

基于改进 U-net 的遥感影像建筑物提取

任欣磊^{1*}, 王阳萍^{1,2,4}, 杨景玉^{1,3}, 高德成⁴

¹兰州交通大学电子与信息工程学院, 甘肃 兰州 730070;

²兰州交通大学计算机科学与技术国家级实验教学示范中心, 甘肃 兰州 730070;

³兰州交通大学甘肃省人工智能与图形图像处理工程研究中心, 甘肃 兰州 730070;

⁴兰州交通大学甘肃省轨道交通装备系统动力学与可靠性重点实验室, 甘肃 兰州 730070

摘要 针对在遥感影像建筑物提取过程中, 建筑物与周围环境信息混淆导致提取精度下降的问题, 提出了一种低维特征信息增强的改进 U 型卷积神经网络(U-net)模型, 用于遥感影像建筑物的提取。借鉴医学影像分割中应用广泛的 U-net 模型对建筑物进行提取; 考虑到在网络传播过程中低维细节信息逐级削弱, 在特征金字塔中的特征图与扩张路径同级上的特征融合前, 先与上一层级的特征图进行融合, 进一步优化了提取结果的边缘提取精度。在覆盖范围约 340 km² 的遥感影像数据集上进行实验, 结果表明本文提出的方法在交并比、像素精度和 Kappa 系数 3 个指标上的均值分别达到 83.9%、92.8% 和 83.6%, 均优于模糊 C 均值、全卷积网络与经典 U-net 方法。

关键词 遥感; 建筑物提取; U-net; 神经网络; 低维特征

中图分类号 P237

文献标识码 A

doi: 10.3788/LOP56.222801

Building Detection from Remote Sensing Images Based on Improved U-net

Ren Xinlei^{1*}, Wang Yangping^{1,2,4}, Yang Jingyu^{1,3}, Gao Decheng⁴

¹ School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China;

² Experiment Teaching Center on Computer Science, Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China;

³ Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphics & Image Processing,
Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China;

⁴ Gansu Provincial Key Laboratory of System Dynamics and Reliability of Rail Transport Equipment,
Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China

Abstract The building environment in urban areas is complex. Achieving high building detection accuracy from remote sensing images is challenging because of the difficulty associated with distinguishing between buildings and the environmental information. To solve this problem, an improved U-type convolutional neural (U-net) network with enhanced low-dimensional feature information is proposed for detecting buildings from the remote sensing images. Initially, a building is detected using the U-net network model typically employed for medical image segmentation. Further, the low-dimensional information is weakened at each step of the network propagation process. Before merging the feature map of a certain level in the feature pyramid with the feature map of the corresponding expansion path level, it is merged with the feature map of the previous level to optimize the detection accuracy of the building edges. According to the experimental results obtained using a dataset of remote sensing images covering a range of approximately 340 km², the proposed method achieves values of 83.9%, 92.8%, and 83.6% for the intersection-over-union (IoU), pixel accuracy, and Kappa coefficient, respectively, demonstrating its superior performance when compared with the fuzzy C-means clustering algorithm, fully convolutional neural

收稿日期: 2019-04-08; 修回日期: 2019-04-28; 录用日期: 2019-05-13

基金项目: 国家自然科学基金(41761082)、甘肃省科技计划(18JR3RA104, 1504FKCA038)、甘肃省教育厅科技项目(2017D-08)、兰州铁路局科研开发计划(201803)、兰州市人才创新项目(2015-RC-86)

* E-mail: 121931236@qq.com

network, and classic U-net methods.

Key words remote sensing; building detection; U-net; neural network; low-dimensional feature

OCIS codes 280.4750; 100.2960; 100.3008

1 引言

遥感影像建筑物提取是从感兴趣区域的遥感数据中准确检测建筑物区域并作定量分析的过程。在城市区域的遥感影像中,建筑物的面积达80%以上,所以建筑物提取在城市规划、城市变化监测、三维建模^[1-3]等方面起着至关重要的作用。目前,用于建筑物提取的传统方法可以归纳为监督分类、知识约束与模板匹配。其中:监督分类方法是先对建筑物的特征进行人工设计,然后对分类器进行训练,最后使用随机森林、条件随机场、支持向量机(SVM)等分类器对实验数据进行建筑物提取;知识约束的方法主要是通过建立约束条件解决假设测试问题,具有代表性的知识约束方法有矩形约束、几何辐射特性、菱形约束^[4]等;模板匹配方法首先建立明确的模板参数对建筑物进行描述,然后根据相关度搜索算法匹配最佳的结果^[5]。城市区域的遥感影像具有建筑物环境复杂多样的特点,随着遥感技术的迅速发展,获取到的影像分辨率不断提高,而传统方法很难满足建筑物提取精度的要求。

随着计算机硬件的快速发展,深度学习的方法在图像处理领域的分类与检测任务中表现出了强大的性能^[6-7],其中以卷积神经网络(CNN)作为代表的深度学习方法在图像级的回归与分类任务中表现良好。但是,建筑物提取是像素级的语义分割任务,使用CNN会使内存开销急剧上升、计算效率低下,并且感知区域受到限制^[8]。据此,全卷积神经网络(FCN)^[9]去除了传统CNN中的全连接层,并对末端的特征图进行反卷积生成与输入图像分辨率一致的分割结果,实现了像素级的分类,并已被应用到遥感领域的目标提取中^[10-12]。但随着网络深度的增加,特征维度增加,细节信息丢失,导致提取结果精度不足。后续又出现了SegNet^[13]、DeconvNet^[14]以及U型卷积神经网络(U-net)^[15]等。其中,U-net以FCN为基础,采用对称的结构设计,融合了低维与高维的特征,在医学图像分割领域达到了较高的提取精度。刘哲等^[16]设计了一种将改进的U-net与Morphsnakes算法结合的计算机断层扫描(CT)图像肝脏分割方法,分割结果得到了增强。Guillaume等^[17]基于U-net模型,用Adam

Optimizer取代随机梯度下降算法,添加批量标准化加速训练,并使用基于Dice系数的损失,实现了对遥感影像建筑物的提取。伍广明等^[18]提出了一种双重约束的改进U-net网络,优化了参数更新过程,提高了建筑物提取精度。现有方法使用U-net对遥感影像进行语义分割任务,分割结果存在边缘平滑、细节信息丢失等问题,对精度造成了一定的影响。

基于以上分析,本文提出了一种改进U-net网络模型的建筑物提取方法,建立了低维特征信息增强的U-net网络模型,增强了网络传播中的低维细节信息,优化了边缘提取结果,提升了提取精度。

2 FCN与U型卷积神经网络

Long等^[9]提出了一种全卷积神经网络,改善了传统CNN对图像的像素级分割效果。CNN在图像级的分类和回归任务中达到了较高的精度水平,通常会在多个卷积层之后连接若干个全连接层,将经过卷积操作后产生的特征图(feature map)映射成为一个固定长度的 N 维特征向量,用 N 维特征向量来预测 N 个类别中每一类的概率值,进而得到输入图像的分类。

遥感影像建筑物提取与上述任务的区别在于,要对输入影像中的每一个像素进行分类,得到像素级的分类结果。尽管CNN能以单个像素为中心来设置滑动窗口,并通过对窗口特征建模来得到像素级的语义分割结果,但因为相邻窗口之间存在重叠区域,会产生大量的冗余信息,所以时间复杂度急剧增加。另外,窗口大小的选择也将是一个难题:窗口太小会丢失目标的上下文信息,降低准确率;窗口太大则会增大计算开销,增加内存负担。

针对以上问题,FCN对CNN进行了改进。如图1所示,FCN与传统神经网络相比,在经过多层卷积和池化操作之后不再使用全连接层构造特征向量预测概率,而是利用反卷积操作将高维特征图进行上采样,得到与输入图像具有相同分辨率的预测结果。这样的网络结构可防止输入图像的空间信息在传播过程中丢失,能对图像中的每一个像素进行预测。另外FCN不用对图像进行逐窗口计算,所以计算效率得到了很大提升。

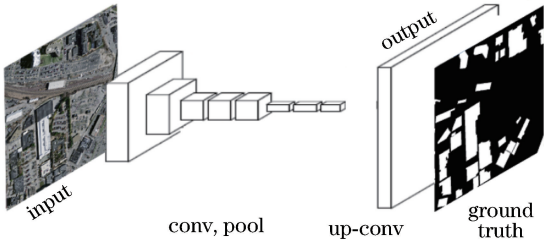


图 1 FCN 网络结构示意图

Fig. 1 Architecture of fully convolutional neural networks

虽然 FCN 的改进能得到与输入图像大小相同的分割结果,但是得到的预测图像往往过于平滑,细节损失比较严重。主要原因是输入图像经过了多次池化操作,使得尾端的神经元能够接收更大范围的信息,即更大的感受野。但同时也导致了图像细节信息的丢失,使提取的结果边缘轮廓效果不理想。因此,为解决上述问题,提高细节信息的提取精度,FCN 将特征金字塔中的低维特征图与反卷积之后的结果融合。

U-net 对这种低维特征与高维特征相融合的思想进行了扩展,如图 2 所示,在 U-net 网络结构中包含向下的收缩路径与向上的扩张路径。在收缩路径中,输入图像经过多次卷积与池化操作,形成从低维到高维的特征金字塔,并最终得到抽象的高维特征图;经过扩张路径上的若干卷积层与上采样层,生成与特征金字塔对应层级的特征图,最后得到与输入图像分辨率相同的像素级分割结果并输出。在扩张路径的上采样过程中,每进行一次反卷积操作,特征图的维度就会降为原来的一半,然后被降维的特征图会通过矩阵级联的方式与特征金字塔中对应层级的特征图进行融合,融合后的特征图既包含了金字塔顶端的高维抽象特征,又引入了各层级中低维特征的细节信息。

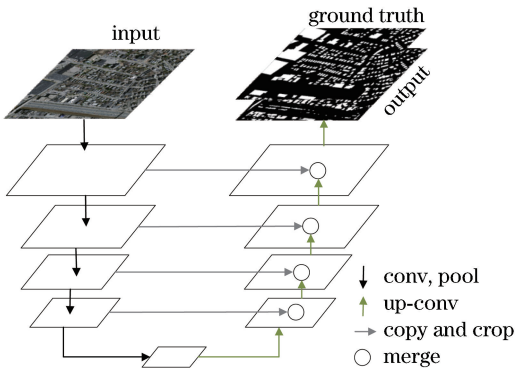


图 2 U-net 网络结构示意图

Fig. 2 Architecture of U-net network

3 低维特征增强的改进 U-net 网络

U-net 最初被应用于医学图像分割领域,在应用于遥感影像处理之后,提取目标的相邻环境变得复杂,导致提取结果的精度不足^[19]。其主要原因是网络深度的增加弱化了低维度的细节信息,导致提取结果的边缘轮廓效果较差。据此,本研究通过引入特征金字塔中较低维特征图,提出一种低维特征增强的改进 U-net 网络结构。图 3 给出本研究使用的网络结构图。网络左半部分通过逐级减少输入数据的空间维度来提取高层特征。每一层内部使用两次 3×3 的卷积操作,卷积核的数目相同。每两层之间使用池化操作进行空间降维,减少了不必要的高频信息,特征图的尺寸也相应地缩小。从低层到高层,卷积核的数目从 64 成倍增多至 1024,特征图从 572×572 缩小至 32×32 。

右半部分自底向上由反卷积组成,逐步恢复图像的细节和空间维度。在低维特征增强的 U-net 模型中,特征金字塔中相应层级上的特征图在与扩张路径同级上的特征融合前,先与上一层级的特征图进行融合。这种网络结构能够在一定程度上增强低维特征,在向后传递过程中能够防止细节信息的丢失。为了消除上采样带来的混淆作用,再使用 3×3 的卷积核进行两次卷积。经过 5 层的重复操作之后,最后一层的输出图像恢复到与输入图像相同的尺寸,并对每个像素进行预测。

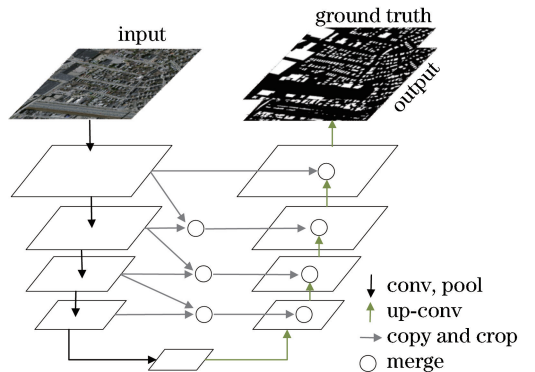


图 3 低维特征增强的 U-net 网络结构示意图

Fig. 3 Architecture of improved U-net network with enhanced low-dimensional features

为了防止梯度消失和梯度爆炸,实验中在每次卷积后引入批量归一化(BN)层,从而能够将前一层输出的特征归一化为均值为 0、方差为 1 的分布,然后将得到的结果输入到激活函数。设该层有 d 维输入 $\mathbf{X} = (x^{(1)} \cdots x^{(d)})$,引入可学习参数 γ, β (分别

表示标准差与均值)。BN 的表达式为

$$\begin{cases} y^{(k)} = \gamma^{(k)} \hat{x}^{(k)} + \beta^{(k)} \\ \gamma^{(k)} = \sqrt{\text{Var}[x^{(k)}]} \\ \beta^{(k)} = E[x^{(k)}] \end{cases}, \quad (1)$$

式中: $y^{(k)}$ 为归一化结果。

因为加入了上一层级的特征图信息,低维信息得到增强,但特征被明显放大,容易导致过拟合,所以在网络中引入 Dropout 层^[20]。在训练过程中,按概率 p 随机选择神经元,保留权重不更新,等下次迭代时再进行更新。使部分节点不工作的策略,可以提高网络的泛化能力,从而防止过拟合。Dropout 的表达形式为

$$y = \begin{cases} \frac{x}{1-p}, & \text{if } u > p, \\ 0, & \text{otherwise} \end{cases}, \quad (2)$$

式中: $u \sim U(0,1)$, U 为伯努利随机数。本实验中 p 取值为 0.25。

遥感影像建筑物提取的流程可分为训练阶段与测试阶段。图 4 为本文基于改进 U-net 网络的遥感影像建筑物提取流程。

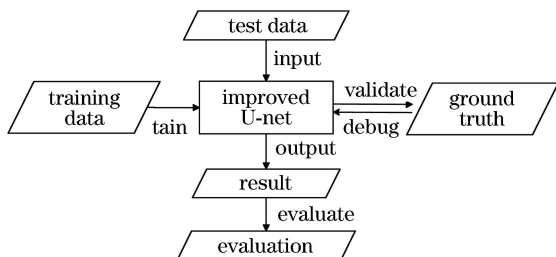


图 4 遥感影像建筑物提取流程

Fig. 4 Flow chart of building detection using proposed method

训练阶段:

1) 根据网络模型,将训练区域的遥感影像数据按 $572 \text{ pixel} \times 572 \text{ pixel}$ 的间距划分成相同大小的子图像,并用相同的方法得到对应的地面真值图。为了保证数据集中正负样本分布均匀,去除其中建筑物像素占比过低的图像^[21-22]。将所得图像分为训练数据与验证数据。

2) 利用反向传播与随机梯度下降算法对改进后的模型进行训练,然后对训练后的模型进行验证与精度评价,根据验证结果反向传播,调整模型参数。重复训练与调整参数,直至模型趋于稳定。

测试阶段:

1) 另选取 4 个不重叠区域的遥感影像,对训练完毕的模型进行测试。将测试区域划分成

$572 \text{ pixel} \times 572 \text{ pixel}$ 大小的测试子图像,再将它们依次输入到模型中进行建筑物提取,得到预测结果。

2) 为了防止由于漏检而在建筑物内部形成的微小空洞及裂缝,对提取结果进行形态学闭运算,得到最终的建筑物提取结果。

3) 最终,根据地面真值数据对提取结果进行精度评价。

4 实验与分析

4.1 实验介绍

以美国马萨诸塞州的遥感影像^[23]作为实验数据,该影像由覆盖面积约 340 km^2 的航拍影像组成,空间分辨率为 0.6 m ,数据的规模能够使网络模型的性能得到充分发挥。其中,划分后的子图像每幅为 $572 \text{ pixel} \times 572 \text{ pixel}$,面积约为 0.33 km^2 。数据集主要涵盖城市和郊区,包含各种规模的建筑物及地表覆盖物。图 5 所示为实验区域的影像。

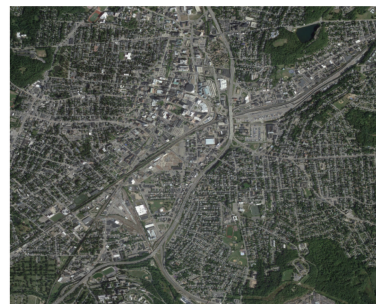


图 5 建筑物提取区域影像

Fig. 5 Image of building detection area

在划分好的遥感影像中,将建筑物占比少于 20% 的影像与对应的地面真值图去除。并按 8:2 的比例分为训练数据与验证数据,另外选取 4 个不重叠区域约 30 km^2 的影像作为测试数据。对比实验中,选用了传统分类方法中的模糊 C 均值聚类算法 (FCM),以及卷积神经网络方法中的 FCN、U-net、本文提出的改进 U-net 网络进行对比评价。

4.2 实验结果与分析

图 6 为使用 4 种方法对影像中的建筑物提取得到的结果。可以看出:FCM 算法提取的建筑物边缘与道路粘连现象比较明显,对道路的误检较高;其他 3 种方法相比较,U-net 与改进 U-net 在建筑物密集的区域,能更好地区分建筑物与背景;与经典 U-net 相比,改进 U-net 在降低误检率的同时提高了边缘的提取精度。

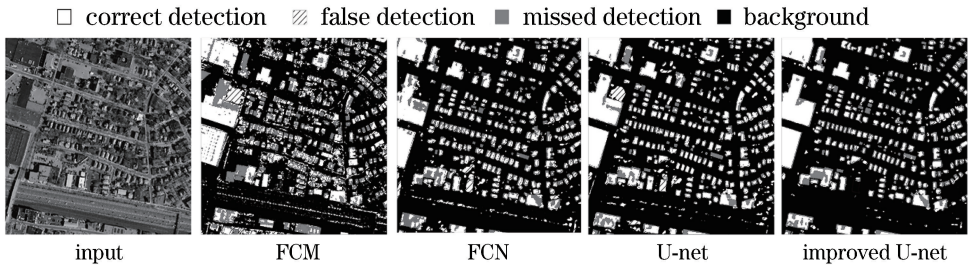


图6 建筑物提取结果

Fig. 6 Results of building detection using different methods

随机选取若干建筑物局部提取结果,如图7所示。从图中可以看出 FCM 对建筑物的形态检测效果较差,有较高的误检率与漏检率,并且椒盐噪声明显。在另外3种卷积神经网络方法中:由于(a)列与(b)列中建筑物表面有遮挡,故 FCN 与 U-net 模型的提取精度明显下降,建筑物内部出现了漏检空洞,改进 U-net 受遮挡干扰较小;(c)列中,FCN 与 U-net 在建筑物内部都出现了明显的漏检,改进 U-net 有效地避免了这种情况的发生;对于如(d)列和(e)

列这种表面特征单一、周围环境简单的建筑物,3种模型均能够较准确地对建筑物进行提取;(f)列中由于存在阴影,3种方法均出现了明显的阴影部分漏检,但改进 U-net 对剩余部分的检测结果较稳定。改进后的 U-net 模型对建筑物内部阴影遮挡部分的识别能力不佳,但与改进前相比,在一定程度上解决了提取结果边缘轮廓平滑的问题,能够更精确地对建筑物进行识别,说明改进的 U-net 增强了网络对细节部分建筑物的提取能力。

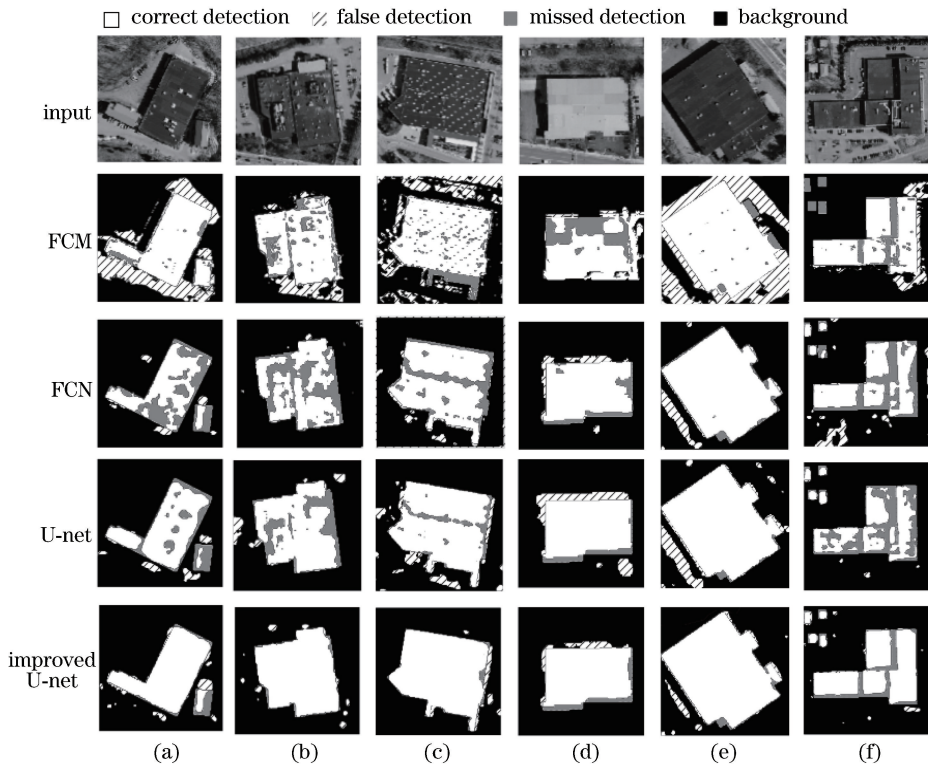


图7 建筑物提取局部结果

Fig. 7 Local results of building detection using different methods

本文采用交并比(IoU)、检测准确率(PA)和 kappa 系数^[24]评价对4种方法的提取结果进行了评价,上述三个指标分别代表检测结果的整体精度水

平、正确部分所占比例系数和提取结果与真值的一致性。交并比、检测准确率和 kappa 系数的计算公式为

$$\left\{ \begin{aligned} f_{\text{IoU}} &= \frac{T_P}{T_P + F_P + F_N} \\ f_{\text{PA}} &= \frac{T_P}{T_P + F_P} \\ f_{\text{kappa}} &= \frac{P_{\text{CC}} - P_{\text{RE}}}{1 - P_{\text{RE}}} \\ P_{\text{CC}} &= \frac{T_P + T_N}{N} \\ P_{\text{RE}} &= \frac{(T_P + F_P)N_C + (F_N + T_N)N_U}{N^2} \end{aligned} \right., \quad (3)$$

式中: T_P 为正确检测出建筑物的像素数; F_P 为误检像素数; F_N 为漏检像素数; T_N 为正确检测的背景像素数; N 为像素总数; N_C 为建筑像素数; N_U 为背景像素数。

表 1 为使用 4 种方法对 4 个测试区域建筑物提

表 1 建筑物识别定量评价比较

Table 1 Quantitative evaluation results of building detection for different methods

Test area	IoU				PA				Kappa			
	FCM	FCN	U-net	Improved U-net	FCM	FCN	U-net	Improved U-net	FCM	FCN	U-net	Improved U-net
1	66.4	67.8	68.8	73.8	83.6	90.7	85.6	92.4	70.3	64.2	73.3	80.0
2	69.3	84.8	87.5	91.6	83.0	89.1	92.9	94.4	72.2	79.9	87.3	88.8
3	73.5	78.5	88.0	90.1	88.4	94.2	94.2	97.3	76.3	73.0	84.5	84.7
4	70.6	79.9	77.2	80.1	95.2	81.7	85.9	87.2	75.0	84.0	78.1	80.8
Mean	69.9	77.7	80.4	83.9	87.5	88.9	89.6	92.8	73.5	75.3	80.8	83.6

5 结 论

将 U-net 应用于遥感影像建筑物提取, 实现了对目标区域建筑物的精确提取。另外, 在复杂环境下的城市区域进行建筑物提取时, 环境信息与建筑物信息容易混淆, 造成提取效果不佳。针对上述问题, 提出一种低维特征增强的 U-net 神经网络, 它可增强网络传递过程中的低维细节信息, 提高模型对建筑物细节的获取能力。在覆盖范围达 340 km² 的遥感影像数据集上进行训练及测试, 并与其他方法进行比较, 结果表明, 本文提出的改进 U-net 模型在 IoU、PA 和 Kappa 系数上的均值分别达到 83.9%、92.8%、83.6%, 均优于 FCM、FCN 及经典 U-net 方法。另外, 建筑物存在的太阳光照阴影、遮挡以及建筑物自身特征差异, 都会对建筑物提取的完整性造成一定影响, 仅考虑像元自身及其局部区域的颜色或亮度特征是不全面的。在今后的工作中, 需针对性地对影像中存在的建筑物阴影、遮挡进行研究, 以提升建筑物的提取效果。

取后 3 项指标的评价结果, 可以看出: FCM 在 3 个指标上均低于其他方法; U-net 模型在 3 个指标上均高于 FCN 模型, 分别提高了 2.7%、0.7%、5.5%; 与经典 U-net 模型相比, 改进 U-net 模型的性能得到了提升, 各指标的均值分别提高了 3.5%、3.2%、2.8%, 均高于其他 3 种方法。

本文实验使用惠普 ProDesk 498 G3 MT 进行深度卷积神经网络的训练。该计算机配备有 NVIDIA GeForce GTX 1050Ti 显卡 (4 GB)、64 位 Ubuntu 操作系统。3 种卷积神经网络模型 (FCN、U-net、改进 U-net) 在训练阶段分别耗时 6.73、11.48、11.82 h。总体上, 改进后的 U-net 网络未显著增加时间成本, 并且训练完毕的模型能在一定程度上提升建筑物的提取结果。

参 考 文 献

- [1] Ferro A, Brunner D, Bruzzone L. Automatic detection and reconstruction of building radar footprints from single VHR SAR images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(2): 935-952.
- [2] Vanegas C A, Aliaga D G, Benes B. Automatic extraction of Manhattan-World building masses from 3D laser range scans [J]. IEEE Transactions on Visualization and Computer Graphics, 2012, 18(10): 1627-1637.
- [3] Lin X G, Zhang J X. Object-based morphological building index for building extraction from high resolution remote sensing imagery [J]. Acta Geodaetica Et Cartographica Sinica, 2017, 46(6): 724-733.
林祥国, 张继贤. 面向对象的形态学建筑物指数及其高分辨率遥感影像建筑物提取应用 [J]. 测绘学报, 2017, 46(6): 724-733.
- [4] McGlone J C, Shufelt J A. Projective and object space geometry for monocular building

- extraction[C]//1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, June 21-23, 1994, Seattle, WA, USA. New York: IEEE, 1994: 54-61.
- [5] Lhomme S, He D C, Weber C, *et al.* A new approach to building identification from very-high-spatial-resolution images[J]. *International Journal of Remote Sensing*, 2009, 30(5): 1341-1354.
- [6] Russakovsky O, Deng J, Su H, *et al.* ImageNet large scale visual recognition challenge [J]. *International Journal of Computer Vision*, 2015, 115(3): 211-252.
- [7] Li C X, Cao L, Zhang Y L, *et al.* Knowledge-based deep reinforcement learning: a review [J]. *Systems Engineering and Electronics*, 2017, 39(11): 2603-2613.
李晨溪, 曹雷, 张永亮, 等. 基于知识的深度强化学习研究综述 [J]. *系统工程与电子技术*, 2017, 39(11): 2603-2613.
- [8] Guo Z L, Shao X W, Xu Y W, *et al.* Identification of village building via google earth images and supervised machine learning methods [J]. *Remote Sensing*, 2016, 8(4): 271.
- [9] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 3431-3440.
- [10] 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.
- [11] Marmanis D, Wegner J D, Galliani S, *et al.* Semantic segmentation of aerial images with an ensemble of CNNs [J]. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2016, III-3: 473-480.
- [12] Song Q S, Zhang C, Chen Y, *et al.* Road segmentation using full convolutional neural networks with conditional random fields [J]. *Journal of Tsinghua University (Science and Technology)*, 2018, 58(8): 725-731.
宋青松, 张超, 陈禹, 等. 组合全卷积神经网络和条件随机场的道路分割 [J]. *清华大学学报(自然科学版)*, 2018, 58(8): 725-731.
- [13] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [14] Noh H, Hong S, Han B. Learning deconvolution network for semantic segmentation[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1520-1528.
- [15] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W, *et al.* Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [16] Liu Z, Zhang X L, Song Y Q, *et al.* Liver segmentation with improved U-Net and Morphsnakes algorithm[J]. *Journal of Image and Graphics*, 2018, 23(8): 1254-1262.
刘哲, 张晓林, 宋余庆, 等. 结合改进的 U-Net 和 Morphsnakes 的肝脏分割 [J]. *中国图象图形学报*, 2018, 23(8): 1254-1262.
- [17] Guillaume C, Cristian B A, Ianis B L. Satellite image segmentation for building detection using U-net[EB/OL]. (2017). [2019-04-07]. <http://cs229.stanford.edu/proj2017/final-posters/5148174.pdf>.
- [18] Wu G M, Chen Q, Ryosuke S, *et al.* High precision building detection from aerial imagery using a U-Net like convolutional architecture [J]. *Acta Geodaetica Et Cartographica Sinica*, 2018, 47(6): 864-872.
伍广明, 陈奇, Ryosuke S, 等. 基于 U 型卷积神经网络的航空影像建筑物检测 [J]. *测绘学报*, 2018, 47(6): 864-872.
- [19] Gao X J, Zheng X D, Liu Z X, *et al.* Automatic building extraction from high resolution visible images based on shifted shadow analysis [J]. *Acta Optica Sinica*, 2017, 37(4): 0428002.
高贤君, 郑学东, 刘子潇, 等. 基于偏移阴影分析的高分辨率可见光影像建筑物自动提取 [J]. *光学学报*, 2017, 37(4): 0428002.
- [20] Gao W, Zhou Z H. Dropout Rademacher complexity of deep neural networks [J]. *Science China Information Sciences*, 2016, 59(7): 072104.
- [21] Zhang X N, Zhong X, Zhu R F, *et al.* Scene classification of remote sensing images based on integrated convolutional neural networks [J]. *Acta Optica Sinica*, 2018, 38(11): 1128001.
张晓男, 钟兴, 朱瑞飞, 等. 基于集成卷积神经网络

- 的遥感影像场景分类[J]. 光学学报, 2018, 38(11): 1128001.
- [22] Peng Y F, Song X N, Zi L L, *et al.* Remote sensing image retrieval based on convolutional neural network and modified fuzzy C-means [J]. Laser & Optoelectronics Progress, 2018, 55(9): 091008.
彭晏飞, 宋晓男, 訾玲玲, 等. 基于卷积神经网络和改进模糊 C 均值的遥感图像检索[J]. 激光与光电子学进展, 2018, 55(9): 091008.
- [23] Mnih V. Machine learning for aerial image labeling [D]. Canada: University of Toronto, 2013: 64-73.
- [24] Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to boosting [J]. Journal of Computer and System Sciences, 1997, 55(1): 119-139.