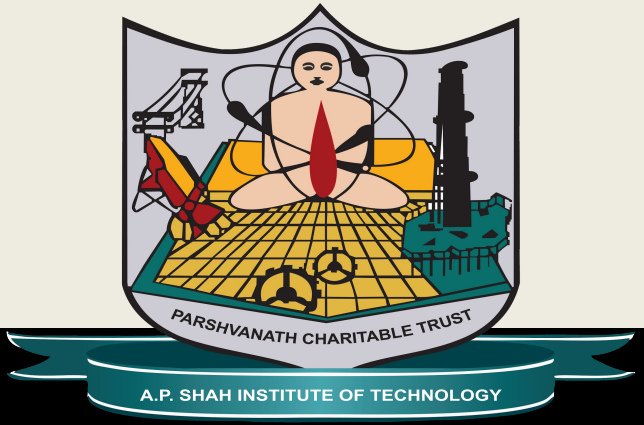


Department of Information Technology
NBA Accredited

A.P. Shah Institute of Technology,
G.B.Road, Kasarvadavli, Thane(W), Mumbai-400615
UNIVERSITY OF MUMBAI
Academic Year 2021-2022

A Project Report on

Real-Time Object Detection

Submitted in partial fulfillment of the degree of
Bachelor of Engineering(Sem-6)

in

INFORMATION TECHNOLOGY

By
Snehal Shanbhag - 19104008
Pranjali Shimpi - 19104017
Akansha Rawat - 19104007

Under the Guidance of
Prof. Sonal Jain
Prof. Charul Singh

1.Project Conception and Initiation

- Object detection is a technology that detects the semantic objects of a class in digital images and videos. One of its real-time applications is self -driving cars.
- In this, our task is to detect multiple objects from an image. The most common object to detect in this application is the car, motorcycle, and pedestrian and other objects in a roadway.
- For locating the objects in the image we use Object Localization and have to locate more than one object in real-time systems.

- There are various techniques for object detection, they can be split up into two categories, first is the algorithms based on Classifications. CNN and RNN come under this category.
- The second category is the algorithms based on Regressions. The YOLO method comes under this category.
- The YOLO algorithm is fast as compared to other classification algorithms.

1.1 Objectives

- To identify and locate one or more effective targets from still image or video data. It comprehensively includes a variety of important techniques, such as image processing, pattern recognition, artificial intelligence and machine learning.
- To automate the recognition and extraction process.
- To detect the object segment from the video frame.
- To classify the features in order to recognize the objects detected.

1.2 Literature Review

Table 2. Comparison between structures of YOLOv3, YOLOv4 and YOLOv5.

	YOLOv3	**YOLOv4**	**YOLOv5**
Neural Network Type	Fully convolution	Fully convolution	Fully convolution
Backbone Feature Extractor	Darknet-53	CSPDarknet53	CSPDarknet53
Loss Function	Binary cross entropy	Binary cross entropy	Binary cross entropy and Logits loss function
Neck	FPN	SSP and PANet	PANet
Head	YOLO layer	YOLO layer	YOLO layer

Table 1. Comparison of YOLO with related works.

Reference	Dataset Used	Algorithms	Findings
Li et al., 2021 [26]	Remote sensing images collected from GF-1 and GF-2 satellites. Training: 826 images. Testing: 275 images. Resolution: 300 × 300, 416 × 416, 500 × 500, 800 × 800, 1000 × 1000	Faster R-CNN YOLO v3 SSD	YOLOv3 has higher mAP and FPS than SSD and Faster R-CNN algorithms.
Benjdira et al., 2019 [12]	UAV dataset Training: 218 Images Test: 52 Images Resolution: 600 × 600 to 1024 × 1024	Faster R-CNN YOLOv3	YOLOv3 has higher F1 score and FPS than Faster R-CNN.
Zhao et al., 2019 [27]	Google Earth and DOTA datasetTraining: 224 Images Test: 56 Images Resolution: 600 × 600 to 1500 × 1500	SSD Faster R-CNN YOLOv3	YOLOv3 has higher mAP and FPS than Faster R-CNN and SSD.
Kim et al., 2020 [29]	Korea expressway dataset Training: 2620 Test: 568 Resolution: NA	YOLOv4 SSD Faster R-CNN	YOLOv4 has higher accuracy SSD has higher detection speed
Dorrer et al., [28]	Custom Refrigerator images Training: 800 Images Test: 70 Images Resolution: NA	Mask RCNN YOLOv3	The detection of YOLOv3 was 3 times higher but the accuracy of Mask RCNN was higher.
Rahman et al., [13]	Custom Electrical dataset Training: 5939 Test: 1400 Resolution: NA	YOLOv4 YOLOv5l	YOLOv4 has higher mAP compared to YOLOv5l algorithms
Long et al., [30]	MS COCO dataset Training: 118,000 Test: 5000 Resolution: NA	YOLOv3 YOLOv4	YOLOv4 has higher mAP compared to YOLOv3
Bochkovskiy et al., [7]	MS COCO dataset Training: 118,000 Test: 5000 Resolution: NA	YOLOv3 YOLOv4	YOLOv4 has higher mAP and fps than YOLOv3
Ge et al., [14]	MS COCO dataset Training: 118,000 Test: 5000 Resolution: NA	YOLOv3 YOLOv4 YOLOv5	YOLOv5 has higher mAP than YOLOv3 and YOLOv5l YOLOv3 has higher FPS than YOLOv4 and YOLOv5l

1.3 Problem Definition

To make a ML project on Real Time Object Detection with the best algorithm such that the project focuses on accuracy of the output, detection of 10 classes or as much as we can.

1.4 Scope

Our topic mainly focuses on Real Time Object Detection related to Road:

- Detection of pedestrians
- Detection of vehicles
- Detection of other roadway objects i.e. traffic signs and lights, etc.

1.5 Technology stack

Software requirements:

- Frontend-Flask, HTML, CSS
- Backend - Python
- OpenCV library
- Yolo v5 algorithm
- Open-source dataset

Hardware requirements:

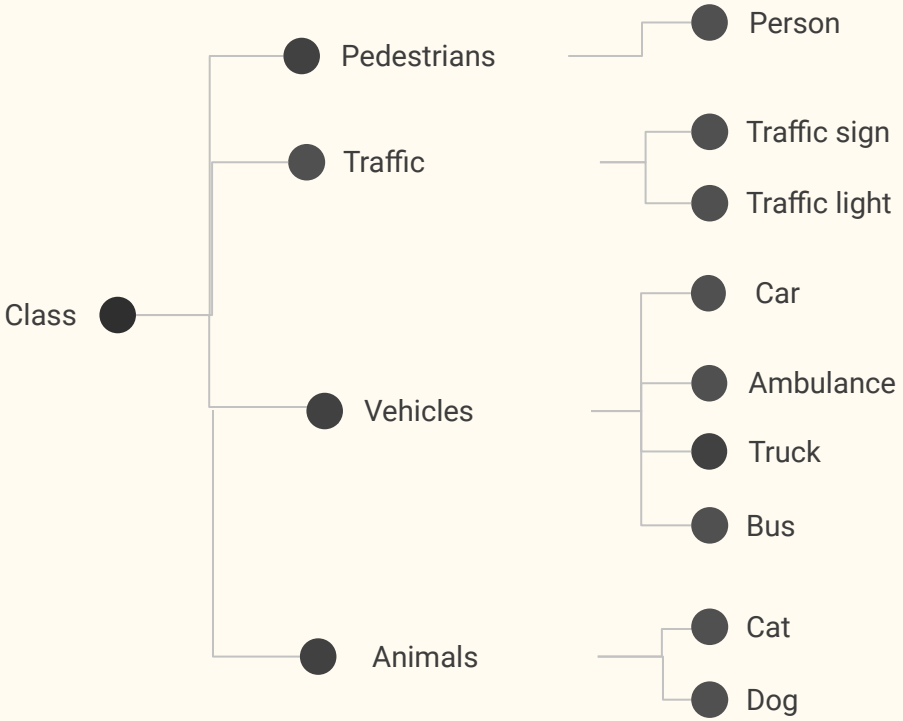
- GPU
- Nvidia driver
- RAM - 8GB or more

2. Project Design

2.1 Proposed System

- The proposed method uses these Yolo v5 to develop a system model which consists of multilayers to classify the given objects into any of the defined classes.
- The schemes then use multiple images and detect the objects from these images, labeling them with their respective class label.

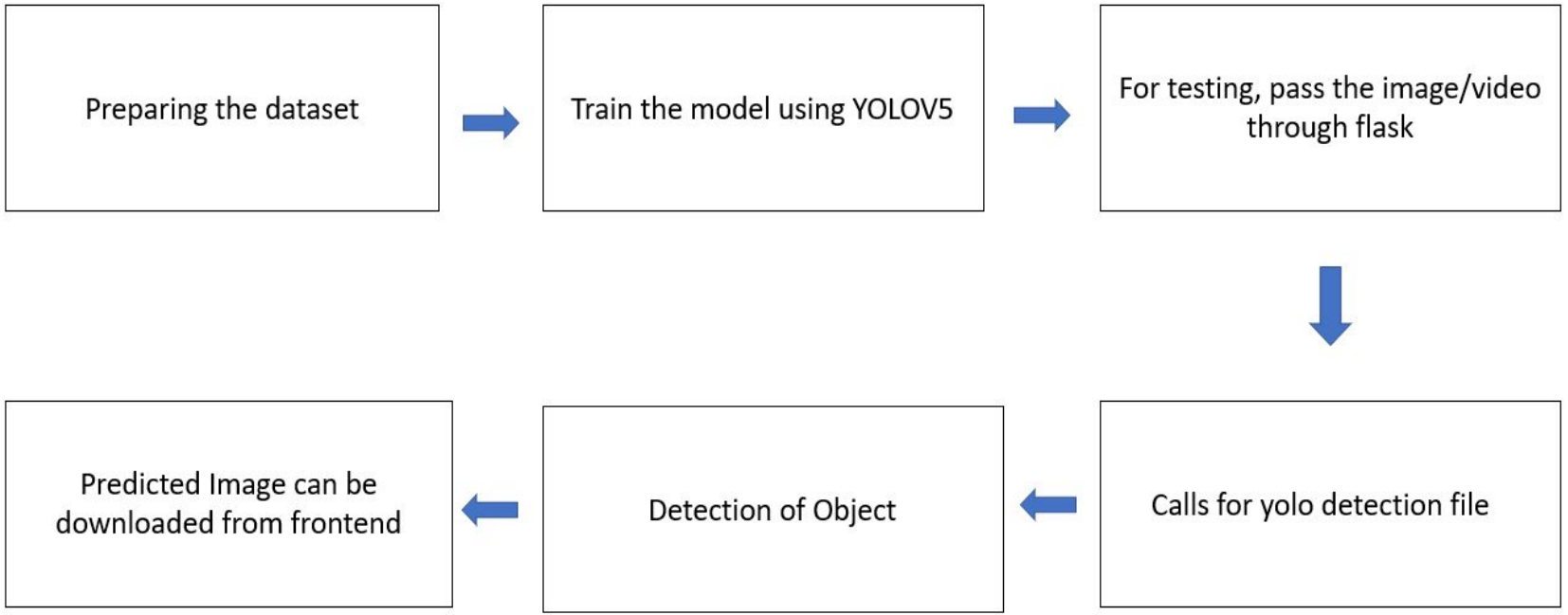
2.1 Proposed System



2.2 Design (Flow Of Modules)

- Flask run command will call app.py
- App.py- It will initiate the process of frontend to take input from users
- Detect.py- This file is responsible for detection of objects
- Download.html and run/detect- This will create a download button on frontend for downloading the output file

2.3 Block Diagram

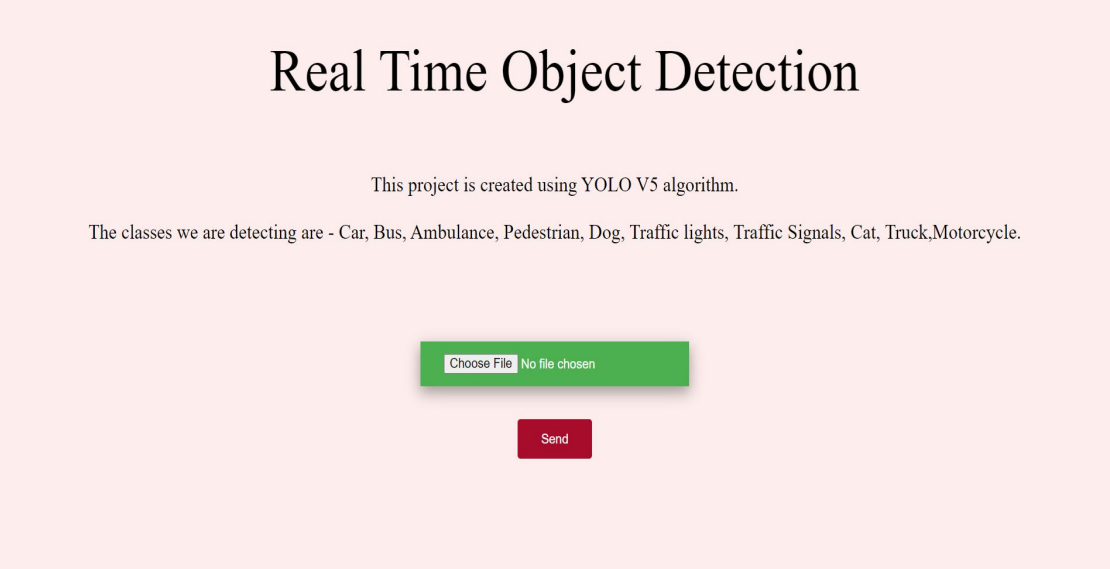


3. Implementation

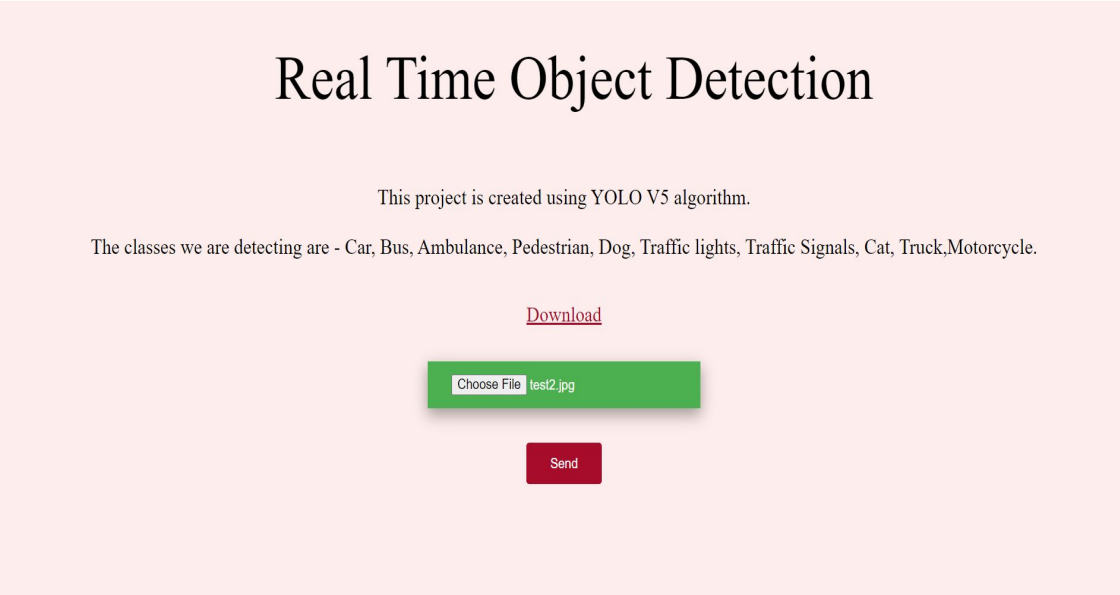
- The outcome of object detection project is to recognize and locate all trained objects in a frame related to road.
- Object detection is implemented in two ways:
  - Video streaming
  - Through images
- User can upload image or video that is to be detected through the web application and get the desired output with certain accuracy

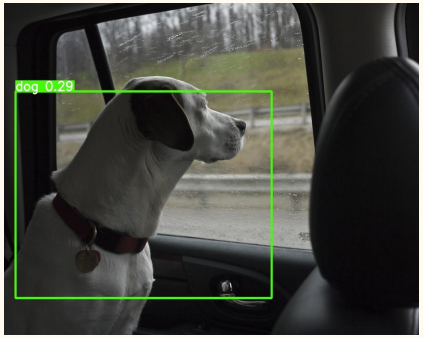
5. Result

Frontend interface:



Output on frontend:




dog 0.29


person 0.
person 0

6. Conclusion and Future Scope

Conclusion

- Being a unified object detection model that is simple to construct and train in correspondence with its simple loss-function, YOLO can train the entire model in parallel.
- YOLO is also better at generalizing Object representation compared with other object detection models and can be recommended for real-time object detection.
- Comparing Yolo v3, v5, and v5 version - Yolov5 turned out to be the best algorithm taking less time and giving max accuracy.

Future Scope

Can be used in :

- Biometric recognition
- Surveillance
- Smart cars
- Lane detection
- Medical analysis

References

- https://leadingindia.ai/internshipproject
- https://opencv.org/multiple-object-tracking-in-realtime/
- https://towardsdatascience.com/implementing-real-time-object-detection-system-using-pytorch-and-opencv-70bac41148f7
- https://www.mdpi.com/1424-8220/22/2/464/pdf
- https://www.researchgate.net/publication/351411017_Real-Time_Object_Detection_Using_YOLO_A_Review

Thank You