

图像分割

- 定义
  - 将数字图像细分为多个图像子区域（像素的集合）的过程，并且同一个子区域内的特征具有一定相似性，不同子区域的特征呈现较为明显的差异
  - 为图像中的每一个像素点分类
- 任务
  - 描述
    - 将图像中每一像素标记后进行热编码，训练网络，进行预测
  - 类型
    - 语义分割
      - 图像中每个像素赋予一个类别标签
    - 实例分割
      - 要区分不同类别的像素，还需要需要对同一类别的不同个体进行区分

- 数据集
  - VOC
  - 城市风光数据集
  - 实例分割和语义分割
- 评价指标
  - 1.像素精度
    - 预测正确的像素个数比上所有的像素个数
  - 2.平均像素精度
    - 获取各个类别的像素精度求平均
  - 3.MIOU
    - 真实为i类并预测为i类的像素比上（真实为i类和预测为i类的并），对各个类别求平均
  - 常用的评价指标

- FCN（全卷积网络）
  - 网络架构
    - 是之后图像分割网络的基石
    - 将分类的CNN网络中最后的全连接层更改为卷积层后，输出与原图像大小相同的分割结果
  - 全卷积部分
    - 1.使用预训练的分类网络
    - 2.分类网络中的全连接层修改1x1的卷积层
    - 3.最终得到一个相对于原图像较小的三维的特征图
  - 上采样部分
    - 反卷积是一种特殊的正向卷积
    - 反卷积法
      - 通俗的讲，就是输入补0+卷积。先按照一定的比例通过补0来扩大输入图像的尺寸，再进行正向卷积即可
  - 跳层连接（skips layers）
    - 将最后一层的预测（有更富的全局信息）和更浅层（有更多的局部细节）的预测结合起来,进行图像的分割
  - 总结
    - 优点
      - 端到端的，可以接受任意大小的输入图像尺寸，比较高效
    - 局限性
      - 得到的结果还是不够精细
      - 对图像中的细节不敏感
      - 没有考虑像素与像素之间的关系

- Unet
  - 网络架构
    - 编码部分
      - 1.重复结构：2个3x3的卷积
      - 2.编码过程中通道数加倍
      - 3.输出
        - 1.下采样，送入下一次编码
        - 2.送入到解码阶段，进行特征的融合
    - 解码部分
      - 1.输入
        - 1.反卷积进行上采样的结果
        - 2.编码进行裁剪之后的结果
      - 2.重复结构：2个3x3的卷积
    - 连接部分
      - 2个3x3卷积

- 实例分割
  - MaskRCNN
    - 在FasterRCNN的基础上进行改进：ROIAlign和mask
  - 流程：
    - 1.特征提取：Resnet+FPN
    - 2.RPN获取候选区域
    - 3.映射到特征图上再进行ROIAlign得到固定维度的特征图
    - 4.进行分类，回归和分割
  - ROIAlign
    - 没有了取整操作，不会丢失像素信息
    - 实现：  
tf.image.crop\_and\_resize
  - Mask分支
    - 全卷积网络
    - 使用反卷积进行上采样，获取分割结果，映射回原图的ROI区域，大于阈值是目标像素，其他就不是
  - 损失函数
    - 分类
      - 交叉熵损失
    - 回归
      - smoothL1损失
    - 分割
      - 1.根据分类结果，选择相对应的分割结果，进行sigmoid激活
      - 2.计算二分类交叉熵损失