

卷积神经网络下的高分二号卫星影像道路提取^①



孙 卓, 李冬伟, 赵泽宾, 张倩倩

(中国矿业大学(北京)机电与信息工程学院, 北京 100083)

通讯作者: 孙 卓, E-mail: sz429543422@163.com

摘 要: 本文针对深度学习神经网络对高分二号遥感影像道路提取时细节信息丢失较多、道路周围环境考虑不充分等情况, 在已有的研究成果上, 提出一种基于全卷积神经网络遥感影像道路提取的改进方案. 方案创新研究了全卷积神经网络的算法原理, 将预调色后的高分二号影像按一定尺寸分幅输出, 将输出图像及标签对应输入于以全卷积神经网络为基础的改进网络, 通过结合残差单元以及增加网络层数得到识别精度较高的道路提取图像. 实验表明, 该方法在同一样本中对高分二号卫星影像道路提取的效果有所提升, 道路的完整性和准确性有所提高.

关键词: 深度学习神经网络; 残差单元; 全卷积神经网络

引用格式: 孙卓, 李冬伟, 赵泽宾, 张倩倩. 卷积神经网络下的高分二号卫星影像道路提取. 计算机系统应用, 2020, 29(11): 128-133. <http://www.c-s-a.org.cn/1003-3254/7511.html>

Road Extraction of GF-2 Satellite Image Based on Convolutional Neural Network

SUN Zhuo, LI Dong-Wei, ZHAO Ze-Bin, ZHANG Qian-Qian

(School of Mechatronics and Information Engineering, China University of Mining and Technology-Beijing, Beijing 100083, China)

Abstract: More details may be lost and considerations of the surrounding environment of the road are inadequate when extract the road from GF-2 remote sensing satellite which based on the deep neural network. Aiming at these problems and based on the existing researches results, this study proposes an improvement proposal which using the full convolutional neural network to extract road from remote sensing images. The scheme innovatively researches the algorithm principle of the full convolutional neural network and outputs the pre-graded GF-2 images in a certain size. Then, the output images and the corresponding labels are input into the improved full convolutional neural network. At last, a road extraction image with higher recognition accuracy is obtained by combining residual unit and increasing the number of network layers. Experiments show that the effect on road extraction of GF-2 satellite images is improved in the same sample, the integrity and accuracy of the road are also improved.

Key words: deep neural network; residual unit; Full Convolutional Neural Network (FCNN)

随着我国遥感卫星的更新换代, 国内高分辨率遥感卫星影像星下点分辨率已经达到亚米级别, 能够识别更为精准的路况信息. 但是目前地理信息库中的道路更新问题大多数仍然采用费时费力的人工更新方式, 而利用遥感影像数据实现道路识别的结果越来越精确, 有逐渐代替人工更新的趋势^[1,2].

目前对遥感影像道路提取多采用传统方式, 如模

板匹配法^[3-6], 特定模型法^[7-10], 多源数据法^[11,12]等. 这些方法对低级别道路的提取效果很难达到生产的要求. 在各类遥感影像道路提取方法中, 基于深度学习领域的高分辨率遥感卫星道路提取方法能够应对各种复杂的遥感影像成像情况, 具有十分广阔的发展前景. 近年来, 基于深度学习的图像分割邻域不断涌现出优秀的算法结构, Long 等^[13]提出使用全卷积神经网络解决图像

① 收稿时间: 2019-12-17; 修改时间: 2020-01-04; 采用时间: 2020-01-22; csa 在线出版时间: 2020-10-29

语义分割问题, U-Net^[14] 网络用于解决小样本简单图像的分割问题, Seg Net^[15] 网络解决多次池化造成位置信息丢失的问题, Deep Lab^[16,17] 系列网络通过组合各网络优化点有效提高语义分割精确度等, 为神经网络在高分辨率遥感卫星道路提取方面的应用打下基础。

在遥感卫星影像道路的提取上, 国内外研究人员对基于深度学习的图像语义分割算法已经有了深入研究, Mnih 和 Hinton^[18] 提出了一种采用受限玻尔兹曼机从高分辨率航拍图像中探测道路区域的方法. 为了获得更好的结果, 采用预处理以减少输入数据的维度, 采用后处理用于去除断开的斑点并填充在道路中的缺口中. Saito 等^[19] 采用卷积神经网络直接从原始遥感图像中提取建筑物和道路. 这种方法比特定道路数据集的 Mnih 和 Hinton 方法^[18] 获得了更好的结果. Cheng 等^[20] 使用两级编码-解码结构的卷积神经网络同时完成道路检测和道路中心线提取任务, 他同时测试了多种神经网络的提取效果, 取得了优秀的综合性能. Zhang 等^[21] 提出使用残差单元结合 U-Net 网络提取道路影像, 得到了较为良好的遥感卫星影像道路提取效果。

由于高分遥感影像标签数据集难以获取且手工制作成本较高, 本文以部分高分二号卫星遥感影像瓦片为样本, 借鉴残差网络加深网络层次的思想, 以在小样本数据集上表现良好的 U-Net 网络为基础, 结合残差单元改善高分遥感卫星道路提取效果。

1 算法基本原理

1.1 U-Net

U-Net 网络由全卷积神经网络发展而来, 详细网络特征见图 1. U-Net 网络分为上采样和下采样两部分, 是典型的编码-解码结构, 下采样通过卷积和池化操作提取图像特征信息, 上采样则通过恢复细节特征和填充空白信息得到边缘分割图. 它继承了全卷积神经网络的简洁、通用的优点, 同时优化了上采样网络流程, 改变了跳跃连接的方式, 采用维度拼接的方式增加了图像信息的“厚度”, 使得卷积核在同样的图层上有了更多的选择, 优化边缘分割的效果。

1.2 残差单元

通过在神经网络中叠加卷积、池化等操作能够得到更加深层的网络, 深层网络能够整合低、中、高不

同层次的输入特征, 通过更深的网络提取更为丰富的特征, 这是由于将神经网络后面若干层作为恒等映射, 这一网络将等同于一个浅层网络. 但是随着层次的加深, 神经网络会产生梯度消失、梯度爆炸、和网络退化等问题, 严重影响网络训练和测试的效果. 通过输入数据的初始化和神经网络正则化层能够实现几十层以内的网络的收敛, 继续增加网络层数, 仍然存在梯度消失或爆炸的问题, 即前若干层网络权值参数更新停滞或过大导致后面的网络层不再起作用. He 等^[22] 通过对传统神经网络深入研究之后发现, 传统神经网络架构层次越深训练集上效果不一定越好, 随着网络层次的加深, 训练变得越来越困难, 网络参数无法再得到优化, 训练效果反而不如浅层网络, 即产生了网络退化问题. 而使用残差网络能够有效改善这一问题。

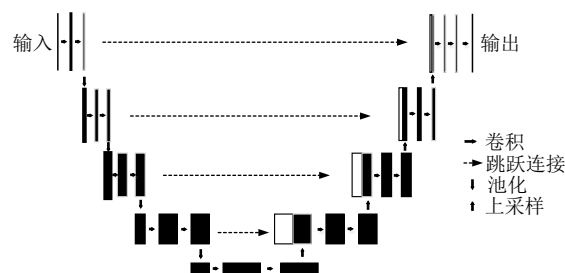


图 1 U-Net 神经网络架构

常规卷积单元和残差单元见图 2. 其中, 线性激活采用 ReLU 函数, 加操作使用 Addition 函数. x 作为输入, $F(x)$ 为输入 x 经过卷积操作之后的输出. 在残差网络结构中, 通过捷径连接的方式将输入 x 与原输出 $F(x)$ 相加作为新的输出 $H(x)$, 即 $H(x)=F(x)+x$, 当 $F(x)=0$ 时, $H(x)=x$, 该卷积层输入等于输出, 形成恒等映射, 则该卷积层只具有完整传递信息的作用, 增加网络深度的同时, 误差不会增加. 由于 $F(x)=H(x)-x$, 网络学习 $F(x)$ 相当于学习输入和输出的差值。

1.3 本文神经网络架构

Zhang 等^[21] 提出了加入残差模块的方法, 其提出的神经网络架构总共有 7 个残差模块, 分为 3 个下采样、3 个上采样和 1 个桥连接. 该网络架构由于网络深度不足, 残差单元对于图像细节补充作用难以显现, 而残差结构增加了冗余信息的传递, 阻碍了对道路细节的提取, 提取效果不理想. 通过增加深度, 使用随机失活 (dropout), 调整网络结构, 本文提出一种改进的神经网络架构, 见图 3。

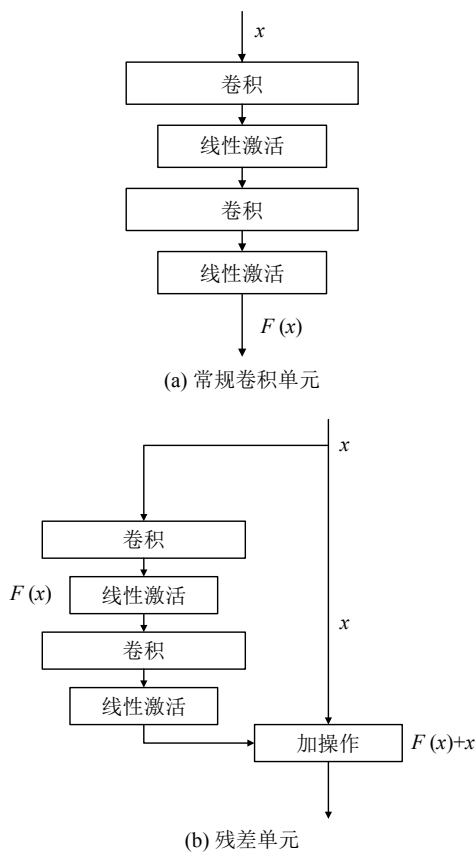


图2 两种结构对比

图3(a)为本文残差单元的基本结构,由两个 3×3 卷积层和一个从输入到输出的恒等映射组成,卷积操作之前加入批量归一化层(BN)和ReLU激活函数. BN层主要作用是把在网络传递中逐渐偏移的输入分布强制拉回到均值为0方差为1的比较标准的正态分布,使得非线性变换函数的输入值落入对输入比较敏感的区域,以此避免梯度消失问题.将激活函数放在卷积操作之前,并加入批量归一化层.经He等研究证明,这一残差单元结构目前是最优的,本文除输入端的残差单元有所差异外,其余部分均采用这一结构.残差单元与单元之间卷积核数量翻倍,具体见表1.

由图3(b)可见,本文神经网络算法中共有11个残差单元,其中有5个下采样层、5个上采样层和1个桥连接层.其中,上采样层中的上采样操作由上采样和一个带有线性激活函数的卷积组成,卷积参数在表1中显示.由于采用的残差单元增加,在桥连接层中加入随机失活,防止网络的过拟合和梯度消失等问题.上采样使用Upsampling函数,跳跃连接的方式为级联,使用Concatenate函数,将下采样和上采样过程中的特征信

息结合在一起,与U-Net的跳跃连接方式相同.在神经网络的输出端采用 1×1 的卷积核将最终级联的图像融合,并采用Sigmoid激活函数将输出值限定在0~1之间.

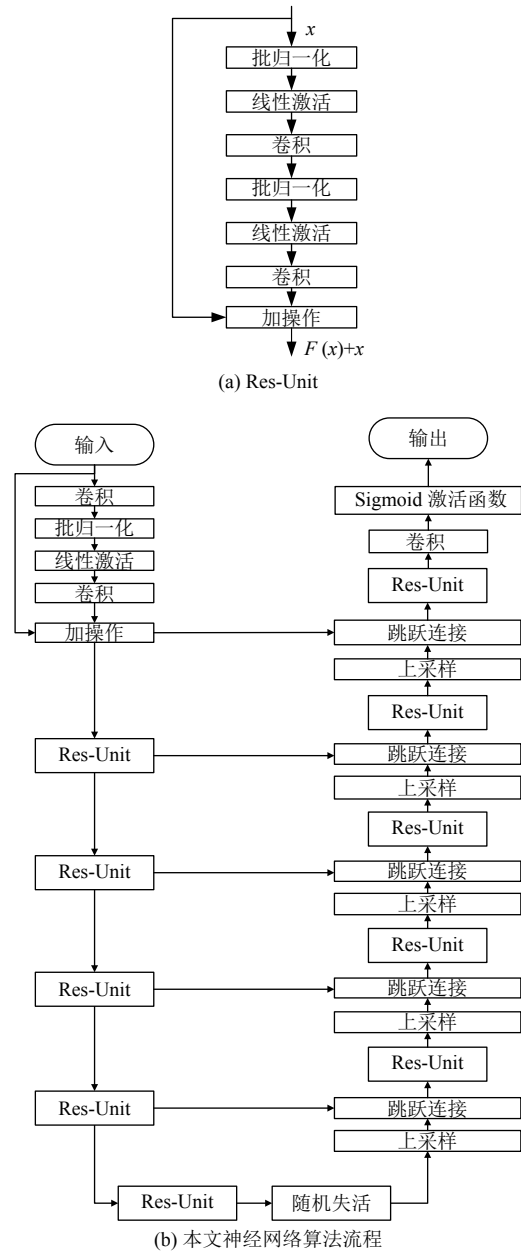


图3 本文神经网络架构

本文神经网络架构中,通过采用卷积步进为2的方式代替池化层的作用,将图像尺寸减半^[22].同时该网络卷积过程中使用了Padding操作保持卷积过程中的图像尺寸不变,取代了U-Net网络中的尺寸裁剪操作,可以保持输入输出图像尺寸不变.

表1 本文卷积网络层次

单元模块	卷积层次	卷积核	卷积步长	图像输出尺寸
输入	-	-	-	640×640×3
层次1	卷积1	3×3/64	1	640×640×64
	卷积2	3×3/64	1	640×640×64
层次2	卷积3	3×3/128	2	320×320×128
	卷积4	3×3/128	1	320×320×128
下采样	卷积5	3×3/256	2	160×160×256
	卷积6	3×3/256	1	160×160×256
层次4	卷积7	3×3/512	2	80×80×512
	卷积8	3×3/512	1	80×80×512
层次5	卷积9	3×3/1024	2	40×40×1024
	卷积10	3×3/1024	1	40×40×1024
桥连接	卷积11	3×3/2048	1	20×20×2048
	卷积12	3×3/2048	1	20×20×2048
层次7	卷积13	2×2/1024	1	40×40×1024
	卷积14	3×3/1024	1	40×40×1024
	卷积15	3×3/1024	1	40×40×1024
层次8	卷积16	2×2/512	1	80×80×512
	卷积17	3×3/512	1	80×80×512
	卷积18	3×3/512	1	80×80×512
上采样	卷积19	2×2/256	1	160×160×256
	卷积20	3×3/256	1	160×160×256
	卷积21	2×2/128	1	320×320×128
层次10	卷积22	3×3/128	1	320×320×128
	卷积23	3×3/128	1	320×320×128
	卷积24	2×2/64	1	640×640×64
层次11	卷积25	3×3/64	1	640×640×64
	卷积26	3×3/64	1	640×640×64
输出	卷积27	1×1/1	1	640×640×1

1.4 阈值滤波

由神经网络模型提取出来的图像像素值在 0~1 之间, 为提升图像提取效果, 滤除道路提取过程中的部分噪声. 经过测试, 自定义二值化阈值 0.94, 当像素值小于 0.94 时设为道路像素, 当像素值大于 0.94 时设为空白像素.

2 遥感图像提取实验

2.1 技术路线

藉由本文提出的网络算法模型, 本文设计总的技术路线见图 4.

神经网络通过训练样本训练网络权重参数实现对测试数据的预测, 假设有一组样本 (x_i, y_i) , x_i 为待训练图像, y_i 为训练标签, 样本像素预测值为 a_i , 样本个数为 N , 则通过不断减少预测值 a_i 与样本标签 y_i 的误差可以实现网络权重参数的训练, 本文采用二分类交叉

熵损失函数, 与 U-Net 网络相同, 为:

$$L = -\frac{1}{N} \sum_{i=1}^N (y_i \log(a_i) + (1 - y_i) \log(1 - a_i)) \quad (1)$$

本文选用 Adam 优化器来训练网络, 这是应用最为广泛的优化器, 训练过程网络参数收敛速度快, 收敛效果好. 优化器初始学习率设为 10^{-5} . 批处理样本大小设置为 2, 迭代期 (epoch) 设置为 100.

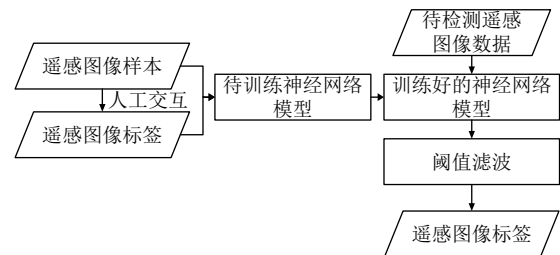


图4 本文技术路线

2.2 实验样本和平台

本次实验采用重庆及上海地区部分高分 2 号遥感影像作为训练数据, 由于单幅影像数据量过大, 现有的硬件设施无法满足处理条件, 故将重庆及上海影像分幅输出, 尺寸定为 640×640 像素, 采用人工交互的方式制作对应标签, 共有图像 1017 张, 其中训练集图像 912 张, 验证集图像 75 张, 测试集图像 30 张, 训练集, 验证集和测试集之间无重复影像, 保证数据独立性. 图像中包括山川、河流、隧道出入口、道路、大型桥梁、建筑、植被、车辆等对象, 具有较强的代表性. 本次实验采用 Windows 10 操作系统, 采用基于 Anaconda 软件的 Tensorflow-GPU, Python 3.6 开发环境. GPU 型号为 NVIDIA Quadro P5000, 其专用内存为 16 GB.

2.3 道路提取评价指标

目前图像分割领域评价分割效果的指标并不统一, 本文除了采用基本的评价指标——准确率外, 采用交并比 (Intersection over Union, IoU) 来作为主要评价指标. 准确率表示预测正确的像素占总像素的比例, 而交并比表示预测图像与标签图像的像素交集占预测图像与标签图像像素并集的百分比, 即预测图像与标签图像两者的重叠部分的像素占重叠与非重叠部分像素总和的百分比. 假设道路类像素值设为 a , 预测图像像素值 o_i 和标签图像像素值 y_i , 则

$$IoU(a) = \frac{\sum_i (o_i == a \cap y_i == a)}{\sum_i (o_i == a \cup y_i == a)} \quad (2)$$

其中, \cap (交集) 表示逻辑和操作, \cup (并集) 表示逻辑或操作. 本文通过统计像素 i 的值求交并比.

2.4 实验结果及比较

本文先使用 987 张图片对 U-Net 网络模型和 Res-Unet 网络模型和本文模型进行训练和验证, 之后从样本库中抽取不同于训练样本的 30 张遥感影像图片作为预

测样本, 图像尺寸同样为 640×640 像素, 使用本文训练好的网络模型对其进行道路提取并评估结果. 之后将分别网络模型替换为 U-net 网络模型和 Res-Unet 网络模型, 训练参数配置与本文保持一致, 使用训练好的模型对预测样本进行道路提取, 经过后处理阈值滤波后得到的道路提取效果见图 5. 道路提取效果评估指标见表 2.

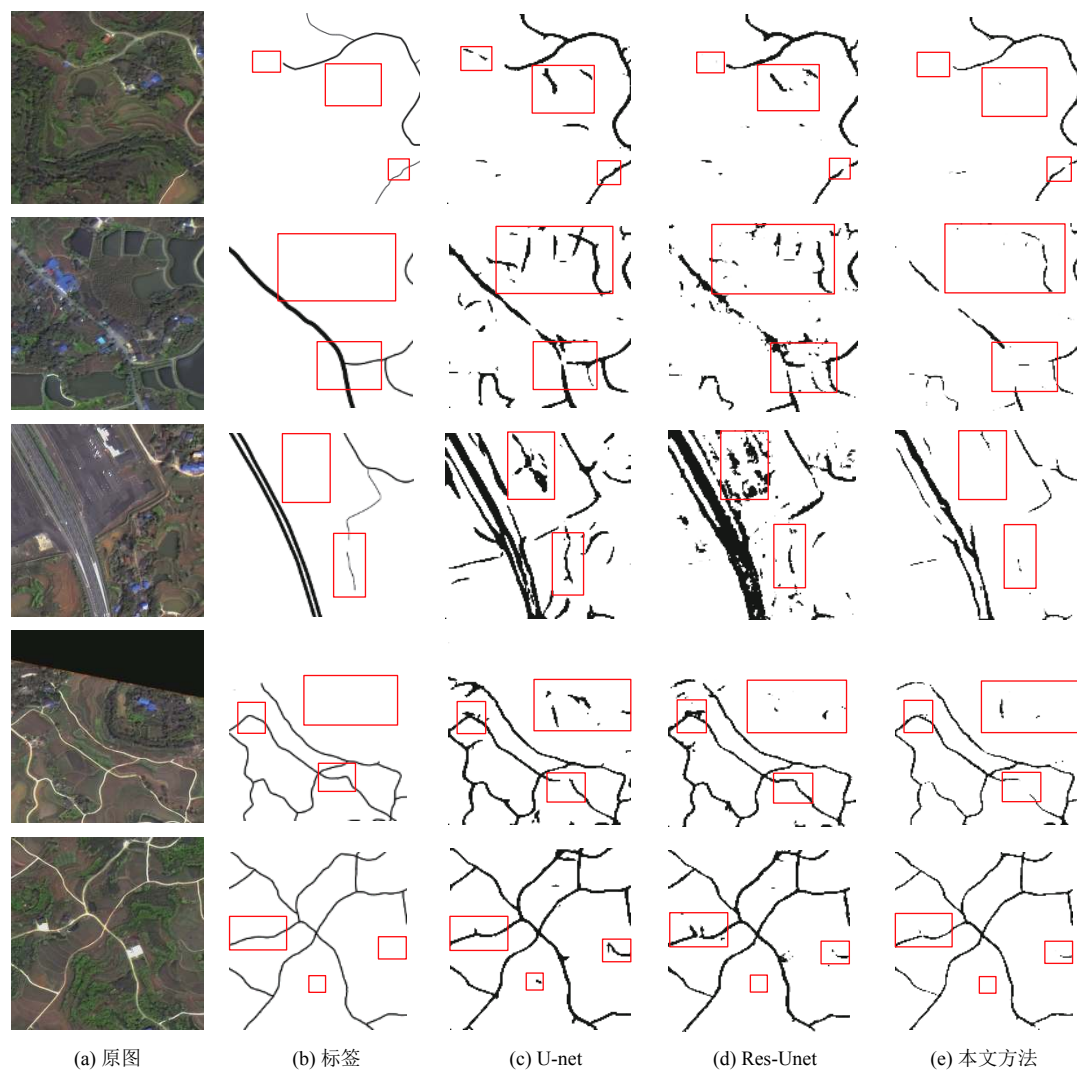


图 5 测试集道路提取效果图

表 2 道路提取效果评估指标

算法模型	准确率	交并比	测试时间(ms/步)
U-net	0.9441	0.9691	207
Res-Unet	0.9499	0.9710	236
本文算法	0.9502	0.9717	344

在图 5 中, 方框位置各种提取方法提取下的道路图像存在显著差异. 以道路标签为标准, U-net (图 5(c)) 提取的道路较 Res-Unet (图 5(d)) 和本文方法 (图 5(e))

出现许多多余的毛刺和非道路线段; Res-Unet 比之 U-net 增加了图像噪声, 使得提取的道路曲线表现粗糙; 本文方法比之 U-net 和 Res-Unet, 所提取道路的与标签相似度更高, 减少了大量毛刺、非道路线段和噪声, 但也存在道路线形中断的问题. 从表 2 的实验数据看出, U-net 的准确率和交并比均优于 Res-Unet 及本文方法, 但由于网络参数的增加, 也牺牲了的图像处理速度.

3 结论

本文通过对各种基于深度神经网络的遥感影像道路提取方法进行分析和实验, 提出深层残差 U-net 网络, 有效提取图像浅层信息的同时增强残差单元对图像细节的补充效果, 提高了图像的提取精度. 但是本文的方法也存在道路截断的问题, 存在改进的空间, 怎样进一步增强道路线形特征是下一步研究的重点. 另外, 整景遥感影像分幅会造成所提取道路的不一致性. 造成道路错位或缺失, 有待进一步的研究解决.

参考文献

- 曹云刚, 王志盼, 杨磊. 高分辨率遥感影像道路提取方法研究进展. 遥感技术与应用, 2017, 32(1): 20–26.
- 王峰萍, 王卫星, 薛柏玉, 等. GVF Snake 与显著特征相结合的高分辨率遥感图像道路提取. 测绘学报, 2017, 46(12): 1978–1985. [doi: [10.11947/j.AGCS.2017.20170393](https://doi.org/10.11947/j.AGCS.2017.20170393)]
- 连仁包, 王卫星, 李娟. 自适应圆形模板及显著图的高分辨率遥感图像道路提取. 测绘学报, 2018, 47(7): 950–958. [doi: [10.11947/j.AGCS.2018.20170596](https://doi.org/10.11947/j.AGCS.2018.20170596)]
- Zhao JQ, Yang J, Li PX, *et al.* Semi-automatic road extraction from SAR images using EKF and PF. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2015, XL-7/W4: 227–230. [doi: [10.5194/isprsarchives-XL-7-W4-227-2015](https://doi.org/10.5194/isprsarchives-XL-7-W4-227-2015)]
- Zang Y, Wang C, Cao LJ, *et al.* Road network extraction via aperiodic directional structure measurement. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(6): 3322–3335. [doi: [10.1109/TGRS.2016.2514602](https://doi.org/10.1109/TGRS.2016.2514602)]
- Chinnathevar S, Dharmar S. FPGA implementation of road network extraction using morphological operator. Image Analysis & Stereology, 2016, 35(2): 93–103.
- 李杰, 冯魁祥, 朱玲玲, 等. 基于阈值标记的分水岭算法遥感图像道路提取. 长春大学学报, 2019, 29(6): 10–14.
- 符喜优, 张风丽, 王国军, 等. 基于模糊连接度的高分辨率 SAR 图像道路自动提取. 计算机应用, 2015, 35(2): 523–527. [doi: [10.11772/j.issn.1001-9081.2015.02.0523](https://doi.org/10.11772/j.issn.1001-9081.2015.02.0523)]
- Kusumandari DE, Munandar A, Redhyka GG. The comparison of GVF snake active contour method and ellipse fit in optic disc detection for glaucoma diagnosis. Proceedings of 2015 International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT). Bandung, Indonesia. 2015. 123–126.
- Perciano T, Tupin F, Hirata Jr R, *et al.* A two-level Markov random field for road network extraction and its application with optical, SAR, and multitemporal data. International Journal of Remote Sensing, 2016, 37(16): 3584–3610. [doi: [10.1080/01431161.2016.1201227](https://doi.org/10.1080/01431161.2016.1201227)]
- 袁鹏飞, 黄荣刚, 胡平波, 等. 基于多光谱 LiDAR 数据的道路中心线提取. 地球信息科学学报, 2018, 20(4): 452–461. [doi: [10.12082/dqxxkx.2018.170634](https://doi.org/10.12082/dqxxkx.2018.170634)]
- 查中亮. 基于多源遥感数据的道路、居民点提取及布局优化研究 [硕士学位论文]. 成都: 四川师范大学, 2018.
- Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. 2015. 3431–3440.
- Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin, Germany. 2015. 234–241.
- Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481–2495. [doi: [10.1109/TPAMI.2016.2644615](https://doi.org/10.1109/TPAMI.2016.2644615)]
- Chen LC, Papandreou G, Kokkinos I, *et al.* DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834–848. [doi: [10.1109/TPAMI.2017.2699184](https://doi.org/10.1109/TPAMI.2017.2699184)]
- Chen LC, Zhu Y, Papandreou G, *et al.* Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany. 2018. 833–851.
- Mnih V, Hinton GE. Learning to detect roads in high-resolution aerial images. Proceedings of the 11th European Conference on Computer Vision. Heraklion. Crete, Greece. 2010. 210–223.
- Saito S, Yamashita T, Aoki Y. Multiple object extraction from aerial imagery with convolutional neural networks. Journal of Imaging Science and Technology, 2016, 60(1): 010402.
- Cheng GL, Wang Y, Xu SB, *et al.* Automatic road detection and centerline extraction via cascaded end-to-end convolutional neural network. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(6): 3322–3337. [doi: [10.1109/TGRS.2017.2669341](https://doi.org/10.1109/TGRS.2017.2669341)]
- Zhang ZX, Liu QJ, Wang YH. Road extraction by deep residual U-net. IEEE Geoscience and Remote Sensing Letters, 2018, 15(5): 749–753. [doi: [10.1109/LGRS.2018.2802944](https://doi.org/10.1109/LGRS.2018.2802944)]
- He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 770–778.