

Cell Instance Segmentation

高易远

清华大学精密仪器系
gaoyy18@foxmail.com

邓雲峰

清华大学工程物理系
dengyf19@mails.tsinghua.edu.cn

田硕

清华大学精密仪器系
tianshuo42@163.com

Abstract

此次任务思路来源于Kaggle competition: Sartorius - Cell Instance Segmentation, aiming at detect single neuronal cells in microscopy images.这是一项针对神经元细胞图像的实例分割任务。我们尝试了基于Mask R-CNN和基于UNet的两种不同方法,经过调整和优化,最终取得了Public leaderboard score: 0.306 (排名169/1413)的结果。

1. Introduction

1.1. Competition Introduction

神经疾病,包括阿尔茨海默病和脑肿瘤等神经退行性疾病,是全球死亡和残疾的主要原因。然而,很难量化这些致命疾病对治疗的反应。

当前一种被广泛采用的检测方法是通过光镜检查神经元细胞,其要求在显微图像中准确且高速地分割单个神经元细胞。因此,计算机需要在神经网络的帮助下,对这些细胞进行精确的实例分割视觉图像处理。

由于神经元细胞具有非常独特、不规则和凹形的形态,目前的网络对神经元细胞的分辨率准确性有限。

为此,Sartorius公司在Kaggle上发布了竞赛:Sartorius - Cell Instance Segmentation, aiming at detect single neuronal cells in microscopy images.这是一项针对神经元及肿瘤细胞图像的实例分割任务竞赛,此次竞赛的主要识别肿瘤细胞为SH-SY5Y神经瘤细胞。

1.2. Overall Plan

在对instance segmentation有一定的了解后,结合kaggle平台上该竞赛的overview与code and discussion,我们小组计划选择Mask R-CNN与UNet网络架构作为本次竞赛的两个尝试方向。

我们计划将总体流程分为三步:

- (1) 赛方数据集规范化

- (2) 神经网络搭建、训练与优化
- (3) Kaggle评估与反馈

2. Background

在实例分割领域,在web of science上以cell instance segmentation进行检索,查阅发表日期与相关度排名靠前的文章,我们发现诸篇论文中提及的网络架构以Mask R-CNN架构居多,而其他网络架构中,UNet++、Signet、PolarMask++等其他架构也成效卓然。这些网络在神经元细胞、心肌细胞等诸多生物细胞图像分割上都取得了较好的效果。

以Sherry Lin, Narges Norouzi等人的研究成果为例[1],其提出了一种将Mask RCNN架构与Shape-Aware Loss相结合的深度学习解决方案,以产生细胞实例分割。文章中提及的网络架构效果优于之前的细胞分割工作,在DIC-C2DH-HeLa数据集上实现了91.91%的IOU,在PhC-C2DH-U373数据集上实现了94.93%的IOU。

与此同时,UNet++也作为高被引论文,展现了其巨大潜力。Zongwei Zhou, Nima Tajbakhsh等人对在医学图像分割领域应用广泛的UNet做出改进[2],修剪后的UNet++模型实现了显著的加速,同时仅表现出适度的性能下降。

其他还有很多表现良好的网络架构,综合复杂度与竞赛平台上的讨论,我们选择了Mask R-CNN与UNet网络架构作为本次竞赛的两个尝试方向。

3. Approach

我们分别对Mask R-CNN网络与UNet网络进行了尝试。

3.1. Mask R-CNN approach

Mask R-CNN [3]在Faster R-CNN [4]的基础上,用ROIAlign替换了ROIpooling,解决了ROIpooling的misalignment的问题,保证原图与feature map, feature map到ROI的像素对齐,提升目标检测精度;增加了一个mask branch,用于分割,对于每个ROI预测k个二分

类的掩码，实现掩码预测与类别预测的解耦。

此次任务中，使用了detectron2 API [5]中的Mask R-CNN实现，选择了Model zoo中ResNet-50 作为 backbone，FPN作为neck，lr sched 3x的COCO Instance Segmentation

Baselines(COCO-InstanceSegmentation/mask_rcnn_R_50_FPN_3x.yaml)。

使用Mask R-CNN的训练的脚本整体框架如下：

Step1: Load the dataset in coco format and visualize;

Step2: Define some util. functions(such as caculate MAP IOU);

Step3: Set the config to import Detectron2 properly;

Step4: Train and visualize results

Mask R-CNN: overview

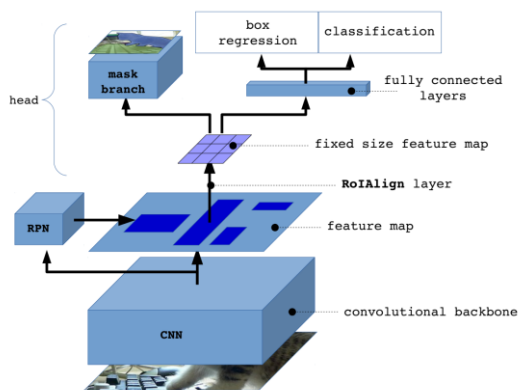


Figure 1: Mask R-CNN structure overview

3.2. UNet Approach

UNet [6]在FCN上提出了更为对称的encoder和decoder结构，即如下图所示左边和右边比较类似，而FCN的decoder只是一个deconvolution的操作，而UNet采取的是upconvolution并类似于encoder增加了卷积层，并且在skip-connection中FCN采取的是加操作，而UNet可从图中看到是copy and crop操作，即裁剪和叠操作。然后最后由最后一层使用1 X 1大小的卷积核，将通道数降低至特定的数量。

我们尝试采取了典型的UNet方法来先对图像进行语义分割，在使用skimage的label方法来将连通区域分割并得到各个轮廓outline来作为实例分割的结果（竞赛不要求进行分类）。

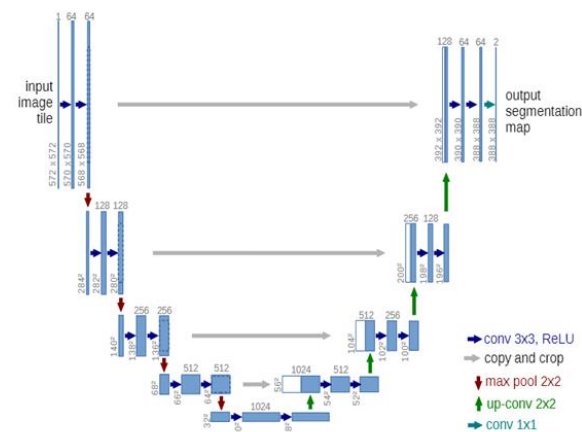


Figure 2: UNet structure overview

4. Experiment

下述实验Mask R-CNN基于Pytorch框架完成，UNet 基于Keras（Tensorflow）框架完成。

4.1. Datasets and Pretreatment

本次比赛主办方提供的数据集包括：

(1)用于训练的train集（其中包含 606 张相位对比度显微镜图像）以及与train集图片对应的train.csv，其中包含id\annotation\cell_type等标注信息，annotation中以RLE方式记录了细胞的mask信息。

(2)用于提交结果的test集（其中包含 3 张相位对比度显微镜图像）以及对应的sample_submission.csv，test集只有用于推断的图片，而没有标注信息，只能在Kaggle平台上提交给主办方进行评估。

(3)用于迁移学习的LIVECell_dataset_2021（其中包含 8 种类型神经细胞图像），以及与之对应的COCO格式的json标注文件。

由于比赛提供的train.csv不便于输入网络进行训练，故而我们先将其转化为coco数据集格式(json文件)，再按照 4:1 的比例从中分割train集与val集，test集由Kaggle平台承担。

实验转化过程如下 [7]:

(1)首先利用pandas库读入train.csv文件中信息，将其中annotation中RLE解码为mask。mask通过coco API与编码函数binary_mask_to_rle转化为uncomposed RLE，并计算对应的area、bbox等信息。原csv文件中的cell_type对应转化为 categories标签，id采用图片名称，汇总后输出至annotation_all.json [8];

(2)将原train文件夹中图片随机打乱后按照 4:1 的比例从中分割train集与val集，对应迁移到新的train文件夹与val夹下；

(3) 读取annotation_all.json, 将train/val文件夹下的图片对应的annotation各自导入annotation_train.json/annotation_val.json, 至此, COCO训练集与评价集分割完成。

同时, 利用coco API本身功能与编写RLE解码函数, 对训练图片与测试图片的处理结果进行了可视化, 便于后期对比。由于比赛不要求对细胞类型做出分割, 故而只需要观察不同的mask即可。

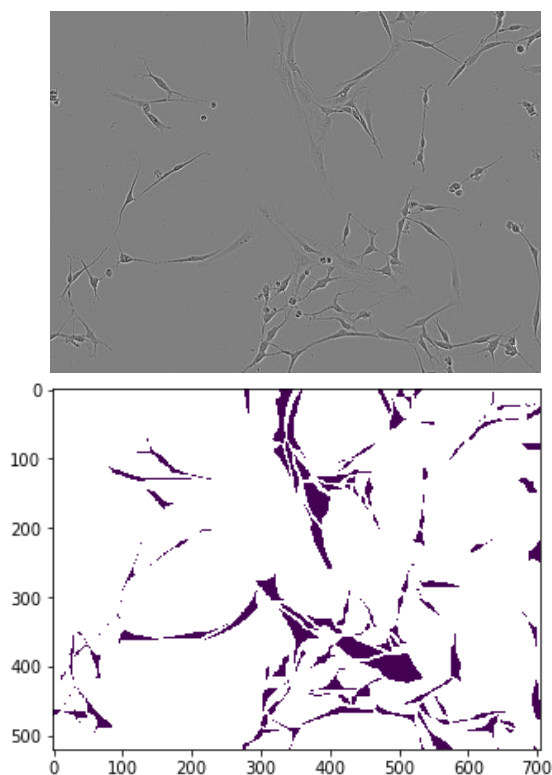


Figure 3: Visualization at train.csv

4.2. Evaluation methods

评价标准是MAP IOU, 对每个test image, 给定一组从0.5至0.95的IoU阈值(0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95), 对每个阈值t计算:

$$\frac{TP(t)}{TP(t) + FP(t) + FN(t)}$$

对每个test image的score是所有IoU阈值t的平均值:

$$\frac{1}{|\text{num of thresholds}|} \sum_t \frac{TP(t)}{TP(t) + FP(t) + FN(t)}$$

最后返回的score是所有test image的score的平均值。

这项竞赛还在进行, 提交后会在public leaderboard(calculated with approximately 41% of the test data)显示score。

4.3. Mask R-CNN experiment

此次任务使用的模型确定后 (3.1 节, Mask R-CNN with ResNet-50 and FPN),

始终确定的config如下:

batch_size为 2, 训练 10k个iteration, 用于验证集上测试的正负例选择阈值为 0.5, 每训练一个epoch (大概 240 个iteration) 在验证集上测试一次。

实验过程中, 训练超参数与训练方法的调整如下:

1. learning rate, 在每次训练中均调整;

2. lr_scheduler: 最终采用WarmupCosineLR;

3. optimizer: 最终采用SGD (SGD是Detectron2 中默认的方法, 使用默认参数), Adam也进行了尝试, 但是效果似乎不如SGD好。关于此, 网络上有一些相关的讨论;

4. Save best (best for val. MAP IOU), 观察训练时的loss, 发现在所有的训练次数中, total_loss (以及各部分loss) 在大概 5k个iteration之后陷入了震荡, 如下图, 暂时不清楚原因; 为了在一定程度上避免这种震荡带来的负面影响, 加入了save best方法。

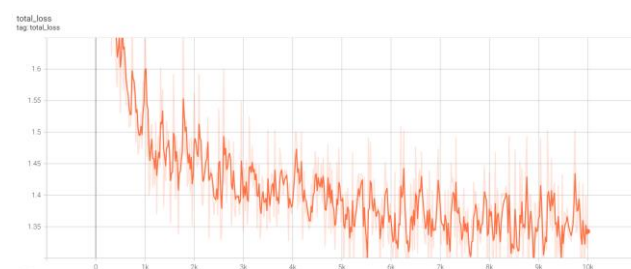


Figure 4: Mask R-CNN total loss during training

5. Data augmentation: 默认的数据增强方法包括resize和随机翻转, 尝试了加入随机更改亮度, 对比度等, 但是在验证集和测试时的效果均略微变差了, 可能的原因是原数据并不包括这些数据增强后的特征。

6. Transfer learning (pretrain with LIVECell dataset), LIVECell is the predecessor dataset to this competition. You will find extra data for the SH-SH5Y cell line, plus several other cell lines not covered in the competition dataset that may be of interest for transfer learning. 我们找到了在LIVECell dataset上预训练的checkpoints (已经预训练了 100000 iterations), 将其作为初始的权重, 减小学习率进行训练, 效果有明显的提高, 推测更多的 (具有细胞特点的) 训练数据以及更多的iteration让模型学习到了更好的特征表示。

7. Visualization in val. dataset: 训练完成后, 可视化在测试集上的predictions并与ground truth相比较, 如下图, 其中左侧是predictions, 右侧是ground truth。明显

发现对于shsy5y和astro这两类细胞检测结果，precision较高，但是recall太低，false negative太多。因此，我们在测试提交时，减小了这两种细胞的正负例选择阈值。

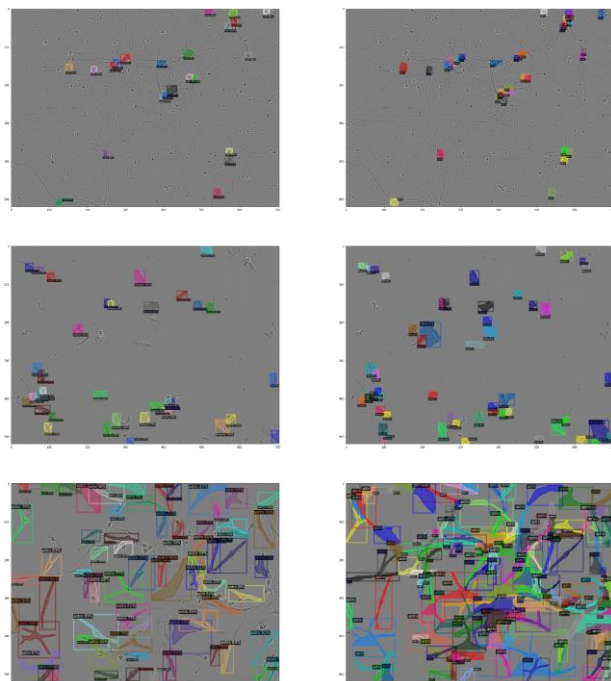


Figure 5: Visualization in val. Dataset (using Mask R-CNN)

实验结果汇总如下：

Results	name	lr	val score	test score
Only 1k iters, no lr_scheduler	1	0.0005	0.20450	NA
10k iters since 3, no lr_scheduler	3	0.001	0.26162	0.284
10k iters, no lr_scheduler, save best since 4	4	0.001	0.26468	0.295
Add WarmupCosineLR since 5	5	0.002	0.26538	0.296
	6	0.003	0.26377	0.296
Use transfer learning since 7	7	0.001	0.27255	0.305
	8	0.0005	0.27048	0.304
Try data augmentation	9	0.0005	0.26937	0.303
Try Adam instead of SGD	10	1.E-06	0.23848	0.278
All train images are used since 11	11	0.0005	NA	0.306
	12	0.0002	NA	0.304

Figure 6: Mask R-CNN experiment results

在kaggle平台上提交的最好的结果是实验名称 11 的实验结果，Public leaderboard score: 0.306 (169/1413)

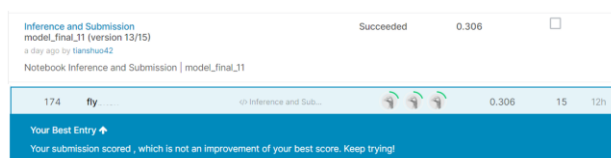


Figure 7: Mask R-CNN best results(till 2021/12/20)

4.4. UNet experiment

采用的config中，首先考虑算力和时间受限，将图片缩放成 256*256，learning_rate 采取默认，batch_size 设为 5，epochs 为 200，optimizer 采用的是 Adam 算法，loss 为 binary_crossentropy，回调采用了 Earlystopper 算法 (patience 设为 10) 和当模型更优时存为 pth 的 checkpoint。其中的 loss 表达式如下：

$$H_p = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i))$$

代码中 dice_coefficient 为描述 mask 的 ground_truth 和 prediction 的匹配度。

最终训练的曲线如下：

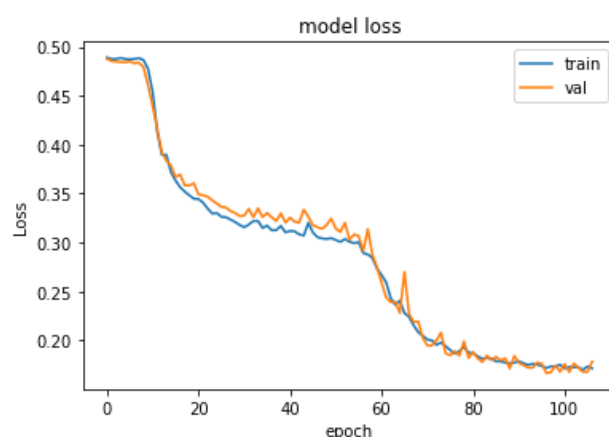


Figure 8: UNet total loss during training

可视化训练结果如下：

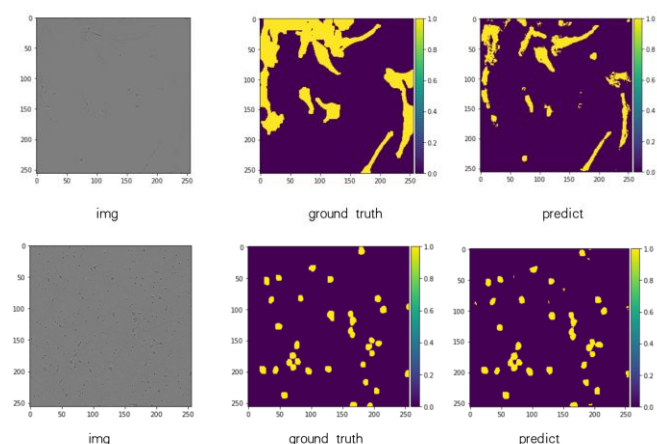


Figure 9: Visualization in val. dataset(using UNet)

5. Conclusion

5.1. Conclusion for Mask R-CNN

主要收获是学习了实例分割的相关知识，直接调用开源cv库中的模型并根据需要更改配置和重写相关的方法，以及训练和可视化的技巧。

为了能够在此次竞赛中获得更高的分数，最有效的方法或许是更改模型结构，Mask R-CNN是一个通用的分割结构，将其略作调整，更适合对细胞进行分割应该是一条可行的路径。此外，还可以进一步尝试Cascade Mask R-CNN, Mask Scoring R-CNN等改进。

5.2. Conclusion for Unet

Unet作为语义分割网络在医学图像情景上表现都十分良好，但在这个实例分割任务中，IoU实际上偏低，对于先语义分割再实例分割不具有竞争力。

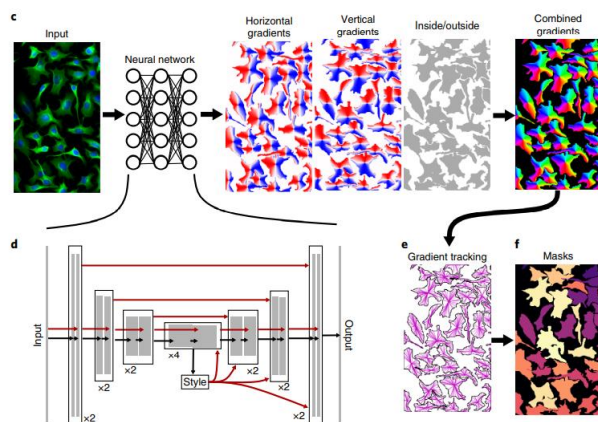


Figure 10: Cellpose structure overview

经过尝试和探索发现了cellpose方法，可以到达 0.307 的LB，其主要思想为先用stimulated diffuse将要拟合的 ground_truth进行染色，然后分别求水平和垂直两个方向的梯度图，然后训练一个神经网络（基于Unet）能将 stimulated diffuse后的图片得出水平和垂直的梯度图，然后再用gradient flow tracking方法得到masks，对于进一步提升UNet性能提供了方向 [9]。

致谢

感谢华为公司提供的免费华为云计算服务！

References

[1] Lin, Sherry, and Narges Norouzi. "An Effective Deep Learning Framework for Cell Segmentation in Microscopy Images", 2021. doi:10.1109/embc46164.2021.9629863.

[2] Zhou, Zongwei, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. "Unet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation". IEEE Transactions on Medical Imaging 39, no. 6 (2020): 1856–67. doi:10.1109/tmi.2019.2959609.

[3] Shaoqing Ren and Kaiming He and Ross B. Girshick and Jian Sun, . "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks".CoRR abs/1506.01497 (2015).

[4] Kaiming He and Georgia Gkioxari and Piotr Dollár and Ross B. Girshick, . "Mask R-CNN".CoRR abs/1703.06870 (2017).

[5] Yuxin Wu, , Alexander Kirillov, Francisco Massa and Wan-Yen Lo, and Ross Girshick. "Detectron2." <https://github.com/facebookresearch/detectron2>. (2019).

[6] Olaf Ronneberger, Philipp Fischer, Thomas Brox,U-Net: Convolutional Networks for Biomedical Image Segmentation

[7] Efficient coco dataset generator, <https://www.kaggle.com/coldfir3/efficient-coco-dataset-generator?scriptVersionId=79100851>

[8] convert-rle-to-bounding-box-x0-y0-x1-y1, <https://www.kaggle.com/eigrad/convert-rle-to-bounding-box-x0-y0-x1-y1>

[9] Stringer, C., Wang, T., Michaelos, M. et al. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods 18,100–106 (2021).