

# 基于卷积神经网络的空心村高分影像建筑物检测方法

李 政<sup>1</sup> 李永树<sup>1</sup> 吴 玺<sup>2</sup> 刘 刚<sup>3</sup> 鲁 恒<sup>4</sup> 唐 敏<sup>5</sup>

(1. 西南交通大学地球科学与环境工程学院, 成都 611756; 2. 四川省土地统征整理事物中心, 成都 610041;  
3. 成都理工大学地球科学学院, 成都 610059; 4. 四川大学水利水电学院, 成都 610065;  
5. 中铁二院工程集团有限公司, 成都 610031)

**摘要:** 基于卷积神经网络(CNN)提出了一种适用于空心村高分影像的建筑物自动检测方法,该方法利用多尺度显著性检测来获取包含建筑物信息的显著性区域,然后通过滑动窗口获取显著性区域内目标样本块,再将这些样本块输入训练好的CNN并结合SVM来实现分类。为检验方法有效性,选取高分影像进行实验,结果表明,显著性检测能够有效地获取主要目标,减弱其他无关目标的影响,降低数据冗余;卷积神经网络能够自动学习高层次的特征,基于CNN对高分影像进行建筑物检测,分类准确度可以达到97.6%,表明该方法具有较好的鲁棒性和有效性。

**关键词:** 空心村; 建筑物检测; 卷积神经网络; 高分影像; 多尺度显著性检测

**中图分类号:** P231.5; TP75      **文献标识码:** A      **文章编号:** 1000-1298(2017)09-0160-06

## Hollow Village Building Detection Method Using High Resolution Remote Sensing Image Based on CNN

LI Zheng<sup>1</sup> LI Yongshu<sup>1</sup> WU Xi<sup>2</sup> LIU Gang<sup>3</sup> LU Heng<sup>4</sup> TANG Min<sup>5</sup>

(1. Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China  
2. Center of Land Acquisition and Consolidation in Sichuan Province, Chengdu 610041, China  
3. College of Earth Sciences, Chengdu University of Technology, Chengdu 610059, China  
4. College of Hydraulic and Hydroelectric Engineering, Sichuan University, Chengdu 610065, China  
5. China Railway Eryuan Engineering Group Co., Ltd., Chengdu 610031, China)

**Abstract:** Accurately obtaining the building information in the hollow village areas is important for hollow village renovation and research. With the rapid development of remote sensing technology, remote sensing image resolution has been greatly improved and the ground targets can be obtained from high-resolution remote sensing image. But the traditional methods based on low-level hand-engineered features or mid-level features have great limitation in complex environment, especially in hollow village areas. So it needs to use high-level features to express. Convolution neural network (CNN) has become one of the important methods of ground object recognition and detection. Based on CNN, a novel automatic building detection method was proposed. Firstly, a multi-scale saliency computation was employed to extract building areas and a sliding windows approach was applied to generate candidate regions. And then a CNN was applied to classify the regions. In order to verify the validity of this method, the high resolution remote sensing image of typical hollow village was selected to construct the building sample library. Finally, the model for building interpretation was experimentally studied based on the sample library. The results showed that multi-scale saliency can effectively get the main target, weaken the impact of other unrelated targets, and reduce data redundancy. The CNN can automatically learn the high level feature, and the classification accuracy (ACC) of this method can reach 97.6%. So the proposed method can be used to detect building and it had high practical value to hollow village research and renovation.

**Key words:** hollow village; building detection; convolution neural network; high resolution remote sensing image; multi-scale saliency test

引言

随着我国工业化、城镇化的快速发展,农村人口非农业化转移速度不断加快,农村常住人口逐渐减少,农村人地关系发生了巨大变化,其中最突出的问题是出现了“人走房空”的“空心房”现象,并逐步由个体出现、扩张,发展为村域成片废弃和整村闲置,产生了农村聚落大规模的“空心化”景观,并由人口空心化逐渐转变为农村人口、土地、产业和基础设施空心化的农村地域空心化,不断发展的结果就是产生了大量的空心村<sup>[1]</sup>。由于缺乏规划引导导致新建房屋分布凌乱、乡村建设用地脱离原有核心结构、无序拓展逐渐空心化,空心村的存在造成土地资源浪费、低效利用以及村落人居环境品质退化等尖锐矛盾<sup>[2]</sup>。因此,面对广袤的农村地区,高效、准确地获取建筑物信息对空心村的研究具有重要意义。

遥感已经广泛应用于农村土地利用调查、测绘以及地理国(省)情监测等领域。特别是无人机高分遥感具有较高的空间分辨率,可以获取高精度的数字高程模型以及高分辨率正射影像,给空心村内信息的智能识别提供了可能性,其应用于空心村建筑物快速检测具有优势<sup>[3-4]</sup>。传统农村建筑物信息获取主要依靠外业大比例尺测图和人工目视解译两种方式。外业大比例尺测图易受到天气等环境因素制约,费时费力;采取人工目视解译方法在遥感影像上获取高精度房屋信息,效率低工作量大<sup>[5-6]</sup>。如何从高分影像中自动检测建筑物是遥感应用的热点问题,研究有效的高分影像建筑物提取方法具有较高的实用价值。早期的建筑物自动提取研究应用的是低层次的特征,比如边缘、线段分割和角点检测。随着高分影像的分辨率越来越高,尽管现有分类方法众多,然而如何在高分影像上进行建筑物自动检测仍然是个难题。

非监督特征学习或者是手动设计的低层次特征在复杂场景中具有一定的局限性,特别是对于地形多变,建筑物被树木遮挡,建筑物本身光谱信息复杂的空心村区域,基于像素或者是面向对象的方法在空心村区域很难获得良好的效果。随着深度学习技术的发展,深层卷积神经网络在图像分类、目标检测、图像语义分割等领域取得了一系列突破性研究成果<sup>[7-10]</sup>。HINTON 等<sup>[11]</sup>在 2006 年提出深度学习,该模型具有更强的表达能力,使卷积神经网络能够更高层次的表达检测目标的特征,进而能够在高分影像上更准确高效地获取建筑物信息。针对传统低层次特征的不足,本文结合卷积神经网络强大的特征学习与分类能力提出一种基于 CNN 的空心村

高分影像建筑物自动检测方法。

1 理论方法

本文基于多尺度显著性检测算法<sup>[12-13]</sup>,获取待检测高分影像中的建筑物显著性区域,降低了数据冗余并且减弱了其他无关目标的影响,然后利用滑动窗口对获取的显著区域进行采样,采集待分类的目标样本块,最后利用训练好的 CNN 网络结合 SVM 进行分类,从而检测得到研究区域内的建筑物。整体流程如图 1 所示。

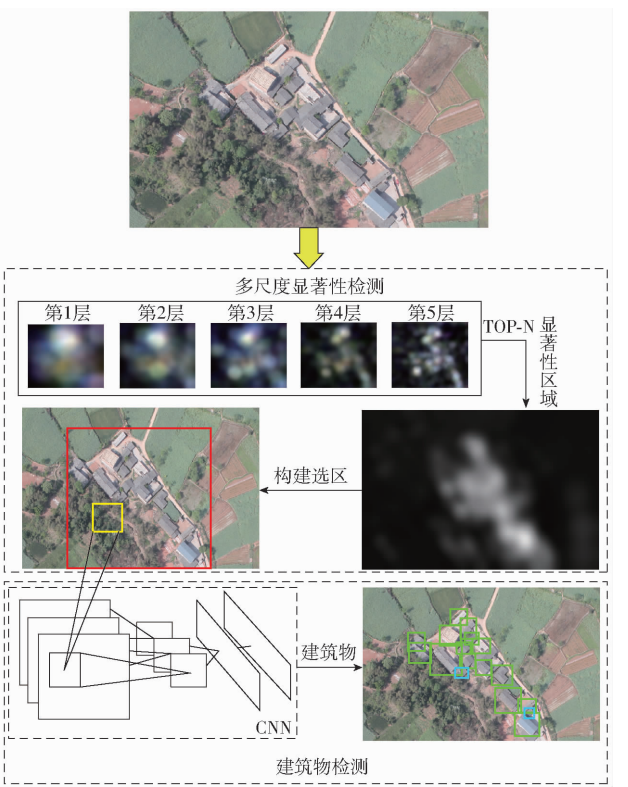


图 1 整体流程图

Fig. 1 Flow chart of proposed method

1.1 多尺度显著性检测

影像分类用于获取场景中有意义的目标,然而由于高分影像的复杂性很难获得理想的分类结果。显著性检测是在大量信息中选择感兴趣区域的过程,这些感兴趣的区域通常包含了影像中的绝大部分信息,因此能够很好地表达影像。本文结合文献[12-13]所提出的多尺度显著性(MS)检测方法来采样。首先对影像进行显著性检测,构建建筑物选区,然后利用滑动窗口在选区内采集样本块。这里定义  $I_{MS}^s(p)$  是利用显著性检测方法获取每个像素  $p$  的显著性。并定义了  $s$  尺度下窗口  $\omega$  的显著性,表达式为

$$MS(\omega, \theta_{MS}^s) = \sum_{\{p \in \omega | I_{MS}^s(p) \geq \theta_{MS}^s\}} I_{MS}^s(p) \frac{|\{p \in \omega | I_{MS}^s(p) \geq \theta_{MS}^s\}|}{|\omega|} \tag{1}$$

式中,  $\theta_{MS}^s$  是特定尺度阈值, 通过在每个尺度  $s$  上对训练样本  $O$  的定位精度进行优化获得。对于每个训练影像  $I$  和尺度  $s$ , 计算显著性图  $I_{MS}^s(p)$  和每个窗口的显著性得分。在 4D 分数空间<sup>[14]</sup> 进行非最大抑制运算获得一组局部最大值窗口  $W_s^{\max}$ , 通过最大化获取最优  $\theta_{MS}^s$ , 表达式为

$$\theta_{MS}^s = \operatorname{argmax} \sum_{o \in O} \max_{\omega \in W_s^{\max}} \frac{|\omega \cap o|}{|\omega \cup o|} \quad (2)$$

式中  $o$ ——训练集  $O$  中的个体

本文首先对输入高分影像进行下采样, 获得 5 层影像金字塔, 然后对获取的每层影像金字塔计算显著图。接着使用窗口  $\omega$  对每层的显著图进行采样, 利用式(1)计算每个窗口的显著性值, 最终选取显著性值最大的 5 个显著性区域来构建建筑物的选区, 有效地降低了数据冗余并且减弱了其他无关目标影响, 具体流程如图 1 所示。接着利用滑动窗口在构建的建筑物选区内采集用于分类的样本块。

## 1.2 基于 CNN 的分类

卷积神经网络(CNN)主要用来识别位移、缩放及其他形式扭曲不变性的二维图形<sup>[15]</sup>。由于 CNN 的特征检测层通过训练数据进行学习, 所以在使用 CNN 时, 避免了显式的特征提取, 而隐式地从训练数据中进行学习; 由于同一特征映射面上的神经元权值相同, 所以网络可以并行学习, 这也是卷积网络相对于神经元彼此相连网络的一大优势<sup>[16-17]</sup>。

特征提取是对影像信息进行抽象, 得到一组可以描述影像的特征向量, 是影像分类的关键, 特征提取决定了最终的分类结果。对于新的数据和任务, 一般人工目视解译可以获得较好的分类结果, 但工作量大, 效率低, 具有主观性强和非定量等缺陷<sup>[18-20]</sup>。中、低层次的特征, 比如 HOVW、SIFT、HOG 等, 在特定的分类和识别任务中具有良好效果。然而, 空心村内的建筑物数据为高分影像, 与普通的自然图像差异极大, 空间光谱变化较大, 因此中低层次的特征提取效果不佳。随着深度学习的深入, 通过输入数据逐级提取从底层到高层的特征, 可以很好地建立从底层信号到高层语义的映射关系, 从而获得建筑物的高层次特征, 能够更好地表达空心村内建筑物<sup>[21-24]</sup>。

卷积神经网络避免了显式的特征取样, 隐式地从训练数据中进行学习, 这使得卷积神经网络明显有别于其他基于神经网络的分类器, 通过结构重组和减少权值将特征提取功能融合进多层感知器, 它可以直接处理灰度图片, 能够直接用于处理基于图像的分类<sup>[25]</sup>。其在图像处理方面有如下优点: 输入图像和网络的拓扑结构能很好吻合; 特征提取和模

式分类同时进行, 并同时在训练中产生; 权重共享可以减少网络的训练参数, 使神经网络结构变得更简单, 适应性更强<sup>[26-28]</sup>。一个基本的卷积神经网络结构可以分为 3 层: 特征提取层、特征映射层以及特征池化层, 通过堆积多个基本网络结构可以形成一个深度卷积网络。

(1) 特征提取层: 特征提取层的作用是, 每个神经元的输入与前一层的局部感受野相连, 并提取该局部的特征。假设输入影像  $I$  是一个二维矩阵, 大小为  $\gamma \times c$ , 利用一个可训练的过滤器组  $K$ , 尺寸为  $w \times w$ , 对其进行卷积运算,  $l$  为步长, 最终得到一个大小为  $[(\gamma - w)/l + 1] \times [(c - w)/l + 1]$  的输出  $Y$ 。

$$y_i = b_i + \sum_j k_{ij} * x_i \quad (3)$$

式中  $x_i$ ——卷积层的输入

$k_{ij}$ ——卷积核参数

$b_i$ ——偏差值

\* 表示卷积运算, 每个滤波器对应一种特定的特征。

(2) 特征映射层: 利用一个非线性函数对过滤层的结果进行映射来保证特征有效性, 得到特征图  $F$ 。

$$f_i = \delta \left( b_i + \sum_j k_{ij} x_i \right) \quad (4)$$

式中,  $\delta$  是激活函数, 常用的激活函数有 tanh、sigmoid 和 softplus。tanh 函数是 sigmoid 函数的一种变体, 其取值范围是  $[0, 1]$ 。线性修正单元 ReLU<sup>[29]</sup> 与生物神经元受刺激后的激活模型最为接近, 而且具有一定的稀疏性, 计算简单, 有助于提升效果。

(3) 特征池化层: 理论上可以通过卷积获得特征以后直接利用这些特征训练分类器, 然而一个中等大小的图像卷积后的特征维度都在百万数量级, 直接训练分类器很容易过拟合 (overfitting), 因此需要对卷积特征进行池化操作, 即下采样。 $F$  为卷积特征图, 将其分成为大小  $m \times m$  的不相交区域, 然后统计这些区域的平均值 (或最大值) 作为池化特征  $P$ , 大小为  $[(\gamma - w)/l + 1]/m \times [(c - w)/l + 1]/m$ 。池化后的特征维度大大降低, 避免了过拟合, 而且具备抗差性。

本文利用 Caffe 框架<sup>[30]</sup> 对 KRIZHEVSKY<sup>[31]</sup> 提出的 AlexNet 进行了训练, 由于本文研究的是二分类问题, 因此将 AlexNet 的 FC8 层用 2 分类输出取代了多分类输出, 图 2 为本文构建的 R-CNN<sup>[32]</sup> 分类流程。

将窗口采集的样本块统一为 256 像素  $\times$  256 像素, 然后送入训练好的 CNN 网络, 结合 SVM 分类器进行分类, 并对结果进行标记, 实现建筑物的快速

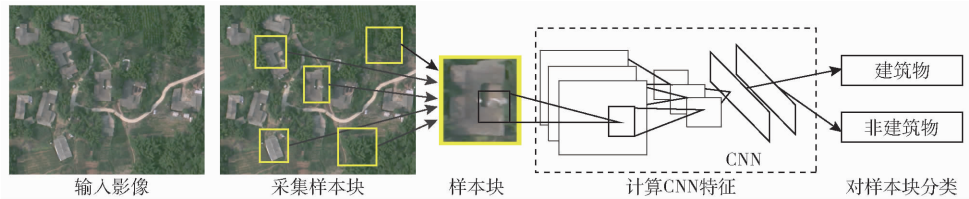


图 2 改进的 R - CNN 分类流程

Fig. 2 Improved R - CNN classification process

检测。

2 实验与分析

2.1 实验数据

研究区位于四川省成都市,包含多个典型空心化村庄。实验采用无人机获取高分影像,空间分辨率为 0.2 m,影像覆盖面积总计 46.48 km<sup>2</sup>。训练样本对分类准确率具有重要影响,而正样本选取是关键。本文利用目视解译的方法,为不同建筑物类型手工标注了大量正样本,并通过旋转、镜像等方式扩充了样本数据,基本包含了空心村区域内各种建筑物类型。随着不同类型、结构建筑物正样本数量的

增加,卷积神经网络能更深层次地表达建筑物特征,进而能够在影像上更准确高效地获取建筑物信息,识别准确率也会相应地提高。对于负样本的选择,本文利用 ArcGIS 软件通过随机选取的方式在空心村区域内获取大量背景负样本,然后通过人工筛选方式选取除建筑物以外的多种对象,包含的非建筑物样本种类越多,识别准确率也会得到提高。样本包括 RGB 3 个波段,并按照标准规范格式进行存储,形成高分影像建筑物样本库。构建的样本库包括建筑物正样本 5 000 个,建筑物负样本 10 000 个,所有样本均缩放到 256 像素 × 256 像素。样本示例如图 3 所示。

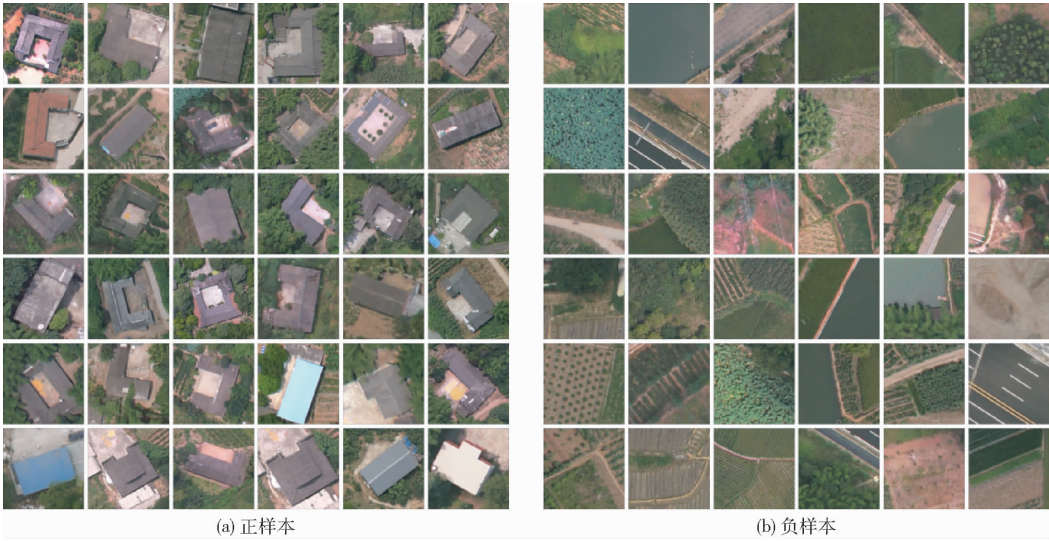


图 3 建筑物样本库示例

Fig. 3 Samples of building sample library

2.2 实验结果与分析

为了验证本文方法的有效性,选取了 HOG 特征及 BOVW 特征在构建的样本库中进行了交叉实验。本文使用混淆矩阵<sup>[33]</sup>来验证构建解译模型的精度,评价指标体系如表 1 所示。

表 1 混淆矩阵相关指标

Tab.1 Related indicators of confusion matrix

| 实际 | 预测          |             | 合计                      |
|----|-------------|-------------|-------------------------|
|    | 1           | 0           |                         |
| 1  | $T_p$       | $F_N$       | $T_p + F_N$             |
| 0  | $F_p$       | $T_N$       | $F_p + T_N$             |
| 合计 | $T_p + F_p$ | $F_N + T_N$ | $T_p + F_p + F_N + T_N$ |

其中  $T_p$  表示将正类预测为正类数;  $T_N$  表示将负类预测为负类数;  $F_p$  表示将负类预测为正类数;  $F_N$  表示将正类预测为负类数。同时定义精确率  $P_{re}$ 、准确率  $A_{cc}$  和召回率  $R_{ec}$  为

$$P_{re} = \frac{T_p}{T_p + F_p} \times 100\%$$
 (5)

$$A_{cc} = \frac{T_p + T_N}{T_p + T_N + F_p + F_N} \times 100\%$$
 (6)

$$R_{ec} = \frac{T_p}{T_p + T_N} \times 100\%$$
 (7)

实验采用支持向量机 (SVM) 作为分类器,采用线性核函数,利用 L2 正则化,取平方根误差 L2 作为

损失函数。训练过程中留 10% 的样本(1 500 个)用于测试。训练结果如下:HOG、BOVW 特征和 CNN 特征交叉验证最佳参数分别为  $\{ 'C': 8\ 000 \}$ 、 $\{ 'C': 6\ 000 \}$  和  $\{ 'C': 3\ 000 \}$ , 关键参数 C 采用 6-fold(6 文件夹)交叉验证获得。它们的混淆矩阵如表 2 所示。

表 2 混淆矩阵  
Tab. 2 Confusion matrix

| 参数          | HOG  |      | BOVW |      | 本文方法 |      |
|-------------|------|------|------|------|------|------|
|             | 建筑物  | 非建筑物 | 建筑物  | 非建筑物 | 建筑物  | 非建筑物 |
| 建筑物         | 376  | 124  | 431  | 69   | 489  | 11   |
| 非建筑物        | 189  | 811  | 117  | 883  | 25   | 975  |
| $P_{re}/\%$ | 66.5 |      | 78.6 |      | 95.1 |      |
| $R_{ec}/\%$ | 75.2 |      | 86.2 |      | 97.8 |      |
| $A_{cc}/\%$ | 79.1 |      | 87.6 |      | 97.6 |      |

从表 2 中可以看出,低层次的 HOG 特征分类准确率为 79.1%,中层次的词袋模型分类准确率可以达到 87.6%,而利用卷积神经网络的分类准确率可以达到 97.6%。说明了卷积特征对于目标具有很好的识别能力。

本文利用卷积神经网络构建了建筑物检测方法,并在未进行样本采集的区域内进行检测,实验在空心村区域内选取了 4 个代表性的区域:地形平坦、地势起伏较大、遮挡较严重以及建筑物密集区域,并另外选取了 2 处非空心村区域高分影像作为对比实验,如图 4 所示。

实验基于 Caffe 框架,使用的是 Intel i7 6700K CPU, 16 G 内存, GeForce GTX 1080 GUPs 的工作

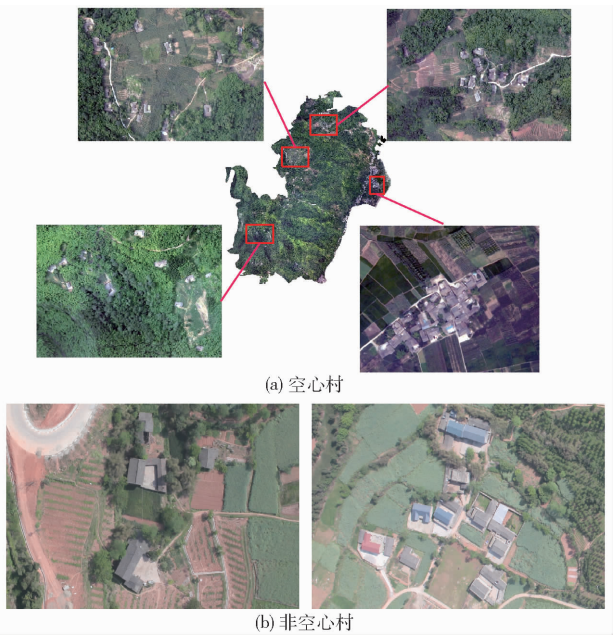


图 4 实验区域  
Fig. 4 Experimental areas

站。各个区域的检测结果如图 5 所示。

利用本文构建的建筑物检测方法对影像进行了分区域检测实验。如图 5a、5b 与 5e 所示,在地形比较平坦、建筑物无遮挡的区域可以获得更好的检测结果。但是对于图 5c,由于地势起伏较大,建筑物遮挡严重,导致检测效果不佳。如图 5d 与 5f 所示,对于建筑物密集区域,检测结果会出现重叠,不能完全地检测出整个建筑物。本文对整个研究区域进行了分块检测,由统计检测结果可知,本文所述方法的整体准确率可以达到 81%,表明该方法具有较好的识别准确率。

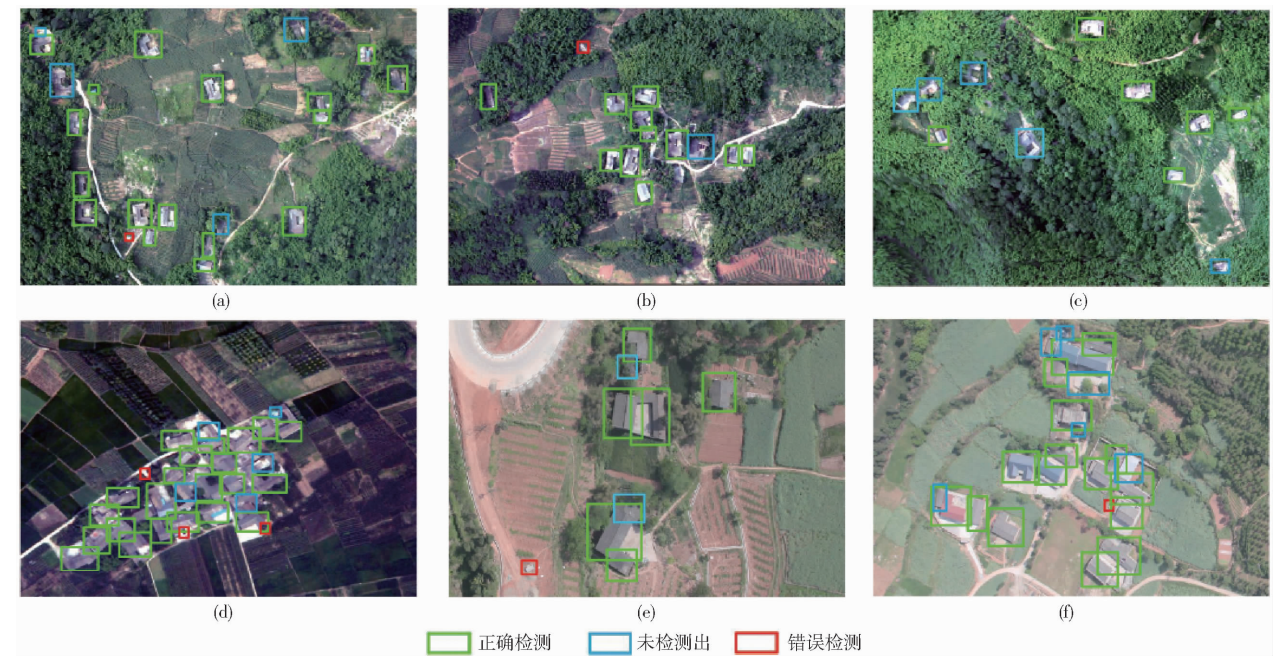


图 5 不同区域检测结果  
Fig. 5 Detection results of different areas

### 3 结束语

基于无人机高分影像构建了建筑物样本库,以空心村为研究对象,利用显著性检验和卷积神经网络提出了一种适用于空心村高分影像的建筑物检测方法。该方法可以准确高效地检测出空心村内建筑物,对于空心村调查研究具有较高的实用价值。

### 参 考 文 献

- 1 陈玉福,孙虎,刘彦随. 中国典型农区空心村综合整治模式[J]. 地理学报,2010,65(6):727-735.  
CHEN Yufu, SUN Hu, LIU Yansui. Reconstruction models of hollowed villages in key agricultural regions of China[J]. Acta Geographica Sinica, 2010, 65(6):727-735. (in Chinese)
- 2 SUN H, LIU Y, XU K. Hollow villages and rural restructuring in major rural regions of China: a case study of Yucheng City, Shandong Province[J]. Chinese Geographical Science, 2011, 21(3): 354-363.
- 3 MAGGIORI E, TARABALKA Y, CHARPIAT G, et al. Convolutional neural networks for large-scale remote-sensing image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(2): 645-657.
- 4 ZHANG L, LI A, ZHANG Z, et al. Global and local saliency analysis for the extraction of residential areas in high-spatial-resolution remote sensing image[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(7): 3750-3763.
- 5 史舟,梁宗正,杨媛媛,等. 农业遥感研究现状与展望[J/OL]. 农业机械学报,2015,46(2):247-260. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20150237&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20150237&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2015.02.037.
- 6 SHI Zhou, LIANG Zongzheng, YANG Yuanyuan, et al. Status and prospect of agricultural remote sensing[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(2):247-260. (in Chinese)
- 7 LE T H A, LE H M, DINH T P. Feature selection in machine learning: an exact penalty approach using a difference of convex function algorithm[J]. Machine Learning, 2015, 101(1-3): 163-186.
- 8 LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- 9 SHIN H C, ROTH H R, GAO M, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning[J]. IEEE Transactions on Medical Imaging, 2016, 35(5): 1285-1298.
- 10 何小飞,邹峥嵘,陶超,等. 联合显著性和多层卷积神经网络的高分影像场景分类[J]. 测绘学报, 2016, 45(9):1073-1080.  
HE Xiaofei, ZOU Zhengrong, TAO Chao, et al. Combine dsaliency with multi-convolutional neural network for high resolution remote sensing scene classification[J]. Acta Geodaetica et Cartographica Sinica, 2016,45(9):1073-1080. (in Chinese)
- 11 TAYLOR M E, STONE P. Transfer learning for reinforcement learning domains: a survey[J]. Journal of Machine Learning Research, 2009, 10(7): 1633-1685.
- 12 HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313(5786): 504-507.
- 13 ALEXE B, DESELAERS T, FERRARI V. What is an object? [C]//2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010: 73-80.
- 14 HOU X, ZHANG L. Saliency detection: a spectral residual approach [C]//2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2007: 1-8.
- 15 ZHANG Q, WANG Y, LIU Q, et al. CNN based suburban building detection using monocular high resolution Google Earth images[C]//2016 International IEEE Geoscience and Remote Sensing Symposium (IGARSS), 2016: 661-664.
- 16 LAWRENCE S, GILES C L, TSOI A C, et al. Face recognition: a convolutional neural-network approach[J]. IEEE Transactions on Neural Networks, 1997, 8(1): 98-113.
- 17 CIRESAN D C, MEIER U, MASCI J, et al. Flexible, high performance convolutional neural networks for image classification[C]//Proceedings of the 22nd International Joint Conference on Artificial Intelligence(IJCAI), 2011: 1-41.
- 18 鲁恒,付萧,贺一楠,等. 基于迁移学习的无人机影像耕地信息提取方法[J/OL]. 农业机械学报,2015,46(12):274-279,284. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20151237&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20151237&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2015.02.037.
- 19 LU Heng,FU Xiao, HE Yi'nan, et al. Cultivated land information extraction from high resolution UAV images based on transfer learning[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015,46(12):274-279,284. (in Chinese)
- 20 LI Y, TAO C, TAN Y, et al. Unsupervised multilayer feature learning for satellite image scene classification [J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(2): 157-161.
- 21 DONAHUE J, JIA Y, VINYALS O, et al. DeCAF: a deep convolutional activation feature for generic visual recognition[C]//Proceedings of the 31st International Conference on Machine Learning, 2014: 647-655.
- 22 GHAFARIAN S. Automatic building detection based on purposive FastICA (PFICA) algorithm using monocular high resolution Google Earth images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 97: 152-159.

Transactions of the CSAE, 2013, 29(8): 36–43. (in Chinese)

- 18 赵武云, 郭康权. 组合式螺旋板齿种子玉米脱粒机工作参数优化[J/OL]. 农业机械学报, 2012, 43(12): 56–61. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20121211&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20121211&journal_id=jcsam). DOI: 10.6041/j.issn.1000-1298.2012.12.011.  
ZHAO Wuyun, GUO Kangquan. Parameters optimization of combined spiral bar tooth thresher for maize seed[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(12): 56–61. (in Chinese)
- 19 宋学娟. 双滚筒大豆种子脱粒机设计研究[D]. 沈阳: 沈阳农业大学, 2013.  
SONG Xuejuan. Design and study on thresher of soybean seed with double-drum[D]. Shenyang: Shenyang Agricultural University, 2013. (in Chinese)
- 20 焦维鹏. 典型大豆力学特性研究与组合式滚筒脱粒装置设计[D]. 沈阳: 沈阳农业大学, 2012.  
JIAO Weipeng. Study on mechanical characteristics of typical soybean and design of thresher with combined roller[D]. Shenyang: Shenyang Agricultural University, 2012. (in Chinese)
- 21 李云雁, 胡传荣. 试验设计与数据处理[M]. 北京: 化学工业出版社, 2008.
- 22 任露泉. 试验设计及其优化[M]. 北京: 科学出版社, 2009.
- 23 SINGH K P, PARDESHI I L, KUMAR M, et al. Optimization of machine parameters of a pedal-operated paddy thresher using RSM[J]. Biosystems Engineering, 2008, 100(4): 591–600.
- 24 康建明, 陈学庚, 温浩军, 等. 基于响应面法的梳齿式采棉机采收台优化设计[J/OL]. 农业机械学报, 2013, 44(增刊2): 57–61. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=2013s212&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=2013s212&journal_id=jcsam). DOI: 10.6041/j.issn.1000-1298.2013.S2.012.  
KANG Jianming, CHEN Xuegeng, WEN Haojun, et al. Optimization of comb-type cotton picker device based on response surface methodology[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(Sup. 2): 57–61. (in Chinese)
- 25 戴飞, 高爱民, 孙伟, 等. 纵轴流锥型滚筒脱粒装置设计与试验[J]. 农业机械学报, 2011, 42(1): 74–78.  
DAI Fei, GAO Aimin, SUN Wei, et al. Design and experiment on longitudinal axial conical cylinder threshing unit[J]. Transactions of the Chinese Society for Agricultural Machinery, 2011, 42(1): 74–78. (in Chinese)
- 26 吴崇友, 丁为民, 石磊, 等. 油菜分段收获捡拾脱粒机捡拾损失响应面分析[J]. 农业机械学报, 2011, 42(8): 89–93.  
WU Chongyou, DING Weimin, SHI Lei, et al. Response surface analysis of pickup losses in two-stage harvesting for rapeseed[J]. Transactions of the Chinese Society for Agricultural Machinery, 2011, 42(8): 89–93. (in Chinese)

(上接第 165 页)

- 21 HUERTAS A, NEVATIA R. Detecting buildings in aerial images[J]. Computer Vision, Graphics, and Image Processing, 1988, 41(2): 131–152.
- 22 GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(1): 142–158.
- 23 CIREGAN D, MEIER U, SCHMIDHUBER J. Multi-column deep neural networks for image classification[C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012: 3642–3649.
- 24 OQUAB M, BOTTOU L, LAPTEV I, et al. Learning and transferring mid-level image representations using convolutional neural networks[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1717–1724.
- 25 YANG J, YANG F, WANG G, et al. Multi-channel and multi-scale mid-level image representation for scene classification[J]. Journal of Electronic Imaging, 2017, 26(2): 023018.
- 26 WEI Y, ZHAO Z, SONG J. Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection[C] // 2004 IEEE International Geoscience and Remote Sensing Symposium, 2004, 3: 2008–2010.
- 27 REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C] // Advances in Neural Information Processing Systems, 2015: 91–99.
- 28 SCHMIDHUBER J. Deep learning in neural networks: an overview[J]. Neural Networks, 2015, 61: 85–117.
- 29 MAAS A L, HANNUN A Y, NG A Y. Rectifier nonlinearities improve neural network acoustic models[C] // Proceedings of the 30th International Conference on Machine Learning (ICML), 2013, 30(1): 1–6.
- 30 JIA Y, SHELHAMER E, DONAHUE J, et al. Caffe: convolutional architecture for fast feature embedding[C] // Proceedings of the 22nd ACM International Conference on Multimedia (ACM), 2014: 675–678.
- 31 KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C] // Advances in Neural Information Processing Systems, 2012: 1097–1105.
- 32 GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580–587.
- 33 HARRINGTON P. Machine learning in action[M]. Greenwich, CT: Manning Publications, 2012: 116–133.