

Real-time Courier service provider detection system using YOLO framework

DATA 606 - Capstone Project

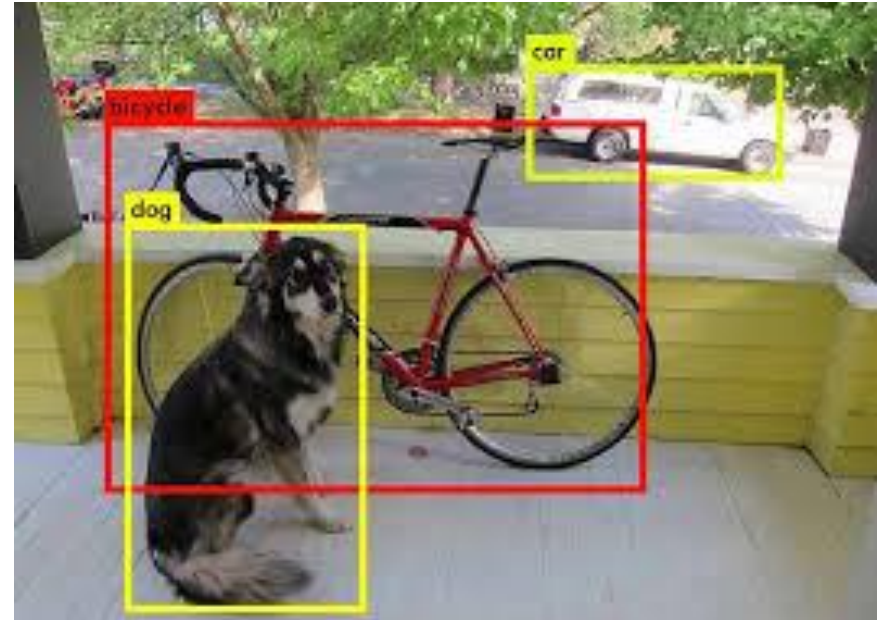
Shrimanth Ajjamane Manohar

Introduction

- Human vision is a vital sense that we possess which enables to see the outside world.
- Highly capable and hence a complex system of our body
- Reproducing the same with the use of computer is the field of computer vision.
- Computer vision is the field of computer science that enables computers to comprehend the digital images and videos in the same way as we do.
- Has applications including video surveillance, self-driving cars, object tracking, image processing in health sectors, etc.

Object Detection

- It is a computer vision technique for locating objects instances of objects in the images or videos.
- Makes use of deep neural networks to produce meaningful results.
- Classifies types of things found in the images and videos.
- This allows us to classify the objects and also locate them at once.



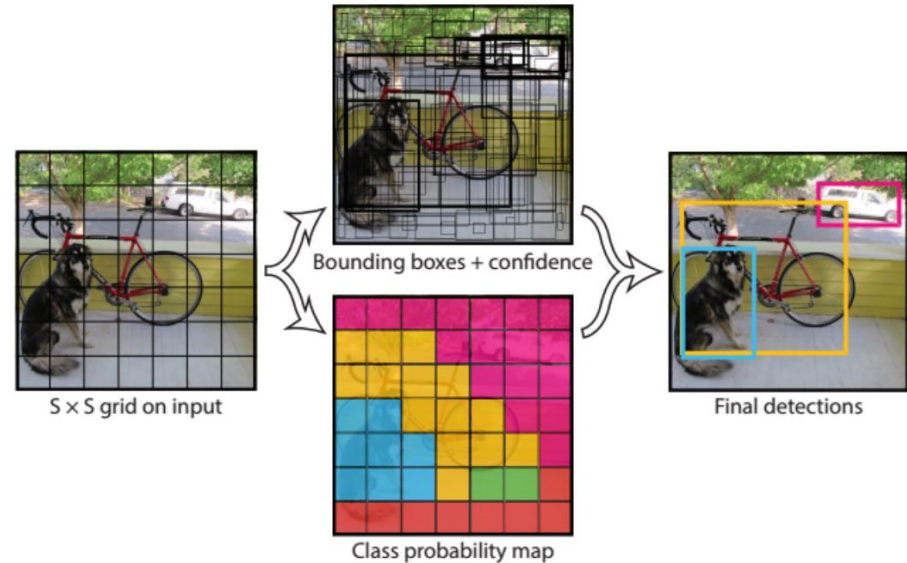
Aim

The aim of this project is to detect various courier packages that arrive at your doorstep by the images or videos captured by the Raspberry Pi camera. This falls under the category of Computer vision tasks and is a subset of object detection.

The generic process of object detection is brought down to detect the various courier service providers like DHL, FedEx, UPS, USPS, Amazon by identifying the logos on the packages along with the bounding boxes. I will be using the YOLO framework for object detection.

YOLO framework

- You only look Once: Unified Real-Time Object Detection
- State-of-the-art model known for its speed and accuracy
- Treats it as a regression problem
- Yolo system divides the input image into an $S \times S$ grid
- each grid cell predicts B bounding boxes, and C class probabilities.
- Predictions are combined



Methodology

1. Data : In order to implement the project, we have to train the model to detect these specific logos in the given image or video. As the data is scenario-specific, I will be creating custom images set for each of these companies that have the target logos in them.
 - Data collection: Scrapping of images with these logos from Google images using selenium. 200 images in each of the classes.
 - Data labeling: Annotations are vital in object detection, so I'm planning of using an open-source tool to label the images.
 - Once the training data is ready it has to be converted in the format how the YOLO framework accepts

Hardware

1. Raspberry Pi 4 Model B Rev 1.2
 - 32 Bit Processor
 - Raspbian GNU/Linux 10 (buster)
 - 4 GB RAM
 - No GPU
2. Pi Camera
 - Module v2
 - Sony IMX219 8-megapixel sensor

