



Bahria University
Discovering Knowledge

FINAL YEAR PROJECT REPORT

**OBJECT DETECTION AND LOCALIZATION
OF CUSTOM OBJECTS IN A VIDEO**

By

ABDUL RAHIM	(43690)
WASI HUSSAIN	(43810)
SAMI GHAFAR	(32724)

SUPERVISED BY
(DR RAHEEL SIDDIQI)

BAHRIA UNIVERSITY (KARACHI CAMPUS)

2019

DECLARATION

We hereby declare that this project report is based on our original work except for citations and quotations which have been duly acknowledged. We also declare that it has not been previously and concurrently submitted for any other degree or award at Bahria University or other institutions.

Signature : Abdul Rahim

Name : Abdul Rahim

Reg No. : 43690

Signature : Wasi Hussain

Name : Wasi Hussain

Reg No. : 43810

Signature : Sami

Name : Sami Ghaffar

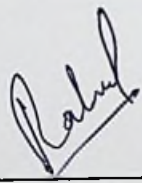
Reg No. : 32724

Date : 10-2-2020

APPROVAL FOR SUBMISSION

We certify that this project report entitled **“OBJECT DETECTION AND LOCALIZATION OF CUSTOM OBJECTS IN VIDEOS”** was prepared by **ABDUL RAHIM WASI HUSSAIN SAMI GHAFAR** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Computer Science (Honours) at Bahria University.

Approved by,

Signature :  _____

Supervisor : Dr RAHEEL SIDDIQI

Date : 10-2-2020

ACKNOWLEDGEMENTS

We would like to thank everyone who had contributed to the successful completion of this project. We would like to express my gratitude to my research supervisor, Dr RAHEEL SIDDIQUI for his invaluable advice, guidance and his enormous patience throughout the development of the research.

In addition, WE would also like to express my gratitude to our loving parent and friends who had helped and given me encouragement.

OBJECT DETECTION AND LOCALIZATION OF CUSTOM OBJECT IN VIDEOS

ABSTRACT

The main objective of this project is to localize and detect the any moving object in a real time video and obviously that can be one of the challenging task in the fields like the computer vision due to several factors such as the centroid or the origin of the object, the location of the object different types of shadings in a video to detect etc.

By studying the Many of the Artificial Intelligence algorithm one of the most efficient that we have found is the Convolutional Neural Network or the CNN that has the good accuracy rate and yet the most perfect one to fit for recognizing any custom trained object in a video, A Module will be developed which will further train and detect the custom objects and all the final implementation will be done using the Programming language called 'Python' IDE (PYcharm).

Just to train a single object there the minimum requirement for a single object is more than the 900 images and if increase the system will get more efficient in detecting that image and by considering that I have noticed that Convolutional Neural Network is the better choice than all other deep learning algorithms there would be one of the training phase and the testing phase as well. Further development is also included in the report.

Object Detection is the process of finding real-world object instances like many real life time objects still images or Videos. It allows for the recongnition , localization, and detection of multiple objects within an image which provide us with a much understanding of an image as a whole. It is commonly used in applications such as image retrieval, security, surveillance, and advanced driver assistance system (ADAS).

We work on different data set and train them to detect an objects , it give us 100% accuracy result through CNN algorithm.

TABLE OF CONTENTS

DECLARATION	ii
APPROVAL FOR SUBMISSION	iii
ACKNOWLEDGEMENTS	vi
ABSTRACT	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	xi
LIST OF FIGURES	xii
LIST OF SYMBOLS / ABBREVIATIONS	xiv
LIST OF APPENDICES	xv

CHAPTER

1	INTRODUCTION	1
1.1	Background	1
1.2	Problem Statement	3
1.3	Aims And Objectives	4
1.4	Scope Of Project	5
2	LITERATURE AND REVIEW	6
2.1	Methods Used in detecting the Object	6
	2.1.1 Appearance based detection	7
	2.1.2 Model based detection	8
	2.1.3 Template based detection	8
	2.1.4 Two stage detection	9
	2.1.5 Unified detection	9

2.1.6	Feature based detection	10
2.1.7	Motion based detection	10
2.1.8	Classifier based detection	11
2.2	Bounding Box	11
2.3	Deep learning	11
2.3.1	Convolutional Neural Network (Concept)	12
2.3.2	Local Receptive Fields	12
2.3.3	Pooling	13
2.3.4	Shared Weights	13
2.3.2	Convolutional Neural network Structure	13
2.4	Network Model	14
2.4.1	ResNet	14
2.4.2	R-CNN	14
2.4.3	Fast R-CNN	16
2.4.4	Yolo	17
3	DESIGN AND METHODOLOGY	18
3.1	System Development Plan Overview	18
3.2	System Architecture	18
3.3	System Components Required	19
3.4	Development Phases	20
3.4.2	Localization Phase	20
3.4.2	Detection Phase	21
3.5	Actor Use Case Diagram	21
3.5.1	User Input	22
3.5.2	System Functionality	22
3.5.3	Gantt Chart	23
4	IMPLEMENTATION	24
4.1	Dataset	24
4.2	Annotation of Dataset	24
4.3	Coding	25

4.3.1	Annotating Images	25
4.3.2	Generating xml	27
4.3.3	Detecting through Camera	28
4.3.4	Steps to be Followed	29
4.3.5	System Requirement	
5	RESULTS AND DISCUSSIONS	30
5.1	BEFORE Detection	31
5.2	After Detection	31
5.3	Common Mistakes	32
5.4	Future Work	
6	CONCLUSION AND RECOMMENDATION	33
6.1	Conclusion	33
6.2	Recommendations	34
	REFERENCES	34