

基于深度学习的遥感图像地物分割方法

沈言善, 王阿川

引用本文:

沈言善, 王阿川. 基于深度学习的遥感图像地物分割方法[J]. 液晶与显示, 2021, 36(5): 733–740.

SHEN Yan-shan, WANG A-chuan. Remote sensing image feature segmentation method based on deep learning[J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(5): 733–740.

在线阅读 View online: <https://doi.org/10.37188/CJLCD.2020-0294>

您可能感兴趣的其他文章

Articles you may be interested in

基于深度学习的磁瓦表面缺陷分割与识别

Segmentation and recognition of magnetic tile surface defects based on deep learning

液晶与显示. 2021, 36(5): 713–722 <https://doi.org/10.37188/CJLCD.2020-0232>

基于深度卷积神经网络的宫颈细胞病理智能辅助诊断方法

Intelligent auxiliary diagnosis method of cervical cytopathology based on deep convolutional neural networks

液晶与显示. 2018, 33(6): 528–537 <https://doi.org/10.3788/YJYXS20183306.0528>

基于U-Net的多图谱标签融合算法

Multi-atlaslabel fusion based on U-Net

液晶与显示. 2019, 34(11): 1091–1103 <https://doi.org/10.3788/YJYXS20193411.1091>

多尺度特征融合与极限学习机结合的交通标志识别

Traffic sign recognition based on multi-scale feature fusion and extreme learning machine

液晶与显示. 2020, 35(6): 572–582 <https://doi.org/10.3788/YJYXS20203506.0572>

基于Mask R-CNN的ORB去误匹配方法

ORB removal mis-matching method based on Mask R-CNN

液晶与显示. 2018, 33(8): 690–696 <https://doi.org/10.3788/YJYXS20183308.0690>

文章编号:1007-2780(2021)05-0733-08

基于深度学习的遥感图像地物分割方法

沈言善, 王阿川*

(东北林业大学 信息与计算机工程学院, 黑龙江 哈尔滨 150000)

摘要:针对传统方法难以同时有效地进行多目标分割,现有基于全卷积神经网络的地物分类方法在复杂场景下分类精度不高的问题,提出了一种基于 U-net 改进的全卷积神经网络 DL-Unet,实现了对遥感图像不同类别地物的有效分割。该网络改进了传统的卷积方式,引入扩张卷积,在不增加网络参数的同时增大感受野。针对遥感影像中地物类别不均衡的问题,使用带权重的交叉熵作为模型的损失函数,有效克服了模型的选择偏好。对预测结果采用了相对多数投票策略进一步提高了各个地物类别的像素准确率(PA)。实验结果表明,该模型在各地物类别的平均像素准确率(MPA)和均交并比(MIoU)上相较于经典 U-net 网络分别提高了 5.94% 和 9.45%,实现了对地物信息的有效分割。

关键词:深度学习;扩张卷积;加权交叉熵;集成学习;语义分割;图像处理

中图分类号:TP391 文献标识码:A doi:10.37188/CJLCD.2020-0294

Remote sensing image feature segmentation method based on deep learning

SHEN Yan-shan, WANG A-chuan*

(College of Information and Computer Engineering,
Northeast Forestry University, Harbin 150000, China)

Abstract: In view of the difficulty of traditional methods for simultaneous and effective multi-target segmentation, the existing ground object classification methods based on fully convolutional neural networks have low classification accuracy in complex scenes, this paper proposes an improved encoder-decoder based on the U-shaped network structure DL-Unet, which realizes the effective segmentation of remote sensing images. This network improves the traditional convolution method and introduces the expanded convolution, which increases the receptive field without increasing the network parameters. Aiming at the problem of imbalance in the classification of features in remote sensing images, weighted cross-entropy is used as the loss function of the model, which effectively overcomes the selection preference of the model function of the model effectively. The relative majority voting strategy is adopted for the prediction results to further improve the pixel accuracy (PA) of each feature category. The experimental results show that the mean pixel accuracy (MPA) and mean intersection over union (MIoU) of this model are improved by 5.94% and 9.45% respectively compared with the classic U-net, which verifies that the method in this paper is an effective remote sensing image classifica-

收稿日期:2020-11-03;修订日期:2020-11-25.

基金项目:黑龙江省自然科学基金(No.C201414);哈尔滨市优秀学科带头人基金(No.2014RFXXJ040)

Supported by Natural Science Foundation of Heilongjiang Province (No.C201414); Harbin Foundation for Good Leaders of Disciplines(No.2014 RFXXJ040)

* 通信联系人, E-mail: wangcal1964@126.com

tion method.

Key words: deep learning; dilation convolution; weighted cross entropy; ensemble learning; semantic segmentation; image processing

1 引言

如何将遥感图像中的地物信息快速准确地进行分割是当前的研究热点之一^[1]。传统的分割方法与地物的一系列特征有关,针对不同地物的特征,采用的方法也不同。针对森林植被分类,李亮等^[2]提出了一种综合光谱、纹理、结构特征的高分辨率遥感影像变化检测方法。针对道路分类有最近邻^[3]、隶属度函数^[4]等方法。上述传统方法只能针对某种地物特性采用相对应的方法,分割的地物种类单一。而高分辨率遥感图像通常包含道路、建筑、水源、植物等具有丰富细节信息的地物目标。针对遥感图像中多种信息特征,传统的分割方法很难同时有效地进行多目标分割。近年来,计算机视觉技术发展迅速,语义分割作为研究热点之一,在卫星遥感图像和医学影像等领域都有着广泛的应用前景。Long 等^[5]第一个提出了用于语义分割的全卷积神经网络(FCN),开启了全卷积网络用于图像语义分割的新篇章。其他优秀的语义分割网络结构还包括 U-net^[6]、SegNet^[7]、MobileNetV2^[8]、PSPNet^[9]等。但是在遥感图像中,受目标类别分布不均衡等因素的影响,常用的语义分割方法针对高分辨率遥感图像的多类别分割问题并不能达到有效目的。

为了解决以上问题,本文借鉴了 U-net 模型结构,设计并提出了一种改进的全卷积神经网络模型用于高分辨率遥感图像的地物分类任务。通过使用扩张卷积作为卷积层之间的级联方式在不增加网络参数的情况下增大感受野,降低训练与预测时的计算量,提高了模型运行的效率。针对样本类别不均衡带来的模型偏好问题,采用加权交叉熵作为模型的损失函数提高了模型的泛化能力。实验结果表明,本文模型总体的分割效果优于 U-Net 和 SegNet 模型,证明了该模型在遥感图像地物分类中的有效性。

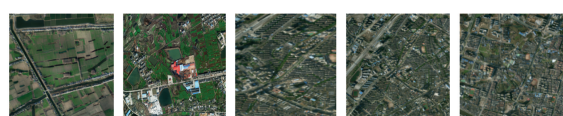
2 相关理论及网络设计

全卷积神经网络(FCN)的提出是在 2016 年,

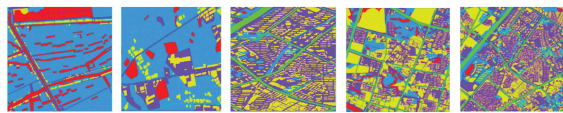
相比较于 CNN 和 VGG 等卷积全连接的网络结构,FCN 将卷积层替代全连接层来处理语义分割问题。全卷积神经网络主要有 3 个特点:一是采用全卷积的方式对输入图像进行特征提取和下采样;二是使用双线性插值的方法进行特征图的上采样;三是使用跳跃连接进行对应层的特征融合。早期基于深度学习的图像分割以 FCN 为核心,旨在重点解决如何更好地从卷积下采样中恢复丢失的信息损失。U-Net 是一个结构对称的 U 型全卷积神经网络,它和 FCN 都是基于 Encoder-Decoder^[10]结构。Encoder 负责特征提取,由卷积操作和下采样操作组成,在这里采用的卷积结构统一为 3×3 的卷积核,padding 为 0, striding 为 1。Decoder 负责样本还原,由反卷积操作和上采样操作组成。反卷积的卷积核也为 3×3 ,上采样采用双线性插值法。Skip Connection 将 U-Net 中的底层特征中的位置信息和深层特征中的语义信息相融合,实现端到端的像素级分类。

2.1 数据集及其预处理

本文采用的数据集为 2015 年南方某地区的高分辨率卫星遥感影像,共包含 5 幅大尺寸的原始图像和 5 幅标签图。该数据集的标签图共标记了 5 种类别的地物目标。为了更好地观察标注情况,我们将人工标记图片可视化如下:紫色表示其他、蓝色表示植被、绿色表示道路、黄色表示建筑、红色表示水体。图 1(a)和(b)分别是公开数据集样本和可视化的标记图。表 1 是本文所用数据集的样本分布统计情况。将数据集用 4:1 的比例分开,分别用作实验的训练集和测试集。由于原始数据集中遥感图像尺寸从 $4\,000 \times 4\,000$ 到 $8\,000 \times 8\,000$ 不等,直接作为网络的输入图像会导致内存溢出而无法训练,因此使用 OpenCV 对其做了相应的数据增强处理。具体地,将用于模型训练的每幅标签图的 4 个地物类别分别截取成仅包含植被、道路、建筑、水体的 4 幅子图。把训练集中每幅原始图像和标签子图随机切割成 256×256 大小的图像块,然后对图像块做了旋转操作、色彩调整^[11]、以及增加噪声操作(Gaussian noise, salt-and-pepper noise)。通过上述处理,新



(a) 原始图像
(a) Original image



(b) 可视化的标签图
(b) Visualized label map

图 1 数据集展示

Fig.1 Data set display

的训练集包含 10 万幅 256×256 大小的子图像。

表 1 数据集样本分布统计

Tab.1 Dataset sample distribution statistics

地物类别	标记值	可视化颜色	样本所占百分比/%
其他	0	紫色	29.5
植被	1	蓝色	32.6
道路	2	绿色	6.5
建筑	3	黄色	24.8
水体	4	红色	6.6

2.2 改进卷积方式

针对遥感图像的语义分割是一种端到端的像素级分类任务,图像分割的精确度一定程度上取决于上下文的信息。传统的卷积计算通过增大卷积核的尺寸来获得更多的感受野,但这样会使网络的参数增多,消耗更多的硬件资源。扩张卷积^[12](Dilated Convolution)是一种不同于传统卷积的卷积方式,能够对输入图像采用具有更大感受野的滤波器提取特征,并且不会增加网络的参数。如图 2 所示,红点表示两种相同尺寸不同卷积方式的卷积核,蓝色的区域表示感受野大小。图 2(a)中扩张率为 1,感受野大小为 3,与传统卷积计算的感受野一样;图 2(b)中扩张率为 2,感受野大小为 7。若采用传统卷积,则需尺寸为 7×7 的卷积核才能获得同样的感受野。对比图 2(a)和(b)可知,相较于传统卷积,扩张卷积在不增加卷积核尺寸的同时能获得更大的感受野。图 3(a)为 2 个标准 3×3 卷积层级联所组成的卷积序列。图 3(b)为将两层级联的卷

积层设置扩张率为 1 和 2 的扩张卷积序列。对比图 3(a)和(b)可以看出,在使用扩张率为 1 和 2 的扩张卷积级联形成卷积序列中,高层神经元在输入层能具备更大感受野。通过图 3(a)的方法将模型中的卷积模块变为扩张卷积模块,能够在增大高层神经元在输入层上感受野的同时保证了神经元的语义信息连续性。

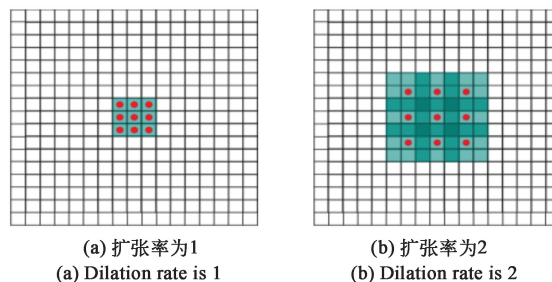
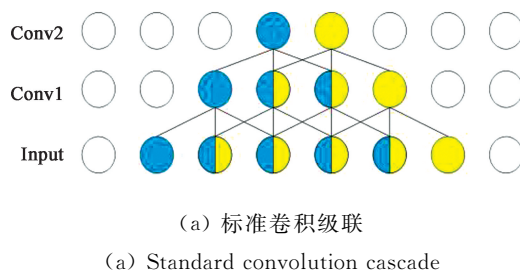


图 2 扩张卷积示意图

Fig.2 Schematic diagram of dilated convolution



(a) Standard convolution cascade

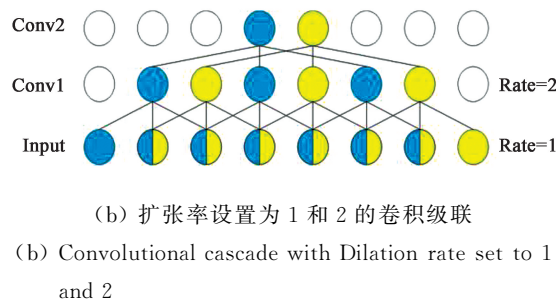


图 3 不同卷积级联

Fig.3 Different convolution cascade

2.3 激活函数的选取

激活函数是神经网络结构中的重要部分,其中经典 U-Net 网络中采用 RELU^[13]函数作为激活函数,其定义如式(1)所示:

$$f(x) = \max(0, x), \quad (1)$$

其函数变化曲线如图 4 所示。RELU 实际上是一个分段函数,当输入为正值时输出不变,但在

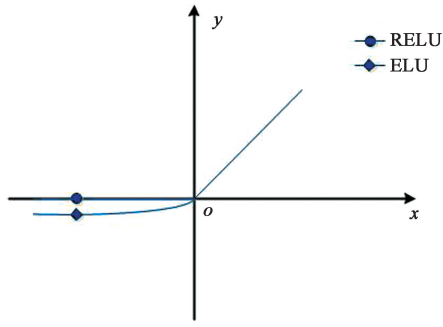


图 4 RELU 和 ELU 的函数图

Fig.4 Function diagram of RELU and ELU

输入为负值时,RELU 是硬饱和的,在反向传播的中梯度始终为 0,权重参数无法更新,会导致神经元死亡。为解决输入为负值时神经元失活问题,将 RELU 函数的前半部分从 0 变为一个指数函数,得到了指数线性单元^[14](ELU)。其定义如式(2)所示:

$$f(x) = \begin{cases} x, & x > 0 \\ \alpha(ex - 1), & x \leq 0 \end{cases}, \quad (2)$$

对于 ELU 的导数来说,如果输入为负值,输出不为常数 0,就可以避免死亡 RELU 问题,同时还保持了一定的稀疏性,可以有效缓解梯度消失问题。

2.4 损失函数的设计

损失函数是遥感图像语义分割任务中重要的组成部分,常用的分类损失均可用作语义分割的损失函数。其中经典 U-Net 网络模型使用的二分类交叉熵作为损失函数,其定义如式(3)所示:

$$\text{Loss} = -[y_{ij} \cdot \lg p_{ij} + (1 - y_{ij}) \cdot \lg (1 - p_{ij})], \quad (3)$$

p_{ij} 是像素 i 被模型预测为 j 类的概率, y_{ij} 是样本实际标签,如果像素点 i 属于 j 类,取值为 1,否则取值为 0。

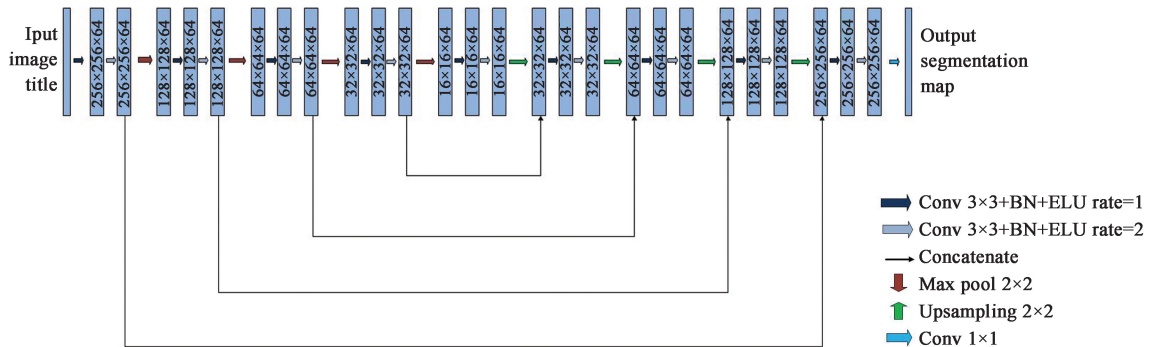


图 5 DL-Unet 模型网络结构

Fig.5 DL-Unet model network structure

本文数据集目标地物类别有 4 类(不包括“其他”类),在表 1 中样本数量较少的道路和水体分别占 6.5% 和 6.6%,而建筑和植被则占到了 24.8% 和 32.6%。使用具有类别不平衡问题的数据作为训练集,网络在训练过程中拟合过的各地物类别的像素数量不同,影响了其对不同类别像素分割的准确率。在图片分类问题中,往往通过数据增强的方式使不同类别图片在数量上实现均衡。但遥感图像地物分类不同于图片分类任务,一幅遥感图像中的标签信息并非单个类别值,而是一幅二维平面上的类别分布图。在实验中发现,仅是通过数据增强的方法增加小类别图片的数量对模型的分割精度没有相应提高,因此对遥感图像的类别不平衡问题应该考虑通过在损失函数中添加权重项进行类别均衡。通过为小类别像素赋予更大的权重值,为大类别赋予较小的权重值,可缩小在模型训练中不同地物类别像素计算损失函数平均值的差距。

本网络采用根据样本不同地物类别比例作为权重项的加权交叉熵作为模型的损失函数,其定义如式(4)所示:

$$\text{Loss} = -[w_{ij} y_{ij} \lg p_{ij} + (1 - y_{ij}) \lg (1 - p_{ij})],$$

$$w_{ij} = \frac{n - n_{ij}}{n}, \quad (4)$$

p_{ij} 是像素 i 被模型预测为 j 类的概率。 y_{ij} 是样本实际标签,如果像素点 i 属于 j 类,取值为 1,否则取值为 0。 w_{ij} 是权重系数, n 为样本总像素数, n_{ij} 为像素点 i 属于 j 类的像素数。

2.5 基于改进 U-net 的遥感图像语义分割模型

本文所改进的 DL-Unet 网络模型结构如图 5 所示。该网络包含 9 个卷积级联结构,每一个

卷积级联设置为图3所示的扩张率为1和2扩张卷积序列,在卷积运算后加入BN层^[15]。在网络的编码器部分,每个卷积层的卷积核尺寸为 3×3 ,过滤器(Filter)深度为64。每两个卷积级联之间有一个窗口大小为 2×2 、步长为2的最大池化层(Maxpooling)。网络的解码器结构与编码器结构完全对称,卷积层的卷积核大小和过滤器深度与编码器部分相同。每两个卷积级联之间有一个窗口大小为 2×2 、步长为2的上采样层(Upsampling)。编码器和解码器部分相对应的卷积层采用跳跃连接的方式融合特征信息。

3 实验与分析

3.1 性能评价

遥感图像地物分类是一种图像分割问题,其本质上是一种图像像素分类。对分类后的遥感图像,可以使用像素准确率(Pixel Accuracy, PA)、平均像素准确率(Mean Pixel Accuracy, MPA)与平均交并比(Mean Intersection over Union, MIoU)等^[16]常见的分类评价指标来评估模型好坏。评价每个像素的预测输出结果,即该像素分类结果取真正例(TP)、假正例(FP)、真反例(TN)、假反例(FN)4种结果中的一种,然后根据这4个指标计算,为了便于解释,假设 $k+1$ 表示类别数量(其中包含一个空类或背景), p_{ij} 表示假正例的像素数量, p_{ji} 表示假负例的像素数量, p_{ii} 表示真正例的像素数量。PA表示预测正确的像素占总像素的比例,定义如式(5)所示:

$$PA = \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}}, \quad (5)$$

$k+1$ 表示类别数量(其中包含一个空类或背景)。

MPA表示分别计算每个类别分类正确的像素数占所有预测为该类别像素比例的平均值,定义如式(6)所示:

$$MPA = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}}, \quad (6)$$

$k+1$ 表示类别数量(其中包含一个空类或背景)。

MIoU的定义为:将标签图像和预测图像看成是两个集合,计算两个集合的交集和并集的比值,将所有类别计算得出的比值取平均。定义如

式(7)所示:

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}}, \quad (7)$$

$k+1$ 表示类别数量(其中包含一个空类或背景)。

3.2 模型训练及测试

实际为每一类地物目标训练了一个二分类模型,共得到植被、道路、建筑、水体的分割模型。在上述的实验环境下对图5所示的网络模型进行训练。实验将数据集按5:1比例进行划分,即75 000张图像作为训练集,25 000张图像作为验证集。模型训练采用批处理方式,每一个batch输入16张图片,完成一个epoch需要6 250个batch,设定模型完成30次epoch。模型基于keras框架使用GPU进行训练,优化器设置为Adam(Adaptive moment estimation),评价指标为准确率(Accuracy),使用加权的交叉熵损失函数。在模型训练的过程中设定监视对象为val_acc,当val_acc最大时自动保存最优权值,保存在最优权重下的模型。训练过程中的学习率设置为0.001。如图6所示,随着迭代次数的增加,训练集和验证集的准确率与损失变化趋于平缓,表明模型达到了最优。

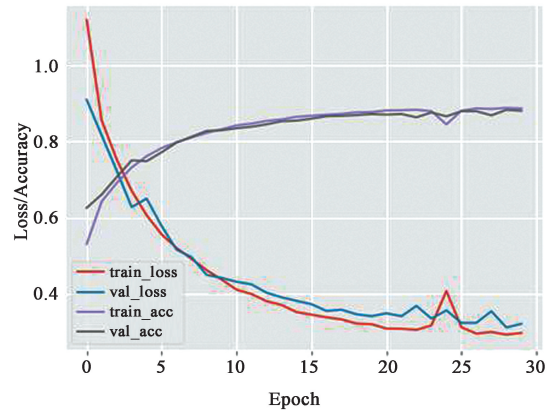


图6 训练过程

Fig.6 Training process

将训练好的各类模型分别对测试图像进行植被、道路、建筑和水体的预测,共得到4幅预测好的子图。由于测试图像尺寸较大,需要将其分割为 256×256 大小的多张图片进行预测,然后根据其位置再进行拼接成一幅完整的预测图。在合并4幅分割好的各类别子图时,使用相对多数投票

的策略,即判定该像素的类别为 4 幅子图中出现次数最多的类别。如果某一像素点类别判定出现次数相同的情况时,则按照植被>道路>建筑>水体的优先级来确定该像素点的类别。遥感图像地物分割的流程如图 7 所示。

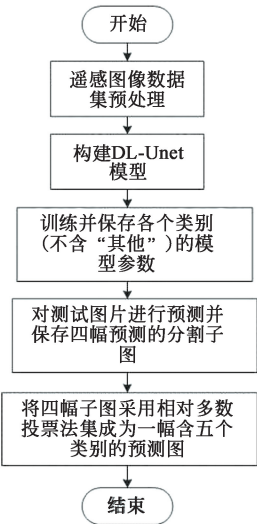


图 7 遥感图像地物分割流程图
Fig.7 Remote sensing image feature segmentation flowchart

将预测出的结果写入灰度图中,像素值介于 0~4 之间,为了便于观察,将每个像素值可视化 为不同的颜色,如表 2 所示。

表 2 数据可视化转换表		
Tab.2 Data visualization conversion table		
地物类别	像素值	可视化颜色
其他	0	紫色
植被	1	蓝色
道路	2	绿色
建筑	3	黄色
水体	4	红色

3.3 实验结果分析

在上述实验条件下分别使用了经典的语义分割网络 FCN-8s、DeconvNet、SegNet、U-Net 和本文提出的方法进行训练和预测,各方法分割完成并可视化的结果如图 8 所示。

不同分割模型的各个地物类别的像素准确率 (PA)如表 3 所示,各个模型在测试集上的平均像

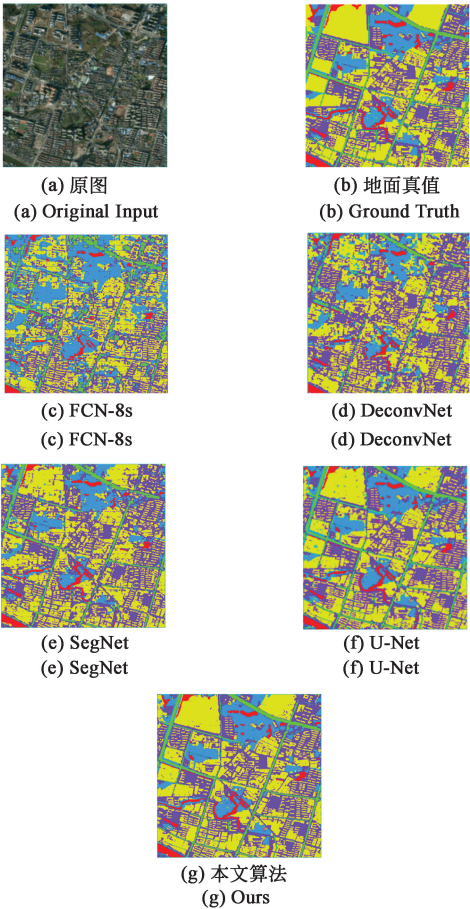


图 8 测试图片分割结果
Fig.8 Test image segmentation results

素准确率 (MPA) 和均交并比 (MIoU) 如表 4 所示。

从预测图和表 3 可以看出,本文的分割方法相较于多分类 U-Net 网络对于植被的像素准确率提高了 3.30%,道路的像素准确率提高了 5.40%,建筑的像素准确率提高了 5.22%,水体的像素准确率提高了 6.42%。从表 4 可以看出,本文的分割方法相较于 U-Net 网络在测试集上的平

表 3 各个地物类别的像素准确率				
Tab.3 Pixel accuracy of each feature category				
方法	植被	道路	建筑	水体
FCN-8s	0.848 0	0.477 2	0.624 5	0.494 2
DeconvNet	0.653 0	0.428 3	0.714 4	0.595 1
SegNet	0.745 2	0.582 8	0.790 4	0.696 7
U-Net	0.906 1	0.743 7	0.879 0	0.776 0
本文方法	0.939 1	0.797 7	0.932 8	0.840 2

表4 不同模型的测试结果

Tab.4 Test results of different models

方法	MPA	MIoU
FCN-8s	0.611 0	0.425 7
DeconvNet	0.600 3	0.465 5
SegNet	0.710 8	0.588 8
U-Net	0.825 0	0.711 6
本文方法	0.884 6	0.806 1

均像素准确率和均交并比分别提高了 5.94% 和 9.45%。

在使用上述的数据集外,我们新增加了 2020 CCF BDCI 遥感影像地块分割比赛的部分数据集来进行实验。该数据集包含 20 000 张分辨率为 2 m/pixel、尺寸大小为 256×256 的原始图和 20 000 张包含建筑、耕地、林地、水体、道路、草地、其他、未标注区域这 8 个类别语义标签的标签图。为了更好地符合本文网络的输入,将耕地、草地、林地归为植被类别,将未标注区域归为其他类。

表5 在 CCF BDCI 数据集上的不同模型的测试结果

Tab.5 Test results of different models on CCF BDCI database

方法	MPA	MIoU
FCN-8s	0.608 0	0.518 9
DeconvNet	0.626 1	0.521 4
SegNet	0.673 2	0.560 6
U-Net	0.676 2	0.562 3
本文方法	0.727 8	0.621 7

按 3.2 节所提的实验条件分别使用了经典的语义分割网络 FCN-8s、DeconvNet、SegNet、U-Net 和本文提出的方法进行训练和预测,各个模型在测试集上的平均像素准确率(MPA)和均交并比(MIoU)如表 5 所示。

从表 4 和表 5 可以看出本文方法在两个不同的遥感图像数据集上的平均像素准确率和均交并比优于 U-Net 和其他 3 种语义分割网络。

4 结 论

本文提出了一种基于 U-Net 网络改进的全卷积神经网络分割方法,对遥感图像实现了像素级别的分类。在保留了 U-Net 优点的同时,采用扩张卷积级联方式能考虑更多图像上下文信息并且不会增加网络的计算量。同时为了克服样本不均衡带来的模型选择偏好问题,本文采用了根据数据集地物类别比例来设置权重项的加权交叉熵作为模型的损失函数。对每一个地物类别都分别训练了一个二分类模型,然后采取相对多数投票策略并设置地物类别优先级的方法有效提升了预测图在各个地物类别的像素准确率(PA)。该模型的预测结果在平均像素准确率(MPA)和均交并比(MIoU)上相较于经典 U-Net 和 SegNet 有一定提升。实验结果表明,本文所提出的方法能够较精确地分割地物目标,达到比较好的效果。

但因硬件设备计算能力有限,模型的 batchsize 参数并不能设置得比较大,这影响了边缘分割的精细度。后续的研究将探索如何在进一步减少模型的参数情况下增加模型边缘分割的精度。

参 考 文 献:

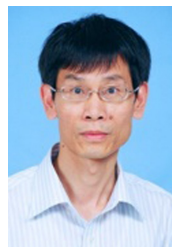
- [1] 林晨曦,周艺,王世新,等.基于变差函数的中高分辨率 SAR 影像农村建筑区提取[J].中国图象图形学报,2016,21(5):674-682.
LIN C X, ZHOU Y, WANG S X, et al. Variogram-based rural build-up area extraction from middle and high resolution SAR images [J]. *Journal of Image and Graphics*, 2016, 21(5): 674-682. (in Chinese)
- [2] 李亮,申学林,李胜,等.一种综合光谱、纹理、结构特征的高分辨率遥感影像变化检测方法[J].测绘通报,2019(S1): 113-118.
LI L, SHEN X L, LI S, et al. A change detection method by combining spectral, textural and structural features for VHR remote sensing imagery [J]. *Bulletin of Surveying and Mapping*, 2019(S1): 113-118. (in Chinese)
- [3] 惠振阳,胡友健.机载 LiDAR 点云中道路的提取方法[J].测绘科学,2017,42(3):70-74.
HUI Z Y, HU Y J. A review on road extraction methods from airborne LiDAR [J]. *Science of Surveying and*

- Mapping*, 2017, 42(3): 70-74. (in Chinese)
- [4] 胡唯, 胡庆武, 艾明耀. 多星形先验和图割的遥感影像交互式分割[J]. 遥感信息, 2016, 31(2): 19-23.
HU W, HU Q W, AI M Y. Interactive remote sensing image segmentation based on multiple star prior and graph cuts [J]. *Remote Sensing Information*, 2016, 31(2): 19-23. (in Chinese)
- [5] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651.
- [6] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C]//*Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich: Springer, 2015.
- [7] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. *Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [8] 周鹏, 姚剑敏, 林志贤, 等. 融合注意力机制的移动端人像分割网络[J]. 液晶与显示, 2020, 35(6): 547-554.
ZHOU P, YAO J M, LIN Z X, *et al.* Mobile-based portrait segmentation network with attention mechanism [J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(6): 547-554. (in Chinese)
- [9] ZHAO H S, SHI J P, QI X J, *et al.* Pyramid scene parsing network [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 2881-2890.
- [10] 张哲晗, 方薇, 杜丽丽, 等. 基于编码-解码卷积神经网络的遥感图像语义分割[J]. 光学学报, 2020, 40(3): 0310001.
ZHANG Z H, FANG W, DU L L, *et al.* Semantic segmentation of remote sensing image based on encoder-decoder convolutional neural network [J]. *Acta Optica Sinica*, 2020, 40(3): 0310001. (in Chinese)
- [11] 陈献明, 王阿川, 王春艳. 基于深度学习的木材表面缺陷图像检测[J]. 液晶与显示, 2019, 34(9): 879-887.
CHEN X M, WANG A C, WANG C Y. Image detection of wood surface defects based on deep learning [J]. *Chinese Journal of Liquid Crystals and Displays*, 2019, 34(9): 879-887. (in Chinese)
- [12] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [C]//*International Conference on Learning Representations*. Caribe Hilton: IEEE, 2016.
- [13] GLOROT X, BORDES A, BENGIO Y. Deep sparse rectifier neural networks [J]. *Journal of Machine Learning Research*, 2011, 15: 315-323.
- [14] CLEVERT D A, UNTERTHINER T, HOCHREITER S. Fast and accurate deep network learning by Exponential Linear Units (ELUs) [C]//*Proceedings of the 4th International Conference on Learning Representations*. San Juan, Puerto Rico: IEEE, 2016.
- [15] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C]//*Proceedings of the 32nd International Conference on International Conference on Machine Learning*. Lille: ACM, 2015.
- [16] 景庄伟, 管海燕, 彭代峰, 等. 基于深度神经网络的图像语义分割研究综述[J]. 计算机工程, 2020, 46(10): 1-17.
JING Z W, GUAN H Y, PENG D F, *et al.* Survey of research in image semantic segmentation based on deep neural network [J]. *Computer Engineering*, 2020, 46(10): 1-17. (in Chinese)

作者简介:



沈言善(1997—),男,湖北天门人,硕士研究生,2018年于东北林业大学获得学士学位,主要研究方向为机器学习、图像处理。E-mail: shenyanshan9792@163.com



王阿川(1964—),男,四川达州人,博士,教授,2011年于东北林业大学获得博士学位,主要研究方向为模式识别、计算机视觉、3S技术。E-mail: wang-cal1964@126.com