

基于 SegNet 网络和概率图模型的工区障碍物提取

胡 敏, 陈 楠, 毕进娜

(中石化石油物探技术研究院有限公司, 南京 211100)

摘要:为适应 5G(第 5 代移动通信技术)智能节点后续投入生产后的高效地震采集工作,利用少量自制训练数据创新性地传统算法与深度学习相结合,提出障碍物自动提取方法。搭建 SegNet 网络,得到粗糙的语义分割结果,利用条件随机场、特征与空间概率融合等概率图模型依次做了边界平滑和噪声消除的优化处理,最终的语义分割结果在各类障碍物上的准确率较高。与单一使用深度网络需要数万级数据来提升泛化能力相比,提出在少量训练集条件下具有较强泛化能力的语义分割方法,从而能够低成本、灵活高效地运用到各种特定的语义分割场景中。

关键词:深度学习;条件随机场;语义分割;障碍物

中图分类号:TP183 **文献标志码:**A **文章编号:**1671-1807(2023)17-0266-07

在石油勘探工作中,选定勘探工区后,要进行观测网设计、车载导航、道路规划等系列工作,这些工作都与工区实际环境相关,需要对无法穿过的物体进行合理的位置偏移,因此在设计过程中需要获取工区各种障碍物的位置信息。传统的工区障碍物位置确定利用软件进行人工获取识别,而人工确定障碍物的工作是不可重复利用的,即需要针对每一个勘探工区都进行相同的工作,对之后的新勘探工区没有再次利用的价值。另一方面,随着无线地震采集节点的使用以及第 5 代移动通信技术(5G)智能节点的研发及后期的市场化,地震采集工作将变得高密度化、高效率化,人工确定工区障碍物分布的方法远远跟不上无线节点的工作效率需求。综上两个方面,有必要研究高效率的工区障碍物识别方法。

人工矢量化工区障碍物是一项大量的机械性且重复的工作,而目前的人工智能在很多领域如自然语言处理、计算机视觉等取得了很好的效果,大大提高相应领域的生产效率^[1-3]。工区障碍物识别主要是需要针对工区的卫星图片辨识出建筑、道路、水体等地震采集节点不能布设的地理位置和所在范围,本质上是对获取的工区卫星图片进行语义分割,分割的任务就是提取像素意义上的障碍物。语义分割算法分为传统图像分割算法和基于卷积

神经网络的语义分割算法两大类^[4]。传统图像分割算法中,图像背景与目标的灰度有明显差异时适合选取阈值法进行分割^[5],图像不同区域灰度或者 RGB 值发生明显突变时适合选取基于边缘的分割方法^[6],即传统的图像分割算法都有其特殊使用情景。而基于卷积神经网络的语义分割算法经过多场景训练集的训练能够具有更强的泛化性能。在文献[7]中,全卷积神经网络(full convolutional neural network,FCN)作为深度学习语义分割第一个网络模型,首次将端到端的卷积神经网络推广到语义分割,设计反卷积层进行上采样,并提出跳跃联合即将对应像素点在深层和浅层处的特征进行求和来进行上采样,但 FCN 的参数太多,占用大量计算机内存,训练比较困难。经典的语义分割网络 U-Net^[8]在 FCN 的基础上进行了改进,主要是用于医学图像上的分割,同样采取了跳跃性连接,但与 FCN 不同的是,U-Net 网络是将深层和浅层对应的特征信息直接进行粘贴联合,能在医学小数据集上取得较好的效果。另一个经典的语义分割网络是 SegNet^[9]网络,与 U-Net 网络结构类似,整体都是前半部分编码后半部分解码,只是没有形成“U”形,该网络不仅复制了编码器的特征,还复制了最大池化指数,使得该网络在内存使用上比前两者更高

收稿日期:2023-04-18

基金项目:中国石油化工股份有限公司科技部项目(P21062-1)。

作者简介:胡敏(1994—),女,河南信阳人,中石化石油物探技术研究院有限公司,助理工程师,硕士,研究方向为深度学习在地震勘探领域的应用;陈楠(1980—),男,四川泸州人,中石化石油物探技术研究院有限公司,高级工程师,研究方向为物探软件开发;毕进娜(1986—),女,山东威海人,中石化石油物探技术研究院有限公司,工程师,硕士,研究方向为物探软件开发。

效。因此 SegNet 从工程方向上考虑是一个比较实用的深度网络。在已有的涉及图像处理的实际工作中,如文献[10]为弥补网络对建筑物特征信息利用的不足将局部二值模式纹理特征与 SegNet 进行结合,文献[11]利用 SE(squeeze-excitation)注意力机制改变特征图权重系数来提高 SegNet 编码能力,文献[12]提出了基于 SegNet 多尺度特征融合的云图分割方法从而提高了云图语义分割任务的泛化能力等工作中 SegNet 都具有良好的指导意义。

综合前述已有的工作不难发现,SegNet 虽然具有很强的工程使用价值,但是都需要针对不同的问题情景结合相应的预处理或者后处理算法,才能得到更好的使用效果。采用 SegNet 网络进行前端的语义分割,但是仅仅依靠深度网络模型达到良好的语义分割效果,需要数万级甚至数十万级的训练数据^[13-14],即在本文的实际任务中需要制备大量的由所用地图软件平台获取的卫星图片标签,针对每种风格的卫星图片要制备的标签数量级达数十万,这与本文的研究目的(节约成本、提高效率)是相违背的,因此希望基于少量样本就能获取良好的语义分割效果,可以结合概率图模型^[8]对结果进行优化处理。

1 基本理论

深度网络学习领域里对图像进行像素级别的语义分割常用的是全卷积神经网络 FCN,以及 FCN 的变种网络。这些网络都有卷积层下采样提取特征和反卷积层上采样恢复原图尺寸的操作,导致图像的部分特征损失,因此语义分割结果出现不同类别物体边界模糊的问题。而深度学习和概率图模型结合是一个趋势,语义分割领域的通用框架^[15-16]是:前端使用 FCN 进行语义分割,得到粗略分割结果,后端使用条件随机场(conditional random fields,CRF)或马尔科夫随机场(Markov random field,MRF)对深度网络的输出结果进行优化,得到较为清晰的分类界限。本文中除了前端深度网络和后端 CRF 组合使用,在此结果上再基于区域连通去除小面积分割噪点,但这个操作会留下未被分类的像素点孔洞,因此利用支持向量机(support vector machine,SVM)得到关于特征的概率与 SegNet 得到的标签空间位置概率进行融合的方法来对孔洞像素点类别分类。本节主要对使用的 CRF、特征与空间概率融合法进行详细介绍。

1.1 全连接条件随机场(CRF)

图像的每个像素 i 基于 SegNet 的结果被标记

为类别标签 x_i ,原图像素点 i 的像素值和空间位置记为该点观测值 y_i ,考虑由所有像素点构成的无向拓扑图,则每个像素点是无向图的节点,再将相邻像素两两建立边联系,此时构成图 1 所示的全连接条件随机场。

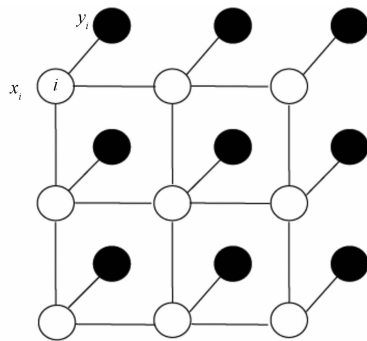


图 1 全连接条件随机场

图像所有像素标签记为 X ,原始图像为全局观测记为 Y ,条件随机场符合吉布斯分布,即 $(X,Y) \sim \text{Gibbs}$:

$$P(X | Y) = \frac{1}{Z(Y)} e^{[-E(X|Y)]} \quad (1)$$

式中: $P(X|Y)$ 为在图像原始观测到的像素值信息的条件下每个像素点进行标签标记的分类结果概率也即条件概率; $E(X|Y)$ 为状态标签 X 的 Gibbs 能量的计算公式; $Z(Y)$ 为归一化因子^[17]。Gibbs 能量的计算公式为

$$E(X | Y) = \sum_i \phi_u(X_i | Y_i) + \sum_{i < j} \phi_p(X_i, X_j | Y_i, Y_j) \quad (2)$$

式中:一元势函数 $\phi_u(X_i | Y_i)$ 由 SegNet 网络结果确定,与像素点 i 在像素特征 Y_i 下标签为 X_i 的概率成反比;二元势函数^[18] $\phi_p(X_i, X_j | Y_i, Y_j)$ 由原图像的像素值和位置信息等特征确定,用来衡量任意两个像素点对在给定相应的观测值下标签为 X_i 和 X_j 的代价,使得原图特征相似的像素点有更大的可能性被分为相同的标签。Gibbs 能量则是衡量当前分割结果的代价,代价越小 Gibbs 概率越大,返回的语义分割结果使得吉布斯概率最大。

1.2 特征与空间概率融合分类法

在 CRF 优化处理基础上,依然存在一些错分的小面积部分,经过连通域处理后小面积部分处于未被指定标签的状态,借鉴高光谱遥感图片中基于少许先验标签的空间概率和光谱特征支持向量机(SVM)概率融合的语义分割法^[19],此时可以基于这些未被分类像素点在 SegNet 已分类结果的空间

信息对应的概率和由 SVM 输出的属于每个类别的特征概率进行融合确定标签类别。设有 L 个类别, 以某个未标注的像素点为中心开统计窗 w , 并设该像素点的标签变量 X_k 取值为 x_k , $x_k \in \{1, 2, \dots, L\}$, 空间概率为

$$p_{w_{x_k}}^{spa} = \frac{p(X_w = x_k)}{\sum_{\gamma=1}^L p(X_w = \gamma)} = \frac{N_{x_k}}{N_w} \tag{3}$$

式中: $P_{w_{x_k}}^{spa}$ 为该点为类别 x_k 的概率; N_{x_k} 为窗 w 内已标注且标签为 x_k 的像素点个数; N_w 为窗内所有像素点个数, 则该像素点为不同标签的概率为 $p_w^{spa} = \{p_{w_1}^{spa}, p_{w_2}^{spa}, \dots, p_{w_L}^{spa}\}$ 。将该像素点经 SVM 得到的类别概率视为特征概率记为 $p_w^{fig} = \{p_{w_1}^{fig}, p_{w_2}^{fig}, \dots, p_{w_L}^{fig}\}$ 。将该像素点的两部分概率进行融合得到 $p_w = \{p_{w_1}, p_{w_2}, \dots, p_{w_L}\}$, 其中 $p_{w_{x_k}} = p_{w_{x_k}}^{spa} p_{w_{x_k}}^{fig}$ 。通常对 $p_{w_{x_k}}$ 取对数得 $\lg p_{w_{x_k}} = \lg p_{w_{x_k}}^{spa} + \lg p_{w_{x_k}}^{fig}$, 本文对 SegNet 已得分类结果赋予较高的权值 0.9, 同时对 SVM 得到的特征概率赋予较小的权值 0.1, 此时有 $\lg p_{w_{x_k}} = 0.9 \lg p_{w_{x_k}}^{spa} + 0.1 \lg p_{w_{x_k}}^{fig}$ 。最大融合概率的对数对应的标签即为像素点 w 的标签, 即

$$x_w = \operatorname{argmax}_k (\lg p_{w_{x_k}}) \tag{4}$$

2 Segnet 语义分割实现

2.1 数据集准备

为了利于后期模型在生产上应用, 选取当前试验阶段使用的奥维地图软件获取卫星图片自制数据集。由于人力有限, 本文中抓取了尺寸大小为 $6\,912 \times 2\,880$ 、分辨率为 18 m 的 RGB 图像(图 2), GPU 限制下需要将图像切割成小图, 这里将切割成 256×256 的小图, 共 297 张, 挑选其中的 16 张回拼成大图用来做测试集, 其余 281 张图片作为原始训练集。使用 Labelme 对 297 张原图进行 5 个类别标注, 分别是房屋(building 标记为 1)、道路(road 标记为 2)、水域(water 标记为 3)、植被(green 标记为 4)及背景(background 标记为 0), Labelme 每标注完一张图片自动生成相应的由边界点坐标组成的 json 文件, 随后编写 Python 语言脚本将 json 文件批量转换成训练使用的灰度标签, 图 3 可视化显示了其中 2 张标签。这里对测试集部分也进行标签标注, 方便之后进行预测效果指标的统计。

2.2 数据增强

由于图像打标签消耗大量人力, 短时间内训练标签数量不足, 为了防止数据过少造成模型过拟合, 需要提高训练数据集丰富性以提高模型的泛化能力。



图 2 奥维地图获取的 tif 图片

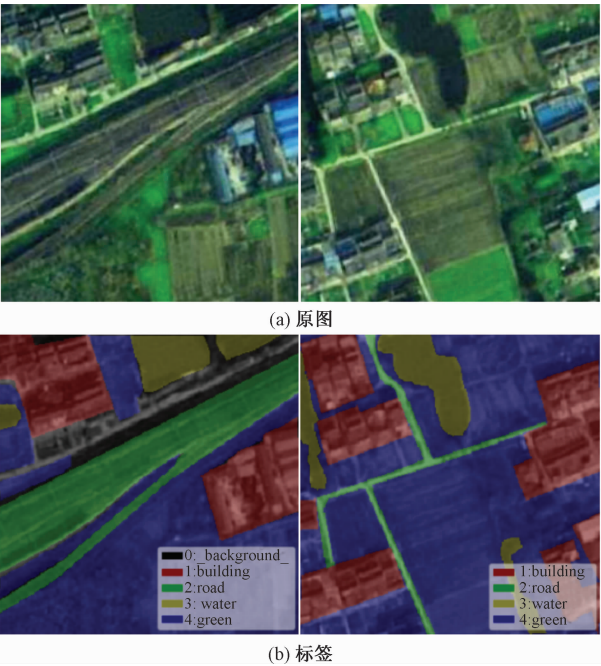


图 3 部分标签制作效果示意

这里对原始训练集和每一对图像和相应的标签做数据增强, 包括旋转角度、调整亮度、调整对比度、调整锐度、调整色度和翻转。经过数据增强之后, 训练集数量扩充到 2 529 张, 数据增强效果如图 4 所示。

2.3 搭建 SegNet 网络

本文使用的 SegNet 网络结构如图 5 所示。输入层是 3 通道的 256×256 图像, 网络分为编码和解码两部分, 前半部分为编码层, 通过多次卷积和池化进行不同深度的特征提取, 后半部分为解码层, 通过多层反卷积和上采样进行原图尺寸恢复。输出层是经过 Softmax 层得到像素点属于各个类别的概率值, 像素点的标签识别为最大概率值对应的类别。搭建的该网络用 2020-CCF-BDCI 提供的 10 万级数据集 CamVid 进行测试, 网络损失率和正确率稳定且分别收敛到 0.002 以下和 0.900 以上。

2.4 模型参数设置

对 281 张原始训练集对应的所有标签进行按类别逐像素统计, 结果见表 1。由统计结果可见, 考虑

的 4 个主要类别中道路的占比严重失衡,该分布下的数据集训练出来的模型很难将道路正确识别出来。因此需要对占比取倒数来作为训练模型时权重参数输入,即 $class_weight = \{129.87, 4.32, 15.04, 4.55, 2.19\}$ 。

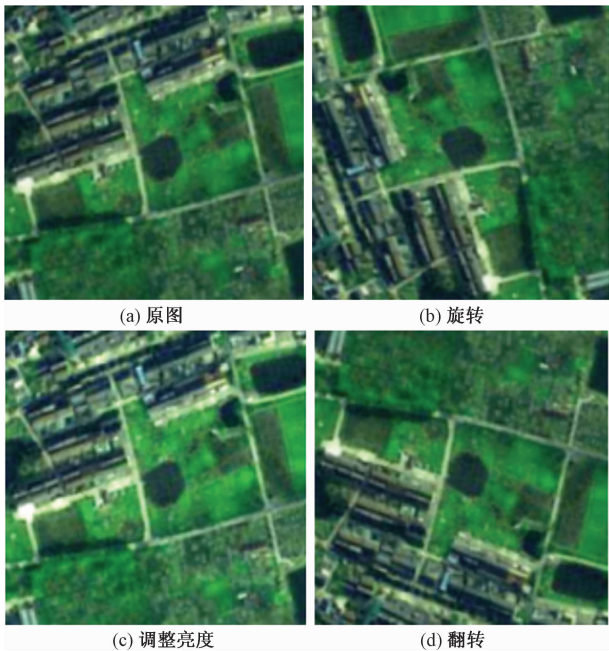


图 4 部分数据增强

表 1 数据集像素类别分布

类别	标签	占比/%
背景	0	0.77
房屋	1	23.16
道路	2	6.65
水域	3	21.99
植被	4	47.42

将增强后的 2 529 对数据集和标签按照 8:2 的比例随机分为训练集和交叉验证集,由于计算机 CPU 算力限制, $batch_size$ 最大设置为 10,则 $steps_each_epoch$ 设置为 202,经过多次实验, $epoch$ 设置为 300、loss 函数选择 $mean_squared_error$ 、优化器选择 $adam$ 时训练和测试效果较好。

2.5 模型效果

搭建的 SegNet 网络在上述数据集和参数设置下训练并且收敛。训练集上损失 $loss=0.0012$,准确率 $accuracy=0.99$;交叉验证集上损失 $loss=0.2420$,准确率 $accuracy=0.93$ 。分别对训练测试集和测试集进行模型预测,为了减少图片切割对边界预测的影响,先对输入的原图进行膨胀填充,然后以滑动步长为 128 将图片切割为 256×256 的小图进行预测,每次取小图预测结果的中心部分 128×128 ,最后得到 SegNet 网络预测结果以预测集为例,如图 6 所示。

3 语义分割结果优化

3.1 道路二分类结果与全分类结果融合

由 SegNet 结果发现,测试集上分类边界不清晰,且对道路这一类别的识别较差,因此需要对 SegNet 的结果进行优化处理。如果直接进行 CRF 处理会将面积较小但识别正确的道路平滑为其他类别,因此需要在 CRF 优化之前先对道路做二分类提取,然后对道路二分类的结果进行膨胀处理,再将二分类的结果覆盖在 SegNet 的五分类结果上,作为下一步 CRF 处理的输入。系列处理如图 7 所示。

3.2 基于 CRF 的优化处理

将上述二分类与五分类融合的结果即图 7(d)作为 CRF 的状态输入,图 6(c)作为 CRF 的观测输

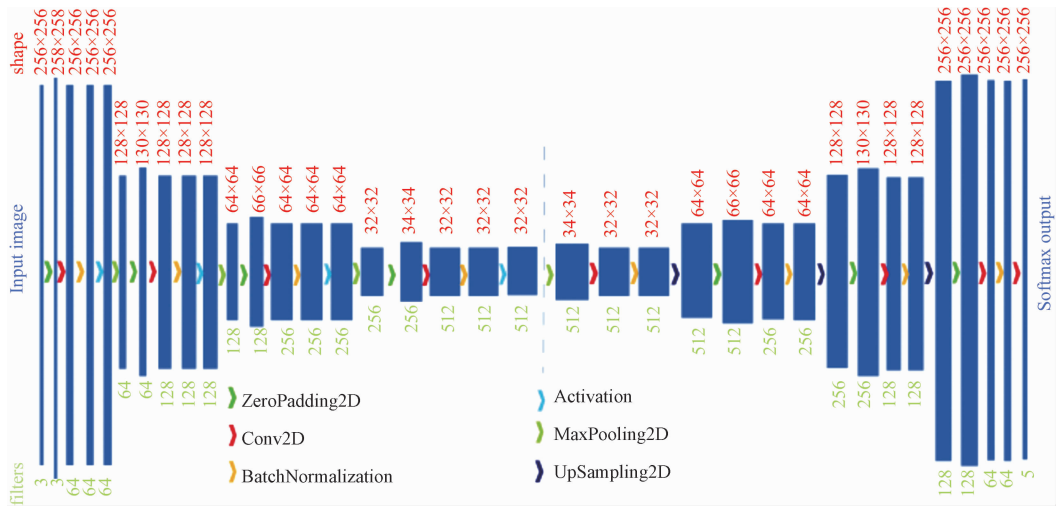


图 5 SegNet 网络结构

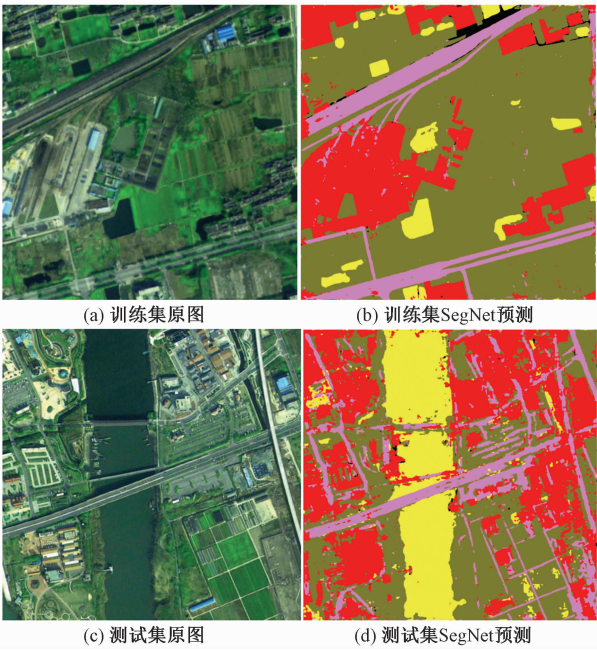


图 6 SegNet 在训练集和测试集上的预测效果

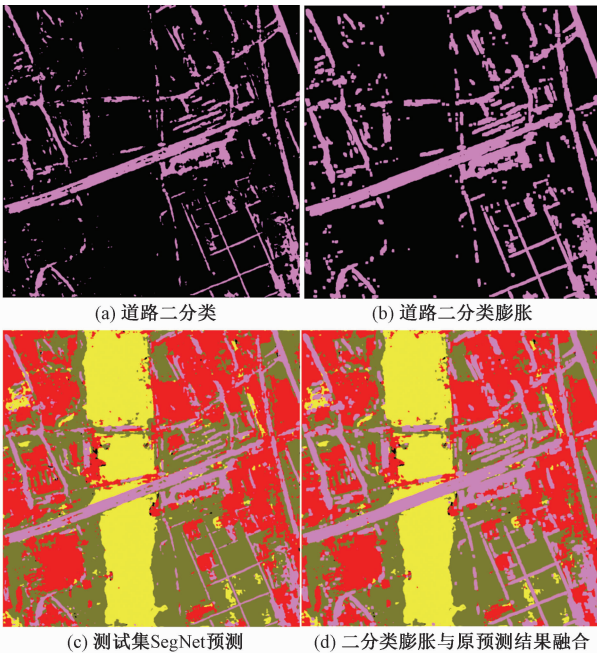


图 7 道路二分类膨胀及与 SegNet 语义分割的融合

入,构成线性条件随机场。将标签状态概率置为 0.7,全局观测原图的 RGB 和位置信息概率置为 0.3,即原图信息在原有结果上对语义分割进行调整。CRF 处理前后效果如图 8 所示,不同类别的分界线较处理之前清晰,但依然存在很多错误分类的小面积噪点。

3.3 基于特征与空间融合的优化处理

定义四邻域连接即认为属于相同连通区域,设置连通区域阈值为 150 像素点,将 CRF 处理之后的

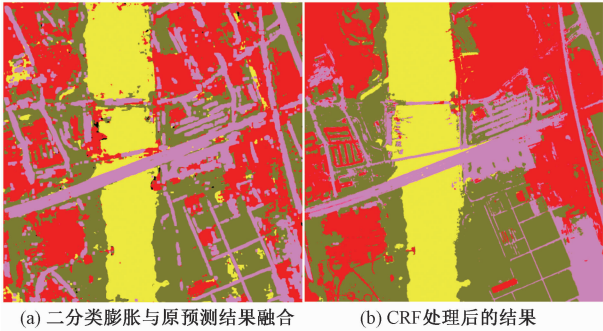


图 8 CRF 处理前后效果对比

语义分割结果中连通区域面积小于 150 像素点的区域删除,得到图 9(b)的结果,留有删除标签之后的未标记孔洞(黑色部分)。度量孔洞处像素点与周围已经被标注的像素点之间 RGB 特征、空间坐标相似性,即基于特征与空间融合的方法对孔洞处像素点进行标签识别,效果如图 9(c)所示,可见比消除杂波之前分割更加干净。

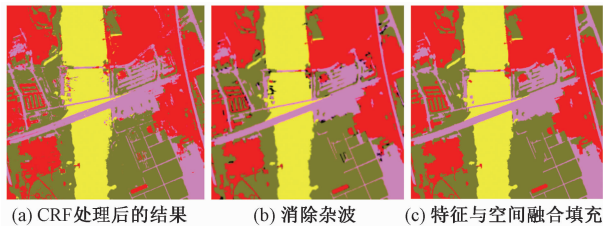


图 9 特征与空间概率融合前后效果对比

4 指标评价

最终的测试集预测结果如图 10(c)所示,测试集原图与最终预测结果的融合如图 10(d)所示。从直观上看预测结果与图 10(b)实际标签较吻合,但评价语义分割效果需要从量化的指标进行,主要用准确率和交并比来衡量。每个类别对应的指标见表 2,水域和植被在准确率和交并比两个指标上表现均很高,道路的交并比最低,说明将其他类识别为道路的占比和将道路识别为其他类别的占比都比较高,这与来源的数据中道路类别占比较低有关,而且从图 10(c)最终预测结果可以发现房屋和道路相互之间被错误分类,这与原图中某些道路段与房屋具有相似或相近的 RGB 特征有关。

表 2 测试集最终语义分割指标

类别	准确率	交并比
房屋	0.832 1	0.780 8
道路	0.867 8	0.667 7
水域	0.938 0	0.889 7
植被	0.890 5	0.787 8

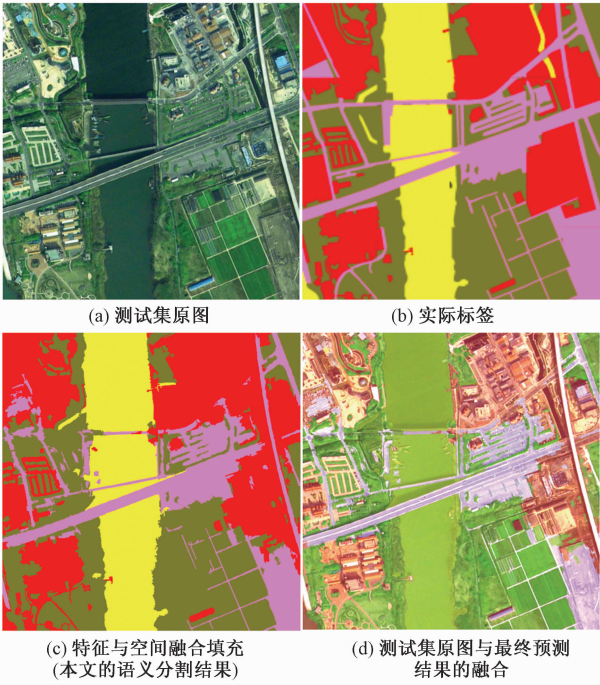


图 10 SegNet 语义分割之后优化处理的效果

5 结论

提出了基于 SegNet 网络并结合条件随机场、特征与空间融合等多种后处理方法进行地形地貌图片的语义分割方法。由于该方法后期会用于生产实践中,因此从生产实践中应用的卫星图片软件获取原图,并使用 Labelme 进行标签制作,完成了 297 张原始图片标签标注工作。在仅有的 281 张训练数据的条件下,概率图模型处理之后的语义分割模型对测试集各类预测准确率都较高,依次有房屋 0.832 1、道路 0.867 8、水域 0.938 0 以及植被 0.890 5,对工区障碍物能进行高效且有效分割提取。因此本文的方法可达到在少量标签下具备较强的泛化能力的效果,只需要少量标签便可以获取较好的语义分割效果,也即能够低成本、灵活地运用到任何特定的语义分割场景中。

参考文献

[1] 侯晓辉,李莉,孙红凯,等.并行多尺度特征增强的小目标检测[J].科技与产业,2023,23(5):178-188.
[2] 杨轲,董兵,吴悦,等.基于改进 YOLOv3 的机场盲区遥感图像目标检测方法[J].科技与产业,2023,23(4):213-218.
[3] 杨利萍,蒋鑫,马跃.基于深度学习的变电站机器人仪表检测研究[J].科技与产业,2021,21(11):333-338.
[4] 何雨岩.基于深度学习的图像语义分割综述[J].长江信息通信,2023,36(1):77-79.
[5] 刘超,蔡文华,陆玲.图像阈值法分割综述[J].电脑知识与技术,2015,11(1):140-145.

[6] 郭锁利,辛栋,刘延飞.近代图像分割方法综述[J].四川兵工学报,2012,33(7):93-96.
[7] 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.
[8] RONNEBERGER O, FISCHER P, BROX T. U-net; convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention- MICCAI 2015: 18th International Conference. Berlin: Springer International Publishing, 2015: 234-241.
[9] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
[10] 谢跃辉,李百寿,高豫川.基于 LBP 纹理 SegNet 网络的灾损建筑物提取[J].北京测绘,2023,37(3):397-401.
[11] 赵江洪,尹利莎,陈先吴,等.一种基于图卷积的车载视频对象语义分割方法[J].测绘科学,2023,48(2):157-167.
[12] 吉茹,张银胜,杨宇龙,等.基于多尺度特征融合的改进型云图分割方法[J].国外电子测量技术,2022,41(11): 37-44.
[13] 徐昭洪,刘宇,全吉成,等.基于 VGG16 预编码的遥感图像建筑物语义分割[J].科学技术与工程,2019,19(17): 250-255.
[14] 代具亭,汤心溢,刘鹏,等.基于 RGB-D 图像和深度学习的场景语义分割网络[J].科学技术与工程,2018,18(20):286-291.
[15] 林娜,王玉莹,郭江,等.基于 SegNet-CRF 的遥感影像水体提取方法[J].测绘与空间地理信息,2023,46(3): 12-16.
[16] 赵昊罡,崔红霞,张芳菲,等.改进 SegNet+CRF 高分辨率遥感影像建筑物提取方法[J].计算机测量与控制, 2023, doi: <https://kns.cnki.net/kcms/detail/11.4762.TP.20230309.1724.043.html>.
[17] LAFFERTY J, MCCALLUM A, PEREIRA F C N. Conditional random fields; Probabilistic models for segmenting and labeling sequence data[C]//Proceedings of the Eighteenth International Conference on Machine Learning. San Francisco: Morgan Kaufmann Publishers Inc, 2001: 282-289.
[18] KRÄHENBÜHL P, KOLTUN V. Efficient inference in fully connected CRFs with Gaussian edge potentials [C]//Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011, Proceedings of a Meeting Held 12-14 December 2011. Granada: NIPS, 2011: 109-117.
[19] LIU J, LU W. A probabilistic framework for spectral-spatial classification of hyperspectral images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(9): 5375-5384.

Obstacle Extraction for Exploration Areas Based on SegNet and Probabilistic Graphical Model

HU Min, CHEN Nan, BI Jinna

(Sinopec Geophysical Research Institute, Nanjing 211100, China)

Abstract: In order to adapt to efficient seismic acquisition after 5G (fifth-generation mobile communication technology) intelligent node being put into production, the scheme of automatically extracting obstacles has been proposed by using a little self-made training data and innovatively combining traditional algorithms with deep learning. A SegNet network has built, obtaining a rough segmentation result. Optimization processing including boundary smoothing and noise elimination, which achieved by utilizing probabilistic graphical model-conditional random fields and feature-space probability fusion respectively. Finally, the segmentation has attained higher accuracy rates on each class of obstacles. Compared with the deep network single using which needs tens of thousands of levels of data to improve the generalization ability, a semantic segmentation method has proposed with strong generalization ability under the condition of a little training data, which can be low-costly, flexibly and efficiently used in various specific semantic segmentation scenarios.

Keywords: deep learning; conditional random fields; semantic segmentation; obstacles