Pedestrian Detection Using Faster R-CNN

Ankit Shrestha

Introduction to Faster R-CNN

- ➔ Faster R-CNN is a deep learning model that helps computers find objects in images.
- ➔ It finds objects by:
  - ◆ First guessing where objects might be (Region Proposals).
  - ◆ Then checking those guesses to identify and improve boxes (Classification & Refinement).
- ➔ It uses two steps:
  - ◆ Region Proposal Network (RPN) generates possible boxes.
  - ◆ Fast R-CNN classifies the contents of those boxes.
- ➔ Both steps share the same features from the image, making the process faster and more efficient.

How Faster R-CNN Works

- ➔ Faster R-CNN starts with a backbone CNN (like ResNet50) that looks at the whole image and creates a feature map.
- ➔ The RPN scans this map and suggests where objects could be—these are called anchors.
- ➔ Then it resizes the proposed regions (ROI Pooling) and sends them to:
  - ◆ A classifier that determines what is inside.
  - ◆ A bounding box regressor that adjusts box positions.
- ➔ The whole model learns together by balancing how well it classifies and how well it finds the boxes.

Advantages

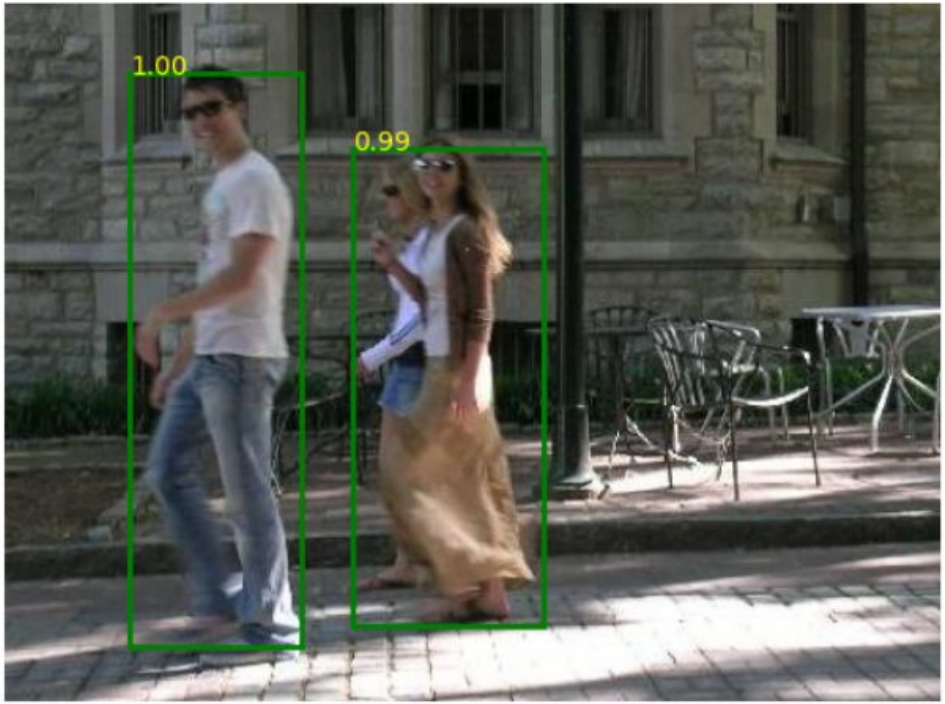
- ➔ Accurate object detection.
- ➔ Fast because of feature sharing.
- ➔ Works for many types of object detection tasks.

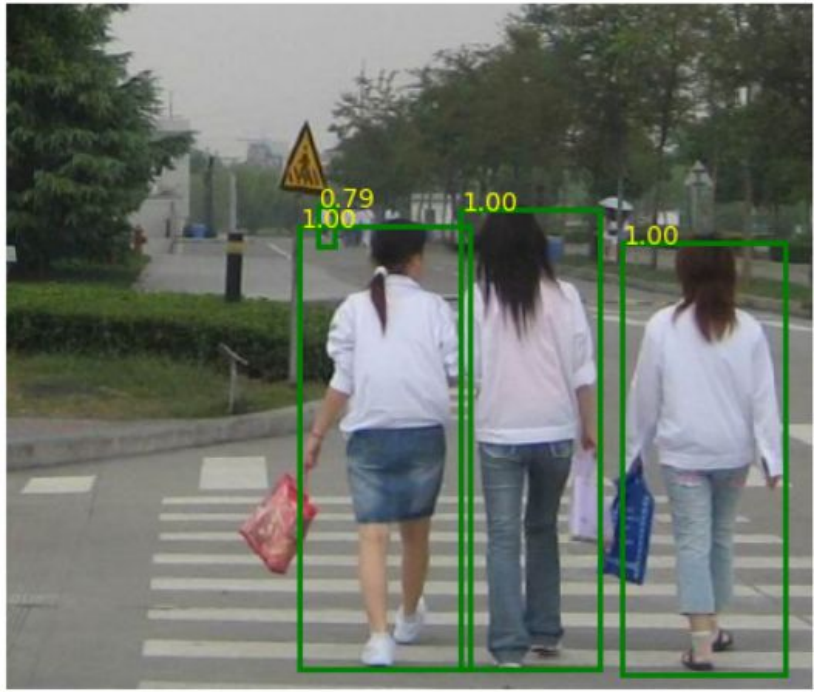
Disadvantages

- ➔ Slower than models that do detection all in one step (like YOLO).
- ➔ Needs more computing power to train.
- ➔ Has many settings to tune which can be complex.

Faster R-CNN for Pedestrian Detection

- ➔ Dataset Used
  - ◆ Penn-Fudan Database containing 170 images with 345 labeled pedestrians, among which 96 images are taken from around University of Pennsylvania, and other 74 are taken from around Fudan University.
- ➔ Split
  - ◆ 80% Train and 20% Test Data


1.00
0.99


0.79
1.00
1.00
1.00

Thank You !