A Mini- Project Report

on

“Object Detection”

Submitted to the

Pune Institute of Computer Technology, Pune

In partial fulfillment for the award of the Degree of

Bachelor of Engineering

in

Information Technology

by

Aditya Chaloo	43302
Amod Dhopavkar	43304
Atharva Joshi	43307

Under the guidance of

Prof. R. R. Chhajed

Department Of Information Technology

Pune Institute of Computer Technology College of Engineering

Sr. No 27, Pune-Satara Road, Dhankawadi, Pune - 411 043.

2020-2021

CERTIFICATE

This is to certify that the project report entitled

Object Detection

Submitted by

Aditya Chaloo	43302
Amod Dhopavkar	43304
Atharva Joshi	43307

is a bonafide work carried out by them under the supervision of Prof. R. R.Chhajed and it is approved
for the partial fulfillment of the requirement of **Computer Laboratory -X** for the award of the Degree of Bachelor of Engineering (Information Technology)

Prof. R. R. Chhajed
Lab Teacher
Department of Information Technology

Dr. A. M. Bagade
Head of Department
Department of Information Technology

Place:
Date:

ACKNOWLEDGEMENT

We thank everyone who has helped and provided valuable suggestions for successfully creating a wonderful project.

We are very grateful to our guide, Prof. R. R. Chhajed, Head of Department Dr. A. M. Bagade and our principal Dr. P. T. Kulkarni. They have been very supportive and have ensured that all facilities remained available for smooth progress of the project.

We would like to thank our professor and Prof. R. R. Chhajed for providing very valuable and timely suggestions and help.

Aditya Chaloo

Amod Dhopavkar

Atharva Joshi

ABSTRACT

Due to object detection's close relationship with video analysis and image understanding, it has attracted much research attention in recent years. Traditional object detection methods are built on handcrafted features and shallow trainable architectures. Their performance easily stagnates by constructing complex ensembles which combine multiple low-level image features with high-level context from object detectors and scene classifiers. With the rapid development in deep learning, more powerful tools, which are able to learn semantic, high-level, deeper features, are introduced to address the problems existing in traditional architectures. These models behave differently in network architecture, training strategy and optimization function, etc

INDEX

LIST OF FIGURES

Introduction

To gain a complete image understanding, we should not only concentrate on classifying different images, but also try to precisely estimate the concepts and locations of objects contained in each image. This task is referred to as object detection [1], which usually consists of different subtasks such as face detection , pedestrian detection and skeleton detection . As one of the fundamental computer vision problems, object detection is able to provide valuable information for semantic understanding of images and videos, and is related to many applications, including image classification , human behavior analysis, face recognition and autonomous driving . Meanwhile, Inheriting from neural networks and related learning systems, the progress in these fields will develop neural network algorithms, and will also have great impacts on object detection techniques which can be considered as learning systems.. However, due to large variations in viewpoints, poses, occlusions and lighting conditions, it's difficult to perfectly accomplish object detection with an additional

Scope

- To build an android app that can be used to detect and determine the objects around you by using the camera module of the phone.
- Use machine learning models to reach a conclusive decision about the object type, and to compute the confidence score in that particular decision.
- Train and test the model/app to check whether it is able to determine different classes of objects correctly.

Objective

- Understand how the machine learning model works, and classifies the various object types.
- Integrate this machine learning model with our android app.
- Enable the object detection and recognition dynamically, i.e, it should give instantaneous results depending on where we're pointing the phone at.
- Minimum resource/performance requirements are desirable.

System Architecture

The SSD[2] approach is based on a feed-forward convolutional network that produces a fixed-size collection of bounding boxes and scores for the presence of object class instances in those boxes, followed by a non-maximum suppression step to produce the final detections. The early network layers are based on a standard architecture used for high quality image classification (truncated before any classification layers), which is called the base network. We then add auxiliary structure to the network to produce detections with the following key features.

- Multi-scale feature maps for detection
  These layers decrease in size progressively and allow predictions of detections at multiple scales.
- Convolutional predictors for detection
  Each added feature layer (or optionally an existing feature layer from the base network) can produce a fixed set of detection predictions using a set of convolutional filters
- Default boxes and aspect ratios
  Allowing different default box shapes in several feature maps lets us efficiently discretize the space of possible output box shapes.

The Model architecture can be seen as follows from Fig1.

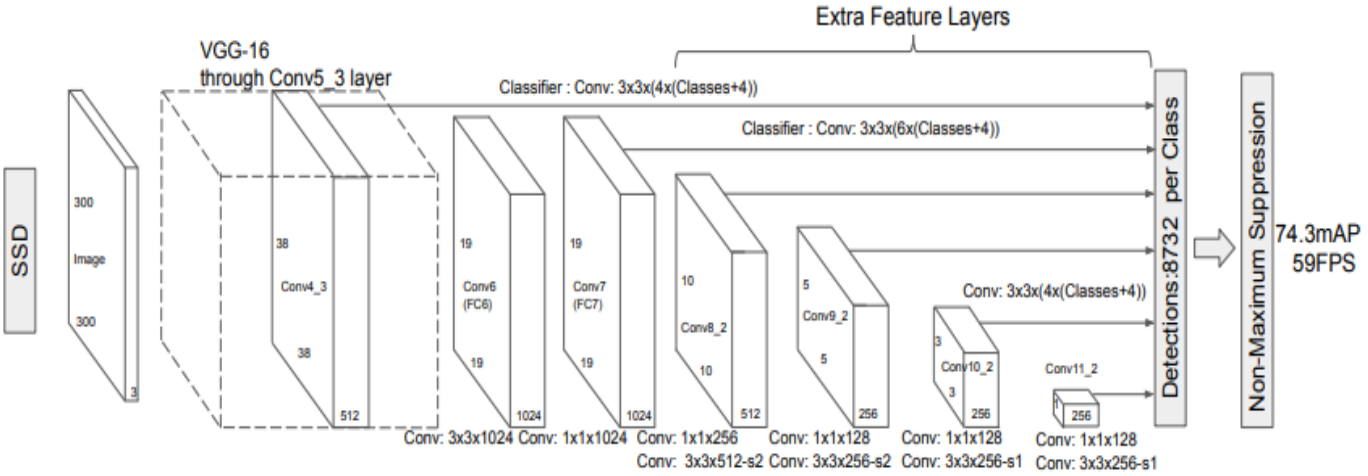

Fig1. Model of SSD

Snapshots

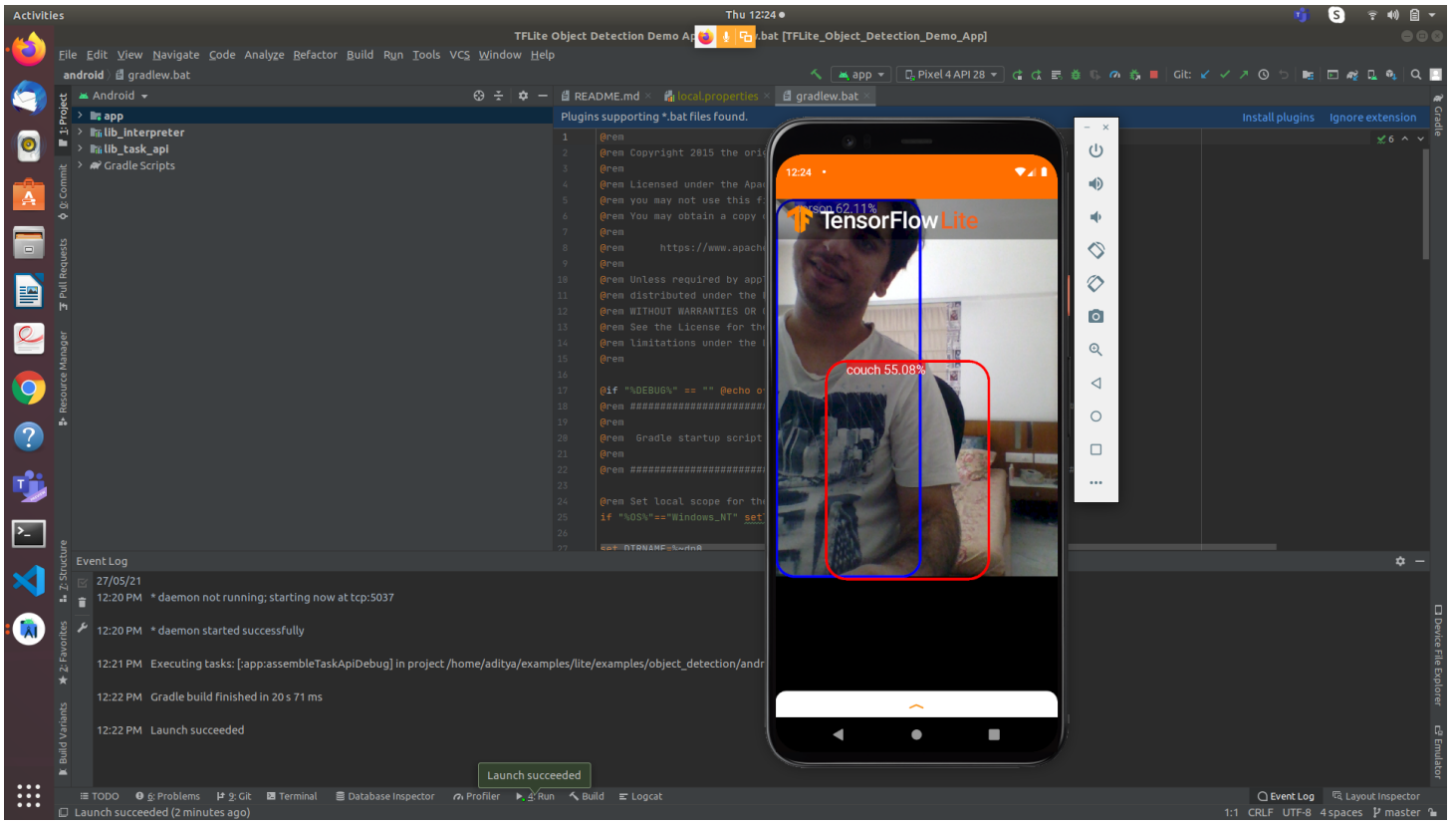

Fig2. Classification and the confidence score when detecting a person

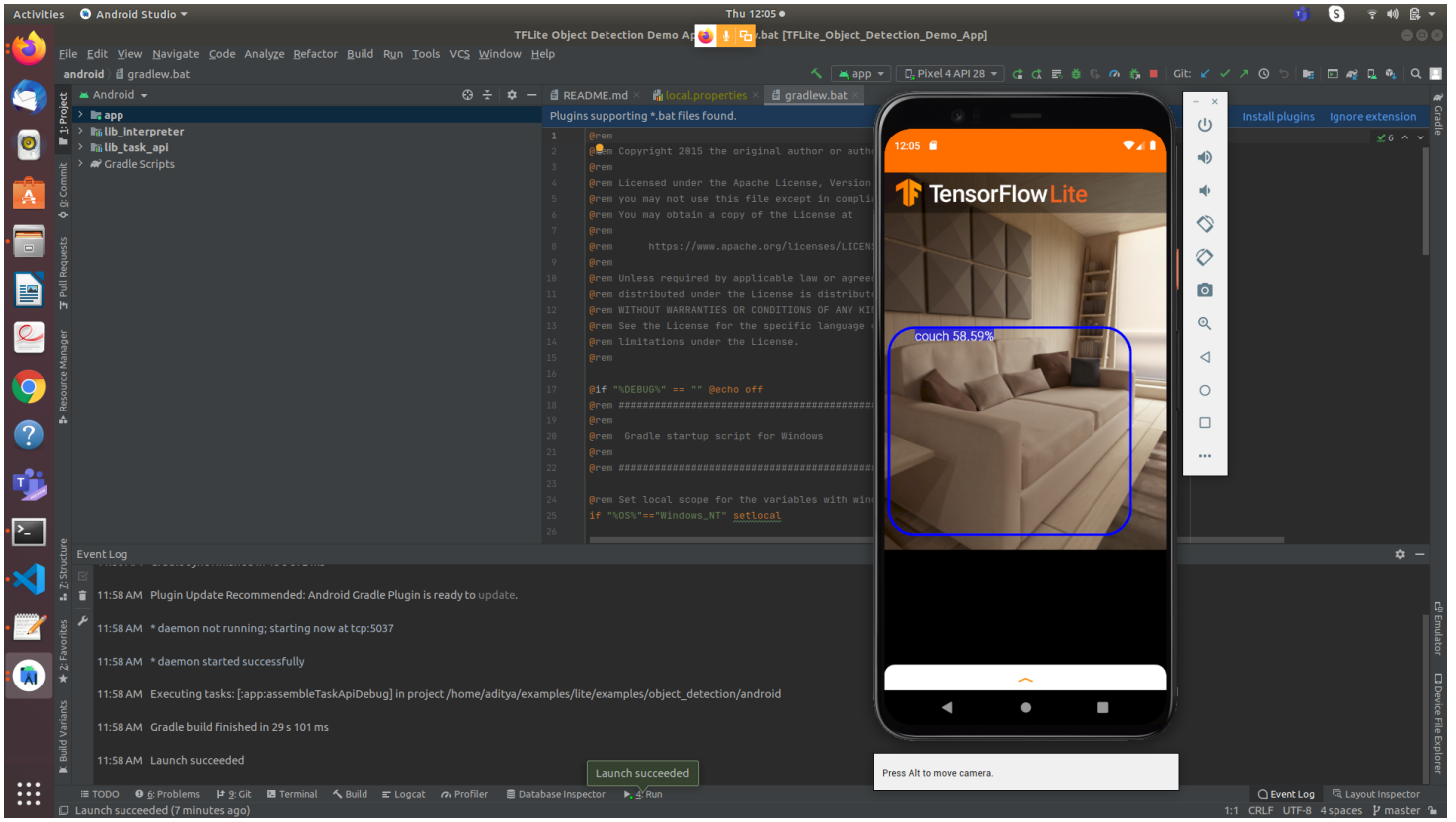

Fig3. Classification and the confidence score when detecting a couch

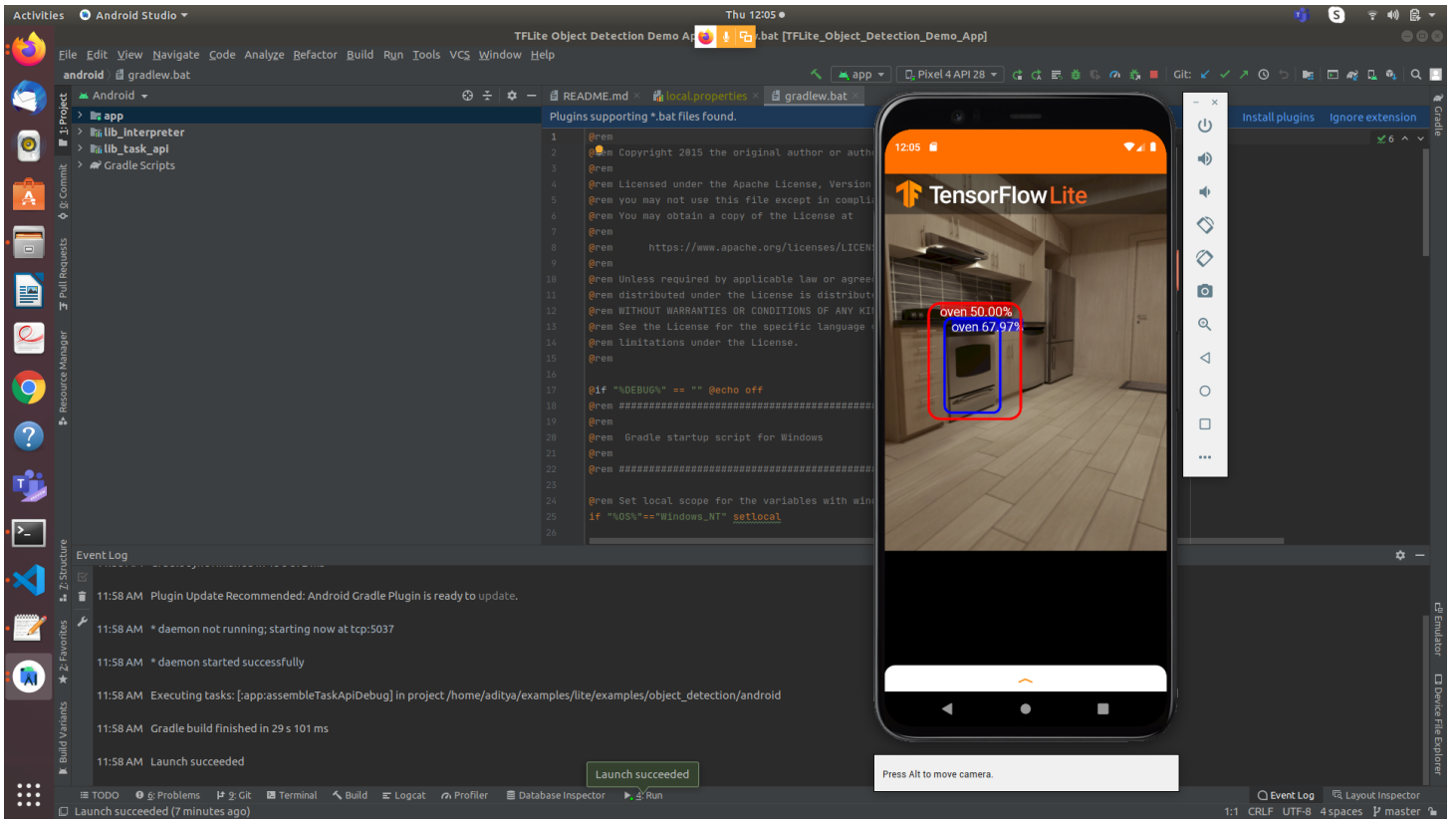

Fig4. Classification and the confidence score when detecting an oven

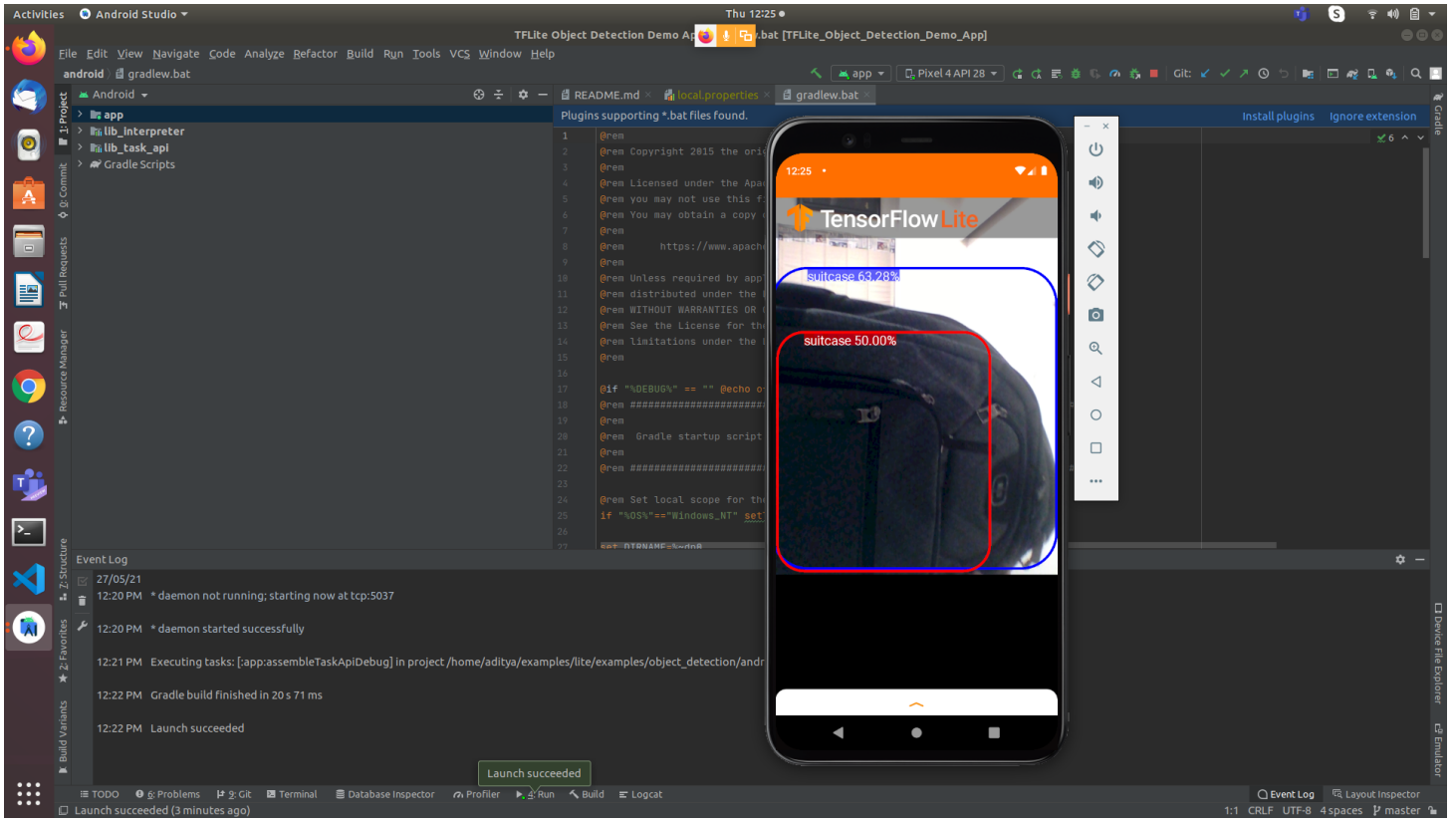

Fig5. Classification and the confidence score when detecting a bag/suitcase

Conclusion

Due to its powerful learning ability and advantages in dealing with occlusion, scale transformation and background switches, deep learning based object detection has been a research hotspot in recent years. This paper provides a detailed review on deep learning based object detection frameworks which handle different sub-problems, such as occlusion, clutter and low resolution, with different degrees of modifications on R-CNN. This project is also meaningful for the developments in neural networks and related learning systems, which provides valuable insights and guidelines for future progress

Future Scope

- To further work on adding additional object classes.
- To improve the training data size, for more accurate detection and recognition.
- To publish the app on the app store.

References

[1] P. F. Felzenszwalb, R. B. Girshick, D. Mcallester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, p. 1627, 2010.

[2] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed. Ssd: Single shot multibox detector. arXiv preprint arXiv:1512.02325, 2015.