

面向高分辨率遥感影像变化检测的混合空间金字塔池化网络

邵攀, 高梓昂

1. 三峡大学 湖北省水电工程智能视觉监测重点实验室, 宜昌 443002;

2. 三峡大学 计算机与信息学院, 宜昌 443002

摘要: 以残差网络为基础, 结合高分辨率遥感影像的特点, 提出一种全新的端到端变化检测网络——混合空间金字塔池化网络 HSPPNet (Hybrid Spatial Pyramid Pooling Network)。HSPPNet 首先将空洞卷积和注意力机制引导卷积并行集成, 构建一种混合空间金字塔池化模块, 以便有效提取高分辨率遥感影像中不同形状尺度的变化对象。然后, 通过定义一种截断—补偿加权交叉熵函数和一种类别级 IoU 函数并将二者集成, 得到一种全新的自适应平衡损失函数, 来降低变化类与未变化类严重不均衡问题对变化检测的影响。最后设计一种简单有效的输入模块, 通过综合考虑两期遥感影像及其差异图来增强变化信息。通过以上3点, HSPPNet 增强了深度学习变化检测的性能。两组常用公开变化检测数据集上的实验结果表明 HSPPNet 可行、有效。

关键词: 遥感, 变化检测, 深度学习, 混合空间金字塔池化, 注意力机制, 损失函数, 截断—补偿加权交叉熵, 类别级 IoU

中图分类号: P237/P2

引用格式: 邵攀, 高梓昂. 2025. 面向高分辨率遥感影像变化检测的混合空间金字塔池化网络. 遥感学报, 29(1): 279–289

Shao P and Gao Z A. 2025. Hybrid spatial pyramid pooling network to change detection for high resolution remote sensing images. National Remote Sensing Bulletin, 29(1): 279–289 [DOI: 10.11834/jrs.20233229]

1 引言

遥感变化检测是通过分析同一地理位置不同时刻获取的遥感影像提取地表变化信息的过程, 在城市研究、自然资源监测、灾害评估等领域有着广泛应用。经过几十年的发展, 已有许多变化检测方法和模型被提出 (Hussain 等, 2013; 佘袁勇 等, 2016)。根据使用的理论基础不同, 现有变化检测技术可大致划分为两类: 传统变化检测和深度学习变化检测。

传统变化检测分为直接比较法和分类后比较法两类 (Chen 等, 2013)。直接比较法通常包含两个主要步骤: (1) 通过比较两时相的遥感影像生成差异图 DI (Difference Image); (2) 通过分析 DI 生成变化检测图。而分类后比较法首先分别对两期遥感影像进行分类, 然后通过逐像素对比分类

结果获取变化信息。传统变化检测, 无论是直接比较法还是分类后比较法, 均是基于手工设计特征实现的。然而手工设计特征对影像的表征能力较弱, 且鲁棒性较差, 使得传统变化检测在某些情况下不能取得理想的检测结果, 特别是对包含更多地物细节的高分辨率遥感影像。

鉴于深度学习在计算机视觉领域的巨大成功, 其在遥感领域受到越来越多的关注, 已有许多基于深度学习的变化检测方法被提出 (Shi 等, 2020)。在这些方法中, 卷积神经网络 CNN (Convolutional Neural Networks) 应用最广泛, 这主要得益于 CNN 强大的特征表征能力。本文主要关注基于 CNN 的深度学习变化检测。基于 CNN 的深度学习变化检测主要通过两种方式输入两期遥感影像 (Zhang 等, 2020): 级联输入和独立输入。级联输入首先将两期影像级联, 即将两期影像在通道维度上进

收稿日期: 2023-06-23; 预印本: 2023-10-31

基金项目: 国家自然科学基金(编号:41901341)

第一作者简介: 邵攀, 研究方向为遥感图像处理。E-mail: panshao@whu.edu.cn

通信作者简介: 高梓昂, 研究方向为遥感图像处理。E-mail: 13355213156@163.com

行拼接(梁哲恒等, 2022), 得到一个通道数翻倍的新影像; 然后将新影像作为输入, 通过单通道网络提取两期影像特征(Daudt等, 2018; Peng等, 2019; Zhang等, 2022; Zhou等, 2020; 顾炼等, 2020)。独立输入则将两期影像作为两个输入, 采用孪生或平行网络结构分别提取两期影像特征(Chen等, 2021, 2022; Chen和Shi, 2020; Liu等, 2020, 2022; Zhang和Shi, 2020; 梁哲恒等, 2022)。虽然通过级联输入和独立输入, 均能够较为有效地实施变化检测, 但两种输入方式都只考虑两期影像自身, 忽略了它们的DI影像; 而DI影像对变化检测任务具有重要价值, 这已在传统的直接比较法变化检测研究中得到证明。

输入两期影像后, 可采用不同的主干网络提取时空特征, 常用的包括: 全卷积神经网络(Chen等, 2021; Daudt等, 2018), VGG(Zhang和Shi, 2020), 残差网络(Chen和Shi, 2020), U-net(Liu等, 2020; 顾炼等, 2020), UNet++(Peng等, 2019; Zhang等, 2022; 王超等, 2023), DeepLabv3+(齐建伟等, 2023)等。为提升变化检测效果, 遥感学者从不同角度将注意力机制引入变化检测来建模影像远距离相关性, 对主干网络所提取的特征进行增强: 比如, Chen和Shi(2020)通过时空注意力; Chen等(2021)通过双注意力; 梁哲恒等(2022)通过多尺度注意力增强特征。最近, Transformer也被引入变化检测建模影像远距离相关性(Chen等, 2022; Liu等, 2022)。针对变化对象尺度多变问题, 遥感学者从不同角度出发, 采用不同策略提取两期影像的多尺度信息: 比如, Zhang等(2022)、王超等(2023)采用多边输出融合和多层级特征差分策略; Cheng等(2021)采用Laplacian金字塔扩展网络; 齐建伟等(2023)采用空洞空间金字塔池化ASPP(Atrous Spatial Pyramid Pooling)结构; 梁哲恒等(2022)采用金字塔网络和ASPP结构获取多尺度信息。ASPP是一种简单有效的获取多尺度信息工具, 然而, 由于ASPP采用空洞卷积, 使得其不能很好地检测线状变化对象和块状变化对象的边界。

深度学习变化检测常用交叉熵计算损失。然而, 由于变化检测任务中未变化区域通常远大于变化区域, 造成变化类与未变化类占比不均衡, 使得传统交叉熵不能很好地适应变化检测模型的

训练。为解决此问题, Cheng等(2021)将Dice损失函数与交叉熵损失函数集成, 使用综合损失函数计算损失。Peng等(2019)通过交叉相乘类别占比对变化类和未变化类进行加权, 使用加权交叉熵损失函数缓解类别不均衡对模型训练的影响。虽然类别占比加权能够有效缓解类别不均衡对模型训练的影响, 但变化类和未变化类可能出现极端不均衡现象(比如类别占比分别为0.99和0.01), 这样会导致过度增强变化类, 不利于模型训练。

为克服前文所述3点不足, 本文提出一种全新的面向高分辨率遥感影像的端到端变化检测网络模型——混合空间金字塔池化网络HSPPNet(Hybrid Spatial Pyramid Pooling Network)。HSPPNet整体上采用编码器—解码器结构, 主要包括以下创新和优势:

(1) 针对级联和独立两种输入方式的不足, 给出一种简单有效的输入模块, 通过综合考虑两期影像及其DI来增强变化信息。

(2) 为克服ASPP的不足, 将空洞卷积和注意力引导卷积有机集成, 构建一种混合空间金字塔池化HSPP(Hybrid Spatial Pyramid Pooling)模块。HSPP能够很好地适应高分辨率遥感影像的多尺度特性: 具有较大感受野的空洞卷积能够有效识别不同尺度的块状变化对象, 而注意力引导卷积能够有效检测线状变化对象和增强块状变化对象边界。

(3) 为解决类别占比加权交叉熵损失函数的不足, 提出一种全新的自适应平衡损失函数: 首先定义一种截断—补偿加权交叉熵损失函数和一种类别级IoU损失函数, 然后将两者进行集成。所提出损失函数分别从全局和局部的角度来缓解变化类与未变化类严重不均衡问题对变化检测的影响。

2 本文方法

如图1所示, 所提出的变化检测网络模型HSPPNet采用编码—解码架构, 主要包括4个部分: 输入模块、下采样模块、混合空间金字塔池化(HSPP)模块和上采样模块。第2.1节给出输入和下采样模块的具体细节; 第2.2节详细介绍HSPP模块; 第2.3节介绍上采样模块和所提出损失函数。

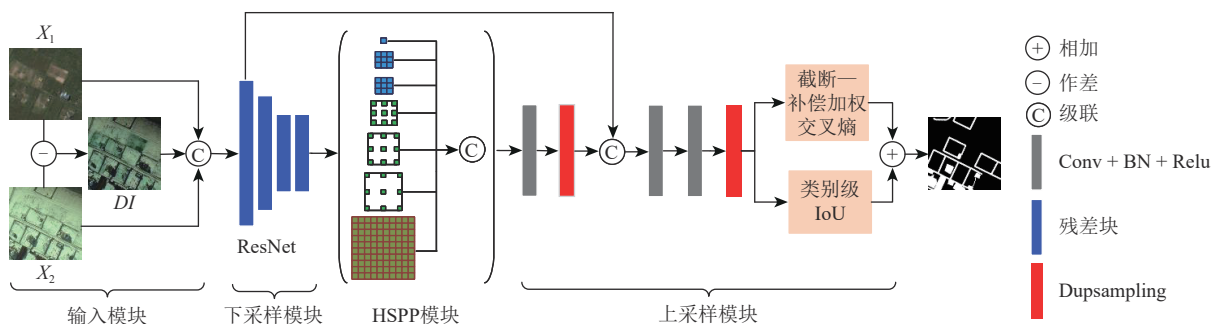


图1 所提出HSPPNet的网络框架

Fig. 1 Architecture of the proposed HSPPNet

2.1 输入和下采样模块

为充分提取变化特征，本文设计一种简单有效的输入模块，来综合考虑两期影像及其DI影像。假设 X_1 和 X_2 表示同一地理位置不同时刻获取的，已经过辐射校正和配准等预处理的两时相高分辨率遥感影像，首先采用差值法来生成两期影像 X_1 和 X_2 的DI影像，即

$$DI = |X_1 - X_2| \quad (1)$$

然后，以级联方式将 X_1 、 X_2 和 DI 集成，记作 X_{in} ： $X_{in} = \text{Concat}(X_1, X_2, DI)$ ， Concat 表示级联操作；最后将 X_{in} 输入到下采样模块。

通过综合考虑性能和效率，本文采用 ResNet50 (He 等, 2016) 进行下采样，提取输入影像 X_{in} 的多层级特征。ResNet50 主要包含 5 个层级模块，分别记作 ConvL1、RconvL2、RconvL3、RconvL4 和 RconvL5，并将 5 个层级模块的输出分别记作 F_1 、 F_2 、 F_3 、 F_4 和 F_5 ；其中 ConvL1 是一个步长为 2 的传统卷积层，而 RconvL2—RconvL5 是由多个残差卷积单元组成的残差块。考虑到遥感变化检测是一个需要逐像素预测的密集分类任务，如果提取的最高层级特征 F_5 尺寸太小，其包含的空间细节就会过少，则不利于影像的恢复与重构。因此，对模块 RconvL5 进行修正，在 RconvL5 中未对 RconvL4 的输出进行下采样，使得特征图 F_5 和 F_4 具有相同的尺寸。同时，为扩大 F_5 的感受野，使用空洞率为 2 的空洞卷积替代 RconvL5 中的传统卷积。ResNet50 的更多细节可参见文献 (He 等, 2016)。

通过下采样模块后，得到 5 个层级的特征图 F_1 — F_5 ，其尺寸分别为原始影像的 1/2、1/4、1/8、1/16 和 1/16。为提取高分影像中不同形状尺度的变化对象，将特征图 F_5 输入到所提出的 HSPP 模块。

2.2 混合空间金字塔池化 (HSPP) 模块

为提取影像中不同尺度的对象，Chen 等 (2018a, 2018b) 将空洞卷积集成到空间金字塔池化结构，针对传统语义分割任务提出一种空洞空间金字塔池化模块 ASPP。空洞卷积在卷积核中引入“空洞”，能够有效扩大感受野，且无需引入额外参数。ASPP 模块包含多个具有不同空洞率的空洞卷积层，能够有效地捕获影像中不同尺度的上下文信息，在传统影像的语义分割任务中取得良好的效果。然而，直接将 ASPP 应用到高分辨率遥感影像变化检测，效果并不理想。这主要是因为 (1) 高分辨率遥感影像中变化对象的形状和尺度比传统影像中对象更加复杂多变；(2) 虽然 ASPP 能够较为有效地检测出不同尺度的块状变化对象，但 ASPP 采用空洞卷积，使得其不能很好地检测线状变化对象和块状变化对象的边界。

针对上述问题，本文利用双注意力引导卷积增强 ASPP 模块，提出一种混合空间金字塔池化模块 HSPP。双注意力机制 DAM (Dual Attention Mechanism) 由文献 (Fu 等, 2020) 提出。DAM 通过分别对空间域和通道域运用自注意力机制，构建空间注意力模块 PAM (Position Attention Module) 和通道注意力模块 CAM (Channel Attention Module)。Fu 等 (2020) 的实验结果表明，DAM 能够有效地建模影像全局依赖关系，显著增强影像中线状对象 (比如树干) 和块状对象 (比如建筑物) 的边界。因此，利用 DAM 增强 ASPP，有望弥补 ASPP 的不足，提升特征提取效果。

所提出的 HSPP 模块结构如图 2 所示，从上到下共包括 7 个分支：第 1 个分支是一个卷积核为 1×1 的传统卷积层，记作 Conv。第 2 个分支是一个空间注意力引导的传统卷积层 (记作 PAconv)：首先

将特征图 F_5 输入到空间注意力模块 PAM，然后对 PAM 的输出进行卷积，卷积核大小为 3×3 。第 3 个分支是一个通道注意力引导的传统卷积层（记作 CAconv）：首先将 F_5 输入到通道注意力模块 CAM，然后进行 3×3 的卷积。第 4—6 个分支是 3 个空洞卷积层，分别记作 Aconv1, Aconv2 和 Aconv3：卷积核大小均为 3×3 ，空洞率分别为 6、12 和 18。第 7 个分支是一个全局最大池化层，记作 GMP。

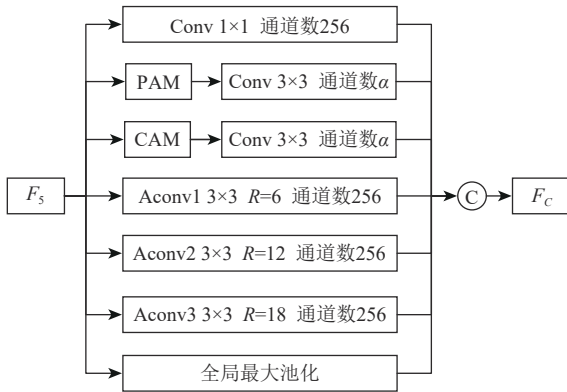


图2 所提出的 HSPP 模块
Fig. 2 The proposed HSPP module

HSPP 模块的第 1、4、5 和 6 分支的卷积核个数采用定值，均设为 256。为灵活调整两种注意力在 HSPP 模块中所起的作用，在 CAconv 和 PAconv 两个分支中将卷积核个数设为变量，通过验证集训练得到。为简化训练过程，CAconv 和 PAconv 分支中采用相同的卷积核个数 α ，并令 α 在集合 $\{32, 64, 128, 256\}$ 中取值。空洞卷积和注意力模块 PAM 与 CAM 的更多细节请参见文献 Chen 等 (2018a) 和 Fu 等 (2020)。特征图 F_5 通过 HSPP 模块 7 个分支处理后，得到 7 种输出，将其级联 (Concat) 得到特征 F_c ：

$$F_c = \text{Concat}(\text{Conv}(F_5), \text{PAconv}(F_5), \text{CAconv}(F_5), \text{Aconv1}(F_5), \text{Aconv2}(F_5), \text{Aconv3}(F_5), \text{GMP}(F_5)) \quad (2)$$

HSPP 模块通过使用不同的卷积核和空洞率，以及全局最大池化，能够有效提取影像中不同尺度的上下文信息；通过 DAM 建模空间和通道全局依赖性，能够有效增强对线状变化对象和块状变化对象边界的提取效果。HSPP 能够很好地提取不同尺度、不同形状的变化对象，效果明显优于 ASPP（参见 3.3.1 节网络结构消融实验）。

2.3 上采样模块和损失函数

(1) 上采样过程。得到 HSPP 模块 7 个分支的

级联特征 F_c 后，需通过上采样将其恢复至原始影像尺寸。为提升对影像的重构能力，本文采用数据依赖型上采样方法 DUpsampling (Data-dependent upsampling (Tian 等, 2020)) 替换双线性插值进行上采样。DUpsampling 在真值标签指导下，通过卷积生成恢复影像分辨率所需的填充部分，重构能力更强，有助于影像细节信息重构。上采样过程 (图 1) 如下：

(1) 将 F_c 输入到卷积层 Uconv1，来初步降低特征通道数并提升网络非线性表达能力。Uconv1 包括 3×3 卷积、批量归一化 BN (Batch normalization) 和 Relu 激活操作，卷积核个数为 256。

(2) 利用 DUpsampling 对步骤 (1) 的输出进行 4 倍上采样，得到特征 F_{u1} 。

(3) 由于低层级特征包含较丰富的空间位置信息，有利于影像细节重构，本步骤通过跳跃连接将 ResNet50 输出的低层级特征 F_2 (2.1 节) 与特征 F_{u1} 进行级联。

(4) 将步骤 (3) 的输出依次通过两个卷积层 Uconv2 和 Uconv3。Uconv2 与 Uconv1 完全相同，Uconv3 通过将 Uconv1 的卷积核个数调整为 128 得到。

(5) 利用 DUpsampling 对步骤 (4) 的输出进行 4 倍上采样，得到与原始影像尺寸相同的特征图，并将其输入 Softmax 层，生成变化概率图。

(2) 所提出的平衡损失函数。一般来讲，在一段时间内发生变化的区域远小于未发生变化区域，导致遥感影像中变化类与未变化类占比不均衡。若使用传统交叉熵损失函数，非常不利于变化检测模型的训练，影响模型性能。为解决此问题，一些学者通过交叉相乘类别占比对变化类和未变化类进行加权，提出加权交叉熵损失函数 L_{BCE} (Peng 等, 2019)：

$$L_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N (r_u y_{ic} \log(p_{ic}) + r_c (1 - y_{ic}) \log(p_{iu})) \quad (3)$$

式中， p_{ic} 和 p_{iu} 分别表示像素 i 属于变化类 c 和未变化类 u 的预测概率； y_{ic} 是一个布尔变量，若像素 i 属于变化类 c ，则 $y_{ic}=1$ ，否则， $y_{ic}=0$ ； $r_u(r_c)$ 表示未变化类 (变化类) 像素数与像素总数 N 的比值，且满足 $r_u+r_c=1$ ； $\log$ 表示自然对数运算。

虽然损失函数 L_{BCE} 能够在一定程度上缓解类别不均衡问题对模型训练的影响，但变化类和未变化类可能出现极端不均衡现象 (比如 $r_u=0.99$,

$r_c=0.01$)；这样会导致变化类权重 r_u 过大，使得 L_{BCE} 过分强调变化类，不利于模型的训练。为解决 L_{BCE} 的不足，本文提出一种截断—补偿加权交叉熵损失函数 L_{TCW} ，通过对类别占比 r_k ($k \in \{u, c\}$) 进行截断补偿来定义变化类 c 和未变化类 u 的权重：

$$w_k = \begin{cases} r_k, & r_k \leq T \\ r_k - a, & r_k > T \\ r_k + a, & r_k < 1 - T \end{cases} \quad k \in \{u, c\} \quad (4)$$

式中， w_k 表示类别 k ($k \in \{u, c\}$) 的截断—补偿权重， $r_{\max}$ 表示 r_u 和 r_c 的最大值。常数 T 表示截断阈值， a 表示对权重的补偿值，本文中令 $T=0.75$ ， $a=0.1$ 。所定义的截断—补偿加权交叉熵损失函数 L_{TCW} 表达式如下：

$$L_{TCW} = -\frac{1}{N} \sum_{i=1}^N (w_u y_{ic} \log(p_{ic}) + w_c (1 - y_{ic}) \log(p_{iu})) \quad (5)$$

L_{TCW} 具有如下优势：(1) 能够根据类别占比自适应调整类别权重；(2) 通过截断—补偿设计能够避免极端权值出现，避免过度增强某一类。

L_{TCW} 逐像素计算损失，从这个角度讲 L_{TCW} 属于局部损失函数。IoU 是另一种经典损失函数，通过比较预测图和真值标签图来计算损失，属于全局损失函数。通过集成全局和局部损失函数，有望得到一个更优的损失函数。然而在变化检测任务中变化区域占比通常较小，直接计算整幅图的 IoU 会使得其对变化类很不敏感。为突出变化类的作用，本文给出基于变化类的类别级 IoU 损失函数 L_{CI} ：

$$L_{CI} = 1 - \frac{G_c \cap P_c}{G_c \cup P_c} \quad (6)$$

式中， G_c 表示真值标签图的变化区域， P_c 表示预测图的变化区域， $\cap$ 和 $\cup$ 表示集合的交和并运算。

本文将 L_{TCW} 和 L_{CI} 集成，得到一种全新的自适应平衡损失函数 $Loss$ ，并基于 $Loss$ 训练所提出的变化检测网络 HSPPNet。

$$Loss = L_{TCW} + \lambda L_{CI} \quad (7)$$

式中， λ 用来调整两种损失函数的权重。

3 实验结果与分析

3.1 实验数据和实验设计

为验证本文方法的有效性，采用两组常用公开数据集进行实验。第1组为CDD数据集 (Lebedev 等, 2018)，真值标签包括全部变化类型。第2组为 Google 数据集 (Peng 等, 2021)：Google 数据

为建筑物专题变化数据，真值标签只包括建筑物变化，用于验证本文方法对专题变化数据的适用性。

CDD 数据集 (Lebedev 等, 2018) 包含 16000 对尺寸为 256×256 像素的影像块，其中 10000 对为训练集，3000 对为验证集，3000 对为测试集。Google 数据集 (Peng 等, 2021) 包含 19 对尺寸从 1006×1168 像素到 4936×5224 像素的影像：首先将这些影像裁剪成 256×256 像素的影像块，然后随机选取 70% 作为训练集，10% 作为验证集，20% 作为测试集。为丰富数据特征，通过对称和旋转操作对两组数据进行增强。

为测试本文方法 HSPPNet 的性能，结合其自身特点组织如下对比试验：(1) 与本文方法最相关的经典语义分割网络 DeepLab v3+ (Chen 等, 2018b) 进行对比。DeepLab v3+ 首先利用深度空洞卷积神经网络提取两期影像级联影像的特征，然后用 ASPP 结构提取多尺度信息。(2) 与本文方法相关的 3 种深度学习变化检测模型进行对比，即 DSIFN (Zhang 等, 2020)、DASNet (Chen 等, 2021) 和 DifUNet++ (Zhang 等, 2022)：DASNet 首先用孪生 CNN 网络分别提取两期影像特征，然后利用双注意力机制 DAM 对所提取特征进行增强，最后基于增强后特征距离生成变化检测图。DASNet 和本文方法都用 DAM 增强特征。DSIFN 亦首先用孪生网络提取特征，然后基于深度监督策略生成变化检测图。DifUNet++ 以 UNet++ 为基础，通过多层级特征差分和多边输出融合策略实施变化检测。DSIFN、DifUNet++ 和本文方法均尝试利用两期影像的多尺度信息：DSIFN 通过深度监督策略，DifUNet++ 通过多层级特征差分和多边输出融合策略，而本文方法则通过设计 HSPP 模块实现。(3) 与两种最新提出的 Transformer 优化 CNN 的变化检测模型 BIT (Chen 等, 2022) 和 MSCANet (Liu 等, 2022) 进行对比。BIT 首先用孪生 CNN 分别提取两期影像特征，然后使用 Transformer 模块基于语义空间对所提取特征进行增强，最后基于两时期特征的差异生成变化检测图。MSCANet 首先用孪生 CNN 提取两期影像的多层级特征，然后基于 Transformer 对多层级特征进行聚合，最后通过多分支分类器检测变化区域。

采用 4 种深度学习变化检测文献中常用的精度指标来定量评价不同方法的变化检测结果：准确

度 Pre (Precision)、召回率 Rec (Recall)、总体准确率 OA (Overall Accuracy) 和 F1 值。

HSPPNet 涉及两个超参数： α 和 λ 。 α 表示注意力引导卷积层 CAconv 和 PAconv 中使用卷积核个数， λ 用于调整所改进两种损失函数的权重。本文通过验证集确定两个超参数的取值：在 CDD 和 Google 数据集上， α 的取值分别为 256 和 32； λ 的取值分别为 0.5 和 0.3。“超参数 α 和 λ 对变化检测结果的

影响分析”和“通过验证集确定超参数的有效性分析”请参见 3.3.2 节超参数消融实验。

3.2 实验结果分析

图 3 和图 4 分别给出 CDD 和 Google 数据集上 4 组典型变化检测结果图：前 3 列分别为 T_1 时刻影像、 T_2 时刻影像以及其真值标签；后 7 列分别为 DeepLab v3+、DSIFN、DASNet、DifUNet++、BIT、MSCANet 和本文方法 HSPPNet 的变化检测图。

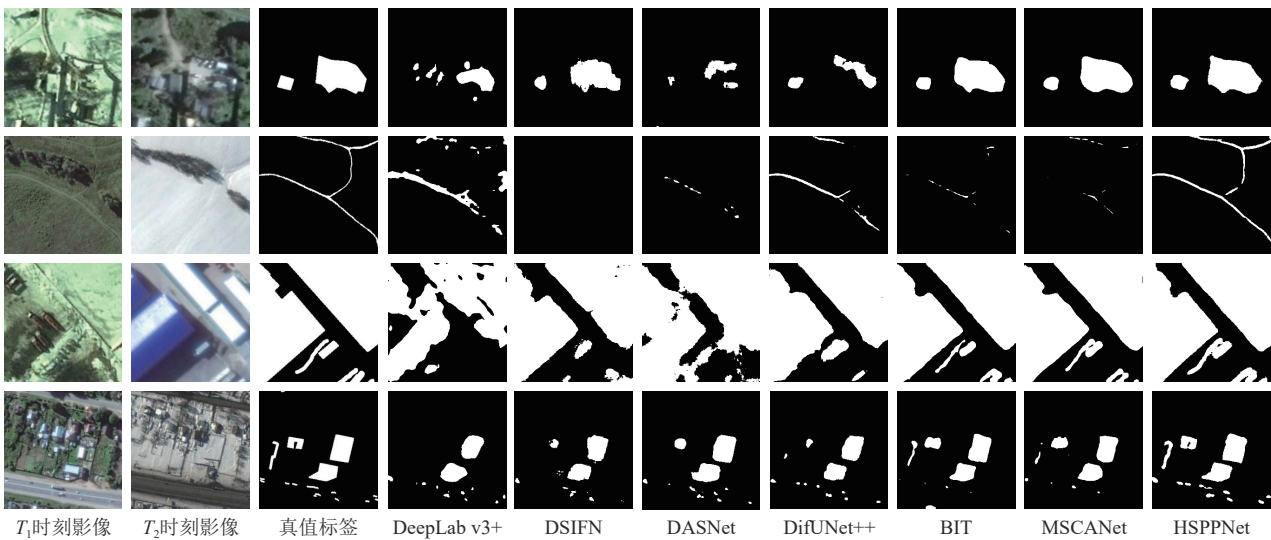


图3 不同方法在CDD数据集上的变化检测结果

Fig. 3 Change detection results from different methods on CDD dataset

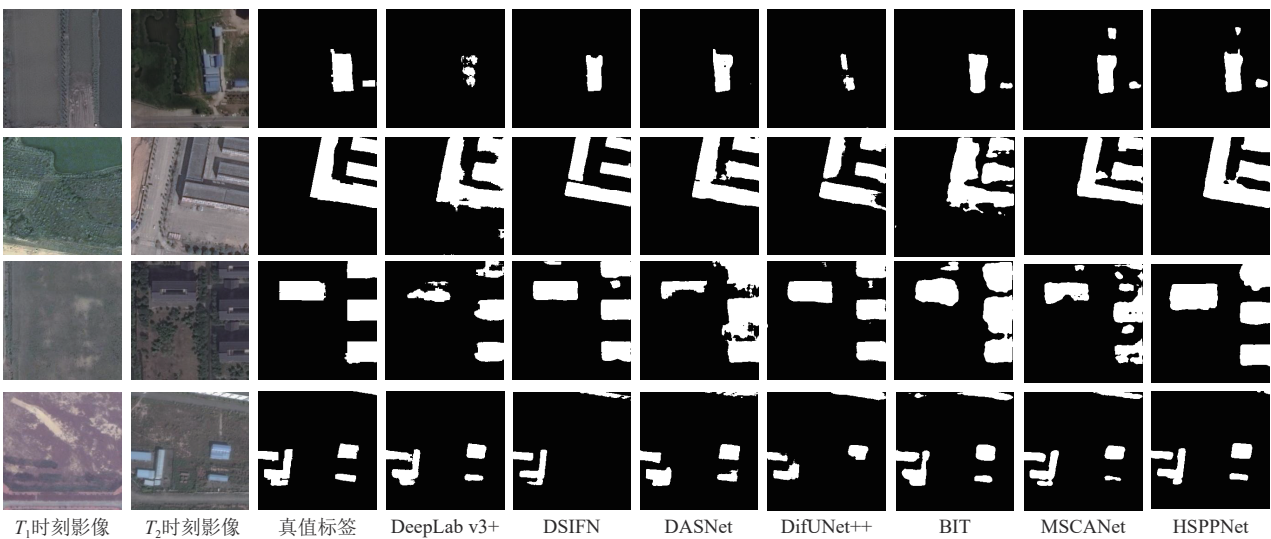


图4 不同方法在Google数据集上的变化检测结果

Fig. 4 Change detection results from different methods on Google dataset

从图 3 可知，对于 CDD 数据集，本文方法 HSPPNet 得到与真值标签最接近的变化检测图：HSPPNet 对影像中道路变化、汽车变化、不同尺寸的建筑物变化，均能比较完整的检测出，且边界

较为准确；表明 HSPPNet 对不同形状尺度的变化对象均有良好的检测效果。而 6 种对比方法表现则不够理想：比如在第 2 行中，6 种对比方法对道路变化的检测均不够完整；而在第 4 行中，6 种对比

方法均有明显的漏检。从图4可知，对于Google数据集，所提出的变化检测网络HSPPNet也得到最优变化检测图：比如在第1行中，DeepLab v3+和DifUNet++具有明显漏检，DSIFN和DASNet遗漏较小建筑物变化，而BIT和MSCANet取得与本文方法类似的结果。在第2行中，虽然7种方法都能检测出建筑物变化的大致轮廓，但与6种对比方法相比，本文方法的检测结果更加完整，边界更加准确。在第3行中，DeepLab v3+和MSCANet具有明显的漏检错误，DASNet和BIT具有较多虚检，DSIFN和DifUNet++取得与本文方法相似的结果。总体来讲，本文方法在Google数据上的变化检测图也更

加准确。

为更加客观地评价不同变化检测方法的检测性能，表1给出不同方法在两组数据集上变化检测结果的定量评价指标。从表1可知，在两组数据集上，本文方法HSPPNet均得到最优定量指标。与其他方法相比，HSPPNet在保证准确度（Pre）的情况下，显著提升了召回率（Rec）与F1值，比如对CDD数据集，HSPPNet的召回率比其他对比方法至少提升4.32%。同时，HSPPNet在两组数据上均得到最高的OA和F1值。比如对CDD数据集，HSPPNet的F1值为0.9758，比其他方法至少提高2.66%。

表1 不同方法变化检测结果的定量分析指标

Table 1 Quantitative indices of the change detection results from different methods

方法	CDD				Google			
	Pre	Rec	OA	F1	Pre	Rec	OA	F1
DeepLab v3+	0.9042	0.8096	0.9684	0.8543	0.9137	0.8311	0.9191	0.8704
DSIFN	0.9496	0.8711	0.9771	0.9030	0.9735	0.8223	0.9346	0.8916
DASNet	0.9220	0.9320	0.9820	0.9270	0.9067	0.8763	0.9300	0.8912
DifUNet++	0.9215	0.9463	0.9806	0.9337	0.9048	0.8571	0.9238	0.8803
BIT	0.9579	0.9407	0.9881	0.9492	0.8926	0.8342	0.9129	0.8624
MSCANet	0.9444	0.9460	0.9870	0.9452	0.8935	0.8826	0.9272	0.8880
HSPPNet	0.9624	0.9895	0.9943	0.9758	0.9142	0.9015	0.9399	0.9076

3.3 消融实验

3.3.1 网络结构消融实验

本文方法HSPPNet的创新主要包括：（1）提出一种简单有效的输入模块；（2）构建一种混合空间金字塔池化模块HSPP；（3）定义一种全新的平衡损失函数。接下来以CDD数据集为例来分别验证输入模块、HSPP和所定义损失函数的有效性。为此，表2分别给出Net-1、Net-2、Net-3-1、Net-3-2和本文方法HSPPNet的变化检测结果：Net-1表示去掉HSPPNet中输入模块得到的网络（Net-1将两期影像级联作为输入）；Net-2表示用传统ASPP模块替换HSPPNet中HSPP模块得到的网络；Net-3-1和Net-3-2分别表示用加权交叉熵损失函数 L_{BCE} （Peng等，2019）和截断—补偿加权交叉熵损失函数 L_{TCW} （式（5））替换HSPPNet中损失函数得到的网络。

（1）通过对比Net-1和HSPPNet的变化检测结果可知：使用所提出的输入模块能够明显增强模型的变化检测性能，使得在4个精度指标上均有较

大提升，验证了输入模块的有效性。

（2）通过对比Net-2和HSPPNet的变化检测结果可知：使用所提出的HSPP模块使得在Pre、Rec、OA和F1值上分别提高3.31%、9.83%、1.45%和6.60%，说明HSPP能够显著提升模型的检测性能。

（3）通过对比Net-3-1、Net-3-2和HSPPNet的变化检测结果可知：本文给出的截断—补偿加权交叉熵损失函数 L_{TCW} （式（5））和类别级IoU损失函数 L_{CI} （式（6））均能提升模型性能，比如两者使得F1值分别提高1.21%和0.69%。

表2 网络Net-1、Net-2、Net-3-1、Net-3-2和HSPPNet的变化检测结果

Table 2 Change detection results from Net-1、Net-2、Net-3-1、Net-3-2 and HSPPNet

网络	Pre	Rec	OA	F1
Net-1	0.9260	0.9625	0.9867	0.9438
Net-2	0.9293	0.8912	0.9798	0.9098
Net-3-1	0.9318	0.9832	0.9897	0.9568
Net-3-2	0.9477	0.9910	0.9926	0.9689
HSPPNet	0.9624	0.9895	0.9943	0.9758

为进一步说明所提出 HSPP 模块相较于现有 ASPP 的优势, 图 5 给出 Net-2 和本文方法 HSPPNet 在 CDD 数据集上的 3 组典型变化检测图。表 3 给出 Net-2 与 HSPPNet 在变化对象边界区域的变化检测结果。Net-2 表示用传统 ASPP 替换 HSPPNet 中 HSPP 得到的网络。变化对象边界区域的获取方式如下: 首先利用 Canny 算子提取真值标签图的边界, 然后通过高斯模糊算子扩充边界宽度。

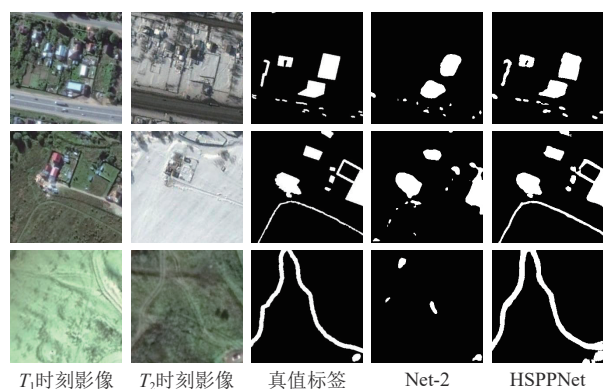


图5 CDD数据集中3组典型影像及其变化检测结果

Fig. 5 Three typical image pairs and their change detection results from CDD dataset

表3 网络Net-2和HSPPNet在变化对象边界区域的变化检测结果

Table 3 Change detection results on boundary areas of change objects obtained by Net-2 and HSPPNet

网络	Pre	Rec	OA	F1
Net-2	0.6929	0.7002	0.6034	0.6482
HSPPNet	0.7601	0.9354	0.8285	0.8387

对比图 5 (c)、(d) 和 (e) 可知: 虽然 ASPP 能够较为有效地检测出不同尺度的块状变化对象, 但其不能很好地识别线状变化对象和块状变化

对象的边界; 而 HSPP 能够有效地提取不同尺度、不同形状的变化对象。实验结果表明, 所提出的 HSPP 在线状变化对象和块状变化对象边界区域的检测效果明显优于 ASPP。从定量的角度, HSPP 在变化对象边界区域上的 Pre、Rec、OA 和 F1 指标比 ASPP 分别提高 6.72%、23.52%、22.51% 和 19.05% (表 3)。

3.3.2 超参数消融实验

(1) 为测试超参数 α 和 λ 不同取值对变化检测结果的影响, 在综合考虑计算成本和验证效果的前提下, 设计如下消融实验: 首先通过初步实验, 初步确定 α 和 λ 的取值; 然后固定其中一个超参数的值, 验证另一个超参数对变化检测结果的影响。(2) 为展示“通过验证集确定超参数取值”的有效性, 本文同时给出验证集和测试集上针对不同 $\alpha(\lambda)$ 取值的变化检测结果。接下来以 CDD 数据集为例进行实验分析。具体如下:

测试超参数 α 时, 固定超参数 λ 的取值为 0.5, 并令 α 在集合 {32, 64, 128, 256} 上取值。表 4 给出本文方法在 CDD 数据验证集和测试集上针对不同 α 取值的变化检测结果。

从表 4 可知: (1) α 的取值对所提出变化检测网络 HSPPNet 的影响较大。对于 CDD 数据, α 取值为 256 时, HSPPNet 变化检测结果最优, 其 F1 值在验证集和测试集上分别达到 0.9799 和 0.9758。(2) 虽然对给定 α 值, 本文方法 HSPPNet 在验证集和测试集上的变化检测结果存在差异, 但 HSPPNet 的变化检测结果随着超参数 α 的变化趋势在验证集和测试集上基本保持一致。因此, 通过验证集确定超参数 α 的取值是可行的。

表4 超参数 α 取不同值时的变化检测结果

Table 4 Change detection results on different values of hyper-parameter α

λ 取值	验证集				测试集			
	Pre	Rec	OA	F1	Pre	Rec	OA	F1
32	0.7925	0.7426	0.9495	0.7668	0.8854	0.8213	0.9675	0.8522
64	0.8217	0.7771	0.9559	0.7988	0.8976	0.7931	0.9660	0.8421
128	0.8914	0.7860	0.9655	0.8354	0.9316	0.9219	0.9833	0.9267
256	0.9732	0.9866	0.9955	0.9799	0.9624	0.9895	0.9943	0.9758

测试超参数 λ 时, 固定超参数 α 值为 256。为充分验证 λ 对变化检测结果的影响, 通过综合考虑计算成本和验证效果, 设计分阶段验证策略: 第 1 阶段采用较大步长, 将步长设置为 0.2, 并令 λ

在集合 {0.1, 0.3, 0.5, 0.7, 0.9} 上取值; 初步测试 λ 对变化检测结果的影响, 并确定超参数 λ 在上述集合上的最优取值 λ_{o_1} 。第 2 阶段采用较小步长, 在 λ_{o_1} 值附近进一步测试超参数 λ 的影响: 将步长

设置为 0.05，并令 λ 在集合 $\{\lambda_{o_1}-0.1, \lambda_{o_1}-0.05, \lambda_{o_1}, \lambda_{o_1}+0.05, \lambda_{o_1}+0.1\}$ 取值；为避免 λ 取值为 0，若 $\lambda_{o_1}=0.1$ ，则需从上述集合中去除 $\lambda_{o_1}-0.1$ 。

对于 CDD 数据集，超参数 λ 在第 1 阶段最优取值为 0.5，即 $\lambda_{o_1}=0.5$ ，故第 2 阶段在集合 $\{0.4, 0.45, 0.5, 0.55, 0.6\}$ 上取值。表 5 给出本文方法在 CDD 数据验证集和测试集上针对不同 λ 取值的变化检测结果。

表 5 超参数 λ 取不同值时的变化检测结果
Table 5 Change detection results on different values of hyper-parameter λ

λ 取值	验证集				测试集			
	Pre	Rec	OA	F1	Pre	Rec	OA	F1
0.1	0.9615	0.9765	0.9930	0.9689	0.9634	0.9713	0.9925	0.9673
0.3	0.9692	0.9716	0.9934	0.9704	0.9721	0.9732	0.9937	0.9726
0.4	0.9661	0.9772	0.9937	0.9716	0.9669	0.9775	0.9936	0.9722
0.45	0.9707	0.9788	0.9942	0.9748	0.971	0.9782	0.9943	0.9746
0.5	0.9732	0.9866	0.9955	0.9799	0.9624	0.9895	0.9943	0.9758
0.55	0.9704	0.9801	0.9944	0.9752	0.9697	0.9809	0.9943	0.9753
0.6	0.9639	0.9822	0.9939	0.9730	0.965	0.9836	0.994	0.9742
0.7	0.9675	0.9805	0.9941	0.9739	0.9671	0.9804	0.9939	0.9737
0.9	0.9641	0.975	0.9932	0.9695	0.9613	0.9751	0.9926	0.9681

从表 5 可知：(1) λ 的取值对所提出变化检测网络 HSPPNet 的影响较小，特别是在区间 $[0.3, 0.7]$ 上。对于 CDD 数据集， λ 取值为 0.5 时，HSPPNet 变化检测结果最优，其 F1 值在验证集和测试集上分别达到 0.9799 和 0.9758。(2) 虽然对给定 λ 值，本文方法在验证集和测试集上的变化检测结果存在差异，但本文方法的变化检测结果随着超参数 λ 的变化趋势在验证集和测试集上基本保持一致。因此，通过验证集确定超参数 λ 的取值是合理的。

3.3.3 模型复杂度分析

本小节分析不同方法的模型复杂度，为此，表 6 给出不同方法的参数量和浮点运算次数 FLOPs (Floating point operations)。

表 6 不同方法的参数量和浮点运算次数
Table 6 The numbers of model parameters and floating point operations for different methods

方法	参数量/M	FLOPs/G
DeepLab v3+	42.52	45.89
DSIFN	35.99	82.35
DASNet	50.27	100.64
DifUNet++	9.1	60.60
BIT	11.94	8.71
MSCANet	14.7	16.59
HSPPNet	57.82	61.39

从表 6 可知：对于参数量，本文方法 HSPPNet 为 57.82 M，与对比方法 DASNet (50.27 M) 接近，

高于其他对比方法。对于 FLOPs 指标，本文方法 HSPPNet 为 61.39 G，与对比方法 DifUNet++ (60.60 G) 接近，高于 DeepLab v3+ (45.89 G)、BIT (8.71 G) 和 MSCANet (16.59 G)，而低于 DSIFN (82.35 G) 和 DASNet (100.64 G)。

4 结 论

针对高分辨率遥感影像变化检测，本文提出一种端到端混合空间金字塔池化网络 HSPPNet。HSPPNet 通过以下 3 点增强变化检测性能：(1) 将空洞卷积和注意力引导卷积集成，构建一种混合空间金字塔池化模块，提升对不同形状和尺度变化对象的检测能力。(2) 提出一种全新的自适应平衡损失函数，来缓解变化类和未变化类严重不平衡对变化检测的影响。(3) 提出一种输入模块，通过综合考虑两期影像及其差异图，来增强变化信息；并利用数据依赖型上采样方法替换双线性插值，提升解码阶段对影像的重构能力。

两组常用公开变化检测数据集的实验结果表明：本文方法 HSPPNet 的检测结果在定性和定量两方面均优于 DeepLab v3+、DSIFN、DASNet、DifUNet++、BIT 和 MSCANet 等 6 组对比方法。

HSPPNet 通过直接级联的方式融合 HSPP 模块 7 个分支的特征，今后将尝试用通道注意力机制来优化 HSPP 模块。

志 谢 论文中的实验受到三峡大学先进计算中心算力的支持。

参考文献(References)

- Chen H, Qi Z P and Shi Z W. 2022. Remote sensing image change detection with transformers. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5607514 [DOI: 10.1109/TGRS.2021.3095166]
- Chen H and Shi Z W. 2020. A spatial-temporal attention-based method and a new dataset for remote sensing Image change detection. *Remote Sensing*, 12(10): 1662 [DOI: 10.3390/rs12101662]
- Chen J, Lu M, Chen X H, Chen J and Chen L J. 2013. A spectral gradient difference based approach for land cover change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 85: 1-12 [DOI: 10.1016/j.isprsjprs.2013.07.009]
- Chen J, Yuan Z Y, Peng J, Chen L, Huang H Z, Zhu J W, Liu Y and Li H F. 2021. 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*, 14: 1194-1206 [DOI: 10.1109/JSTARS.2020.3037893]
- Chen L C, Papandreou G, Kokkinos I, Murphy K and Yuille A L. 2018a. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4): 834-848 [DOI: 10.1109/TPAMI.2017.2699184]
- Chen L C, Zhu Y K, Papandreou G, Schroff F and Adam H. 2018b. Encoder-decoder with atrous separable convolution for semantic image segmentation//15th European Conference on Computer Vision. Munich: Springer: 833-851 [DOI: 10.1007/978-3-030-01234-2_49]
- Cheng H Q, Wu H Y, Zheng J, Qi K L and Liu W X. 2021. A hierarchical self-attention augmented Laplacian pyramid expanding network for change detection in high-resolution remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 182: 52-66 [DOI: 10.1016/j.isprsjprs.2021.10.001]
- Daudt R C, Le Saux B and Boulch A. 2018. Fully convolutional Siamese networks for change detection//Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP). Athens: IEEE: 4063-4067 [DOI: 10.1109/ICIP.2018.8451652]
- Dian Y Y, Fang S H and Yao C H. 2016. Change detection for high-resolution images using multilevel segment method. *Journal of Remote Sensing (in Chinese)*, 20(1): 129-137 (佘袁勇, 方圣辉, 姚崇怀. 2016. 多尺度分割的高分辨率遥感影像变化检测. *遥感学报*, 20(1): 129-137) [DOI: 10.11834/jrs.20165074]
- Fu J, Liu J, Tian H J, Li Y, Bao Y J, Fang Z W and Lu H Q. 2020. Dual attention network for scene segmentation//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE: 3141-3149 [DOI: 10.1109/CVPR.2019.00326]
- Gu L, Xu S Q and Zhu L Q. 2020. Detection of building changes in remote sensing images via FlowS-Unet. *Acta Automatica Sinica*, 46(6): 1291-1300 (顾炼, 许诗起, 竺乐庆. 2020. 基于FlowS-Unet的遥感图像建筑物变化检测. *自动化学报*, 46(6): 1291-1300) [DOI: 10.16383/j.aas.c180122]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hussain M, Chen D M, Cheng A, Wei H and Stanley D. 2013. Change detection from remotely sensed images: from pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, 80: 91-106 [DOI: 10.1016/j.isprsjprs.2013.03.006]
- Lebedev M A, Vizilter Y V, Vygolov O V, Knyaz V A and Rubis A Y. 2018. Change detection in remote sensing images using conditional adversarial networks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2: 565-571 [DOI: 10.5194/isprs-archives-XLII-2-565-2018]
- Liang Z H, Li X, Deng P, Sheng S and Jiang F Q. 2022. Remote sensing image change detection fusion method integrating multi-scale feature attention. *Acta Geodaetica et Cartographica Sinica*, 51(5): 668-676. (梁哲恒, 黎宵, 邓鹏, 盛森, 姜福泉. 2022. 融合多尺度特征注意力的遥感影像变化检测方法. *测绘学报*, 51(5): 668-676) [DOI: 10.11947/j.AGCS.2022.20200540]
- Liu J F, Chen K M, Xu G L, Sun X, Yan M L, Diao W H and Han H Z. 2020. Convolutional neural network-based transfer learning for optical aerial images change detection. *IEEE Geoscience and Remote Sensing Letters*, 17(1): 127-131 [DOI: 10.1109/LGRS.2019.2916601]
- Liu M X, Chai Z Q, Deng H J and Liu R. 2022. A CNN-Transformer network with multiscale context aggregation for fine-grained cropland change detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15: 4297-4306 [DOI: 10.1109/JSTARS.2022.3177235]
- Peng D F, Bruzzone L, Zhang Y J, Guan H Y, Ding H Y and Huang X. 2021. SemiCDNet: a semisupervised convolutional neural network for change detection in high resolution remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 59(7): 5891-5906 [DOI: 10.1109/TGRS.2020.3011913]
- Peng D F, Zhang Y J and Guan H Y. 2019. End-to-end change detection for high resolution satellite images using improved UNet++. *Remote Sensing*, 11(11): 1382 [DOI: 10.3390/rs11111382]
- Qi J W, Wang W F, Zhang L and Wang G Y. 2023. Building change detection of remote sensing image based on improved DeepLabV3+. *Bulletin of Surveying and Mapping*, (4): 145-149 (齐建伟, 王伟峰, 张乐, 王光彦. 2023. 基于改进DeepLabV3+算法的遥感影像建筑物变化检测. *测绘通报*, (4): 145-149) [DOI: 10.13474/j.cnki.11-2246.2023.0119]
- Shi W Z, Zhang M, Zhang R, Chen S X and Zhan Z. 2020. Change detection based on artificial intelligence: state-of-the-art and challenges. *Remote Sensing*, 12(10): 1688 [DOI: 10.3390/rs12101688]
- Tian Z, He T, Shen C H and Yan Y L. 2020. Decoders matter for semantic segmentation: data-dependent decoding enables flexible

- feature aggregation//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE: 3121-3130 [DOI: 10.1109/CVPR.2019.00324]
- Wang C, Wang S, Chen X, Li J Y and Xie T. 2023. Object-level change detection of multi-sensor optical remote sensing images combined with UNet++ and multi-level difference module. *Acta Geodaetica et Cartographica Sinica*, 52(2): 283-296 (王超, 王帅, 陈晓, 李俊勇, 谢涛. 2023. 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. *测绘学报*, 52(2): 283-296) [DOI: 10.11947/j.AGCS.2023.20220202]
- Zhang C X, Yue P, Tapete D, Jiang L C, Shangguan B Y, Huang L and Liu G C. 2020. A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 166: 183-200 [DOI: 10.1016/j.isprsjprs.2020.06.003]
- Zhang M and Shi W Z. 2020. A feature difference convolutional neural network-based change detection method. *IEEE Transactions on Geoscience and Remote Sensing*, 58(10): 7232-7246 [DOI: 10.1109/TGRS.2020.2981051]
- Zhang X W, Yue Y Z, Gao W X, Yun S, Su Q, Yin H L and Zhang Y N. 2022. DifUNet++: a satellite images change detection network based on UNet++ and differential pyramid. *IEEE Geoscience and Remote Sensing Letters*, 19: 8006605 [DOI: 10.1109/LGRS.2021.3049370]
- Zhou Z W, Siddiquee M M R, Tajbakhsh N and Liang J M. 2020. UNet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging*, 39(6): 1856-1867 [DOI: 10.1109/TMI.2019.2959609]

Hybrid spatial pyramid pooling network to change detection for high resolution remote sensing images

SHAO Pan, GAO Zi'ang

1. Hubei Key Laboratory of Intelligent Vision Monitoring for Hydroelectric Engineering, China Three Gorges University, Yichang 443002, China;

2. College of Computer and Information Technology, China Three Gorges University, Yichang 443002, China

Abstract: Change Detection (CD) is the process of identifying the land cover changes of interest by analyzing bitemporal remote sensing images acquired in the same geographical area at different times. As a popular research topic in remote sensing, CD plays an important role in many practical applications, such as urban research, environmental monitoring, and disaster assessment. With the rapid development of deep learning, its applications have expanded to the remote sensing CD task, leading to the development of numerous deep learning-based CD methods. Although the existing deep learning-based CD methods have their own advantages and can implement the CD task effectively, their applicability and robustness need to be further analyzed. How to design an appropriate network model to obtain more reliable and accurate CD results is still an open problem. Currently, two main challenges remain for CD: (1) The changed objects existing in remote sensing images have various scales and shapes. (2) The proportions of the changed and unchanged classes are often significantly unbalanced, i.e., the number of the changed pixels is usually much less than that of the unchanged pixels. To deal with these two challenges, this study proposes a novel end-to-end CD network for high-resolution remote sensing images based on a residual network, i.e., Hybrid Spatial Pyramid Pooling Network (HSPPNet). First, a Hybrid Spatial Pyramid Pooling (HSPP) module is built by integrating atrous convolutions and attention mechanism-guided convolutions in parallel to effectively extract the changed objects with different shapes and scales from high-resolution remote sensing images. Then, an adaptive balancing loss function is presented to alleviate the effects of the serious imbalance between changed and unchanged classes on CD. The loss function is constructed by defining a truncation - compensation weighting cross entropy function and a class-level IoU function and integrating them. Finally, a simple but effective input module that considers the bitemporal remote sensing images and their difference image is designed to enhance the change information. Owing to the above three points, the proposed HSPPNet method enhances the performance of deep learning CD. Experiments are carried out with two public CD datasets, namely, the CDD and Google datasets, to evaluate the performance of the proposed HSPPNet method. Six state-of-the-art deep learning methods are used as the comparative methods. Quantitative analysis of CD results is done on four evaluation metrics, namely, precision, recall, overall accuracy, and F1-score. For the CDD and Google datasets, the overall accuracy/F1-score values of the proposed HSPPNet method are 0.9943/0.9758 and 0.9399/0.9076, respectively, which are superior to those of the six comparative methods. The experimental results on two public CD datasets demonstrate that the proposed HSPPNet is effective and feasible. In addition, the results of ablation experiments show that the HSPP module, the adaptive balancing loss function, and the input module proposed in this work can enhance the CD performance effectively.

Key words: remote sensing, change detection, deep learning, hybrid spatial pyramid pooling, attention mechanism, loss function, truncation-compensation weighting cross entropy, class-level IoU

Supported by National Natural Science Foundation of China (No. 41901341)