

Research on Urban Building Identification and Classification Combining Aerial Survey and Remote Sensing and Deep Learning

Jian Xu Lin Zhao

Zhongshui North Survey and Design Research Co., Ltd., Tianjin, 300000, China

Abstract

This research focuses on the integration of aerial survey and remote sensing technology and deep learning methods to improve the ability of automatic identification and classification of urban buildings. Through the comprehensive utilization of high-resolution aerial images and lidar data, a set of effective data fusion strategy is designed, which provides rich multimodal information for the deep convolutional neural network, and further enhances the feature extraction and classification performance of urban buildings. This study provides a new perspective and a way to solve the urban building identification and classification problem combining aerial remote sensing and deep learning.

Keywords

aerial survey and remote sensing; deep learning; urban building identification; data fusion; multi-modal information

航测遥感与深度学习相结合的城市建筑物识别与分类研究

许健 赵琳

中水北方勘测设计研究有限责任公司, 中国·天津 300000

摘要

本研究聚焦于航测遥感技术与深度学习方法的融合, 以提升城市建筑物的自动识别与分类能力。通过综合利用高分辨率航拍影像和激光雷达数据, 设计了一套有效的数据融合策略, 为深度卷积神经网络提供了丰富的多模态信息, 进一步增强了城市建筑物的特征提取与分类性能。本研究为航测遥感与深度学习相结合的城市建筑物识别与分类问题提供了新的视角和解决途径。

关键词

航测遥感; 深度学习; 城市建筑物识别; 数据融合; 多模态信息

1 引言

随着城市化进程的不断加速, 城市建筑物的规模和数量呈现出爆发式增长, 城市信息的获取和管理日益成为当代社会的紧迫任务。城市建筑物的识别与分类对城市规划、资源管理、环境监测、灾害应对等领域至关重要。然而, 传统的城市建筑物识别方法在面临大规模、高密度城市区域的挑战时, 常常表现出局限性, 难以满足快速和准确的需求。航测遥感技术已经成为城市信息获取的重要手段之一, 其通过卫星、飞机或其他平台获取的高分辨率影像和激光雷达数据, 为城市建筑物的识别提供了丰富的信息源。然而, 仅仅依赖航测遥感数据进行建筑物识别在复杂城市环境下依然

具有挑战性, 因为影像可能受到阴影、遮挡、反射等因素的干扰。因此, 如何提高城市建筑物的自动识别与分类性能一直是遥感领域的研究重点之一。

2 航测遥感与深度学习融合

2.1 航测遥感技术概述

航测遥感技术是一种通过航空器、卫星或其他遥感平台获取地球表面信息的科学与技术领域。它的应用范围涵盖了地理信息系统(GIS)、城市规划、环境监测、资源管理、农业、林业、地质勘查等众多领域^[1]。主要数据源包括高分辨率航拍影像和激光雷达数据。

高分辨率航拍影像能提供详细的地表信息, 包括建筑物、道路、绿地等。这些影像通常具有高空间分辨率和丰富的视觉信息, 适合用于城市建筑物的定位和外观识别。然而, 它们容易受到天气条件、光照变化、遮挡和反射等因素的干扰。

【作者简介】许健(1984-), 女, 中国天津人, 本科, 高级工程师, 从事航测遥感技术、GIS开发和遥感AI模型等研究。

激光雷达数据则是通过发射激光脉冲来测量地表距离，从而创建高精度的地形和高程模型。激光雷达数据具有独立于光照的优势，适用于建筑物高度估计和地形分析。但它们缺乏颜色和纹理信息，对建筑物的识别有一定的局限性。

2.2 深度学习在遥感中的应用

深度学习在遥感领域的应用已经产生了深远的影响。深度卷积神经网络（CNN）等深度学习模型在遥感图像处理中表现出卓越的性能，主要得益于其强大的特征学习能力。CNN可以自动学习图像中的特征，这对于城市建筑物识别至关重要。传统的图像处理方法通常需要手工设计特征提取器，而深度学习模型可以从大规模数据中学习到最优的特征表示。这意味着无需复杂的特征工程，深度学习模型可以直接处理航测遥感数据^[2]。另一个深度学习的优势是其在大数据背景下的泛化能力。深度学习模型可以处理大规模数据集，从而更好地适应城市建筑物的多样性和复杂性。

3 数据融合与特征提取

3.1 多模态数据的获取与处理

多模态数据的获取和处理是城市建筑物识别与分类研究中的核心环节。多模态数据通常包括高分辨率航拍影像和激光雷达数据，它们在城市建筑物特征提取和分类中提供了丰富的信息。高分辨率航拍影像的获取是多模态数据的关键之一。这些影像通常由卫星、无人机或飞机获取，具有较高的地面分辨率，能够捕捉建筑物的外观和纹理信息。在数据获取过程中，需要考虑飞行高度、相机参数、飞行路线等因素，以确保数据质量和全覆盖性。传感器校正和地理配准也是必要的，以将不同数据源的信息融合在一起。激光雷达数据则提供了地形和高程信息。激光雷达通过发射激光脉冲并测量其返回时间来获取地表上的三维坐标。这为建筑物的高度估计和地面的数字高程模型（DEM）创建提供了有力支持。激光雷达数据在城市建筑物识别中常常用于消除阴影和遮挡效应，提高分类准确性。

数据获取之后，对多模态数据的预处理至关重要。这包括图像校正、去噪、镶嵌、地理坐标校正等步骤。图像校正用于解决图像畸变和变形问题，以确保影像的准确性。去噪则有助于降低影像中的噪声，提高建筑物边缘的清晰度。镶嵌是将多幅航拍影像拼接成连续的覆盖区域，以获得更广泛的覆盖范围。地理坐标校正则是将不同数据源的地理坐标对齐，以便后续的数据融合和分析。

3.2 数据融合策略

数据融合是多模态数据处理的关键步骤，旨在将高分辨率航拍影像和激光雷达数据有机地结合起来，以提供更全面、准确的信息以支持城市建筑物的识别与分类。以下是数据融合策略的关键方面：

①多尺度信息整合：高分辨率影像和激光雷达数据常常具有不同的空间分辨率。数据融合应考虑将这些数据按照

合适的比例进行缩放，以便它们在相同的尺度下进行比较和融合。这可以通过插值、下采样或上采样等方法实现。

②特征提取与融合：数据融合的关键目标之一是提取多模态数据的共享特征。这包括建筑物的外观特征（从影像中提取）和高程信息（从激光雷达数据中提取）。深度学习方法可以用于自动学习和提取这些特征，例如使用卷积神经网络（CNN）来对影像进行特征提取，同时使用特定的算法从激光雷达数据中提取高程特征。

③数据融合方法：数据融合可以采用多种方法，如级联、融合网络或多模态融合方法。级联方法是将两个或多个模态的特征层叠在一起，以创建一个混合特征。融合网络是专门设计的神经网络，用于有效地融合多模态数据。多模态融合方法通常采用加权融合，将不同模态的特征赋予不同的权重，以最大程度地减小信息丢失。

3.3 深度卷积神经网络设计

深度卷积神经网络（CNN）设计在城市建筑物的识别与分类研究中扮演着至关重要的角色。通过精心设计深度CNN模型，可以实现对城市建筑物的高效特征提取和精确分类。以下是深度CNN网络设计的详细考虑因素：

网络架构的选择至关重要。不同网络架构，如ResNet、VGG、Inception等，都拥有各自的优势和适用性。选择适当的网络架构需要考虑建筑物识别任务的复杂性和数据规模。例如，在处理大规模城市区域的数据时，采用深层网络（如ResNet）可能更合适，因为它们能够捕获更丰富的特征表示。多尺度处理是提高城市建筑物识别性能的关键因素。城市建筑物的尺寸和形状差异巨大，因此网络应具备处理不同尺度信息的能力。这可以通过引入多尺度卷积核或多尺度池化层来实现，以确保网络能够捕捉不同规模建筑物的特征。数据增强是提高模型鲁棒性的重要手段^[3]。在城市建筑物识别任务中，建筑物可能出现在不同的角度、光照条件和遮挡情况下。因此，数据增强技术如随机裁剪、翻转、旋转、缩放和亮度调整等，有助于模型更好地适应不同场景，提高其泛化性能。

4 实验与结果分析

4.1 实验设计

本研究的实验旨在评估所提出的深度学习模型在城市建筑物识别与分类任务中的性能。使用了多模态数据，包括高分辨率航拍影像和激光雷达数据。实验分为以下几个步骤：

数据准备：采集了包括不同城市区域的多模态数据，包括航拍影像和激光雷达数据。这些数据包括不同类型的建筑物，如住宅、商业和工业建筑。具体如表1所示。

数据预处理：对数据进行了校正、配准、镶嵌、去噪和地理坐标校正等预处理步骤，以确保数据的质量和一致性。

表 1 数据准备

| 区域 | 数据类型 | 建筑物类型 | 数据来源 | 数据格式 | 数据日期 |
|-----|------|-------|------------------|---------|-----------|
| 北京市 | 卫星影像 | 住宅 | Gaofen-2 | GeoTIFF | 2023/1/15 |
| 北京市 | 航拍数据 | 住宅 | DJI Phantom 4 | JPEG | 2023/2/10 |
| 北京市 | 雷达数据 | 住宅 | TerraSAR-X | HDF5 | 2023/3/20 |
| 上海市 | 卫星影像 | 商业 | ZY-3 | GeoTIFF | 2023/1/20 |
| 上海市 | 航拍数据 | 商业 | Yuneec Typhoon H | JPEG | 2023/2/15 |
| 上海市 | 雷达数据 | 商业 | Sentinel-1 | NetCDF | 2023/3/25 |
| 广州市 | 卫星影像 | 工业 | Gaofen-1 | GeoTIFF | 2023/1/25 |
| 广州市 | 航拍数据 | 工业 | Parrot Anafi USA | JPEG | 2023/2/20 |
| 广州市 | 雷达数据 | 工业 | RADARSAT-2 | HDF5 | 2023/4/5 |

数据分割：数据集分为训练集、验证集和测试集，以进行模型训练、调优和性能评估。

模型训练：采用了所提出的深度卷积神经网络模型，使用训练集数据进行模型训练。还使用了迁移学习技术，利用在大规模图像数据上训练的预训练模型进行初始化。

性能评估：使用测试集数据对模型的性能进行评估，包括准确性、召回率、F1 分数等指标。

4.2 实验结果

表 2 列出了模型在测试集上的性能结果，包括准确性、召回率、F1 分数。混淆矩阵是用于评估分类模型性能的表格，它显示实际类别与模型预测类别的交叉情况，用于计算各种性能指标。

表 2 模型在测试集上的性能结果

| 指标 | 住宅区 | 商业区 | 工业区 | 平均 |
|-------|------|------|------|------|
| 准确性 | 0.92 | 0.88 | 0.94 | 0.91 |
| 召回率 | 0.94 | 0.87 | 0.95 | 0.92 |
| F1 分数 | 0.93 | 0.88 | 0.94 | 0.91 |

4.3 结果分析

从上述数据可以看出，所提出的深度学习模型在城市建筑物识别与分类任务中表现出色。模型在测试集上实现了

高准确性（91%）、高召回率（92%）和高 F1 分数（91%）。这表明模型在不同建筑物类别的识别中都具有出色的性能。

5 结论

在本研究中，成功结合了航测遥感与深度学习，致力于解决城市建筑物识别与分类的重要问题。通过充分利用多模态数据、深度卷积神经网络设计、特征提取和建筑物分类方法在大规模城市区域的数据集上实现了卓越的性能。这些研究成果不仅为城市规划、资源管理和环境监测等领域提供了有力支持，还为航测遥感与深度学习的结合提供了实际案例和方法。未来，继续探索和改进这一领域的技术，可以不断推动城市建筑物识别与分类的研究和应用，以更好地应对城市化发展的挑战。

参考文献

- [1] 宋廷强.基于深度学习的高分辨率遥感图像建筑物识别[J].计算机工程与应用,2020(8):26-34.
- [2] 史文旭,鲍佳慧,姚宇.基于深度学习的遥感图像目标检测与识别[J].计算机应用,2020,40(12):5.
- [3] 赵伍迪,李山山,李安,等.结合深度学习的高光谱与多源遥感数据融合分类[J].遥感学报,2021,25(7):14.