



基于深度学习的卫星图像道路分割算法

张新华¹, 黄梦醒², 张雨¹, 李玉春², 单怡晴¹, 冯思玲²

(1. 海南大学 计算机与网络空间安全学院, 海口 570228; 2. 海南大学 信息与通信工程学院, 海口 570228)

摘要: 针对道路分割时存在的梯度消失问题, 构建基于U-Net的卫星道路图像语义分割模型。通过密集连接模块减少梯度消失, 并引入空间空洞金字塔结构保留更多的图像特征, 在学习深层次特征信息时采用注意力监督机制, 提取道路要素的特征信息。在卫星图像道路数据集上的测试结果表明, 与FCN、SegNet、U-Net算法相比, 该算法模型的准确率、召回率和精确率指标分别达到96.3%、96.9%和96.6%, 能够有效地对道路元素进行准确分割。

关键词: 深度学习; 道路分割; 密集连接模块; 空间空洞金字塔结构; 注意力监督机制

开放科学(资源服务)标志码(OSID):



中文引用格式: 张新华, 黄梦醒, 张雨, 等. 基于深度学习的卫星图像道路分割算法[J]. 计算机工程, 2021, 47(10): 306-313.
英文引用格式: ZHANG X H, HUANG M X, ZHANG Y, et al. Road segmentation algorithm for satellite images based on deep learning[J]. Computer Engineering, 2021, 47(10): 306-313.

Road Segmentation Algorithm for Satellite Images Based on Deep Learning

ZHANG Xinhua¹, HUANG Mengxing², ZHANG Yu¹, LI Yuchun², SHAN Yiqing¹, FENG Siling²

(1. School of Computer Science and Cyberspace Security, Hainan University, Haikou 570228, China;

2. School of Information and Communication Engineering, Hainan University, Haikou 570228, China)

[Abstract] The existing road segmentation methods generally suffer from gradient disappearance, and are limited in feature utilization and semantic segmentation accuracy. To address the problem, a U-Net-based semantic segmentation model is proposed for satellite road images. The dense connection module is used to reduce gradient disappearance. Then the Atrous Spatial Pyramid Pooling (ASPP) structure is introduced to retain more image features. Finally, the attention monitoring mechanism is used when learning deep-level feature information, so as to more effectively extract the feature information of road elements. The test results on a road data set from satellite images show that compared with FCN, SegNet and U-Net algorithms, the proposed algorithm improves the accuracy to 96.3%, recall rate to 96.9% and precision to 96.6%. This algorithm can segment the road elements accurately.

[Key words] deep learning; road segmentation; dense connection module; Atrous Spatial Pyramid Pooling (ASPP); attention monitoring mechanism

DOI: 10.19678/j.issn.1000-3428.0059312

0 概述

道路是影响自动驾驶技术的重要因素, 道路的精确度影响地图的准确性, 但是由于社会的快速发展, 道路规划的频率也越来越快^[1-3]。目前获取道路的方式有: 使用车载激光扫描设备, 通过遍历扫描信息的方式获取路面信息, 但是人力物力耗资巨大; 使用航空或者卫星图像, 通过技术方法获取道路信息, 能够节省大量的人力和时间^[4-6]。

近年来, 深度学习在计算机视觉、图像分类及图像分割等领域取得重大的成果。深度学习特殊的网络结构可以通过神经元将提取的特征值传递, 每一层对上一层传递的特征再进行提取学习继续传递, 从而提取最优的特征值^[7-9]。在图像语义分割领域的方法包括AlexNet、FCN、U-Net、SegNet以及ResNet。AlexNet实现了在卷积神经网络(Convolutional Neural Network, CNN)中成功应用ReLU、Dropout和LRN(Local Response Normalization)等Trick^[10]; FCN将

基金项目: 国家重点研发计划(2018YFB1703403, 2018YFB1404400); 海南省重点研发计划(ZDYF2019020); 海南省高等学校科学研究项目(Hnky2019-22)。

作者简介: 张新华(1995—), 女, 硕士研究生, 主研方向为计算机视觉、图像处理; 黄梦醒(通信作者), 教授、博士; 张雨, 副教授、博士; 李玉春, 博士研究生; 单怡晴, 硕士研究生; 冯思玲, 副教授、博士。

收稿日期: 2020-08-20 **修回日期:** 2020-10-27 **E-mail:** 993958129@qq.com

AlexNet、VGG 全连接层改为卷积层,FCN^[11]下采样对图像进行特征提取,对上采样进行反卷积,保证输出图像和输入图像大小相同;NOH^[12]等在FCN网络架构的基础上,提出学习一个多层反卷积网络代替简单的双线性插值;BADRINARAYANAA等^[13]提出SegNet基于编码器-解码器架构,卷积层和池化层组成了编码器,卷积层和上采样层组成了解码器。编码器的作用是提取特征图像,解码器的作用是将特征图像返回到和输入图像大小一致。

SegNet与反卷积网络相同点是网络结构相似,不同点是SegNet去除了网络中间的2个全连接层,并采用Batch Normalization方法和Softmax分类器;SegNet的优点是效率高、消耗内存少,缺点是精度不高;FCN和SegNet网络中引入池化层,优点是在减小图像大小的同时增大感受野,缺点是部分位置信息丢失。

YU等^[14]设计一种专用于图像像素预测的网络,该网络不包含池化层,并且卷积层采用扩张卷积。该网络的优点是扩张卷积增大卷积核的感受野,可以将捕获的图像多尺度上下文信息融合,提高像素预测的准确率。

ResNet^[15-16]残差网络内部的残差块使用跳跃连接,优点是减少随着神经网络深度增加引起的梯度消失问题。RONNEBERGER等^[17]提出的U-Net网络,利用编码器与解码器之间的级联操作将图像高层信息与浅层信息融合,优点是避免高层语义信息的丢失,尽可能性地保留图像特征。深度学习可以通过计算机实现目标对象的分割和提取,并且通过特殊的网络模型学习较低层次和较高层次的特征,具有较高的学习效率^[18]。深度学习对图像特征要素的分割,可以通过大量的数据实验取得较好的实验效果。因此,可利用深度学习算法对卫星图像道路分割提取特征图。

针对道路分割时存在混淆和识别不清等问题,本文在U-Net基础上,提出一种基于密集连接块^[19]、ASPP^[20]和CBAM^[21]的卫星道路图像语义分割算法。在融合图像特征信息时通过注意力监督机制输入图像特定部分,以提高卫星图像道路分割的精确度。

1 基于U-Net结构的道路语义分割模型

考虑到U-Net网络在图像语义分割和密集连接网络方面的优越性,本文构建基于U-Net结构的道路语义分割模型。其网络结构由编码器、解码器、密集连接模块、ASPP和CBAM组成。

1)编码器部分。包括密集连接模块和最大池化层。密集连接模块通过卷积层来提取图像的语义特征,最大池化层对图像的特征信息进行下采样操作。

2)密集连接模块。对上一层的图像特征再利用;引入ASPP增加图像特征信息的感受野,提高图像特征的鲁棒性;引入CBAM在学习深层次的特征信息时进行注意力监督,可有效提取道路要素的特征信息。

3)解码器部分。包括密集连接模块和反卷积

层。反卷积层对图像特征信息进行上采样操作,使得输入图像和输出图像大小不变。

1.1 算法流程

算法实现过程包括网络设计、网络搭建、获取数据、对数据进行预处理和扩大数据^[22]。算法流程如图1所示。

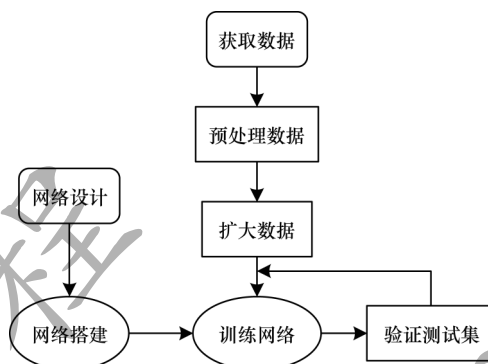


图1 本文算法流程

Fig.1 Procedure of this proposed algorithm

1.2 网络结构

本文网络结构在原始U-Net网络基础上引入密集连接模块、ASPP和CBAM模块。其网络结构如图2所示。

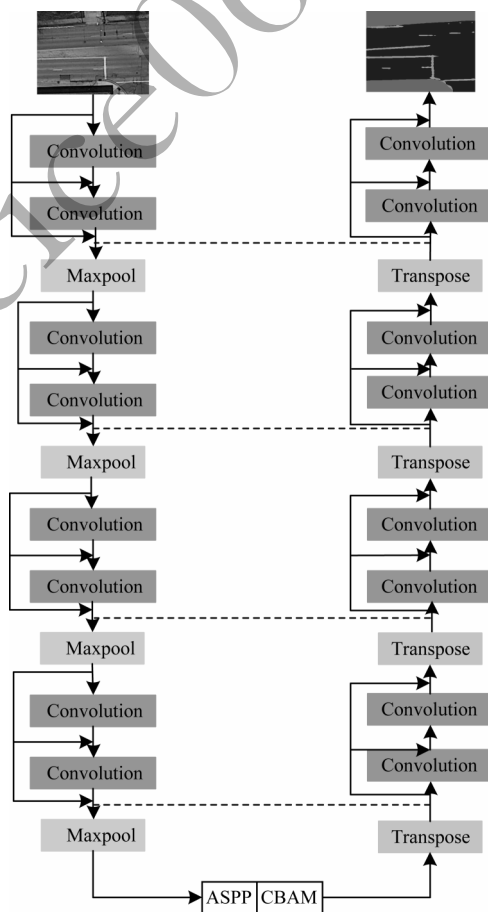


图2 本文模型网络结构

Fig.2 Network structure of proposed model

输入图像大小为 480×480 , 密集连接模块包括 2 次卷积和 2 次特征融合, 图像大小保持不变为 480×480 。经过最大池化层, 图像大小为原来的 $1/2$, 即 240×240 。将密集连接模块和最大池化操作记为 1 个操作单元。然后进行 3 次操作单元, 图像大小变成 30×30 的特征图。在对特征图进行反卷积(上采样)前引入 ASPP 和 CBAM, ASPP 可以融合更深层次的图像细节信息; CBAM 不影响特征图的大小, 网络学习将通道和空间注意力模型的特征图和权重图点乘融合, 然后将融合后结果图输入到反卷积层(上采样)中。在上采样时进行反卷积, 图像大小扩大为 2 倍, 即 60×60 , 再经过密集连接模块, 图像大小仍为 60×60 。最后进行 3 次操作单元(上采样与密集连接模块), 图像大小恢复为 480×480 。本文算法模型中各层的参数设置如表 1 所示。

表 1 网络各层的参数设置

Table 1 Parameter settings of each layer of the network

模块类型	特征图尺寸	尺寸/步长
输入	480×480	—
密集连接模块	480×480	$[3 \times 3 \text{ Conv-64}] \times 2$
最大池化层	240×240	$2 \times 2/2$
密集连接模块	240×240	$[3 \times 3 \text{ Conv-128}] \times 2$
最大池化层	120×120	$2 \times 2/2$
密集连接模块	120×120	$[3 \times 3 \text{ Conv-256}] \times 2$
最大池化层	60×60	$2 \times 2/2$
密集连接模块	60×60	$[3 \times 3 \text{ Conv-512}] \times 2$
最大池化层	30×30	$2 \times 2/2$
ASPP	30×30	$[1 \times 1 \text{ Conv-256}] \times 2$
CBAM	30×30	$[3 \times 3 \text{ Conv-512}] \times 2$
反卷积层	60×60	$2 \times 2/2$
密集连接模块	60×60	$[3 \times 3 \text{ Conv-512}] \times 2$
反卷积层	120×120	$2 \times 2/2$
密集连接模块	120×120	$[3 \times 3 \text{ Conv-256}] \times 2$
反卷积层	240×240	$2 \times 2/2$
密集连接模块	240×240	$[3 \times 3 \text{ Conv-128}] \times 2$
反卷积层	480×480	$2 \times 2/2$
密集连接模块	480×480	$[3 \times 3 \text{ Conv-64}] \times 2$

1.2.1 密集连接

在深度学习网络中, 随着网络深度的加深, 梯度消失问题会愈加明显, 本文参考 DenseNet 中密集连接的概念, 引入密集连接模块。在保证网络中层与层之间最大程度信息传输的前提下, 将所有层连接起来。为能够保证前馈的特性, 每一层将之前所有层的输入进行拼接, 之后将输出的特征图传递给之后的所有层。

操作过程为: 输入图像大小为 480×480 , 密集连接模块包括 2 次卷积和 2 次特征融合, 卷积核的大小为 3×3 , 步长为 2, 卷积核的个数为 64; 经过 4 次密集连接模块和池化操作后, 得到 30×30 大小的特征图。密集连接模块如图 3 所示。

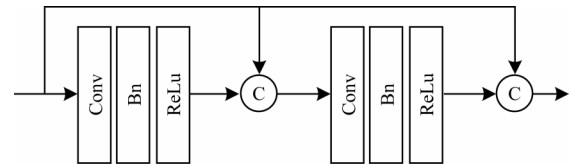


图 3 密集连接模块

Fig.3 Densely connected modules

每个卷积层与激活函数之间添加批量归一化层 (Batch Normalization, BN), 即在每次随机梯度下降时对每个批处理的数据做归一化操作, 使得输出特征图中每个通道数据的均值为 0, 方差为 1, 减少了网络学习过程中梯度消失的问题, 引入 ReLU 激活层对特征进行激活。

1.2.2 空洞空间金字塔结构

在上采样前引入带有空洞卷积的空间金字塔结构 (ASPP)。空洞空间金字塔结构基本要素是不同扩张率的空洞卷积。空洞卷积的优点是: 在不增加参数的情况下, 增大卷积核的感受野, 在下采样不损失信息的情况下, 卷积输出范围更大的空间信息使得网络保留了更多的图像特征。本文的空洞卷积的卷积核为 3×3 , 扩张率为 2、4、8、16 的空洞卷积替换了原来的普通卷积, 使得该模型分割道路图像更加清晰, 如图 4 所示。

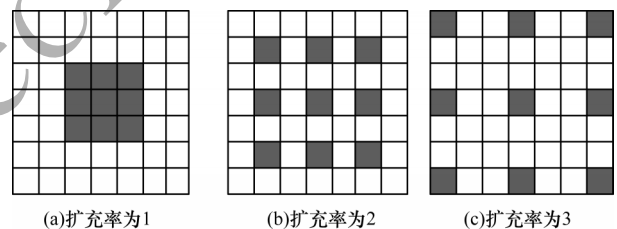


图 4 空洞卷积模型

Fig.4 Dilated convolution model

ASPP 对输入的特征图进行 6 次卷积操作。第 1 次卷积: 卷积核大小为 1×1 ; 第 2 次~第 5 次卷积为: 卷积核大小为 3×3 , 扩张率为 2、4、8、16 的空洞卷积替换了原来的普通卷积, 获得多尺度特征信息。引入平均池化, 对输入的特征图进行全局均值池化后, 将特征图送入大小为 1×1 卷积核, 使用 BN 操作和上采样到图像原始大小。然后将融合 6 个多尺度的特征图送入大小为 1×1 卷积层中, 最后将输出特征图送入到注意力机制模型中。ASPP 网络结构如图 5 所示。

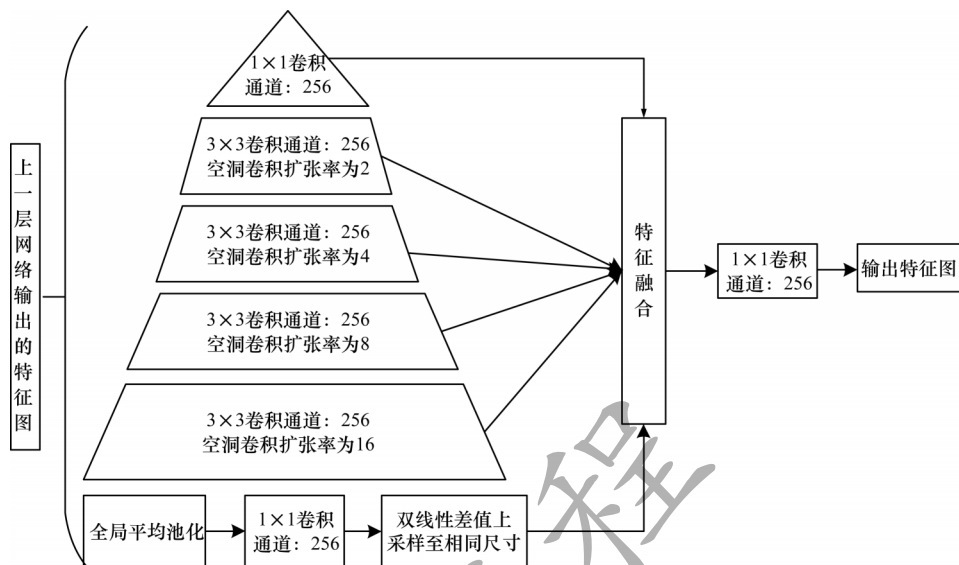


图5 ASPP网络结构

Fig.5 ASPP network structure

1.2.3 注意力机制模块

在ASPP后及上采样前引入CBAM, CBAM模块(包括通道注意力和空间注意力2个子模块)如图6、图7和图8所示。

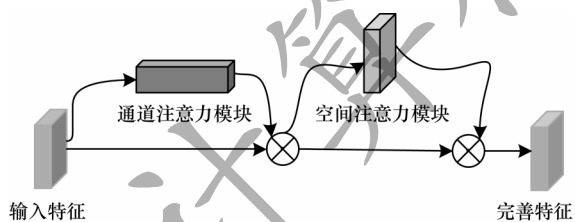


图6 注意力模型总体结构

Fig.6 Overall structure of attention model

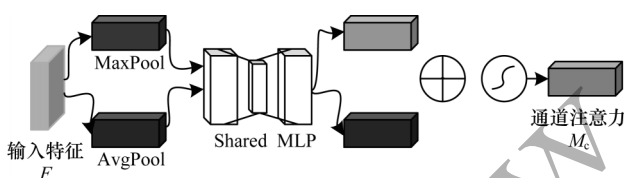


图7 通道注意力子图

Fig.7 Channel attention subgraph

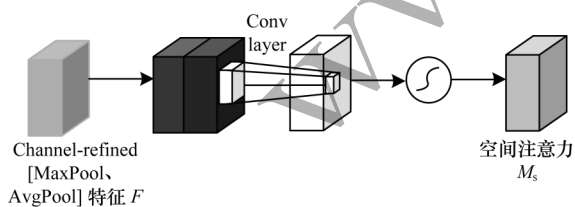


图8 空间注意力子图

Fig.8 Spatial attention subgraph

输入特征图经过通道注意力子模块和空间注意力子模块后,将2个子模块输出结果相乘后的特征图送入编码阶段进行上采样。CBAM通过空间和通

道学习更深层次的特征信息时进行注意力监督,可以有效提取道路要素的特征信息,其实现过程如下:

$$F' = M_c(F) F \otimes F'' = M_s(F') F' \quad (1)$$

其中: F 经过通道注意力模块和空间注意力模块操作后的特征图分别为 F' 和 F'' ; \otimes 表示逐元素相乘。

通道注意力模块实现过程为:对输入特征图 F 进行全局平均池化(AvgPool)和全局最大池化(MaxPool),结果分别为 a_1 、 b_1 ,然后将 a_1 、 b_1 通过多层感知机(Multi-Layer Perceptron, MLP)得到特征要素 A 、 B 相加得到 C_1 ;通道注意力特征是 C_1 与 F 融合的结果。运算过程如下:

$$\begin{aligned} M_1(F') &= (\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \\ M_c(F') &= (W_1(W_0(a_1)) + W_1(W_0(b_1))) M_c(F') = \\ &= (A + B) \\ M_c(F') &= (C_1) \end{aligned} \quad (2)$$

其中: W_0 和 W_1 为MLP的两层参数; σ 为sigmoid激活函数; W_0 需要使用Relu函数激活。

空间注意力模块实现过程为:对输入特征图进行平均池化和最大池化的结果为 a_2 、 b_2 ,然后将 a_2 和 b_2 融合进行 3×3 卷积操作得到特征图 C_2 ;空间注意力特征是 C_2 与 F 融合的结果。运算过程如下:

$$\begin{aligned} M_s(F) &= (f([\text{AvgPool}(F); \text{MaxPool}(F)])) M_s(F) = \\ &= (f(a_2; b_2)) \\ M_s(F) &= (C_2) \end{aligned} \quad (3)$$

其中: f 表示 3×3 卷积操作。

1.3 坐标点数据

本文对获取的卫星道路图像信息优化,淘汰不存在道路要素的图像。OpenStreetMap首先根据经纬度截取道路图像的坐标值,然后根据端点坐标数据定位

在 Mapbox 中相应的位置,使用 Labelme 对数据进行标注。为解决获取的图像尺寸不同、图像内容不清晰、数据量较大、训练速度慢等问题,将截取的图像通过预处理调整大小为 480×480 ,提高模型的训练速度^[23]。本文采用随机翻转、图像加噪、增加图像亮度等方法来增强数据,避免网络出现过拟合现象,更好地训练网络模型^[24]。卫星道路图像预处理过程如图 9 所示。

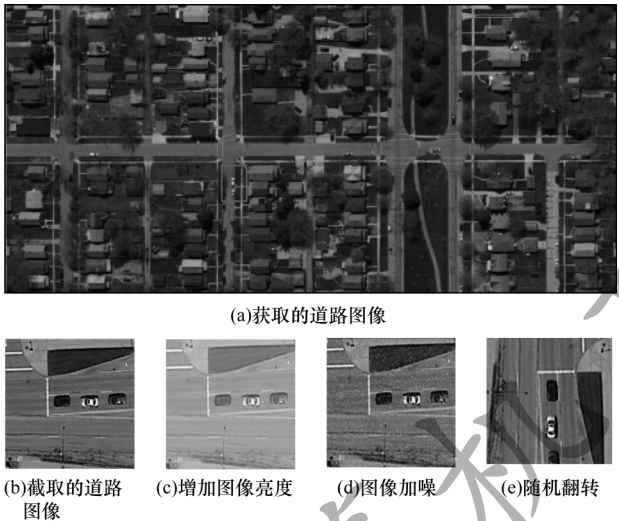


图 9 数据预处理过程
Fig.9 Data preprocessing process

2 实验

2.1 实验方法

将每张卫星道路图像的像素点分为 3 类,如表 2 所示。实验数据、实验环境和参数的设置如表 3 所示。

表 2 语义分割类别	
Table 2 Semantic segmentation categories	
语义分割要素	标记
道路面	0
道路线	1
背景	2

表 3 参数设置			
Table 3 Parameter settings			
实验数据	数量	实验环境参数	
		框架	TensorFlow
训练集	3 000	初始学习率	0.001
验证集	1 000	batch size	10
测试集	1 000	epoch	20

2.2 性能指标

采用准确率(Accuracy)、召回率(Recall)和精确率(Precision)3 个性能指标对本文算法分割效果进行评估^[25]。U-Net 和本文算法在训练过程中精准率

和损失函数分析结果如图 10、图 11 所示。

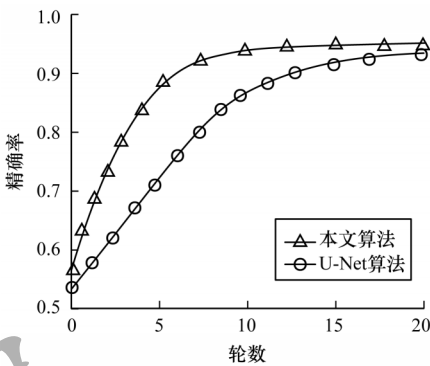


图 10 网络训练阶段精度比变化
Fig.10 Accuracy ratio change of network training

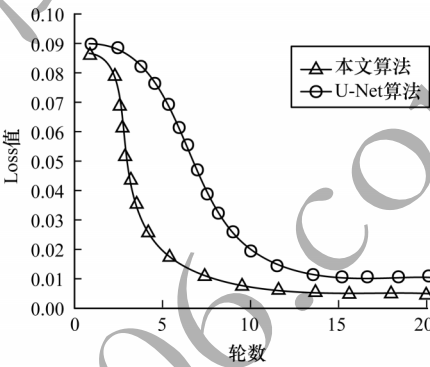


图 11 网络训练阶段 Loss 变化
Fig.11 Loss changes of network trainingstage

从图 10 可以看出:在训练过程中经过 20 轮迭代,本文算法的准确率永远高于 U-Net,证明本文算法分割道路图像的可行性。

从图 11 可以看出:在训练过程中经过 20 轮迭代,U-Net 的损失曲线永远高于本文算法,本文算法的损失曲线比 U-Net 提前达到稳定值,损失值越小,表示模型的拟合效果越好;损失曲线越光滑,表示模型的鲁棒性越强;根据曲线变化可得本文具有较高的鲁棒性。

2.3 实验对比与分析

实验 1 研究不同改进策略对道路图像语义分割性能影响。实验结果如表 4 所示,其中粗体为最优值。

表 4 不同策略的结果对比			
Table 4 Results comparison of different strategies			
策略	准确率	召回率	精确率
U-Net	0.955	0.947	0.952
U-Net+密集连接模块	0.957	0.950	0.955
U-Net+ASPP	0.958	0.952	0.954
U-Net+CBAM	0.959	0.956	0.955

从表 4 可以看出:在 U-Net 基础上分别引入密集连接模块、ASPP 和 CBAM,准确率分别提升了 0.2、0.3 和 0.4 个百分点,召回率分别提升了 0.3、0.5 和 0.9 个百分点,精确率分别提升了 0.3、0.2 和 0.3 个百分点。通过性能指标结果可得:引入密集连接模块、ASPP 和 CBAM 能够有效利用图像上下文信息,提升网络模型的性能。特别是在下采样时引入密集连接模块可以保留更多的图像信息,图像分割结果的精确率达到 95.5%。实验结果如图 12 所示。

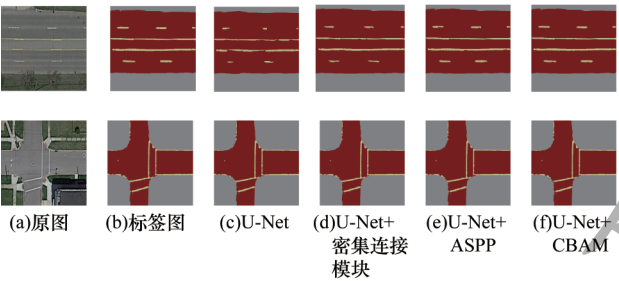


图 12 不同改进策略实验结果

Fig.12 Experimental results of different improvement strategies

从图 12 可以看出:在 U-Net 网络基础上分别引入密集连接模块、ASPP 和 CBAM 分割的道路元素比原始的 U-Net 更清晰,但仍存在道路线和道路面分割中断和未分割的情况,说明分别引入 3 种不同的改进策略对道路元素的分割存在缺陷,对道路元素分割的准确性仍需要提高。

实验 2 研究不同改进策略相互结合对道路图像语义分割性能影响。实验结果如表 5 所示,其中粗体为最优值。

表 5 不同改进策略结合的结果对比

Table 5 Results comparison of different improvement strategies combining

策略	准确率	召回率	精确率
U-Net+密集连接模块+ASPP	0.960	0.953	0.956
U-Net+密集连接模块+CBAM	0.961	0.958	0.959
U-Net+ASPP+CBAM	0.962	0.962	0.962
本文算法	0.963	0.969	0.966

从表 5 可以看出:在 U-Net 模型中引入密集连接模块、ASPP 和 CBAM,任意两者结合,准确率、召回率和精确率都比分别引入 3 种改进策略有所提高。实验结果表明,同时结合 3 种改进策略的准确率、召回率和精确率具有更优异的表现,验证了本文算法的有效性。特别是在 U-Net 引入密集连接模块的基础上再分别引入 ASPP 和 CBAM 后,准确率比 U-Net 引入密集连接模块提高了 0.3 和 0.4 个百分点。实验结果如图 13 所示。

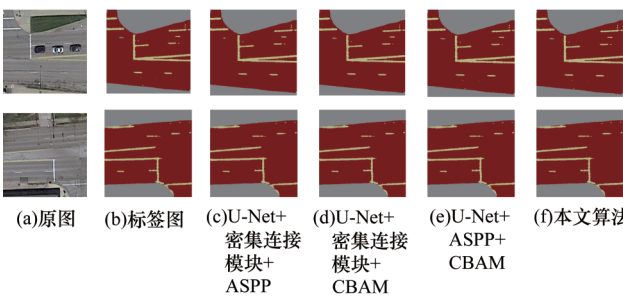


图 13 不同改进策略相结合的实验结果

Fig.13 Experimental results of different improvement strategies combining

从图 13 可以看出:将改进策略相互结合更大限度地读取了图像的特征信息,提高道路图像语义分割的结果,但仍存在对部分道路线元素分割能力较差。根据实验结果对比可得,U-Net 结合 ASPP 和 CBAM 模块对道路语义分割效果更佳,原因是 ASPP 融合多尺度的图像特征信息,而且 CBAM 提高网络对目标区域的关注。在此基础上本文算法同时结合 3 种改进策略,能够清楚地分割出道路线、道路面和背景,尤其是提高道路线的分割准确性更为明显。

实验 3 研究空洞卷积扩张率对算法性能的影响。本文研究在 ASPP 模块中采用正常卷积和空洞卷积对道路语义分割结果的影响。实验结果如表 6 所示,其中粗体为最优值。

表 6 空洞卷积扩张率对网络性能的影响

Table 6 Impact of the dilated convolution expansion rate on network performance

空洞卷积扩张率	准确率	召回率	精确率
$C_1=1,1,1$	0.958	0.959	0.960
$C_2=3,5,7,9$	0.960	0.964	0.962
$C_3=2,4,8,16$	0.963	0.969	0.966
$C_4=4,8,16,32$	0.961	0.967	0.964

从表 6 可以看出:正常卷积(C_1)在一个小区域内获取图像特征信息,分割效果较差;空洞卷积(C_2, C_3, C_4)在不增加参数的情况下,增大卷积核的感受野,卷积输出范围更大的空间信息,使得网络保留了更多的图像特征。因此,采用空洞卷积的分割效果要优于正常卷积。为验证扩张率对网络性能的影响,ASPP 采用不同扩张率组合。实验结果表明,随着扩张率的增加,语义分割的准确率、召回率和精确率均有所提高。但是当空洞卷积为 C_4 时,图像特征信息的关联性减弱,降低了分割效果。因此,本文采用 ASPP 空洞卷积的扩张率为 2、4、8、16。

实验 4 研究本文算法在速度方面的对比实验。实验结果如表 7 所示,其中粗体为最优值。

表7 本文算法速度对比

Table 7 Comparison of algorithm speed in this paper

算法	样本数量	准确率	训练时间 /h	验证速度 /(frame·s ⁻¹)	测试速度 /(frame·s ⁻¹)
U-Net	1 057	0.949	18.23	22.37	18.28
本文算法	1 057	0.957	18.74	20.18	17.64
U-Net	3 000	0.955	48.47	21.88	18.35
本文算法	3 000	0.963	50.02	20.24	17.79

当训练样本为1 057张图像时,本文算法的训练时间比U-Net多0.51 h,验证速度慢2.19 frame/s,测试速度慢0.64 frame/s。当训练样本为3 000张图像时,本文算法的训练时间比U-Net多1.55 h,验证速度慢1.64 frame/s,测试速度慢0.56 frame/s。分析结果表明:网络模型计算量越大,语义分割的速度越慢,但语义分割的准确率越高,证明了本文算法在道路图像语义分割方面的有效性。

实验5 研究不同算法对道路图像语义分割的性能指标对比结果。实验结果如图14所示。

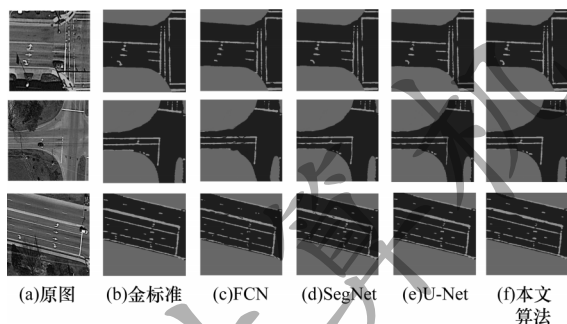


图14 不同算法的道路分割结果

Fig.14 Road segmentation results of different algorithms

从图14可以看出:本文算法融合了多尺度的图像并专注于特定的图像信息,提高了道路语义分割精确度,可以清晰地识别出道路图像中的道路面、道路线和背景3个要素。

将不同算法对道路图像的3种要素(道路面、道路线和背景)语义分割的准确率、召回率和精确率进行对比。实验结果如表8所示。

表8 性能指标对比分析

Table 8 Comparative analysis of performance indicators

性能指标	算法	道路线	道路面	背景	平均
准确率	FCN ^[5]	0.934	0.951	0.959	0.948
	SegNet ^[6]	0.941	0.953	0.964	0.953
	U_Net ^[9]	0.953	0.955	0.957	0.955
	本文算法	0.959	0.967	0.964	0.963
召回率	FCN ^[5]	0.929	0.943	0.951	0.941
	SegNet ^[6]	0.936	0.945	0.952	0.944
	U_Net ^[9]	0.939	0.949	0.954	0.947
	本文算法	0.969	0.968	0.969	0.969
精确率	FCN ^[5]	0.949	0.944	0.949	0.947
	SegNet ^[6]	0.951	0.947	0.95	0.949
	U_Net ^[9]	0.954	0.948	0.953	0.952
	本文算法	0.967	0.963	0.969	0.966

横向对比表8性能指标结果可得:道路面和背景的分割结果好于道路线。主要原因是3个要素在实验数据中的覆盖率不同,在模型训练过程中根据要素阈值进行筛选,从而出现3个要素语义分割不等价的情况。

纵向对比表8性能指标结果可得:本文算法语义分割结果好于其他语义分割模型。FCN保证输出图像和输入图像大小一样,语义分割准确率达0.948,但是在下采样过程中丢失了图像部分特征信息。SegNet引入归一化方法和Softmax分类器,但池化层的存在导致了图像位置信息丢失,SegNet比FCN道路图像语义分割的精确度提高了0.5个百分点。U-Net通过编码器与解码器之间的级联操作融合图像高层信息与浅层信息,U-Net比FCN和SegNet减少语义信息的丢失,保留了图像细节。U-Net语义分割的准确率为0.959,比FCN和SegNet语义分割准确率提高了0.7和0.2个百分点。本文算法相比U-net对道路图像语义分割平均准确率提升了0.8个百分点。本文算法相比U-Net在道路线、道路面和背景语义分割准确率分别提升了0.6、1.2、0.7个百分点,主要是因为本文算法在U-Net网络的基础上引入密集连接模块,ASPP可以融合多尺度的图像特征信息,而注意力机制模型对关键信息的权值加重,提高神经网络对目标区域的关注,专注于输入图像特定部分。

3 结束语

本文设计一种改进的U-Net网络结构,该网络结构在下采样时引入密集连接模块,在上采样前引入ASPP和CBAM,从而对卫星图像道路元素做出准确的分割。实验结果表明,与FCN、SegNet和U-Net等传统网络相比,该网络对道路要素分割的准确率更高。下一步将继续对网络结构进行优化,提高网络运行速度和语义分割的精确度,以更清晰地识别出道路图像信息。

参考文献

- [1] 王海,蔡英凤,贾允毅,等. 基于深度卷积神经网络的场景自适应道路分割算法[J]. 电子与信息学报,2017,39(2): 263-269.
WANG H, CAI Y F, JIA Y Y, et al. Scene adaptive road segmentation algorithm based on deep convolution neural network[J]. Journal of Electronics and Information, 2017, 39(2): 263-269. (in Chinese)
- [2] 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.
SONG Q S, ZHANG C, CHEN Y, et al. Road segmentation based on combined total convolution neural network and conditional random fields [J]. Journal of Tsinghua University, 2018, 58(8): 725-731. (in Chinese)
- [3] 杨飞,王欢,金忠. 基于融合分层条件随机场的道路分割模型[J]. 机器人,2018,40(6): 803-816.

- YANG F, WANG H, JIN Z. Road segmentation model based on fusion hierarchical conditional random fields[J]. Robot, 2018, 40(6): 803-816.
- [4] ONUEAM A, CHOI N, KIM T, et al. Adaptive road segmentation of OpenStreetMap data using Haversine formula[J]. Information and Communication Engineering, 2018, 10(1): 297-299.
- [5] MA W, YANG Q, WU Y, et al. Double-branch multi-attention mechanism network for hyperspectral image classification[J]. Remote Sensing, 2019, 11(11): 1307-1321.
- [6] ZAIMBASHI A, LI J. Tunable adaptive target with kernels in Colocated MIMO[J]. IEEE Transactions on Signal, 2020, 45(12): 1500-1514.
- [7] YU A, WANG L. Using the wide-range attention U-Net for road segmentation[J]. Remote Sensing Letters, 2019, 10(5): 506-515.
- [8] HAO H, YANG D F, WANG S C, et al. Road segmentation of cross-modal remote sensing images using deep segmentation network and transfer learning[J]. International Journal of Robotics Research and Application, 2019, 46(3): 384-390.
- [9] MISTYI V H, MAKWANA R M. Road segmentation and tracking on Indian road scenes[J]. International Journal of Computer Aided Engineering and Technology, 2019, 11(4/5): 480-494.
- [10] FAN R, BOCUS M J, DAHNOUN N. A novel disparity transformation algorithm for road segmentation[J]. Information Processing Letters, 2018, 140: 18-24.
- [11] CHENG G L, WU C R, YU H, et al. Recognizing road from satellite images by structured neural network[J]. Neurocomputing, 2019, 356: 131-141.
- [12] LIU Y, JIN W. Rethinking labelling in road segmentation[J]. International Journal of Remote Sensing, 2019, 40(22): 8359-8378.
- [13] ZHANG X X, ZHU X. Efficient and de-shadowing approach for multiple vehicle tracking in aerial video via image segmentation and local region matching[J]. Journal of Applied Remote Sensing, 2020, 14(1): 14503-14517.
- [14] CHENG G L, ZHU F Y, XIANG F M, et al. Accurate urban road centerline extraction from VHR imagery via multiscale segmentation and tensor voting[J]. Neurocomputing, 2016, 205: 407-420.
- [15] ZAIMBASHI A, LI J. Tunable adaptive target detection with kernels in colocated MIMO Radar[J]. IEEE Transactions on Signal Processing, 2020, 41(3): 1500-1514.
- [16] CORONA E, ALENY G, GABAS A, et al. Active garment recognition and target grasping point detection using deep learning[J]. Pattern Recognition, 2018, 74: 629-641.
- [17] YU C Q, WANG J B, PENG C, et al. BiSeNet: bilateral segmentation network for real-time semantic segmentation[C]//Proceedings of the 15th IEEE European Conference on Computer Vision. Berlin, Germany: Springer, 2018: 334-349.
- [18] 周飞, 唐建, 杨成松, 等. 基于混合自动编码器道路语义分割方法研究[J]. 计算机工程与科学, 2019, 41(8): 1453-1458.
- ZHOU F, TANG J, YANG C S, et al. Research on road semantic segmentation method based on hybrid automatic encoder[J]. Computer Engineering and Science, 2019, 41(8): 1453-1458. (in Chinese)
- [19] HU J, SHEN L, SUN G. Squeeze-and-excitation[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2018: 7132-7141.
- [20] 闫昭帆, 李雨冲, 严国萍. 基于多特征融合和条件随机场的道路分割[J]. 计算机系统应用, 2020, 29(3): 240-245.
- YAN Z F, LI Y C, YAN G P. Road segmentation based on multi-feature fusion and conditional random field[J]. Computer System Application, 2020, 29(3): 240-245. (in Chinese)
- [21] 张凯航, 冀杰, 蒋骆, 等. 基于SegNet的非结构道路可行驶区域语义分割[J]. 重庆大学学报, 2020, 43(3): 79-87.
- ZHANG K H, JI J, JIANG L, et al. Semantic segmentation of driving area of unstructured road based on SegNet[J]. Journal of Chongqing University, 2020, 43(3): 79-87. (in Chinese)
- [22] 莫虹平. 基于递归神经网络的高分辨率航拍图像道路分割与提取[D]. 厦门: 厦门大学, 2017.
- MO H P. Road segmentation and extraction from high-resolution aerial images based on recurrent neural network[D]. Xiamen: Xiamen University, 2017. (in Chinese)
- [23] FU F F, WANG Y, LAI F C, et al. Efficient rain-fog model for rain detection and removal[J]. Journal of Electronic Imaging, 2020, 29(2): 23020-23032.
- [24] YANG F, RAO Y T. Vision-based intelligent vehicle road recognition and obstacle detection method[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2020, 34(7): 15-28.
- [25] LAN M, ZHANG Y P, ZHANG L F, et al. Global context based automatic road segmentation via dilated convolutional neural network[J]. Information Sciences, 2020, 535: 156-171.