

引用格式: Liu Zhao, Zhao Tong, Liao Feifan. Research on Gap-repairing Algorithm based on Convolutional Neural Network Prediction Result[J]. Remote Sensing Technology and Application, 2021, 36(2): 247-255. [刘钊, 赵桐, 廖斐凡. 基于卷积神经网络预测结果的缝隙修复算法研究[J]. 遥感技术与应用, 2021, 36(2): 247-255.]  
doi: 10.11873/j.issn.1004-0323.2021.2.0247

# 基于卷积神经网络预测结果的缝隙修复算法研究

刘 钊, 赵 桐, 廖斐凡

(清华大学 土木工程系交通与地球空间信息研究所, 北京 100084)

**摘要:**随着卷积神经网络技术的发展,近来的研究越来越注重于准确率的提升以及语义信息的完善。其中Mask R-CNN网络是对Faster R-CNN进一步改进后的实例分割网络,在高分遥感图像地物识别具有良好的分割效果。但由于卷积神经网络只能用小瓦片图像进行训练和预测,而导致预测结果存在较大的语义信息误差。面对这种问题,提出了针对卷积神经网络预测结果缺陷的缝隙修复算法,即先使用OverlapSize算法改善预测结果与真实结果的匹配程度,再通过PostGIS数据库中的相关函数填补缝隙,使小瓦片能真正拼接成完整大图。研究及实验结果表明:该算法能够很好地改善图像语义信息,具有实用性。

**关键词:**卷积神经网络;实例分割;Mask R-CNN;缝隙修复算法

**中图分类号:**P237 **文献标志码:**A **文章编号:**1004-0323(2021)02-0247-09

## 1 引言

我国的土地幅员辽阔,土地上的地物类型多种多样,这些信息在高分辨率遥感图像上有非常清晰和直观的展示。对于应用部门而言,掌握不同地物的分布情况不仅可以了解一个城市乃至一个国家的现状及发展情况,还可以借此规划城市或国家未来的发展方向。

自从2012年Hinton等<sup>[1]</sup>学者提出AlexNet网络并取得成功之后,卷积神经网络算法就逐渐取代传统的图像处理方法,成为了遥感图像处理领域的研究热门。2014年R-CNN<sup>[2]</sup>实例分割网络的提出将图像分割推到了一个新的高度,这是深度学习首次应用于目标检测的领域。实例分割网络和U-Net<sup>[3]</sup>、SegNet<sup>[4]</sup>等成熟的语义分割网络最大的不同点在于,实例分割不仅仅要在像素层面将属于物体的前景和背景分离出来,还要在像素层面给出每个物体单独的像素并给予边界框定位。自从实例分割网络首次被引入到像素级的语义分割起,诸多学者对

其进行了改革。SPP-Net<sup>[5]</sup>在R-CNN网络的基础上加入了金字塔池化,降低了输入图片尺寸大小对结果的影响;Fast R-CNN<sup>[6]</sup>减少了R-CNN中繁琐的池化过程,加入RoIPooling,采取单尺度池化以加快网络的运行速度;Faster R-CNN<sup>[7]</sup>网络摒弃了R-CNN和Fast R-CNN中消耗大量时间的Selective Search,用区域建议网络(RPN)实现了端到端的训练,在大大提升训练速度的同时也很好地改善了区域提取的准确率。

2017年底何凯明等<sup>[8]</sup>学者提出了Mask R-CNN网络,该网络在Faster R-CNN的基础上进行了进一步的改进,并且取得了更好的分割效果。Mask R-CNN的创新点在于用RoIAlign代替了RoIPooling,提升了边界框的准确率;同时在损失函数部分增加了Mask分支,引入了特征金字塔<sup>[9]</sup>网络,学习到了更多特征信息,同时也提高了预测的准确率。

Nie等<sup>[10]</sup>和Zhang等<sup>[11]</sup>根据船舶细长不好辨认的特点,基于改进的Mask R-CNN对船舶进行了目

收稿日期:2020-06-20;修订日期:2021-03-28

作者简介:刘 钊(1967—),男,江苏镇江人,副教授,主要从事GIS及其应用、云GIS、时空大数据及遥感图像处理等方面的研究。

E-mail:liuz@tsinghua.edu.cn

通讯作者:赵 桐(1996—),男,北京人,硕士研究生,主要从事深度学习、地理信息科学研究。E-mail:zhaot18@mails.tsinghua.edu.cn

标检测;Zhao等<sup>[12-13]</sup>应用该网络提取出了高分辨率卫星影像中的建筑物并取得了很好的效果;Zhang等<sup>[14]</sup>通过改善的Mask R-CNN网络对数据集进行了实例分割,并取得了比原网络更理想的结果。在其他的一些实验研究中,田锦等<sup>[15]</sup>基于Mask R-CNN进行了地面标识检测的实验;白宝林<sup>[16]</sup>对网络进行了微调,提高了网络效率并对车辆进行识别;邓璇元等<sup>[17]</sup>引入了可变性的卷积和可变性的池化层,以应对基于环视相机的识别问题;陈茗杨等<sup>[18]</sup>增大了RoIAlign层的分辨率,增加了掩膜(Mask)分支里图片的尺寸,提升了算法的效率;张旭东等<sup>[19]</sup>在Mask R-CNN的基础上引入Depth特征,使网络学习到了深度信息。

使用Mask R-CNN卷积神经网络进行遥感影像中地物的分割,该网络是目前效果最好的实例分割网络,通过FPN<sup>[20]</sup>、RoIAlign以及Mask branch(掩码层分支)三项改进可极大提升识别准确率。相比于传统语义分割网络,该网络具有更高的精度,并能够准确分辨出每一个不同个体及相对应的位置。Mask R-CNN网络已经成为实例分割领域中各种网络的佼佼者。

上述各种研究中,关注点主要集中在分割结果和真实值的契合度以及网络的准确率上,较少关注到矢量图像的语义信息。另外,上述研究均基于一些照片和简单地物,预测难度不高。如果范围扩大到整个城市的高分遥感影像,其预测难度就会大大增加。这主要因为上述研究的输出结果为栅格格式,而栅格格式无法统计语义信息,故必须要转为矢量结果。目前可以将栅格格式转化为矢量格式的唯一方法是GDAL开源栅格空间数据转换库(之后统称为GDAL库)。但是当GDAL库在处理过大的图像、图像中有大量锯齿或零碎像素点时,函数会出现卡退、无法执行的情况,从而无法得到理想的矢量格式结果。

基于卷积神经网络的图像分割都是以小的图像瓦片为单位,即通过训练好的模型预测图像瓦片,并把预测后的瓦片拼合成一张完整的大图。如果输出格式为栅格,则在拼合的过程中可以生成一幅完整的大图。但是GDAL库的拼合函数在矢量合成时仅仅是简单的图片拼合,并没有实际合成一个整体。虽然这对图像的质量没有很大影响,但当输出格式为矢量并进行矢量信息统计时会产生很大误差。所以,如何修复矢量预测结果中的缝隙,

就成为了一个关键性的问题。

在修复缝隙的算法中,最著名的DeepCrack<sup>[21]</sup>算法意在识别裂缝并以最快的速度预测单幅图像,其出发点和目标与本文不同。本文试图通过裂缝修复,达到预测大图并得到完整矢量语义信息的效果,并使该方法在实例分割网络上具有普适性。

为此,本文提出了针对卷积神经网络预测结果缺陷的缝隙修复算法,并以Mask R-CNN网络为例进行了应用。该算法的目的是解决卷积神经网络在预测矢量结果并拼合成完整大图过程中会出现缝隙的问题。预测时先使用overlapsize算法缩小瓦片之间的缝隙,同时改善预测结果与真实结果的匹配程度;再通过PostGIS中的函数,选出与矢量瓦片边界相交的矢量地块,识别交点的坐标并完成合并,最后把不和边界相交的矢量地块添加进来,生成最后的结果。此方法生成的矢量大图速度快、精度高、质量不受影响,且大大方便了日后的统计工作。

## 2 Mask R-CNN网络原理及缝隙修复算法研究

### 2.1 Mask R-CNN网络基本原理

Mask R-CNN网络总体框架<sup>[8]</sup>和预测流程分别如图1和图2所示。

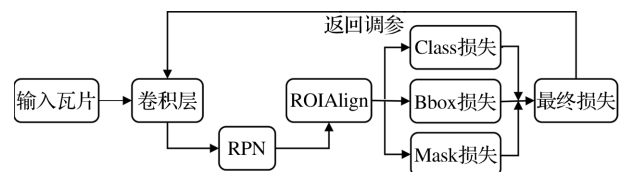


图1 Mask R-CNN网络框架图

Fig.1 Mask R-CNN network framework

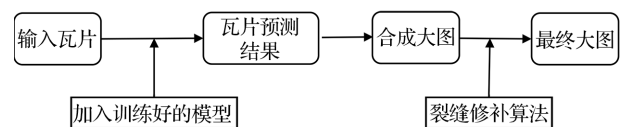


图2 Mask R-CNN网络预测流程图

Fig.2 Mask R-CNN network prediction process

该网络相对于其他卷积神经网络是否拥有更好的分割效果,关键在于FPN、RoIAlign和损失部分的Mask分支这3部分的改进工作。

FPN (Feature Pyramid Network)是特征金字塔网络的缩写。FPN建立了每一层之间的横向连接,将深层的语义信息和浅层的特征有效地结合在一起;使得较低的层次在语义丰富的同时又增加了

细节。

该网络的第二项创新在于 RoIAlign 对 RoIPooling 的替代。Fast R-CNN 和 Faster R-CNN 网络中的 RoIPooling 在下采样的过程中进行了两次量化取整的操作,这样的取整会使边界框定位发生偏移。RoIAlign 算法取消了这两步量化的操作,将之改为双线性插值:即每一个 RoI 框的 4 个顶点的坐标都由它周围 4 个网格点的坐标经过线性插值得到,这样就大大增加了 Bounding box 线框的准确性。

该网络的第三项改进是增加了与分类回归和 Bounding box 回归并行的 Mask branch(掩码层)。掩码层的主要作用在于可以对每一个 RoI 精确预测类别。有多少类别就会生成多少 Mask,第  $i$  个类别的 RoI 只会与第  $i$  个 Mask 有关,不会受到其他 Mask 的影响。

## 2.2 缝隙修复算法研究

缝隙修复算法由两个流程组成,分别为第一步

的 Overlapsize 算法以及第二步的填补缝隙算法。后文将针对这两个流程的原理做详细阐述。

卷积神经网络的预测过程同样把未经训练、待预测的栅格影像瓦片作为输入,加入在训练过程中调整好的参数模型后,输出得到相应的预测结果。然而预测为矢量格式的缺陷在于,一些地物如果被裁剪为很多细碎的部分,并分布在周围的若干瓦片之中时,有很大几率不会被模型识别并预测出正确结果,造成识别误差。

实验将缝隙修复算法在预测过程中加入 Overlapsize 算法,即在预测之前先将待预测栅格影像瓦片扩大同等宽度的一圈,待模型预测扩大后的图像以后,再对预测结果进行裁剪,使其恢复原来的大小。这种改进的好处在于预测过程中那些瓦片里被切割细碎的小地物不会被忽视,从而可得到更为准确的分割结果。Overlapsize 算法的原理如图 3 所示。

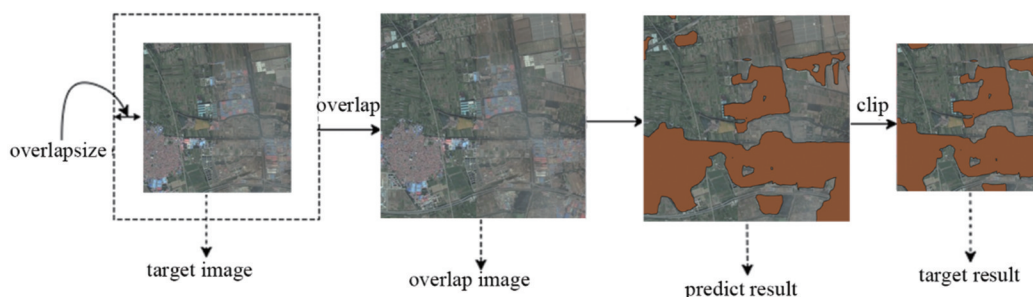


图3 Overlapsize 算法原理图

Fig.3 Principle of Overlapsize algorithm

Overlapsize 算法可以改善预测过程中小地块的识别问题,并且对栅格图预测和矢量图预测一并适用。然而,Overlapsize 算法尽管改善了预测结果的质量,却不能解决矢量预测结果中因存在缝隙,导致语义信息有误的问题。故实验在合成大图后提出缝隙修复算法的第二步—填补缝隙算法,目的是修补矢量大图中的裂缝,使之合成一张语义信息完整的矢量大图。该算法基本原理是通过在本地连接服务器,并且在远程服务器上利用 PostGIS 函数,使用 SQL 语句执行缝隙修补。在执行算法之前,首先需要在预测图片的同时生成每张矢量瓦片包围的线框矢量,并且将矢量瓦片和线框瓦片通过 gdal\_merge 算法分别合成为两个待处理的大图。这两张大的矢量图将会作为缝隙修复算法的输入。

从连接数据库到最后生成拼合的矢量结果,整个过程共分为 5 个步骤。第一步为数据库连接,具体操作方法为输入服务器的主机地址、想要连接的

数据库名称、数据库用户名、用户名密码以及主机的端口号,通过 python 中的模块函数实现服务器的连接。第二步为读取文件并创建数据表。用 Python 中的函数读取线框大图以及矢量大图,并通过 SQL 语句在远程数据库中创建数据表,将读取的线框大图以及矢量大图信息传输到新创建的矢量表中。第三步为核心算法,即通过 PostGIS 中的函数填补矢量大图中的缝隙。第四步为导出矢量,即将执行后生成的结果数据表导出为矢量文件,并保存到本地的路径中。最后一步为断开服务器连接。操作方法为将数据库中所有相关数据表归类,并用 Python 的函数断开数据库连接,结束整个算法的流程。

算法流程的第三个步骤为核心算法处理,该步骤通过 PostGIS 函数达到消除缝隙的目的。算法核心部分的流程如图 4 所示。

填补缝隙算法流程中,首先通过 PostGIS 数据库中的 st\_buffer 函数获取缝隙缓冲区的几何对象,



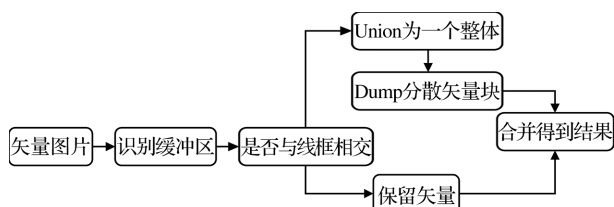


图4 缝隙修复算法流程图

Fig.4 Gap-repairing algorithmic flow

并设置缓冲区阈值为—浮点数(本例中为0.001 m)。如果将阈值设置过大,尽管会加快实验速度,但会导致原本不相关的两块田地识别为同一类,从而引起矢量结果崩溃;如果将阈值设置过小,会大大减慢实验速度,并且在阈值小于某一临界值后,可能造成因为阈值小于缝隙宽度而不能将同一地块的两部分合并。本函数的输入为经过 Overlapsize 算法以后的矢量图片,输出为半径为阈值大小的缓冲区,从而获取到矢量缝隙所在位置。

其次,填补缝隙算法在此处设置判断语句,判断每一个小的矢量瓦片中,诸多矢量多边形是否与线框,即矢量瓦片的边界处存在交集,并把有交集和没有交集的两类矢量多边形分类存放在两个数据表中。如果瓦片中的矢量多边形与缝隙不存在交集,则保持原矢量多边形;针对与线框有交集的多边形,缝隙修复算法采用 `st_union` 函数进行合并,将所有零碎的矢量多边形合而为一。但由于 `st_union` 函数执行合并后所有矢量多边形合并成为了一个矢量,违背了将每一个地块独立开来,生成独立矢量信息的基本原则,故算法需再运用 `st_dump` 函数进行打散。打散以后,每一个地块即可被独立分开,且缝隙也得到了填补。最后,该算法再把和线框没有交集的矢量多边形加入,经过合并以后生成最后的结果。中间核心的处理部分代码如下文所示。

```

with
  line as (select st_buffer( 线框数据表 .geom,
0.001) as geom from 线框数据表),
  block_intersect as (select block.* from 矢量数据
表 block, line where st_intersects (block.geom, line.
geom)),
  block_union_result  as  (select  (st_dump
(st_union  (st_buffer  (st_makevalid  (geom)  ,
0.0001))))).geom from block_intersect),
  block_not_intersect as (select 矢量数据表 .

```

```

geom from 矢量数据表 where frame_sql.id not in
(select id from block_intersect)),

```

```

  result as (select * from block_union_result
union select * from block_not_intersect)
  select * from result

```

### 3 实验过程与结果分析

#### 3.1 实验设置及实验流程设计

实验将针对农田地块进行实例分割,对数据集中的农田地块进行标注后使用 Mask R-CNN 网络进行训练,并在预测过程中加入 Overlapsize 算法进行对比分析。然后,在加入 Overlapsize 算法得到结果的基础上再加入填补缝隙算法,通过实验结果对比体现出缝隙修复算法的优异性能。

实验的深度学习模型基于 PyTorch 搭建,使用 Mask R-CNN 网络作为基础框架,硬件环境采用 Nvidia GeForce GTX 1080 GPU 显卡。采用公开的 coco-2014 数据集和 ImageNet 数据集对网络进行预训练,将经典的 ResNet-50+FPN 作为特征提取的骨架部分。实验开始时统一将输入图片的大小设置为  $256 \times 256$  像素,初始学习率设置为 0.002 5,权重衰减设置为 0.000 1。

实验的第一步是图像裁剪。所有的原始图片都要切割成相应输入尺寸,在本实验中为长宽为  $256 \times 256$  像素的图像瓦片。除了原始的影像以外,还要生成与之对应的真实值 (Ground Truth) 标签。并将所有处理好的图像瓦片一并放入网络当中。

第二步是特征提取,即通过 Mask R-CNN 初始的骨架网络提取出图像的浅层和深层特征。Mask R-CNN 普遍采用的 backbone 是 FPN,因为它可以融合特征映射并得到理想的输出。输出的结果输入至 RPN 和 ROIAlign。

第三步是分类回归。根据 RPN 网络本身的特性,对所有边界框进行筛选并挑出 300 个候选边界框放入 ROIAlign。输出得到的结果作为损失函数的输入进入损失函数,并且根据三类不同的损失函数对最终的结果进行分类回归。

第四步是返回调参。根据上一步得到的损失函数 Loss 值的反馈,整个网络框架会返回到卷积的部分调参,对网络整体进行完善。

第五步是预测结果。训练结束以后会得到已完成调参的实验模型。根据模型和测试集进行网

络预测,得到最后的预测结果。

### 3.2 实验数据来源

实验采用美国堪萨斯州高分辨率的正射遥感图像作为实验的数据源,原因在于该地区农田形状相对规范,并且地形复杂多样,便于进行深度学习的研究。本实验下载了美国堪萨斯州2019年9月初的无云10 m分辨率哨兵2号影像,并采用人工精修过的已有堪萨斯州农作物矢量文件作为与影像

相对应的标签图。

实验首先通过裁剪得到了4 000余张256像素 $\times$ 256像素大小的影像,其中随机选取80%作为训练集,其余作为验证集。测试集遥感影像选取2019年9月初美国堪萨斯州的另一无云10 m分辨率哨兵2号影像地块,目的是保持地物类型相同的同时不与训练集及验证集存在交集。训练集中的部分影像样本如图5所示。

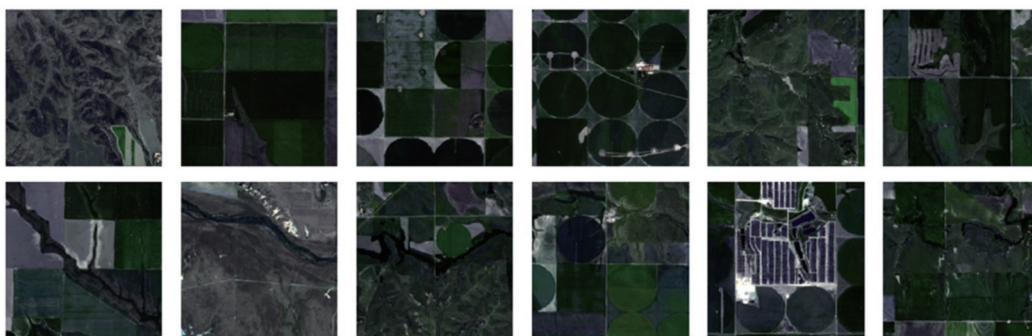


图5 农田实例分割数据集示例

Fig.5 Farmland instance segmentation dataset examples

### 3.3 实验结果与分析

实验使用4 000余张256 $\times$ 256大小的影像进行训练。设置网络总共迭代90 000次,其中每20次显示一次结果。实验的初始学习率设为0.002 5,并在每30 000次迭代以后缩小为原先的0.1倍。每一次显示结果均输出4个不同参数,分别为:总损失、分类损失、边界框损失、掩膜损失。由于每一次迭代得到的损失值有不稳定性,为了得到大体变化趋势,本实验每隔1 000次迭代提取一次网络的平均损失值数据,得到如图6~图9中4个参数的变化曲线。同时,Mask R-CNN网络会在迭代循环结束以后给出象征训练精度的3个参数,如表1所示。

从图6~图9可以看出,曲线总体平滑,前期斜

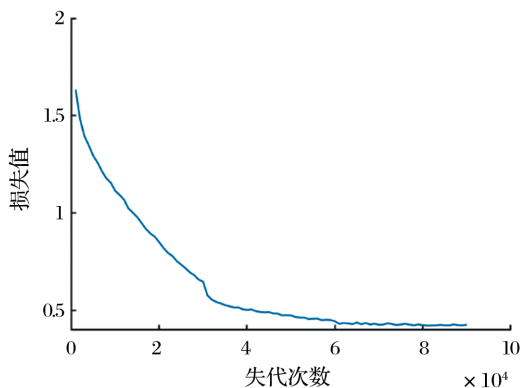


图6 总损失值变化曲线

Fig.6 Varying curve of total loss value

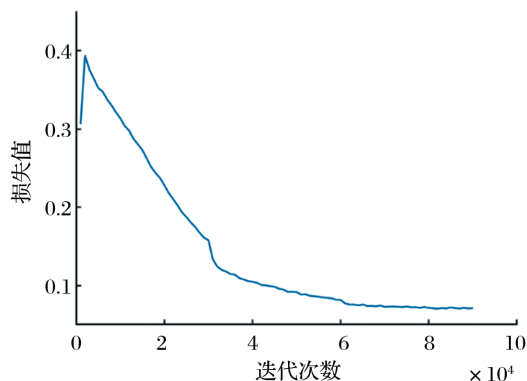


图7 分类损失值变化曲线

Fig.7 Varying curve of classified loss value

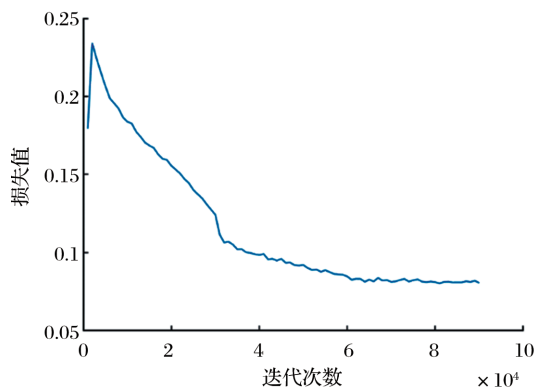


图8 边界框损失值变化曲线

Fig.8 Varying curve of bounding box loss value

率较高,后期斜率逐渐降低并最终趋于稳定。收敛以后的总损失值已经下降到了0.42~0.43,这说明

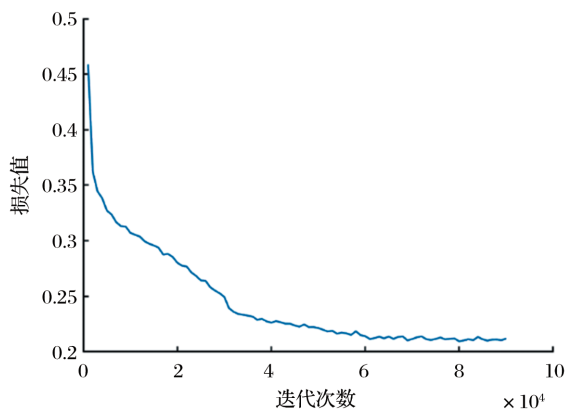


图 9 掩膜损失值变化曲线

Fig.9 Varying curve of mask loss value

各项损失值都已经被抑制到了一个很小的数值。由于此次实验的 Mask R-CNN 网络是已预训练的网络,以及数据量足够大,故不会出现过拟合的情况。从表 1 可知,实验得到的 3 个精度参数 (AP,

表 1 网络训练精度

Table 1 Network training accuracy

精度	AP	AP <sub>50</sub>	AP <sub>75</sub>
Mask R-CNN/%	33.6	55.2	35.3
本实验/%	34.0	55.5	35.4

AP<sub>50</sub>, AP<sub>75</sub>) 分别比 Mask R-CNN 网络训练得到的相应参数高出了 0.4、0.3 和 0.1 个百分点。这说明人工精修过的堪萨斯州农作物训练集拥有相当高的精度,并且印证了 Mask R-CNN 网络对于颜色很相近的草地、山体和农田地块依然具有很高的区分度。

第一组对比实验是基于 Overlapsize 算法。图 10 为 Overlapsize 算法实验效果对比图,其中图 10(a) 为未添加 Overlapsize 的预测矢量结果;图 10(b) 为添加 Overlapsize 的预测矢量结果。

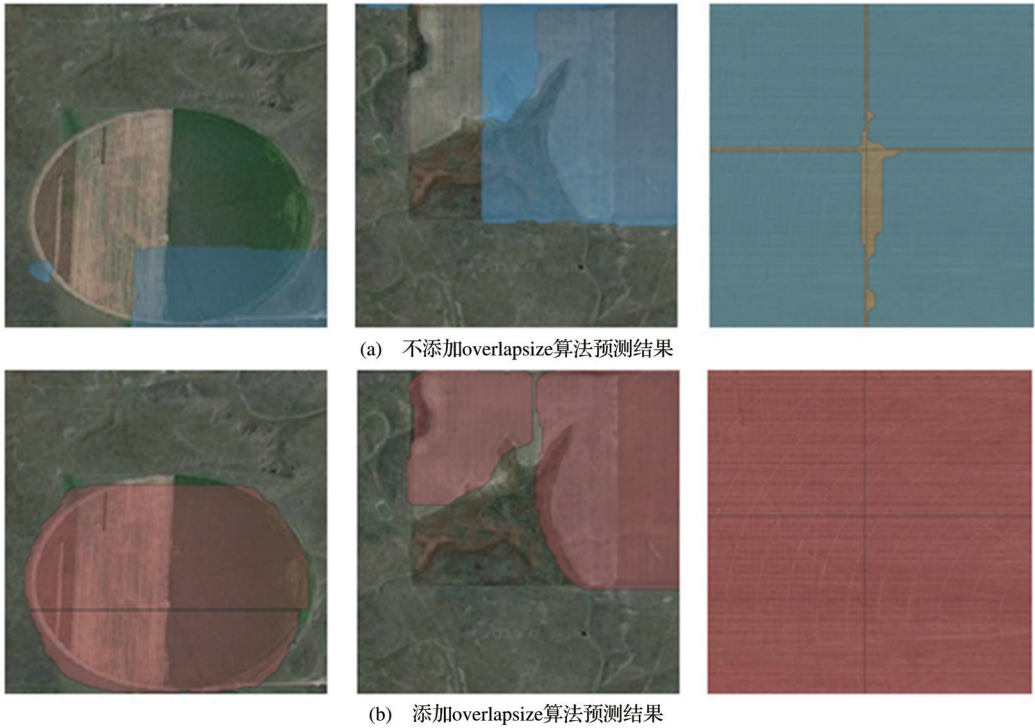


图 10 Overlapsize 算法实验对比图

Fig.10 Experimental comparison of overlapsize algorithm

从图 10 的对比中可以看出,有些地方二者的预测结果都很理想,而有些地方(图 10(b)左一、左二)加入 Overlapsize 算法的预测结果要明显好于不加入 Overlapsize 算法的结果。如果没有 Overlapsize 算法,在预测结果中会存在两点主要缺陷:一是当某地物(本实验中的农田地块)被分割成为多个瓦片图像时,被分割后的某些碎片会因为面积太小而

不容易被网络识别出来,那么在合成为完整影像以后,重新拼凑好的地物将会缺失一角,从而带来很大误差;二是当某一个地物被分割成为多个瓦片图像后,被分割后的碎片都被网络识别出来,那么在拼合以后一个完整的地物上面就会出现多个边界框,除此之外还会有很宽的缝隙,这会影响到视觉效果。而 Overlapsize 算法通过“先放大再裁剪”的



步骤巧妙地规避了这两点缺陷,使得一个完整良好的分割结果能够展示出来。

然而,Overlapsize算法也造成负面影响。因为Overlapsize算法在执行时裁剪掉了最外面的一圈预测结果,使得预测得到的大图比输入的大图少了一圈最外面的影像信息,导致了信息的丢失。如果图片尺寸较小,该算法会给结果带来很大的影响。图片的尺寸越大,相对应的影响也越小。由于本文研究的数据集是堪萨斯州,图像尺寸较大,所以该算法的缺陷给结果带来的损失可以忽略不计。

从图10(b)中可以看出,Overlapsize算法的加入使得网络在预测测试集农田地块的时候很少出现漏识别的情况,并且从效果上看,农田边界与实际的边界契合很好。不足之处在于因农田的颜色与山体过于接近,网络存在少量错误识别与漏识别,但是不影响整体的效果。但前面的分析仅仅基于图片的预测效果,当进行后续的矢量统计工作

时,该方法的缺陷就会暴露出来,即很多的农田地块由一块被切割成了多块。为此,本实验中进一步地加入填补缝隙算法,并进行对比分析。加入填补缝隙算法前后的结果对比如图11所示。

图11中棕色矢量多边形为未添加填补缝隙算法的预测结果;灰色矢量多边形为添加填补缝隙算法的预测结果。矢量图的大小为63 MB,测试集影像的分辨率为10 m,宽高分别为11 397像素 $\times$ 10 979像素。该算法在此条件下的运行时长为4 min。对比之下,GDAL函数运行较小,曲面平滑的栅格图时拥有很快的速度,但当曲面不平滑且矢量图面积很大时,GDAL函数会发生卡崩,从而无法预测出矢量图。运行时长虽然与本机电脑中央处理器性能相关,但该填补缝隙算法都可以在远小于GDAL函数的时间内完成转矢量并输出理想的矢量结果。说明该算法完全可以胜任更大数据量的处理,不会出现由栅格转矢量的过程中容易发生的卡退现象。



图11 缝隙修复算法对比图

Fig.11 Comparisons of gap-repairing algorithm

从图片的效果看,除了个别的预测存在边界框拼合方面的误差,几乎所有的农田田块都完成了合并,并且矢量图的质量没有受到任何影响。加入缝隙修复算法以后,既可以保证图像在预测过程中的质量不受影响,又可以得到完整的矢量大图,并在数据量非常大的时候仍然能在最短时间内完成任务。说明该算法在进行大数据量影像预测时是十分有效的。

现阶段处理大影像的通用方法是分块处理并在最后进行拼合。这种做法虽然分块越小单块的处理时间就越少,但会导致总块数大大增加,这与预测出 $256 \times 256$ 的小瓦片再进行拼合的方法没有本质上的差别。除此之外,无论是进行何种分块操

作,如果想要统计矢量信息,最后都会不可避免地涉及到矢量的拼合工作。在缝隙客观存在的基础上,GDAL函数无法从自身实现这种缝隙拼合。本文研究的缝隙修复算法的优势就在于无论处于什么情况,该算法均可实现图片一次性输入得到矢量图,并能在很短时间内运行出准确结果;另外由于该算法与网络预测并不矛盾,所以对任意实例分割网络均可使用。因此,本算法具有普遍性、适用性和实用性,可进行推广与应用。

## 4 结论

针对基于Mask R-CNN网络的实例分割中预测结果出现较大语义信息误差的问题进行了深入

研究,提出了解决该问题的缝隙修复算法,通过 overlapsize 算法以及 PostGIS 数据库中的函数填补缝隙,完善了地块图像的语义信息。

对比实验表明,相比于传统 GDAL 算法,该算法处理效率极高,可在很短时间内完成大幅矢量图的处理;同时,算法并不会对网络本身的分割精度造成负面的影响;另外,算法具有很好的实用性,能够应用于几乎全部的神经网络处理图像案例。在解决大型矢量图的语义信息问题方面,该算法具有良好的应用前景。

#### 参考文献 (References):

- [1] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks[C]// Advances in Neural Information Processing Systems, 2012: 1097-1105.
- [2] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [3] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]// International Conference on Medical Image Computing and Computer-assisted Intervention. Springer, Cham, 2015: 234-241.
- [4] 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.
- [5] He K, Zhang X, Ren S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [6] Girshick R. Fast r-cnn[C]// Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [7] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-time Object Detection with Region Proposal Networks [C]// Advances in Neural Information Processing Systems, 2015: 91-99.
- [8] He K, Gkioxari G, Dollár P, et al. Mask R-CNN[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [9] Lin T Y, Dollár P, Girshick R, et al. Feature Pyramid Networks for Object Detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [10] Nie S, Jiang Z, Zhang H, et al. Inshore Ship Detection based on Mask R-CNN[C]// IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2018: 693-696.
- [11] Zhang Y, You Y, Wang R, et al. Nearshore Vessel Detection based on Scene-mask R-CNN in Remote Sensing Image [C]// 2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC). IEEE, 2018: 76-80.
- [12] Zhao K, Kang J, Jung J, et al. Building Extraction from Satellite Images Using Mask R-CNN with Building Boundary Regularization[C]// CVPR Workshops. 2018: 247-251.
- [13] Wen Q, Jiang K, Wang W, et al. Automatic Building Extraction from Google Earth Images Under Complex Backgrounds based on Deep Instance Segmentation Network[J]. Sensors, 2019, 19(2): 333. doi:10.3390/s19020333.
- [14] Zhang C, Li Z, Liu Q, et al. Toward Efficient Simultaneous Detection and Segmentation[C]// 2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM), IEEE, 2018: 1-5.
- [15] Tian Jin, Zhang Chi, Wang Yongsun, et al. Ground Mark Detection based on Mask R-CNN[C]// The 22nd Annual Meeting of Network New Technologies and Applications of China Computer Users Association Network Application Branch 2018.[田锦, 张弛, 王永森, 等. 基于 Mask R-CNN 的地面标识检测[C]// 中国计算机用户协会网络应用分会 2018 年第二十二届网络新技术与应用年会论文集, 2018.]
- [16] Bai Baolin. Vehicle Identification and Detection based on Improved Mask R-CNN[D]. Anhui: Anhui University, 2018. [白宝林. 基于改进的 Mask R-CNN 的车辆识别及检测[D]. 安徽: 安徽大学, 2018.]
- [17] Deng Liuyuan, Yang Ming, Wang Chunxiang, et al. Surround View Cameras based Instance Segmentation Method for Autonomous Vehicles [J]. Huazhong University of Science and Technology(Natural Science Edition), 2018, 46(12): 24-29.[邓珧元, 杨明, 王春香, 等. 基于环视相机的无人驾驶汽车实例分割方法[J]. 华中科技大学学报(自然科学版), 2018. 46(12): 24-29.]
- [18] Chen Mingyang, Zhao Zhigang, Pan Zhikuan, et al. Image Instance Segmentation based on Deep Learning[J]. Journal of Qingdao University (Natural Science Edition), 2019, 32(1): 46-50, 54.[陈茗杨, 赵志刚, 潘振宽, 等. 基于深度学习的图像实例分割[J]. 青岛大学学报: 自然科学版, 2019, 32(1): 46-50, 54.]
- [19] Zhang Xudong, Wang Yuting, Fan Zhiguo, et al. RGB-D Multi-class Instance Segmentation based on Double Pyramid Feature Fusion Model[J]. Control and Decision, 2020, 35(7): 1561-1568.[张旭东, 王玉婷, 范之国, 等. 基于双金字塔特征融合网络的 RGB-D 多类实例分割[J]. 控制与决策, 2020, 35(7): 1561-1568.]
- [20] Lin T Y, Dollár P, Girshick R, et al. Feature Pyramid Networks for Object Detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 2117-2125.
- [21] Liu Y, Yao J, Lu X, et al. DeepCrack: A Deep Hierarchical Feature Learning Architecture for Crack Segmentation [J]. Neurocomputing, 2019, 338: 139-153.



## Research on Gap-repairing Algorithm based on Convolutional Neural Network Prediction Result

Liu Zhao, Zhao Tong, Liao Feifan

*(Institute of Traffic Engineering and Geospatial Information, Department of Civil Engineering,  
Tsinghua University, Beijing 100084, China)*

**Abstract:** With the development of convolutional neural network technology, recent research has paid more attention to the improvement of accuracy and the improvement of semantic information. Mask R-CNN network is a further improved segmentation network of Faster R-CNN. It has a good segmentation effect in high-resolution remote sensing image feature recognition. However, since the convolutional neural network can only be trained and predicted with small tile images, there is a large semantic information error in the prediction results. Faced with this problem, this paper proposed a gap-repairing algorithm based on the defect of prediction result of convolutional neural network. The approach use overlapsize algorithm to improve the matching degree between the prediction result and the ground-truth result at first. Then fill the gap through the correlation function in the PostGIS database to repair the small tile, which can make it be spliced into a complete picture. The research and experiment results showed that the algorithm could improve the image semantic information well and has practicability.

**Key words:** Convolutional neural network; Instance segmentation; Mask R-CNN; Gap-repairing algorithm