

图像分割

作者: Calvin

QQ: 179209347

Mail: 179209347@qq.com

介绍

笔记简介:

- 面向对象: 深度学习初学者
- 依赖课程: **线性代数, 统计概率**, 优化理论, 图论, 离散数学, 微积分, 信息论

知乎专栏:

<https://zhuanlan.zhihu.com/p/693738275>

Github & Gitee 地址:

https://github.com/mymagicpower/AIAS/tree/main/deep_learning

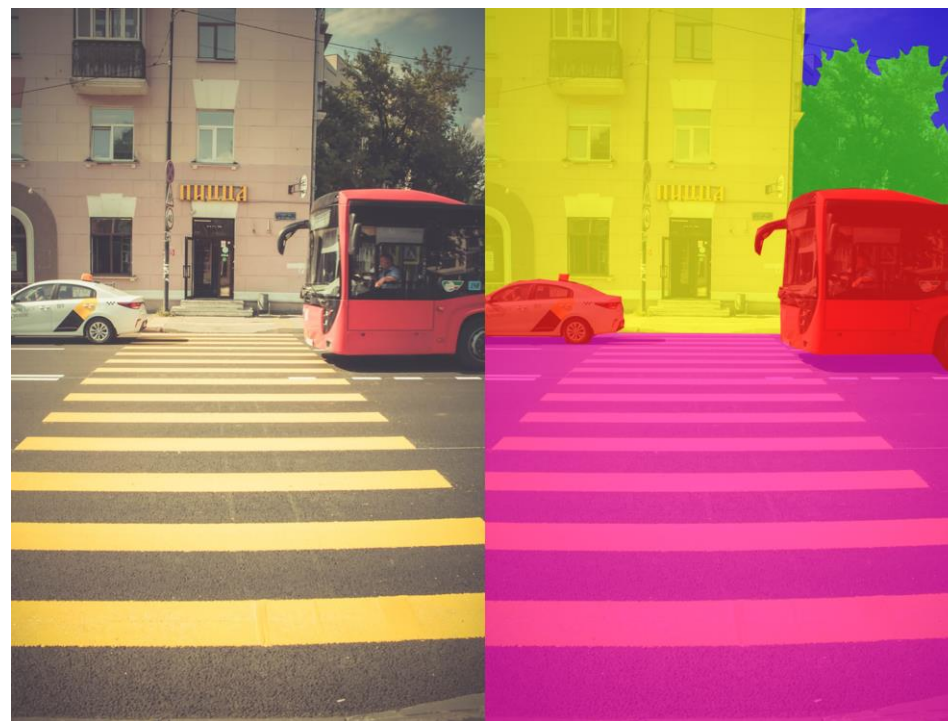
https://gitee.com/mymagicpower/AIAS/tree/main/deep_learning

* 版权声明:

- 仅限用于个人学习
- 禁止用于任何商业用途

图像分割 (Image Segmentation)

图像分割是计算机视觉领域中的重要任务，旨在将数字图像划分成多个具有语义信息的区域或对象。图像分割在许多应用中发挥着关键作用，如医学图像分析、自动驾驶、视频处理等。



图像分割算法概要

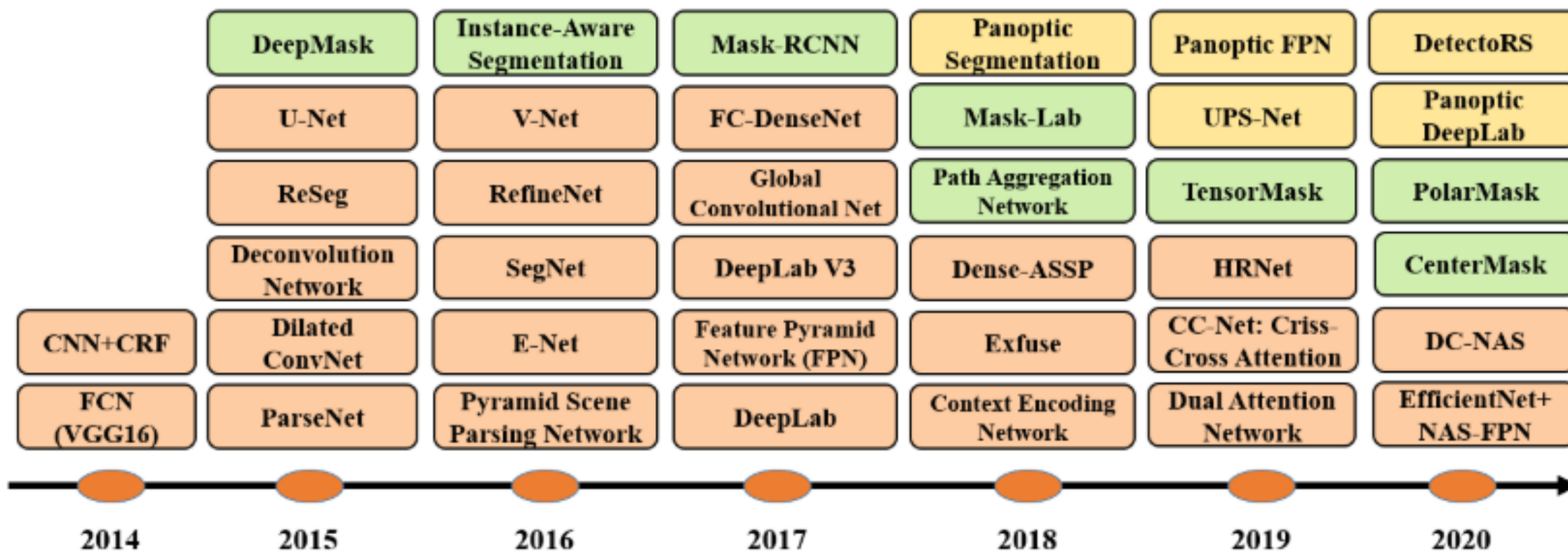
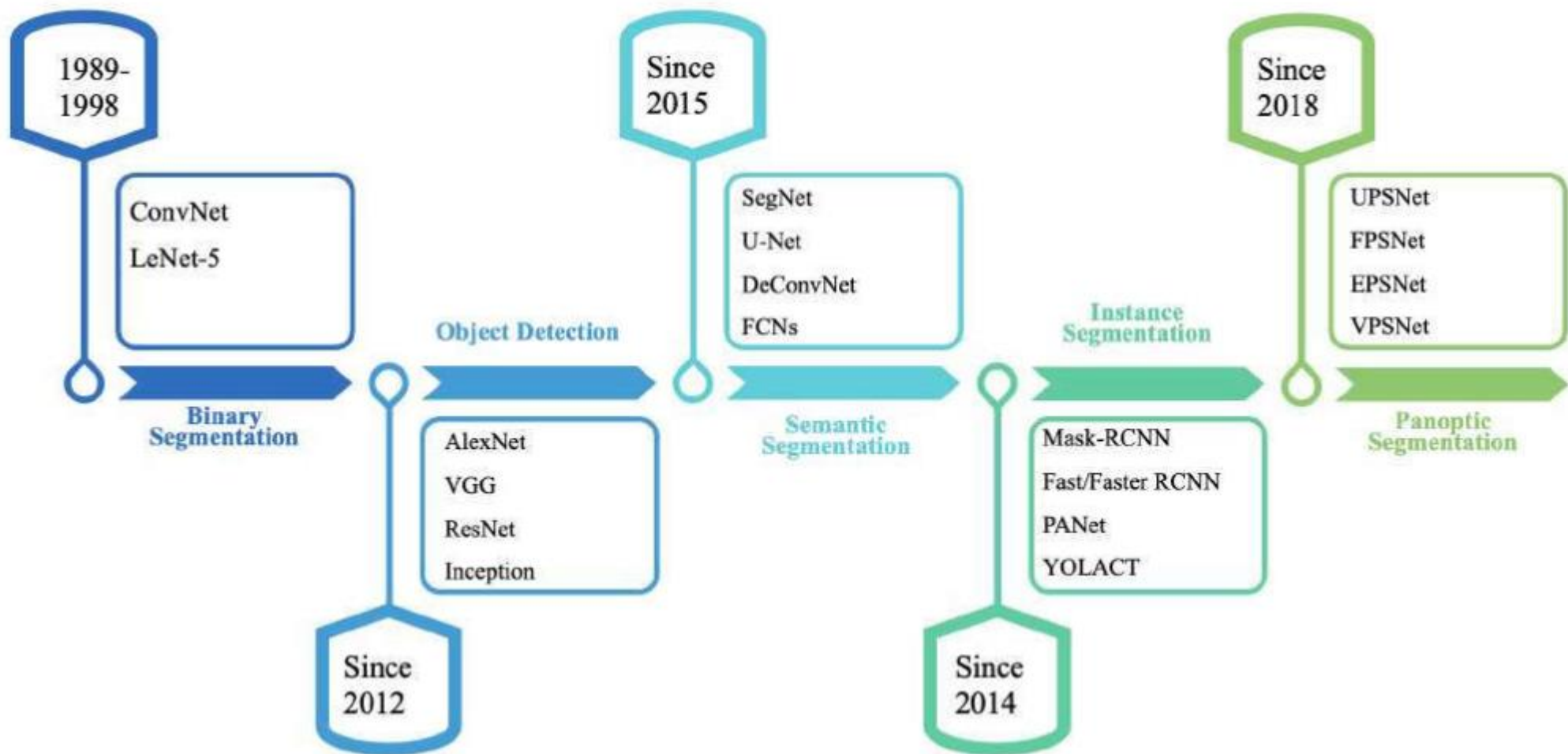


Image Segmentation Using Deep Learning: A Survey
<https://arxiv.org/pdf/2001.05566>

图像分割算法概要



Panoptic Segmentation: A Review
<https://arxiv.org/pdf/2111.10250>

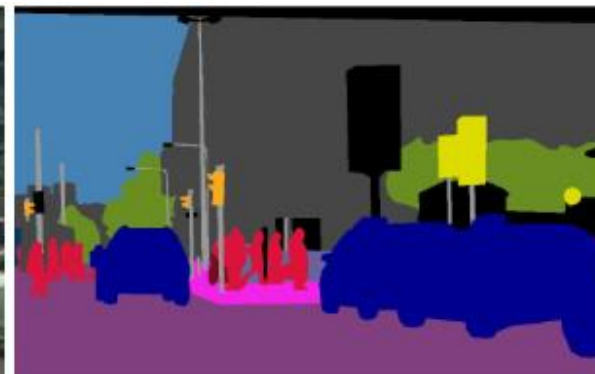
图像分割分类

图像分割主要分为三大领域：语义分割（Semantic Segmentation）、实例分割（Instance Segmentation）、全景分割（Panoptic Segmentation）：

- **语义分割**：每个像素对应一个类标签。同一类会被定义成一个区域块，不区分其中单个物体。
- **实例分割**：每个对象的掩码和类标签。区分单个物体以及单个物体所属的类型，无法识别的都作为背景。
- **全景分割**：每像素类+实例标签。相当于在语义分割的基础上，增加单个实例的区分。



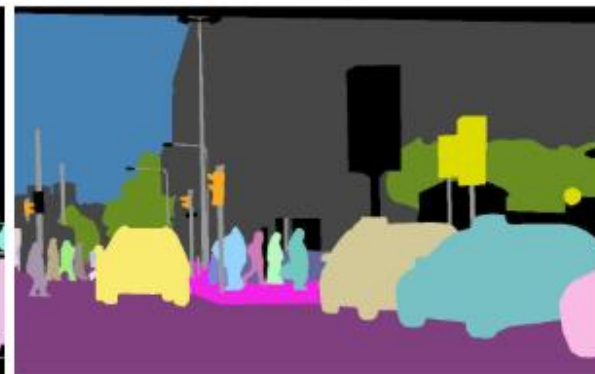
(a) image



(b) semantic segmentation



(c) instance segmentation



(d) panoptic segmentation

Panoptic Segmentation
<https://arxiv.org/pdf/1801.00868>

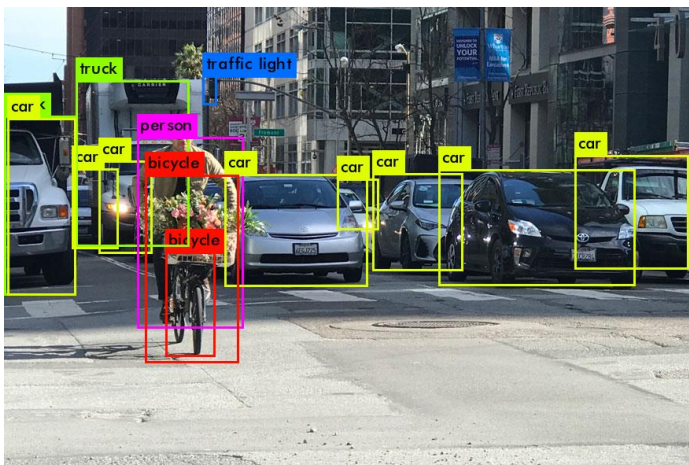
图像分割与目标检测的关系

共同点：图像分割和目标检测都是图像理解的重要任务，都涉及对图像内容的理解和解释。

区别：

- 图像分割更注重对图像中每个像素的语义理解，通常输出像素级别的标签；
- 目标检测更注重在图像中定位和识别特定目标，通常输出目标的位置和类别信息；
- 分割拥有更准确的位置信息。

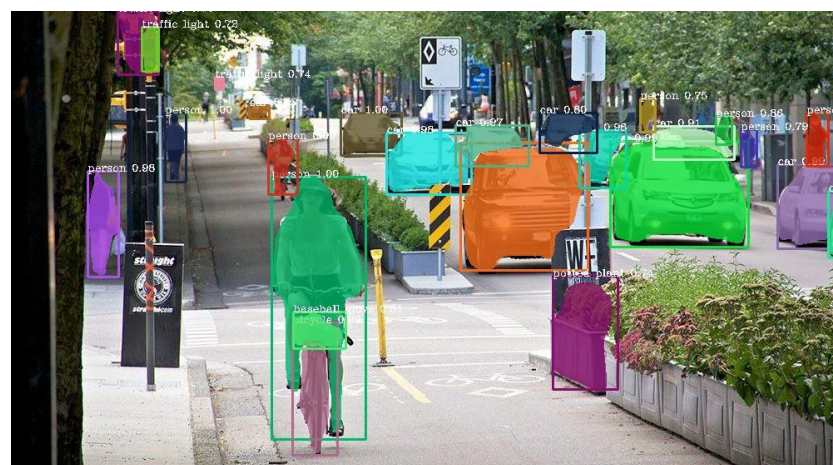
关联：在许多场景下，图像分割和目标检测可以相互促进，例如在目标检测中可以利用图像分割的结果来提高检测的准确性和鲁棒性。



检测任务



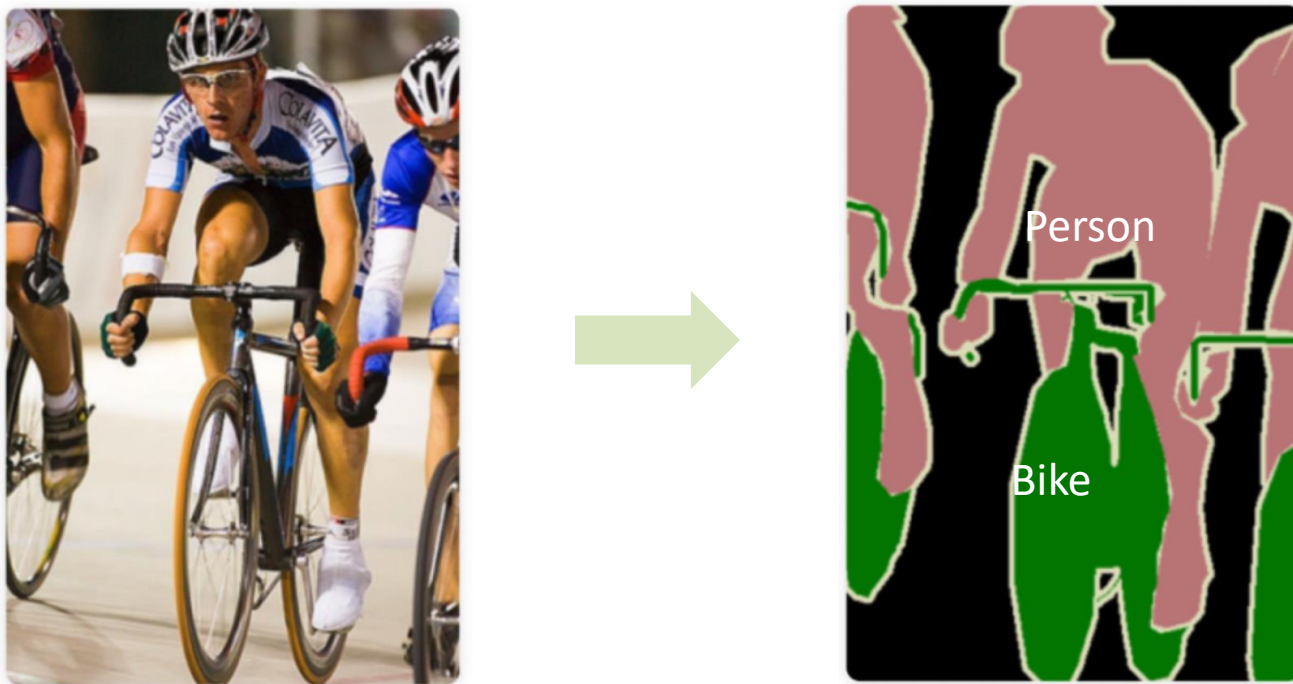
分割任务



检测 + 分割任务

图像分割 - 标签

类标签：将图像中的像素标记，使属于相同类型对象的像素被分配相同的标签。



语义分割

语义分割 (Semantic Segmentation)：每个像素对应一个类标签。同一类会被定义成一个区域块，不区分其中单个物体。

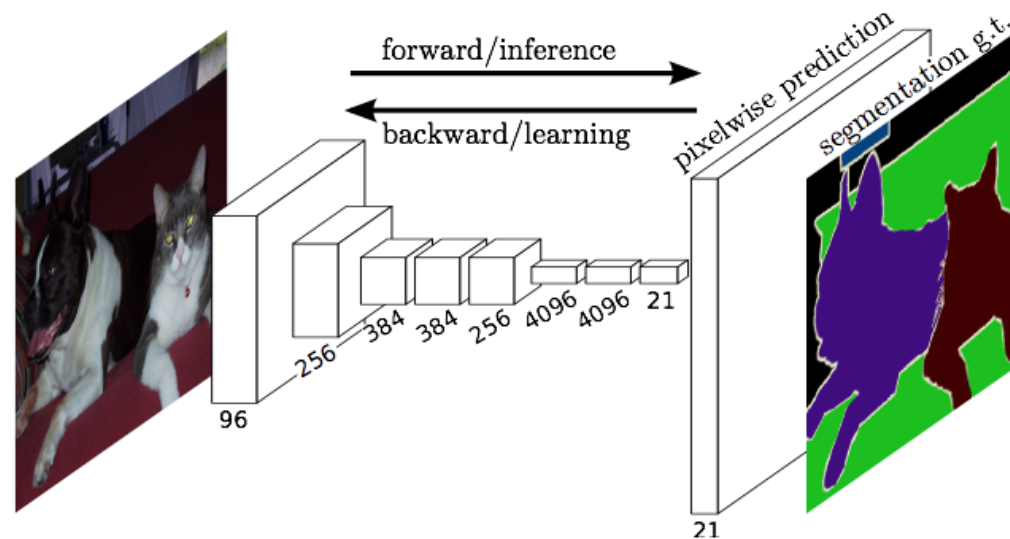
- **任务**：将图像中的每个像素分配到预定义类别中，从而实现对图像的像素级别理解和分类。
- **特点**：不区分不同物体的实例，只关注像素所属的语义类别，例如人、车、树等。
- **应用**：自动驾驶、医学图像分析、视频分析等领域。



语义分割 – 算法 – FCN简介

全卷积网络 (Fully Convolutional Networks, FCN) 是 UC Berkeley 的 Jonathan Long 等人于 2015 年提出的用于图像语义分割的一种框架。

该网络从图像级别的分类进一步延伸到像素级别的分类，解决了语义级别的图像分割问题。



Paper:

Fully Convolutional Networks for Semantic Segmentation

<https://arxiv.org/pdf/1411.4038>

Code:

<https://github.com/pytorch/vision/blob/main/torchvision/models/segmentation/fcn.py>

- **输入**可以为任意尺寸的彩色图像
- **全卷积部分**为一些经典的CNN网络（如VGG，ResNet等），并把最后的全连接层换成卷积，用于提取特征，形成热点图；
- **反卷积部分**则是将小尺寸的热点图上采样得到原尺寸的语义分割图像。
- **输出**与输入尺寸相同，通道数为 n （目标类别数）+1（背景）。

由于是分类问题，损失函数使用交叉熵损失函数。



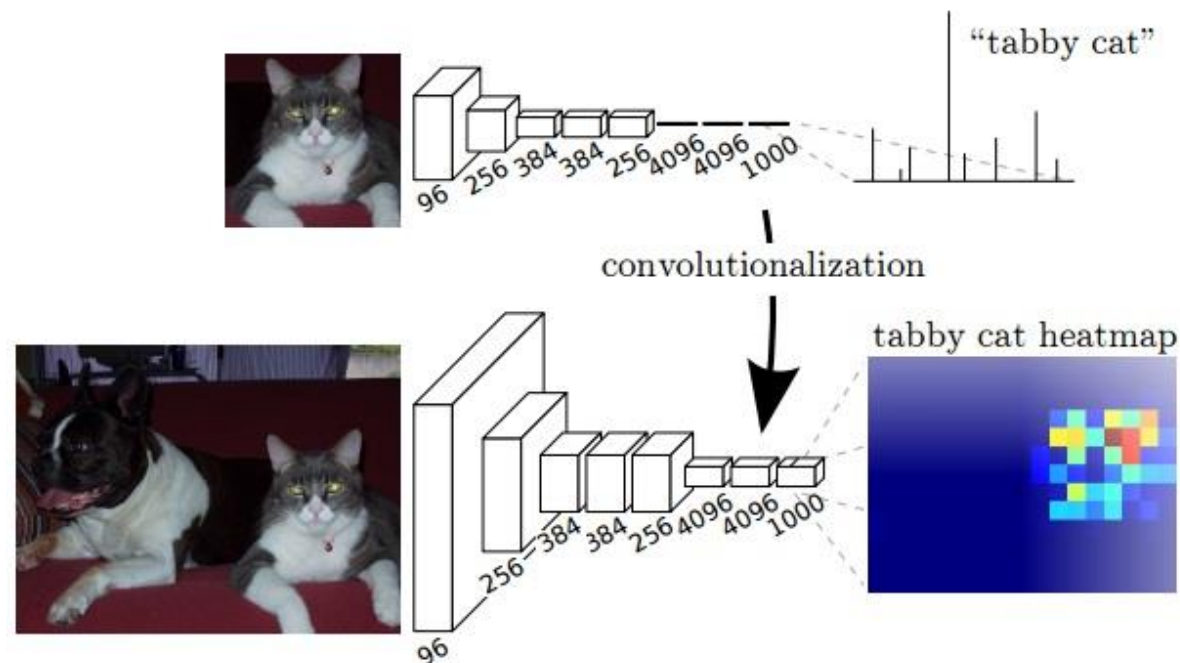
语义分割 – 算法 – FCN

FCN的优点:

- **空间信息保留**: 相比于传统的全连接层, FCN使用卷积层来替代, 可以保留输入图像的空间信息, 有利于更好地理解图像内容。
- **适应不同尺寸的输入**: FCN可以接受不同尺寸的输入图像, 因为卷积操作是位置不变的, 这使得模型更具有灵活性。

FCN的缺点:

- **计算量大**: 由于FCN需要处理整个图像的像素级别预测, 导致模型的计算量较大, 训练和推理时间较长。
- **空间信息不足**: 在处理大尺寸图像时, FCN可能会出现空间信息不足的问题, 导致分割结果不够精确。
- **空间一致性**: 未考虑像素与像素之间的关系, 缺乏空间一致性等。

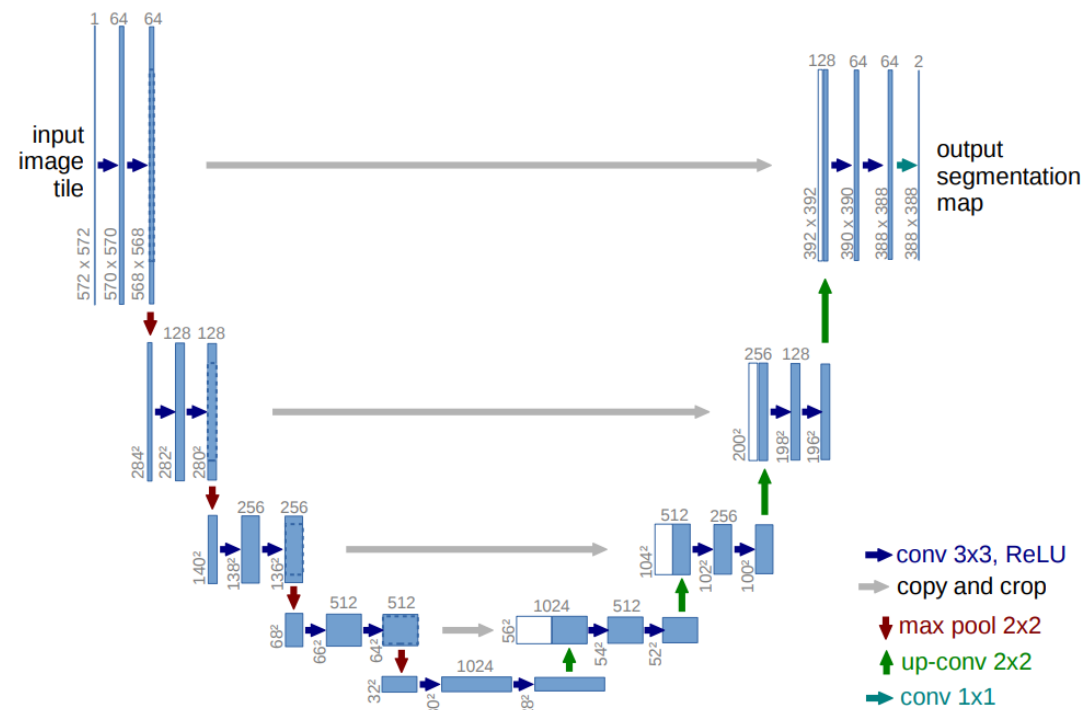


语义分割 – 算法 – U-Net

U-Net是一种用于图像分割的深度学习架构，由Olaf Ronneberger等人于2015年提出。它得名于其U形状的网络结构。U-Net在医学图像分割等领域取得了很大成功。

U-Net的特点包括：

- **U形结构**：U-Net由对称的编码器（下采样路径）和解码器（上采样路径）组成。
- **跳跃连接**：跳跃连接有助于解决梯度消失问题，同时帮助网络学习到更好的特征表示。这些连接使网络能够利用不同层级的特征信息。
- **损失函数**：在训练过程中，U-Net通常使用像素级别的损失函数，如交叉熵损失函数或Dice系数损失函数，以便更好地适应图像分割任务。
- **应用领域**：除了医学图像分割外，U-Net还被广泛应用于各种图像分割任务，如道路检测、卫星图像分析等。



Paper:

U-Net: Convolutional Networks for Biomedical Image Segmentation
<https://arxiv.org/pdf/1505.04597v1>

Code:

<https://github.com/milesial/Pytorch-UNet>

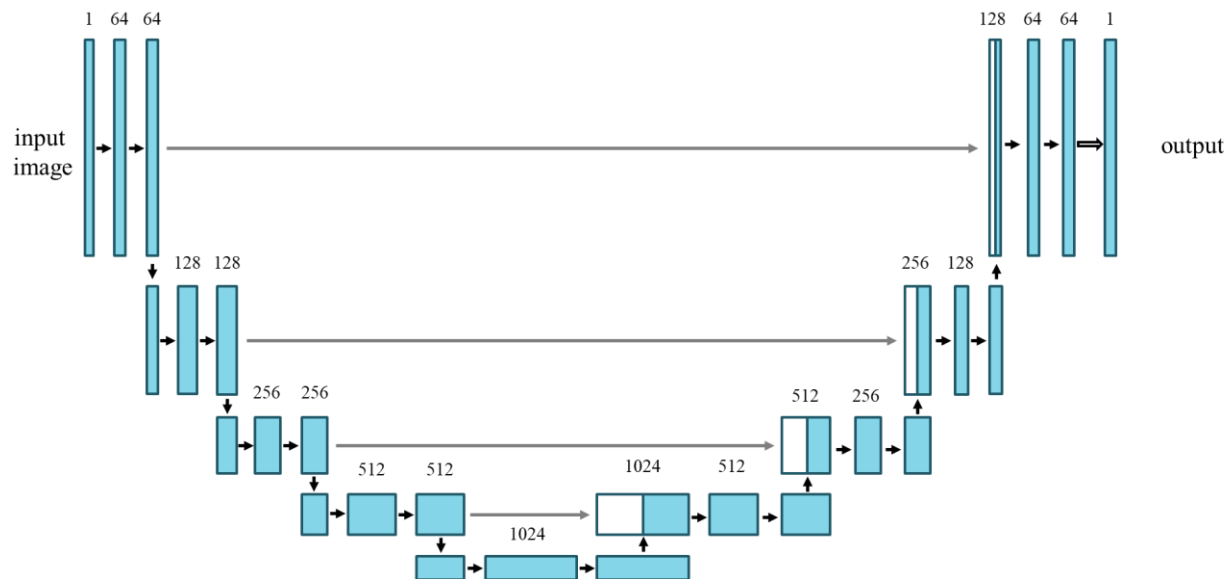
语义分割 – 算法 – U-Net

网络结构:

- **编码器:** 其中编码器部分用于捕获输入图像的上下文信息, 而解码器部分则用于将上采样的特征映射转换为与输入图像相同大小的分割结果。
- **解码器:** 解码器部分使用上采样操作来将特征图的尺寸恢复到原始输入图像的尺寸。
- **跳跃连接:** 将编码器中的特征图与解码器中对应的特征图连接起来。这有助于在解码器阶段保留更多的空间信息, 从而改善分割结果的准确性。
- **最终分类层:** 在U-Net的最后, 通常会添加一个1x1的卷积层, 将特征图映射到最终的分割结果。这一层通常使用softmax激活函数来输出每个像素属于每个类别的概率。

损失函数:

由于是分类问题, 损失函数使用交叉熵损失函数。

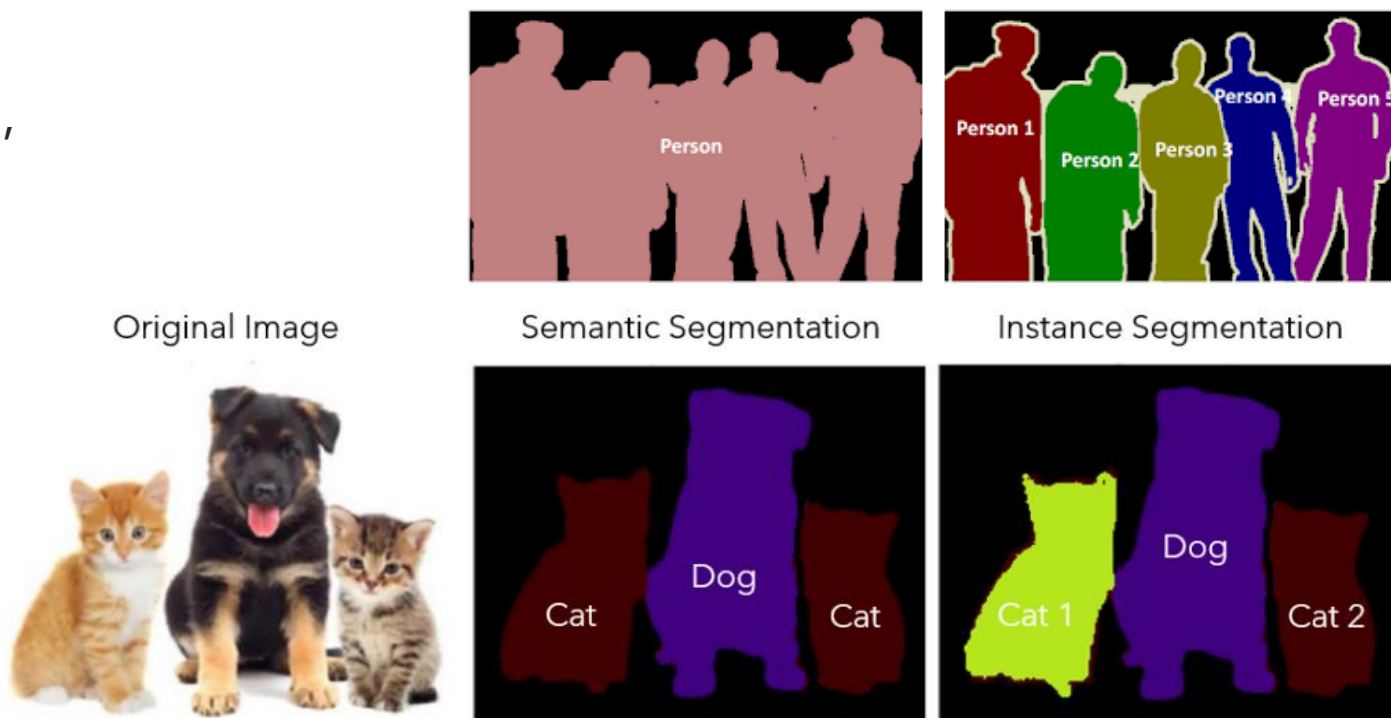


- ➔ Conv 3×3, ReLU
- ➡ copy and crop
- ↓ Max Pooling 2×2
- ↑ Up Conv 2×2
- ⇒ Conv 1×1

实例分割

实例分割 (Instance Segmentation)：在图像中同时识别和分割出不同物体的实例，即对每个物体进行像素级别的分割，并为每个实例分配唯一标签，无法识别的都作为背景。

- **任务**：在图像中同时识别和分割出不同物体的实例，即对每个物体进行像素级别的分割，并为每个实例分配唯一的标识符。
- **特点**：区分不同物体的实例，能够准确地标记出图像中的每个物体。
- **应用**：目标检测、视频分析、人体姿态估计等领域。

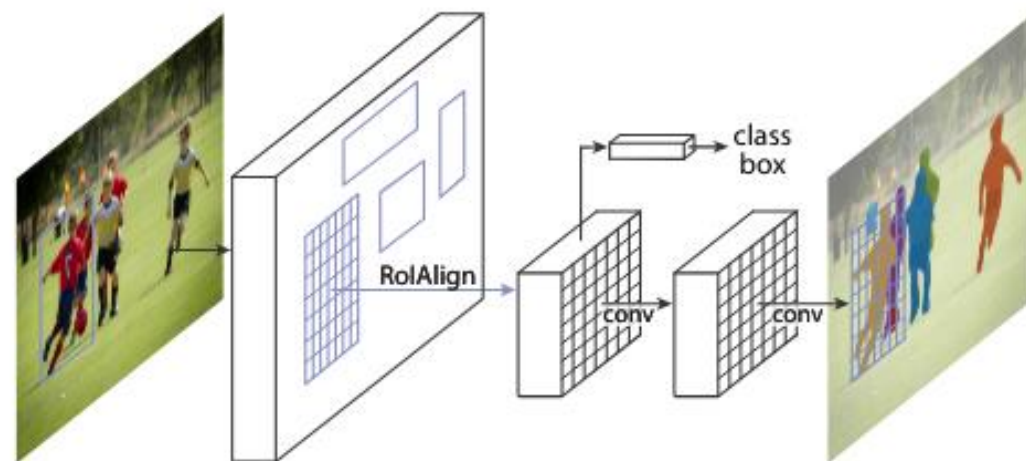


实例分割 – 算法 – Mask R-CNN

Mask R-CNN是一种用于实例分割（Instance Segmentation）的深度学习模型，它是在Faster R-CNN的基础上进行改进而来的。Mask R-CNN不仅可以检测图像中的对象，还可以为每个检测到的对象生成一个像素级的分割掩模（mask），从而实现对对象的精确分割。

Mask R-CNN的工作流程如下：

- 输入图像经过骨干网络提取特征。
- RPN根据这些特征生成候选对象边界框。
- 对每个候选对象边界框，Mask Head生成一个二值掩模，用于精确分割对象。



Paper:

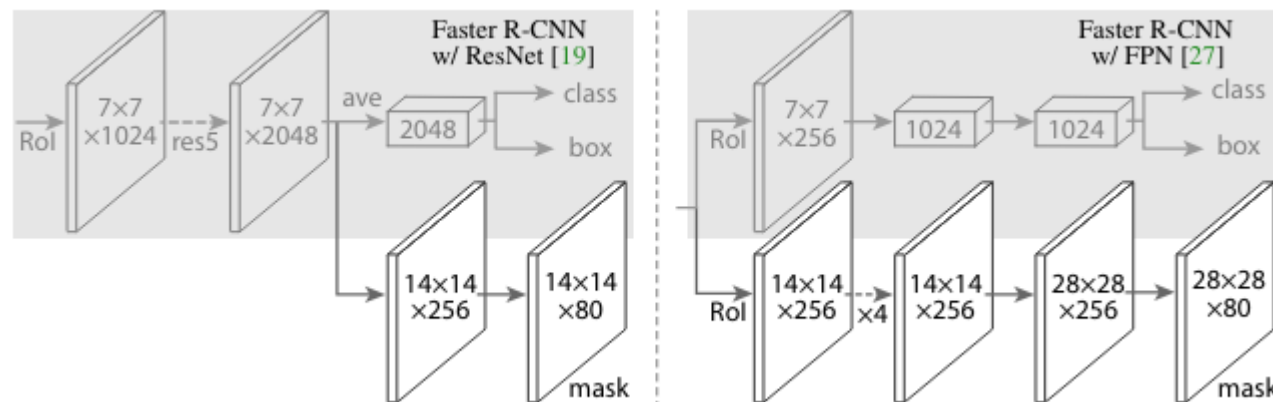
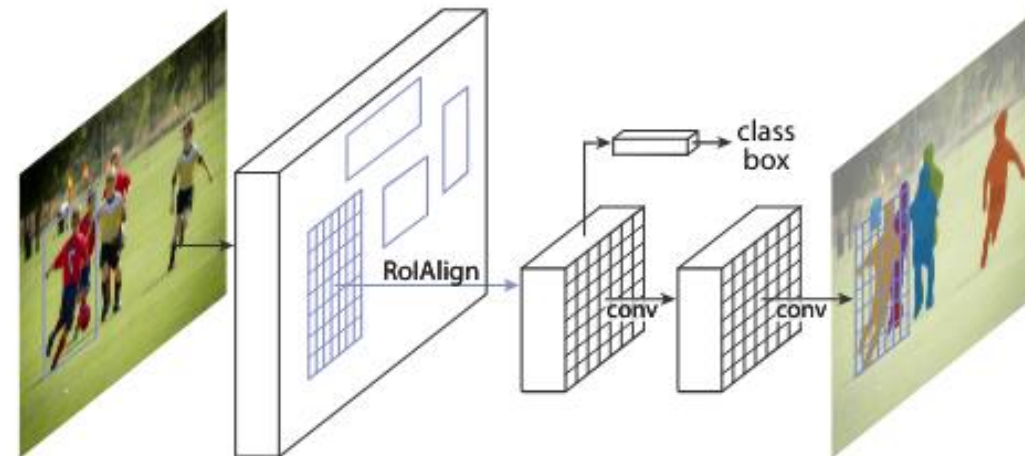
Mask R-CNN

<https://arxiv.org/pdf/1703.06870v3>

实例分割 – 算法 – Mask R-CNN

网络结构：

- **骨干网络 (Backbone Network)**，通常使用ResNet或者其他类似的卷积神经网络作为特征提取器。ResNet可以是：ResNet-50, ResNet-101, ResNeXt-50, ResNeXt-101；
- **区域建议网络 (Region Proposal Network, RPN)**，用于生成候选对象边界框；
- **头部网络 (Mask Head)**，用于为每个候选对象生成像素级的分割掩模。头部网络，包括边界框识别（分类和回归）+ mask预测。头部结构见右下图：



Head Architecture

全景分割

全景分割 (Panoptic Segmentation)：每像素类+实例标签。相当于在语义分割的基础上，增加单个实例的区分。

- **任务**：结合了语义分割和实例分割的优点，旨在同时对图像中的所有像素进行语义类别的标记，并将所有物体实例分割出来。
- **特点**：不仅考虑了物体的语义类别，还能够区分不同物体的实例，从而提供了更全面的图像理解。
- **应用**：场景理解、智能交通、视频分析等领域。

Code:

<https://github.com/elharroussomar/Awesome-Panoptic-Segmentation>

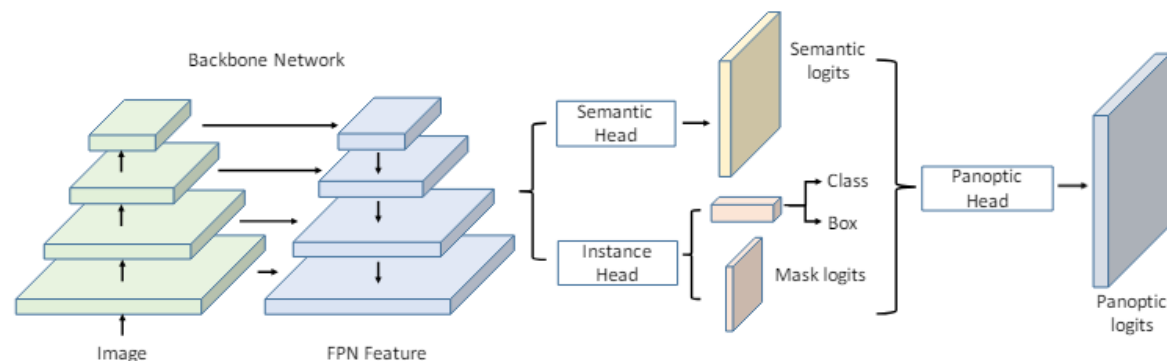


全景分割 – 算法 – UPSNet

UPSNet (Unified Panoptic Segmentation Network) 是一种用于图像分割任务的深度学习模型。它是由香港中文大学的研究人员提出的, 旨在解决全景分割 (panoptic segmentation) 任务, 即将实例分割 (instance segmentation) 和语义分割 (semantic segmentation) 结合起来的问题。

UPSNet的主要特点包括:

- **统一的架构:** 将实例分割和语义分割任务统一到一个网络中。
- **多尺度特征融合:** 利用多尺度特征融合的方法, 有效地捕获不同尺度下的语义信息, 提高了分割的准确性。
- **全景分割:** 在一张图像中同时标记出所有的物体实例和像素级的语义标签。



Paper:

UPSNet: A Unified Panoptic Segmentation Network
<https://arxiv.org/pdf/1901.03784>

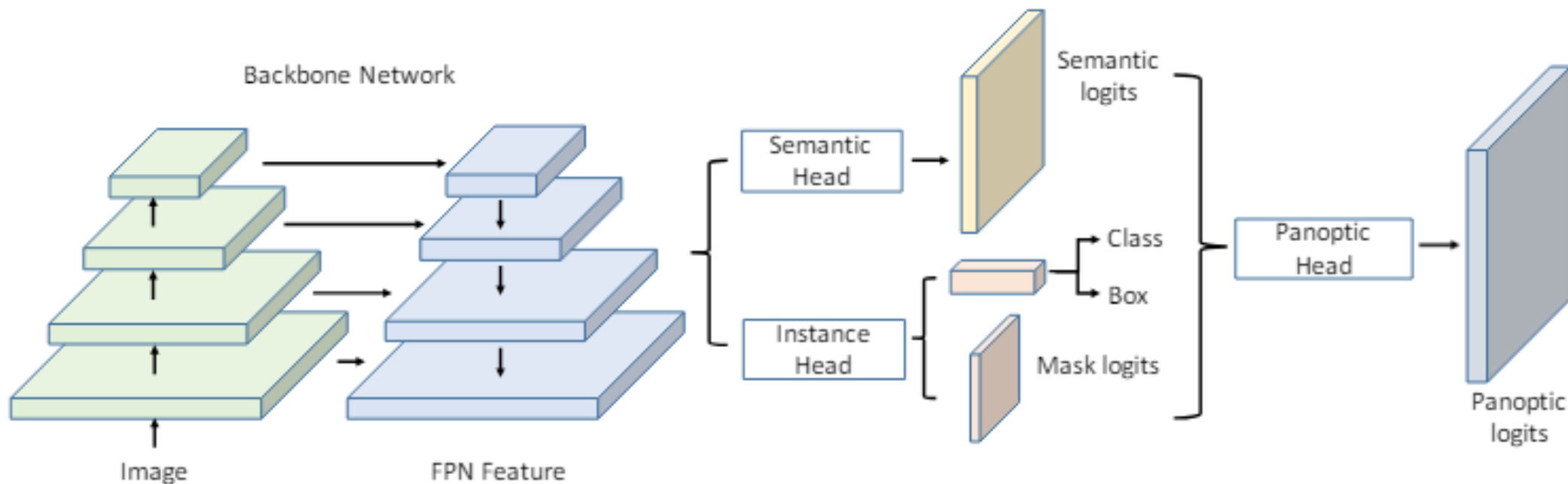
Code:

<https://github.com/uber-research/UPSNet>

全景分割 – 算法 – UPSNet

网络结构：

- **骨干网络 (Backbone Network)** 和Panoptic FPN网络结构非常类似，同样使用FPN作为backbone；
- **Semantic Head** 部分使用FCN；
- **Instance Head** 部分使用RPN、RCNN和Mask；
- **Panoptic Head** 对Semantic Head和Instance Head分支的输出使用concat整合。



图像分割应用场景

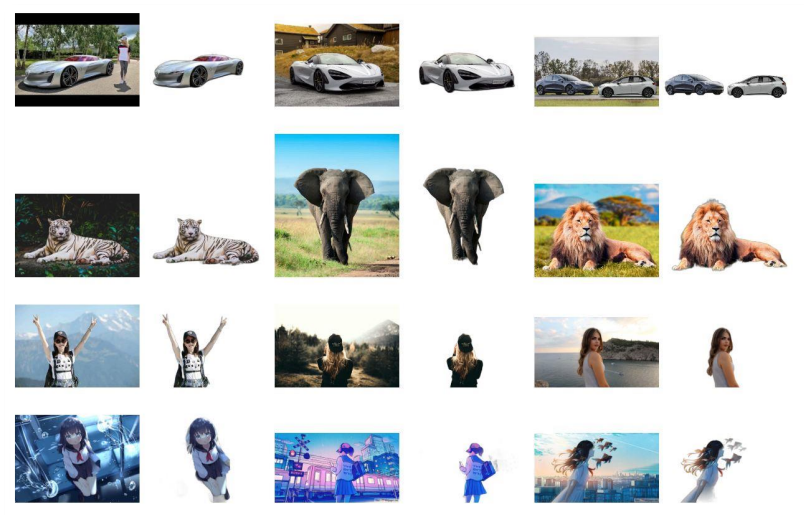
图像分割在许多应用中发挥着关键作用，如医学图像分析、自动驾驶、视频处理、抠图等。



医学图像分析

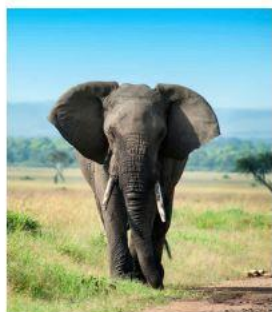


自动驾驶



一键抠图


图像分割应用场景 - 一键抠图





动漫一键抠图


无需手动绘制边界，大大提高了抠图的效率和精准度。应用场景如：


- 广告设计
- 影视后期制作
- 动漫创作等

 通用一键抠图【大模型】

 通用一键抠图【中模型】

 通用一键抠图【小模型】

 人体一键抠图

 动漫一键抠图

Dashboard / 动漫一键抠图



本地图片

选取文件
 上传抠图

图片格式: JPG(JPEG), PNG

图片链接

<https://aias-home.oss-cn-beijing.aliyuncs.com/assets/anime-girl.jpg>

在线抠图

人体一键抠图

人体一键抠图可以将人体从背景中抠出，形成一个透明背景的人体图像。这项技术应用场景：

- 社交媒体应用
- 电子商务
- 广告和营销
- 艺术创作
- 教育和培训

通用一键抠图【大模型】

通用一键抠图【中模型】

通用一键抠图【小模型】

人体一键抠图

动漫一键抠图

图片链接



在线抠图



Thank

You