

练习 1：杆塔检测

任务介绍

- ◆ 背景：输电杆塔是电力网络的重要组成，由于长期处在暴露环境中，常出现倾斜、变形等异常情况。因此，有效检测巡检中拍摄到的杆塔目标，是保证其正常运行必不可少的环节。在本次练习中，将完成对杆塔这一电力场景中的基本目标的检测。
- ◆ 目标：检测图像中的杆塔目标。
- ◆ 输入：单张图像。
- ◆ 输出：杆塔目标包围框，及判定置信度。

数据集

- ◆ 共计 1280 张图像：训练集 896 张，验证集 128 张，测试集 256 张。
- ◆ 包含 tower 单类别，表示杆塔。
- ◆ 下载链接：<https://pan.baidu.com/s/1JmcC8w-l25ZGPuVrY-nefg>（提取码：juzb）

训练集

- 图像数量：896 张
- 图像位置：JPEGImages/train/*.jpg
- 标注文件：Annotations/train/*.xml
- 用途：训练杆塔检测模型

验证集

- 图像数量：128 张
- 图像位置：JPEGImages/val/*.jpg
- 标注文件：Annotations/val/*.xml
- 用途：用于验证模型的检测性能，使用 eval.py 进行本地测试

测试集

- 图像数量：256 张
- 图像位置：JPEGImages/test/*.jpg
- 用途：用于最终测试，未提供标注文件

评测指标

- ◆ AP：杆塔类别的 AP
 - AP：在不同 Recall 下的平均 Precision
 - 参考资料：<https://zhuanlan.zhihu.com/p/48693246>

标注文件格式

每张图像对应一个 xml 格式的标注文件。下面是一个 xml 文件的示例：

```
▼<annotation>
  ▼<size>
    <width>592</width>
    <height>800</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  ▼<object>
    <name>tower</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    ▼<bndbox>
      <xmin>229</xmin>
      <ymin>140</ymin>
      <xmax>386</xmax>
      <ymax>660</ymax>
    </bndbox>
  </object>
  ▼<object>
    <name>tower</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    ▼<bndbox>
      <xmin>43</xmin>
      <ymin>260</ymin>
      <xmax>171</xmax>
      <ymax>652</ymax>
    </bndbox>
  </object>
</annotation>
```

在当前图像上标注了 2 个 object，均为 tower（杆塔）目标。<size>标签指定了图像的大小，包括宽度、高度以及通道数；每对<object>标签对应图像中的一个标注实例；<name>标签指定了当前实例的分类名；<bndbox>标签指定了当前实例包围框的坐标信息（包含左上顶点与右下顶点的横纵坐标）。

检测结果保存格式

每张图像对应一个 txt 文件，用于存放检测结果。具体说明如下：

- txt 的文件名与图像的文件名相同。
- 每行对应一个检测实例（即检测到的杆塔目标）。
- 每行格式为 predicted_class confidence_score left top right bottom（通过空格分开）。其中，predicted_class 为预测的类别名（本练习中即为 tower），confidence_score 为预测的置信度，left、top、right 以及 bottom 为预测实例 bounding box 的顶点坐标。

作业说明

- 使用训练集中的数据训练模型，使用验证集进行验证。在有把握的情况下，可以重新划分训练、验证集，或采取适当的数据增强策略。
- 对测试集中的所有图像进行杆塔检测，提交检测结果，命名为“学号.zip”，存放测试集中各图像对应的检测结果文件。