

UNET与FPN相结合的遥感图像语义分割

王曦, 于鸣, 任洪娥

引用本文:

王曦, 于鸣, 任洪娥. UNET与FPN相结合的遥感图像语义分割[J]. 液晶与显示, 2021, 36(3): 475–483.

WANG Xi, YU Ming, REN Hong-e. Remote sensing image semantic segmentation combining UNET and FPN[J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(3): 475–483.

在线阅读 View online: <https://doi.org/10.37188/CJLCD.2020-0116>

您可能感兴趣的其他文章

Articles you may be interested in

Faster-RCNN和Level-Set结合的高分遥感影像建筑物提取

High-resolution remote sensing image building extraction combined with Faster-RCNN and Level-Set

液晶与显示. 2019, 34(4): 439–447 <https://doi.org/10.3788/YJYXS20193404.0439>

基于Faster R-CNN的仪表识别方法

Instrument recognition method based on Faster R-CNN

液晶与显示. 2020, 35(12): 1291–1298 <https://doi.org/10.37188/YJYXS20203512.1291>

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>

基于深度学习的木材表面缺陷图像检测

Image detection of wood surface defects based on deep learning

液晶与显示. 2019, 34(9): 879–887 <https://doi.org/10.3788/YJYXS20193409.0879>

基于纹理抑制和连续分布估计的显著性目标检测方法

Significant target detection method based on texture inhibition and continuous distribution estimation

液晶与显示. 2015, 30(1): 120–125 <https://doi.org/10.3788/YJYXS20153001.0120>

文章编号:1007-2780(2021)03-0475-09

UNET 与 FPN 相结合的遥感图像语义分割

王 曦¹, 于 鸣^{1,2*}, 任洪娥^{1,2*}

(1. 东北林业大学 信息与计算机工程学院, 黑龙江 哈尔滨 150040;

2. 黑龙江省林业智能装备工程研究中心, 黑龙江 哈尔滨 150040)

摘要:针对传统的遥感图像分割方法效率低下,复杂场景下分割精细度不够,以及 UNET 模型对于图像中包含的较小目标以及较大目标的边缘分割效果不佳等问题,本文提出了一种 UNET 结构与 FPN 结构相结合的方法,提升 UNET 模型整合多尺度信息的能力,同时辅以最能够更好地捕捉目标边缘的 BLR 损失函数,提升 UNET 模型对目标边界的分割效果。实验结果表明,本文所使用的方法有效提升了语义分割的精度,较好地缓解了对小尺度目标和大尺度目标边缘分割不佳的问题。该方法对目标边缘分割更精准,达到更好的分割效果。

关键词:深度学习;UNET;FPN;BLR

中图分类号:TP751.1 **文献标识码:**A **doi:**10.37188/CJLCD.2020-0116

Remote sensing image semantic segmentation combining UNET and FPN

WANG Xi¹, YU Ming^{1,2*}, REN Hong-e^{1,2*}

(1. College of Information and Computer Engineering,
Northeast Forestry University, Harbin 150040, China;

2. Heilongjiang Forestry Intelligent Equipment Engineering Research Center,
Harbin 150040, China)

Abstract: The traditional remote sensing image segmentation method is inefficient and the segmentation fineness is not enough in complex scenes. The UNET model is well-known for its good segmentation effect, but it does not perform well for the smaller objects contained in the image and the edge segmentation of larger objects. In order to solve this problem, a method combining UNET structure with FPN structure is proposed in this paper to improve the ability of UNET model to integrate multi-scale information. At the same time, the BLR loss function which can better capture the edge of the target edge is used to improve the segmentation effect of UNET model on the target boundary. The experimental results show that the method used in this paper effectively improves the accuracy of semantic segmentation and alleviates the problem of poor edge segmentation of small-scale targets and

收稿日期:2020-08-07;修订日期:2020-09-04.

基金项目:中央高校基本科研业务费专项资金(No.2572017PZ10)

Supported by Central University Basic Research Funding for Specialized Projects Funded (No. 2572017PZ10)

* 通信联系人, E-mail: yuming@nefu.edu.cn; nefu_rhe@163.com

large-scale targets. The target edge segmentation can be more accurate to achieve better segmentation results.

Key words: deep learning; UNET; FPN; BLR

1 引言

卷积神经网络(Convolution Neural Networks, CNN)^[1]以其局部权值共享的特殊结构以及良好的容错能力、并行处理能力和自学习能力被广泛应用于语义分割^[2]、对象检测^[3]、人脸识别^[4]、图像分类^[5]等诸多计算机视觉领域^[6]。以图像语义分割为例,Long J 和 Shelhamer E 首次将 CNN 引入了图像语义分割问题,提出了著名的 FCN^[7]结构。FCN 将深度学习扩展到语义分割领域。FCN 在提取特征方面与 CNN 基本一致。不同的是,FCN 为每一个像素分配一个值,用于表示该像素属于特定类别的概率。相比使用手工设计的特征描对图像进行编码,通过反卷积实现上采样将 CNN 获取的编码恢复到原有图像尺寸,采用端到端的训练方式,简化了语义分割问题过程,得到了更优的准确率。Ronneberger O 和 Fischer P 等人在 FCN^[8]基础上提出了 UNET 结构,使用了编码器解码器结构并加入了横向连接, f 反卷积使得从图像特征编码到最后的预测结果更加平滑,横向连接为解码器增加了来自浅层具有局部信息的特征,丰富了解码器阶段包含的信息,进一步提升了效果。Chen L C 和 Papandreou G 提出了 DeepLab V3 模型,在模型中对空间金字塔池化模块(ASPP)进一步改进,使用具有不同感受野的空洞卷积提取多尺度语义信息,通过 1×1 卷积进行整合,提升了模型整合多尺度信息的能力,将语义分割的效果带上了新台阶。

UNET 模型最早用于医学影像分割,该模型以其简洁的结构以及显著的效果受到了国内外学者的关注。李忠智等人提出基于 UNET 的船舶检测方法,通过在 UNET 中加入多变输出融合策略算为基础在模型中加入特征融合^[9];周锐烨等人提出 PI-UNET 用于异质虹膜精确分割,使用深度可分离卷积替换传统卷积对模型进行加速,在上采样阶段同时使用反卷积与二次先行插值,弥补因深度可分卷积带来的精度问题^[10];蒋宏达

等人提出 I-UNET 用于皮肤病图像分割,在 UNET 中加入 Inception 提升 UNET 对多尺度信息的捕捉能力^[11]。本文将对 UNET 模型出现的问题进一步探究,通过两种方式对其进行改进,提升模型的语义分割精度。

高分辨率遥感影像语义分割^[12]问题是语义分割的一个重要分支。梁华等人^[13]利用深度学习技术对航空对地小目标进行检测。对遥感影像精确的分割以及对建设数字型、智慧型城市有着重要意义。语义分割是将图像中属于相同类别的像素聚类为一个区域,对图像更加细致地了解^[14]。本文在语义分割^[15]基础上,探究语义分割模型在遥感影像分割上的应用。针对分割过程中存在的图像边缘分割精细程度不佳的问题,本文提出了一种改进 UNET 模型,通过加入具备整合多尺度信息的 FPN 结构,丰富每一个像素点在分类时所需要的信息。同时使用能够更好捕捉目标边缘^[16]的损失函数,缓解上述模型在目标边缘分割不佳的问题。实验证明,在加入上述两种改动后,模型对目标边缘的分割精细程度有所提升,得到了更优的分割效果。

2 研究技术与方法

UNET 是通过反卷积与通道维度连接进行整合;反卷积上采样与通道连接的优点在于可以得到更加平滑的结果,但是经过 conv + relu 的非线性变换,通过该层传递的浅层信息传递到下一层的量将会减少。我们希望每一层都可以像 ASPP 那样尽可能多的掌握不同尺度语义信息,很显然,UNET 的方式不能满足对图像中目标边缘附近区域精准分割。所以本文主要从以下两个方面着手进行改进:

(1)通过在 UNET 中加入 FPN 结构,提升 UNET 整合多尺度语义信息的能力,丰富其像素点标签分类所使用的特征中包含的信息。

(2)对损失函数进行改进,相比对每个点独立进行预测的多分类交叉熵,本文使用了能够更好地捕捉目标边缘的边界标签松弛损失函数。

2.1 改进的 UNET 模型

为提升 UNET 模型的分割效果,本文对 UNET 模型的整体结构进行了改进,在 UNET 结构中引入 FPN 结构,充分利用 UNET 编码器中包含多尺度信息的优势。

FPN 结构又称特征金字塔网络,这是一种广泛应用于目标检测任务中的结构,FPN 结构能有效整合来自编码器部分的多尺度语义信息。FPN 主要由自底向上流程、自顶向下流程和横向连接 3 部分构成,如图 1 所示。

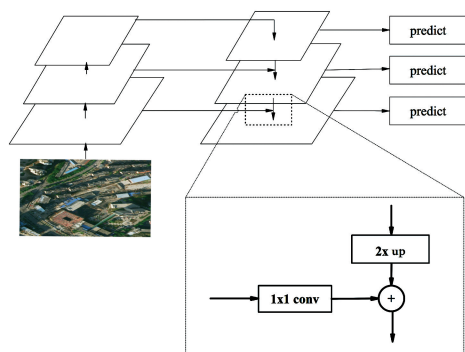


图 1 FPN 网络结构

Fig.1 Diagram of FPN network structure

自底向上流程即为卷积神经网络在前向计算时提取具备不同尺度特征的过程。自顶向下流程即对前述所提取每步卷积特征进行上采样处理,保证处理后的高层卷积特征维度与自底向上的低层特征维度相同。随后将高级特征图与横向连接传递的低层特征图对应元素进行相加,得到具备多尺度信息的融合特征。低层次特征由于经历的卷积次数相对于高层特征较少,因此具有较多的纹理信息,而高层次特征因为经过多次卷积滤波,具有更高级的语义特征。经过对高层的特征上采样至同等尺寸,与低层次特征进行相加,使用低层次特征包含的纹理和细节信息,进一步对高层特征做补充,丰富了融合后的特征图所包含的信息。

UNET 使用的反卷积上采样可得到相对平滑的结构特征,但是经过卷积加非线性激活层的变换后,使原始特征无法得到较好的保留,所以在 UNET 中通过跳层连接传入来自浅层的特征对原始特征进行弥补。图像中每个像素点的准确分类要依靠来自不同尺度的语义信息,而

UNET 的跳层连接传递的特征较为单一。因此本文引入了 FPN 结构。此外 FPN 的上采样方式为二次线性插值,相比反卷积操作,对原始特征保留更完整,加和方式特征整合保证了传递给解码器的每一层都包含尽可能多的多尺度信息。对比 FPN 结构与 UNET 结构,发现二者具有相似性,FPN 结构中的横向连接可使用 UNET 中的横向连接实现。这一特征为 FPN 结构扩展 UNET 模型提供了便利。充分利用 UNET 中结构的同时,极大地利用 UNET 模型中包含的不同尺度信息。在加入 FPN 结构后,UNET 整体结构如图 2 所示。

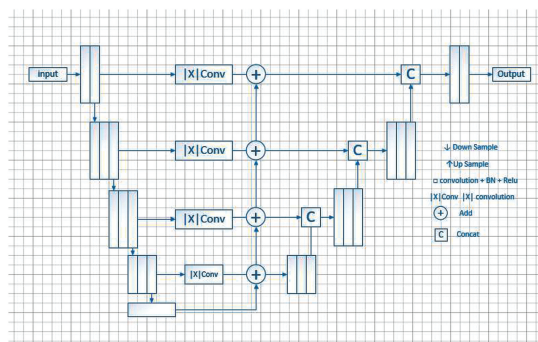


图 2 加入 FPN 结构后的 UNET 网络结构

Fig.2 Diagram of UNET network structure after joining FPN structure

FPN 是通过二次线性插值和对应元素相加的方式进行整合;而 UNET 则通过反卷积与通道维度连接进行整合。由于二次线性插值是一个线性操作,对原始的浅层信息没有更改,仅仅通过线性插值增加了特征图的尺寸,对应元素相加的方式保证了传递给每个解码器层的特征都是原始浅层特征加和的方式,避免了因反卷积导致的信息丢失,但是二次线性插值的缺点在于会引入噪声,无法获得平滑的结果,因此我们选择在保留 UNET 的基础上加入 FPN 结构。

2.2 BLR 损失函数

深度学习技术的加入快速推动了语义分割的发展,图像中目标包含的大部分像素都能够被准确分类,但对于边缘部分的分割效果依然有限,探索这一问题的本质在于,深度学习技术的主要方法论是将认为加工得到的目标理解为在一种分布上采样得到的结果,通过训练方式使模型去逼近

这种分布,我们无法做到完全准确的拟合,因此对于模型的预测就存在一些出入。此外对于目标边缘上实为两目标相交处,所以此处的像素点类别存在二义问题。针对这一问题,国内外学者的解决方法主要通过更高质量的标注数据和更加复杂的模型去提升分割效果。本文将从损失函数入手,对 UNET 模型进行改动,本文使用了能更好捕捉目标边缘的边界标签松弛损失函数^[15] (Boundary Label Relaxation, BLR) 来缓解这一问题带来的影响。将边界像素点定义为具有不同标记邻居的任何像素,如图 3 所示。

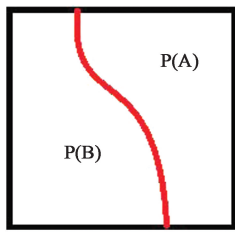


图 3 边界像素点

Fig.3 Border pixels

图 3 中红色曲线代表边缘,边缘的两侧标签分别为 A 与 B。为了产生更好边缘的预测情况,可以选择两种策略,一种方式为直接对边缘的形状进行建模,但是图像中包含的边缘形状各异,统一对边缘进行建模比较复杂。为此本文选择第二种方式,即对图像边缘两侧的区域进行最大似然估计,上文提到了边缘像素点定义为具有不同标记邻居的任何像素,因此只需要将边缘两侧的区域进行最大似然估计,便可找到边缘区域。又因边缘两侧的标签分布相互独立,因此对两侧区域的概率描述可表示为:

$$P(A \cup B) = P(A) + P(B) . \quad (1)$$

对该项进行负对数似然估计,则有:

$$-\log [P(A \cup B)] = -\log [P(A) + P(B)] , \quad (2)$$

考虑目标边缘具有多变性,为保证每一个目标边缘都能得到理想的语义分割效果,将图片切割成 3×3 大小的图像块。此时每个图像块上只包含较小部分边缘或者不包含边缘,降低了对整张图像使用梯度下降法求解时的难度。整张图片是由大小相同的图像块构成,因此对整张图片的负对数似然估计为:

$$L = \sum_{i=1}^M -\log \sum_{c \in N_i} P(c) . \quad (3)$$

该函数即为上文提到的边界标签松弛(BLR)损失函数,其中 M 代表每张图片中包含的图像块个数, N_i 代表每个图像块, c 代表图像块中包含的像素点。相比传统的交叉熵损失函数,优化降低 BLR 损失函数在保证准确预测图像中包含的每个点的类别标签的同时,提升模型对目标边缘的语义分割效果。在模型训练中只需要使用梯度下降法对上式进行最小化,便可得到比较理想的效果。

2.3 数据集与数据处理

本文选择了 CAMVID 数据集和“CCF 卫星影像的 AI 分类与识别竞赛”数据集。

(1)CAMVID 数据集:该数据由剑桥大学收集整理和标注。数据集由一组道路街景图像构成,其中包含 367 个训练图像和 233 个测试图像(白天和黄昏场景)组成,分辨率为 360×480 。挑战在于划分 11 类,如道路、建筑物、汽车、行人、标志、电线杆、人行道等。

(2)“CCF 卫星影像的 AI 分类与识别竞赛”数据集:训练集中包含 5 张尺寸为 $3\,357 \times 6\,166$ 的高清分辨率遥感图片,部分影像如图 4 所示。为了克服数据极其稀少的问题,选择充分利用图片具有的大尺寸高分辨率性质对数据集进行了扩充。首先从 5 张图片中随机选取 4 张,使用过采样策略在训练集上的每张图片上随机选点,因为原始图像的像素为 $3\,357 \times 6\,166$,其中横向像素为 3 357 个,纵向像素为 6 166 个,所以一共可以取到约为 18×10^6 个像素点。将每个随机选中的点作为图像块中的左上角,分别

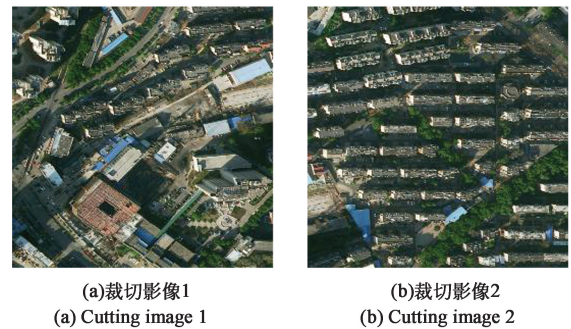


图 4 CCF 卫星影像分类与识别比赛中的裁切得到的影像

Fig.4 AI classification and CCF satellite influenced part of the images in the competition

从左上角的像素点向右和向下延伸 256 个像素点,这样就得到一个以 256 个像素为长度切割出 256×256 的图像块,使用切割得到的图像块构成训练集。将剩余的一张遥感图像使用相同的采样策略,并使用采样得到的图像块作为验证集。

本文对数据集使用了数据增强策略,采用高斯平滑、随机添加噪声点、颜色抖动、图像旋转与缩放等操作对训练集进行扩充,测试集与验证集未做数据增强。

3 实验与结果分析

3.1 训练过程

实验在 Ubuntu 16.0.4 的环境下,使用 Tesla v100GPU 作为硬件环境。为了获得更好的结果,使用了 Xavier 方法对模型中包含的参数进行初始化。在优化算法上选择了具备自适应调节步长能力的 Adam 方法,其中学习率初始值设置为 0.002 批(batch size),大小设置为 16,共迭代了 6 000 次。模型训练过程如图 5 所示。

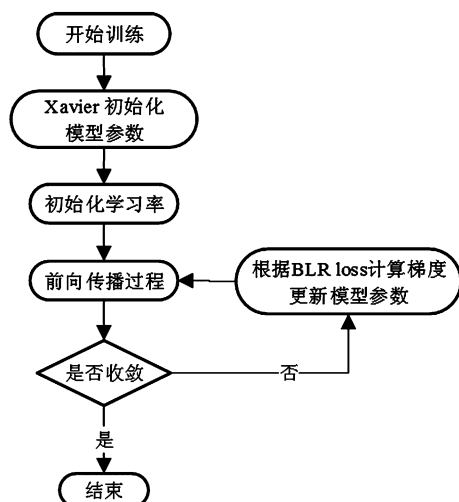


图5 模型训练过程

Fig.5 Model training process

此外,为了防止模型出现过拟合,本文在模型中使用了批归一化(Batch Norm),并且在训练阶段加入了系数为 0.000 05 的 L2 正则项,训练阶段所使用的损失函数如下:

$$L = - \sum_i^M \log \sum_{c \in N_i} P(c) + \lambda \sum |w_i|^2. \quad (4)$$

在模型推理(inference)阶段,将测试集中的图像调整为相应尺寸,依次送入训练好的改进 UNET 模型中,得到最终的预测结果。

3.2 对比实验

本文方法在上述 2 个数据集上与 FCN、SEGNET^[17]、UNET、DeepLabV1^[18]、DeepLabV2^[19]以及目前较为先进的 DeepLabV3 进行了多组实验以验证其语义分割效果。使用了常用的像素精度(PA)和均交并比(MIOU)作为衡量标准:

(1)PA 是最简单的度量,为标记正确的像素占总像素的比例,数学表达如下:

$$PA = \frac{\sum \text{TurePositive}}{\sum \text{predictConditionPositive}}. \quad (5)$$

(2)MIOU 为语义分割的标准度量。均交并比计算模型预测得到的标注图与人工标注图的交集与并集之比。在每个类上计算 IOU,之后取平均,其中 Class Number 为包含的类别。数学表达如下:

$$MIOU = \frac{1}{\text{ClassNumber} + 1} \sum \frac{\text{GroundTruth} \cap \text{prediction}}{\text{GroundTruth} \cup \text{prediction}}. \quad (6)$$

在 CAMVID 测试集上得到的分割结果如表 1 所示,不同模型耗时分析如表 2 所示。对比表 1 中的各数值,本文改进的模型具备了与 DeepLabV3 相近的能力。在平均像素精度和平均交并比上都明显超越以往的 FCN 模型、SEGNET 模型、DeepLabV1、DeepLabV2。对比表 2 各数值,本文改进后模型的消耗时间快于 DeepLabV1、DeepLabV2 和 DeepLabV3。

表 1 CAMVID 结果对比(%)

Tab.1 Comparison of CAMVID result(%)

方法	MPA	MIOU
FCN	66.9	55.9
SEGNET	71.2	60.1
Deep Lab V1	71.5	60.7
Deep Lab V2	72.3	62.5
Deep Lab V3	74.8	65.9
改进 UNET+BLR	74.9	65.6

表 2 CAMVID 不同模型耗时分析

Tab.2 Time consuming analysis of different models of CAMVID

方法	时间/ms
FCN	673
SEGNET	732
Deep Lab V1	1 210
Deep Lab V2	976
Deep Lab V3	793
改进 UNET+BLR	758

在 CCF 影像分类数据上的精度指标如表 3 所示,不同模型消耗时间如图 4 所示。对比表 3 的各个数据可以清晰地观察到在 UNET 中加入 FPN 结构后模型的平均像素精度和均交比都有提升,在引入 BLR 损失函数后模型的均交并比又一次得到了提升,得到了与 DeepLabV3 相接近的效果。对比表 4 的各个数据值,不同模型耗时同 CAMVID 测试集。

表 3 CCF 影像分类结果精确度 (%)

Tab.3 Classification result accuracy of CCF image(%)

方法	MPA	MIOU
FCN	69.1	75.7
SEGNET	73.0	79.5
UNET	72.4	78.9
DeepLabV1	73.9	79.9
DeepLabV2	74.7	80.1
DeepLabV3	76.7	85.1
改进的 UNET	74.8	83.6
改进的 UNET+ BLR	75.2	85.3

表 4 CCF 影像不同模型耗时分析

Tab.4 Time consuming analysis of different models of CCF image

方法	时间/ms
FCN	679
SEGNET	793
UNET	951
DeepLabV1	1 432
DeepLabV2	1 214
DeepLabV3	870
改进的 UNET	892
改进的 UNET+ BLR	805

在建立的两个数据集上与其他先进方法的时间对比如表 2 和表 4 所示。由表可知, DeeplabV1 模型耗时最长,造成这一现象的主要原因为条件随机场的参数量大,同时推理速度慢。相比其他方法如 FCN,SEGNET,改进的 UNET +BLR 速度稍逊一筹,原因在于改进 UNET +BLR 参数量增加,相比 DeepLabV3,本文方法具有更快的速度,同时具有相似的效果。

为了更加明显地展示加入 BLR 损失函数带来的交并比提升,本文将改进的 UNET 和加入 BLR 后的改进 UNET 分割结果进行了可视化,如图 6~9 所示。图 6(a)为原始图片,(b)为改进的 UNET 的预测结果,(c)为改进 UENT+BLR 的预测结果,(d)为人工标注的真实值。对比图 6 (b)与图 6 (c),注意到加入 BLR 损失函数后, UNET 模型使得目标边缘的分割结果得到了大幅度提升,缓解了原始 UNET 中对目标边缘分割效果不佳的问题。同时整体分割结果相比于改进的 UNET 结构,分割结果效果更佳。实验证明,在

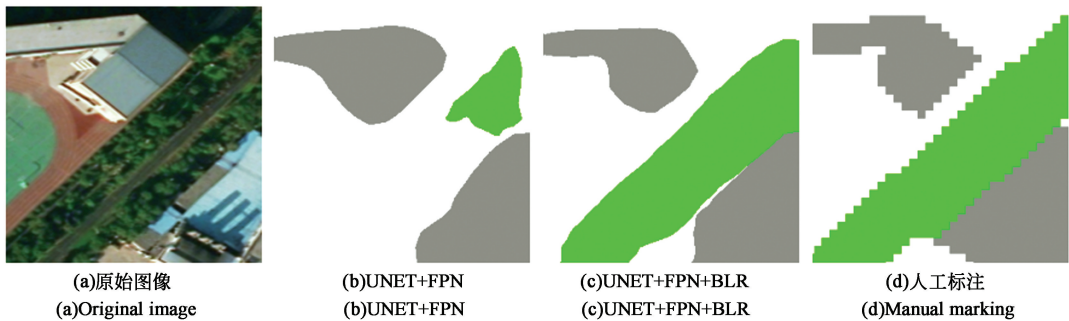


图 6 CCF 影像分割结果

Fig.6 Segmentation results of CCF image

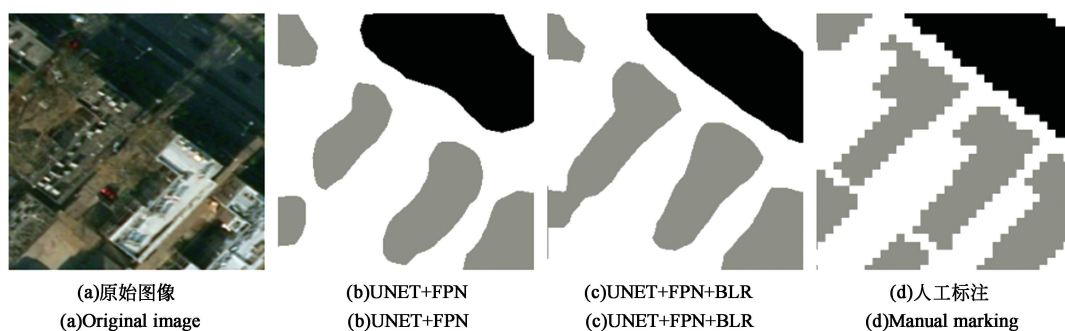


图7 CCF 影像分割结果 1

Fig.7 CCF affect segmentation results 1

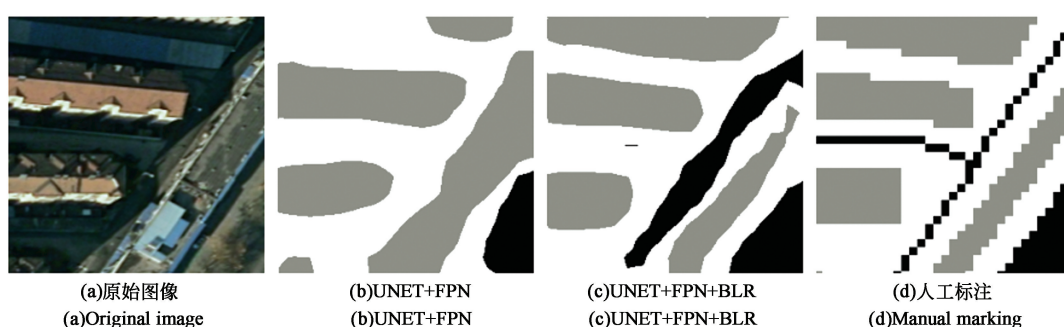


图8 CCF 影像分割结果 2

Fig.8 CCF affect segmentation results 2

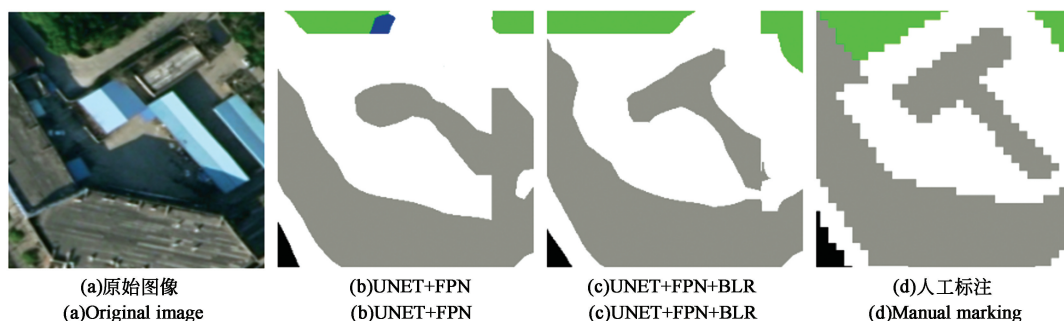


图9 CCF 影像分割结果 3

Fig.9 CCF affect segmentation results 3

增加 FPN 结构和使用 BLR 损失函数的基础上,相比于原始的 UNET 结构,本文方法不仅在均交并比和像素类别预测精度上有大幅度提升,同时使得目标边界的预测结果得到了明显改善。

4 结 论

本文对 UNET 模型在目标边缘的语义分割精细度不够理想方面做出改进,通过在 UNET 中加入 FPN 结构,提升 UNET 对多尺度语义信息

的整合能力,同时引入能更好地捕捉目标边缘的边界标签松弛损失函数 BLR。在数据量较少的条件下,使用了将大尺度的遥感图像分割为 256×256 的图像块,同时通过旋转平移、添加高斯噪声、高斯平滑、椒盐噪声等方式有效地扩充了数据量。实验结果表明,改进的 UNET 模型在提升语义分割效果的同时,使得目标边缘的分割效果也得到了大幅度改善。未来将尝试 RESNET 等更加优秀的骨干网络,或在模型中加入空洞卷积,获得性能及速度的进一步提升。

参 考 文 献:

- [1] LECUN Y, BOTTOU L, BENGIO Y, *et al.* Gradient-based learning applied to document recognition [J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [2] REN S Q, HE K M, GIRSHICK R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks [C]//*Proceedings of the 28th International Conference on Neural Information Processing Systems*. Montreal, Canada: NIPS, 2015: 91-99.
- [3] 曹玉良, 明廷锋, 贺国, 等. 基于深度学习的离心泵空化状态识别[J]. 西安交通大学学报, 2017, 51(11): 165-172.
CAO Y L, MING T F, HE G, *et al.* Cavitation state recognition of centrifugal pump based on deep learning [J]. *Journal of Xian Jiaotong University*, 2017, 51(11): 165-172. (in Chinese).
- [4] FARFADE S S, SABERIAN M J, LI L J. Multi-view face detection using deep convolutional neural networks [C]//*Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*. Shanghai: ACM, 2015: 643-650.
- [5] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]//*Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe, USA: ACM, 2012: 1097-1105.
- [6] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
ZHOU F Y, JIN L P, DONG J. Review of convolutional neural network [J]. *Chinese Journal of Computers*, 2017, 40(6): 1229-1251. (in Chinese)
- [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] LU Y, CHEN Y R, ZHAO D B, *et al.* Graph-FCN for image semantic segmentation [C]//*Proceedings of the 16th International Symposium on Neural Networks*. Moscow, Russia: Springer, 2020: 97-105.
- [9] 周锐桦, 沈文忠. PI-Unet: 异质虹膜精确分割神经网络模型的研究[J/OL]. 计算机工程与应用. (2020-08-06)[2020-09-04]. <https://www.cnki.net/KCMS/detail/11.2127.TP.20200806.1353.016.html>.
ZHOU R Y, SHEN W Z. PI-Unet: research on a neural network model for accurate segmentation of heterogeneous iris [J/OL]. *Computer Engineering and Applications*. (2020-08-06)[2020-09-04]. <https://www.cnki.net/KCMS/detail/11.2127.TP.20200806.1353.016.html>.
- [10] 李忠智, 尹航, 左剑凯, 等. 基于 UNet++ 与多边形输出融合算法的船舶检测方法[J/OL]. 计算机工程. (2020-08-07)[2020-09-04]. <https://www.cnki.net/KCMS/detail/31.1289.TP.20200806.1722.004.html>.
LI Z Z, YIN H, ZUO J K, *et al.* Ship detection method based on UNet++ and multilateral output fusion algorithm [J/OL]. *Computer Engineering*. (2020-08-07)[2020-09-04]. <https://www.cnki.net/KCMS/detail/31.1289.TP.20200806.1722.004.html>.
- [11] 蒋宏达, 叶西宁. 一种改进的 I-Unet 网络的皮肤病图像分割算法[J]. 现代电子技术, 2019, 42(12): 52-56.
JIANG H D, YE X N. An improved skin disease image segmentation algorithm based on I-Unet network [J]. *Modern Electronics Technique*, 2019, 42(12): 52-56. (in Chinese)
- [12] CHEN X L, GIRSHICK R, HE K M, *et al.* Tensormask: a foundation for dense object segmentation [C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. Seoul, Korea (South): IEEE, 2019: 2061-2069.
- [13] 梁华, 宋玉龙, 钱锋, 等. 基于深度学习的航空对地小目标检测[J]. 液晶与显示, 2018, 33(9): 793-800.
LIANG H, SONG Y L, QIAN F, *et al.* Detection of small target in aerial photography based on deep learning [J]. *Chinese Journal of Liquid Crystals and Displays*, 2018, 33(9): 793-800. (in Chinese)
- [14] 陈彦彤, 李雨阳, 吕石立, 等. 基于深度语义分割的多源遥感图像海面溢油监测[J]. 光学精密工程, 2020, 28(5): 1165-1176.
CHEN Y T, LI Y Y, LÜ S L, *et al.* Research on oil spill monitoring of multi-source remote sensing image based on deep semantic segmentation [J]. *Optics and Precision Engineering*, 2020, 28(5): 1165-1176. (in Chinese)
- [15] ZHU Y, SAPRA K, REDA F A, *et al.* Improving semantic segmentation via video propagation and label relaxa-

- tion [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, CA, USA: IEEE, 2019: 8848-8857.
- [16] CAI Z W, VASCONCELOS N. Cascade R-CNN: high quality object detection and instance segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, doi: 10.1109/TPAMI.2019.2956516.
- [17] 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.
- [18] CHEN L C, PAPANDREOU G, KOKKINOS I, *et al.* Semantic image segmentation with deep convolutional nets and fully connected CRFs [C]//*Proceedings of International Conference on Learning Representations*. San Diego, United States, 2015.
- [19] CHEN L C, PAPANDREOU G, KOKKINOS I, *et al.* DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.

作者简介:



王 曦(1995—),女,黑龙江哈尔滨人,硕士研究生,2017 年于哈尔滨师范大学获得学士学位,主要从事图像识别与智能控制方面的研究。E-mail: qq0224wx@163.com



任洪娥(1962—),女,吉林白山人,博士,教授,2009 年于东北林业大学获得博士学位,主要从事图像识别与智能控制方面的研究。E-mail: nefu_rhe@163.com



于 鸣(1981—),男,黑龙江哈尔滨人,博士研究生,讲师,2006 年于东北林业大学获得硕士学位,主要从事图像识别与智能控制方面的研究。E-mail: yuming@nefu.edu.cn