

文章编号: 2095-2163(2022)10-0031-06

中图分类号: TP391.1

文献标志码: A

基于注意力门残差网络的遥感影像道路提取

李文书, 李绅皓, 赵朋

(浙江理工大学 信息学院, 杭州 310018)

摘要: 针对遥感影像道路提取出现的无关噪声多, 道路不连续问题, 本文通过改进 U-Net 提出了基于注意力门残差网络的道路提取算法。首先, 编码器部分引入残差块传递原始特征, 在保证网络深度的同时, 使梯度能够有效传递; 其次, 在连接层使用多尺度空洞卷积特征提取模块, 来充分挖掘图像中的多尺度特征信息; 最后, 用注意力门将浅层网络信息和反卷积信息融合实现解码, 以抑制浅层噪声特征。使用的数据集包括 Massachusetts Roads Dataset 数据集和 CVPR DeepGlobe 2018 道路提取挑战赛数据集。实验结果表明, 该算法可以有效提升道路分割的效果。

关键词: 道路提取; 遥感影像; 残差网络; 门控卷积; U-Net

Road extraction from remote sensing images based on attention gate residual network

LI Wenshu, LI Shenhao, ZHAO Peng

(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

[Abstract] Aiming at the problem that there are many independent noises and discontinuity in road extraction from remote sensing images, a semantic segmentation algorithm of Residual Attention U-Net is proposed by improving U-Net. Firstly, the encoder introduces the original character of residual block transfer, which can guarantee the depth of the network and make the gradient transfer efficiently. Then, multiscale dilated convolution extraction module is used in the connection layer to fully mine the feature information in the image. Finally, the decoder uses the attention gate to fuse the shallow network information with the deconvolution information to suppress the shallow noise characteristics. The used datasets include the Massachusetts Roads Dataset and the CVPR DeepGlobe 2018 Road Extraction Challenge dataset. The experimental results show that the algorithm can effectively improve the effect of road segmentation.

[Key words] road extraction; remote sensing images; residual network; gated convolution; U-Net

0 引言

随着卫星遥感技术迅猛发展, 遥感影像的获取也日益便捷。遥感技术直接或间接运用于军事侦察、地球资源监测、城市道路规划、农业灾害预警、应急救援等多个领域。遥感图像具有分辨率高、覆盖广、获取便捷的优势, 已经成为目前更新地理信息数据的主要来源。道路是地理信息系统的重要组成部分, 在政治、经济、地理、军事等领域发挥着重大作用。在数字地图时代, 研究地图更新技术以保证地图内容的即时性、准确性和可靠性有着十分重要的意义。而遥感图像中道路信息易受建筑物阴影、车辆、植被等非道路信息干扰, 导致人工提取遥感图像中的道路费时费力。每天产生海量卫星遥感数据,

完全依靠人力进行道路提取显然是不可行的。因此, 亟需研究一种能自动便捷地从遥感影像中获取道路的方法。

当前的道路提取方法主要包括模板匹配、知识驱动、面向对象和深度学习等 4 类。Pudaruth^[1] 开发了一种多形状、多角度的模板匹配算法, 从 4 个方向提取道路, 最终将 4 幅独立的图像融合成一幅图像, 从而得到路网。Cao 等人^[2] 设计了一种基于均值漂移的道路中心匹配算法, 可以较好地提取道路中心线。Treash 等人^[3] 利用边缘检测器提取道路边缘, 认为平行的边缘极有可能为道路, 而后采用边缘链接算法提取道路。Guo 等人^[4] 使用美国地理调查局 (USGS) 的地理数据来消除与道路不相关的杂波, 继而使用均值-偏移聚类方法, 对卫星图像的道路

基金项目: 国家自然科学基金 (31771224, 61603228); 国家科技部重点研发计划重点专项课题 (2018YFB1004901); 浙江省自然科学基金 (LY17C090011, LGF19F020009)。

作者简介: 李文书 (1975-), 男, 博士, 教授, 主要研究方向: 图像处理、认知建模、虚拟实现、物联网集成开发; 李绅皓 (1994-), 男, 硕士研究生, 主要研究方向: 计算机视觉、数据挖掘; 赵朋 (1997-), 男, 硕士研究生, 主要研究方向: 计算机视觉。

通讯作者: 李文书 Email: Charlie@zstu.edu.cn

收稿日期: 2022-03-07

进行分割。Zhu 等人^[5]先用灰度形态学特征监测路网轮廓,然后利用线段匹配法监测基本路网,最后基于形态学方法对检测的基本路网进行处理。金静等人^[6]对遥感影像进行双边滤波,平滑细节信息并保留道路边缘信息,然后使用模糊 C 均值算法分割图像以得到独立的地物对象,最终提取出遥感影像中的道路。Mnih 等人^[7]首次将深度学习应用于遥感影像道路提取任务,并通过实验证明了深度学习对解决卫星图像道路提取任务的可行性。张永宏等人^[8]提出基于全卷积神经网络的多源高分辨率遥感道路提取算法,将遥感图像的 Canny 特征信息、RGB、Gray 和标签图放入全卷积神经网络(FCN)进行训练,在遥感影像道路提取这一任务上取得了不错的效果。Chen 等人^[9]提出 SegNet 语义分割网络,SegNet 网络首次采用编解码器架构,在解码器上采样时可以更好地保留图像的轮廓细节,在遥感影像道路分割上取得了较好效果。总之,模板匹配方法易受几何变换影响,参数也很敏感,对于复杂图像提取过后仍需要大量的人工处理,仅适用于场景简单的遥感图像。知识驱动的方法对所提知识和道路的匹配性有较高要求,不适合大范围推广。面向对象的道路提取方法虽然涌现出很多成果,但在其分割准确性上却面临严峻挑战。研究可知,深度学习通过在数据中学习模型,近年来已展现出在道路分割领域中的独特优势。

综合前述分析,研究提出了一种基于注意力门残差网络的遥感影像道路提取网络模型 RAU-Net (Residual Attention U-Net)。研究把残差块传递原始特征引入编码器,在保证网络深度的同时,也确保了梯度的有效传递。使用多尺度空洞卷积特征提取 (Multiscale dilated convolution extraction, MDCE) 模块,多尺度地挖掘图像中的特征信息。最后,用注意力门将浅层网络信息和反卷积信息融合实现解码,将浅层信息有选择地融入深层网络,抑制浅层噪声特征,进一步提高了道路分割的精确性。

1 网络结构

U-Net 网络模型分为编码器、解码器、链接器三部分。其中,编码器通过卷积神经网络提取特征,通过链接器传递给解码器,解码器通过卷积特征和浅层跳跃连接特征融合并上采样进行语义分割。对于道路提取任务,道路的狭长特性导致这样的方法仍无法解决道路细节丢失问题,同时跳跃连接会给提取结果带来与道路无关的噪点。本文网络模型主体

框架受 U-Net 模型启发,但对 U-Net 的构成做了大量调整,将 U-Net 编码器改为 ResNet34 编码器,加入了 MDCE 模块和融合注意力门的跳跃连接解码器,在此基础上进一步提出了 RAU-Net。RAU-Net 的组成如图 1 所示,这是一种端到端的网络结构,输入高分辨率卫星图像,最终输出道路提取的结果。

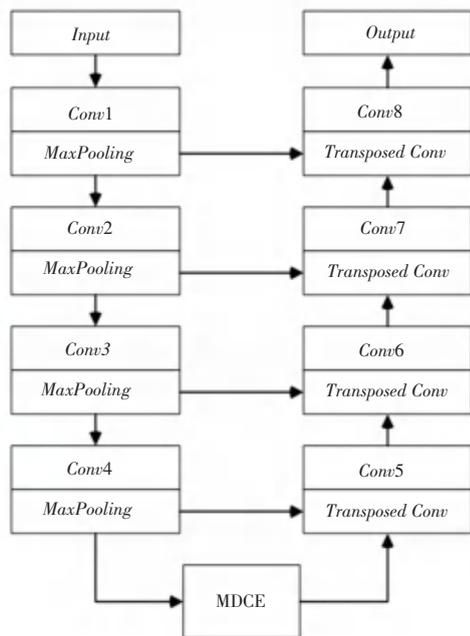


图 1 RAU-Net

Fig. 1 RAU-Net

1.1 ResNet34 编码器

RAU-Net 编码器部分不再采用 U-Net 卷积编码器,而是用 ResNet34 作为编码器。理论上,网络模型越深,能提取的语义信息越多。为了尽可能提取出更多的道路细节,就需要加深 U-Net 的编码器部分网络。但一味地加深网络不仅会使模型参数过多、难以训练,还会造成梯度不稳定、出现梯度消失^[10]或者梯度爆炸^[11]问题。针对这种情况,He 等人^[12]首次提出了残差网络。残差网络是由一系列残差块组成。一个残差块的设计如图 2 所示。

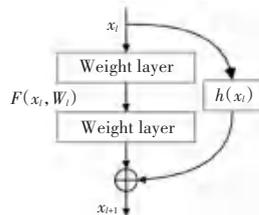


图 2 残差块示意图

Fig. 2 Residual block diagram

单独一个残差块可以表示为:

$$y_l = h(x_l) + F(x_l, W_l) \quad (1)$$

$$x_{l+1} = f(y_l) \quad (2)$$

其中, $h(\cdot)$ 表示直接映射; $f(\cdot)$ 是激活函数; x_i 为原始输入; x_{l+1} 为残差块输出, 也是下一个残差块的输入。本文使用 ResNet34 网络作为模型的编码器结构, 替换掉原来 U-Net 的普通卷积编码器。

1.2 MDCE 模块

为了适应道路提取任务的特殊性, 使网络能学习到不同尺度的道路特征信息, 文中在模型连接层引入多尺度卷积提取模块 MDCE (Multiscale dilated convolution extraction)。多尺度卷积提取模块的设计结构如图 3 所示。由图 3 可知, MDCE 采用 4 个扩张率不同的空洞卷积模块, 空洞卷积膨胀率分别设置为 1、2、4、8。不同的膨胀率, 可以保证充分提取多尺度的特征信息。最终 4 种膨胀率空洞卷积提取的结果并联入 3×3 的融合卷积层, 提升了网络捕捉多尺度语义信息的能力。

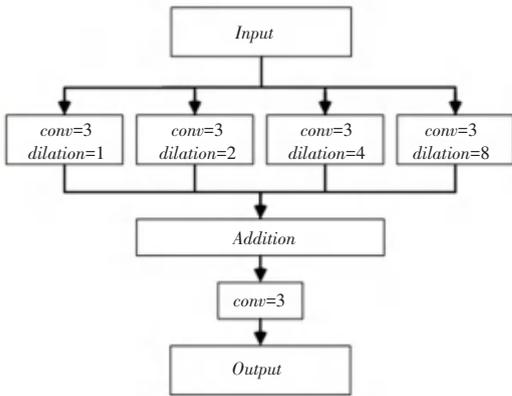


图 3 多尺度卷积提取模块

Fig. 3 Multiscale dilated convolution extraction model

1.3 融合注意力门的跳跃连接解码器

随着编码器编码、语义信息增多, 特征图尺寸会不断减少, 但是本文的提取任务却需要得到一张与原分辨率相同的图像。由图 1 可知, 本文的网络输入图像是 $1\ 024 \times 1\ 024 \times 3$, 经过编码器和 MDCE 模块后, 数据维度变为 $32 \times 32 \times 512$ 。为了达到提取效果, 研究提出了一种结合注意力门的解码器结构, 能够将编码器浅层的语义信息, 与 MDCE 反卷积后的结果相融合, 同时将特征图加以扩张, 从而将图像转换为 $1\ 024 \times 1\ 024 \times 1$ 维度。

由于道路提取任务的特殊性, 研究的目标结果是细长的道路, 而在经过编码器编码后和道路相关的一些语义信息已经缺失, 后续无法复原, 导致最终解码后的提取图片道路出现断裂情况。而浅层特征图中并非所有的信息对提取都有帮助, 浅层特征图

中非道路信息的噪声数据过多, 如果类似 U-Net 直接做跳跃链接, 提取到的道路图像会出现很多非道路信息。本次研究在跳跃连接之前, 加入了注意力门 (Attention Gate)。注意力门会抑制浅层特征图中的无关特征响应, 在增加极少的额外计算量的同时, 带来模型准确性的显著提高。注意力门结构如图 4 所示。

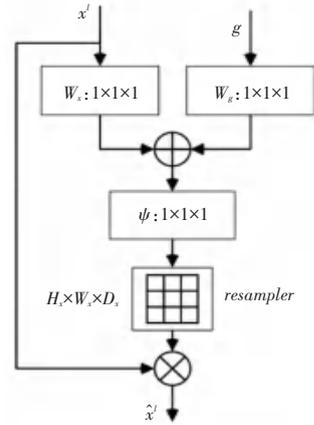


图 4 注意力门

Fig. 4 Attention gate

由图 4 可知, g 是来自浅层网络的特征图, x^l 是网络上一层上采样后的特征图, 其运算结果与 g 进行融合。

2 实验数据与损失函数

为了验证提出模型的有效性和泛化能力, 实验采用了马萨诸塞州道路公开数据集 (Massachusetts Roads Dataset) 和 CVPR DeepGlobe 2018 数据集。在深度学习中, 训练样本数量不足很容易导致过拟合, 并不能真实反映网络的有效提取能力, 现对数据集进行增强。

实验最初道路数据集共有 1 172 组卫星影像, 每组影像含有一张卫星影像和一张道路标签, 每个像素对应的真实地理空间大小为 1 m。单张图像尺寸为 1 500 像素 \times 1 500 像素。实验中, 分析发现此数据集的一些影像标注并不准确, 经过筛选后剔除了标注不准确的图像。选出标注准确的图像后, 共有 826 组。显而易见, 本数据集中的 826 组影像并不足以进行训练, 为了增加数据集中的数据量, 还要对数据集进行扩充。本文采用镜像翻转、旋转翻转等方法, 变换后共得到 3 200 组影像。最终, 将数据集中的 2 900 组影像用于训练, 300 组影像用于测试, 300 组影像用于验证。数据扩充示例如图 5 所示。

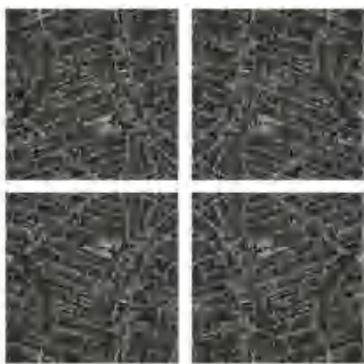


图5 数据扩充示例

Fig. 5 Data expansion examples

CVPR DeepGlobe 2018 道路提取挑战赛数据集共 6 226 组卫星影像,单张图像尺寸为 1 024 像素 * 1 024 像素,每个像素对应的真实地理空间大小为 0.5 m。同时提供了验证集和测试集,共计 2 344 张卫星图像,这些图像并未提供道路标签。为验证本算法的有效性,将包含道路标签的 6 226 组卫星影像重新划分出测试集和验证集。至此,数据集中有 5 000 组影像用于训练,613 组影像用于测试,613 组影像用于验证。

Dice 系数是一种集合相似度度量函数,用于计算 2 个样本的相似度,取值范围是 0~1。数值越大,说明 2 个集合相交元素越多,彼此之间越相似。*Dice* 损失由 Milletari 等人^[13]引入图像分割领域,在样本不均衡时表现优于传统的交叉熵损失。*Dice* 系数表示和 *Dice* 损失的数学定义分别如下:

$$Dice(\hat{y}, y) = \frac{2 \sum_{i=1}^m y^{(i)} \hat{y}^{(i)}}{\sum_{i=1}^m (y^{(i)} \hat{y}^{(i)} + \hat{y}^{(i)} \hat{y}^{(i)})} \quad (3)$$

$$DiceLoss(\hat{y}, y) = 1 - Dice(\hat{y}, y) \quad (4)$$

3 实验与分析

3.1 实验环境

为验证提出 RAU-Net 网络在道路提取任务中的有效性,基于 Tensorflow2.3.0 平台,采用 Ubuntu 18.04 LTS 64 位操作系统,搭载 Intel E5-2620 CPU, NVIDIA GTX2080Ti GPU,显存 11 G。相关参数见表 1。

表 1 参数设置表

Tab. 1 Parameters setting table

参数名称	参数值
学习率	0.001
批尺寸	2
训练次数	300
损失函数	<i>Dice</i> 损失
优化器	<i>Adam</i>

3.2 评价指标

卫星影像道路任务是语义分割任务的一种,本实验采用语义分割通用的评估方法,即准确率 (*Precision*)、召回率 (*Recall*) 以及 F_1 -score。其中,准确率和召回率分别反映正确提取道路占预测样本比例和道路提取完整度, F_1 -score 为准确率和召回率的调和均值。对应数值可由如下公式计算得出:

$$pre = \frac{TP}{TP + FP} \quad (5)$$

$$rec = \frac{TP}{TP + FN} \quad (6)$$

$$F_1 - score = \frac{2 \cdot pre \cdot rec}{pre + rec} \quad (7)$$

3.3 实验结果分析

基于本文所增强的实验数据集,分别训练 U-Net、MultiResUNet^[14]、RAU-Net,并通过准确率、召回率、 F_1 -score 进行定量分析。分析结果见表 2 和表 3。由表 2 得知,在同等实验条件和数据集的情况下,RAU-Net 在各项指标上均有一定提升,在道路提取研究上也有着优异表现。

表 2 不同模型在 Massachusetts Roads Dataset 的结果

Tab. 2 Results of different models at Massachusetts Roads Dataset

Model	<i>Prec</i>	<i>Rec</i>	F_1
U-Net	0.721	0.726	0.723
MultiResUNet	0.739	0.765	0.752
RAU-Net	0.746	0.793	0.769

表 3 不同模型在 CVPR DeepGlobe 2018 Roads Dataset 的结果

Tab. 3 Results of different models at CVPR DeepGlobe 2018 Roads Dataset

Model	<i>Prec</i>	<i>Rec</i>	F_1
U-Net	0.730	0.764	0.747
MultiResUNet	0.723	0.779	0.749
RAU-Net	0.748	0.787	0.767

图 6 展示了各模型在 Massachusetts Roads Dataset 和 CVPR DeepGlobe 2018 的道路提取结果。图 6 中,前 3 张为 Massachusetts Roads Dataset 中的图片,后 2 张为 CVPR DeepGlobe 2018 中的图片。由图 6 可以看出,虽然遥感影像地表复杂,受建筑物、河流、车辆影响,但是 RAU-Net 都较好地提取出了主干道路。U-Net 由于直接将浅层网络提取信息融合到后层网络,导致生成结果中有很多不连通的噪点。虽然 MultiResUNet 采用 res-path 取代了 U-Net 中的跳跃链接,但其预测结果却仍然存在错位和缺失的问题。RAU-Net 通过添加 MDCE 模块

充分利用不同维度信息,并且加入了融合注意力的跳跃连接解码器,消除提取结果中道路易出现断裂

问题,同时抑制了结果中易出现的噪声问题,提高了模型精度。

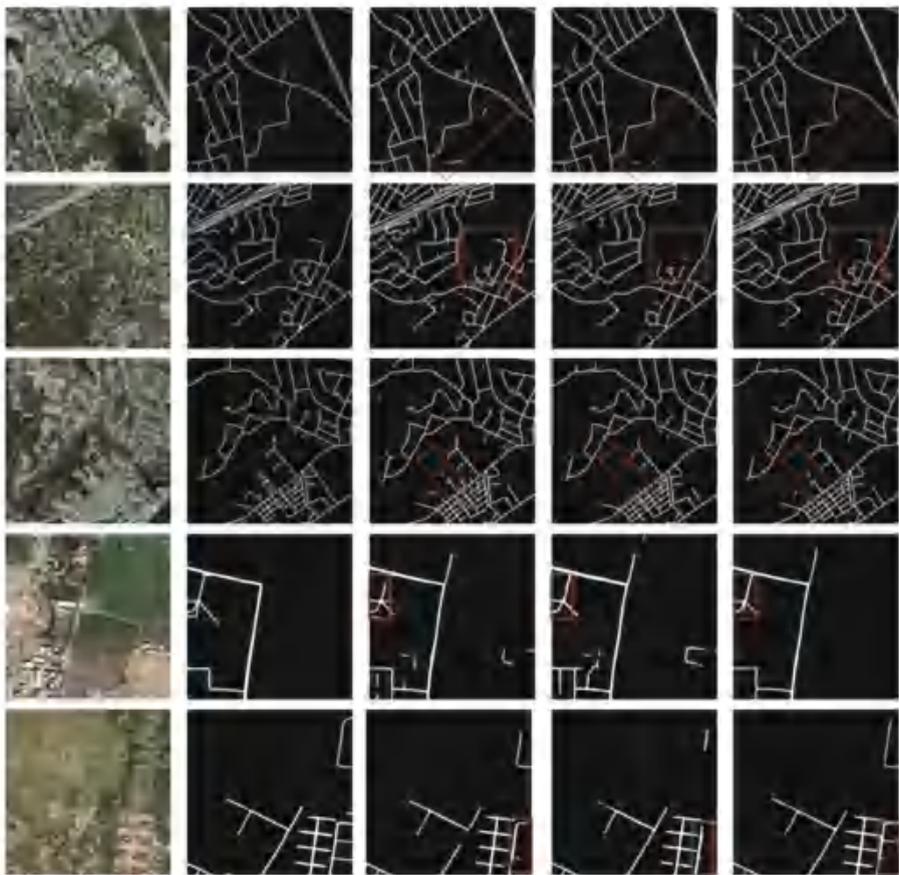


图 6 可视化结果

Fig. 6 Viewable results

4 结束语

将遥感影像分割为道路和非道路,属图像解译问题,随着深度学习的流行,图像解译任务也取得了长足进步。本文提出了一种 RAU-Net 网络进行语义分割的方法,在 3 项关键的分割效果量化指标,即准确率(Precision),召回率(Recall)以及 F_1 -score 上均优于经典提取算法,取得先进结果。RAU-Net 受到 Unet 网络架构和残差网络、注意力机制、空洞卷积等多种网络结构启发,在卫星图像道路提取任务中具有重要的理论和实用意义。后续研究中,尝试将单纯的道路提取任务拓展为更多地物,多种地物同时提取,相辅相成,以进一步提升该方法准确性。

参考文献

[1] PUDARUTH S. Extraction of roads from remotely sensed images using a multi-angled template matching technique [C]// Proceedings of the Third International Symposium on Computer Vision and the Internet. Jaipur, Japan: ACM, 2016: 21-29.

[2] CAO Fanzhi, ZHU Shulong, ZHU Baoshan, et al. Tracking road centerlines from remotely sensed imagery using mean shift and Kalman filtering [J]. Acta Geodaetica et Cartographica Sinica, 2016, 45(2): 205-212, 223.

[3] TREASH K, AMARATUNGA K. Automatic road detection in grayscale aerial images [J]. Journal of Computing in Civil Engineering, 2000, 14(1): 60-69.

[4] GUO D, WEEKS A, KLEE H. Robust approach for suburban road segmentation in high-resolution aerial images [J]. International Journal of Remote Sensing, 2007, 28(2): 307-318.

[5] ZHU C, SHI W, PESARESI M, et al. The recognition of road network from high-resolution satellite remotely sensed data using image morphological characteristics [J]. International Journal of Remote Sensing, 2005, 26(24): 5493-5508.

[6] 金静, 党建武, 王阳萍, 等. 面向对象的高分辨率遥感影像道路提取算法[J]. 兰州交通大学学报, 2017, 36(01): 57-61.

[7] MNIH V, Geoffrey E. Hinton. Learning to detect roads in high-resolution aerial images [C]// DANIILIDIS K, MARAGOS P, PARAGIOSaragios N. Computer Vision - ECCV 2010. ECCV 2010. Lecture Notes in Computer Science. Berlin/Heidelberg: Springer, , 2010, 6316: 210-223.

[8] 张永宏, 夏广浩, 阚希, 等. 基于全卷积神经网络的多源高分辨率遥感道路提取[J]. 计算机应用, 2018, 38(07): 2072-2075.