TODDLER MONITORING SYSTEM IN VEHICLE BY USING ARTIFICIAL INTELLIGENCE

KOK JIA QUAN

BACHELOR OF MECHATRONIC ENGINEERING WITH HONOURS
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2021

TODDLER MONITORING SYSTEM IN VEHICLE BY USING ARTIFICIAL INTELLIGENCE

KOK JIA QUAN

A report submitted
in partial fulfillment of the requirements for the degree of
Bachelor of Mechatronic Engineering with Honours

Faculty of Electrical Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2021

DECLARATION

I declare that this thesis entitled"TODDLER MONITORING SYSTEM IN VEHICLE BY USING ARTIFICIAL INTELLIGENCE is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : Quan

Name : KOK JIA QUAN

Date : 5 JULY 2021

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

APPROVAL

I hereby declare that I have checked this report entitled "TODDLER MONITORING SYSTEM IN VEHICLE BY USING ARTIFICIAL INTELLIGENCE and in my opinion, this thesis it complies the partial fulfillment for awarding the award of the degree of Bachelor of Mechatronics Engineering with Honours

Signature : .....

Supervisor Name : IR. DR. ZAMANI BIN MD. SANI
PENSYARAH KANAN
Fakulti Kejuruteraan Elektrik
Universiti Teknikal Malaysia Melaka

Date : 5 / 7 /2021

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

DEDICATIONS

To my beloved mother and father

ACKNOWLEDGEMENTS

During the preparation of this report, I was in contact with many researchers, academicians, seniors and friends. They have contributed to my understanding and thought. I want to express my appreciation to those who have helped me to finish the report. Firstly, I would like to express my most tremendous gratitude to my main project supervisor, IR. DR. ZAMANI BIN MD. SANI, for all the guidance, patience, and advice while completing the project. His guidance is beneficial and helps me finish this report within the allocated time.

Moreover, I would also like to express my appreciation to my panels, PROFESOR MADYA DR. MARIAM BINTI MD GHAZALY and DR. MOHD FAID BIN YAHYA for advice and critics. After they provided me sincere advice, I managed to improve my work and identify my mistakes, which contributed to the report completion. Besides, I am thankful to University Teknikal Malaysia Melaka (UTeM) for funding this project.

Next, I want to thank my family members for the moral support they have provided to me while completing the project. Furthermore, the help and support from my fellow postgraduate seniors must be recognized. Lastly, I would like to appreciate the others who have contributed their time and assistance to me on various occasions. Still, it is impossible to list them all in this limited space.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

ABSTRACT

Nowadays, the vehicle becomes the main transportation and the safety of toddler need to be concerned. There is much news regarding the death of a toddler in the vehicle, such as heatstroke and accident caused by parents' carelessness. This project is to develop a monitoring system which able to detect a toddler's presence in a car and classify the seatbelt condition of the toddler. The toddler monitoring system is responsible for monitoring toddler conditions to ensure the toddler is safe. The system will protect the toddler by detecting the seatbelt condition of the toddler and alert the driver when the toddler is left inside the car after the vehicle is off. The monitoring system is vision-based and using machine learning algorithm to detect toddlers and seatbelts. The environment inside the car is emphasized in the project. Nvidia Jetson Nano is chosen as the microcontroller due to its powerful performance. The Faster R-CNN, SSD-Mobilenet and SSD-Inception algorithm with Tensorflow is used and compared for the detection and classification. The Faster R-CNN is not applicable on Jetson Nano due to the high computational requirement and Jetson Nano unable to use it to process the image and hence the performance only compared from SSD-Mobilenet and SSD-Inception. From the result, it is observed that SSD-Inception gives better performance with 77.70% of accuracy, 97.92% of precision and 77.47% of recall when detecting the toddler while SSD-Mobilenet perform better when detecting the class of seatbelt with 82.98% of accuracy, 99.07% of precision and 77.37% of recall. From the observation, the performance of both neural networks are just slightly difference and hence the decisive factor is the frame per second (FPS) of neural network when running in real-time. The SSD-Mobilenet detecting in real-time with 8.5 FPS which faster than SSD-Inception with 5.7 FPS

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

ABSTRAK

Pada masa kini, kereta menjadi pengangkutan utama dan keselamatan kanak-kanak haruslah diberi perhatian. Terdapat banyak perkara baru mengenai kematian kanak-kanak di dalam kereta seperti panas dan kemalangan yang disebabkan oleh kecuaian ibu bapa. Projek ini adalah untuk mengembangkan sistem pemantauan yang dapat memeriksa kehadiran kanak di dalam kereta dan mengklasifikasikan keadaan tali pinggang keledar kanak. Sistem pemantauan kanak-kanak bertanggungjawab untuk memantau keadaan balita dan memastikan kanak berada dalam keadaan keselamatan. Sistem ini akan melindungi kanak-kanak dengan memeriksa keadaan tali pinggang keledar kanak-kanak dan memberi amaran kepada pemandu apabila kanak ditinggalkan di dalam kereta setelah kenderaan dimatikan enjin. Sistem pemantauan berasaskan penglihatan dan menggunakan algoritma pembelajaran mesin untuk memeriksa kanak dan tali pinggang keledar. Persekitaran dalam kereta akan ditekan dalam projek.. Nvidia Jetson Nano dipilih sebagai pengawal mikro kerana prestasinya yang hebat. Algoritma Faster R-CNN, SSD-Mobilenet dan SSD-Inception dengan Tensorflow digunakan dan dibandingkan untuk memeriksa dan klasifikasi.Faster R-CNN tidal dapat dilangsung di Jetson Nano kerana keperluan komputasi yang tinggi dan Jetson Nano tidak dapat menggunakannya untuk memproses gambar dan oleh itu prestasi hanya dibandingkan dengan SSD-Mobilenet dan SSD-Inception. Dari hasilnya, diperhatikan bahawa SSD-Inception memberikan prestasi yang lebih baik dengan ketepatan 77.70%, ketepatan 97.92% dan penarikan 77.47% ketika semasa mengesan kelas kanak-kanak kecil dan SSD-Mobilenet berprestasi lebih baik ketika mengesan kelas tali pinggang keledar dengan ketepatan 82.98% , Ketepatan 99.07% dan penarikan semula 77.37%. Dari pemerhatian, prestasi kedua-dua rangkaian saraf hanya sedikit perbezaan dan oleh itu faktor yang menentukan adalah bingkai sesaat (FPS) rangkaian saraf ketika berjalan dalam masa nyata. SSD-Mobilenet mengesan dalam masa nyata dengan 8.5 FPS yang lebih pantas daripada SSD-Inception dengan 5.7 FPS

TABLE OF CONTENTS

PAGE

DECLARATION

APPROVAL

DEDICATIONS

ACKNOWLEDGEMENTS **2**

ABSTRACT **3**

ABSTRAK **4**

TABLE OF CONTENTS **5**

LIST OF TABLES **7**

LIST OF FIGURES **8**

LIST OF SYMBOLS AND ABBREVIATIONS **10**

LIST OF APPENDICES **11**

CHAPTER 1 INTRODUCTION **12**
1.1 Background 12
1.2 Motivation 13
1.3 Problem Statement 14
1.4 Objectives 15
1.5 Scope 15
1.6 Organisation of Thesis 16

CHAPTER 2 LITERATURE REVIEW **17**
2.1 Introduction 17
2.2 Computer Vision 17
2.3 Image Pre-processing 18
2.4 Machine Learning 18
2.5 Method for Human Detection 21
2.5.1 Background Subtraction Method 21
2.5.2 Feature Description Method 23
2.5.3 Region-Based Convolutional Neural Networks(R-CNN) 26
2.5.4 Single-Shot Detector (SSD) 27
2.5.4.1 SSD-Mobilenet 27
2.5.4.2 SSD-Inception 28
2.6 Summary 30
2.7 Gap from Literature Review 33

CHAPTER 3 METHODOLOGY **34**
3.1 Introduction 34
3.2 Project Overview 34

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

3.3 Toddler Monitoring System in Vehicle Flowchart 35
3.4 Gantt Chart 37
3.5 Hardware Components 38
3.5.1 NVIDIA Jetson Nano Developer Kit 38
3.6 System Design Configuration 40
3.6.1 Setup of Camera Location 40
3.6.2 Algorithm Development 41
3.6.3 Design of Alert System 46
3.7 Experiment Description 47
3.7.1 Experiment 1: Confusion Matrix of Models 47
3.7.2 Experiment 2: Performance Testing With Different Neural Network 48
3.8 Summary 48

CHAPTER 4 RESULT AND DISCUSSION 49
4.1 Introduction 49
4.2 Result Overview 49
4.2.1 Dataset Preparation 49
4.2.2 System Training Monitoring 50
4.2.3 Image Result With Bounding Box 52
4.2.4 Alert System 54
4.3 Experiment Result 56
4.3.1 Experiment 1: Confusion Matrix of Models 56
4.3.2 Experiment 2: Performance Testing With Different Neural Network 57

CHAPTER 5 CONCLUSION 60
5.1 Conclusion 60
5.2 Future Work 61

Reference 62

APPENDICES 71

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

LIST OF TABLES

Table 2-1: Different of Supervised And Unsupervised Learning[25] 20

Table 2-2: Comparison of Background Subtraction Method 23

Table 2-3: Comparison of Feature Description Method 25

Table 2-4:Mobilenet Architecture[55] 28

Table 2-5: Pros and Cons of Human Detection Method 30

Table 2-6: Summary of Literature Journal 32

Table 3-1: Gantt Chart for FYP1 37

Table 3-2: Gantt Chart for FYP2 37

Table 3-3: Comparison of Performance of Different Platform[64] 39

Table 3-4: Data collected in experiment 1 47

Table 3-5: Data Required in Experiment 2 48

Table 4-1: Sample of Image Result of Different Neural Network 52

Table 4-2: Result of Backseat Condition 54

Table 4-3: Confusion Matrix of SSD-Mobilenet 56

Table 4-4: Confusion Matrix of SSD-Inception 56

Table 4-5: Performance of Neural Network 57

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

LIST OF FIGURES

Figure 1-1: Number of Road Accidents in Malaysia (2012-2018)[5] 13

Figure 1-2: Car Accident Death Statistic (2006-2015)[6] 13

Figure 2-1: Flow of Machine Learning 19

Figure 2-2: Convolutional Neural Network Layers[30] 21

Figure 2-3: Implementation of Gaussian Mixture Model for Skin Modelling[32] 22

Figure 2-4: Non-Parametric Model for Background Subtraction[34] 23

Figure 2-5: HOG Result 24

Figure 2-6: mAP Comparison of RCNN and Another Model[48] 26

Figure 2-7: Speed Comparison of Different Version of R-CNN[51] 27

Figure 2-8: Inception Building Block[57] 29

Figure 2-9: Comparison of Accuracy of Different Models[59] 31

Figure 2-10: Comparison of frames per second (FPS) Implementing The Different Models Using Input Images With Different Resolutions.[59] 31

Figure 3-1: Project Overview Flowchart 35

Figure 3-2: System Flowchart of Toddler Monitoring System in Vehicle. 36

Figure 3-3: Nvidia Jetson Nano Overview 40

Figure 3-4: Location of Camera Inside The Car 41

Figure 3-5: Flowchart of Monitoring System Design 41

Figure 3-6: Pre-processing of Image (Splitting and Cropping) 42

Figure 3-7: Labelling Process and Generated Annotation File 43

Figure 3-8: Annotation information 43

Figure 3-9: Image labelled 44

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Figure 3-10: Valid Data for Machine Learning 44

Figure 3-11: Tensorboard for Tracking of Learning Process 45

Figure 3-12: Connection of Alert System 46

Figure 4-1: Dataset Splitting 49

Figure 4-2: Faster R-CNN loss 50

Figure 4-3: SSD-Mobilenet loss 50

Figure 4-4: SSD-Inception loss 51

Figure 4-5: The Performance of Neural Network Toward The Toddler Class 57

Figure 4-6: The Performance of Neural Network Toward The Seatbelt Class 58

Figure 4-7: Average Precision of SSD-Mobilenet 58

Figure 4-8: Average Precision of SSD-Inception 59

UTeM
اونيۏرسيتي تيکنيکل مليسيا ملاک
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

LIST OF SYMBOLS AND ABBREVIATIONS

CNN	-	Convolutional Neural Network
R-CNN	-	Regional Convolutional Neural Network
SSD	-	Single Shot Detector
ML	-	Machine Learning
AI	-	Artificial Intelligence

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

LIST OF APPENDICES

Appendix A: Specification of Jetson Nano 71

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

CHAPTER 1

INTRODUCTION

1.1 Background

In this era, vehicle become the main transportation for humans. In this case, people emphasize the safety feature in car specification like the car's structure, car seat, airbag, etc. However, the most significant factor to safety is people ignore the behaviour of humans inside the vehicle. Human activity recognition (HAR) is a topic that aims to determine the activities of a person based on sensor or video observation[1].

Recognition of human behaviour in vehicles is becoming increasingly important. Nowadays, the development of autonomous is drastic. The more control the car has, the more we need to know about the person behind the wheel, especially if they are expected to take control from automation[2]. Due to the development of the autonomous vehicle recently, a lot of research on the surrounding monitoring of vehicles was done. However, the tracking of a passenger is still less to be concerned about. The accident is sometimes unpredictable and unpreventable and not only because of the carelessness or drowsiness of the driver. What people can do in an accident is knowing how to survive in the accident. According to the child accident prevention trust organization, twelve children under ten are killed or injured as passengers in cars every day [3]. Due to a large number of statistics, the safety of the toddler in the car has received much emphasis from parents. Inside car safety recognizes system for toddler is a system that recognizes the activity and the pose did by toddlers and to predict whether the toddler made any dangerous motion in the car. There are a lot of the hazard action can be done by toddler due to their childishness and ignorance. The focus on this research is to develop a relatively new system that can recognize the toddlers and their activity in vehicle.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

1.2 Motivation

The statistics data from the World Health Organization (WHO) show that approximately 1.35 million people die from road traffic crashes each year and most of the victims are aged between 5 to 29 years old[4]. Road traffic injuries are the leading cause of death among children and young adults. Figure 1-1 shows the number of car accident death statistics and Figure 1-2 shows the statistic of car accidents and death.

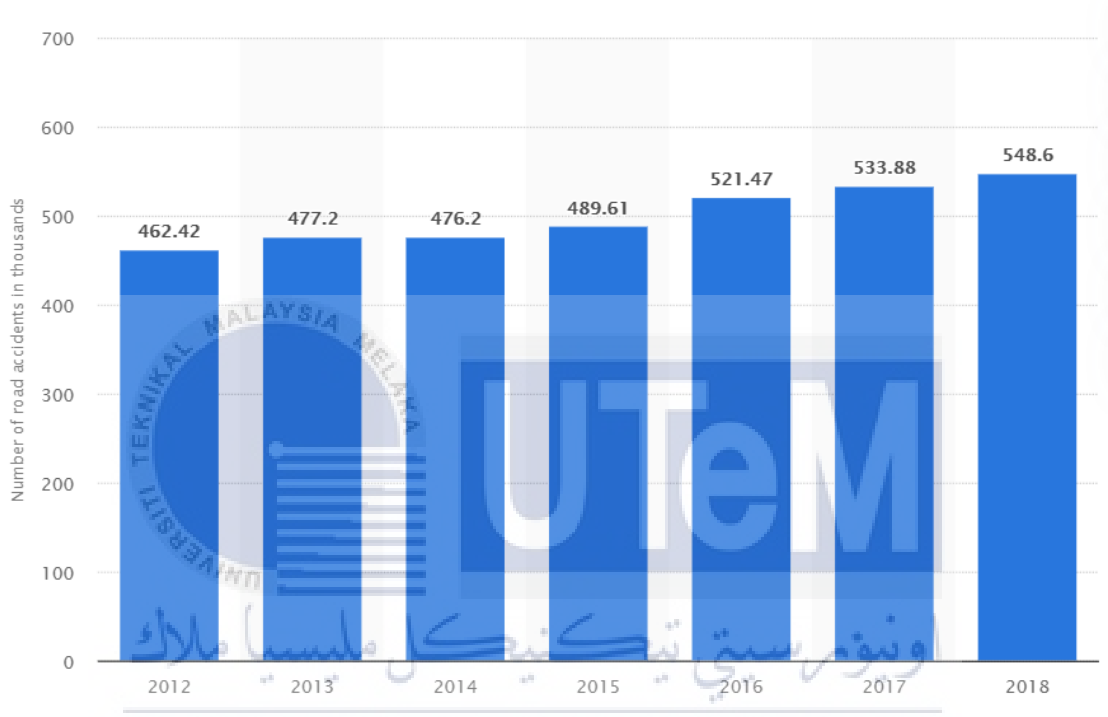

Figure 1-1: Number of Road Accidents in Malaysia (2012-2018)[5]

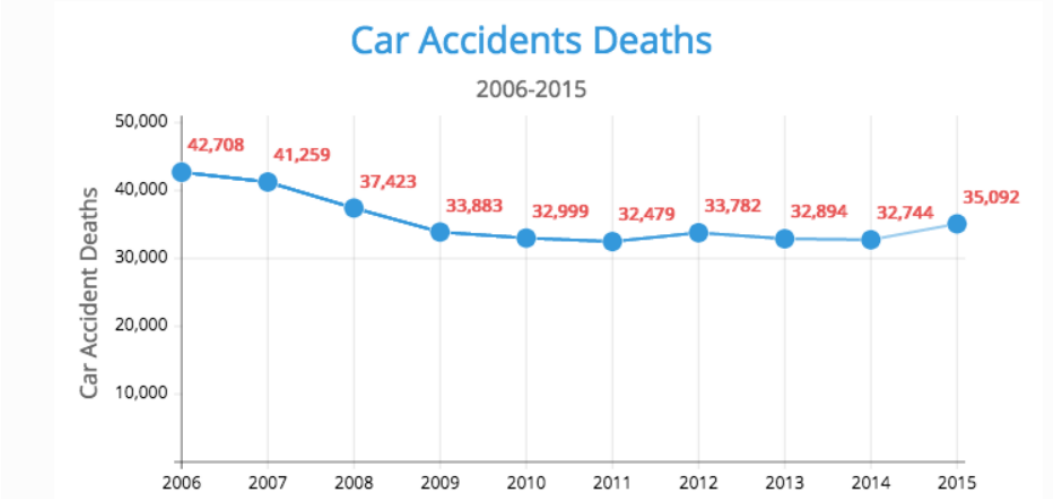

Figure 1-2: Car Accident Death Statistic (2006-2015)[6]

From the past research found that the seatbelt is one of the essential elements to keep passengers safe. Some of the toddlers are unable to tie the seatbelt by themselves, and parents forget to fasten their seatbelt properly[7]. The function of the seatbelt is to reduce the risk of death during an accident by preventing passengers and drivers from being ejected. Seatbelts can reduce the risk of dying by 45% and cut the risk of serious injury by 50%[8].

Hence, the main factor that causes the major children to die in the accident is that they can not or do not know how to sit properly behind the car. The inside car safety monitoring system for toddlers is a system to monitor toddlers' activity to ensure that toddlers are in safe condition with the seatbelt.

1.3 Problem Statement

In this bustling modern time, parents are busy and rushing and hence more attention needs to be paid to the safety of toddler in vehicle. The occurrence of toddler and baby without fastening the seatbelt is dangerous because they will probably move or leave the seat and it is really dangerous for toddler's life.

UTeM
اونيۏرسيتي تيکنيکل ملسيا ملاك
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

The first problem statement of this study is that parent cannot get the alert signal when the toddler is not fastening the seatbelt inside the car. When rushing, the parent would put the toddler into the car and then start traveling without checking the toddler's seatbelt condition. At the moment, there is no alert signal to tell the parent that the toddler is in danger because the toddler is not using the seatbelt. For example, J. Dillon et al. proposed a seatbelt reminder to produce visual and auditory signals to alert the driver and passenger about the unfastened seatbelt[9].

The second problem statement is that some toddlers able to unfasten the seatbelt due to unwillingness and discomfort even though the parent had fastened the seatbelt for them. This kind of action would harm their life. An alert signal is important to remind the parent to fasten the seatbelt of toddler.

However, based on past research, there is very little research about toddler safety. Most of the past research related to inside car safety focuses on the driver but not on the safety of passengers in the backseat. The research that is most relevant to toddler safety is that the inside car human pose recognition[10]. There is no monitoring of the situation of the toddler and cannot recognize whether the condition of the toddler is safe.

1.4 Objectives

i. To develop a monitoring system which able to detect baby/toddler presence in the car
ii. To develop a monitoring system which able to classify the condition of a toddler using the seatbelt
iii. To compare the performance with different neural networks (Faster R-CNN, SSD-mobilenet, SSD-Inception).

1.5 Scope

The project will focus on the following segments:

i. The system is designed to be used inside the car.
ii. The system is designed as vision-based and the location of the camera is set in front of the toddler.
iii. The system is designed under good illumination.
iv. The parameter analyzed is only the detection rate and accuracy of the neural network.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

1.6 Organisation of Thesis

Chapter 1 of this report gives a brief introduction to the project, including the motivation, objectives, and problem statement. The background of toddler safety and the importance of the monitoring system are described. This project aims to solve all the problem statements stated.

Chapter 2 is the literature review for related works done by other researchers. It gives the basic knowledge needed to complete this project. The methods and techniques used in problem-solving are compared。

In chapter 3, the proposed methodology to solve the problem of developing the toddler monitoring system is described. The methods to carry out experiments to verify the proposed solution are discussed.

Chapter 4 is the results and discussion of the project outcome in terms of the design and development of the toddler monitoring system. The results of the experiments and simulations are thoroughly analyzed and discussed.

Chapter 5 is the conclusion of the project and suggestions for future works. The overall process of the design and development of toddler monitoring system is discussed.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter describes the technical theories involved in the project and reviews past related works and research. The theoretical background of computer vision, machine learning, human detection techniques and age prediction are discussed. The related research about the type of methods used for human detection has also been written.

2.2 Computer Vision

The recognition of human action is the main issue in computer vision due to its complexity and uncertainty[11]. Computer vision can be defined as a field of artificial intelligence that trains the computer to understand the visual world[12]. Computer vision tasks include acquiring, processing, analyzing, and understanding the images. It can identify all the objects or people in an image using a combination of information can be done with moderate success[13]. The computer can accurately identify and differentiate objects by using digital images capture from cameras and learning models. Due to the advances in deep learning and neural networks, computer vision has been able to imitate humans in some tasks related to detect and label objects[14]. It is all about pattern recognition by training a computer on how to know the visual data. The computer is feed with numerous images with labels and subject to software techniques or algorithms. This allows the computer to find out all the elements that relate to the label. One of the well-known implementations of computer vision is the autonomous vehicle which is a self-driving car[15][16]. In the world of autonomous vehicles, computer vision is often called as "perception" because cameras are the primary but not the only tool that a vehicle uses to perceive its environment.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2.3 Image Pre-processing

Image pre-processing is a technique to perform operations on an image to acquire an enhanced image or retrieve useful information from it[17]. It is defined as a type of signal conditioning by using an image as an input and output with an image or characteristic related to the image. Pre-processing is a general name for operations with images at the lowest level of abstraction with both input and output are intensity images[18]. By definition, image pre-processing involves those processes that start with an image and end with an image[19]. It can be defined as a type of signal conditioning with image input. The process can enhance, compress and study images on a microscopic level. After the software completing the task, a processed image file or an analytical report about the image will be received. The example of pre-processing of an image included grayscale transformation, resizing, cropping etc. For example, Jun-Juh Yan used the grayscale image processing technique to capture the eye of the driver by using edge detection[20].

Some advanced pre-processing comes from machine learning such as edge detection, background subtraction and using the processed data in deep learning. These methods can be directly trained to detect humans or use as a pre-processing approach. For example, Shiyang Yan et al. used GMM to subtract the background and feed the data to deep-CNN[21]. Seyed et al. use edge computation to locate the human location in the image and feed the data into the CNN model for training[22].

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2.4 Machine Learning

Benefit from the father of deep learning, Frank Rosenblatt, father of deep learning, a perceptron was initially simulated on an IBM 704 computer. This led to the construction of an electronic machine[23]. Machine learning is a subfield of artificial intelligence that allows machines to learn from existing data or experiences without being clearly programmed[24]. Machine learning focuses on developing computer programs that can access data and learn from the data collected themselves. Machine learning can be applied in various fields because the machine learning algorithm uses statistics to identify a pattern in a massive dataset. The data can be numbers, words, images, signals etc. The first step of machine learning is to generate example data by collecting and preparing data. Then, the prepared data will be feed

into the machine for training. After the training process, a model will be deployed. The model may be improved by generating more example data. Figure 2-1 shows the flow of machine learning.

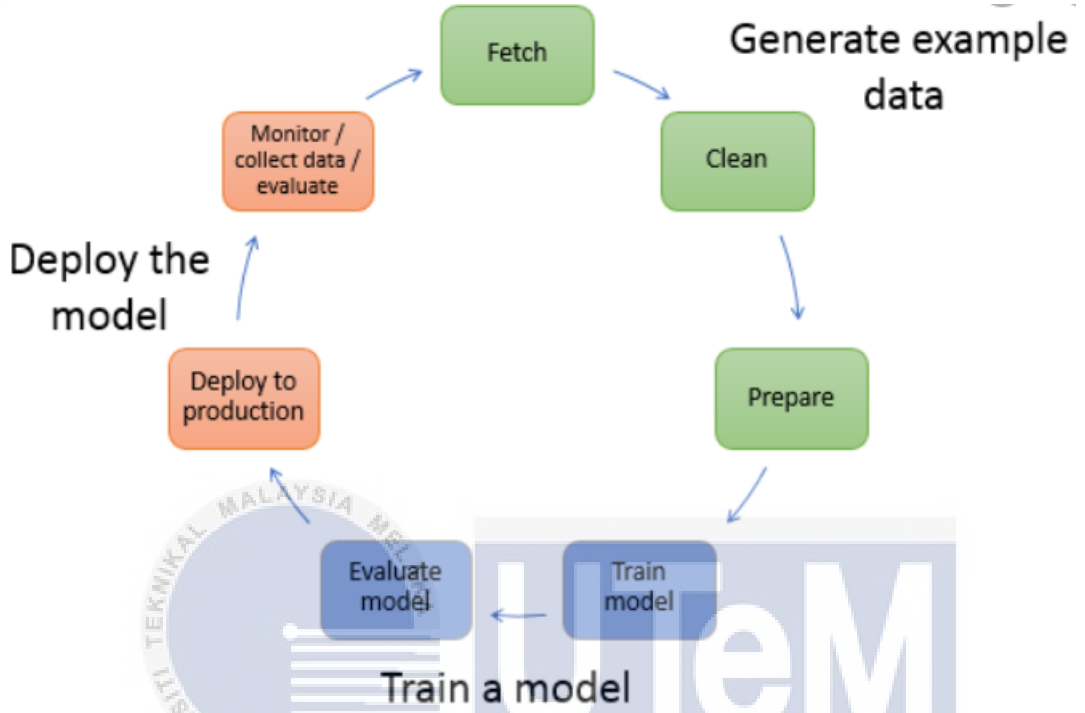

Figure 2-1: Flow of Machine Learning

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Machine learning can be categorized into supervised learning and unsupervised learning[25]. In supervised learning, the data is labeled with annotation to inform the machine exactly what patterns it should look for, while there is no label in unsupervised learning. In unsupervised learning, the machine will just look for whatever patterns can be found. The unsupervised technique is not popular to use because of the less obvious applications. As shown in Table 2-1, different of supervised and unsupervised are listed.

Table 2-1: Different of Supervised And Unsupervised Learning[26]

Supervised Learning	**Unsupervised learning**
Supervised learning algorithms are trained using labeled data.	Unsupervised learning algorithms are trained using unlabelled data.
The supervised learning model takes feedback directly to check if it is predicting correct output or not	The unsupervised learning model does not take any feedback
Input data is provided to the model along with the output	Only input data is provided to the model
The goal of supervised learning is to train the model so that it can predict the output when it is given new data.	The goal of unsupervised learning is to find the hidden patterns and useful insights from the unknown dataset.
Supervised learning model produces an accurate result.	Unsupervised learning model may give less accurate results as compared to supervised learning.

Support vector machine (SVM) is a supervised machine learning model with learning algorithms that can analyze data for classification and regression analysis. For example, Qixiang Ye et al. proposed the PL-SVM method to detect a human in images, which is challenged by the view and posture variation problem[27]. Besides, Yahia Said et al. and Houssein Ahmed et al. show that SVM can perform effectively in human detection with the aids of Histograms of Oriented Gradients (HOG) [28].

The k-nearest neighbour algorithm is another example of supervised machine learning. k-nearest neighbour algorithm (k-NN) is a non-parametric machine learning approach that was first developed by Evelyn Fix and Joseph Hodges in 1951[29] and then expanded by Thomas Cover[30]. K-nearest neighbour algorithm is usually used for classification and regression. Sani et al. used k-NN to recognize the activity of human

Convolutional neural network (CNN) is supervised learning and it is a class of deep neural network that is used in conjunction with a deep learning platform. In the variety of neural networks, the convolutional neural network is the main class for