	0.898 4	0.866 0
	✓	✓	✓	—	0.899 3	0.918 2	0.881 3
	✓	✓	✓	✓	0.910 8	0.921 6	0.900 2

实验结果表明,在数据集进行直方图匹配后,对主干网络改进的基础网络 Precision、Recall 和 F1 指数分别为 0.897 5、0.806 3、0.849 4;添加双通道特征融合模块后 Precision、Recall 和 F1 指数分别提升了 0.010 2、0.043 9、0.028 6;同时添加双通道特征融合模块和增强型 ASPP 后,评价指标 Precision、Recall 和 F1 指数在添加双通道特征融合模块的基础上分别提升了 0.000 4、0.014 5、0.007 8;当同时添加双通道特征融合模块、EASPP 并引入特征对齐模块后,评价指标 Precision、Recall 和 F1 指数在前者基础上分别提升了 0.012 0、0.017 7、0.015 2;在前三者的基础上引入全连接 CRF 后,评价指标 Precision、Recall 和 F1 指数分别提升 0.003 2、0.017 0、0.010 2,经过整体改进后的网络相较于原始网络的 Precision、Recall 和 F1 指数分别提升了 0.025 8、0.093 1、0.061 8。数据集未进行直方图匹配时,整体指标均低于直方图匹配后的指标,说明对数据集进行直方图匹配后可以进一步提升模型的精确度。综合分析可知,本文网络的改进是有效的,充分说明了本文方法的优越性。同时本文在 LEVIR-CD 测试集上实验,消融实验可视化结果如图 7 所示。

观察图 7 可知,仅对主干网络进行改进,检测出的变化建筑会出现严重的黏连,检测结果不独立且边缘粗糙不平整,甚至出现错检和漏检;添加双通道特征融合模块后,减少了小目标错检和漏检,但整体检测效果还不是很好;在 EASPP 和特征对齐模块的加持下基本解决了错检和漏检的问题,有效改善了目标建筑物的边界黏连,提高了整体精度,说明 EASPP 和特征对齐模块对变化特征的精准提取和恢复轮廓细节作用明显;添加了全连接 CRF 的检测分割效果相较于前四者更好,建筑物像素定位更加精准,边界更清晰平滑,与标签图像更接近。

2)对比实验

为验证本文方法在变化检测任务中的有效性,本文将经典的变化检测网络 BIT、STANET 和 Unet++、Deep-

LabV3+、LGE-NET 与本文模型在 LEVIR-CD 验证集上进行对比,结果如表 5 所示。

表 5 不同模型评价指标对比

Method	F1 指数	Precision	Recall
Unet++ ^[6]	0.875 0	0.896 0	0.873 0
DeepLabV3+	0.848 3	0.865 3	0.832 1
BIT ^[22]	0.899 6	0.905 0	0.894 2
STANET ^[23]	0.881 0	0.850 0	0.914 0
LGE-NET ^[8]	<u>0.910 6</u>	<u>0.922 8</u>	0.898 7
本文	0.911 2	0.923 3	<u>0.899 4</u>

注:最优值加粗;次优值下划线表示

Unet++通过密集跳跃获得不同的语义级别多尺度特征图,但对小尺寸建筑不敏感,边界存在模糊问题,其检测的 F1 指数为 0.875 0;DeepLabV3+网络中的 ASPP 特征采样不够密集,上下文信息有限和建筑物边界定位不准确,使检测完整性较差,此网络的 F1 指数为 0.848 3;BITH 和 STANET 均为建筑物变化检测网络,BIT 通过二进制编码和不变性理论实现变化检测,STANET 则通过空间和时间注意力机制来提高检测准确性,其 F1 指数在 LEVIR-CD 验证集上分别为为 0.899 6、0.881 0;LGE-NET 采用局部—全局编码结构,有效地提取变化区域局部和全局信息,且通过多个卷积层和池化层学习高级语义特征,其检测的 F1 指数为 0.910 6,本文方法的 F1—score、Precision 均是最优值,而 Recall 仅次于 STANET,是因为在测试集中包含了大量与训练集不同的特殊样本或场景,同时测试样本设置较少,算法在测试过程中没有充分适应,图像中的一些重要特征没有被良好地捕获,从而导致相对较低的 Recall 值。但相较于目前最新检测方法,本文 Precision、Recall 和 F1 指数分别提高了 0.005、0.000 7、0.000 6。

本文方法与 Unet++、DeeoLabV3+、STANET 和 LGE-NET 的可视化分割结果如图 8 所示。其中,Deep-

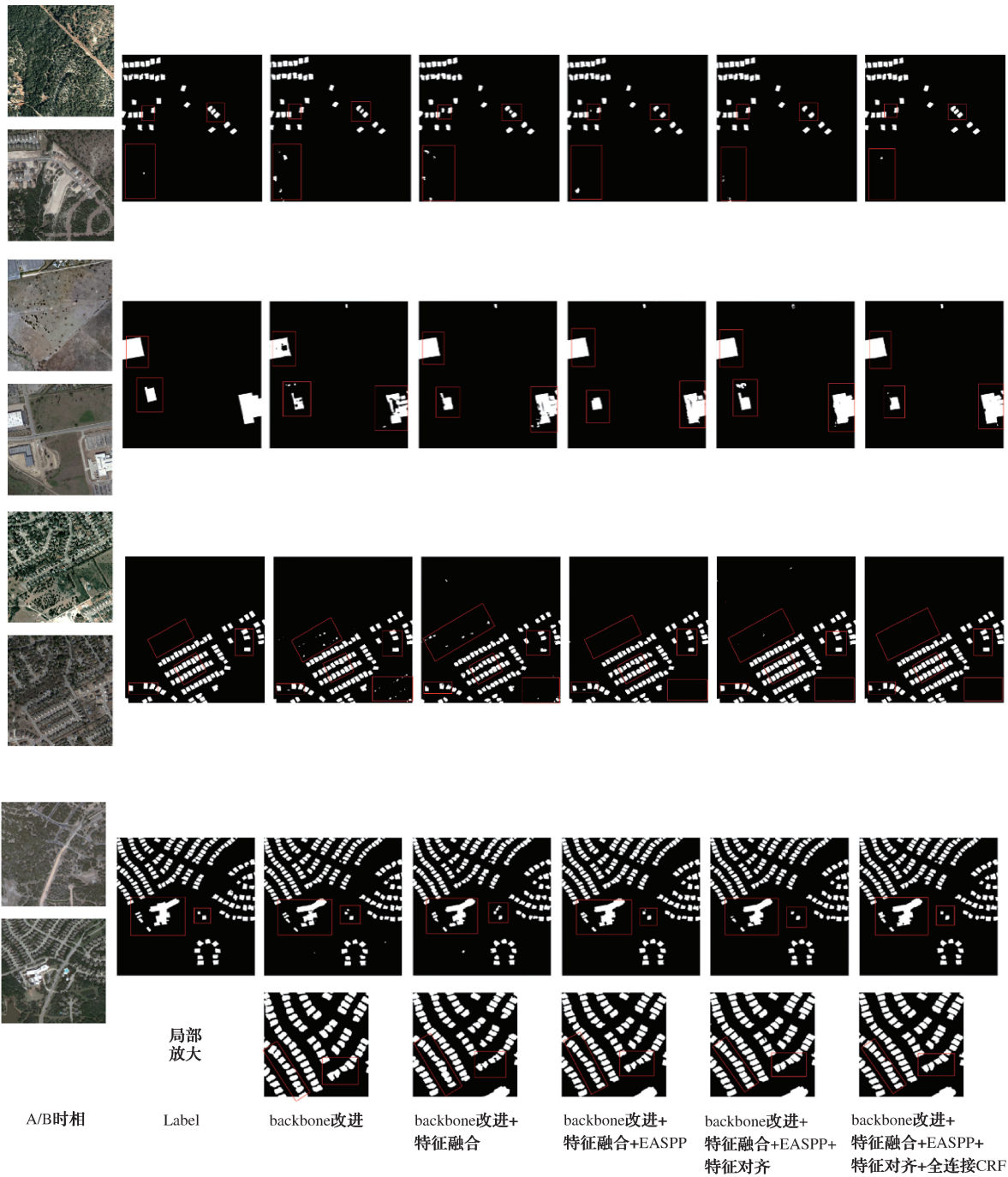


图 7 消融实验检测结果

LabV3+和 Unet++ 的检测效果较差,存在小建筑物之间严重的黏连问题,同时出现错分和漏分现象,建筑物轮廓不够平滑,对于复杂建筑物的检测不完整。STANET和 LGE-NET 在整体检测效果上有所提升,减少了建筑物之间的黏连程度并提高了检测精度,但仍存在一定程度的

错分和漏分现象。相比之下,本文方法的检测效果较好,能更完整地检测出多边复杂建筑和小目标建筑,黏连程度较低,边缘更清晰。本文通过改进结构,充分利用多尺度特征提高了整体预测的精准度,可视化结果优于其他改进模型充分证明本文方法的有效性。

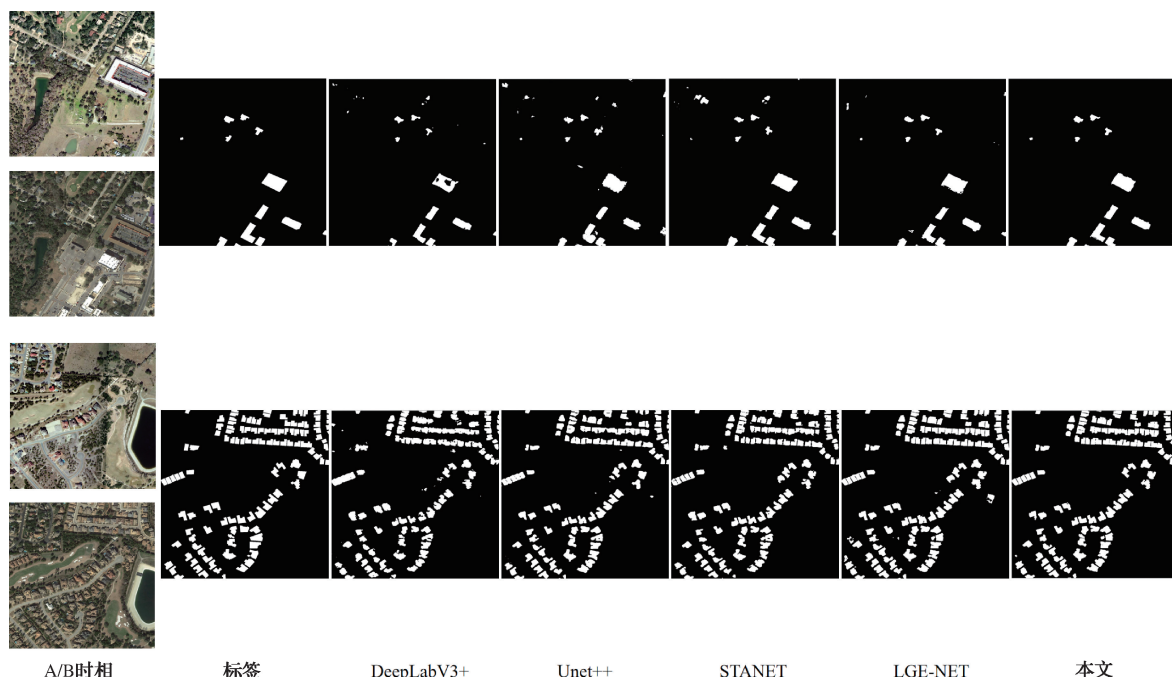


图8 LEVIR-CD数据集检测结果对比

3 结 论

本文提出了一种适用于遥感建筑物变化检测的DeeLabV3+孪生网络,使用 Ghost Module 的 ResNeXt50 (32×4d)作为主干网络提取特征,并通过注意力机制的双通道融合模块提高分割结果的精准度。此外,对空洞空间卷积池化金字塔进行增强处理以提升信息感知能力,并引入特征对齐模块和全连接 CRF 对检测结果进行补充和细化。实验结果表明,本文方法能够有效地改善小建筑物变化检测的表现。在未来工作中将重点研究用轻量化的结构提高分割结果的准确性和鲁棒性。

参 考 文 献

- [1] 唐文博. 基于卷积神经网络的高分辨率多光谱遥感图像上的城区建筑物变化检测技术[D]. 杭州:浙江大学, 2023.
- [2] 唐璎, 刘正军, 杨懿, 等. 基于特征增强和 ELU 的神经网络建筑物提取研究[J]. 地球信息科学学报, 2021, 23(4): 692-709.
- [3] SHI W, MIN Z, ZHANG R, et al. Change detection based on artificial intelligence: State-of-the-art and challenge[J]. Remote Sensing, 2020, 12(10): 1688.
- [4] 吴海平, 黄世存. 基于深度学习的新增建设用地信息提取试验研究——全国土地利用遥感监测工程创新探索[J]. 国土资源遥感, 2019, 31(4): 159-166.
- [5] 王民水, 孔祥明, 陈学业, 等. 基于随机补片和 DeeLabV3+ 的建筑物遥感图像变化检测[J]. 吉林大学学报(地球科学版), 2021, 51(6): 1932-1938.
- [6] 朱节中, 陈永, 柯福阳, 等. 基于 Siam-UNet++ 的高分辨率遥感影像建筑物变化检测[J]. 计算机应用研究, 2021, 38(11): 3460-3465.
- [7] 陈良轩, 于海洋, 李英成, 等. 一种融合注意力机制的建筑物变化检测模型[J]. 测绘科学, 2022, 47(4): 153-159.
- [8] 高建文, 管海燕, 彭代锋, 等. 基于局部-全局语义特征增强的遥感影像变化检测网络模型[J]. 地球信息科学学报, 2023, 25(3): 625-637.
- [9] MALINI A, PRIYADHARSHINI P, SABEENA S. An automatic assessment of road condition from aerial imagery using modified VGG architecture in faster-RCNN framework[J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(6): 11411-11422.
- [10] JAMALI A, MAHDIANPARI M, BRISCO B, et al. 3DUNetGSFormer: A deep learning pipeline for complex wetland mapping using generative adversarial networks and swin transformer [J]. Ecological Informatics: An International Journal on Ecoinformatics and Computational Ecology, 2022, 72:101904.
- [11] XIE S, GIRSHICK R, PIOTR D, et al. Aggregated residual transformations for deepneural networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 1492-1500.
- [12] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589.

- [13] 孟俊熙, 张莉, 曹洋, 等. 基于 Deeplabv3+ 的图像语义分割算法优化研究[J]. 激光与光电子学进展, 2022, 59(16): 161-170.
- [14] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [15] 符樊博, 高玮玮, 单明陶, 等. 融合注意力机制与重影特征映射的无人机交通场景目标轻量级语义分割[J]. 电子测量与仪器学报, 2023, 37(3): 21-28.
- [16] 张定军, 廖明潮, 高拉劳. 基于 GhostNet 的轻量级桥梁裂缝图像语义分割算法[J]. 公路, 2023, 68(4): 246-255.
- [17] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [18] PENG X, YIN Z, YANG Z. Deeplab_v3_plus-net for image semantic segmentation with channel compression [C]. 2020 IEEE 20th International Conference on Communication Technology (ICCT). IEEE, 2020: 1320-1324.
- [19] 陈丽勇, 孙显, 王宏琦. 一种基于 TSVM-MRF 的变化检测方法[J]. 国外电子测量技术, 2015, 34(7): 32-36.
- [20] 顾佼佼, 翟一琛, 姬嗣愚, 等. 基于 BERT 和知识蒸馏的航空维修领域命名实体识别[J]. 电子测量技术, 2023, 46(3): 19-24.
- [21] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [22] 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: 1-14.
- [23] 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.

作者简介

郭江, 硕士研究生, 主要研究方向为图像处理与可视化。

E-mail: 1366486324@qq.com

辛月兰(通信作者), 博士, 教授, 博士生导师, 主要研究方向为图像处理、模式识别与智能信息处理。

E-mail: xinyue001112@163.com

王庆庆, 硕士研究生, 主要研究方向为图像处理与可视化。

E-mail: 1595981173@qq.com

王浩臣, 硕士研究生, 主要研究方向为图像处理与可视化。

E-mail: 1114601734@qq.com

盛月, 硕士研究生, 主要研究方向为图像处理与可视化。

E-mail: 2270330841@qq.com