行人检测

行人检测，就是将一张图片中的行人检测出来，并输出bounding box级别的结果。而如果将各个行人之间的轨迹关联起来，就变成了行人跟踪。而行人检索则是把一段视频中的某个感兴趣的人检索出来。行人再识别是行人分类问题。

目前行人检测领域的工作大致可被归为以下四类：

第一类是将传统的检测方法Boosting trees 和 CNN 结合起来。Is Faster R-CNN Doing Well for

Pedestrian Detection?（ECCV 2016）使用RPN 提取 proposal，同时提取卷积特征，然后使用 Boosting trees进行二次分类，性能得到了很大的提升。

第二类是解决多尺度问题，例如在视频数据中人的尺度变化问题。Scale-aware Fast R-CNN for Pedestrian Detection（ECCV2016）颜水成教授课题组提供了一种解决方法：训练两个网络，一个网络关注大尺度的人，另一个网络关注小尺度的人，在检测时将两个网络进行加权融合得到最终的结果，这样能使性能得到很大的提升；UCSD 在 ECCV 2016上有一个类似的工作，提出在高层提取大尺度人的特征，在低层提取小尺度人的特征，这样能保留尽量多的信息量，使得对小尺度的行人也有较好的检测效果。

第三类是使用语义分割信息来辅助行人检测。Semantic Channels for Fast Pedestrian Detection（CVPR2016）首先对整个图像进行语义分割，然后将分割的结果作为先验信息输入到检测网络中（包括传统的 ICF 网络，以及现在常用的CNN），这样可以通过对整体环境的感知来提高检测的效果。

第四类是解决黑暗条件下样本问题。Learning Cross-Modal Deep Representations for Robust Pedestrian Detection（CVPR2017）。训练时，给出一个多模态数据集，采样CNN学习非线性映射，对RGB图和热图之间关系建模。然后所学习得到的特征表示被迁移到第二个深度网络里，其接受RGB图像作为输入，并输出检测结果。测试时则不需要热量数据且仅考虑第二个网络。

数据集

Caltech dataset（11）

KITTI（12、13）