

深度学习-语义分割篇

作者：神秘的wz

语义分割前言

- 什么是语义分割
- 暂定的学习规划
- 语义分割任务常见数据集格式
- 语义分割得到结果的具体形式
- 语义分割常见评价指标
- 语义分割标注工具



语义分割前言

常见分割任务

语义分割	(semantic segmentation)	FCN
实例分割	(Instance segmentation)	Mask R-CNN
全景分割	(Panoramic segmentation)	Panoptic FPN



语义分割



实例分割

注意边缘以及忽略的像素

语义分割前言

暂定的学习规划

Model	mIoU	Global Pixel Acc	Inference on CPU (sec)	Params (M)
LR-ASPP MobileNetV3-Large	57.9	91.2	0.3278	3.22
DeepLabV3 MobileNetV3-Large	60.3	91.2	0.5869	11.03
FCN MobileNetV3-Large (not released)	57.8	90.9	0.3702	5.05
DeepLabV3 ResNet50	66.4	92.4	6.3531	39.64
FCN ResNet50	60.5	91.4	5.0146	32.96

<https://pytorch.org/blog/torchvision-mobilenet-v3-implementation/#semantic-segmentation>

语义分割前言

语义分割任务常见数据集格式

PASCAL VOC



PNG图片(P模式)

注意边缘以及忽略的像素

语义分割前言

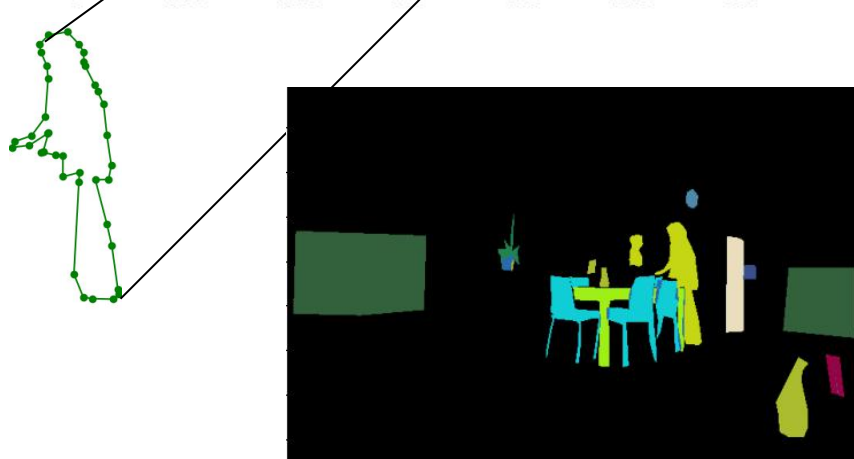
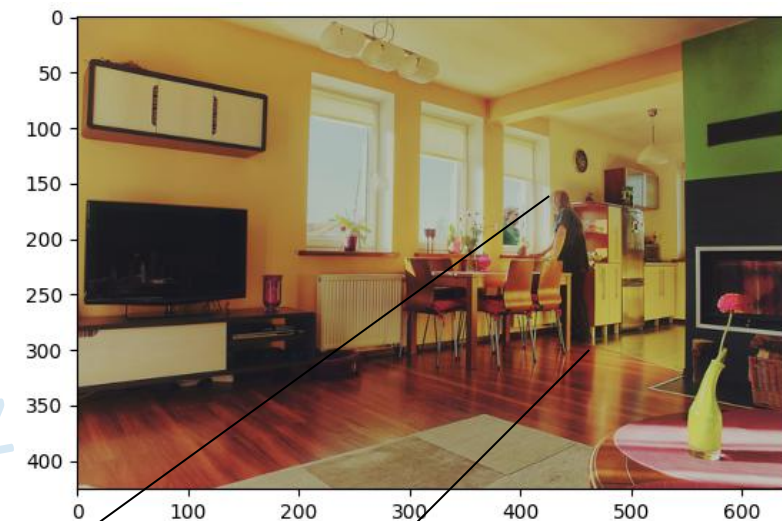
语义分割任务常见数据集格式

MS COCO

polygons

针对图像中的每一个目标都记录了多边形坐标

[428.19, 219.47, 430.94, 209.57, 430.39, 210.12, 421.32, 216.17, 412.8, 217.27, 413.9, 214.24, 422.42, 211.22, 429.29, 201.6, 430.67, 181.8, 430.12, 175.2, 427.09, 168.06, 426.27, 164.21, 430.94, 159.26, 440.29, 157.61, 446.06, 163.93, 448.53, 168.06, 448.53, 173.01, 449.08, 174.93, 454.03, 185.1, 455.41, 188.4, 458.43, 195.0, 460.08, 210.94, 462.28, 226.61, 460.91, 233.76, 454.31, 234.04, 460.08, 256.85, 462.56, 268.13, 465.58, 290.67, 465.85, 293.14, 463.38, 295.62, 452.66, 295.34, 448.26, 294.52, 443.59, 282.7, 446.06, 235.14, 446.34, 230.19, 438.09, 232.39, 438.09, 221.67, 434.24, 221.12, 427.09, 219.74]



语义分割前言

语义分割得到结果的具体形式



palette

mask蒙板 (加上调色板之后的效果)



每个像素数值
对应类别索引

bilibili: 霹雳吧啦WZJ

语义分割前言

常见语义分割评价指标

$$\text{Pixel Accuracy (Global Acc)} \quad \frac{\sum_i n_{ii}}{\sum_i t_i}$$

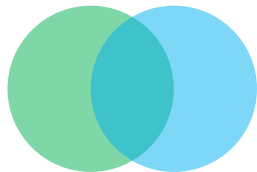
$$\text{mean Accuracy} \quad \frac{1}{n_{cls}} \cdot \sum_i \frac{n_{ii}}{t_i}$$

$$\text{mean IoU} \quad \frac{1}{n_{cls}} \cdot \sum_i \frac{n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}}$$

n_{ij} : 类别i被预测成类别j的像素个数

n_{cls} : 目标类别个数(包含背景)

$t_i = \sum_j n_{ij}$: 目标类别i的总像素个数(真实标签)



语义分割前言

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

真实标签
(忽略边缘)

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

预测标签

真实标签

	0	1	2	3	4
0					
1					
2					
3					
4					

预测标签

语义分割前言

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

真实标签

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

预测标签

真实标签

	0	1	2	3	4
0	16	0	0	2	0
1					
2					
3					
4					

预测标签

语义分割前言

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

真实标签

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

预测标签

真实标签

	0	1	2	3	4
0	16	0	0	2	0
1	1	3	0	0	0
2					
3					
4					

预测标签

语义分割前言

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

真实标签

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

预测标签

真实标签

	0	1	2	3	4
0	16	0	0	2	0
1	1	3	0	0	0
2	1	1	16	0	0
3					
4					

预测标签

语义分割前言

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

真实标签

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

预测标签

真实标签

	0	1	2	3	4
0	16	0	0	2	0
1	1	3	0	0	0
2	1	1	16	0	0
3	0	0	0	12	0
4					

预测标签

语义分割前言

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

真实标签

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

预测标签

真实标签

	0	1	2	3	4
0	16	0	0	2	0
1	1	3	0	0	0
2	1	1	16	0	0
3	0	0	0	12	0
4	2	0	0	2	8

预测标签

语义分割前言

真实标签

预测标签	0	1	2	3	4
0	16	0	0	2	0
1	1	3	0	0	0
2	1	1	16	0	0
3	0	0	0	12	0
4	2	0	0	2	8

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

真实标签

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

预测标签

$$\text{global_accuracy} = \frac{16 + 3 + 16 + 12 + 8}{64} \approx 0.859$$

语义分割前言

真实标签

	0	1	2	3	4
0	16	0	0	2	0
1	1	3	0	0	0
2	1	1	16	0	0
3	0	0	0	12	0
4	2	0	0	2	8

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

真实标签

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

预测标签

cls0_acc

$$\frac{16}{20}$$

cls1_acc

$$\frac{3}{4}$$

cls2_acc

$$\frac{16}{16}$$

cls3_acc

$$\frac{12}{16}$$

cls4_acc

$$\frac{8}{8}$$

语义分割前言

真实标签

预测标签

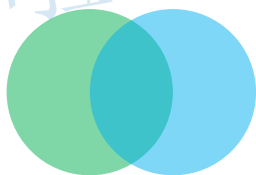
	0	1	2	3	4
0	16	0	0	2	0
1	1	3	0	0	0
2	1	1	16	0	0
3	0	0	0	12	0
4	2	0	0	2	8

0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3
0	0	4	4	4	4	0	0
0	0	4	4	4	4	0	0

真实标签

0	0	1	1	1	0	0	0
0	0	2	2	1	0	0	0
2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
0	3	3	3	3	3	3	3
0	4	3	3	3	3	3	4
0	0	4	4	4	4	4	0
0	0	4	4	4	4	4	0

预测标签



cls0_iou

$$\frac{16}{20 + 18 - 16}$$

cls1_iou

$$\frac{3}{4 + 4 - 3}$$

cls2_iou

$$\frac{16}{16 + 18 - 16}$$

cls3_iou

$$\frac{12}{16 + 12 - 12}$$

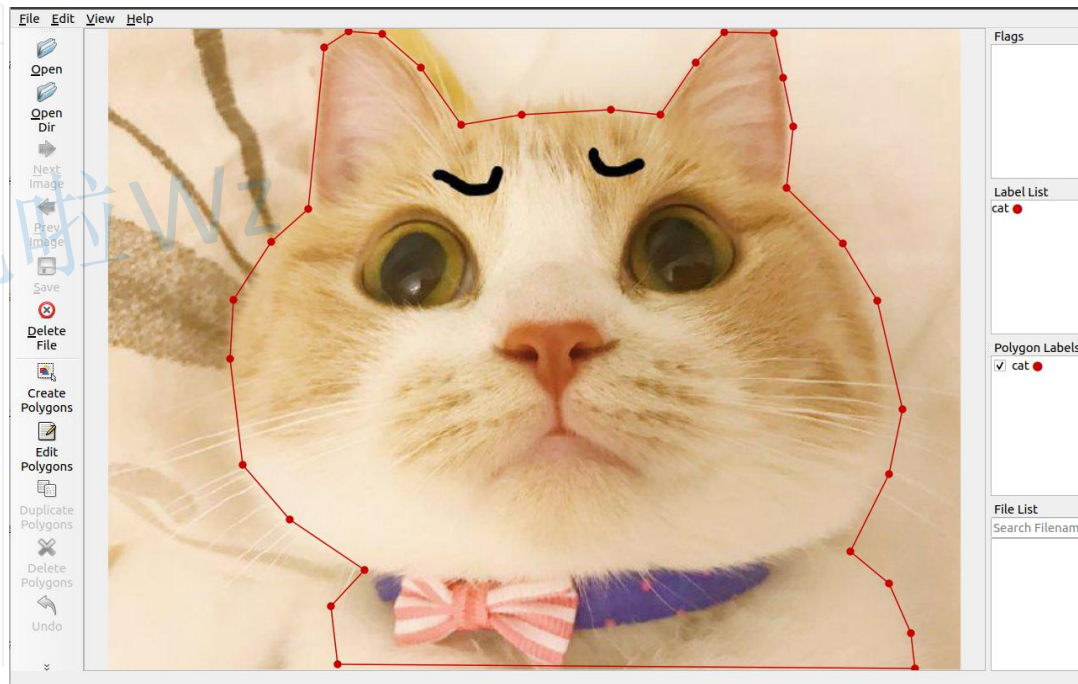
cls4_iou

$$\frac{8}{8 + 12 - 8}$$

语义分割前言

语义分割标注工具

Labelme <https://github.com/wkentaro/labelme>



参考博文: https://blog.csdn.net/qq_37541097/article/details/120162702

语义分割前言

语义分割标注工具

EISeg <https://github.com/PaddlePaddle/PaddleSeg>

EISeg

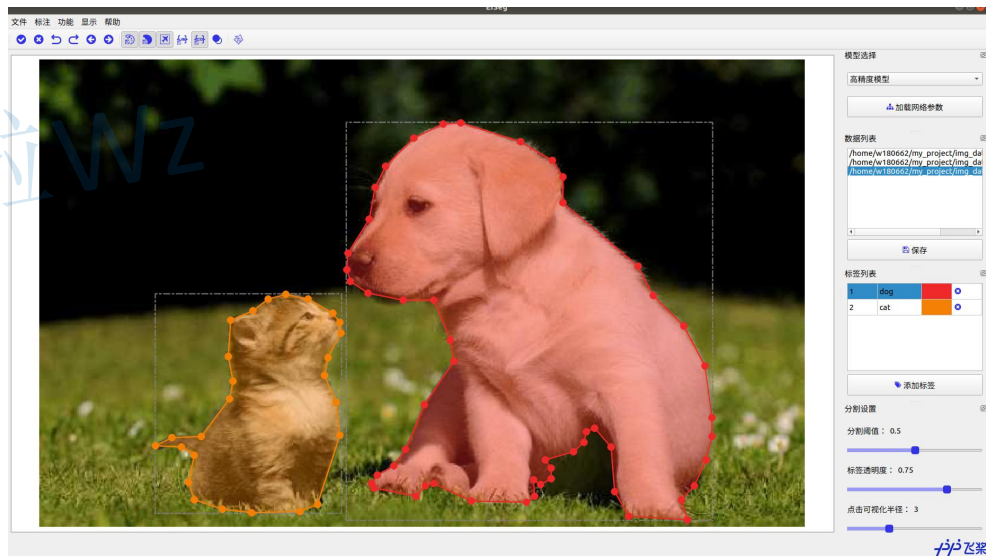
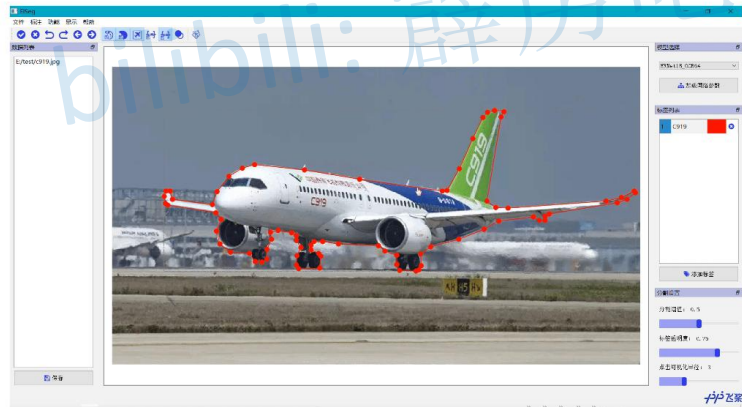
python 3.6+ license Apache 2

最新动向

- 支持多边形编辑，上线更多功能，最新EISeg 0.3.0推出。

介绍

EISeg(Efficient Interactive Segmentation)是基于飞桨开发的一个高效智能的交互式分割标注软件。涵盖了高精度和轻量级等不同方向的高质量交互式分割模型，方便开发者快速实现语义及实例标签的标注，降低标注成本。另外，将EISeg获取到的标注应用到PaddleSeg提供的其他分割模型进行训练，便可得到定制化场景的高精度模型，打通分割任务从数据标注到模型训练及预测的全流程。



参考博文: https://blog.csdn.net/qq_37541097/article/details/120154543

沟通方式

1.github

<https://github.com/WZMIAOMIAO/deep-learning-for-image-processing>

2.bilibili

<https://space.bilibili.com/18161609/channel/index>

3.CSDN

https://blog.csdn.net/qq_37541097/article/details/103482003