

改进 SeMask 主干网络的高分辨率遥感影像变化检测模型^①



陈海文, 王璐, 徐中荣, 崔璐璐, 罗维

(青海大学 计算机技术与应用系, 西宁 810016)

通信作者: 王璐, E-mail: wlqgl@126.com

摘要: 随着城市化进程的加速和人口不断增加, 土地资源的利用和管理变得愈发重要. 高分辨率遥感影像技术的发展为土地覆盖类别变化检测提供了新的途径. 目前, 多数遥感影像变化检测任务主要针对显著建筑物的变化检测, 缺少对土地覆盖类别变化检测任务的研究, 本研究基于公开数据集, 对更多土地覆盖类别变化情况进行标注. 在原语义分割主干网络的基础上结合孪生网络结构, 提出适用于土地覆盖类别变化检测任务的检测模型, 该模型在网络的特征提取阶段加入变化引导模块, 以辅助网络关注两时相影像中的变化信息, 并在网络不同阶段加入通道信息交互模块, 以增强不同特征图的信息融合. 同时, 在特征提取阶段最后一层加入特征对齐模块, 以缓解下采样过程导致的特征偏移. 在土地覆盖类别变化检测数据集上的实验结果表明, 本文提出的方法可以有效提取影像中的变化信息, 并提高分割精度.

关键词: 深度学习; 遥感影像; 变化检测; Transformer

引用格式: 陈海文, 王璐, 徐中荣, 崔璐璐, 罗维. 改进 SeMask 主干网络的高分辨率遥感影像变化检测模型. 计算机系统应用, 2023, 32(11): 167-174. <http://www.c-s-a.org.cn/1003-3254/9304.html>

Improvement of High-resolution Remote Sensing Image Change Detection Model Based on SeMask Backbone Network

CHEN Hai-Wen, WANG Lu, XU Zhong-Rong, CUI Lu-Lu, LUO Wei

(Department of Computer Technology and Application, Qinghai University, Xining 810016, China)

Abstract: As urbanization accelerates and the population continuously increases, the utilization and management of land resources have become increasingly important. The development of high-resolution remote sensing technology provides a new approach for detecting land cover changes. Currently, most remote sensing image change detection tasks mainly focus on detecting significant changes in buildings, and there is a lack of research on detecting changes in land cover categories. In this study, based on a public dataset, more land cover change scenarios are annotated. Combining the original semantic segmentation backbone network with a Siamese network structure, this study proposes a detection model suitable for tasks of detecting changes in land cover categories. The model incorporates a change guidance module in the feature extraction stage to assist the network in focusing on change information in the two temporal images. A channel information interaction module is added at different stages of the network to enhance the fusion of information from different feature maps. Additionally, a feature alignment module is added to the last layer of the feature extraction stage to alleviate feature offset caused by downsampling. Experimental results on a dataset of detecting changes in land cover categories demonstrate that the proposed method can effectively extract change information from the image and improve the segmentation accuracy.

^① 基金项目: 2021 年度青海省科技厅自然科学基金青年基金 (2021-ZJ-952Q)

收稿时间: 2023-05-07; 修改时间: 2023-06-06; 采用时间: 2023-06-26; csa 在线出版时间: 2023-09-15

CNKI 网络首发时间: 2023-09-19

Key words: deep learning; remote sensing image; change detection; Transformer

随着城市化进程的加速和人口的不断增加,土地资源的利用和管理变得愈发重要.其中,土地覆盖类型变化是土地资源管理和环境保护的核心问题之一,通过检测土地覆盖类型变化,人们可以更好地了解土地资源的利用情况和环境变化的发展趋势.然而,传统的土地覆盖类别变化检测方法需要投入大量的人力和物力,并且存在着精度低、效率低等问题.高分辨率遥感影像技术的发展为土地覆盖变化检测提供了新的途径.

由于高分辨率遥感影像可以提供更为细致的地表信息,因此被广泛应用于城市规划^[1],生态环境监测^[2],生态价值估算^[3],农业生产估值^[4]等诸多领域,为各行业提供强有力的数据支持和技术服务.遥感影像变化检测是利用同一场景、不同时相的两幅或多幅遥感影像进行定量分析从而确定地物变化情况的过程.它的主要价值体现在以下几个方面:一是环境检测和资源管理,土地覆盖类型的变化情况与生态环境的稳定性、健康度息息相关,通过对遥感影像土地覆盖类型变化信息的提取与分析,可以帮助监测环境变化.二是土地利用规划和决策支持,土地覆盖变化检测可以帮助制定土地利用规划和决策支持.三是经济发展和市场分析,土地覆盖变化检测可以用于评估经济发展和市场趋势,例如城市房地产开发和农业生产等.因此,遥感图像土地覆盖类别变化检测对人类社会发展和生态环境保护具有重要的意义.

遥感影像土地覆盖变化检测主要包括3个步骤:图像预处理、差异图生成和差异图分析.图像预处理主要包括影像去噪、几何矫正、大气矫正、影像配准等步骤.传统机器学习方法的差异图生成主要使用图像差值法^[5]、图像比值法^[6]和基于特征检测法,如纹理^[7]、颜色^[8]等特征.差异图分析方法主要有支持向量机(support vector machine)^[9]和贝叶斯网络分类模型^[10]等.基于传统机器学习的变化检测方法差异图生成和分析是独立的两部分工作,缺乏自动化的工作方式,而且基于有限特征表示的差异图,只能提取到图像的浅层特征,不利于提高变化检测精度.

随着深度学习方法的出现,变化检测任务进入高速发展阶段.检测过程被统一为端到端的模式,简化检测流程,同时能够提取更深层次的图像特征,有利于检

测精度的提高.

遥感影像变化检测任务是一种二元语义分割任务,通过提取同一地点两个时相影像特征,分割出变化与未变化两个类别.许多学者在语义分割主干网络基础上对两幅影像特征进行提取,并研究多种特征融合方式完成变化检测任务.Daudt等人^[11]首次将全卷积神经网络应用在变化检测任务,在特征提取阶段分别使用两时相影像合并输入网络的方式和两幅影像的特征提取网络参数共享的方式提取影像特征,在上采样阶段分别在两幅影像不同阶段将特征图合并上采样和特征图做差取绝对值上采样两种方式,研究了以上几种不同网络结构在OSCD数据集^[12]和AC数据集^[13]中的分割表现.STA-Net^[14]使用ResNet^[15]网络作为主干网络提取两个时相影像特征,在特征融合阶段,基于时空注意力模块和金字塔时空注意力模块融合两时相特征,并在LEVIR-CD^[16]建筑物变化检测数据集中取得87.3%的F1分数.DASNet^[17]提出基于通道注意力构建通道间的关系,并在传统损失函数的基础上增加改变像素和未改变像素的权重来优化类别损失函数更加倾向关注变化的类别.季顺平等^[18]在全卷积神经网络的基础上结合空洞卷积提出全空洞卷积神经网络,能够顾及矢量数据中地物不同尺度和不同精细程度,在武汉市土地覆盖变化实验中变化检测精度达到74.1%.

尽管上述网络在特征提取阶段能够同时提取两时相影像特征形成差异特征,全自动地完成变化检测任务,但由于卷积运算的固有局限性,基于卷积神经网络(CNN)的主干网络对图像全局进行建模时,很难学习到距离相对较远的语义信息.近年来,受到Transformer^[19]在自然语言处理领域取得巨大成功地启发,很多学者借鉴Transformer网络结构,并将其应用在计算机视觉领域.Dosovitskiy等人^[20]提出Vision Transformer(ViT)模仿自然语言处理领域句子的输入方式,将图片裁切为互不重合的图像块序列输入网络中,完成图像的全局建模.在一些大型图像数据集中取得了与CNN网络相当的分精度.但直接对图像全局建模也带来了计算量过大的问题,Liu等人^[21]提出Swin Transformer采用多个窗口内部计算自注意力,不同窗口之间进行

信息交流的方式,对图像分层建模,降低模型计算量. Swin Transformer 主干网络被广泛应用图像分割领域,对比其他方法在多个数据集中有较大的精度优势. Jain 等人^[22]在 Swin 的基础上提出 SeMask Transformer,针对语义分割任务,在主干网络的不同阶段加入语义模块帮助网络在特征提取阶段关注图像的语义特征.

本研究设计了适用于遥感影像土地覆盖类别提取任务的变化检测网络,采用特征提取能力更强的 SeMask Transformer 主干网络,同时为了使网络在特征提取阶段关注两幅影像的变化信息,在网络的不同阶段利用语义模块引导网络关注两幅影像的变化特征.最后,为了帮助两幅影像的信息融合并缓解下采样过程中导致的特征偏移问题,增加了信息交互模块和特征对齐模块.此外,目前遥感影像变化检测研究多是针对显著建筑物变化检测任务,相关数据集如 LEVIR-CD^[16]、S2Looking^[23]等,而遥感影像土地覆盖类别变化检测任务的数据集相对较少.本文构建了遥感影像土地覆盖类别变化检测数据集以支持土地覆盖类别变化检测方法的研究.具体而言,本文的贡献可以概括为:(1)在 Google 显著建筑物变化检测数据集^[24]的基础上对林地、草地、水体和裸地等土地覆盖类型变化情况进行再标注,构建 Google 土地覆盖类别变化检测数据集.(2)结合孪生网络结构^[25]采用参数共享的方式提取两时相影像特征,在每个阶段加入变化引导模块,帮助网络关注两幅影像中变化信息.(3)在主干网络中加入通道信息交互模块和特征对齐模块,增强两幅影像的特征的信息交互,并缓解特征提取过程中导致的特征偏移问题.

1 研究流程

本文提出的基于 SeMask 变化检测模型,整体流程可以分为3个阶段:(1)数据预处理,在 Google 显著建筑物变化检测数据集的基础上对更多土地覆盖变化类别进行标注包括水体、道路、草地和林地等,并将数据集划分为训练集、验证集和测试集.(2)模型设计,针对变化检测任务特点,在原始网络的基础上对主干网络结构和损失函数进行改进,增强边缘检测效果和两时相影像的特征融合.(3)模型训练,将标注后的数据集在模型上进行多轮迭代训练,过程中我们选用多个指标来评估模型的性能.如图1所示.

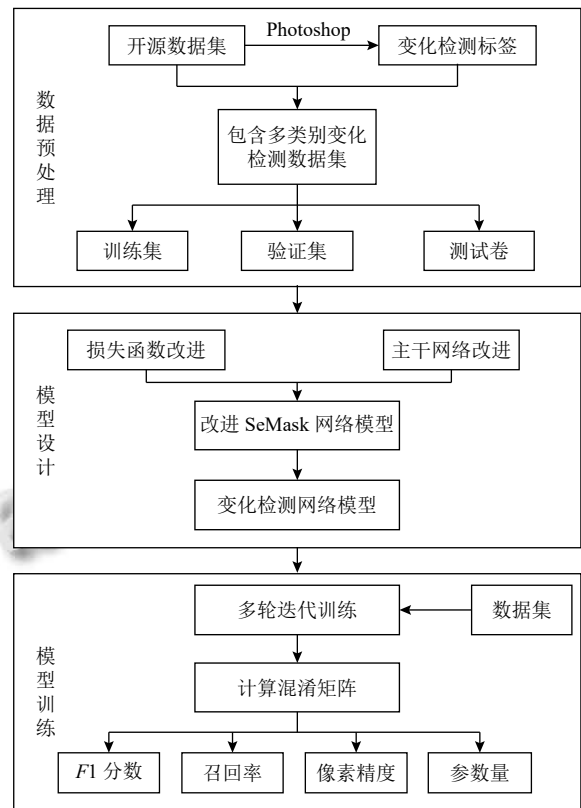


图1 研究流程

2 数据集构建

目前针对土地覆盖类别变化检测数据集较少,多数变化检测数据集仅标注了建筑物的变化情况,为了完成多种土地覆盖类别变化检测任务,我们在原始建筑物变化检测数据集的基础上对多种土地覆盖类型的变化进行标注,构建了 Google 土地覆盖类别变化检测数据集.原始 Google 数据集^[24]是一个显著建筑物变化检测数据集,主要标注了双时相影像大型建筑物的变化情况.原始影像数据来自 2006–2019 年间中国广州市的郊区,共 19 对包含了季节变化的图像对,大小为 $1006 \times 1168 - 4936 \times 522$ 像素,包含红、绿、蓝这 3 个波段,空间分辨率为 0.55 m.为了研究多种类型的土地覆盖类别变化检测,在原数据集标注的基础上对其他几种土地覆盖类型变化情况进行标注,主要包括:建筑、道路、裸地和水域等多种土地覆盖类别的变化情况.

标注过程中,为了准确对照同一地点的土地覆盖类型,首先将原始影像裁剪为 256×256 大小不重叠的块,再使用 Photoshop 工具同时打开两幅影像采用目视

解译的方法对比变化情况, 在新建的空白图层上对发生变化的区域进行标注, 最后因为发生变化的区域相对较少, 为了平衡样本比例, 对部分完全未发生变化的影像进行剔除, 得到共 1 000 对图像数据. 数据集被随机划分样本为训练集 (80%, 800 张), 验证集 (10%, 100 张) 和测试集 (10%, 100 张). 部分数据集影像原始标签以及新增标注后标签对比图如图 2 所示.

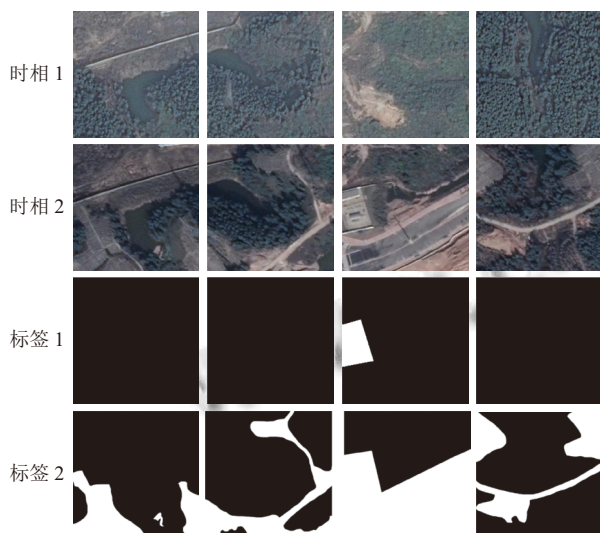


图 2 Google 数据集部分影像及对应的原始标签 1 和新标签 2

3 研究方法

遥感影像变化检测任务, 是一种结合双时相影像特征的语义分割任务. 与语义分割任务单幅影像对应一个类别标签不同, 变化检测任务是一对影像对应一个变化标签. 根据变化检测类型, 变化检测可以分为二进制变化检测和语义变化检测两类, 二进制变化检测时语义变化检测的基础. 其中, 二进制变化检测, 即检测同一场景内, 土地覆盖类型发生变化的区域, 不区分具体的变化类别. 语义变化检测, 不仅需要检测出发生变化的区域还要给出不同位置变化类别, 如草地变为耕地或草地变为道路等. 本研究主要探究遥感影像土地覆盖类别二进制变化检测任务. 通过结合孪生网络结构^[25]对原始特征提取网络进行改进, 使网络在特征提取阶段能够同时提取两时相影像的特征图, 并生成差异特征, 最终完成变化检测任务.

变化检测任务的重点在于两幅影像差异特征的提取. 本研究根据变化检测的任务特点, 在 SeMask 主干

网络的基础上进行改进, 提出了适用于遥感图像土地覆盖类别变化检测的特征提取网络. 在网络的不同阶段增加了两时相影像通道信息交换模块和特征对齐模块, 并且为了在特征提取阶段引导网络关注两时相影像的变化部分, 在每个阶段增加了变化引导模块. 网络的上采样阶段简单地使用了特征金字塔结构对提取的特征图进行上采样. 整体网络框架如图 3 所示. 本研究提出的适用于遥感影像土地覆盖类别变化检测网络, 主要由 4 个部分组成, 分别为特征提取模块、变化引导模块、通道交换和特征对齐模块.

特征提取模块使用 Swin Transformer^[21] 的特征提取块, 在每个特征提取块内有个 N_A 可位移窗口注意力块, 具有线性计算复杂度, 并使用跨窗口连接来处理非重叠区域之间的信息交互, 这样的设计在获得图像级特征建模的同时降低计算复杂度. 为了计算 Transformer 层的自注意力, 将特征图 X 转化为: Q 、 K 、 V , 它们分别是查询矩阵 (query)、键矩阵 (key) 和值矩阵 (value), 具有相同的维度. 最后为了表示图片的位置关系, 引入相对位置嵌入 (relative position embedding, RPE)^[26], 自注意力计算表达式为:

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{C}} + RPE\right)V \quad (1)$$

变化引导模块是原 SeMask Transformer^[22] 的语义层, 主要为了帮助网络在特征提取阶段就关注两幅影像中发生变化的部分. 与原始主干网络不同, 网络将每个阶段的输出的两个特征图 f_1 和 f_2 , 做差并取绝对值后为 $|f_1 - f_2|$ 输入到变化引导模块, 变化引导模块输出的特征图为 $N \times 2 \times H \times W$, 其中 N 为批量大小, 2 为通道数, H 和 W 为不同阶段特征图的大小. 在训练阶段, 为了有效监督网络, 避免长距离的上采样导致梯度消失, 将标签图进行下采样到与特征图相同的大小进行损失计算. 标签图有一定的空间冗余, 对其下采样不会造成很大的损失, 将标签图下采样到不同阶段的分辨率, 对变化引导特征进行直接监督.

3.1 通道信息交互模块

遥感图像变化检测是结合双时相影像的变化检测, 为了帮助两个特征图像特征融合, 在网络第 2 阶段和第 3 阶段引入通道交互模块^[27], 两个特征图部分通道互相交换, 通过交互后的两个特征分支之间的分布更加相似, 如图 4 所示.

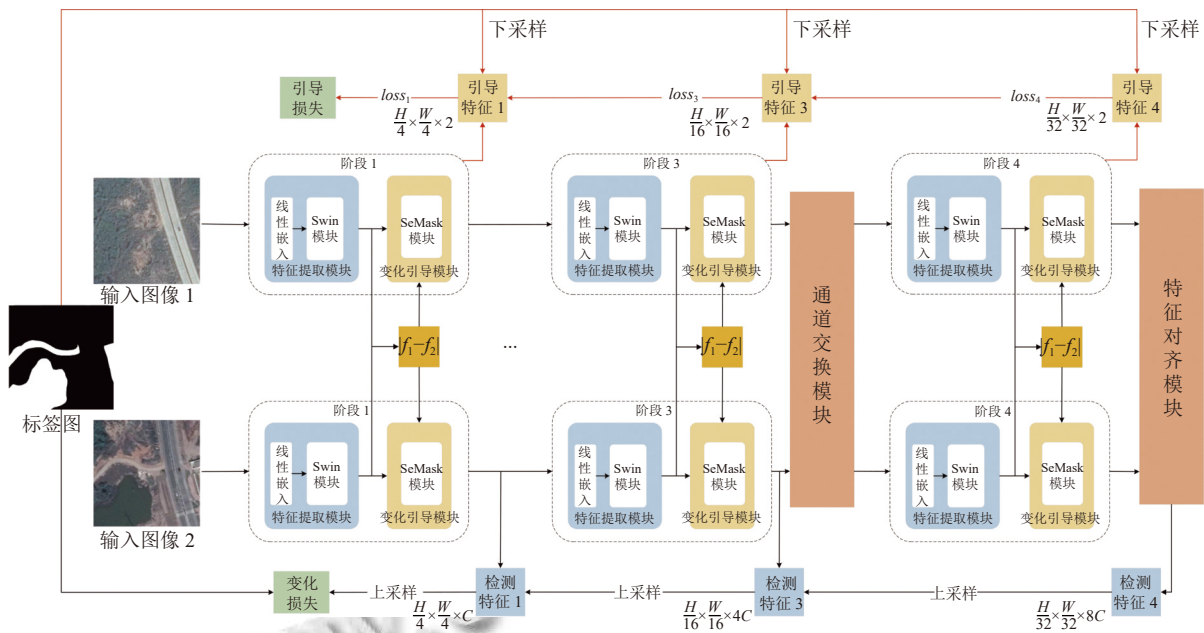


图3 变化检测网络整体框架

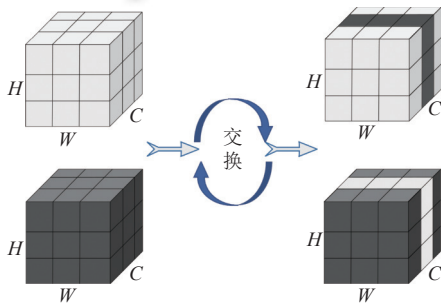


图4 通道交互模块

3.2 特征对齐模块

在主干网络的特征提取过程中, 图像经过多次池化操作分辨率不断下降, 在提取到高层语义特征的同时也导致图像特征偏移问题. 偏移后的特征图在上采样过程中会发生更大的位置偏移, 这成为限制网络检测精度的一大因素.

本研究针对特征位置偏移问题, 在模型第4阶段引入特征对齐模块 (FDAF)^[27]. 具体来说, 主要通过一个可变形重采样网络获得两幅图像各自的位置偏移, 对原始特征图进行位置矫正. 数学公式表示为:

$$x = \text{Concat}(x_0(p + \Delta p_0) - x_1, x_1(p + \Delta p_1) - x_0) \quad (2)$$

其中, x 表示对齐后输出的图像特征, x_0 和 x_1 分别表示两时相的特征图像, p 表示原始特征图的特征位置, Δp 表示通过重采样网络获得的特征偏移矫正量. Δp 的获得

可以用式 (3) 表示.

$$\Delta p = \text{FlowNet}(\text{Concat}([x_0, x_1]) \quad (3)$$

3.3 损失函数

为了提高模型的分割精度, 获得更精确的边缘分割结果, 在交叉熵损失函数的基础上增加了边缘损失函数 L_{edge} ^[28], 损失函数定义如下:

$$L = L_{\text{edge}} + L_{\text{ce}} \quad (4)$$

其中, L_{ce} 代表交叉熵损失函数, L_{edge} 代表边界损失函数.

交叉熵损失函数 (cross-entropy loss function), 也称为 *Softmax* 损失函数, 是语义分割任务中常用的损失函数, 可以用于多类别分割任务, 计算公式表示为:

$$L_{\text{ce}} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{i,j} \log(p_{i,j}) \quad (5)$$

其中, N 是样本数, C 是类别数, $y_{i,j}$ 是样本 i 属于类别 j 的真实标签, $p_{i,j}$ 是模型对样本 i 属于类别 j 的预测概率.

边界损失函数为了强调分割对象的边界区域, 通过计算像素点与变化标签的距离为位于边界的像素点赋予更高的权重, 使网络在优化过程中重点关注模型的边界区域. 具体权重计算方法如下:

$$w_{\text{edge}}(c) = \frac{\alpha}{n} \sum_{k \in N} |L(k) - L(c)| \quad (6)$$

其中, $w_{\text{edge}}(c)$ 代表点 c 的权重, α 是一个系数, N 代表点

c 邻域的点, N 在实验中被设置为7, c 表示区域 N 的圆心点, $L(k)$ 代表点 k 的标签值, 1表示变化, 0表示未变化.

4 实验设计与评价指标

4.1 软硬件环

本次遥感图像土地覆盖类别变化识别实验的硬件环境, 使用了4块内存容量为11 GB的GeForce GTX 1080 Ti显卡. 深度学习平台为PyTorch 1.8.1, Python 3.8和CUDA 10.2.

4.2 训练设置

优化器: 本研究设计的变化检测网络使用AdamW优化器^[29]和WarmupPolyLR学习率迭代策略, 设置初始学习率为 10^{-2} , 权重衰减为 $6 \cdot 10^{-5}$. 对比方法使用随机梯度下降(stochastic gradient descent, SGD)优化器, 设置初始学习率为 10^{-2} , 权重衰减设置为 $5 \cdot 10^{-5}$.

批大小(batchsize): 所有模型批统一设置为8, 训练轮次设置为1000轮.

4.3 评价指标

为了从多个角度对模型进行评估, 本实验选用召回率(R)、像素精度(PA)、 $F1$ 分数和参数量作为评价指标. 计算公式分别为:

$$R = TP / (TP + FN) \quad (7)$$

$$PA = TP / (TP + FP) \quad (8)$$

$$F1 = \frac{2}{R^{-1} + P^{-1}} \quad (9)$$

其中, TP 为正确预测变化像素数; FP 为非变化预测为变化的像素数; FN 为将变化预测为非变化像素数; TN 为将非变化预测为变化像素数. PA 为像素精度, 即正确预测类别的像素数量占总像素数量的比例, 反映了变化检测的可靠程度; R 为召回率, 即预测正确的变化像素占总真实变化像素数量的比例, 反应变化检测的完整程度. $F1$ 为精确率(precision)与召回率的平均值, 综合表示模型的分割精度. 模型参数量是考察模型规模的重要参数, 代表处理相同输入图像的大小情况下, 模型推理所需的参数数量, 其值越低代表模型规模越小, 轻量化程度越高.

5 实验结果与分析

5.1 Google土地覆盖类别变化检测数据集实验结果

将几个先进的变化检测网络和本研究方法在Google

土地覆盖类别变化检测数据集中进行实验, 数据集按照8:1:1的比例划分训练集、验证集和测试集, 使用测试集的分割成绩作为最终的分割结果, 使用召回率(R)、像素精度(PA)和 $F1$ 分数作为评价指标. 表1显示了各个网络的分割结果.

表1 Google土地覆盖类别变化检测数据集评价表

网络	参数量(M)	R (%)	PA (%)	$F1$ (%)
FC-EF ^[11]	1.83	59.34	77.28	68.23
FC-Siam-Diff ^[11]	1.83	60.29	80.71	71.12
FC-Siam-Conc ^[11]	2.03	57.22	81.13	69.10
STA-Net ^[14]	12.2	88.62	81.34	82.4
DASNet ^[17]	16.25	85.14	88.31	85.57
本文方法	26.2	86.8	88.4	87.45

基于全连接卷积神经网络的前3个方法, 采了多种特征结合方式, 拥有较小的参数规模, 但 $F1$ 分数相对其他方法较低. 其中, FC-EF采用将两时相影像合并输入到全连接神经网络方法, 提取变化特征, 最终 $F1$ 分数为68.23%. FC-Siam-Conc采用分开提取两时相特征并在上采样阶段将不同阶段特征图进行合并后上采样方法提取变化特征, 最终 $F1$ 分数为69.1%. FC-Siam-Diff同样采用分别提取影像特征, 并对相同尺寸的特征图做差取绝对值处理来突出变化特征, 获得了该网络结构最高 $F1$ 分数为71.12%. STA-Net和DASNet两个网络均采用ResNet主干网络, 并在上采样阶段采用不同的特征融合策略完成变化检测任务. STA-Net网络在上采样阶段分别基于注意力机制和多尺度方法提取特征图中不同尺度的变化特征, 获得最高的召回率88.62%. DASNet在基于注意力机制的基础提取特征图的同时加入了带有权重的损失函数, 使网络更加关注变化类别, 最终以16.25M的参数量获得85.57%的 $F1$ 分数. 本文方法主要在主干网络部分通过增强特征信息交互和位置偏移矫正来提高模型对变化信息的提取能力, 在上采样阶段仅采用最基本的插值上采样方法. 最终使用26.2M参数, 获得87.45%最高分数, 与DASNet相比提高了2.8%.

5.2 实验结果可视化

本文方法和几个对比方法在Google土地覆盖类别变化检测数据集可视化结果如图5所示. 图像1和图像2分别代表同一个地区两个时相的遥感影像, 标签图分别用白色和黑色表示变化和未变化区域. 从图5可以看到基于全连接神经网络的方法, 如FC-EF和

FC-Siam-Diff, 只能在较大范围内分割变化和未变化区域, 对于小的变化区域如区域 2 和区域 3 存在变化区域未能检测出来的情况, 或者碎片化的检测结果, 整体分割精度较差. 相比之下, STA-Net 和本文提出的网络

模型表现更好, 与标签图相比, 虽然在图像边缘细节方面还有提升空间, 但整体差距不大. 尤其在区域 1、2 和 3 中, 本文的分割方法基本能够分割出细小的变化区域, 如细小的道路和房屋等区域.

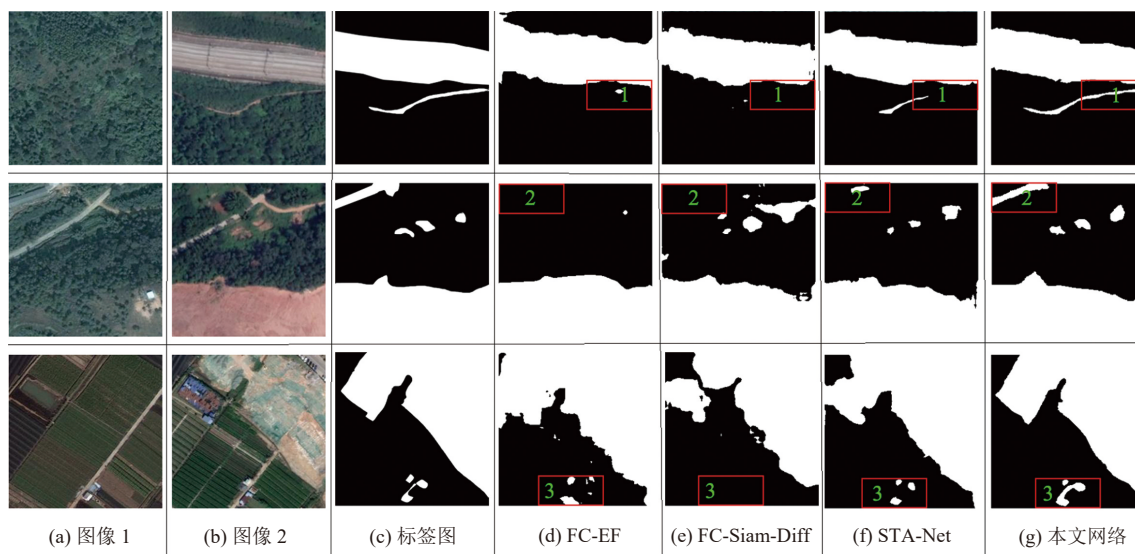


图5 Google 土地覆盖类别变化检测数据集部分分割结果可视化

通过对多个方法的可视化分割结果, 可以看到本文的变化检测方法利用信息交互模块和特征对齐模块增强两时相影像的信息交互, 充分提取两幅影像的差异特征, 获得较好的整体分割结果. 通过加入边缘损失函数强调变化边缘特征信息, 提高网络的边缘检测能力, 获得了最优的边缘分割结果.

6 结论

本文以深度学习 SeMask 语义分割主干网络为基础, 采用孪生网络结构, 共享主干网络参数以降低模型规模. 在特征提取阶段, 加入特征通道信息交互模块增强两幅影像的特征信息交互. 此外, 通过特征对齐模块缓解特征偏移现象, 并在损失函数中增加边缘损失函数以提高边缘分割性能. 实验结果显示, 改进后的网络在遥感影像变化检测任务中取得了较好的分割结果, 尤其对于房屋和道路等相对较小变化目标的检测能力有所提升. 与 DASNet 相比, 本文方法的 $F1$ 分数提高了 2.8%. 然而, 改进后的网络模型参数量较大. 因此, 下一步的研究工作可以探索基于注意力机制的两时相影像信息融合方法, 改进特征融合并降低模型参数量, 以提高检测精度. 此外, 有限的数据集标注数量也限制

了土地覆盖类型变化检测模型的发展, 因此可以从如何获得更多高质量的变化标签出发, 进一步提高变化检测精度.

参考文献

- 1 宋刚贤. 基于遥感影像的城市违规建设用地监测研究 [硕士学位论文]. 南京: 南京农业大学, 2008.
- 2 李博. 基于卫星遥感与 IBIS 模型的藏北高原草地动态变化研究 [硕士学位论文]. 兰州: 兰州大学, 2018.
- 3 柳晶辉, 邵芸, 黄初冬. 基于遥感影像的城市森林分类提取及生态价值估算研究. 地理与地理信息科学, 2007, 23(4): 33-36. [doi: 10.3969/j.issn.1672-0504.2007.04.009]
- 4 杨蜀秦, 宋志双, 尹瀚平, 等. 基于深度语义分割的无人机多光谱遥感作物分类方法. 农业机械学报, 2021, 52(3): 185-192. [doi: 10.6041/j.issn.1000-1298.2021.03.020]
- 5 Gao JP, Xu CB, Zhang L, et al. Infrared image change detection of substation equipment in power system using Markov random field. Proceedings of the 2017 International Conference on Computing Intelligence and Information System (CIIS). Nanjing: IEEE, 2017. 332-337.
- 6 李龙凯. 面向地理国情监测的城市遥感影像变化监测分析方法研究. 中国住宅设施, 2022, (7): 64-66. [doi: 10.3969/j.issn.1672-5093.2022.7.zgzzss202207022]

- 7 付宝晶, 李自立. 基于多特征融合的遥感图像河流提取. 中国农村水利水电, 2022, (12): 53–58. [doi: [10.12396/znsd.220470](https://doi.org/10.12396/znsd.220470)]
- 8 王光辉, 李建磊, 王华斌, 等. 基于多特征融合的遥感影像变化检测算法. 国土资源遥感, 2018, 30(2): 93–99.
- 9 Vapnik VN. The Nature of Statistical Learning Theory. New York: Springer, 1995.
- 10 Ruiz JAM, Riaño D, Arbelo M, *et al.* Burned area mapping time series in Canada (1984–1999) from NOAA-AVHRR LTDR: A comparison with other remote sensing products and fire perimeters. Remote Sensing of Environment, 2012, 117: 407–414. [doi: [10.1016/j.rse.2011.10.017](https://doi.org/10.1016/j.rse.2011.10.017)]
- 11 Daudt RC, Le Saux B, Boulch A. Fully convolutional siamese networks for change detection. Proceedings of the 25th IEEE International Conference on Image Processing (ICIP). Athens: IEEE, 2018. 4063–4067.
- 12 Daudt RC, Le Saux B, Boulch A, *et al.* Urban change detection for multispectral earth observation using convolutional neural networks. Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2018. 2115–2118.
- 13 Benedek C, Sziranyi T. Change detection in optical aerial images by a multilayer conditional mixed Markov model. IEEE Transactions on Geoscience and Remote Sensing, 2009, 47(10): 3416–3430. [doi: [10.1109/TGRS.2009.2022633](https://doi.org/10.1109/TGRS.2009.2022633)]
- 14 Bi HB, Lu D, Zhu HH, *et al.* STA-Net: Spatial-temporal attention network for video salient object detection. Applied Intelligence, 2021, 51(6): 3450–3459. [doi: [10.1007/s10489-020-01961-4](https://doi.org/10.1007/s10489-020-01961-4)]
- 15 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 770–778.
- 16 Chen H, Shi ZW. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection. Remote Sensing, 2020, 12(10): 1662. [doi: [10.3390/rs12101662](https://doi.org/10.3390/rs12101662)]
- 17 Chen J, Yuan ZY, Peng J, *et al.* DASNet: Dual attentive fully convolutional Siamese networks for change detection in high-resolution satellite images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 1194–1206. [doi: [10.1109/JSTARS.2020.3037893](https://doi.org/10.1109/JSTARS.2020.3037893)]
- 18 季顺平, 田思琦, 张驰. 利用全空洞卷积神经网络进行城市土地覆盖分类与变化检测. 武汉大学学报·信息科学版, 2020, 45(2): 233–241. [doi: [10.13203/j.whugis20180481](https://doi.org/10.13203/j.whugis20180481)]
- 19 Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.
- 20 Dosovitskiy A, Beyer L, Kolesnikov A, *et al.* An image is worth 16x16 words: Transformers for image recognition at scale. Proceedings of the 9th International Conference on Learning Representations. ICLR, 2021.
- 21 Liu Z, Lin YT, Cao Y, *et al.* Swin Transformer: Hierarchical vision transformer using shifted windows. Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021. 9992–10002.
- 22 Jain J, Singh A, Orlov N, *et al.* SeMask: Semantically masked transformers for semantic segmentation. arXiv: 2112.12782, 2022.
- 23 Shen L, Lu Y, Chen H, *et al.* S2Looking: A satellite side-looking dataset for building change detection. Remote Sensing, 2021, 13(24): 5094. [doi: [10.3390/rs13245094](https://doi.org/10.3390/rs13245094)]
- 24 Peng DF, Bruzzone L, Zhang YJ, *et al.* SemiCDNet: A semisupervised convolutional neural network for change detection in high resolution remote-sensing images. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(7): 5891–5906. [doi: [10.1109/TGRS.2020.3011913](https://doi.org/10.1109/TGRS.2020.3011913)]
- 25 Bromley J, Guyon I, LeCun Y, *et al.* Signature verification using a “Siamese” time delay neural network. Proceedings of the 6th International Conference on Neural Information Processing Systems. Denver: Morgan Kaufmann Publishers Inc., 1993. 737–744.
- 26 Raffel C, Shazeer N, Roberts A, *et al.* Exploring the limits of transfer learning with a unified text-to-text Transformer. The Journal of Machine Learning Research, 2020, 21(1): 140.
- 27 Fang S, Li KY, Li Z. Changer: Feature interaction is what you need for change detection. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5610111.
- 28 Chen HJ, Pu FL, Yang R, *et al.* RDP-net: Region detail preserving network for change detection. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5635010.
- 29 Loshchilov I, Hutter F. Decoupled weight decay regularization. Proceedings of the 7th International Conference on Learning Representations. New Orleans: ICLR, 2019.

(校对责编: 牛欣悦)