

Acme Robotics Human Detector-Tracker

Naitri Rajyaguru, Mayank Joshi

Abstract-- Human detection and tracking is very renowned problem as it is an important part of many practical applications. In moving objects like robot where camera is installed, it becomes very crucial to detect obstacles and navigate through the environment safely. In this proposal, we present a real time human tracking and detection system which is more robust, accurate and lightweight for deployment. The implementation is designed as such for general obstacle avoidance but here we will specifically deal with humans as obstacles and plan to give location of human in robot reference frame.

I. INTRODUCTION

In real time scenario where videos are considered for either surveillance or autonomous navigation, the problem of detecting obstacles which are dynamic like humans is very difficult and hence crucial. Human detection is a part of object detection where primary class “human/person” is detected in images or video frames. In robotics, human detection is very necessary as humans are dynamic obstacles and no one would want a robot colliding with a human. Therefore, for navigation, a robot needs to detect and identify humans in its field of view. We propose a human detection and tracking module, which uses computer vision algorithms to perceive the environment. The idea is to generalize and to abstract algorithmic details from the system for better testing and results. The implementation incorporates object-oriented design where we provide three modes for operation i.e., training, testing and run-real time. It also incorporates the auto calibration mode where user doesn't need to calibrate camera for obtaining real time location of human.

II. SYSTEM IMPLEMENTATION

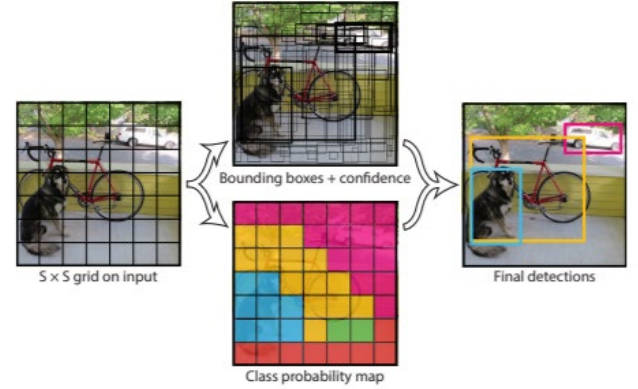
In this section, the details of our components in implementation are explained.

Detection Module

We use YOLO that is You Only Look Once real-time object detection system. YOLO is extremely fast and has many versions by far. It takes a single image for entire neural network. The network divides the image into small regions and predicts bounding boxes and confidence scores for those boxes. This reflects that how confident is the model about containing an object in the bounding box and how accurate is the prediction. It can recognize several objects in a single frame. We use MS COCO dataset that is Common Objects in Context, which is a

large-scale object detection, segmentation captioning dataset. It contains 164k split images.

Figure 1: The YOLO model [1]



Tracker Module

Object tracking is an application where the program takes initial detections and develops unique identification for each initial detection and tracks the objects as they move around in the frame. The tracker module will have detected human bounding box and its confidence value. Motion predictor is used to predict subsequent positions of each tracked object. Feature predictions are calculated using similarity scores between detection couplets. OpenCV based object tracker like KCF tracker is used. [6]

Pose transform module

This module provides output as location of human in robot reference frame. This helps to know where the object in real world and there are several applications based on this.

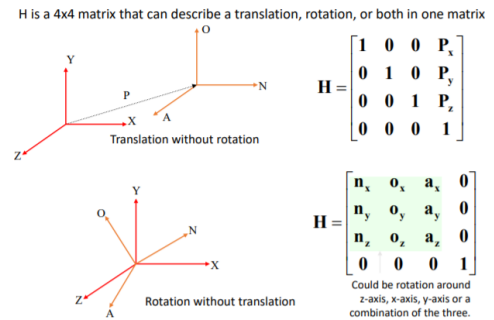


Figure 2: Homogenous transform for Pose [7]

Activity Diagram for Implementation

For building the proposed design, we present the flow of functionality as given below.

[6] <https://viso.ai/deep-learning/object-tracking/>

[7] <https://www.cs.cmu.edu/~16311/current/schedule/ppp/Le>
c17-FK.pdf

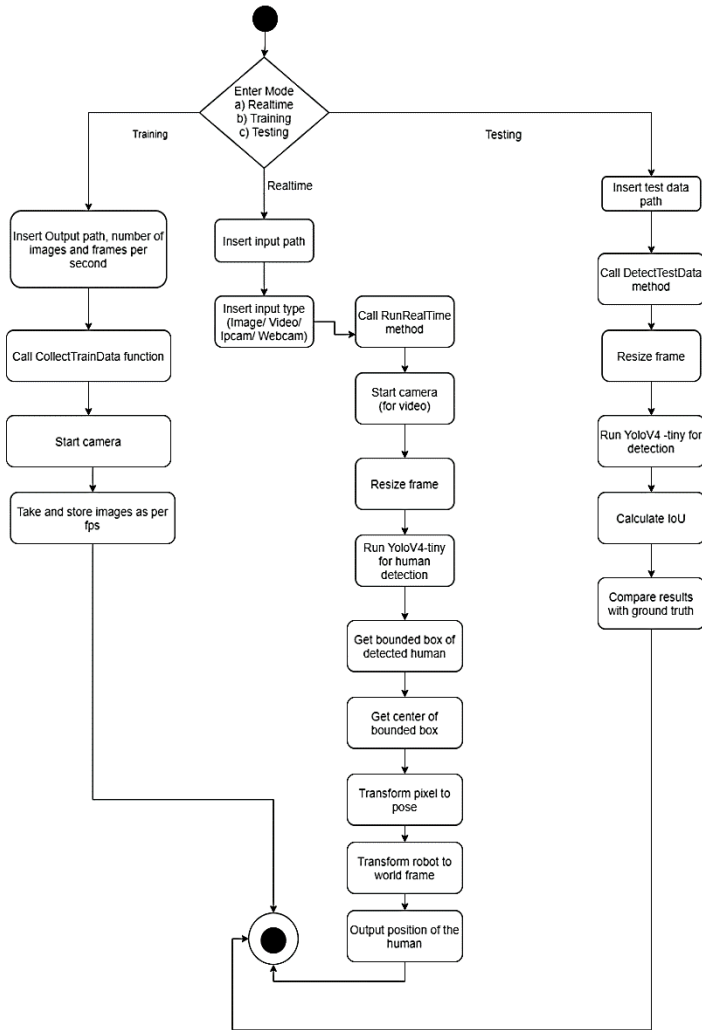


Figure 3: Activity Flow Diagram

III. REFERENCES

- [1] You Only Look Once: Unified, Real-Time Object Detection Joseph Redmon, Santosh Divvala, Ross Girshick , Ali Farhadi
- [2] YOLOv3: An Incremental Improvement Joseph Redmon Ali Farhadi
- [3] Object Detection with Deep Learning: A Review Zhong-Qiu Zhao, Member, IEEE, Peng Zheng
- [4] A Survey of Deep Learning-based Object Detection Licheng Jiao, Fellow, IEEE, Fan Zhang, Fang Liu, Senior Member, IEEE, Shuyuan Yang, Senior Member, IEEE, Lingling Li, Member, IEEE, Zhixi Feng, Member, IEEE, and Rong Qu, Senior Member, IEEE
- [5] Real-Time Human Detection, Tracking, and Verification in Uncontrolled CameraMotion EnvironmentsMohamed Hussein Wael Abd-Almageed Yang Ran Larry DavisInstitute for Advanced Computer StudiesUniversity of Maryland