

OBJECT DETECTION FOR GENERAL OBJECT TYPE
USING MOBILENET-SSD NEURAL NETWORK DATABASE



UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2021

B051710085

BACHELOR OF MANUFACTURING ENGINEERING (Hons.)

2021 UTeM



اونيورسيتي تيكنيكل مليسيا ملاك

UNIVERSITI TEKNIKAL MALAYSIA MELAKA



OBJECT DETECTION FOR GENERAL OBJECT TYPE USING MOBILENET-SSD NEURAL NETWORK DATABASE

Submitted in accordance with requirement of the Universiti Teknikal Malaysia Melaka
(UTeM) for the Bachelor Degree of Manufacturing Engineering (Hons.)



LAI JIUN HONG

B051710085

960711-01-5267

FACULTY OF MANUFACTURING ENGINEERING

2021

BORANG PENGESAHAN STATUS LAPORAN PROJEK SARJANA MUDA

Tajuk: **OBJECT DETECTION FOR GENERAL OBJECT TYPE USING MOBILENET-SSD NEURAL NETWORK DATABASE**

Sesi Pengajian: **2020/2021 Semester 2**


Saya **LAI JIUN HONG (960711-01-5267)**

mengaku membenarkan Laporan Projek Sarjana Muda (PSM) ini disimpan di Perpustakaan Universiti Teknikal Malaysia Melaka (UTeM) dengan syarat-syarat kegunaan seperti berikut:


1. Laporan PSM adalah hak milik Universiti Teknikal Malaysia Melaka dan penulis.
2. Perpustakaan Universiti Teknikal Malaysia Melaka dibenarkan membuat salinan untuk tujuan pengajian sahaja dengan izin penulis.
3. Perpustakaan dibenarkan membuat salinan laporan PSM ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. *Sila tandakan (√)

- SULIT** (Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysiasebagaimana yang termaktub dalam AKTA RAHSIA RASMI)
- TERHAD** (Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/ badan di mana penyelidikan dijalankan)
- TIDAK TERHAD**

Disahkan oleh:



Alamat Tetap:
J2350, JALAN JASIN 77300,
MERLIMAU, MELAKA.
Tarikh: 2nd July 2021



SHARIMAN BIN ABDULLAH
Cop Rasmi *Pensyarah*
Fakulti Kejuruteraan Pembuatan
Universiti Teknikal Malaysia Melaka
Tarikh: 5/7/2021

*Jika Laporan PSM ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh laporan PSM ini perlu dikelaskan sebagai SULIT atau TERHAD.

DECLARATION

I hereby, declared this report entitled "Object Detection For General Object Type Using MobileNet-SSD Neural Network Database" is the result of my own research except as cited in references.

Signature

: 

Author's Name

: LAI JIUN HONG

Date

: 2nd JULY 2021



APPROVAL

This report is submitted to the Faculty of Manufacturing Engineering of Universiti Teknikal Malaysia Melaka as a partial fulfilment of the requirement for Degree of Manufacturing Engineering (Hons). The member of the supervisory committee is as follow:



ABSTRAK

Pada era globalisasi ini, teknik pengesanan objek terhadap komputer telah berkembang dengan secara mendadak di bidang eletrik dan elektronik industri. Selain itu, tugas utama bagi teknik pengesanan objek adalah untuk mengklasifikasikan kategori objek dan mengenalpasti kedudukan objek dalam gambar atau video tertentu. Baru-baru ini, kemajuan dalam pembelajaran mendalam telah meningkatkan kecekapan teknik pengesanan objek dalam sektor kelajuan dan ketepatan. Tambahan pula, kemajuan dalam rangkaian neural yang mendalam bagi meningkatkan prestasi yang tinggi dalam seni bina supaya sistem tertanam dapat ditingkatkan dengan lebih pesat. Projek ini bertujuan untuk mengkaji kesesuaian sistem pengesanan objek beroperasi di Raspberry Pi 4 Model B. Seterusnya, Raspberry Pi juga dipanggil sebagai papan komputer tertanam yang terkenal. Demi membandingkan prestasi algoritma SSD dan algoritma MobileNet-SSD, terutamanya dari segi ketepatan. Justeru itu, eksperimen yang bersesuaian hendaklah dilakukan dengan menggunakan parameter yang berbeza, iaitu ukuran gambar input (280 x 280 piksel, 320 x 320 piksel dan 360 x 360 piksel) dan menggunakan penjarakan antara kamera dan objek yang berbeza (0.6 meter hingga 1 meter). Selanjutnya, terdapat lima jenis objek yang berkenaan akan menjalani eksperimen tersebut iaitu tetikus, telefon bimbit, alat kawalan jauh, komputer riba dan papan kekunci. Berdasarkan hasil keseluruhan yang diperolehi, hal ini dapat disimpulkan bahawa algoritma MobileNet-SSD mempunyai prestasi yang lebih baik, dengan mencapai purata kadar pengesanan hampir 76% dibandingkan dengan algoritma SSD, yang hanya memiliki 65% purata kadar pengesanan dalam sistem pengesanan. Jadi, algoritma MobileNet-SSD lebih sesuai untuk menjalankan sistem pengesanan objek pada masa nyata di Raspberry Pi dengan ketepatan yang tinggi. Oleh itu, Raspberry Pi sesuai digunakan sebagai peralatan dalam sistem pengesanan objek masa nyata. Bukan itu sahaja, ia juga dapat menggantikan dengan komputer tradisional demi mengurangkan kos penyelenggaraan dan secara tidak langsung meningkatkan kecekapan kerja.

ABSTRACT

Object detection has been experienced a dramatically technological improvement in the computer vision area. Besides that, the object detection technique is the combination of object classification and object localisation. The main task for object detection is to classify the object categories and identify the object's position in a certain image or video. Recently, advances in deep learning have significantly enhanced the efficiency of object detection techniques in speed and precision. This has allowed modern desktop computer systems to conduct extremely effective object detection at the real-time level. Other than that, the advancement in creating the high performance of deep neural network architectures for the embedded system has been increasing rapidly. Besides that, this project aims to investigate the appropriateness of operating object detection on the Raspberry Pi 4 Model B. Next, Raspberry Pi also is known as a popular embedded computer board. By comparing the performance of the SSD algorithm and MobileNet-SSD algorithm, especially on the accuracy, there is one experimental framework that should be carried out by using the different parameters, which is the size of input images (280 x 280 pixels, 320 x 320 pixels and 360 x 360 pixels) and the range distance between camera and object (0.6 meters until 1 meter). Furthermore, five types of objects will undergo in the experiment: a computer mouse, cell phone, remote, laptop, and keyboard. Based on the overall result obtained, it can be concluded that the MobileNet-SSD algorithm has better performance, which has an almost 76% average detection rate compared with the SSD algorithm, which only has a 65% of average detection rate in the detection system. Hence, the MobileNet-SSD algorithm is more suitable for running real-time object detection on Raspberry Pi with high accuracy and precision than the SSD algorithm. Apart from that, Raspberry Pi is suitable to use as hardware in a real-time object detection system. It can replace the traditional desktop system to reduce the maintenance fee and indirectly improve work efficiency.

DEDICATION

Only

my beloved father, Lai Fuh Chyan

my appreciated mother, Tan Hong Leng

my adored sister and brother, Lai Pei Shan and Lai Jiun Liang

for giving me moral support, money, cooperation, encouragement and also understandings

Thank You So Much & Love You All Forever



اونيورسيتي تيكنيكل مليسيا ملاك

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

ACKNOWLEDGEMENT

I should like to acknowledge and thank all those who gave me the time and the opportunity to complete this study. A special thanks to my supervisor, Dr Shariman Bin Abdullah, whose help, giving suggestions and encouragement, helped me have a clear mindset on how to efficiently carry out my work progress and finish my project on time. Without his kind direction and proper guidance, this project would have been a never success.

Furthermore, I want to give special thanks to my coursemates who offered me a lot of inspiration and mental cooperation to complete this project. With the support and encouragement given by my friends and my family, I overcame the tensions and problems I faced when I was completing this report.

In conclusion, I would like to thank everybody who was important to this Final Year Project report; I would like to express the apology that I could not mention each of you personally.

TABLE OF CONTENT

ABSTRAK	i
ABSTRACT	ii
DEDICATION	iii
ACKNOWLEDGEMENT	iv
TABLE OF CONTENT	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST ABBREVIATIONS, SYMBOLS AND NOMENCLATURES	xii
CHAPTER 1 INTRODUCTION	1
1.1 Background of Study	1
1.2 Problem Statement	2
1.3 Objectives	3
1.4 Scope	3
1.5 Importance Of The Study	4
1.6 Organization Of Report	4
1.7 Summary	6
CHAPTER 2 LITERATURE REVIEW	7
2.1 Object detection in Computer Vision	7
2.2 Deep Neural Network In Deep Learning	12
2.2.1 Convolutional neural network (CNN) in Deep Neural Network	14

2.3	Tensorflow In Object Detection	21
2.4	Dataset Related To Object Detection	22
2.4.1	PASCAL Visual Object Classes (VOC) dataset	25
2.4.2	ImageNet dataset	30
2.4.3	Open Images dataset	37
2.4.4	Microsoft Common Objects in Context (MS-COCO) dataset	42
2.5	Components Of Object Detection System	50
2.5.1	Raspberry Pi	50
2.5.2	Raspberry Pi Camera Module	52
2.6	Image Processing Software Of Computer Vision System	53
2.6.1	MATLAB	54
2.6.2	LabVIEW	56
2.6.3	OpenCV	57
2.7	Summary	59
CHAPTER 3 METHODOLOGY		62
3.1	Project Planning	62
3.1.1	Project flow chart	63
3.2	Project Design And Development	65
3.2.1	List of Components	66
3.2.2	Design of Object Detection System Hardware	68
3.3	Implementation Of Software And Programming Code	70
3.3.1	OpenCV software	71
3.3.2	PyCharm software	72
3.3.3	Python	73
3.4	Experimental Setup	74
3.4.1	Programming Setup	76

3.5	Bill Of Material (B.O.M)	82
CHAPTER 4 RESULT AND DISCUSSION		84
4.1	Result Of Object Detection Process	84
4.1.1	Interpretation on the Result of Computer Mouse and Cell Phone	85
4.1.2	Interpretation on the Result of Remote and Laptop	88
4.1.3	Interpretation on the Result of Keyboard	90
4.2	Analysis and Discussion	92
4.2.1	Colour Analysis	95
CHAPTER 5 CONCLUSION AND RECOMMENDATIONS		96
5.1	Conclusion	96
5.2	Future Recommendations	97
5.3	Sustainability	98
5.4	Complexity	98
5.5	Lifelong Learning	99
5.6	Entrepreneurship	100
REFERENCES		101
APPENDICES		106
A	Specifications Of Object Detection System Hardware	106
B	Coding Of Object Detection	109
C	Gantt Chart For Fyp 1	110
D	Gantt Chart For Fyp 2	111

LIST OF TABLES

2.1	Dimensional of the multi-layer feature maps (MobileNet-SSD)(Chiu et al., 2020)	19
2.2	Performance of detection networks (Pehlivan et al., 2020).	20
2.3	Comparison between MobileNet-SSDv2 network and MobileNet-SSD network (Chiu et al., 2020).	21
2.4	Four datasets for object detection (Liu et al., 2020).	23
2.5	Characteristic of four datasets (Liu et al., 2020).	24
2.6	Summary of the performance of object detection datasets in each challenge (Liu et al., 2020).	25
2.7	Tasks for each type of principal challenge (Everingham et al., 2010, 2015).	26
2.8	Tasks for each type of subsidiary challenge (Everingham et al., 2010, 2015).	27
2.9	Challenge tasks in ImageNet dataset (Russakovsky et al., 2015).	31
2.10	Comparison between four datasets for object detection	39
2.11	Comparison Table of MATLAB, LabVIEW and OpenCV(Anaz, 2020; Sharma, 2017)(Retrieved from https://opencv.org/platforms/).	59
3.1	List of components	66
3.2	Detection accuracy for each model by using different parameters.	75
3.3	Bill of Material	83
4.1	Result of Computer Mouse.	86
4.2	Result of Cell Phone.	87
4.3	Result of Remote.	89
4.4	Result of Laptop.	90
4.5	Result of Keyboard.	91

LIST OF FIGURES

2.1	Example of image classification (Wu et al., 2020).	8
2.2	Example of object detection (Wu et al., 2020).	9
2.3	Process of object detection (Vahab et al., 2008).	10
2.4	Example of semantic segmentation (Wu et al., 2020).	11
2.5	Example of instance segmentation (Wu et al., 2020).	11
2.6	Deep neural network.	13
2.7	Architecture of the convolutional neural network (Galvez et al., 2019).	14
2.8	Framework of MobileNet architecture (Chiu et al., 2020).	16
2.9	Framework of Single Shot Detector (SSD) (W. Liu et al., 2015).	17
2.10	MobileNet-SSD network architecture (Chiu et al., 2020).	18
2.11	Pyramid structure of Mobile-SSD detector (Ali et al., 2019)	18
2.12	Learning system task (Pehlivan et al., 2020).	19
2.13	Dataflow graph of TensorFlow	22
2.14	PASCAL VOC dataset (20 Classes) (Liu et al., 2020).	28
2.15	Sample of annotation images in PASCAL VOC dataset (Everingham et al., 2015).	29
2.16	ILSVRC dataset (200 Classes) (Liu et al., 2020).	30
2.17	Easiest and hardest classes in the image classification task (Russakovsky et al., 2015).	32
2.18	Classification accuracy versus the average scale of an object.	32
2.19	Easiest and hardest classes in the single-object localization task (Russakovsky et al., 2015).	34
2.20	Localization accuracy versus the average scale of an object.	34
2.21	Easiest and hardest classes in the object detection task (Russakovsky et al., 2015).	36
2.22	Average precision versus the average scale of an object.	36
2.23	Sample of annotations images for each task (Kuznetsova et al., 2017).	37
2.24	Open Images Detection Challenge dataset (500 classes) (Liu et al., 2020).	37

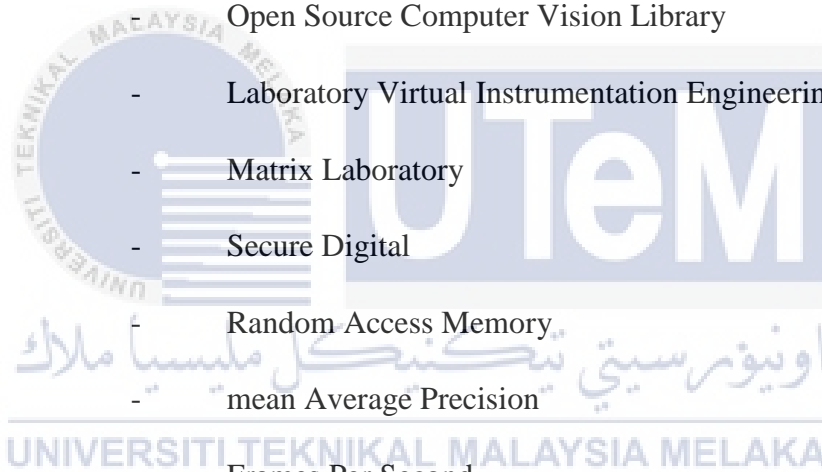
2.25 Graph of the percentage of images versus the number of object per images (Kuznetsova et al., 2017).	39
2.26 Graph of the number of images versus the number of object per images (Kuznetsova et al., 2017).	40
2.27 Graph of number of boxes versus class sorted position (Kuznetsova et al., 2017)	41
2.28 Graph of the percentage of images versus the number of distinct classes per image (Kuznetsova et al., 2017).	41
2.29 Graph for the number of images and number of distinct classes per image (Kuznetsova et al., 2017).	42
2.30 MS-COCO dataset (80 Classes) (Liu et al., 2020).	43
2.31 91 categories icons in MS-COCO dataset (Lin et al., 2014).	43
2.32 Sample of annotation images in MS-COCO dataset (Lin et al., 2014).	44
2.33 Image classification	45
2.34 Object localization	46
2.35 Semantic pixel-level segmentation.	47
2.36 The number of instances per category for MS-COCO dataset and PASCAL VOC dataset (Lin et al., 2014).	47
2.37 Review of the several object recognition datasets indicating the number of object categories and the number of instances per category (Lin et al., 2014).	48
2.38 Graph of the number of categories per images for each dataset (Lin et al., 2014).	49
2.39 Graph of the number of instances per images for each dataset (Lin et al., 2014).	49
2.40 Range of instance sizes for each dataset (Lin et al., 2014).	50
2.41 Raspberry Pi 4 Model B (Pi, 2020).	50
2.42 Specification of Raspberry Pi 4 Model B (Pi, 2020).	51
2.43 Raspberry Pi Camera Module v2.	52
2.44 Camera Module v1 vs Camera Module v2 vs HQ camera.	53
2.45 Icon of MATLAB	54
2.46 Design flow in image processing using MATLAB (Sharma, 2017).	55
2.47 Icon of LabVIEW	56
2.48 Icon of OpenCV	57
2.49 Implementation of image processing on a mobile platform (Deepthi & Sankaraiah, 2011).	58

3.1	Flow Chart For the Project.	64
3.2	Project Design and Development Flow Chart	65
3.3	Object Detection System.	68
3.4	Four major elements in OpenCV (Neelima & Saravanan, 2014).	71
3.5	Importing the dataset, configuration file, and weight file.	77
3.6	Coding for import MS-COCO dataset.	77
3.7	Coding for the import configuration file and weight file.	77
3.8	Coding for parameter setting of the detected objects.	78
3.9	Apply non-maximum suppression in the function.	78
3.10	Coding for webcam setup and output setup.	79
3.11	Computer hardware setup.	80
3.12	Raspberry Pi hardware setup.	81
4.1	Result of object detection for a computer mouse.	87
4.2	Result of object detection for cell phone.	87
4.3	Result of object detection for a remote.	89
4.4	Result of object detection for laptop.	90
4.5	Result of object detection for the keyboard	92
4.6	Graph of average detection rate versus input size.	94
4.7	Graph of average detection rate versus distance.	94
4.8	Examples of the multicolour object.	95

LIST ABBREVIATIONS, SYMBOLS AND NOMENCLATURES

KKM	-	Kementerian Kesihatan Malaysia
CCTV	-	Closed-Circuit Television
CV	-	Computer Vision
AI	-	Artificial Intelligence
ML	-	Machine Learning
IDE	-	Integrated Development Environment
DNN	-	Deep Neural Network
ANN	-	Artificial Neural Network
CNN	-	Convolutional Neural Network
FPN	-	Feature Pyramid Network
TPUs	-	Tensor Processing Units
CPUs	-	Central Processing Units
GPUs	-	Graphics Processing Units
SSD	-	Single Shot Detector
R-CNN	-	Region-based Convolutional Neural Networks
R-FCN	-	Region-based Fully Convolutional Network
PASCAL VOC	-	PASCAL Visual Object Classes
CIFAR	-	Canadian Institute For Advanced Research
MS-COCO	-	Microsoft Common Objects in Context
SUN	-	Scene UNderstanding

Caltech	-	California Institute of Technology
ILSVRC	-	ImageNet Large Scale Visual Recognition
MNIST	-	Modified National Institute of Standards and Technology
ReLU	-	Rectified Linear Units
BLOB	-	Binary Large Object
NMS	-	Non-maximum Suppression
USB	-	Universal Serial Bus
HDMI	-	High Definition Multimedia Interface
LED	-	Light Emitting Diode
OpenCV	-	Open Source Computer Vision Library
LabVIEW	-	Laboratory Virtual Instrumentation Engineering Workbench
MATLAB	-	Matrix Laboratory
SD	-	Secure Digital
RAM	-	Random Access Memory
mAP	-	mean Average Precision
FPS	-	Frames Per Second
MB	-	Megabyte
GB	-	Gigabyte
%	-	Percentage
m	-	meter
ms	-	millisecond
s	-	second
V	-	Volt
W	-	Watt



VAC	-	Volt of Alternating Current
P	-	Pixel
A	-	Ampere
Lm	-	Lumen
SoC	-	System On a Chip
GHz	-	Gigahertz
LPDDR	-	Low-Power Double Data Rate
GPIO	-	General-Purpose Input/Output
OS	-	Operating System
HD	-	High Definition



CHAPTER 1

INTRODUCTION

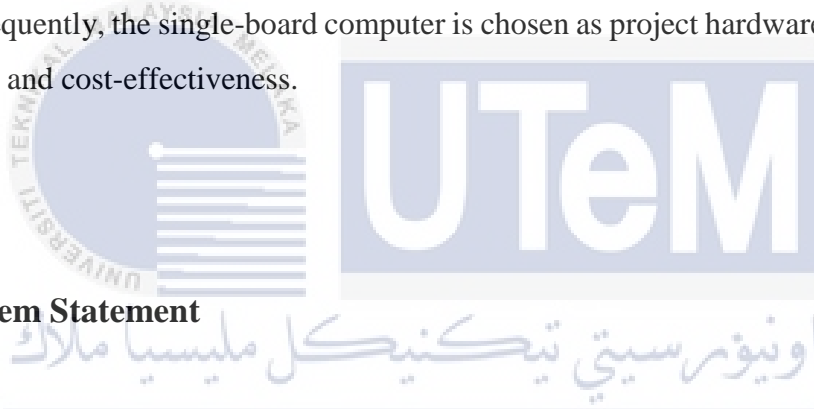
1.1 Background of Study

In an era of globalization, the "pandemic" word is always heard everywhere. Almost the whole world on Earth are concerned about this global issue. This pandemic disease will affect the body health of a person and cause death if the person is in a critical situation. From the statistic posted by Kementerian Kesihatan Malaysia (KKM), since 10 December 2020, there are almost more than 1000 cases in one day. According to KKM, the total active cases in Malaysia until 24 December 2020 in 100,318 cases and 468 disease clusters. Besides that, the pandemic has not only caused the global economic downturn, but even some companies will face the fate of bankruptcy.

Furthermore, the global unemployment rate is increasing day by day due to the pandemic issue. Indirectly the crime will also increase due to the current situation. According to the crime index presented by the Department of Statistics Malaysia, the crime rate in 2019 is rising to 256,6 compared to 273,8 in 2018, per 100,000 population. There are a few methods that can be implemented to prevent or even reduce the risk of crime—for example, using a surveillance system like CCTV or simple camera drones for aerial inspection. As additional information, object detection is a computer vision technology that can apply to the surveillance system to recognize and locate certain objects like animals, signboard, vehicles, furniture, person, and so on in an image or video.

The evolution of deep learning has been introduced convolutional neural networks as a subset of deep neural networks which able to transfer large image datasets to computers. In addition, it will teach the computer to identify or track the object that has been seen in the videos or images with high accuracy, indirectly improve the performance and efficiency. By the way, a few challenges that the manufacturing industry needs to overcome, such as visual tasks, need a huge amount of computing power to run, the limitation of hardware resource on the device, and running object detection in real-time.

In addition, Raspberry Pi is a small single-board computer that the object detector can implement to maintain a real-time frame rate performance with high precision. On the other hand, the manufacturing industry's issue is that the traditional computer system provides high computing power than the single-board computer. Still, it is too expensive, and the size is big. Consequently, the single-board computer is chosen as project hardware due to its size, performance and cost-effectiveness.



1.2 Problem Statement

The recent years have seen many exciting developments in the field of the automation industry. Real-time object detection is such a more challenging task in computer vision applications. Previously, Single Shot Detector (SSD) algorithm is used to perform real-time object detection. However, the setup for the detection system is a huge expense for the industry due to its need for high computing power to support and run the program. Other than that, the performance of the detection system is also low accuracy and lower speed. Therefore, the MobileNet-SSD algorithm has been used in this project to increase the accuracy and speed of the object detection system when performing real-time tasks. In addition, the single-board computer will replace the traditional computer system to develop a low costs detection system with high accuracy and high speed, indirectly improving the performance of the object detection system.

1.3 Objectives

The objectives of the project are important in order to ensure that the research is carried out in order to resolve the issue which has been investigated. Objectives are shown as below:

1. To implement MobileNet-SSD algorithm into a low-cost vision detection system using OpenCV in order to create high accuracy and precision cost-effectiveness detection system.
2. To improve detection system with more cost-effectiveness detection system using a single board computer.

1.4 Scope

The scope of this project will be fixed according to the requirements from the objectives. By referring to the first objective, this project focused on creating high accuracy and precision cost-effectiveness detection system in the field of electric and electronic industry. However, there are too many types of electronic devices are currently available in the surroundings. Following are the scopes and limitations of the project:

- Five types of electronic devices: computer mouse, cell phone, remote, laptop and keyboard
- Size of an input image: 280 x 280 pixels, 320 x 320 pixels and 360 x 360 pixels.
- Range of distance between the camera and the object: 0.6m, 0.7m, 0.8m, 0.9m, and 1m.
- Network architecture: MobileNet-SSD algorithm

1.5 Importance Of The Study

Based on the statistic of the crime index in Malaysia, the crime rate is increasing day by day to a critical level. This issue should be carried out as fast as possible in order to reduce the crime rate. In this way, this study is important for people who have the capability to create its own home security system by using the Raspberry Pi with an object detector. It not only will reduce the cost of installing the CCTV surveillance system, but it also can increase the safeness of the members in order to reduce the risk of crime.

Besides that, the detection system can also be applied in small industries to trace the working environment of employees. For example, a production line delay due to the operator having an accident in line. In this way, the data collected in the detection system supports the operator and uses an effective way to solve the issue.

In addition, this study also important for the visually impaired. By using the object detection system to help a person who has a low vision problem to recognize the appearance and shape of the object accurately. Other than that, Google Lookout is an application on Google Pixel smartphone using the camera to detect the object and give the information to the user. This kind of application is very useful for the visually impaired.

1.6 Organization Of Report

Chapter 1 introduces the background of this project with respect to the title of this project. The problem statements are established through a range of research methods and literature reviews in books, the Internet, newspapers and the current global question. After the problem had been identified, there are a few objectives that should be carried out to solve the problem. This project's scope is limited so that the data can be obtained and analyzed accurately when conducting the experiment research for this project. Chapter 1 also