

基于深度学习的城市区域变化检测技术研究

廖永皓^{1*}, 廖明伟^{2,3#}, 梅承华⁴

¹江西省自然资源政策调查评估中心, 江西 南昌

²江西省国土空间调查规划研究院, 江西 南昌

³自然资源部大湖流域国土空间生态保护修复工程技术创新中心, 江西 南昌

⁴南昌市测绘勘察研究院有限公司, 江西 南昌

收稿日期: 2025年9月24日; 录用日期: 2025年10月14日; 发布日期: 2025年10月27日

摘要

城市日新月异的发展, 如何保障地理信息公共服务地图数据的鲜活性和现势性成为了刻不容缓的瓶颈问题。因此探索利用深度学习的方式自动发现变化区域、变化地物, 也即进行不同时相影像的变化检测, 提升矢量数据更新效率、更新效果非常必要而且效果良好。

关键词

深度学习, 城市区域, 变化检测, 分类

Discussion on Urban Area Change Detection Based on Deep Learning

Yonghao Liao^{1*}, Mingwei Liao^{2,3#}, Chenghua Mei⁴

¹Jiangxi Provincial Natural Resources Policy Survey and Evaluation Centre, Nanchang Jiangxi

²Jiangxi Institute of Land Space Survey and Planning, Nanchang Jiangxi

³Technology Innovation Center for Land Spatial Ecological Protection and Restoration in Great Lakes Basin MNR, Nanchang Jiangxi

⁴Nanchang Surveying and Mapping Research Institute Co., Ltd., Nanchang Jiangxi

Received: September 24, 2025; accepted: October 14, 2025; published: October 27, 2025

Abstract

The rapid urbanization necessitates urgent solutions for ensure the timeliness and accuracy of public

*第一作者。

#通讯作者。

文章引用: 廖永皓, 廖明伟, 梅承华. 基于深度学习的城市区域变化检测技术研究[J]. 图像与信号处理, 2025, 14(4): 505-513. DOI: 10.12677/jisp.2025.144046

geographic information service map data. This study explores the application of deep learning methodologies for automated change detection in multi-temporal remote sensing images. The proposed approach effectively identifies modified regions and evolving geographic features, thereby enhancing both the efficiency and accuracy of vector data updates.

Keywords

Deep Learning, Urban Areas, Change Detection, Classification

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 研究背景

地理信息服务是为了吸引更多潜在的用户，提高地理信息系统的利用率而建立的一种面向服务的模式，用户可以通过互联网按需获得和使用地理数据和计算服务，如地图服务、空间数据格式转换等。地理信息服务的目标是让任何人、任何时间、任何地点都可以获取任何空间信息。近年来，随着信息时代的发展，人们对地理信息日益增长的社会需求，地理信息服务早已成为服务政府、企业、大众的重要方式；全国组织开展的数字城市、天地图建设，为不同用户提供了各种地图的实时在线服务。

随着时间的推移、社会经济的飞速发展、城市化进程的加快，城市变化越来越频繁[1]，如何保障地理信息公共服务地图数据的鲜活性和现势性成为了刻不容缓的瓶颈问题。而传统更新手段已经完全不能满足行业应用对现势性的强烈要求。因此，为了提供有效优质的公共服务给政府、公众，充分发挥地理信息公共服务平台服务的最大效益，将城市的发展状况快速体现在地理信息公共服务地图上，迫切需要一种快速稳定的城市更新体系来支撑广大用户的服务需求。

2. 研究内容

对于地理信息公共服务平台矢量数据的更新，首先需要及时获取完整的变化信息；目前采用的方式是将原有矢量数据与经过校正后的遥感影像套合比对，目视判定地物变化情况，并把需要进行更新的内容逐一找出，然后对相应数据进行增加补充或删除修改。地物变化判别主要根据标志物的光谱、纹理、形状特征对影像中地物类别及范围进行判读。个人的经验直接影响目视判读准确度。更新作业时对作业人员熟练掌握卫片判读有较高要求，需尽量避免因为错误判读而导致错误更新或更新遗漏。而且通过人工比对查看发现城市地物变化的工作量太大、地物判读准确性受到限制。因此探索研究自动发现变化区域、变化地物，也即进行不同时相遥感影像的变化检测技术，提升矢量数据更新效率、更新效果非常必要。

随着对地观测技术、信息技术、人工智能的发展，变化检测方法的演变不断进行[2]。利用深度学习算法对高分辨率遥感影像的像素进行语义分割，赋予每个像素类别意义，可以解决传统方法中难以提取抽象特征、复杂场景解析能力较差的问题[3]。变化检测分为分类前比较方法和分类后比较方法两种[4]。本项目采用分类后检测的两步法实现变化检测。在分类精度较高的情况下，变化检测也有较高精度。为了减少在人工标注样本中的人力和物力，采用迁移学习法把已标记样本或模型通过一定的方式迁移到待分类影像中。由于不同的大气条件、光照条件、季节干扰以及传感器自身矫正等因素，导致卫星数据质量差异较大[5]，会导致已标记的样本和待分类的影像在数据分布上存在偏移，因此本文利用一种基于中心损失的对抗学习框架对影像进行分类，减少数据分布偏移和人工标注样本的工作量。在高精度分类结

果基础上,对两时相影像预测类别结果进行比较得到最终的变化检测结果。

3. 研究方法

我们提出的变化检测方法分为两个阶段:第一阶段为场景分类,采用基于中心损失的对抗学习框架进行场景分类;第二阶段为变化检测,通过对两时相分类结果进行比对生成变化图。

第一阶段分类模型采用基于中心损失的对抗学习框架,其核心结构由三部分组成:源域特征提取器、目标域特征提取器和域分类器。特征提取器的骨干网络采用 ResNet101;为了提高源域特征提取器的精度,采用困难三元组损失函数,为了提高从源域到目标域的迁移适应性,采用对抗损失和中心损失进行约束,整体结构如图 1 所示。

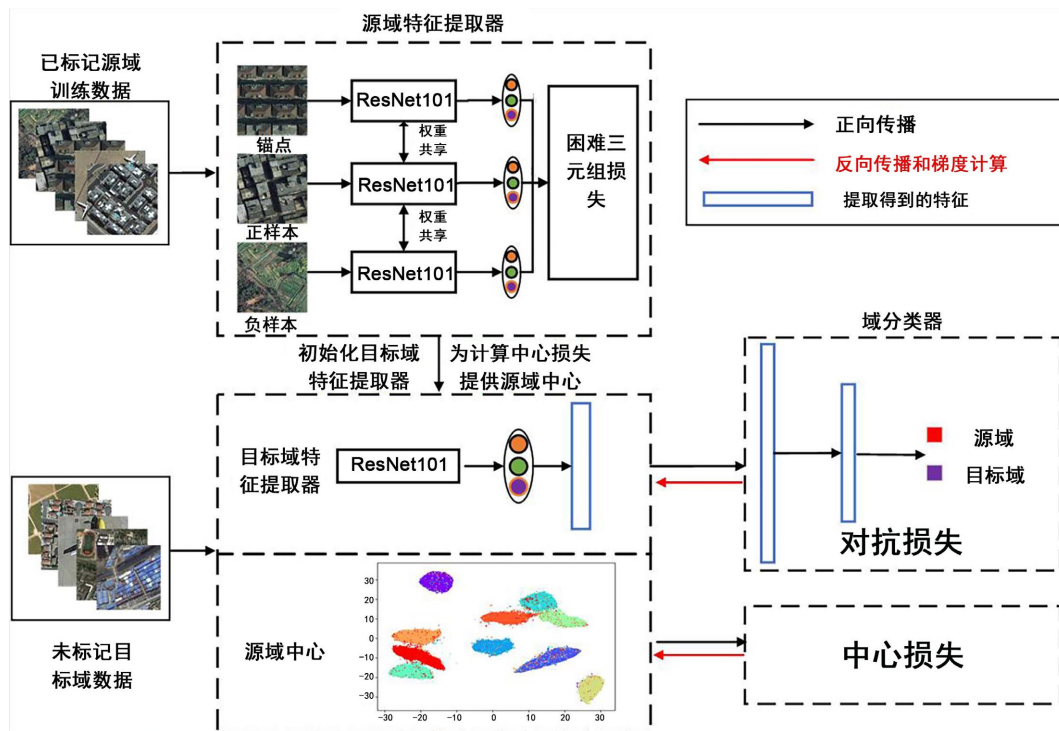


Figure 1. Scene classification based on the center loss adversarial learning framework

图 1. 基于中心损失对抗学习框架的场景分类

下面将特征提取骨干网络、源域训练损失、对抗和中心损失三个方面进行介绍。

1) 骨干网络

源域与目标域特征提取器均采用 ResNet101 为骨干网络的三元组网络(Triplet Network),用于提取高层语义特征。域分类器为一个二分类卷积神经网络,用于区分特征来自源域还是目标域。在骨干网络的选择上,ResNet101 凭借其深层残差结构,能有效缓解梯度消失问题,提升特征表达能力,尤其在处理高分辨率遥感影像时表现出优异的分类性能,适用于本项目中复杂城市场景的特征提取任务。

其中,ResNet101 的神经网络结构如图 2 所示。

2) 源域训练损失

损失函数的确定基于最困难三元组损失和半困难三元组损失函数,三元组中的困难和半困难类别取

决于锚点样本的特征与正、负样本的特征之间的欧式距离。

当两个特征欧氏距离达到最大值时，表示为“最困难”的三元组类别：

$$\max \left\| f_{\theta_s}(a_j) - f_{\theta_s}(p_k) \right\|^2 - \left\| f_{\theta_s}(a_j) - f_{\theta_s}(n_m^l) \right\|^2 \quad (3.1)$$

其中： α_j^i ：锚点(Anchor)，代表第 i 个类别中的第 j 个样本的。

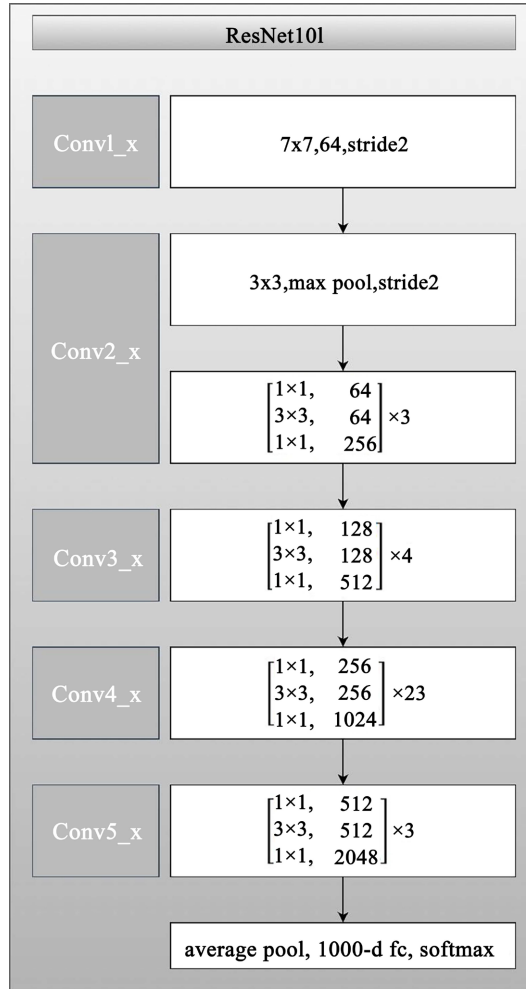


Figure 2. ResNet101 neural network

图 2. ResNet101 神经网络

p_k^i ：正样本(Positive)， k 从 1 到 j 遍历，max 表示着选取的是与锚点 α_j^i 同属于第 i 类的所有样本中，距离锚点最远，即“最困难”的正样本。

n_m^l ：负样本(Negative)， m 从 1 到 l (但不能等于锚点的类别 i) 遍历。

$f_{\theta_s}()$ ：表示某样本的特征。例如 $f_{\theta_s}(a_j^i)$ 表示锚点样本的特征。

当正、负样本特征与锚点样本间的欧氏距离在设定区域内时，表示为“半困难”的三元组类别：

$$\left\| f_{\theta_s}(a_j^i) - f_{\theta_s}(p_k^i) \right\|^2 < \left\| f_{\theta_s}(a_j^i) - f_{\theta_s}(n_m^l) \right\|^2 < \left\| f_{\theta_s}(a_j^i) - f_{\theta_s}(p_k^i) \right\|^2 + m \quad (3.2)$$

其中, α_j^i 、 p_k^j 、 n_m^l 仍表示锚点样本, 正样本和负样本。

m : 人工设定的边界阈值。

最困难三元组损失和半困难三元组损失可以分别提升三元组网络对困难样本和半困难样本的辨识能力。本文采用结合最困难三元组损失和半困难三元组损失的困难三元组损失函数进行约束训练:

$$L_H(\theta_s) = L_{BH}(\theta_s) + \beta L_{BS}(\theta_s) \quad (3.3)$$

其中, 其中权衡参数 β 表示半困难三元组的重要性, $L_{BH}(\theta_s)$ 和 $L_{BS}(\theta_s)$ 分别代表最困难三元组损失函数和半困难三元组损失函数, 这两种损失函数分别如式(3.4)与式(3.5)所示:

最困难三元组损失函数如下:

$$L_{BH}(\theta_s) = \sum_{i=1}^C \sum_{j=1}^K \max \left(m + \max_{k=1, \dots, K} \|f_{\theta_s}(\alpha_j^i) - f_{\theta_s}(p_k^j)\|^2 - \min_{\substack{l=1, \dots, C \\ k=1, \dots, K}} \|f_{\theta_s}(\alpha_j^i) - f_{\theta_s}(n_m^l)\|^2, 0 \right) \quad (3.4)$$

其中, C : 批次中的类别数。

K : 每个类别中的样本数。通常, 批次由 C 个类别, 每个类别 K 个样本组成, 因此总批次大小为 $C \times K$ 。

θ_s 为深度网络的参数, 即需要训练的参数。

半困难三元组损失函数如下:

$$L_{BS}(\theta_s) = \sum_{i=1}^C \sum_{j=1}^K \sum_{k=1, k \neq j}^K \sum_{l=1, l \neq i}^C \sum_{m=1}^K \max \left(\|f_{\theta_s}(\alpha_j^i) - f_{\theta_s}(p_k^j)\|^2 - \|f_{\theta_s}(\alpha_j^i) - f_{\theta_s}(n_m^l)\|^2 + m, 0 \right) \quad (3.5)$$

其中, C 、 K 、 θ_s 、 α_j^i 、 p_k^j 、 n_m^l 定义与最困难三元组损失函数一致。

3) 对抗损失与中心损失

传统的对抗网络框架包含源域特征提取器、目标域特征提取器和域分类器三个部分。对抗网络的中心思想是通过对抗训练促使特征提取器学习域不变特征。具体而言, 域分类器是一个二分类的判别器, 其核心任务是准确判断输入的特征向量是来自源域还是目标域。该分类器通常由若干全连接层构成, 其最终通过 Sigmoid 激活函数输出布尔值, 表示输入特征属于目标域的概率。在训练过程中, 域分类器的目标是最小化其自身的分类误差; 而特征提取器的目标则是生成能够欺骗域分类器的特征, 即最大化域分类器的误差。

该方法在确保源域提取到的特征有足够强的区分性的同时, 通过中心损失和域分类器反向传播减少源域和目标域特征提取器之间的特征分布差异, 使得源域提取的特征能够较好的区分目标域的影子。在现有对抗网络框架中加入中心损失函数, 在减少源域样本和目标域样本特征分布差异的同时一定程度上保持特征的辨识能力。

中心损失为目标域特征形成与源域特征相似的聚类中心, 提高目标域特征的辨识能力。

中心损失函数如下:

$$L_C(\theta_T) = \sum_{i \in T} \min_j \|f_{\theta_T}(x_i) - C_j\|^2 \quad (3.6)$$

其中: θ_T : 目标域的特征提取器参数。

x_i : 目标域样本。

C_j : 源域样本。

对抗损失为域分类器和目标域特征提取器的对抗, 减少源域目标域特征分布差异。对抗损失函数如下:

$$L_A(\theta_T, \theta_D) = \min_{\theta_D} \max_{\theta_T} - \sum_{i \in S} \log D_{\theta_D}(f_{\theta_S}(X_{S_i})) - \sum_{i \in T} \log(1 - D_{\theta_D}(f_{\theta_T}(X_{T_i}))) \quad (3.7)$$

其中: θ_T, θ_D : 特征提取器与域分类器参数。

max: 针对 θ_T ，特征提取器的目标是最大化这个损失，即生成让域分类器无法区分来源的特征(域不变特征)，从而欺骗判别器。

min: 针对 θ_D ，域分类器的目标是最小化这个损失，即提高自己区分特征来自源域还是目标域的能力。

$f_{\theta_S}(X_{S_i})$ ：特征提取器从源域样本中提取的特征。

$f_{\theta_T}(X_{T_i})$ ：特征提取器从目标域样本中提取的特征。

本文使用融合中心损失的对抗网络框架：

$$L(\theta_T, \theta_D) = L_A(\theta_T, \theta_D) + \frac{\lambda}{2} L_C(\theta_T) \quad (3.8)$$

其中， λ 为超参数权重，用于平衡两项损失的贡献。

第二阶段变化检测，在获得两时相影像的高精度分类结果后，通过逐像素比对类别标签，生成变化检测二值图，最终输出变化检测结果。

4. 研究过程

4.1. 试验数据

试验数据取自新余市同一地域的 2011 和 2018 年获取的天地图 19 级影像，每张影像尺寸为 20,000*20,000 像素，分割为了 256*256 像素的瓦片图像。共得到 6974 对影像。由人工标注，每对影像匹配一个高精度变化检测标签。如图 3 所示。



Figure 3. 2011 and 2018 image pairs and change detection labels

图 3. 2011 年和 2018 年影像对及其变化检测标签

4.2. 场景样本集准备

源域样本集为现有航空影像样本集，共包含 NWPU-RESISC45、AID 和 Jinmen 三个样本集。三个样本集均为现有经典的已有标注的航空数据集，每个样本集中与新余影像共有的类别被抽取出来，类别包含裸地、农田、森林、工业区、公路、河流、居民区、操场、其他，共计 9 个类别，每个数据集中用于训练的样本数量如表 1 所示。样本影像使用程序自动配对，每 3 张影像形成一个三元组。

目标域数据为两张同一地区不同时相的新余影像，分别称作 A0 和 B0。获取时间分别为 2018 和 2011 年。两张影像大小均为 20,000 × 20,000，空间分辨率为 0.2 米，分块数为 6975。对每张影像每类人工标注 20% 样本用于调整方法的参数，对剩下的 80% 的影像块进行分类。

Table 1. Class-Wise sample counts in the source domain dataset**表 1.** 源域样本集每类样本数量

类别	NWPU-RESISC45	AID	Jinmen
裸地		310	286
农田	700	370	90
公路	700		128
工业区	700	390	89
其他	700	390	23
操场	700	370	
居民区	700	290	143
河流	700	410	167
植被	700	250	180

4.3. 场景分类模型训练

训练过程如下：

- 1) 把已标记样本影像三张作为一组，组成锚点、正样本和负样本。其中正样本和锚点为一个类别，负样本和锚点属于另一种不同类别。
- 2) 对锚点、正样本和负样本分别训练一个神经网络，三个神经网络权重共享，利用困难三元组损失计算反向传播的梯度，更新神经网络的权重。
- 3) 利用训练好的源域特征提取器来初始化目标域特征提取器，并为目标域特征提取器提供源域中心。
- 4) 利用目标域特征提取器结合源域特征提取器对每张影像进行来自源域或者目标域分类，反向传播计算梯度并更新目标域特征提取器的权重。
- 5) 利用目标域特征提取器结合源域中心计算中心损失，反向传播计算梯度并更新目标域特征提取器的权重。
- 6) 利用源域特征提取出的特征训练分类器，用于区分目标域特征提取器得到的目标域特征。

4.4. 场景分类实验

基于中心损失的对抗学习框架对新余市 2011 年和 2018 年的影像分别进行场景分类。2018 年影像分类结果如图 4 所示，在新余 2018 年影像中工业区、居民区和河流等类别上取得了优于 95% 的分类精度，在除“其它类别”之外的类别上取得了优于 90% 的分类精度。其它类别包含了集装箱、停车场和篮球场等类别，其它类别中某些影像肉眼难以区分其类别。

2011 年影像分类结果如图 5 所示，在新余 2011 年影像中农田、裸地、森林和居民区等类别上取得了优于 95% 的分类精度，在除“其它”、操场和河流之外的类别上取得了优于 90% 的分类精度。其它类别包含了集装箱、停车场和篮球场等类别，其它类别中某些影像肉眼难以区分其类别，操场由于影像数量较少，仅错分一张。河流类别与裸地、农田和森林等类别在光谱上差异较小，容易与其它类别错分。

5. 变化检测结果和结论

由于天地图各层级的影像和矢量数据都是严格配准好的，在上述分类结果的基础上，可以直接比较两时相影像类别，得到变化图。如图 6 所示。

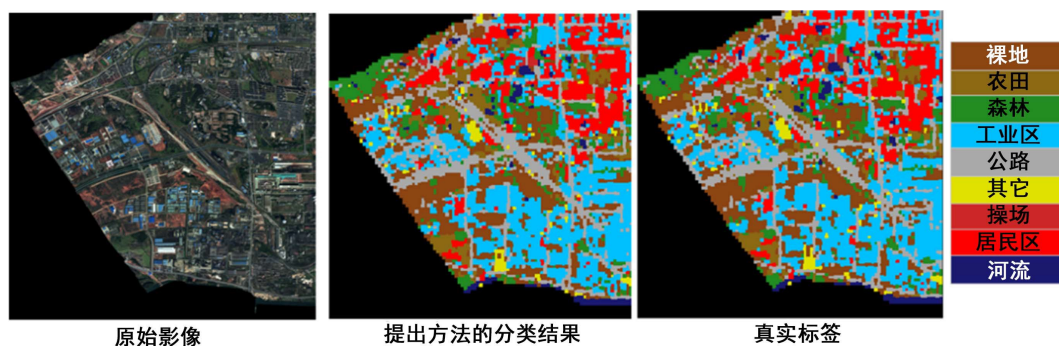


Figure 4. Image classification results of Xinyu city in 2018

图 4. 新余 2018 年影像分类结果

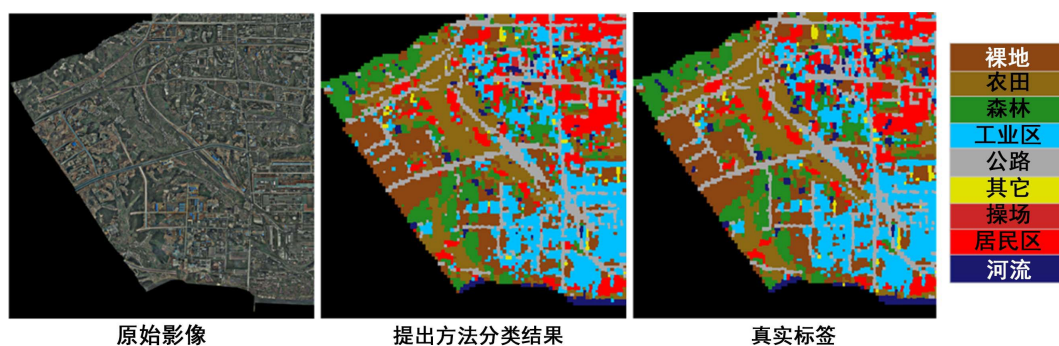


Figure 5. Image classification results of Xinyu city in 2011

图 5. 新余 2011 年影像分类结果

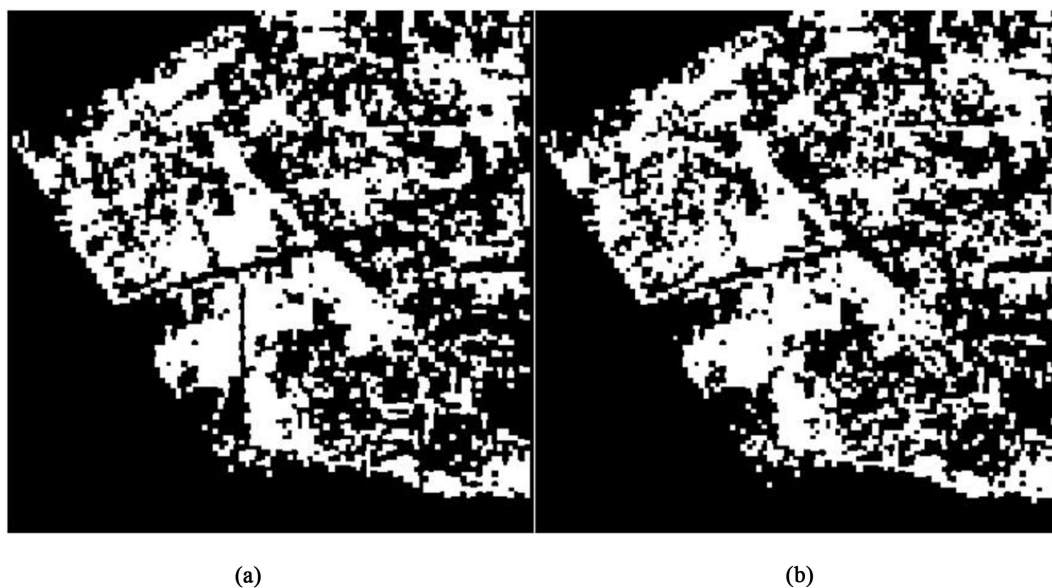


Figure 6. Change detection results of Xinyu experimental area (a) Ground truth (b) Predicted results

图 6. 新余试验区域影像变化检测结果(a)人工标注真值(b)实验预测结果

我们得到的变化检测结果目视精度较好，量化指标见表 2。准确率 93%，召回率 95%，精度 90%，kappa 系数 86%。因此，采用深度学习的分类后检测的两步法实现城市区域的变化检测有效且精准。

Table 2. Accuracy assessment of change detection for the Xinyu experimental area
表 2. 新余试验区域变化检测精度评价

准确率	召回率	精度	kappa 系数
0.9302	0.9509	0.9093	0.8604

尽管本研究提出的基于深度学习的城市区域变化检测方法取得了较高的精度，但是仍存在一些局限性，值得在未来进一步研究和完善。

本研究方法将半困难三元组网络损失与对抗学习框架下的中心损失结合，同时解决域适应中的分布对齐和特征判别性问题；仅需较少数量的目标域标注数据即可实现高精度分类，适用于实际应用中标注成本高的场景。

但研究仍存在一些局限性：由于包含三元组采样、对抗训练和中心损失计算，训练时间较长，因此本方法不适合实时或资源受限场景。同时，本研究方法主要依赖视觉特征相似性，无法有效处理组成相同但空间布局不同的类别(如工业区和住宅区等)，导致在这些类别上易混淆；本方法基于场景分类后的比较，对于集装箱、停车场和篮球场等小型建筑物等小尺度变化目标的检测能力有限。这是由于分类网络更侧重于提取图像的全局和上下文特征，而对精细的像素级局部特征的感知能力有所不足。

在未来的工作中，可以引入空间结构信息以更好区分空间布局相似的类别、探索更轻量化的网络结构以降低计算成本；以及进一步研究无监督域适应(完全无目标域标注)场景下的应用，可探索结合像素级语义分割或目标检测网络，设计多尺度变化检测框架，以同时保证大范围变化的效率和细小变化的精度。

基金项目

江西省科技计划(20171BBE50062)。

参考文献

[1] 万冉冉, 陈娟, 廖明伟, 刘异, 庞超. 基于 Siamese 卷积神经网络的影像瓦片变化检测技术[J]. 测绘通报, 2020(4): 96-100.

[2] 孟琮棠, 赵银娣, 向阳. 基于卷积神经网络的遥感图像变化检测[J]. 现代测绘, 2019, 42(5): 1-5.

[3] 张雯, 曾豆豆. 基于深度学习的城市地理国情地表覆盖分类研究[J]. 测绘与空间地理信息, 2021, 44(1): 112-115.

[4] 杨晓丽. 多时相遥感影像的变化检测[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2011.

[5] 吕浩博. 基于深度学习的长时间序列城市制图与变化检测研究[D]: [博士学位论文]. 北京: 清华大学, 2018.