

基于改进的SegNet分割网络的遥感图像分割

雷竞雄

成都理工大学数理学院, 四川 成都

收稿日期: 2022年10月9日; 录用日期: 2022年11月8日; 发布日期: 2022年11月16日

摘要

城市遥感图像的建筑物分割技术, 对城市的土地规划、资源利用、灾害防治具有广泛的应用价值, 对此提出了一种基于改进的SegNet分割网络的城市遥感图像分割方法。1) 为了解决所需数据量过大、网络层数过深等问题, 模型在SegNet网络结构的基础上, 删减了部分卷积层, 使网络结构更加简洁, 运行速度更快。2) 针对SegNet网络保留特征信息不足、图像分割边缘粗糙的缺点, 针对这一问题, 我们设计了一个新的多尺度特征提取模块, 通过三个不同尺度的卷积核提取目标信息。以SegNet为基线, 通过用该模块替换跳过连接, 我们提出了一种多尺度特征提取SegNet。该方法可以对跳过连接中的浅层特征信息进行二次特征提取, 细化细节信息, 缩小低级特征与高级特征之间的语义差距。它不仅可以提高网络提取多尺度特征信息的能力, 从更大的范围到更多的层次来提取遥感图像中建筑物的边缘细节信息, 而且可以增加跳过连接的数量, 以减少网络过度拟合。多尺度模块实验结果表明, 所提出的方法与现有的FCN、U-Net、SegNet网络相对比, 准确率和交并比有明显的提升, 并且对遥感图像的城市建筑物分割边缘有了较好的改善。

关键词

语义分割, 遥感图像, 建筑物, SegNet网络, 卷积神经网络

Remote Sensing Image Segmentation Based on Improved SegNet Segmentation Network

Jingxiong Lei

College of Mathematics and Physics, Chengdu University of Technology, Chengdu Sichuan

Received: Oct. 9th, 2022; accepted: Nov. 8th, 2022; published: Nov. 16th, 2022

Abstract

The building segmentation technology of urban remote sensing image has a wide application value for urban land planning, resource utilization and disaster prevention and control. This paper proposes a segmentation method of urban remote sensing image based on improved SegNet segmentation network. 1) In order to solve the problems such as too large amount of data required and too deep number of network layers, the model has deleted some convolution layers based on the SegNet network structure, making the network structure simpler and faster. 2) In view of the shortcomings of SegNet network, such as insufficient feature information and rough edge of image segmentation, we designed a new multi-scale feature extraction module to extract target information through three convolution kernels of different scales. Taking SegNet as the baseline, we propose a multi-scale feature extraction SegNet by replacing the skip connection with this module. This method can extract secondary features from shallow feature information in the skip link, refine the details, and narrow the semantic gap between low-level features and high-level features. It can not only improve the ability of the network to extract multi-scale feature information, extract the edge details of buildings in remote sensing images from a wider range to more levels, but also increase the number of skipped connections to reduce network over fitting. The experimental results of multi-scale modules show that the proposed method has significantly improved the accuracy and intersection/merge ratio compared with the existing FCN, U-Net and SegNet networks, and has better improved the urban building segmentation edge of remote sensing images.

Keywords

Semantic Segmentation, Remote Sensing Images, Buildings, SegNet Network, Convolutional Neural Network

Copyright © 2022 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. 引言

图像的语义分割技术是指根据图像中表达的语义信息对图像的每一个像素进行分类,使得每一像素都有其对应的标签,从而得到具有逐像素语义注释的分割图像。高分辨率的城市遥感图像在城市的监测、规划、灾害防治等方面具有积极的意义,因此对城市遥感图像进行逐像素图像分割研究是必要的。当前,深度学习方法在图像处理上有广泛的应用,特别是基于卷积神经网络(CNN)的方法,在图像处理领域取得了显著效果,并对遥感图像处理的影响越来越大,例如:FCN [1]、U-Net [2]、HRNet [3]、SegNet [4]等方法。陈梦[5]、宋延强[6]等学者利用深度学习处理高分辨率遥感图像,从而实现建筑物与建筑物震害识别。林志斌[7]、王蓝玉[8]等学者使用 U-Net、Deeplab V3+网络进行高分辨率遥感图像地物分割。徐胜军[9]、张书瑜[10]等学者将传统的图像处理方法与深度学习相结合实现遥感图像建筑物分割。最近关于 U-Net、SegNet 的改进受到了各方学者的热爱。DIResUNet [11]通过集成初始模块、修改的残差块和密集全局空间金字塔池(DGSP)模块,改进了 U-Net 的分割精度。Li [12]等人在综合了 U-Net 不同层生成的多尺度特征,以设计多尺度跳跃连接和非对称卷积 U-Net (MACU-Net),在 U-Net 的基础上添加多尺度特征提取思想,目标的分割效果得到了较大提升。Chaitra Dayananda [13]与 Xu [14]等人在 SegNet 的基础上进

行改进用于医学图像分割。Nagaraj Yamanakkanavar [15]等人利用多尺度侧输入对判别信息的提取，提升网络的分割精度与效率。

SegNet 卷积编码 - 解码网络的核心思想是通过编码器的池化操作获得最大池化索引值，这为解码阶段提供了像素的空间位置关系。这种对称的编码 - 解码网络模型虽然在一定程度上能够获得精细的分割结果，但也存在一定弊端，即在解码时利用的信息只含有编码阶段最后一个池化层的特征，且该层特征图分辨率过小，保留的特征信息较少。为了解决 SegNet 网络保留特征信息不足、图像分割边缘粗糙的问题，本文在 SegNet 网络的基础上，结合传统的图像分割方法——高斯金字塔进行改进，最终提出了本文的模型。具体改进如下：1) 结构上的删减调整，以防止网络层数的深度带来的过拟合；2) 在传统的图像分割方法思想中，引入多尺度分割模块，从不同维度提取多重信息。

本文主要分为 4 章。第一章为引言部分。第二章主要介绍本文方法的主要原理。第三章说明了实验数据、实验环境及实验结果。第四章介绍了主要本文实验数据的缺陷问题及未来的改进方向。

2. 方法原理

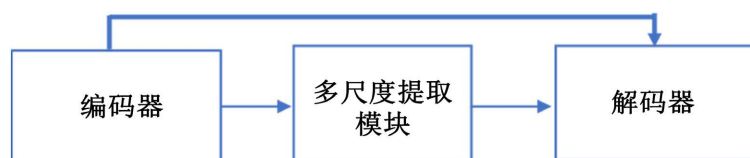


Figure 1. Network structure

图 1. 网络结构图

本模型主要包含三个部分：编码器、多尺度处理模块、解码器。如图 1 所示，1) 编码器：用于提取特征。2) 多尺度特征提取模块：对已获取的下采样浅层语义信息采用不同的卷积核大小进行卷积操作，用于提取和融合不同尺度、维度的深层特征，以保留更多的信息。3) 解码器，采用跳跃连接融合高低语义信息进行解码，并输出语义分割图。

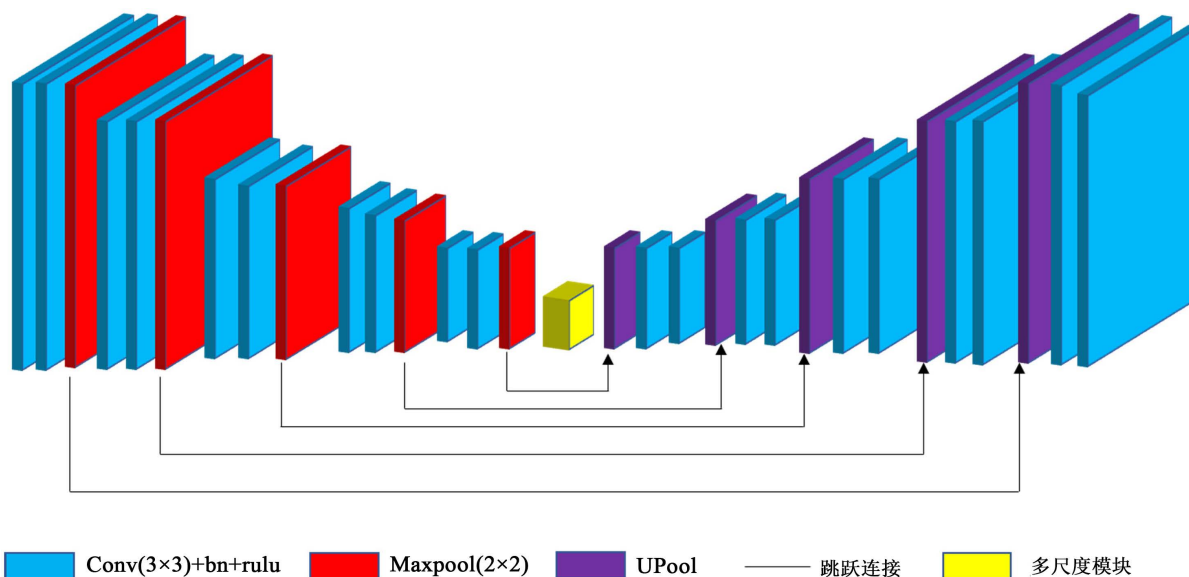


Figure 2. Improved SegNet segmentation network

图 2. 改进的 SegNet 分割网络

编码器、解码器保留了 SegNet 网络的主体结构，但考虑到 SegNet 网络进行分割时的不足，本文对 SegNet 网络进行了改进。具体网络结构如图 2 所示，其编码器主要包含三层结构，每层包含两个卷积层和一个最大池化层，共包含 10 个卷积层和 5 个池化层。解码器与编码器呈镜像关系，也包含 10 个卷积层和 5 个反池化层。其中，卷积层的卷积核大小设置为 3×3 ，步长为 1。同时，在每个卷积层之后，采用 Batch Normalization 算法使输入到下一层网络的数据与原始输入数据具有相似分布，激活函数使用稀疏且易于计算的 ReLU 函数。编码器中的池化操作采用大小为 2×2 的最大池化，同时保存了最大池化的索引位置。解码器中的反池化层利用编码器池化层所获得的最大池化索引将尺寸较小的特征图扩充，从而获得稀疏特征图，扩充的位置进行补 0 填充，再针对稀疏特征图进行卷积操作，使得反池化操作中填充的 0 值发生变化，稀疏特征图像变得致密。

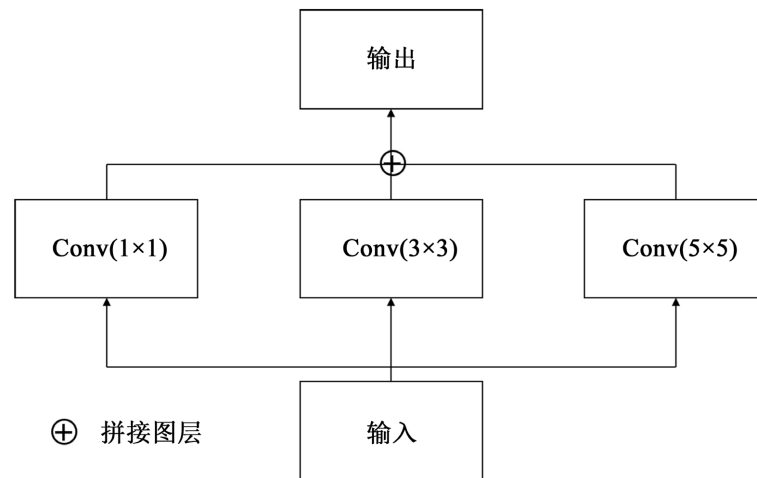


Figure 3. Multi-scale extraction module
图 3. 多尺度提取模块

本文的多尺度特征模块是在图像中用多尺度的表达方式来解释图像的有效结构，它在各种图像处理中都有应用。如图 3 所示，本文提出的多尺度特征模块包含三个平行卷积核。滤波器的大小为 1×1 、 3×3 和 5×5 。滤波器大小为 1×1 的卷积层相当于同等映射，充分保留原始信息。其余不同大小的滤波器进一步提取不同尺度的丰富特征信息，将三个滤波器的输出拼接起来，然后作为上采样块的输入。通过这种方式，底层特征信息被多尺度特征模块处理，然后输入到深度神经网络中。即使在深度网络中，我们也可以获得丰富的、不同尺度的浅层语义信息。与 SegNet 网络的池化索引相比，我们的多尺度特征模块为解码器提供了更准确的空间位置信息和更丰富的语义信息，对遥感图像的分割更准确，在缺失分割、误分类、边缘保留、抗噪声干扰等方面效果更好。于是，我们采用大小不同的卷积核对含有噪声干扰的深层次抽象特征进行不同尺度的特征提取来弥补 SegNet 网络池化索引保留的特征信息较少的问题。

3. 实验

3.1. 数据集和训练

数据集来自于 University of Toronto 的 Massachusetts 建筑遥感图像[16]，包含 32 幅超高分辨率遥感图像，其中 28 幅用于训练，4 幅用于测试，图像的尺寸均为 1500×1500 像素，并且数据集标签已知。数据内容划分为 2 种常见的土地覆盖类别，分别为城市建筑物与其他背景，如图 4 所示，展示了其中一幅。在进行实验时，首先对数据进行预处理，预处理过程包括数据切割和数据增强、加噪，具体操作如

下: 1) 将数据集中的图像均匀地切割成尺寸为 256×256 像素的 800 张训练数据集 100 张测试数据集。
2) 对得到的训练图像进行数据增强、加噪, 即对训练集中的每幅图像进行上下左右翻转, 或旋转 90° 、 180° 和 270° 并添加随机噪声, 真实模拟实际情况。最终获得的训练集中包含 800 幅训练图像, 测试集中包含 100 幅测试图像。



Figure 4. Sample dataset (a) remote sensing image, (b) label
图 4. 数据集样例(a) 遥感图像, (b) 标签

为了更好地评价网络的分割效果, 我们使用训练集的准确率与交并比作为模型的评价指标。在训练过程中, 本文采用随机梯度下降算法更新网络权值, 损失函数使用交叉熵损失函数, 迭代次数设置为 200。

3.2. 实验环境

本文方法及对比方法使用 PyTorch 实现网络主干。在训练中, 优化器为 Adam, 学习率为 0.0001, 批量大小为 2, 权重衰减默认为 0, betas 为 0.5 和 0.9999。所有的实验都是在处理器为 Intel(R) Core(TM) i7-9700k CPU @ 3.60 GHz 和 NVIDIA GeForce RTX 2060 实现的。

3.3. 实验分析与讨论

为了进一步验证本文方法的有效性, 本文选取了 U-Net、FCN、SegNet 作为对比方法。模型的输入均为尺寸大小为 256×256 的 RGB 三通道彩色图像, 输出是与输入图像大小相同的预测标签图, 并且迭代次数相同。

表 1 中给出了本文方法与对比方法对相同数据集使用同样迭代次数进行实验所获得的准确率与交并比数值对比表。本文方法相对于 SegNet 方法, 准确率与交并比有了明显的提升。U-Net 网络数据最高, FCN 网络的交并比最低。

Table 1. Comparison of accuracy and cross/merge ratio values of different models
表 1. 不同模型的准确率与交并比数值对比

评价指标/模型	本文方法	SegNet	U-Net	FCN
训练准确率	0.9800	0.9737	0.9864	0.9745
训练交并比	0.9048	0.8784	0.9349	0.8606

图 5 展示了本文所提出的方法与对比方法的比较。图 5(b)是图 5(a)的标签, 可见建筑物较为密集, 且有一定的规律。由图 5 可见, 本文所提出的分割方法不仅保证了建筑物分割的准确率, 而且对建筑物的边缘信息也进行了有效的保留。对比方法 FCN 训练的准确率较高, 但其分割结果不理想, 空间信息与图像信息丢失较为严重。分割网络 SegNet, 虽然建筑物的位置信息保留较好, 但边缘细节保留不完整,

并且将图像的某些干扰信息错误识别为建筑物。U-Net 分割网络与 SegNet 分割网络类似，其对边缘信息保留不完整。

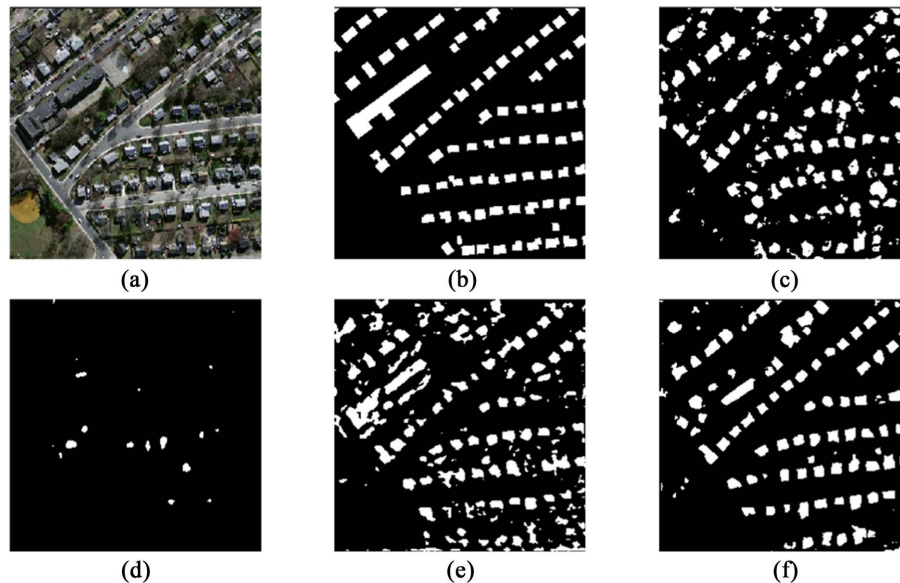


Figure 5. Comparison of segmentation results of different methods (a) original image; (b) labels; (c) U-Net segmentation results; (d) FCN segmentation results; (e) SegNet segmentation results; (f) improved SegNet segmentation results

图 5. 不同方法的分割结果比较(a) 原图; (b) 标签; (c) U-Net 分割结果; (d) FCN 分割结果; (e) SegNet 分割结果; (f) 改进 SegNet 分割结果

4. 总结

本文对 SegNet 分割网络做了结构上的简化和多尺度模块的添加，有效改善了城市遥感图像建筑物的分割效果，图像的多尺度特征得到有效保留。但实际生活中，包括但不限于：天气、光照条件、阴影物遮挡等各种影响，数据集中并没有包含以上各种情况的遥感图像。在未来，将进一步考虑 SegNet 分割网络能否与传统的图像处理方法相交融，增加模型的迁移能力，提升网络的性能。

参考文献

- [1] Zhuang, J., Yang, J., Gu, L., *et al.* (2019) Shelfnet for Fast Semantic Segmentation. *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, Seoul, 27-28 October 2019, 847-856. <https://doi.org/10.1109/ICCVW.2019.00113>
- [2] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W.M., *et al.*, Eds., *Medical Image Computing and Computer-Assisted Intervention MICCAI 2015*, Springer International Publishing, Cham, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [3] Sun, K., Xiao, B., Liu, D., *et al.* (2019) Deep High-Resolution Representation Learning for Human Pose Estimation. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 5686-5696. <https://doi.org/10.1109/CVPR.2019.00584>
- [4] Badrinarayanan, V., Kendall, A. and Cipolla, R. (2017) SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 2481-2495. <https://doi.org/10.1109/TPAMI.2016.2644615>
- [5] 陈梦. 基于深度学习的建筑物震害遥感识别研究[D]: [硕士学位论文]. 北京: 中国地震局地震预测研究所, 2019.
- [6] 宋廷强, 李继旭, 张信耶. 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, 56(8): 26-34.

-
- [7] 林志斌, 黄智全, 颜林明. 基于 U-Net 的高分辨率遥感图像地物分类[J]. 电子质量, 2020(11): 69-76.
- [8] 王蓝玉. 基于 Deeplab V3+网络的遥感地物图像语义分割研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2020.
- [9] 徐胜军, 欧阳朴衍, 郭学源, Khan Taha Muthar. 基于多尺度特征融合模型的遥感图像建筑物分割[J]. 计算机测量与控制, 2020, 28(7): 214-219.
- [10] 张书瑜. 基于深度学习和多尺度多特征融合的高分辨率遥感地表覆盖分类研究[D]: [博士学位论文]. 杭州: 浙江大学, 2020.
- [11] Priyanka, Sravya, N., Lal, S., *et al.* (2022) DIResUNet: Architecture for Multiclass Semantic Segmentation of High Resolution Remote Sensing Imagery Data. *Applied Intelligence*, **52**, 15462-15482. <https://doi.org/10.1007/s10489-022-03310-z>
- [12] Li, R., Duan, C., Zheng, S., *et al.* (2022) MACU-Net for Semantic Segmentation of Fine-Resolution Remotely Sensed Images. *IEEE Geoscience and Remote Sensing Letters*, **19**, Article ID: 8007205. <https://doi.org/10.1109/LGRS.2021.3052886>
- [13] Dayananda, C., Choi, J.Y. and Lee, B. (2022) A Squeeze U-SegNet Architecture Based on Residual Convolution for Brain MRI Segmentation. *IEEE Access*, **10**, 52804-52817. <https://doi.org/10.1109/ACCESS.2022.3175188>
- [14] Xu, X.M., Wang, Y.X., Liang, Y., *et al.* (2022) Retinal Vessel Automatic Segmentation Using SegNet. *Computational and Mathematical Methods in Medicine*, **2022**, Article ID: 3117455. <https://doi.org/10.1155/2022/3117455>
- [15] Yamanakkanavar, N., Choi, J.Y. and Lee, B. (2022) SM-SegNet: A Lightweight Squeeze M-SegNet for Tissue Segmentation in Brain MRI Scans. *Sensors*, **22**, Article No. 5148. <https://doi.org/10.3390/s22145148>
- [16] Mnih, V. (2013) Machine Learning for Aerial Image Labeling. University of Toronto, Toronto.