THE HONG KONG
POLYTECHNIC UNIVERSITY
香港理工大學
Pao Yue-kong Library
包玉剛圖書館

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

Pao Yue-kong Library, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

http://www.lib.polyu.edu.hk

A NOVEL DRIVE-THROUGH VEHICLE PROFILING SYSTEM

LI SHEK PING

MPhil

The Hong Kong Polytechnic University

2022

The Hong Kong Polytechnic University

Department of Industrial and Systems Engineering

A Novel Drive-through Vehicle Profiling System

LI Shek Ping

A thesis submitted in partial fulfilment of the requirements

for the degree of

Master of Philosophy

August 2021

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma in any other University or Institute, except where due acknowledgement and reference have been made in the text.

_____(Signed)

__LI Shek Ping_____(Name of student)

Abstract

As commercially available off-the-shelf (COTS) vehicle profiling products have many limitations and restrictions, a better solution is expected from this industrial research project.

This research aims to develop a novel integrated drive-through vehicle profiling system that automatically classifies incoming vehicles based on multiple factors by critical features of different car type. This system should be able to accurately measure the length and width of moving vehicles and to recognize some critical features to be assigned to appropriate size categories.

Many possible solutions have been proposed in the past decades. The main technological approaches for detection can be divided into two: sensor-based detection methods and vision-based detection methods. Sensor-based detection methods collect different types of data to perform tasks. Vision-based detection methods can solve complex tasks, such as face detection, traffic sign detection and pedestrian detection, etc. With low price tag and easy installation, a vision-based sensor is a natural solution for detection.

The tools for this research are based on the latest technologies, such as Light Detection and Ranging (LiDAR), Computer Vision (CV), geomagnetic sensor, and the

deep learning technique. In order to cater for the needs of Smart City development, they can also provide all the profiling data that can be shared with other systems to form big data.

This research will develop the theory and methodology for a new vehicle profiling system design with lower power consumption, more flexibility, and higher cost-effectiveness. The new system is expected to be easier to install and to generate significant savings on maintenance and total cost of ownership.

Keywords: vehicle profiling, vehicle classification, multiple sensor system, smart city

Acknowledgements

This research project is funded by RF Tech Limited and the University-Industry Collaboration Programme of the Innovation and Technology Commission, The Government of the Hong Kong Special Administrative Region (HKSAR) of the People's Republic of China, which is under Teaching Company Associate (TCA) scheme UIT/140, and it is one of the on-going projects in RF Tech Limited. RF Tech Limited provides all the necessary equipment reasonably available to support research related activities. The R&D department of the RF Tech Limited also involves in this project to provide suitable supporting. The Company is also responsible for the management and administration of this project, as well as the development of the system.

I want to express my deepest thankfulness to my supervisor Dr. K.M. Yu. During my studies, although my works were often not gratified and sluggish, he still tried his best to provide all round support, gave me advice to handle my difficulties and how to proceed next. Besides, I would also like to thank my industrial supervisor Ms. Posy Y.C. Yeung. The motivation of this project was initiated by her, she provided a platform for me to learn and enhance my ability in both academic and real working experience together. Last but not the least, I am very grateful to everyone who had helped and assisted me during the whole study period.

Table of Contents

List of Figures

List of Tables

Abbreviations

AI	Artificial intelligence
Cap.	Capitulus (Latin), means chapter
CNN	Convolution neural network
COCO	Common Objects in COntext
COTS	Commercially available off-the-shelf
CPU	Central Processing Unit
CV	Computer vision
ETC	Electronic toll collection
FoV	Field of vision
GLOH	Gradient location-orientation histogram
GPU	Graphics Processing Unit
ICT	Information and communication technology
IoT	Internet of things
LGV	Light good vehicle
LiDAR	Light Detection and Ranging
MCDC	Multiple-criteria decision-making
MEMS	Micro-electro-mechanical system
MFI	Magnetic field index
PSM	Pages segmentation modes
PVC	Polyvinyl chloride
RFID	Radio Frequency Identification
RGB	Red-Green-Blue, refers to colour
RNN	Recurrent neural network
ROI	Region of interest
RPN	Region proposal network
SBC	Single board computer
SGM	Semi-global matching
SIFT	Scale Invariant Feature Transform
SSD	Single shot multibox detector
SUV	Sport Utility Vehicle
SVM	Support vector machine
TRIZ	теория решения изобретательских задач, in Russian
TOF	Time-of-flight
UTMC	Urban Traffic Management and Control
YOLO	You Only Look Once

Publications arising from the thesis

Paper submitted

S.P. LI, K.M. YU, Y.C. YEUNG, Vehicle Classification System Patent Review with TRIZ

Paper in preparation

S.P. LI, K.M. YU, Y.C. YEUNG, Side-view dimensional profiling of drive-through vehicle using LiDAR, camera and geomagnetic sensor

Chapter 1 Introduction

1.1 General background

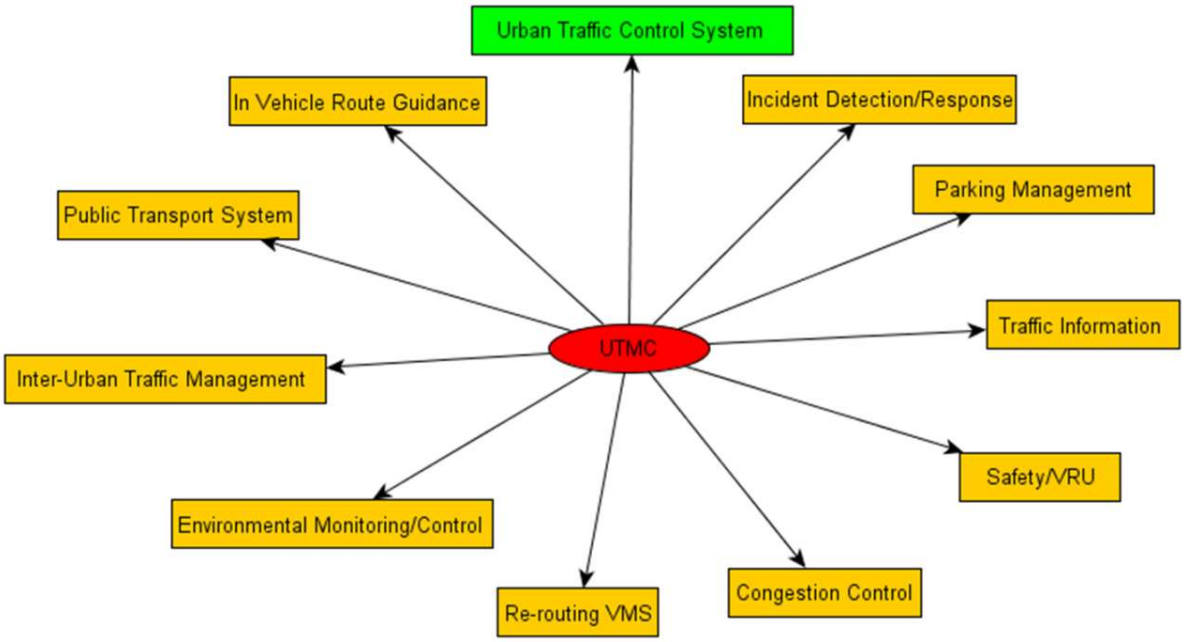


Figure 1.1: Schematic illustration of an Urban Traffic Management and Control (UTMC) system
(Hamilton et al., 2013)

In recent years, the development of smart city has been a topic of much discussion. To remain competitive, many cities, including Hong Kong, have built a strategic blueprint to move toward a smart city, as it can have better management and more efficient allocation of resources in the community. A smart city demands safer and more efficient public services through an intelligent integration of state-of-the-art technologies; thus, automation service will be one of the key operations in the future (Keung et al., 2018). The main idea of a smart city is to make resource allocation and management more efficient through artificial intelligence and data analytics. The Urban Traffic Management and Control (UTMC) is important to build a smart city. UTMC consists of several parts, as shown in Figure 1.1, such that different applications can benefit from using the UTMC system and share information with each other to maximize their efficiency. This research attempts to contribute to the urban traffic control system. It is mainly divided into three parts as shown in Figure 1.2, i.e., sensor networks, traffic management, and safety.

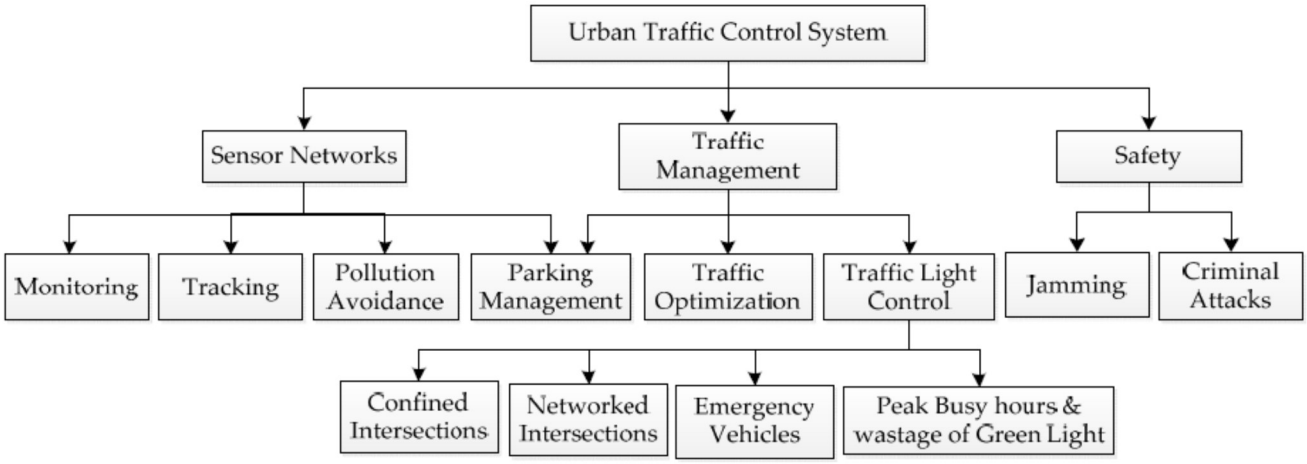


Figure 1.2: Hierarchical functionality of wireless sensor network-based urban traffic management system (UTMC)
(Nellore & Hancke 2016)

As shown in Figure 1.3, many sensors have been proposed to be used in traffic management. For example, video imaging is one of the common sensors which can profile and identify the objects directly from the Red-Green-Blue (RGB) image. Together with deep learning algorithm, the application can be much more robust and diverse. In addition, induction loop and Radio Frequency Identification (RFID) are also common sensors being used in transportation management. The induction loop can detect if any vehicle is moving around the detection area, while RFID can be looking for any registered information about a certain vehicle with a database.

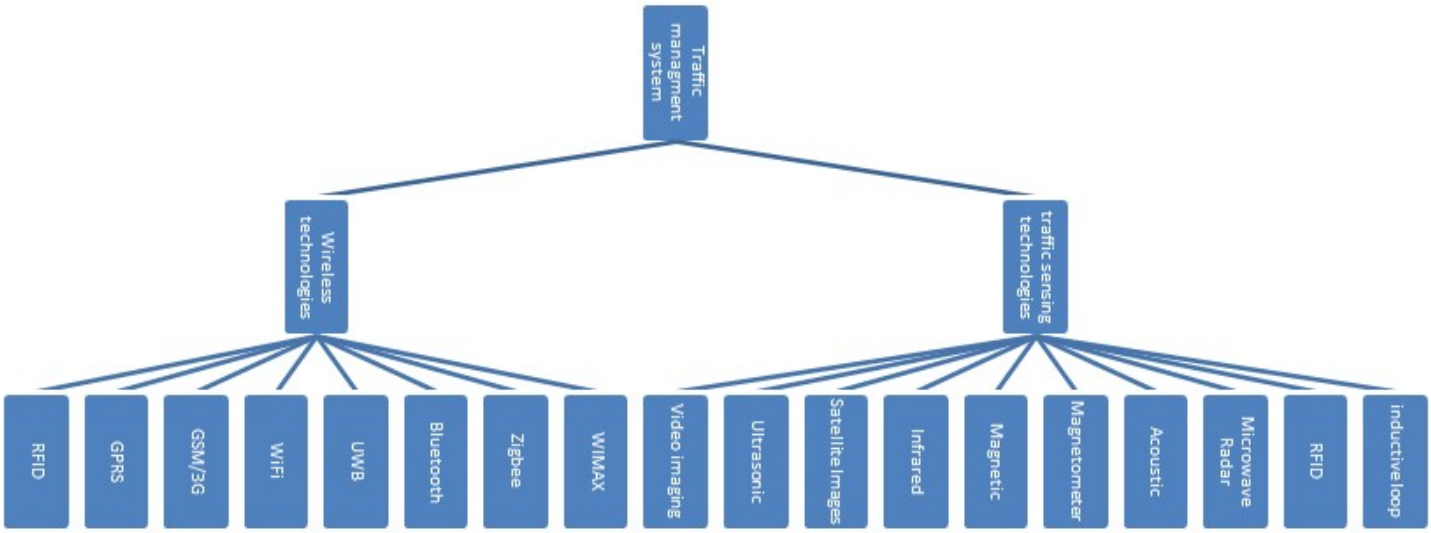


Figure 1.3: Sensors and wireless technologies available for traffic management (Alkhatib, Hnaif & Kanan, 2019)

Sensor networks can help to acquire information on road conditions instantly; this information can, thus, be further used for traffic management and safety control. The use of a vehicle profiling system can help deduce the type of vehicle. This classification information can be employed in a statistical fashion such as road planning for construction and maintenance, road surveillance and traffic control. Specifically, these data are significant for automatic toll charging systems.

The research focuses on developing a vehicle classification methodology through profiling. The expected research result is creating a novel drive-through vehicle profiling system. Though there are several facial recognition systems being used nowadays, a mature vehicle classification system has not been developed yet, as the features of different vehicles are not similar like human faces. Moreover, the environment of a vehicle profiling system is usually worse than facial recognition. Therefore, it is difficult to use the exact same technique used in facial recognition to perform vehicle profiling.

Most profiling systems only use a camera and has limitation to work properly in certain situation such as bad weather. As commercially available off-the-shelf (COTS) vehicle

profiling products have several limitations and restrictions, a better solution is expected from this science-based project. The research aims to combine certain critical sensors and adjust their detection algorithm to complement the weakness of a single camera sensor system.

In Hong Kong, different types of vehicles have their own tunnel rates. Before the vehicle reaches the charging counter, if there is a classification system can identify the type of vehicle and sets the correct rate to be charged. The driver immediately pays the charge through contactless payment cards such as the Octopus card, once they reach the counter. In this case, no staff is required to stay at the charging counter to collect the payment, thus speeding up the payment process, and create an automatically tunnel troll service without any devices or tags inserted into their car.

As this system can collect a considerable amount of data about the type of car passing though the road, a valuable traffic flow data set can be obtained. The analysis of such data can have other applications. For example, decide what advertisement should be placed on a particular traffic road. If there are many buses passing the road, the system may focus on which products can be advertised to the people on the bus. This can help to increase efficiency in advertisement expenditures.

1.2 Problem Statement

There are reputable brands associated with the Vehicle Profiling System, such as sensors produced by Sick AG that have been implemented for traffic management. However, they are less flexible, less cost-effective and require complicated integration of unattended solutions. As a result, they are not very successful in the global market.

Most of the above market products are usually formed based on the ground-based laser, which full name is light amplification by stimulated emission of radiation, scanning data and an analytical approximation of the vehicle profile. The contemporary vehicle profiling system can only recognize small cars, full-size cars, light trucks and Sport Utility Vehicles (SUVs).

However, this classification is mainly based on a simple four-parameter description of the vertical profile of the vehicles. There are no high accuracy vehicle classification and recognition methods for categorizing vehicles to support traffic flow applications.

The technology used in face login authorization is not suitable for all profiling cases. There are limitations to this technology: the object should be placed near the camera so that all its details. For instance, if the user of a laptop is far away from the device, they cannot pass the face login authorization and the system displays a message to inform the user to get closer to the camera. Besides this, it is also not suitable for capturing a large object. To capture the complete picture of a large object, it cannot be placed too close to the camera. Based on the above limitations, this technology is not suitable for vehicle profiling. As the size of a vehicle is much larger than human, the distance between the camera and the detected vehicle is longer; therefore, the image captured may have different degrees of distortion.

1.3 Significance of the Research

To remain competitive, many cities including Hong Kong have a strategic blueprint to move towards smart city, as it can lead to better management and more efficient allocation of resources in the community. A smart city demands safer and more efficient public services through intelligent integration of state-of-the-art technologies; thus, automation service is one of the key areas of a smart city. Simply speaking, latest information and communication technology (ICT) are used for data collection, which is connected to the Internet of Things (IoT) and processed by artificial intelligence and/or data analytic algorithms for better management and more efficient allocation of resources in the community, such as traffic and transportation systems, power plants, water supply networks, waste management, law enforcement, information systems, schools, libraries, hospitals, and others.

For instance, extensive research for the LiDAR technology has been done for autonomous or self-driving vehicles because of its multi-faceted advantages that include

improved safety, reduced congestion, lower emissions and greater mobility. More recently, research has begun exploring the feasibility of using LiDAR data for transportation applications, including infrastructure, emergency and environmental mapping along corridors. An initial investigation on the performance of extracting vehicles from LiDAR data, together with other types of data, and then categorizing them has proved that it is a valuable traffic flow information (Lovas, Toth & Barsi, 2004).

1.4 Research Objectives

Research in the field of profiling system mainly focuses on programming rather than hardware. However, this research still needs to attempt to consider the hardware aspect as well, in order to have an equipment that is cost-effective and yet has a reasonable accuracy rate.

This research focuses on developing smart vehicle profiling solutions for the smart government and citizens. The work to be conducted includes devising new algorithms with a theoretical basis as well as customizing the state-of-the-art embedded information technology, computer vision and deep learning.

Briefly, this research aims to design and develop a profiling method to investigate the factors of rectification to reduce the degree of distortion, to design and develop a multiple sensor architecture and a circuit unit to assist in the identification of the vehicle type, to design and apply an ROI deep learning algorithm to limit the range of analysis, and to evaluate in real life scenarios about the application of vehicle classification.

The profiling system in this research is proposed to be applied in Hong Kong. After profiling the vehicles, classification of the vehicle type can be carried on. The list of vehicle types to be classified is according to the different toll rates for vehicles that go through a road tunnel in Hong Kong. According to the Transport Department of Hong Kong, toll rates for different tunnels are not the same, and the vehicles are divided into 10 types:

1) Motorcycles, motor tricycles
2) Private cars
3) Taxis
4) Public light buses
5) Private light buses
6) Light goods vehicles,
7) Medium goods vehicles
8) Heavy goods vehicles
9) Public and private single-decked buses
10) Public and private double-decked buses

According to Cap. 374 Road Traffic Ordinance: 2 Interpretation (https://www.elegislation.gov.hk/hk/cap374), the definition of the above vehicles is:

- *"**bus** means a motor vehicle constructed or adapted for the carriage of a driver and more than 19 passengers and their personal effects"*
- *"**light bus** means a motor vehicle constructed or adapted for use solely for the carriage of a driver and not more than 19 passengers and their personal effects, but does not include an invalid carriage, motor cycle, motor tricycle, private car or taxi"*
- *"**public light bus** means a light bus, other than any private light bus, which is used or intended for use for hire or reward"*
- *"**private light bus** means—*

(a) a school private light bus; or

(b) a light bus (other than a school private light bus) used or intended for use—

(i) otherwise than for hire or reward; or

(ii) exclusively for the carriage of persons who are disabled persons and persons assisting them, whether or not for hire or reward"

- *"**private car** means a motor vehicle constructed or adapted for use solely for the*

carriage of a driver and not more than 7 passengers and their personal effects but does not include an invalid carriage, motor cycle, motor tricycle or taxi"

- *"**taxi** means a motor vehicle which is registered as a taxi under this Ordinance"*
- *"**light goods vehicle** means a goods vehicle having a permitted gross vehicle weight not exceeding 5.5 tonnes"*
- *"**medium goods vehicle** means a goods vehicle having a permitted gross vehicle weight exceeding 5.5 tonnes but not exceeding 24 tonnes"*
- *"**heavy goods vehicle** means a goods vehicle having a permitted gross vehicle weight exceeding 24 tonnes but not exceeding 38 tonnes"*
- *"**motor cycle** means a two-wheeled motor vehicle with or without a sidecar"*
- *"**motor tricycle** means a three-wheeled motor vehicle other than—*

a motor cycle with a sidecar; and

a village vehicle"

There are private bus and public bus in fact, however, since the tunnel charges for both are the same, no further distinguish is needed for this research.

In addition, Hong Kong tunnels have an electronic toll collection (ETC) service provided by Autotoll Limited, which is the only one to provide this service. Vehicles with an Autotoll Tag do not have to stop and pay cash at toll booths. A unique tag is matched with particular information in the database of the system, which includes information on the type of vehicle. In this research, the profiling system has been designed such that no tags are required to be installed in the vehicle, and no information has to be matched and found from the database.

1.5 Project scope

The aim of this project is to develop a profiling system with the classification function used in UTMC for a future smart city. There are many literatures and patents related to vehicle profiling systems. This project will analyse those works and investigate its innovation trend, and thus propose a novel profiling system. This profiling system will collect RGB images, point cloud and magnetic flux density by camera, LiDAR and geomagnetic sensor respectively. When a vehicle enters the profiling area, the system will profile it and deduce the type of vehicle by analyzing all the data collected by sensors.

1.6 Thesis organization

This thesis is divided into five chapters, and each chapter is organized as below:

- Chapter 1: Introduces the project background, problem statement, significances of the research, objective and organization of the works.
- Chapter 2: Reviews the related knowledge, including the profiling system proposed in the past, object detection in 2D images by camera and deep learning, depth measurement, LiDAR technology and geomagnetic sensor through a comprehensive literature review.
- Chapter 3: Evaluates the patents in public databases and identifies a design trend with the TRIZ technique and gives an overview of the implementation of the proposed solution using LiDAR data, images as well as the measured magnetic flux density.
- Chapter 4: Describes a prototype made for simulations and presents the experimental results, with a comparison with other solutions, discussion on the limitations of the proposed solution, and suggestions any directions for further development.
- Chapter 5: Concludes the thesis and summarizes the contributions of the proposed method.

Chapter 2 Literature Review

2.1 Review of vehicle profiling system

In most of the profiling systems, objects are scanned using a single or multiple cameras within a period. Numerous other devices have been suggested for object scanning in these systems. For vehicle profiling, the commonly used sensors include video camera, ultrasonic device, inductive loop, geomagnetic sensor and infrared camera (Klein, 2002). The ways in which data are collected vary, as each sensor has its own set of strengths and weaknesses with respect to cost, accuracy and performance. Based on the type of data collected, several vehicle classifications have been proposed by different researchers. Examples include the axle-based methods (Martin, Feng, & Wang, 2003), the vehicle sound-based methods (Nooralahiyan et al., 1997), the vehicle-length-based methods (Meta & Cinsdikici, 2010), the video-based methods (Zhang, Avery & Wang, 2007), and the radar-based methods (Urazghildiiev et al., 2007). In this research, a novel profiling system using multiple sensors (camera, LiDAR, and geomagnetic sensor) as well as applying both vehicle-length-based method and video-based method is proposed. A more technical review about the proposed solution is discussed in the following sections, and its general idea is illustrated in Figure 2.1.

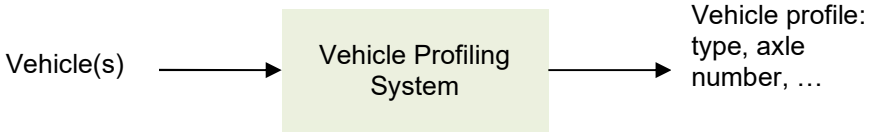


Figure 2.1: Vehicle profiling system

2.2 Object detection in RGB images

A camera is a commonly used device in object detection. In transportation engineering, a camera can take a photo of vehicles which contain some information about the vehicles. One of its applications nowadays is in roadside surveillance systems, used to validate the accuracy of parking-gap estimation and system performance, aiming at facilitating traffic and fleet management for smart mobility. This surveillance system insists on measuring the length and

height of the vehicles. The proposed mounting position of the camera is on streetlights or nearby facilities, with a 45°-90° viewing angle (Ho et al. 2019), as shown in Figure 2.2.

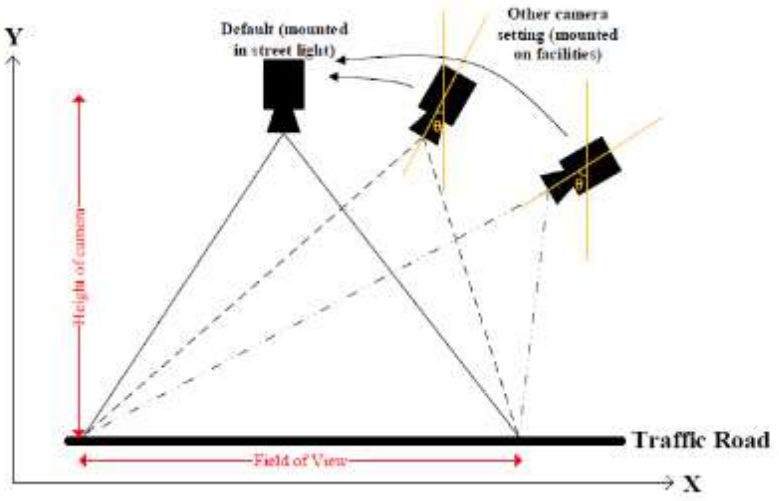


Figure 2.2: Illustration of various camera positions
(Ho et al., 2019)

However, the technological challenge is that the performance of this approach is affected when camera lenses and different lighting conditions cause the barrel distortion and severe blurring in photos are caused by camera lenses and different lighting conditions. It In addition, it is also difficult to overlap nearly exact images under non-identical exposures to produce seamless stitching results. This may greatly affect the data analysis of the vehicles. Therefore, dimension measurement of the camera is not preferred.

Machine learning on images is one of the key technologies of object recognition, as it allows the system to learn how to classify objects. It is a field of study that gives computers the ability to learn without being explicitly programmed.

Finding of vehicle features can be obtained as recognition problem. Local Photometric Descriptors (usually called Local Descriptor) are N-dimensional feature vectors that are used to describe special/colour information of a certain point in an image. A cross-correlation between feature vectors is useful to compute the correspondence of two points between two

images. For example, a basic Local Descriptor stores the RGB value of a pixel in the image. However, the state of image pixels of the same object changes from time to time as a result of the change in lighting condition and viewing direction. Therefore, an image pixel descriptor is impractical to the process of object recognition. In recent years, different researchers have tried to develop a new Local Descriptor that is invariant to scaling, rotation, noise and illumination. Some remarkable descriptors are Scale Invariant Feature Transform (SIFT), Gradient location-orientation histogram (GLOH) and Moment Invariants. The invariant characteristic provides a foundation for a robust, practical image retrieval method in Book Cover Recognition Problem. These robust descriptors find a wide range of application in the machine vision area, including object recognition, image retrieval and video data mining. However, this algorithm is strong for detecting same object with the same feature. The same types of vehicles but of different model will have different appearance. In consideration with the image noise, the performance of using local descriptor-related method is poor as it cannot detect exactly the same object as in training.

Although it is not suitable for vehicle classification, it can be considered for detecting part of vehicle features. The detection of a wheel can result in counting the axles of a vehicle, which is one of the important features in vehicle profiling. There are several findings with respect to axle counting. Some methods propose utilizing several laser beams on the close surface of the road (Xiang, Otto & Wen, 2008) or optic pairs (Ueda et al., 1997). If the wheels pass across them, the laser beam or optic pairs are occluded, allowing for the number of axles to be counted. However, the error rate of these methods is too high as its noise is significant, snow or dirt will highly affect its performance, and it requires an additional cost of the system. LiDAR is another tool that can be possibly used instead of laser beams for the purpose of counting axles (Sato, Aoki & Takebayashi, 2014). However, while the strengths of LiDAR include measuring the depth of an object, LiDAR data may sometimes fail to identify whether the object is a wheel. Therefore, LiDAR is not a robust solution. Another proposed method to

use physics sensor such as a capacitive sensor or pressure sensor. Detection can be done on the top of the sensors, as the wheel is passing through; however, these types of sensors have a short life cycle and require an additional cost of the system. To minimize the physical problem, counting axles on an images base is a better solution.

In fact, all the machine learning approaches mentioned above need to define the main feature of an object first and then use another technique such as support vector machine (SVM) for the classification. However, machine learning approaches have their own limitations. Even if the amount of data increases, the performance will not improve proportionally. With the rapid growth of information technology, the quality and quantity of the data set obtained and the computing power of machines are much better than before. It is a waste to not use the benefit of technology. Recently, a machine learning method, namely deep learning, has been developed. The heart of deep learning approaches is utilizing data as much as possible, to boost its performance. With the aid of Graphics Processing Unit (GPU) and a set of data images from the Internet, the performance of approached based on deep learning is better than traditional machine learning approaches in most cases as described in Figure 2.3, which is why the field of object recognition has become becoming popular in recent years.

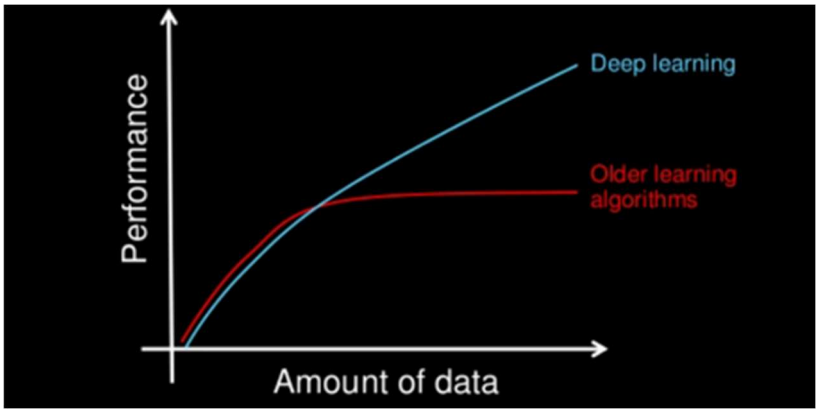


Figure 2.3: Deep learning vs. older learning algorithms
(Brownlee, 2020)

In most research and application nowadays, the pioneers are mainly using a camera to capture images for recognition purposes, and their main methodology is to use deep learning

approaches, which is one sub-group in machine learning. Deep learning tends to result in higher accuracy, but has higher hardware requirements or requires a longer training time, and performs exceptionally well on machine perception tasks that involve unstructured data such as blobs of pixels or text. As deep learning is a powerful algorithm, research in this technology has won the Turing Award in 2018. Compared with classic object detection, the deep learning algorithm does not need to detect the accurate characteristic of the object, and can simply focus on some common features. Thus, it can greatly reduce processing time and increase the accuracy rate.

There are several deep learning algorithms in object detection, for example, Faster R-CNN (Region-based convolution neural network), SSD (Single shot multi-box detector) and YOLO (You only look once). Each of them has their own advantages and disadvantages, and so people may choose one of them to suit their application or for further development. These algorithms are commonly used for different purposes, and their performance evaluation is done using five metrics: precision, recall, F1 score, quality and processing time. For Faster R-CNN and YOLO, both algorithms are comparable in precision, which means that both have a high capability to correctly classify car object in the image. However, YOLO is more capable of extracting all the cars and its processing time is faster (Benjdira et al., 2019).

Up to now, there have been five versions for YOLO. The first version was released in 2016. YOLO is an end-to-end object detection algorithm; it means that once an image is input, the next step gives out the location of the bounding box and detection result directly, without the process of handling region proposal in faster R-CNN, and hence, the detection speed in YOLO is faster. Moreover, this benefit can help to achieve the requirement of real-time detection in this research.

The basic procedure is shown in Figure 2.4. The input image will be divided into S×S grid cells, and each cell will predict B boundary boxes. Each boundary box contains 5 elements: (x, y, w, h) and a box confidence score. The confidence score reflects how likely the box contains an object and how accurate the boundary box is.

Each cell has C-conditional class probabilities. The conditional class probability is the probability that the detected object belongs to a particular class (one probability per category for each cell). So, YOLO's prediction has a shape of (S, S, B×5 + C). To make a final prediction, it keeps those with high box confidence scores (greater than 0.25) as its final predictions.

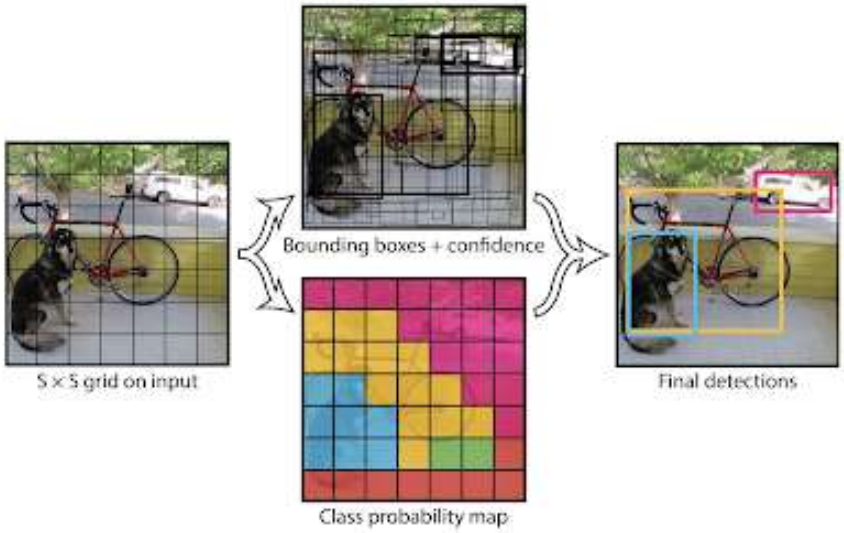


Figure 2.4: The working model of YOLO
(Redmon et al., 2016)

2.3 Depth measurement

The depth represents the distance between the sensor and the object. This information is important as it can be used to calculate the dimension of an object. Since there is a wide application range, there is a huge motivation of developing a cost-effective and high accuracy depth estimation solution. Most of the depth estimation is image-based.

In the history of depth estimation from image, Semi-Global Matching (SGM) is one of the classical algorithms, where the depth map is generated from a pair of stereo images (Hirschmuller, 2007). The estimation process mainly divided into four steps which are matching cost computation, cost aggregation, disparity computation, and disparity refinement. A matching cost computation occurs by inputting stereo images. These matching costs are refined within a certain neighbourhood in cost aggregation. Depth estimation can then be derived from matching the cost, and finally processing a refinement by removing peaks,

increasing the accuracy by sub-pixel interpolation or interpolating gaps. However, this method has errors occurring at ill-posed regions and streaking artefacts in the depth map.

Zbontar and LeCun (2016) combine deep learning with SGM in depth estimation. This method matches patches using CNN instead of matching pixel. First, the cost volume of the obtained patches from CNN output is computed. The cost volume of disparity image is then refined by using SGM.

If the training data set is sufficient, the superiority of the deep learning algorithm becomes more evident, and hence full deep learning is expected, which comes to the idea of end-to-end learning. A research study proposed using a convolutional network by stereo images for the entire learning task (Mayer et al., 2016). They had introduced over 35000 pairs of stereo images with ground truth disparity, scene flow and optical flow for training. The main limitation of supervised learning is the training process where labelled data, ground truth depth information and cumbersome pre-process analysis are required, and it consumes a huge amount of resources.

Apart from a supervised learning method, there is an unsupervised leaning method as well. There are two main training approaches for the unsupervised method; one is using a monocular video; and the other is using synchronized pairs of stereo images. Garg et al. (2016) proposed an unsupervised learning method by CNN for single view depth estimation. The weakness of most deep CNN learning algorithms is the requirement for a set of annotated ground-truth depth data. In their work, they only need a pair of stereo images to train their system. One of the stereo images is passed on to CNN to generate a predicted inverse depth map. After inverse warping between this map and another stereo image, a wrap image is formed. By comparing the original image and the warped image, a reconstruction error can be deduced. The acquisition of this training data can be through any stereo images captured by a normal camera, and does not need any depth sensor or manual annotation. The main contribution of this method is that their network learns a deep CNN for monocular estimation without requiring

a pre-training stage or annotated ground-truth depth. The limitations of this method are that it is slow and inefficient, the loss is high, and the depth map has a "texture-copy" artefacts problem. A paper proposed another modified method (Godard, Aodha & Brostow, 2017). Their algorithm can deal with the "texture-copy" artefacts problem. The left image is used to generate both the left and the right depth maps. An inverse warping is then performed with the original right and left image respectively, to finally form both the left and right warped images. This method improves quality and increases robustness. A paper proposed a further modified self-supervised learning method which trains on pairs of stereo images simultaneously and symmetrically (Goldman, Hassner & Avidan, 2019). It shows how a network trained on stereo images can naturally be used for monocular depth estimation at test time. In some cases, this monocular disparity estimation of this method even outperforms supervised systems.

However, stereovision suffers from some serious problems that cause lower accuracy of depth estimation. First, the accuracy often depends on the distance of the objects. Second, the estimation is poor in a low-light environment. Third, a high computational resource is required for processing the estimation. Last but not the least, the image-base method is still not accurate enough.

2.4 LiDAR technology

Other than image-based method estimation, there are, in fact, other methods which can measure the depth of an object. Light-based depth estimation is one of the most famous methods, and the sensor usually employed is LiDAR (Werner, 2014). LiDARs were first used in military applications to detect and identify vehicles by their range profiles (Sun et al., 2018). It is a surveying method that measures distance to a target by illuminating the target with a laser light and measuring the reflected light with a sensor, thereby analysing the accurate distance between LiDAR and target. In fact, each light pulse received can locate a point on the reflection surface, and by getting those points reflected by an object or vehicle, its dimension

or shape can be generated. Nowadays, LiDAR is widely used for object perception and recognition as 1D, 2D, and 3D variants in different industries, for instance, industrial manufacturing, traffic systems, logistics, and autonomous vehicle field.

Generally, there are a few types of laser range finding principles nowadays: time-of-flight (TOF), triangulation and phase shift measurement. All these principles are established not only based on three characteristics but also on the advantages of laser; straight propagation of unobstructed light, known velocity of light travelling in space and easy detection as only one wavelength of light is generated from laser.

To begin with, range finding through TOF is extensively used in commercial products. Basically, the concept is to measure the duration between each laser pulse from being emitted to being received.

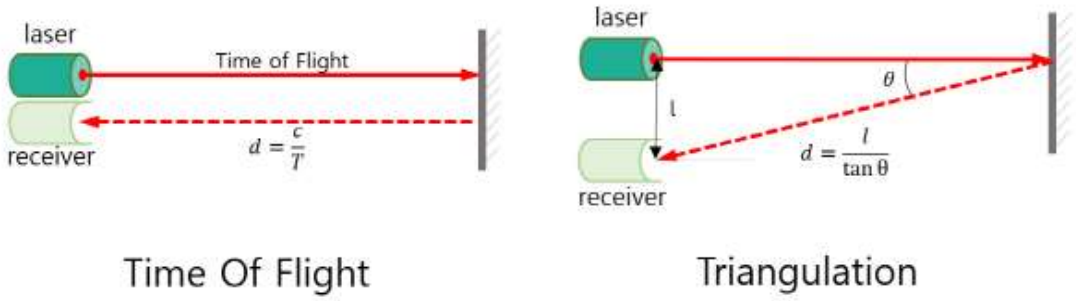


Figure 2.5: Working principle of LiDAR
(https://steemit.com/technology/@rnjena/low-cost-solid-state-2d-lidar)

In Figure 2.5, if applying the ToF method, the laser and receiver should be very close. When the traveling time (T) is obtained, where c is the speed of light (3×10^8 ms^{-1}), the distance (d) can be calculated. On the other hand, for the triangulation method, the position of receiver should be known, and hence can be a bit far away, the θ indicates that there is an angular distance between the laser emitter and the receiver. This separation is small enough that there will be no adverse effect on the accuracy during the range-finding process (Angelopoulou & Wright Jr, 1999). By the above method, a 3D shape of objects can be generated, it provide direct geometric information of the environment (Wulf & Wagner, 2003).

Unlike TOF, there is another method that a continuous wave laser is used instead of laser pulses. This continuous wave acts as a carrier wave that modulates the intensity at a specific frequency. The idea is to compare the difference in phase between emitted and received signals (Hu et al., 2011). This range can be determined by the following equation:

$$R = \frac{c}{2f} \cdot \frac{\Delta\varphi}{2\pi} = \frac{\lambda}{2} \cdot \frac{\Delta\varphi}{2\pi} \tag{2.1}$$

where c is the speed of the light, f is the modulation frequency, λ is the wavelength and Δφ is the phase shift. The following diagram describes the general idea of wave modulations in phase-shift measurement (Pfeifer & Briese, 2007), as shown in Figure 2.6.

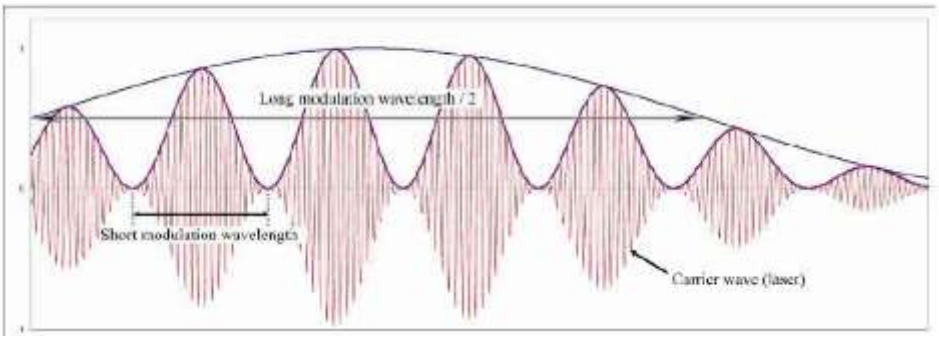


Figure 2.6: Wave modulations in phase-shift measurement
(Pfeifer & Briese, 2007)

Examples of world leading vendors are Velodyne (http://velodynelidar.com/) and Quanergy (http://quanergy.com/) from America and Canada respectively; and Slamtec (https://www.slamtec.com/cn/) in the Mainland China.

Compared to other 3D sensing techniques, such as stereo cameras and radar, LiDAR sensors can provide high resolution and highly accurate measurements of the surroundings under various weather conditions.

There are research studies that use LiDAR for vehicle classification as well. A research study supposed to mount two LiDARs in a side-fire configuration next to the road (Coifman & Lee, 2011), as shown in Figure 2.7. After distinguishing between vehicle or non-vehicle, the

vehicle body information is extracted from the two LiDARs. Accurate LiDAR data can deduce the vehicle length and height, as well as the speed of the vehicle. It insists on the dimensional variation along the vehicle body, so the authors propose several features including Vehicle length (VL), Vehicle height (VH), Detection of middle drop (DMD), Vehicle height at middle drop (VHMD), Front vehicle height (FVH), Front vehicle length (FVL), Rear vehicle height (RVH) and Rear vehicle length (RVL). By setting a suitable threshold of each feature, a decision tree is formed, and the classification is simply based on it.

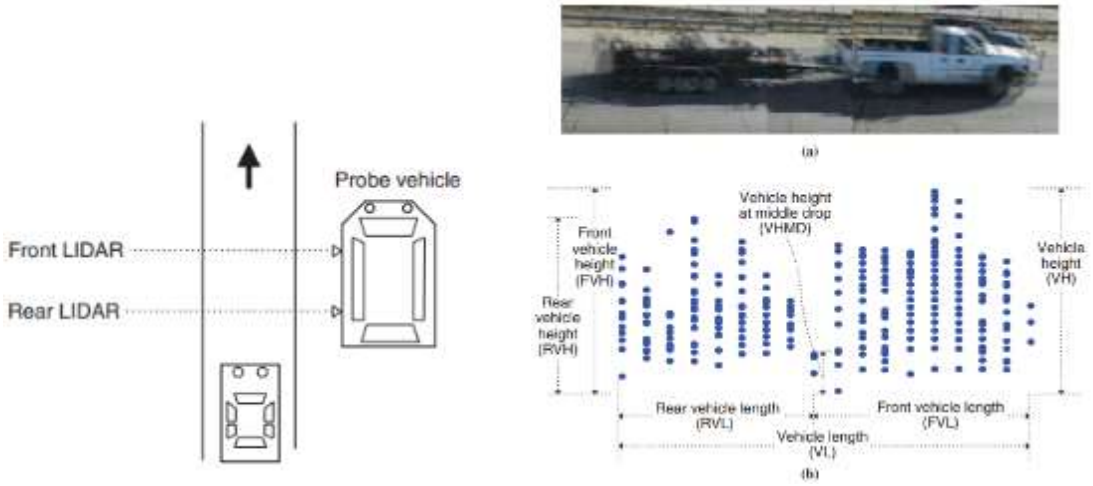


Figure 2.7: Measuring of vehicle dimension by two LiDARs

Besides dimension measuring, LiDAR data can generate a 3D model of an object, and there is an idea for using the deep learning algorithm on a 3D model, similar to object detection on 2D images. The data of a 3D model usually are a cloud of points; the shape of an object is presented by several points. LiDAR is one of the point-capturing devices used nowadays. Based on a technique similar to YOLO in 2D images, there is a research study that use a modified YOLO to detect 3D points of cloud (Simony, Milz, Amende & Gross, 2018).

Apart from a modified version of YOLO, another point cloud deep learning algorithm called Voxelnet is proposed (Zhou & Tuzel, 2018). An example is illustrated in Figure 2.8. It is a famous algorithm for LiDAR object detection. The general working procedure is to use a

recurrent neural network (RNN). It will scan the target area to form a point cloud image, then bound the target object by a bounding box.

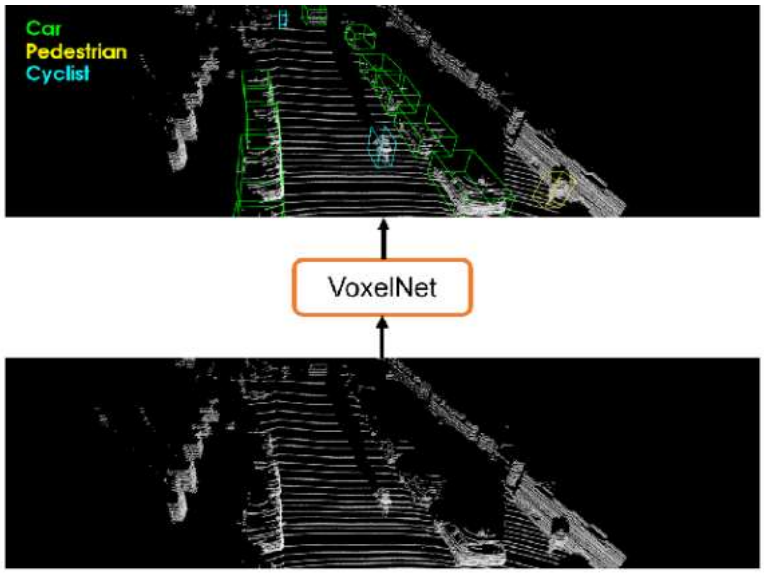


Figure 2.8: Bounding from LiDAR point cloud data by VoxelNet
(Zhou & Tuzel, 2018)

The featured learning network takes a raw point cloud as input, partitions the space into voxels, and transforms points within each voxel to a vector representation characterizing the shape information. The space is represented as a sparse 4D tensor. The convolution middle layers process the 4D tensor to aggregate spatial context. Finally, a region proposal network (RPN) generates 3D detection. The above workflow is shown in Figure. 2.9.

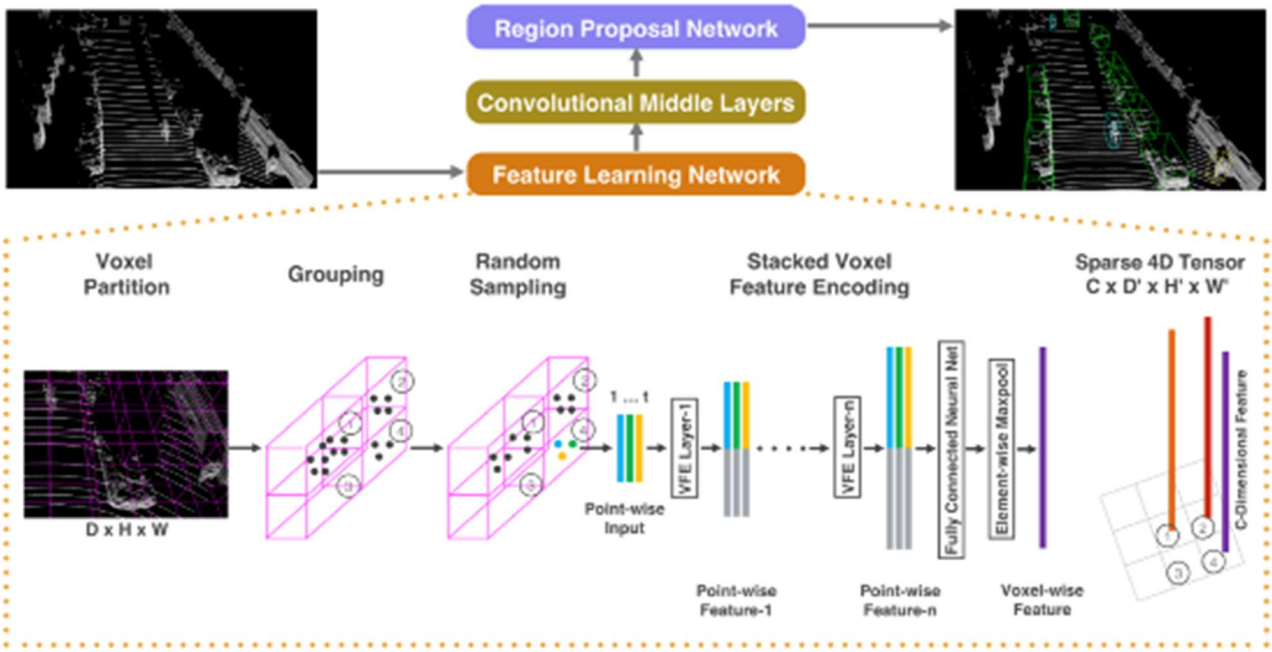


Figure 2.9: Workflow of the VoxelNet
(Zhou & Tuzel, 2018)

The LiDAR can provide a point cloud and its geometry can be adequately modelled to form a height histogram of the car. After some data processing such as histogram smoothing, a "picture" of the car can be generated, as shown as Figure 2.10.

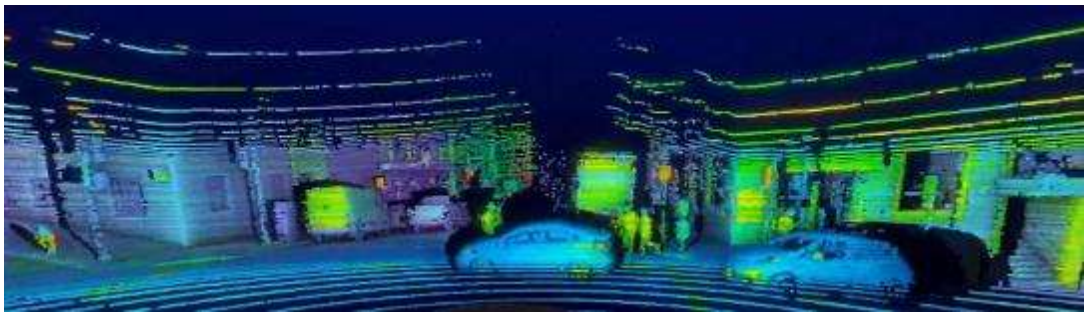


Figure 2.10: Point cloud data
(https://velodynelidar.com/vls-128.html)

Similar to the deep learning in images, a set of point cloud is needed for training purposes. However, point cloud labeling is more complex than images. Figure 2.11 shows an example of labeling some point cloud data, the 3D bounding box should bind the target object and give an annotation.

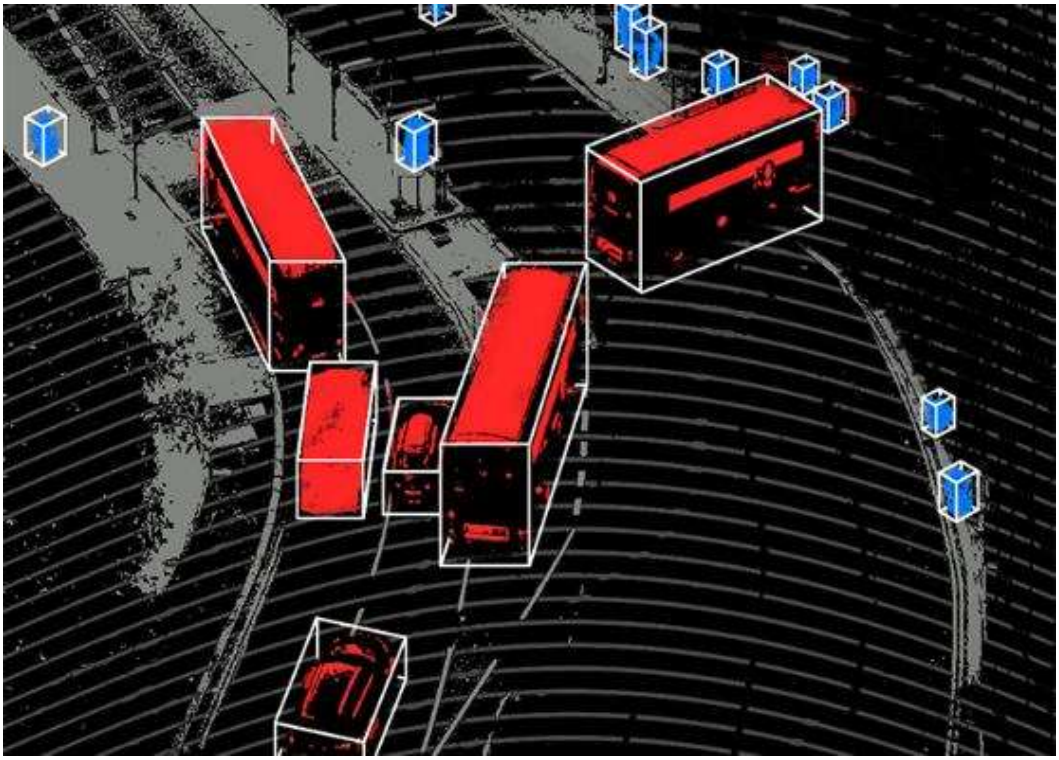


Figure 2.11: 3D point cloud annotation
(https://www.cloudfactory.com/computer-vision)

In fact, there is a significant problem in 3D deep learning. As the performance of deep learning is directly affected by the quality and quantity of the training data, the actual performance in 3D is technically worse than in 2D images. The data of 3D point clouds are hard to collect. Unlike 2D images, 3D points cloud cannot be captured by general camera. Lots of 2D image training data can be found on the Internet, which the 3D point cloud data does not. The photos of buses, taxis and motorcycles can be easily found, but there is either less 3D points cloud in all types of vehicles or none. As a result, it takes a huge amount of time and resources to collect enough training data which is not cost-effective. Therefore, it is better than using LiDAR just for direct dimension measuring but not deep learning.

2.5 Geomagnetic sensor and magnetic field

Besides camera and LiDAR, there are researches proposed to use geomagnetic sensor for vehicle profiling. The Earth has a huge magnetic field surrounding it. Some scientists

suggest that some animals have a “magnetic feeling” and use it for guiding some actions, like birds usually know the southerly direction and fly toward the south to escape cold weather. A geomagnetic sensor can measure the magnetic field of the Earth. Since metal can interfere with magnetic flux density, a larger metal result in greater change in this density. Therefore, as all the vehicles are made of metal, the magnetic sensor can classify vehicles by measuring the change of magnetic flux density.

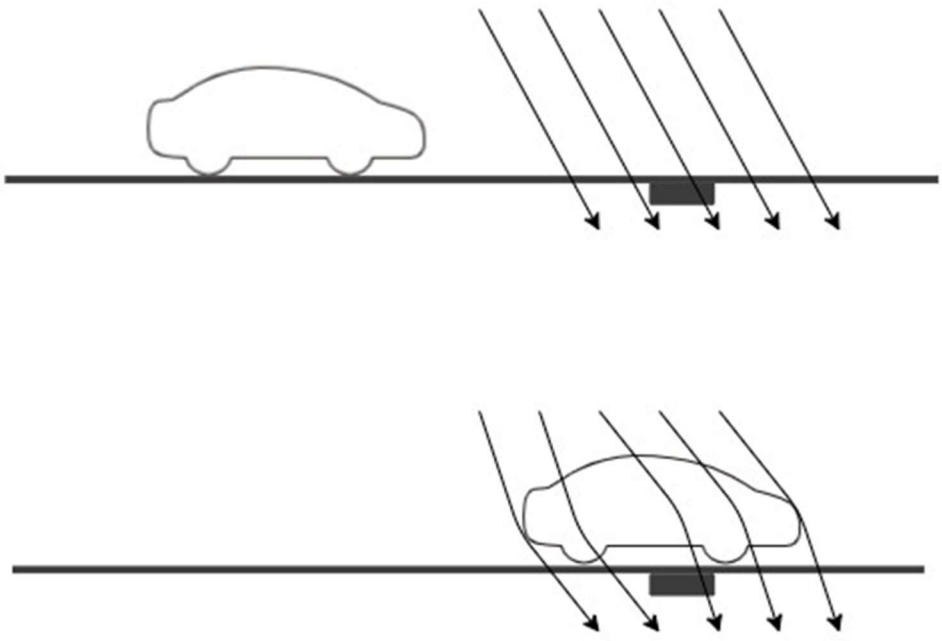


Figure 2.12: The magnetic flux density is different between before and after the vehicle across

In Figure 2.12, at the beginning, the geomagnetic sensor measures the magnetic flux density without vehicles, and this value acts as a reference value. When a vehicle reaches the detecting area, the value of magnetic flux density will increase. Different vehicle sizes must have different increasing values. The type of vehicle can be identified by calculating the differences between increasing value and reference value. The mapping table between vehicle type and magnetic flux density change should be further evaluated by analysing more testing data.

The main advantages of using the geomagnetic sensor are on the manufacturing cost, maintenance cost, installation difficultness and weather resistant issue (Bugdol et al., 2014). However, a limitation of the geomagnetic sensor is that it cannot distinguish type of vehicles with similar size in a high accuracy rate, a related result as shown in Table 2.1 (Kaewkamnerd et al., 2010). This is because their material is comprised of nearly the same mass of metal, and the change in the magnetic flux field is, therefore, similar. On the other hand, if it is used for detecting the existing status of a vehicle, it is a good application. Once the magnetic field flux has changed, it insists that a vehicle lays on the road. Though nowadays, a parking detecting system has used a geomagnetic sensor to sense vehicles.

Table 2.1: Classification of results into 4 types of vehicles

Type	Classification Result (%)
Motorcycle	100
Car	82.46
Van	78.57
Pickup	65.71

Chapter 3 Methodology and Implementation

3.1 Patents of vehicle profiling and classification

After vehicle profiling, to make the system marketable in daily life, a classification action should be carried on. In transportation engineering, vehicle classification is a labour intensive and an error-prone task. Automated vehicle classification system is a much-discussed topics nowadays. The main purpose of this system is to identify the type of passing vehicles automatically by analysing different kinds of data from a different sensor(s). Vehicle classification is one part of a future smart city (Haferkamp et al., 2017), as it has many potential smart applications, for instance, it can be used in traffic safety analysis (Alkheder et al., 2013), automated toll payment (Suryatali & Dharmadhikari, 2015), road maintenance, road infrastructure management (Masino et al., 2017), and other applications related to road traffic. All of these applications can greatly increase the efficiency of resource allocation of city related road management.

As one can foresee the benefit of the vehicle classification system, there have been many proposed vehicle classification systems with different algorithms, but each of them may have its limitations. To fulfil different requirements, people try to keep on optimizing existing solutions or design a new solution, in order to resolve different kinds of limitations and improve adaptability in various environments. Therefore, there are patents related to the vehicle classification system. In addition, vehicle classification is a relatively new discipline. The earliest patent is published in 1974, the patent number of which is US3794966.

These patents are investigated by the popular теория решения изобретательских задач (in Russian, TRIZ) method. TRIZ is a theory of inventive problem solving which is developed from the research of nearly 200 thousand patents by the Soviet Union inventor Genrich Altshuller (Altshuller et al., 1999). He believes that invention has a track to follow, as most innovation is based on previous inventions and improved by known knowledge and technology. By analysing such a large number of patents, a contradiction matrix is designed and it can be used to check the most likely inventive principles applied in certain conditions.

However, the matrix was developed around the year 1956 when electronics and computer technology were still in their infancy (Souchkov, 2015). A new version of a contradiction matrix was proposed by Mann (2010). In this section, the new contradiction matrix will be used.

The raw data extracted from the patents are summarized in Table 3.1. The contradicting parameters of TRIZ have been identified and tabulated in Table 3.2. The inventive principles identified from the patents and those recommended from the contradiction matrices are compared in Table 3.3.

Table 3.1: Patent extracted information

Item	Patent			Improving features	Worsening features	Solution Inventor Used
	Number	Title	Publication Date			
1	US3794966	Automatic vehicle classification and ticket issuing system	1974/2	Increase the degree of automation by automatic distribution of tickets	Not robust enough as the classification results are affected by tailgating, variable length of vehicles and traffic patterns	Adding a barrier to the traffic lane and controlling the flow of vehicles through feedback
2	US4789941	Computerized vehicle classification system	1988/12	Classifying vehicles automatically with a treadle arrangement and counting axles	Axis counting by a treadle is not robust due to a wide variety of vehicle configurations and hostile working environments	Profiling the vehicles' height by ultrasonic additionally
3	US5392034	Vehicle classification system using profile	1995/2	Classifying vehicles automatically based on detected information	The system fails in some cases if license plates cannot be detected	Profiling the shape through optical sensor
4	US5446291	Method for classifying vehicles passing a predetermined waypoint	1995/8	Detect the characteristic of vehicles automatically by using a seismic detector	A distinction cannot be made between specific vehicle types within a generic type	Profiling the height by using laser
5	US5619616	Vehicle classification system using a passive audio input to a neural network	1997/4	More information can be obtained for classification purposes by using radar, ultrasound	The performance is affected by the environment such as weather and light conditions	Using passive audio sensors instead of active sensor like ultrasound
6	US5717390	Doppler-RADAR based automatic vehicle-classification system	1998/2	Improving the classification accuracy	Needs to place an identification label for each vehicle and hence is difficult to apply	Scan the vehicles' shape by using radar
7	US6014447	Passive vehicle classification using low frequency electromagnetic emanations	2000/1	Collect less amount and types of sensed data for enhancing efficiency	The sensed data may not characterize the vehicle with sufficient specificity to classify it	Only use acoustic sensor by improving its data processing algorithm

Table 3.1 (Con't): Patent extracted information

Item	Patent			Improving features	Worsening features Number	Solution Inventor Used
	Number	Title	Publication Date			
8	US2002 0140924	Vehicle classification and axle counting sensor system and method	2002/ 10	Improving the classification accuracy by using a video and collecting other types of data like axis counting	Highly increases the installation cost by collecting axis counting data from treadle	Profiling and axis counting by a scabbing time-of-flight laser rangefinder
9	US6828 920	System and method for classifying vehicles	2004/ 12	Decrease manufacturing cost by reducing the number of installations of RF transponders	Reliability decreases or may even not work as not enough information collected	Use a single inductive loop to generate a field for electrically sensing vehicles
10	US6865 518	Method and device for classifying vehicles	2005/ 3	More refined classification is required	More sensors are needed	Obtaining a digitized signal from the electromagnetic signal from one inductive loop, which is digitized, sequenced, and time-stamped data.
11	US6894 233	Systems and methods for classifying vehicles	2005/ 5	Utilizing the existing assets in detecting aspect	Require extensive renovations and/or road construction for their installation	Using other field of sensor like capacitive sensor together with existing assets
12	US2005 0267657	Method for vehicle classification	2005/ 12	Decrease manufacturing cost by reducing the number of installations of any affixing items	Reliability decreases or may even not work as the apparatus is not sufficient	Use image capture instead of installing anything inside the vehicle
13	US8311 343	Vehicle classification by image processing with laser range finder	2012/ 11	Improving the precision of measurement in profiling	The distance between the sensor and vehicle needs to be fixed to fix the measuring reference	Scanning the vehicles by a camera with an additional rangefinder
14	US2015 0269444	Automatic classification system for motor vehicles	2015/ 9	Make it possible to measure physical characteristics of the vehicles by using multiple sensors on various features	Not implemented in a simple or cost-effective manner	Capturing images separately by multiple cameras from birds-eye view and three-quarter front view

Table 3.1 (Con't): Patent extracted information

Item	Patent			Improving features	Worsening features Number	Solution Inventor Used
	Number	Title	Publication Date			
15	US9239955	Method and system for vehicle classification	2016/1	To ensure the images of vehicles are valid for classification process	The area of obstruction caused by overlapping of vehicles cannot be too large	Capturing the images from the rear view of vehicles
16	US9361798	Vehicle classification system and method	2016/6	Improving the accuracy by using dual inductive loops	Some locations are only allowed a single inductive loop embedded in the road	Generating a signal waveform from a signal in a single inductive loop generated by a passing vehicle
17	US9466000	Dynamic Bayesian networks for vehicle classification in video	2016/10	Automation classification by image, video streams and computer	Several problems occur such as occlusion, tracking a moving object, shadows, rotation, lack of color invariance	Classifying from rear view instead of side view
18	US9519060	Methods and systems for vehicle classification from laser scans using global alignment	2016/12	Improving the performance on both scope and efficiency	Requires intense human labor	Using data to generate the shape of the vehicle instead of length
19	US9683836	Vehicle classification from laser scanners using fisher and profile signature	2017/6	By using laser scan, it can be placed in sensitive locations such as bridges	Limited in scope and efficiency	Use of a linear classifier with more complex features than just the high-level features or the raw profiles
20	US9977972	3-D model based method for detecting and classifying vehicles in aerial imagery	2018/5	Obtaining geometric relation by 3D approach	Faulty edge detection, lack of scene contrast, blurry imagery, scene clutter and noise	Introduce a 3D method in aerial imagery
21	US9984704	Vehicle classification system and method	2018/5	Count the number of axis automatically by piezoelectric sensor	Piezoelectric sensor broken frequently	Use of contactless non-mechanical sensor, i.e., sound sensor
22	US10345449	Vehicle classification using a recurrent neural network (RNN)	2019/7	Reducing the installation cost or cost of apparatus	Reliability decreases or may even not work as apparatus are not sufficient to collect enough data	Analysis of the vehicles' types by using GPS track data from the driver's mobile phone

Table 3.2: Corresponding TRIZ contradicting parameters

	Improving Feature		Worsening Feature	
Patent Item	**Technical Parameter**	**Name/description**	**Technical Parameter**	**Name/description**
1	43	Automation	35	Reliability/Robustness
2	43	Automation	35	Reliability/Robustness
3	43	Automation	35	Reliability/Robustness
4	43	Automation	35	Reliability/Robustness
5	49	Ability to detect/measure	32	Adaptability/Versatility
6	35	Reliability/Robustness	34	Controllability
7	44	Productivity	35	Reliability/Robustness
8	35	Reliability/Robustness	41	Manufacturability
9	41	Manufacturability	35	Reliability/Robustness
10	50	Measurement precision	45	System complexity
11	49	Ability to detect/measure	41	Manufacturability
12	41	Manufacturability	35	Reliability/Robustness
13	50	Measurement precision	32	Adaptability/Versatility
14	50	Measurement precision	45	System complexity
15	28	Loss of information	5	Area of moving object
16	50	Measurement precision	41	Manufacturability
17	43	Automation	46	Control complexity
18	35	Reliability/Robustness	43	Automation
19	41	Manufacturability	50	Measurement precision
20	49	Ability to detect/measure	35	Reliability/Robustness
21	43	Automation	36	Reparability
22	41	Manufacturability	35	Reliability/Robustness

Table 3.3: Inventive principles used

Patent Item	TRIZ Matrix (Mann's version) Recommended Inventive Principles	Inventive Principle Inventor Used	
		Parameter	Name of Inventive Principle
1	12, 28, 23, 7, 35, 31	23	Feedback
2	12, 28, 23, 7, 35, 31	28	Mechanics Substitution / Another Sense
3	12, 28, 23, 7, 35, 31	28	Mechanics Substitution / Another Sense
4	12, 28, 23, 7, 35, 31	28	Mechanics Substitution / Another Sense
5	1, 26, 13, 23, 35, 15, 19	13	The Other Way Round
6	28, 1, 40, 29, 3, 19, 13	28	Mechanics Substitution / Another Sense
7	3, 1, 35, 10, 14, 24, 39, 9	3	Local Quality
8	28, 10, 35, 4, 40	28	Mechanics Substitution / Another Sense
9	2, 3, 35, 9, 28, 27, 33	28	Mechanics Substitution / Another Sense
10	3, 35. 10, 27, 1, 13, 28, 26	35	Parameter Change
11	5, 28, 37, 13, 2, 29, 11, 24	28	Mechanics Substitution / Another Sense
12	2, 3, 35, 9, 28, 27, 33	28	Mechanics Substitution / Another Sense
13	35, 2, 19, 13, 24, 6, 1	2	Separation / Taking out / Extraction
14	3, 35, 10, 27, 1, 13, 28, 26	1	Segmentation
15	28, 26, 17, 25, 30, 16, 1, 37, 7	17	Dimensionality change / Transition into a new dimension
16	3, 25, 28, 13, 35, 24, 1	35	Parameter Change
17	28, 3, 4, 17, 37, 10	17	Dimensionality change / Transition into a new dimension
18	35, 10, 1, 13, 28, 17, 27	17	Dimensionality change / Transition into a new dimension
19	35, 12, 1, 6, 28, 15, 13, 24, 29	35	Parameter Change
20	28, 40, 26, 1, 35, 2, 8, 10	17	Dimensionality change / Transition into a new dimension
21	13, 35, 4, 2, 28, 37, 17	28	Mechanics Substitution / Another Sense
22	2, 3, 35, 9, 28, 27, 33	28	Mechanics Substitution / Another Sense

From Table 3.3, it is observed that the most used inventive principle is “Mechanics Substitution/Another Sense”, 11 out of 22. Use of treadles and identification items were the most common classification methods in the past. A treadle can count the number of vehicles’ axles if the vehicles pass on it. The sensor will sense the force acting on it and then determine how many axles pass across, by analysing the duration of the acting force and the time difference between the separated acting forces. Further, it can also predict the type of vehicles. However, a treadle is not a robust sensor as the data collected by it cannot conclude a high accuracy classification result. In addition, its life cycle is not long. In order to deal with the limitation of the treadle, Patents 2 and 8 suggest changing the treadle to other sensors in different fields, i.e., measure vehicles’ dimensions by using ultrasound and count the number of axles by TOF laser range-finder. Patent 11 has a similar case. Rather than changing the treadle to another field of sensor, the inventors propose adding a capacitive sensor and together work with the existing treadle system, which can utilize the existing assets in the aspect of detection. The same problem also occurs in Patent 21’s piezoelectric-sensor. A non-contact sound sensor is thereby preferred by the inventors.

Some sensors are used in prior methods, but the reliability is not acceptable. In Patent 3, the prior method uses multiple sensors, but if one of them fails, especially in detecting the content of license plates, the system mostly fails in classification. Therefore, the inventor proposes using an optical sensor as it has a larger fault tolerance. Another prior method mentioned in Patent 4 uses a seismic sensor, but its performance is poor since distinction cannot be made between specific vehicle types and a generic type. Hence, a laser sensor is suggested by the inventors.

Besides, some classifications require identification items be installed in vehicles, like an RFID transponder or barcode, which contains some information about that vehicle. A scanning device can check this information as vehicles pass across it. However, as described in Patents 6, 9 and 12, installation of any affixed items in vehicles induces very high costs and

low operability. It may technically not be possible to install affixed items on every vehicle. Hence, the patents proposed changing the field of sense into others, i.e., radio-wave field from radar, the magnetic field from inductive loop, and image-based camera.

Another approach proposed by Patent 22 employs GPS tracking of vehicles. A vehicle classification cloud system can receive the travelling record of different vehicles and then predict their type by analysing the travelling data. This method can greatly reduce the cost with respect to system installation and hardware, as it does not need physical sensors along the traffic road. In addition, instead of active sensors, which emit a pulse to the objects to be detected and reflect back to the sensors, a passive audio sensor is preferred since it can reduce the effect of environmental noise as proposed in Patent 5.

As described above, there are many different fields of sensors that can be used in classifying vehicles, and it comes down to an idea of creating a classification system by combining various outputs of sensors. However, in Patent 7, the inventors stated that it might not be a practical idea to gather all kinds of data in some application or environment. In this situation, only a few data are valid and the system may not characterize the vehicle with sufficient specificity to classify it. Therefore, the inventors applied the “Local Quality” inventive principle. A new algorithm is proposed to enhance the analysis of the data from an acoustic sensor(s) and to enable it to function in locally optimized conditions. This suggestion can decrease the handling of too many kinds of data.

On the other hand, too few sensors may result in low accuracy or other limitation. In Patent 13, the classification process is only based on measuring dimensions of vehicles by the camera. If the distance between the camera and the different vehicles varied too much, the measuring dimension will differ a lot. Therefore, the “Separation/Taking out/Extraction” inventive principle is applied in this case, so that the function of dimension measuring and distance measuring are separated.

In other patents, other inventive principles are used. In Patent 1, the proposed method adds a barrier to separate the vehicles passing through the lane one by one to prevent the problem of overlapping. This involves the “Feedback” inventive principle. In Patent 14, it is mentioned that the prior arts have used laser devices together with thermal imaging cameras or TOF cameras. However, the implementation is neither simple nor cost-effective. The inventors thereby proposed to capture the images separately by using multiple common cameras from birds-eye view and three-quarter front view. This incorporates the “Segmentation” inventive principle.

Sometimes the raw data is changed into another form for further analysis, which employs the “Parameter Change” inventive principle. In Patents 10 and 16, multiple sensors may not be applicable due to system complexity or environmental issue, and hence designing a valid classification system with only one sensor is preferred. By generating a signal waveform or obtaining a digitized signal from the electromagnetic signal which is then digitized, sequenced, and time-stamped, the collected data can be much more useful when compared with a set of raw signals. Patent 19 suggested using a linear classifier with more complex features rather than the high-level features or raw profiles such as width and height. Fisher vectors, extracted from “patches” of the profile, and profile feature, obtained by computing the integrity of the profile, can be used to improve the scope and efficiency of the system with laser scanners.

An overlapping problem is considered in Patent 15, as the image of the vehicle will be obstructed by other vehicles and the influence level will be based on the size of vehicles. The proposed solution captures the images of vehicles from the rear-view instead of the side-view by employing the “Dimensionality change/Transition into a new dimension” inventive principle. A similar situation is mentioned in Patent 17. It states that if the classification system uses a camera to capture images automatically, it is hard to prevent the problem of occlusion, tracking a moving object, shadows, rotation, and lack of colour invariance, and hence it proposes a solution changing the capturing direction to rear-view. The same inventive principle

is also used in Patent 18. As classification by length is not sufficient, the inventors proposed to classify the vehicles by 3D instead of 2D.

There is one case which cannot be predicted by both the classical and the new matrices. In Patent 20, the inventors point out that the classification of the 2D methods can only make a coarse utilization of the geometric relations and is ill suited. In this scenario, 3D models become indispensable. However, 3D methods also have their technical problems such as faulty edge detection, lack of scene contrast, blurry imagery, scene clutter, and noise. In this case, the inventors suggest a 3D method in aerial imagery instead of ground-level imagery, which also belongs to the "Dimensionality change/Transition into a new dimension" inventive principle.

In addition to recommending inventive principles, TRIZ also predicts technical developments and possible patterns of the products by using evolution trend. It can provide hints and innovation trends for further development of existing products or systems. There are 8 evolution trends and 35 evolution trends described by classical TRIZ (Altshuller, Shulyak & Rodman, 2005) and Mann's revised TRIZ (Mann, 2003) respectively. As found in the previous section, the classical TRIZ may no longer be applicable in technology nowadays. Therefore, the evolutionary trend of Mann's revised TRIZ is used in the vehicle classification system forecasting.

However, not all the evolutionary trends are suitable for analysing all the systems, as each system will have its own evolutionary trends. In the vehicle classification system, 8 evolution trends from Mann's revised TRIZ are identified to be more applicable. They are "Geometric Evolution (Lin)", "Mono-Bi-Poly (Various) – Interface", "Sense interaction", "Design point", "Controllability", "Human Involvement", "Market Evolution", and "Customer Purchase Focus". They are examined with reference to the 22 selected patents as in the follow section.

1) Sense interaction

Trend: 1 sense → 2 senses → 3 senses → 4 senses → 5 senses

From the investigated patents, it can be observed that there are many proposed sensors in the vehicle classification system. For example, force sensors like treadle and capacitive sensor; electromagnetic wave sensors like camera, laser sensor, radar and Infra-red sensor; magnetic sensors like geomagnetic sensor and inductive loop sensor; sound sensors like audio sensor, ultrasound sensor and acoustic sensor. Most of the patents are designed to have one or two sensors in their system only due to the system complexity issue. In the future, if the computing power of the system is improved, more sensors with different sense fields will be used, as more sense fields will result in more information. In addition, a higher computing power can process more data, and hence the performance of the system can be improved.

2) Geometric Evolution (Lin)

Trend: Point → 1D Line → 2D Plane → 3D Surface

The detecting geometry is changed from “Point” to “3D Surface”. Treadle is a point sensor, as it can only sense the force at a point. Laser sensor can measure the length and height of vehicles. In recent technology, 2D models of vehicles can be captured in detail by a camera. Together with a camera and the aid of other sensors, or other 3D sensors, a 3D model of vehicles can be obtained or generated. Therefore, the evolutionary trend of “Geometric Evolution (Lin)” is at the end stage, and the future invention should be focusing on 3D model development.

3) Mono-Bi-Poly (Various) – Interface

Trend: Mono-System → Bi-System → Tri-System → Poly-System

The systems in the patents are focused on vehicle classification, as its performance still has to improve. In fact, other useful systems can be combined to give an all-round smart system. Several road-related functions can be added into the system for better road management, such as detecting the speed of vehicles, measuring vehicles' weight, and capturing the number of the license plate.

4) Design point

Trend: Single Operating Point → Two Operating Point → Many Operating Point → Continuous Re-optimization

The operating point is dependent on the type of sensors. For a camera, it can detect different points of interest (POI), which is at the final stage of "Design point" evolution trend. However, for some sensors such as magnetic sensors or force sensor, it usually operates at a single point to detect one vehicle. Therefore, the status of this evolution trend varies with different classification systems due to the use of different sensors. The development trend should be similar to the camera's trend, as it can operate across a range of area and find a POI for classification.

5) Controllability

Trend: No feedback → Addition of feedback → System with learning/adaptive → Feed-forward → Self-adaptive, autonomous system

The feedback function can improve the performance of the system. The status is generally at the second "Addition of feedback" stage. Some of the classification systems try to add some control components to prevent overlapping issues. In the future, it is expected to have

more learning function or self-adaptive functions such that they can identify some unknown cases to prevent fault detection.

6) Human Involvement

Trend: Human → Human + Tools → Human + Powered Tools → Human + Semi-Automated Tools → Human + Automated Tools → Automated Tools

The main object of using a vehicle classification system is to perform the vehicle classification function automatically. Based on the above patents, the status is at the "Human + Automated Tools" stage or even tends to the "Automated Tools" stage. The main problem with the vehicle classification system not being fully automated relates to the robustness. This is because system faults may cause troubles like money disputes or traffic safety issues, and hence human will still need to partially be involved in controlling the system. In the future, if the technology develops further, this system should be fully automated.

7) Market Evolution

Trend: Commodity → Product → Service → Experience → Transformation

The status of the current classification system is between "product" and "service". Most of the inventors or patent owners most likely sell their system just for a vehicle classification function. Only few of them sell it with complete services such as toll charging and traffic management. It can be foreseen that if the technology develops further, this system could be sold in packages, or as a united system.

8) Customer Purchase Focus

Trend: Performance → Reliability → Convenience → Price

The main purpose of research and invention in most of the patents is still focused on performance and reliability, to function more or less on the same level as humans. In the future, the customer will demand the system to be more convenient to use, for instance, simpler implementation and more user-friendly control interface, and ultimately, lowest prices.

A radar plot summary of the above trends is depicted in Figure 3.1. The percentage of unfilled area shows the estimated potential for improvement for vehicle classification up to 2020, which is 72.92%.

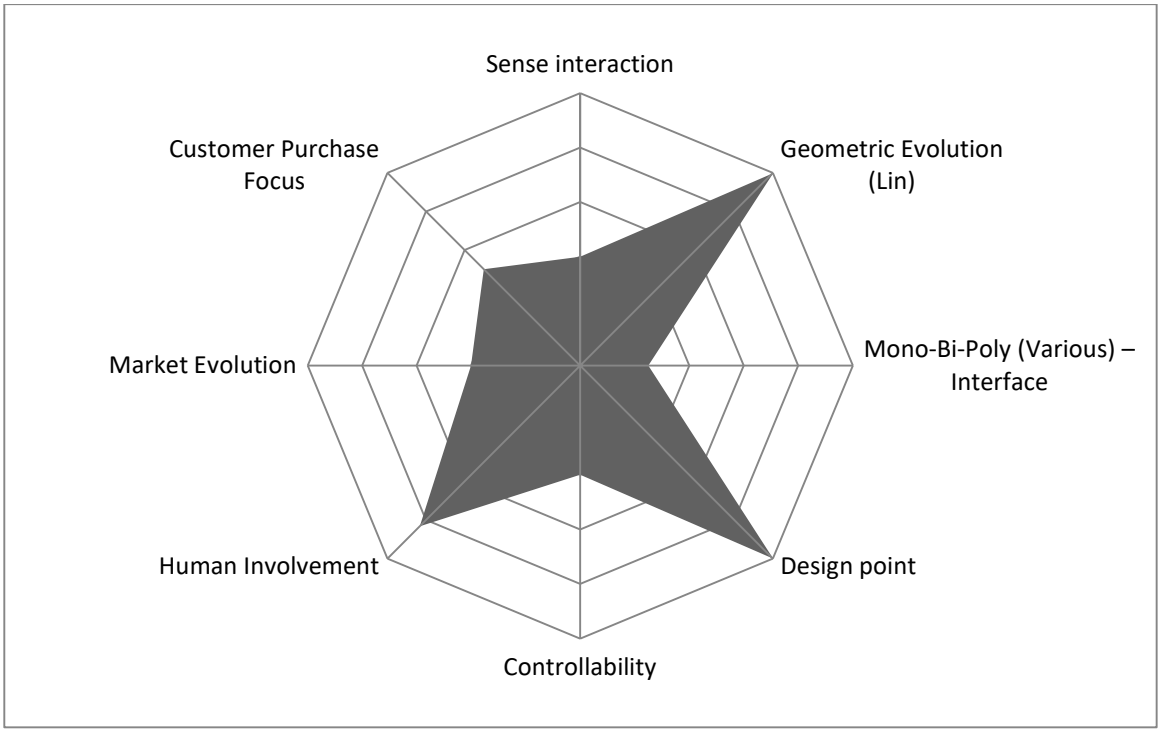


Figure 3.1:1 Evolution potential of vehicle classification

In those patents, the vehicle classification system focuses not only on the degree of automation, but also insists on its robustness, accuracy, manufacturability, production cost, environment adaptability and/or diversity of classification types. However, each invention may have its limitation, and hence, it may not be possible to apply only one particular system in all

applications.

In the past, vehicle classification was processed manually. Sensor technology was not developed at that time. It could not capture enough vehicles' characteristics, and hence it was not possible to build an applicable automatic vehicle classification system due to accuracy and robustness issues. Beginning from around the 1950s, the Third Industrial Revolution and the Fourth Industrial Revolution occurred within the engineering industry. One of the main development trends was to accelerate automation by using the latest electronic and information technology, to make a complex and flexible system involving automation technology and many other areas. The development of automated vehicle classifications started during this period. From Table 3.1, it can be observed that most of the patents in the latter half of the 20th century are focused on improving the degree of system automation, and aimed at reducing reliance on humans, however, provide an acceptable performance. After the automation developed relatively, the improving factors were changed to accuracy, robustness, or manufacturability in recent years. In order to have a wild range of applications, a feasible vehicle classification system should have high classification accuracy and should be adaptable in various environments such as rainy days and night-time, as these are the key factors that determine the performance of the whole system. Another vital factor is the number of categories that the system can identify. There are a large number of vehicle types in the world, where different countries may have different types of vehicles. The classification category of the system depends on its application Some applications may have to classify the type of the passing vehicles, while some applications may only need to classify the size of vehicles.

In order to achieve the above goals, some modifications to the existing systems are expected. From Table 3.2, the inventive principle of "Mechanics Substitution/ Another Sense" is the most famous method. As sensor technology has greatly improved recently, people should consider the type of sensor which was not suitable in the past. Inventors can also consider other fields of sense to profile the vehicle, such as light field, magnetic field, and microwave field.

While these fields can provide more information, they are not limited to only one field. Multiple sensors across different fields can make the classification system more adaptable and robust in different environments. If the cost of each sensor is not high, multiple sensors can be a possible solution in the future with advances in vehicle classification systems. This also agrees with the evolution trend proposed by TRIZ.

Therefore, this research is trying to design a novel profiling system with multiple sensors that can perform a robust classification function. Multiple sensors can increase the reliability of a recognition system. Each of the sensors must have its own pros and cons. Combining different sensors into one system can maximize the pros and compensate for the advantages of each sensor. With more data collected by the system, the accuracy of a recognition system in different environments can be enhanced. It is just like a human being senses the environment with six senses instead of only one sense.

3.2 Proposed vehicle profiling system

In fact, nowadays, object detection systems or profiling systems insist on the computer vision technique with the deep learning algorithm. This research is trying to lighten the role of deep learning while remedying the limitation by using other sensors, i.e. geomagnetic sensor and LiDAR. The proposed profiling system is working with sensor-bases object detection to see if it can increase the efficiency and stability of the profiling system. Geomagnetic sensor, camera and LiDAR are the main devices used for acquiring data to perform a real-time vehicle profiling system.

3.3 Hardware equipment

The main objective of the hardware aspect of this work of research is cost-effective and simple installation, so the system has enough market potential.

There are three sensors in this system, i.e., camera, LiDAR and geomagnetic sensor, as well as a single board computer (SBC) to receive and compute raw data from all the sensors.

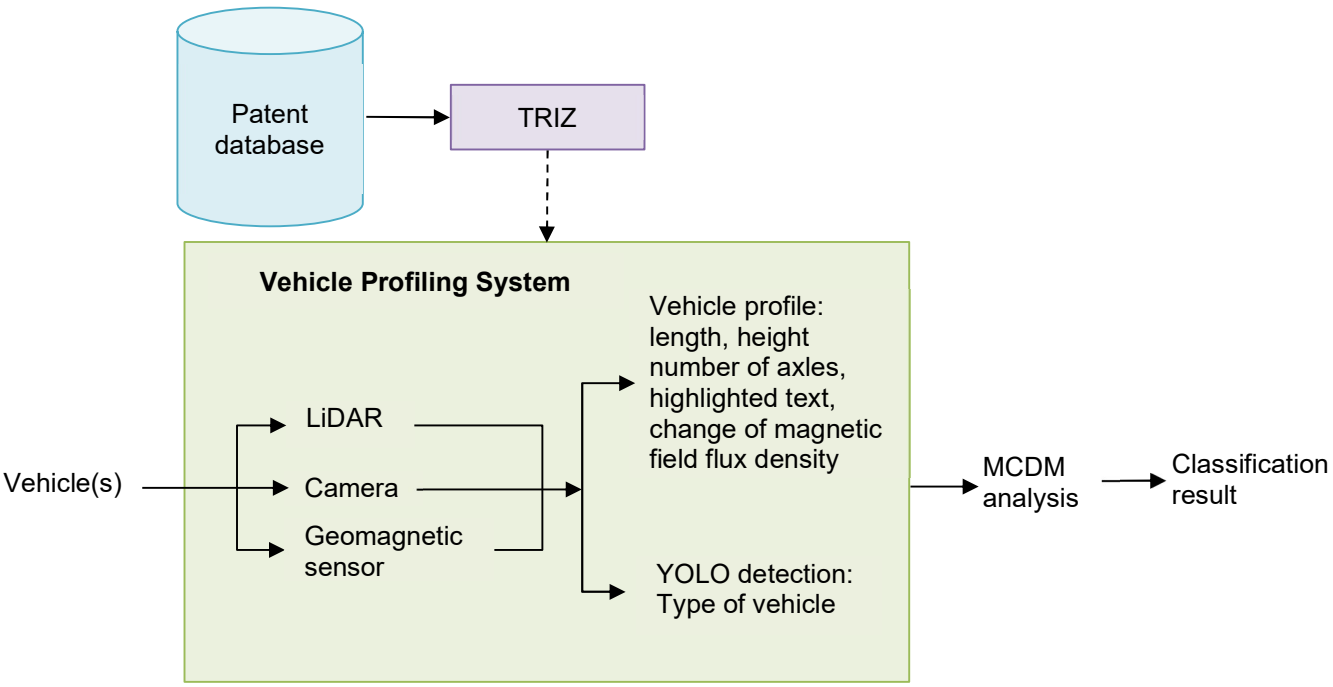


Figure 3.2: The proposed vehicle profiling system

As shown in Figure 3.2, all the three sensors, i.e. LiDAR, camera and geomagnetic sensor, are connected to an SBC either with a wire or wirelessly. Each sensors have its own task. The LiDAR can measure the distance from the vehicle to itself by calculating the time of flight of the light pulse, and together with the camera information, the dimension of the vehicle can be deduced. The geomagnetic sensor measures the magnetic flux density all the time; if a vehicle is passing across it, the magnetic field flux density will greatly change according to the size of the vehicle. The image captured by the camera can be used to directly identify the number of axles and text on the vehicle's side body. With the above profiling, a classification result can be generated. In addition, by using YOLO, another classification result can also be generated by different routines. These two results are further processed with the help of

multiple-criteria decision-making (MCDM) to attain a final classification result.

3.3.1 Speed control

One of the main concerns in profiling is the movement of the vehicle. For instance, all the sensors fail to detect or measure a fast-moving object, the geomagnetic sensor fails to pick up on the change in the magnetic field due to the limited sampling time, the camera fails to capture a clear picture due to low frame per second and the LiDAR fails to collect valid data due to distortion. In order to ensure valid object measurement and capturing, the most robust method is to limit the speed of the vehicle while it is passing through the profiling system. There are two possible ways in which the driver can be advised to slow down the speed of their vehicle. The first is to place a speed limit sign along the roadside to let the driver know that he should slow down his vehicle. Figure 3.3 shows the speed limit sign in Hong Kong, and the number inside the sign indicates the highest speed the vehicle can take in km per second. This sign should be placed near the beginning of the road so that the speed can decrease to a certain value before the vehicle enters the profiling system. However, this method cannot ensure that all drivers would comply with the speed restriction. The second method concerns installing a speed bump on the lane. It is a physical equipment that uses vertical deflection to slow vehicle traffic. If the driver does not slow down his vehicle, he will feel a huge damping and damaging to his vehicle. These penalties can ensure that drivers move their vehicles slowly when they are passing across.

Figure 3.3: Speed limit sign along the road

In this work of research, there is a geomagnetic sensor that needs to be placed under a passing vehicle. In order to reduce the manufacturing cost, it is proposed that the sensor is placed inside a speed bump rather than digging into the ground. The speed bump should have enough space to house the geomagnetic sensors. Moreover, the speed bump should not be made of metal, in order to prevent the blocking effect of the wireless signal emitted or received from the sensors. Figure 3.4 shows one possible speed bump made of polyvinyl chloride (PVC). The height of the slot inside is 20mm, which is enough to accommodate the geomagnetic sensors and the battery.


Figure 3.4: Speed bump that can store some sensors or cables

3.3.2 Geomagnetic senor

As mentioned in the previous sections, the function of geomagnetic sensors is to detect the existence of a vehicle and to classify the vehicle based on its size. Therefore, there are three purposes for using geomagnetic sensors in this study. First, it will act as an event-trigger device for the profiling system to indicate when the vehicle is presented. The geomagnetic sensors will obtain a ground value, which is in the situation of no vehicles passing, after starting the profiling system. The geomagnetic sensors will continuously measure the magnetic field for every single period. If a change in the magnetic field is larger than the threshold value, it indicates that a vehicle is approaching the profiling system; Secondly, after the event-trigger, the geomagnetic sensors will record the value of the magnetic field until it drops lower than the threshold value again. The recorded value will then be sent to SBC for further data analysis. To enlarge the observability of the change of magnetic flux density, the magnetic field index (MFI) will be introduced to compare the measuring result directly.

$$\left\|\vec{B}\right\| = \sqrt{B_x^2 + B_y^2 + B_z^2} \tag{3.1}$$

$$B_{dist} = \left\|\vec{B}_t - \vec{B}_0\right\| \tag{3.2}$$

$$MFI = \left(\frac{B_{dist}}{S}\right)^2 \tag{3.3}$$

where $\vec{B}$ is magnetic flux density, t implies the time for measuring, 0 indicates the ground value, and S is the sensitivity coefficient. Finally, the profiling system will define a period having the highest MFI measured, and this indicates a valid interval where further data analysis will focus on this period.

3.3.3 Camera

The profiling system needs to profile the whole shape of the vehicle in the side view. In this case, multiple camera setups with key images stitching may be necessary. Images are captured by many cameras at almost the same time, and the images are rectified by rotation,

translation, cutting edge and resizing to a certain resolution. However, this method causes a high installation cost and complicates data processing. The latter may also increase the error rate. To decrease the processing time and minimise the error rate, a single set of sensors to detect one vehicle is preferred over multiple sets, as it can reduce pre-processing work, such as reforming the whole profile.

Together with geomagnetic sensors recording the magnetic value of the vehicle, the camera will record a video or continuous frames while the vehicle passes through the profiling system. These videos should contain certain information about the vehicle, i.e., text printed on the vehicle's body, number of axles, and appearance of the body. The camera should be high revolution to capture the detail of the vehicles, and it should be at least Full HD: 1080p image resolution (1,920×1,080 pixels).

3.3.4 LiDAR

A LiDAR is used to measure the depth of the vehicle from LiDAR itself, so the dimension of the vehicle can be calculated. As the sensor part of the profiling system is supposed to install on the side of the lane, only the vehicle's length and height remain as points of concern. The dimension along the vehicle's body is usually not identical; take a taxi for example.

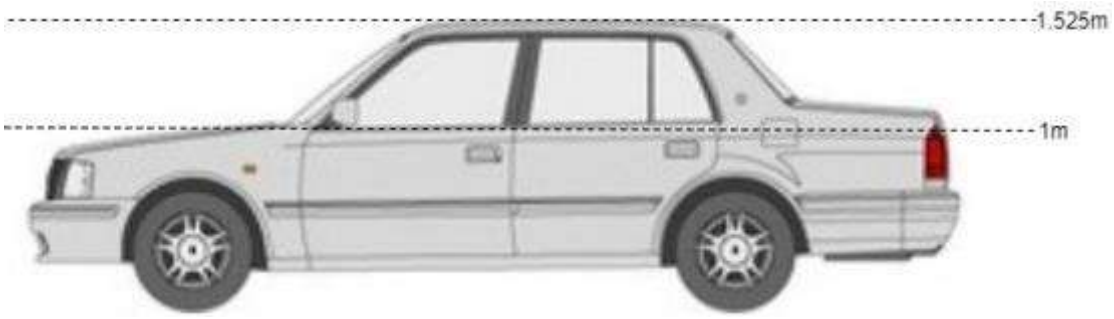


Figure 3.5: A taxi with a height level mark

From Figure 3.5, it can be observed that the length between 1m height below and 1m height above are much different, but the valid length or height is defined as the longest part of the vehicle. On the other hand, the variation in length along the vehicle body may be one of the classification features for some types. Further modification of the algorithm in this study can investigate this aspect in the future. Moreover, the flatness and the material of the object surface are two of the factors that affect the performance of LiDAR. If the surface is made of glass, i.e., windows, the LiDAR light pulse may penetrate through it and no light pulse will be reflected to the LiDAR. Hence, fault detection occurs.

Nowadays, there are mainly two Beam-steering technology methods for 3D LiDAR, which are the spinning LiDAR and the mechanical scanning LiDAR. When in operation, part of the spinning LiDAR is rotating so that the light pulse can be emitted to different directions. Hence, it has the advantage of 360° coverage, but its price is higher. Mechanical scanning LiDAR uses a mirror to redirect a single laser to different directions. The mirror is driven by a unique system called micro-electro-mechanical system (MEMS). It usually has a limited FoV, but its price is lower than the spinning LiDAR. In this project, a mechanical scanning LiDAR is more suitable, as the scanning rage does not need to be 360° and the cost is one important consideration of the system.

3.3.5 Computing apparatus

A computing apparatus is needed for deep learning training and system operation. The hardware requirement for the training purpose is very high. Normally, it requires a powerful GPU instead of using CPU for training; otherwise, the training process is too slow. Therefore, the choice of GPU affects the performance and efficiency of training. Without GPU, the whole deep learning may take several months to process. A suitable GPU can compress the duration into several days, hours or even minutes. Based on the latest information available in 2020, the acceptable GPU with the lowest specification is GTX 1050 Ti (4GB) or the online GPU Colab

from Google. On the other hand, if there is a trained data set already for use, then the GPU and the training process is unnecessary, unless the user requires a better performance.

For the operating computer, a single-board computer (SBC) is recommended. An SBC is a complete computer built on a single circuit board. Its size is smaller than a desktop computer, and it can hence be installed quite simply.

3.4 Sensor's position

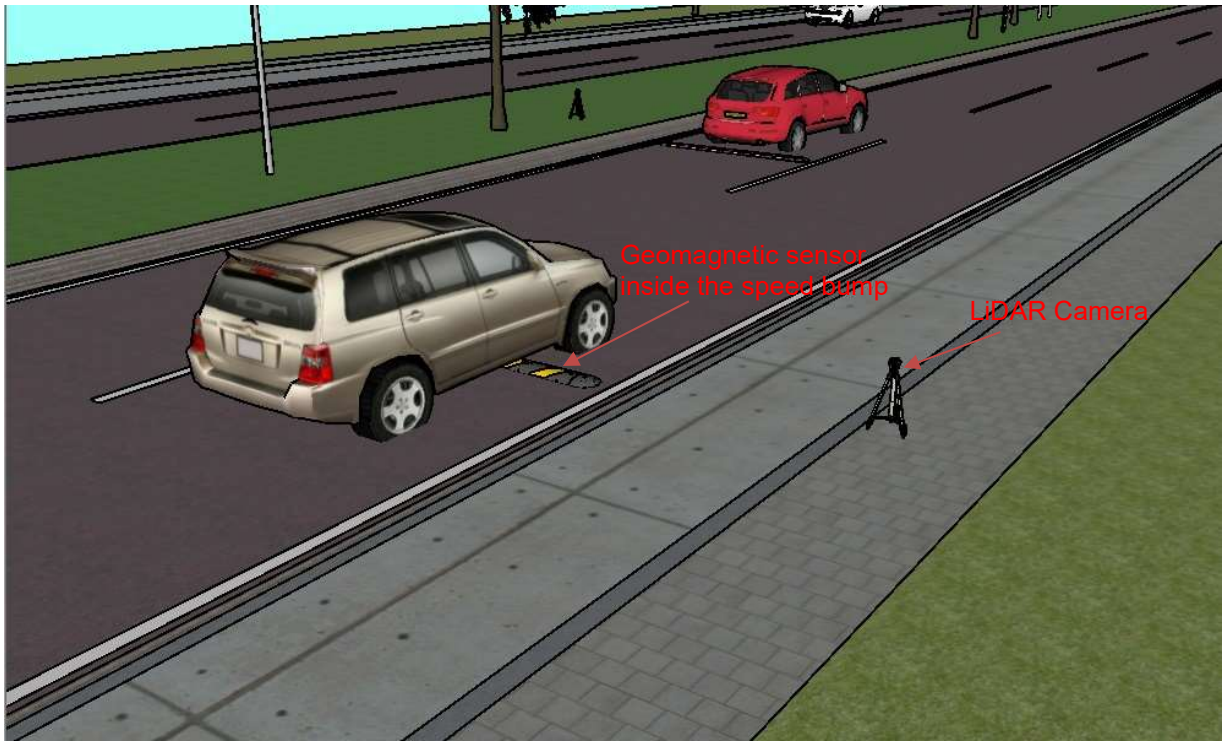


Figure 3.6: Illustration of the proposed profiling system

The proposed profiling system is configured as shown in Figure 2.6. The sensors should be placed in suitable positions so that the profiling system can supervise the whole target area and fully profile any vehicles. The faulty position of sensors will result in the failure of the entire profiling system. In fact, all sensors have their limited field of view (FoV) or detecting range; hence, the design sensor's position should be considered the limitation for all the sensors used.

There are three camera classes based on the FoV, which is classified by the lens. The first is normal lens, where the FoV is about 50-55° and it is nearly the same as the FoV of human eyes; therefore, the image taken with this lens is similar to human vision. The second is telephoto lens, where the FoV is usually less than 30°. This lens is mainly used in long distance observation, as it has a larger depth of field. The third one is a wide-angle lens, where the FoV is usually over 60°. To capture a large object or a broad area, a wide-angle camera is more suitable than the others. Since the size of vehicles is relatively big, this kind of lens is suitable for vehicle profiling.

Similar to the camera, LiDAR also has its own angle of view. Commonly, there are two types of LiDAR, depending on their usage. The first one is an all-around LiDAR. Its FoV is 360° and it can scan in all directions; hence, it can detect its surroundings. This type of LiDAR is mainly used in autonomous vehicles to check if there are any obstacles blocking the travel path. The second one is similar to a camera and it has a limited FoV. This type of LiDAR is suitable for detecting a specific area in detail.

The camera and the LiDAR will be installed closer to each other in order to minimise the installation space. It will be assumed that they are close enough so that the origin of their FoV is the same, and hence, the system FoV will be dominated by the narrow one between the camera and the LiDAR. For better performance, the camera FoV and LiDAR FoV should be similar, as it can maximise the information obtained from both. Figure 3.7 and 3.8 are the top-view and side-view of the sensor with vehicle and lane.

$$\mathrm{FoV}_{\mathrm{system}} = \min(\mathrm{FoV}_{\mathrm{camera}}, \mathrm{FoV}_{\mathrm{LiDAR}}) \tag{3.4}$$

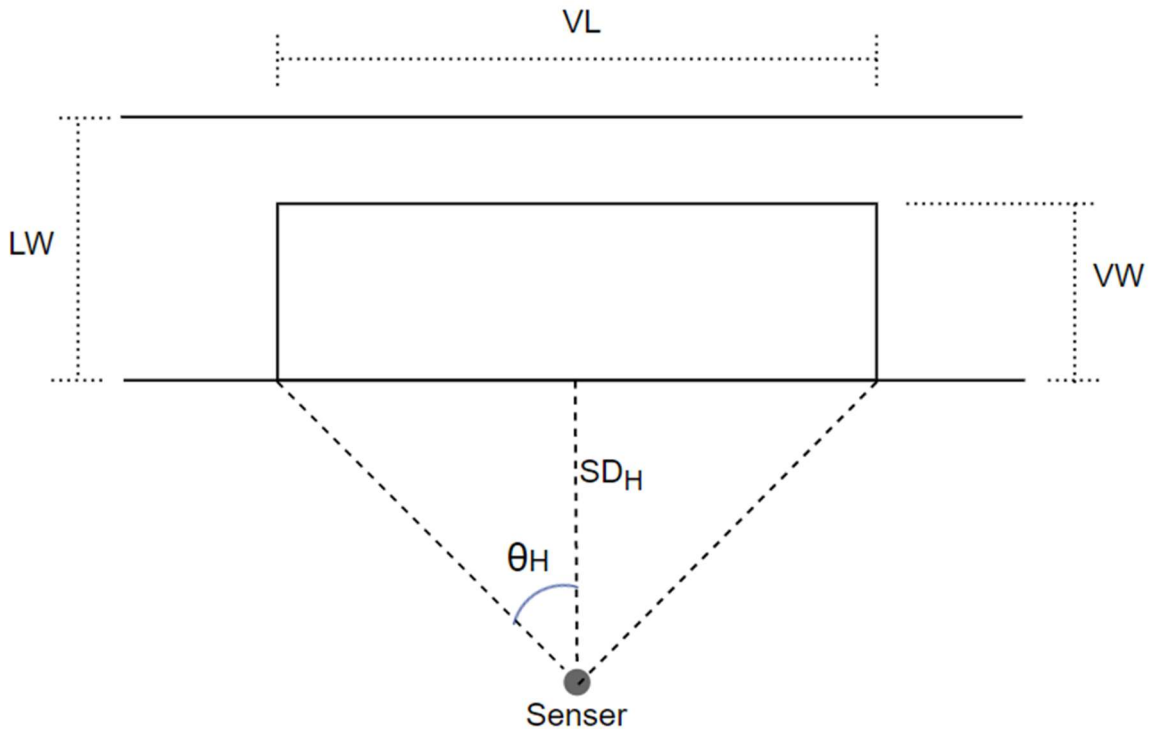


Figure 3.7: Top-view of a sensor with expected horizontal FoV.

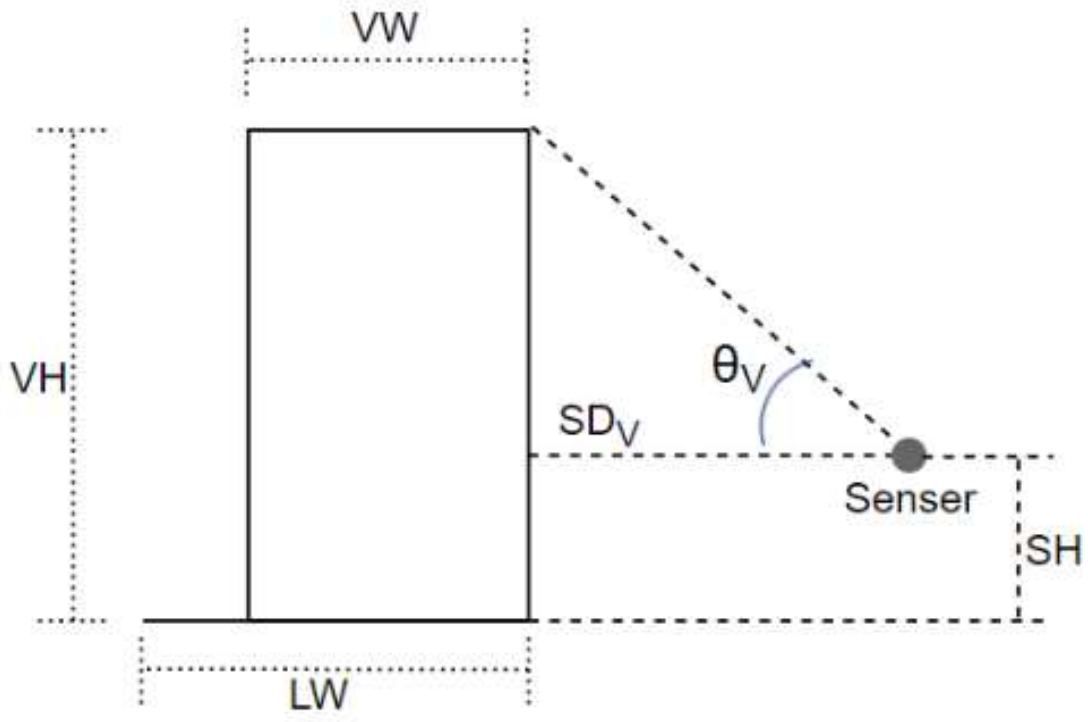


Figure 3.8: Side-view of a sensor with expected vertical FoV

VL: Vehicle length

VW: Vehicle width

VH: Vehicle height

LW: Lane width

SH: Sensor height

SD_H: Distance between sensor and the near rim of the lane for horizontal

SD_V: Distance between sensor and the near rim of the lane for vertical

θ_H: Half of horizontal field of view (HFoV)

θ_V: Half of vertical field of view (VFoV)

$$\tan(\theta_H) = \frac{VL/_2}{SD_H} \tag{3.5}$$

$$\tan(\theta_V) = \frac{VH-SH}{SD_V} \tag{3.6}$$

Table 3.4: Cap 374A Road Traffic (Construction and Maintenance of Vehicles) Regulations, FIRST SCHEDULE OVERALL DIMENSIONS OF VEHICLES

Vehicle		Overall Length	Overall Width	Overall Height
Private Car		6.3 metres	2.3 metres	2.0 metres
Taxi		6.3 metres	2.3 metres	2.0 metres
Invalid Carriage		6.3 metres	2.3 metres	2.0 metres
Light Bus		7.5 metres	2.3 metres	3.0 metres
Bus	Single-decked	12.0 metres	2.5 metres	3.5 metres
	Double-decked	12.0 metres	2.5 metres	4.6 metres
	Articulated	15.0 metres	2.5 metres	3.5 metres
Light Goods Vehicle		10.0 metres	2.5 metres	3.5 metres
Medium Goods Vehicle		11.0 metres	2.5 metres	4.6 metres
Heavy Goods Vehicle	Rigid	11.0 metres	2.5 metres	4.6 metres
	Articulated	16.0 metres	2.5 metres	4.6 metres
Special Purpose Vehicle		12.0 metres	2.5 metres	4.6 metres
Tricycle		-	1.1 metres	-
Trailer		13.5 metres	2.5 metres	4.6 metres
Pedestrian-controlled Vehicle		4.3 metres	1.6 metres	-

The general size limitation of vehicles in Hong Kong is mentioned in Table 3.4. Considering the case for the largest measuring one for most cases, excluding the one that is articulated, the length and height are 12m and 4.6m, respectively. Assume that the sensors are placed at 1m above the ground and at most 2m from the near rim of the lane, so that the camera can capture a good quality image. On substituting the above parameter into equation 3.5 and 3.6, the half of HFoV and VFoV are 71.6° and 60.1° respectively.

3.5 Supervising area

Besides the function of vehicle classification, another possible application is to supervise all the vehicles on the traffic road, and hence the total number can be identified, together with the types of vehicles. Therefore, there is a need for deciding the number of sensors used to cover a certain supervising area. This question is similar to the art gallery problem, which is a kind of visibility problem. An art gallery is supposed to find the minimum number of guards needed for observing its entirety. Specifically, this question is one involving computational geometry. Given a simple polygon with n vertices or sides, how many points that can look around 360° in the polygon are needed if every point in the polygon is to be covered. According to Chvátal's art gallery theorem (Chvátal, 1975), if the detecting direction of the sensors is 360°, the upper bound on the minimal number of points is n/3, where n is the number of vertices of the polygon.

If the above case is changing the guard to sensors used in a profiling system, the number of sensors required for monitoring a whole polygon area should be n/3 at most, where n is the number of vertices of the detected area. However, one of the main assumptions of the above thermo is that the detecting range of sensors is unlimited and goes in all directions, so that it can travel far away from the source. In reality, as discussed in the section above, all the sensors have their limited FoV and detecting range, and the sensors should be able to profile the whole vehicle, i.e., the dimension or full body picture. Figure 3.9 shows an illustration position of sensors where there is a distance SD between the sensors and the road. It should be noted that the side-view on both right and left are supposed to be identical.

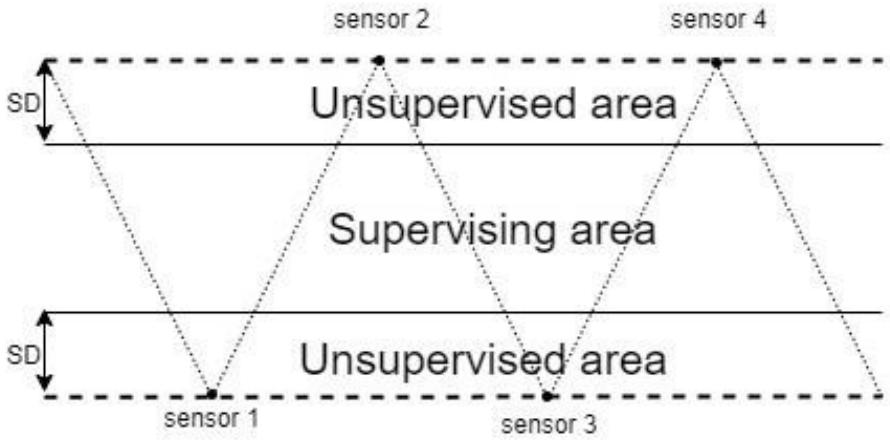


Figure 3.9: Supervising areas of multiple sensors with limited FoV

Like the art gallery problem (Abdelkader et al., 2015), finding the exact number of minimum sensors necessary for supervising the whole polygon even the sensors can be located in the interior of the polygon for any degree of FoV α, where $0<\alpha<360°$, is an NP-hard problem. In fact, the sensors used in this study have a limited FoV, and assume it is $45° <\alpha \leq 180°$. In this case, there is proof that for $\alpha = (180° - \varepsilon)$ where $\varepsilon>0$, at least $2n/3–1$ sensors needed; and for $\alpha = 180°$, at least $n/3$ sensors are enough to monitor (Tóth, 2002). This information is useful as it gives a reference of hints for calculating the manufacturing and installation cost.

3.6 Image processing for camera image

In view of a general design for taking large objects, wide angle lens can be used to capture an image, so that it can increase the flexibility of the system to be applicable in different environments. On the other hand, Barrel distortion is associated with wide-angle lens; so, the image should be rectified before it is passed on for further processing. However, Figure 3.10 depicts that the edge (shown in the red circle) on the left is stretched to the square form (shown in the red circle) on the right in the rectification procedure. This process will decrease the image quality or cause loss of detail on the edge area, as the rectification procedure of the skewed images involves a number of steps related to rotation and translation. A typical example is as shown in Figure 3.11.

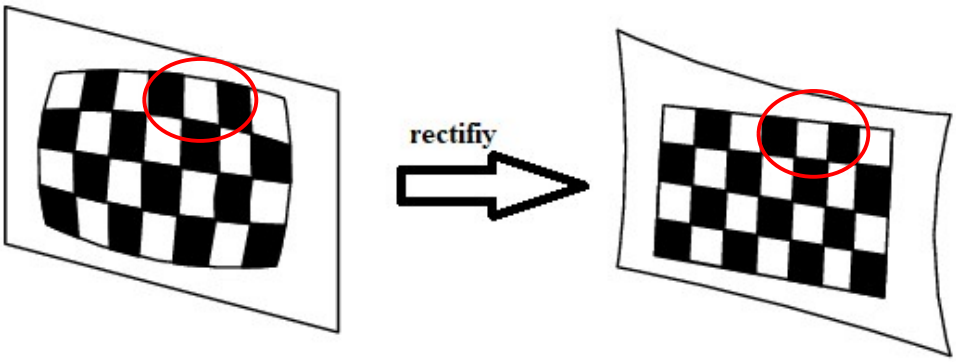


Figure 3.10: Problem of rectification
(Florez, 2010)

Figure 3.11: Before rectification (left) and after rectification (right)
(Hughes et al., 2009)

There are two types of distortion, radial distortion and tangential distortion. The rectification formulas are as follows (Bradski, 2000):

$$x_0 = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \tag{3.7}$$

$$y_0 = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \tag{3.8}$$

$$x_1 = x_0 + (2p_1 xy + p_2(r^2 + 2x^2)) \tag{3.9}$$

$$y_1 = y_0 + (p_1(r^2 + 2y^2) + 2p_2 xy) \tag{3.10}$$

where k_1, k_2, k_3, p_1 and p_2 are the five distortion parameters, and $r^2 = x^2 + y^2$. x, y are the original position of each pixel, while x_0, y_0 are the rectified position from radial distortion and x_1, y_1 are the rectified position from both radial and tangential distortion. As different cameras have their own distortion characteristics, the five distortion parameters given above should be

applied for different cameras, and they can be found by utilizing the calibration method with the pre-rectification position and the post-rectification position of three pixels. In the market nowadays, the output image of some wide-angle cameras has already rectified.

3.7 Dimension measuring and calculation

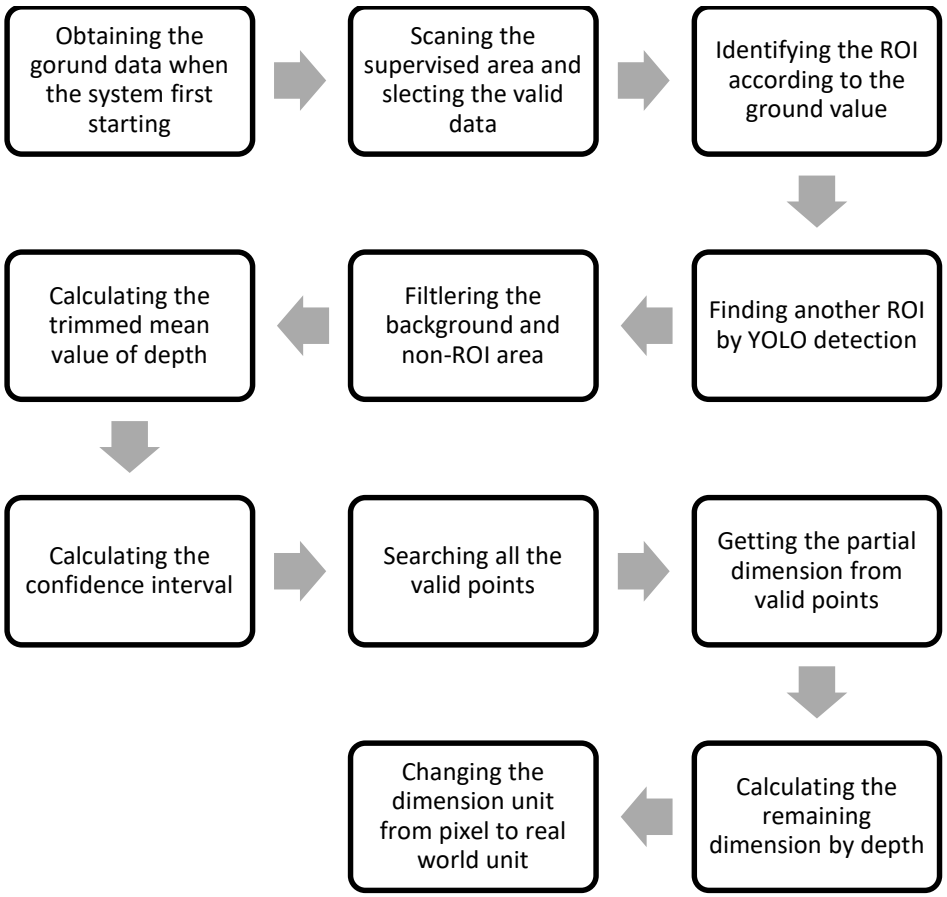


Figure 3.12: Procedure for finding the dimension of the vehicle

Figure 3.12 is the general workflow of dimension determining. A point cloud is collected by the LiDAR. Let P be a point cloud contains n points. With respect to the LiDAR position as the origin, x_n, y_n, and d_n represent the location of point n in x-coordinate and y-coordinate and its depth, respectively.

$$P_n = [p_1\ p_2 \cdots p_n] = \begin{bmatrix} X_n \\ Y_n \\ D_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \\ d_1 & d_2 & \dots & d_n \end{bmatrix} \tag{3.11}$$

where $(x_n, y_n, d_n) \in \mathbb{R}^3$

Actually, d represents the depth of the detected point in its z-coordinate, not to be confused with the distance, as descripted in Figure 3.13.

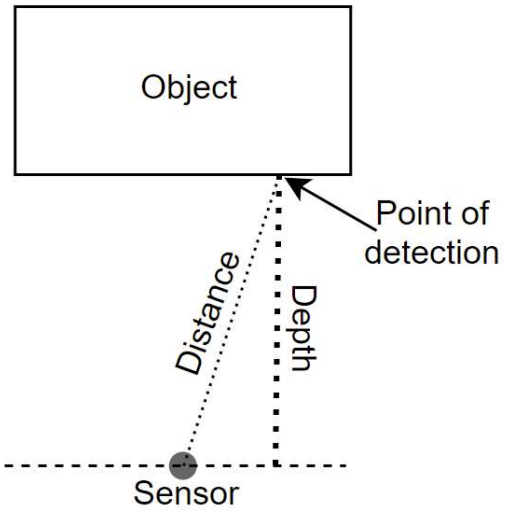


Figure 3.13: Distance and depth illustration on top view

In fact, the LiDAR is going to scan in all directions within its field of view, including the lane shown in Figure 3.14, which will affect the determination of the height of the vehicle. Figure 3.15 shows the depth maps where the lane and the vehicle merge into each other since their values are very close. As a result, it is hard to determine the edge of the vehicle.

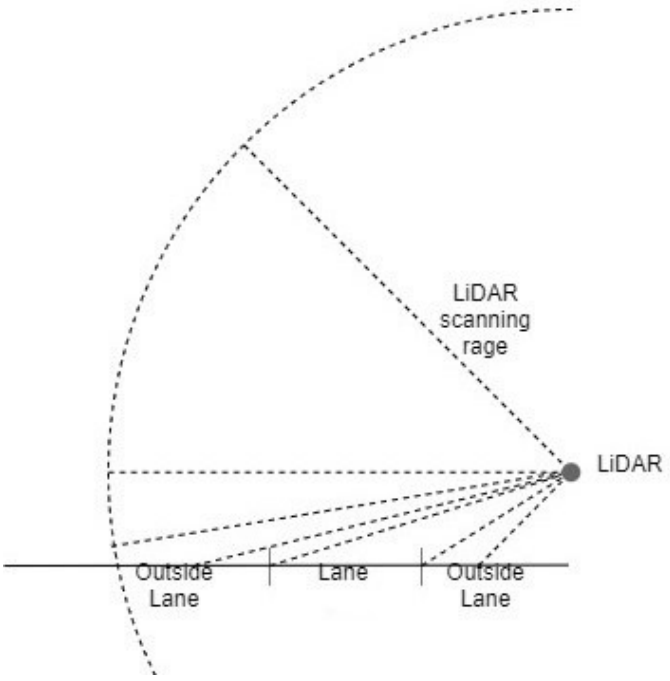


Figure 3.14: The detection range of the LiDAR

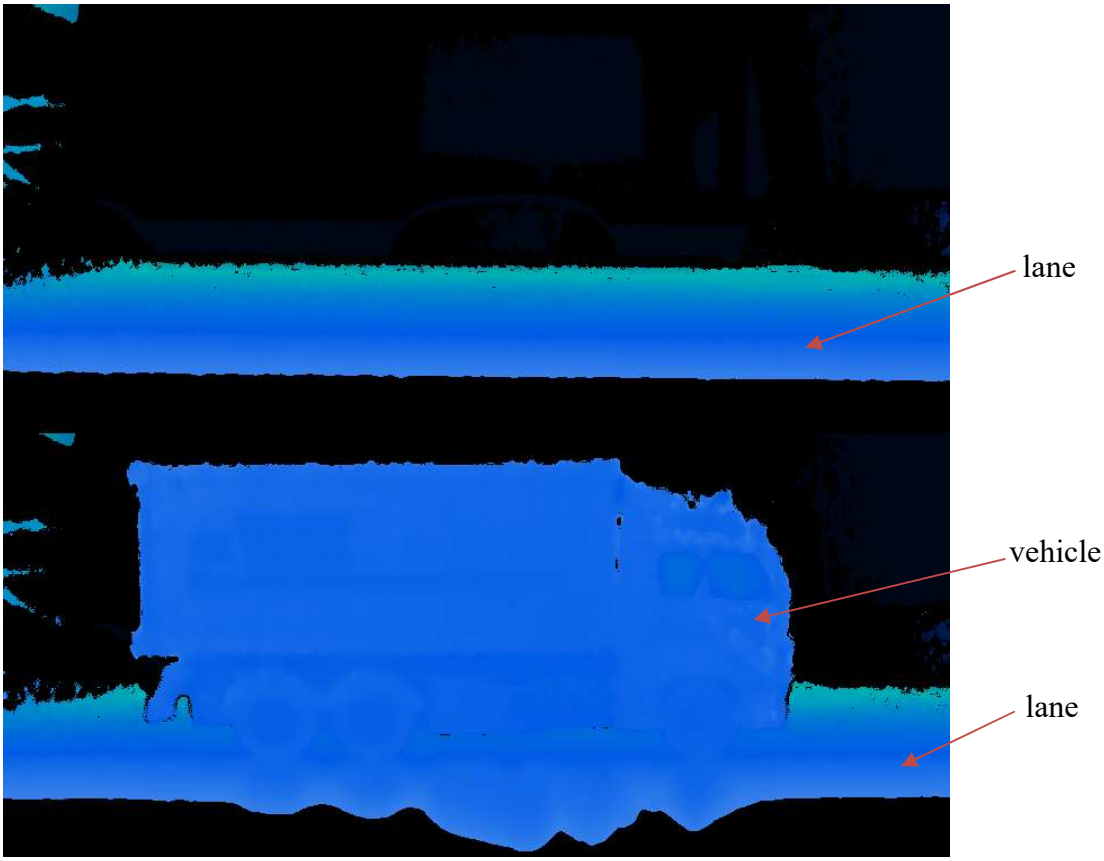


Figure 3.15: Depth map without vehicle (Top); depth map with a vehicle (Bottom)

To prevent improper detection, the height profiling is divided into two parts: *ab* and *bc*, as shown in Figure 3.16, where SH and SF are known.

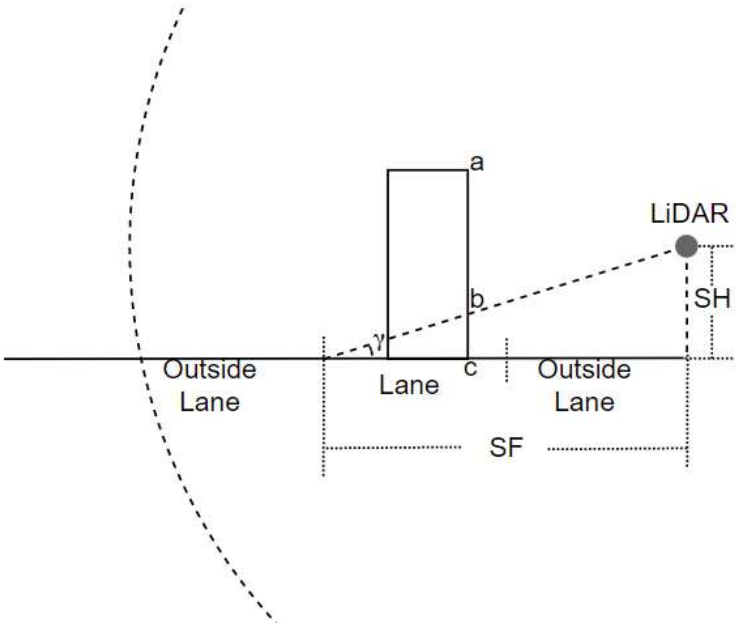


Figure 3.16: A vehicle is passing on the lane; ac is the upper part of the vehicle, while bc is the lower part

While the profiling system begins to work, the LiDAR should scan the background and acquires its depth map corresponding to the ground environment. After that, a point cloud of the ground data is obtained.

There are several points where the value of d is equal to that of SF, which indicate the far rim of the lane. A mean value of these points in the y-axis is defined by $\tilde{y}_{lane}$, which is a threshold level to indicate which part belongs to ab or bc. An angle γ can be calculated from the actual position of the LiDAR.

$$\tilde{y}_{lane} = \frac{y_{l_1} + y_{l_2} + \cdots + y_{l_r}}{r} \tag{3.12}$$

where $l_r \in n$ and $1 \leq r \leq n$

$$\gamma = \tan^{-1} \frac{SH}{SF} \tag{3.13}$$

After the above preparing procedures, the system will start to operate normally, raw data from LiDAR and the camera are collected continuously. The system will only focus on analysing valid data which is defined as the data taking at the moment of the highest magnetic flux density measured by geomagnetic sensor. Take an example as shown in Figure 3.17.

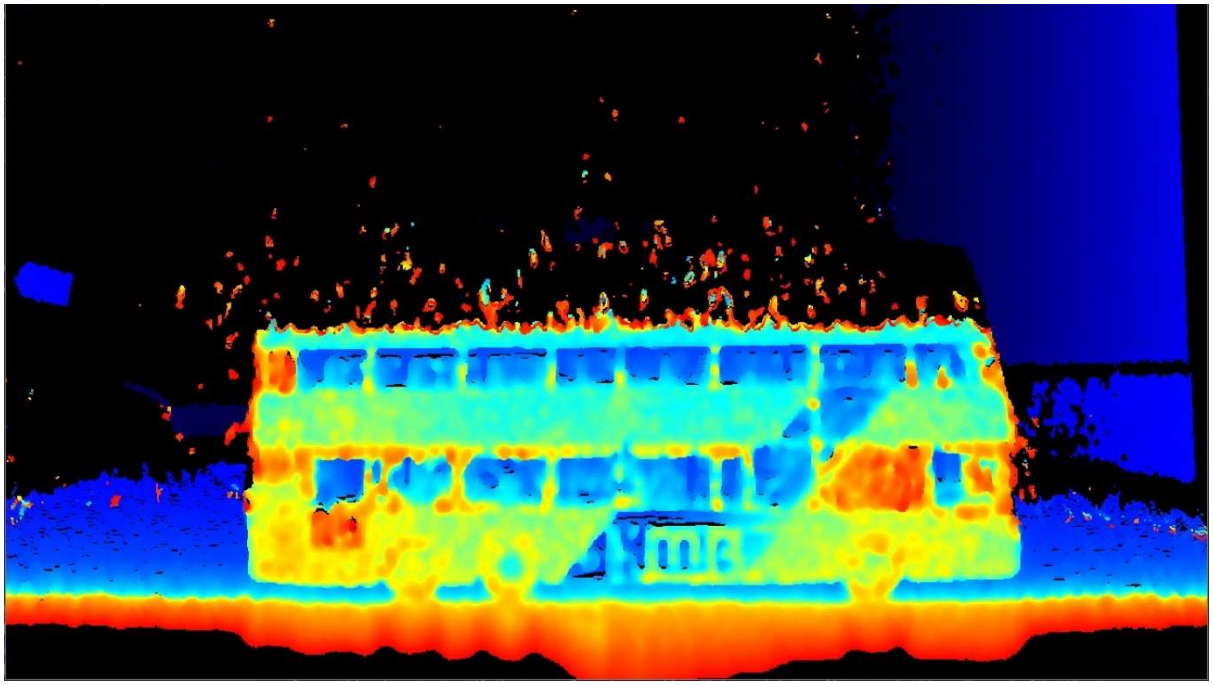


Figure 3.17: Depth map of a double-decked bus

Base on the depth map generated by valid data, the region of interest (ROI) will be identified according to the ground value.

$$D_n^{(ROI)} = D_n \times \begin{bmatrix} L_1 & 0 & \cdots & 0 \\ 0 & L_2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & L_n \end{bmatrix} \tag{3.14}$$

$$L_n = \begin{cases} 1 & if\ d_n \geq SF \\ 0 & if\ d_n < SF \end{cases} \tag{3.15}$$

where L_n is an $n \times n$ diagonal matrix.

While the vehicle is passing into the scanning area, the YOLO detection will also give another ROI. This is represented by a point cloud of YOLO bounding box B^{yolo}, which indicates that a single vehicle is detected. It should be noted that for each time of detection, the bounding box does not always bound the whole vehicle. The bounding box is applied to the $D^{(ROI)}$ to narrow down the processing range and form with a new ROI $D^{(ROI*)}$.

$$B^{(yolo)} = \begin{bmatrix} x_{b_1} & x_{b_2} & \cdots & x_{b_k} \\ y_{b_1} & y_{b_2} & \cdots & y_{b_k} \end{bmatrix} \tag{3.16}$$

$$D_n^{(ROI*)} = \begin{cases} D_n^{(ROI)} & if\ n = b_k \\ 0 & if\ n \neq b_k \end{cases} \tag{3.17}$$

where $b_k \in n$ and $1 \leq k \leq n$.

Figure 3.18 shows the above bounding boxes of ROI. After locating them, some background error and region of non-interest will be removed accordingly. In addition, vehicle windows are transparent or partially transparent to the light pulse of LiDAR where the depth measurement is not valid. Hence, these regions will be filtered by setting a threshold value.

Let F be the set of points filtered by T from $P_n^{(ROI*)}$, where T is a threshold value of background error detection and the region of non-interest. Figure 3.19 shows the depth map after the above filtering procedures.

$$F = [p_{f_1} \quad p_{f_2} \quad \cdots \quad p_{f_j}] = P_j^{(Filtered)} \qquad \forall d_n^{(ROI*)} > T \tag{3.18}$$

where $f_j \in n$ and $1 \leq j \leq n$.

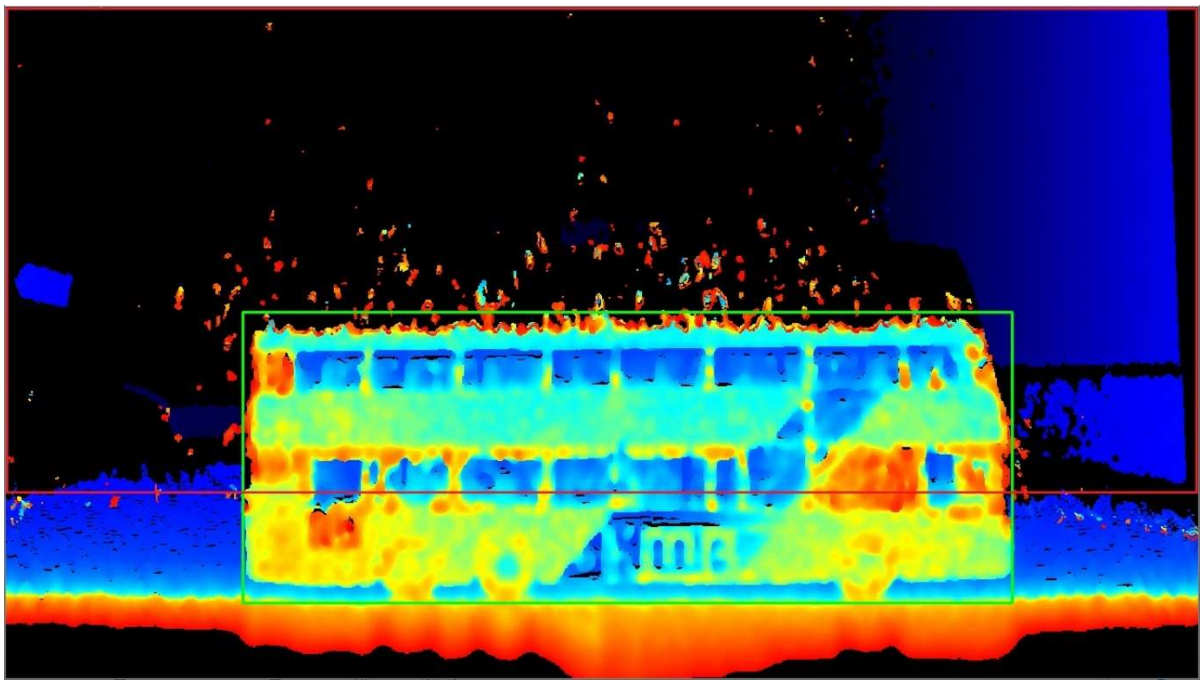


Figure 3.18: Depth map with ROI according to ground data (red box) and YOLO (green box)

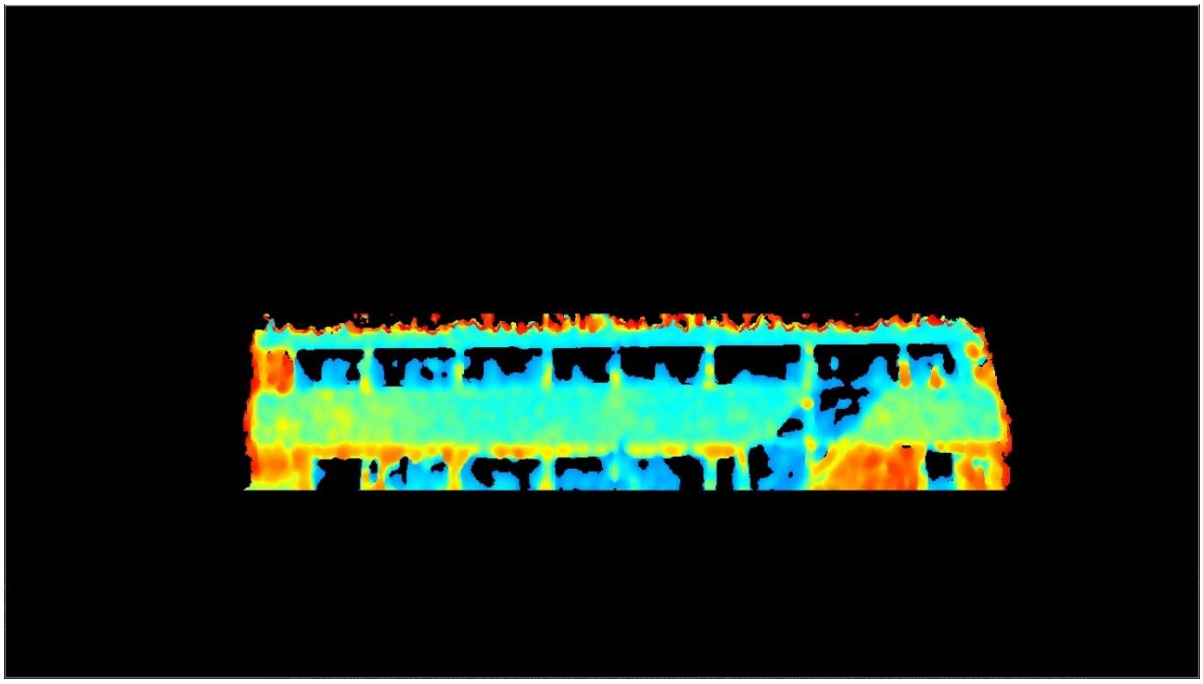


Figure 3.19: Depth map after filtering

Finally, the average depth of the vehicle can be calculated. To reduce the error from the edge detection, the trimmed mean is taken:

$$d_{mean} = \frac{1}{K-2p} \sum_{j=p+1}^{K-j} d_{f_j} \tag{3.19}$$

where $p = 0, 1, 2, \ldots, (K-1)/2$ and K is an odd number.

By using the average depth, a valid point cloud of a vehicle can be found. Assume that the confidence level is α. The population standard deviation of d_{mean} and the margin of error (ME) can be obtained by the calculation given below.

$$\sigma_d = \sqrt{\frac{\sum_{i=1}^{j}(d_{f_i} - d_{mean})^2}{j}} \tag{3.20}$$

$$ME = Z_{\alpha/2} \times \frac{\sigma_d}{\sqrt{j}} \tag{3.21}$$

A valid point of a vehicle can be defined within a range $d_{mean} \pm ME$, which is called a confidence interval. The profiling system will start from a point in F, e.g., (x_i, y_i), in which the depth magnitude is nearest to d_{mean}. All nearby points are then searched, i.e., $(x_i + r, y_i + s)$, where r and s is an integer number which constitute the searching range to see if it is a valid point. If this point is valid, the searching process will repeat with respect to this point until no other valid points are detected. It should be noted that the area of lane, i.e. all points of $y \leq \tilde{y}_{lane}$, will not be searched. Please be noted that the above procedures can obtain all the points belongs to the vehicle outside the YOLO bounding box and remove all the error points within the confidence interval but not belongs to the vehicle.

Next, all the valid point will transform to a depth map, as show in Figure 3.20, and apply Canny edge detector to find the edge of whole vehicle except inside the lane area, as shown in Figure 3.21.

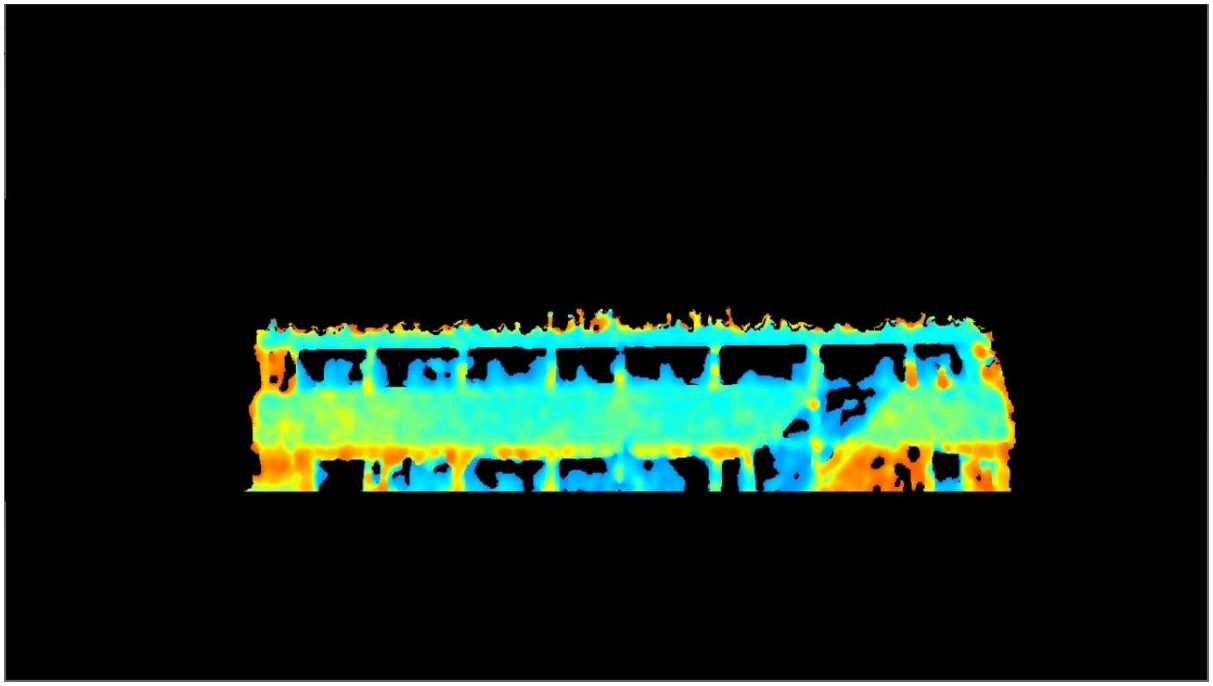


Figure 3.20: Depth map generating by searching valid points

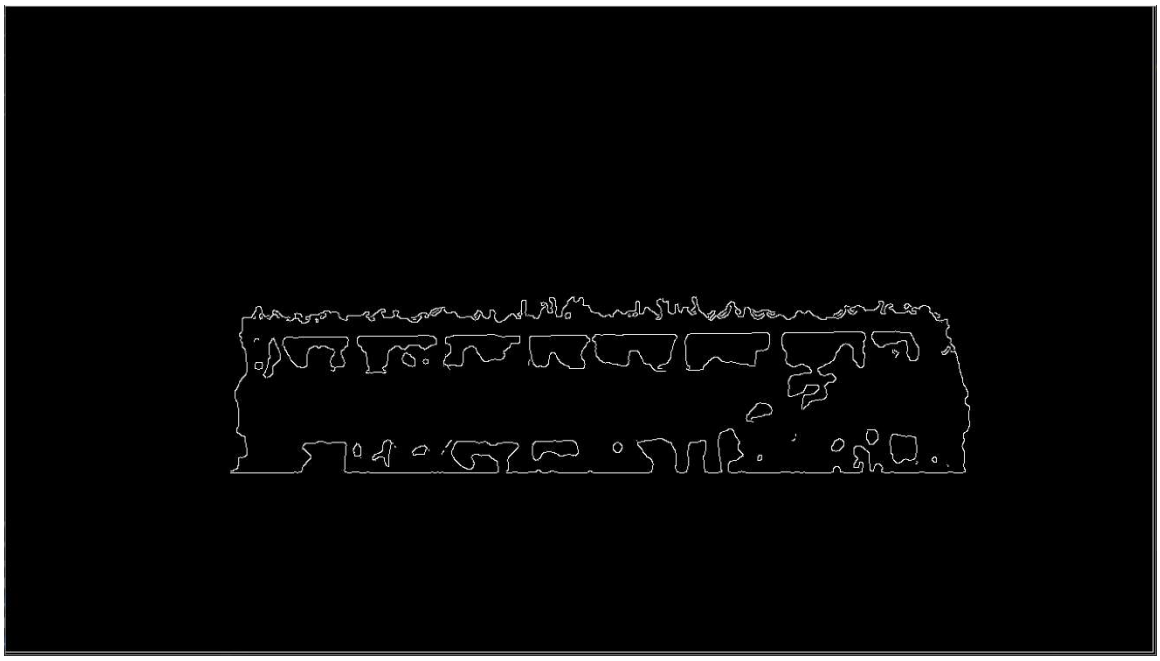


Figure 3.21: The edge of partial vehicle

Afterwards, a set of points (A) will be obtained from the points of edge and a shade of part of the vehicle profile will be derived.

$$A = [p_{a_1} \quad p_{a_2} \quad \cdots \quad p_{a_m}] \tag{3.22}$$

The dimension in pixel, length (L') and height (H'), of the vehicle can be calculated by

the formula given below.

$$L' = \max_{1<i<m} x_{a_i} - \min_{1<i<m} x_{a_i} \tag{3.23}$$

$$H' = ab + bc = \max_{1<i<m} y_{a_i} - \tilde{y}_{lane} + (SF - d_{mean}) \tan\gamma \tag{3.24}$$

Finally, the dimension in the real world, length (L) and height (H), of the vehicle can be calculated by the formula given below.

$$L = kL'd_{mean} \tag{3.25}$$

$$H = kH'd_{mean} \tag{3.26}$$

where k is a scaling coefficient.

The assumption of the above algorithm is that the height of the vehicle is higher than the position of the LiDAR. Otherwise, the whole vehicle is out of the ROI and the system cannot distinguish between the lane and the vehicle.

3.8 Workflow of the profiling system

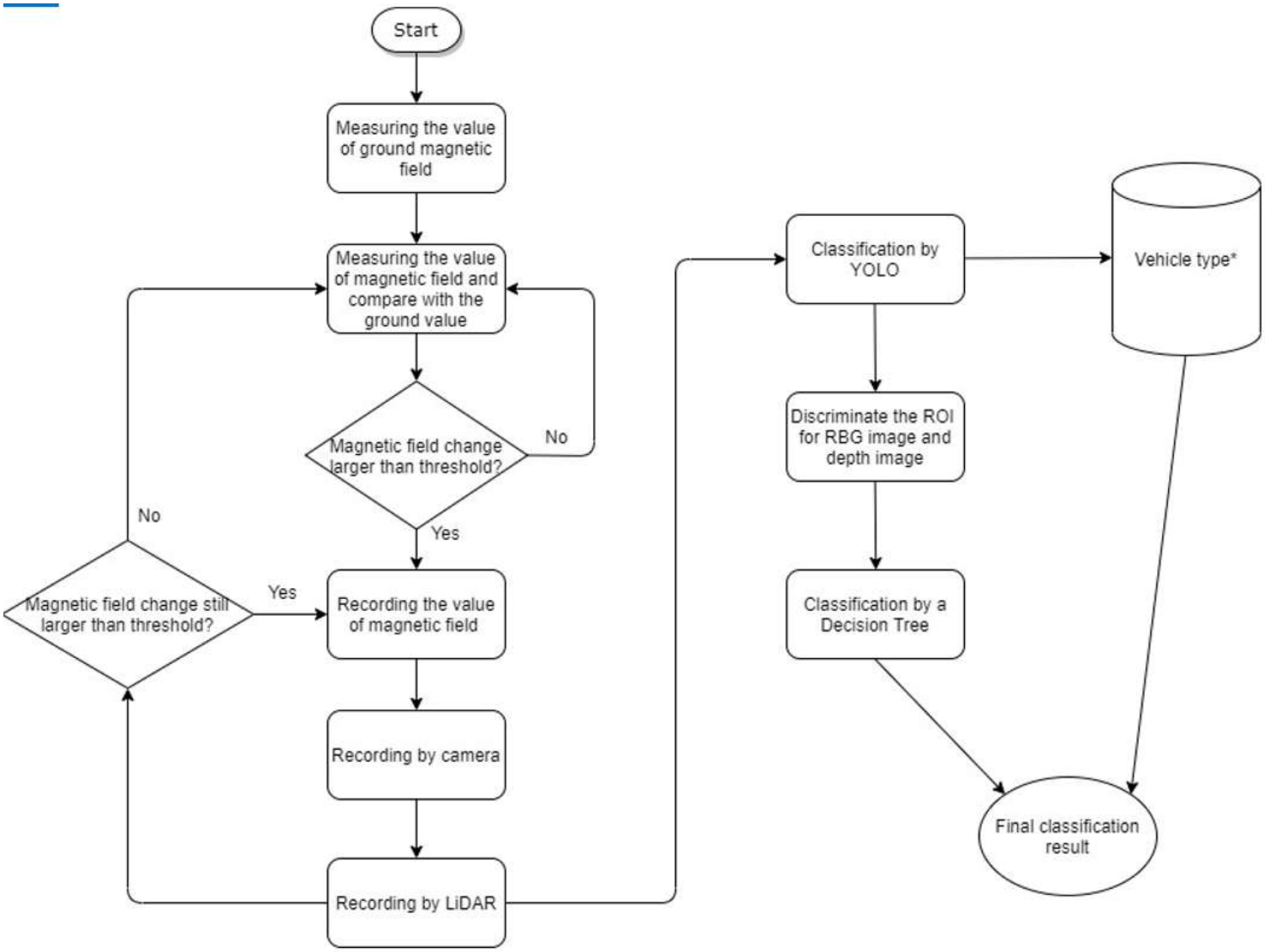


Figure 3.22: Workflow of the profiling system

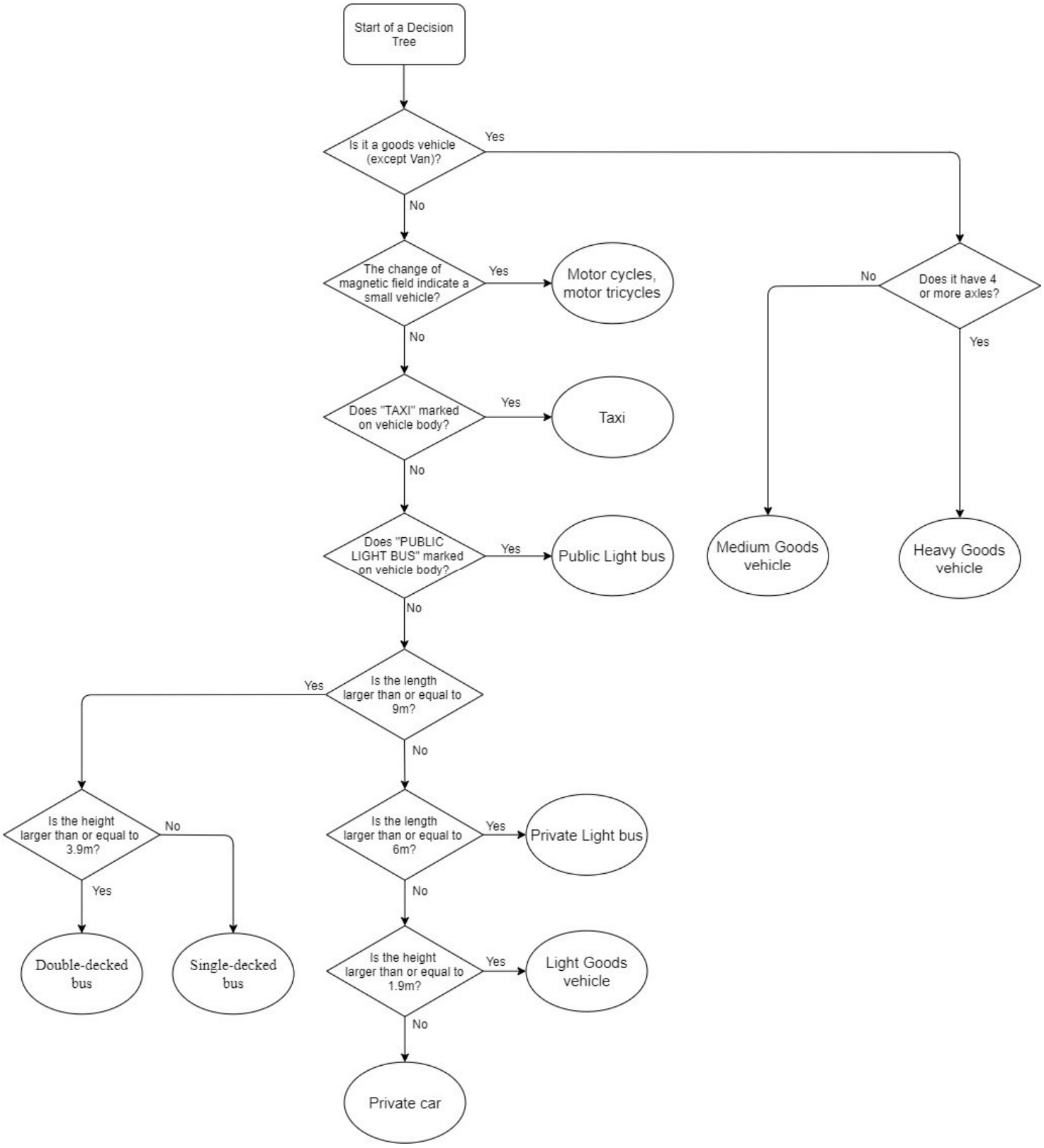


Figure 3.23: The decision tree for classification

The data processing of the proposed profiling system and program workflow are illustrated in Figure 3.22 and 3.23. To begin with, the start of the system is to get the ground value the data obtained from the geomagnetic sensor and the LiDAR. A ground value is an offset value and is supposed to be the lowest value for all time. In order to obtain this value, the user should ensure that no vehicle or unusual object, especially metallic, is placed inside the detection range; otherwise, the reference ground value is useless and will result in fault detection while operating the profiling system. After that, the system can be worked accordingly. Upon starting the profiling system, the geomagnetic sensor will scan the magnetic field continuously.

When SBC receives the raw data from the geomagnetic sensor, a measured magnetic field will be calculated and compared with a threshold value. The threshold value is a defined value that indicates the existence of any vehicles. If the geomagnetic sensor senses that the change in magnetic flux density is larger than the threshold value, it means that there is a vehicle approaching the geomagnetic sensor. The LiDAR and the camera will start to record the point cloud and RGB images. The frame per second is set to 30, and so, hundreds of frames will be obtained for further analysis. After sensing, the system will start classifying the type of vehicle. As the magnetic field value of a motorcycle is much smaller than others, it can used to identify if it is a motorcycle first.

Next, the RGB image will first undergo analysis by YOLO. The system will start identifying any vehicles in the frame. This action can identify the Region of Interest (ROI) for processing the consequence. After that, the YOLO will start determining whether it is a goods vehicle. If it is a goods vehicle, the RGB images will be checked by the Hough circle algorithm to count the axles; normally, if it has 4 or more than 4 axles, a heavy goods vehicle can be identified. If it is not a goods vehicle, the RGB images will be checked by Tesseract to identify some keyword, e.g., ‘TAXI’ and ‘PUBLIC LIGHT BUS’, and hence, a taxi or a public light bus can be identified.

In addition, the point cloud measured by the LiDAR can be used for calculating the dimension, and hence the system can distinguish among a light bus, a double-deck bus, a single deck bus, a private car and a light goods vehicle.

3.8.1 Magnetic Field measurement

According to the literature and the patents discussed in the previous section, the geomagnetic sensor is strong enough to identify motorcycles/motor tricycles. As mentioned in the previous sections, the measurement of motorcycles/motor tricycles by the LiDAR is relatively poorer than that of other vehicles. Therefore, in the first step of classification analysis, the profiling system can identify motorcycles/motor tricycles using the magnetic field data.

3.8.2 Word identification

There is a difficulty in distinguishing between a taxi and a private car. In fact, the same vehicular model can be used as a taxi or a private car depending on the owner registration under the government. Figure 3.24 is an example, where the private car model is the Toyota Prius. Not only is their shape similar, but the color of both is also nearly the same. It is hard to distinguish between them by size or appearance. Therefore, the system should further focus on other possible features that form the main differences between a taxi and a private car.

Figure 3.24: A taxi (left) and a private car (right)

According to Cap. 374A Road Traffic (Construction and Maintenance of Vehicles) (https://www.elegislation.gov.hk/hk/cap374A) Regulations no. 45,

Taxis to have illuminated signs and markings

Every taxi shall—

(a) have fitted on the top of its roof an illuminated sign, of a type approved by the Commissioner, which at all times during the hours of darkness when the taxi is available for hire displays the word "TAXI" so that it is clearly visible from the front and the rear of the vehicle; and (L.N. 258 of 1984)

(b) *be plainly marked in English and Chinese writing of uniform size not less than 100 millimetres in height on the outside of the vehicle on both the near and off sides with the word "TAXI" and the characters "的士".*

In brief, each taxi should have a sign on the top, as shown in Figure 3.25, and a word "TAXI" printed on both sides of its body, as shown in Figure 3.26.

Figure 3.25: Taxi sign on the top of taxi

Although the aforementioned is one of the main differences between a taxi and a private car, it is not a good and reliable feature to distinguish between them. The camera in this study is designed to be installed on the side of the road, the word 'TAXI' on the sign cannot be captured. Actually, compared with the vehicle's body, the tag is relatively small; and it is sometimes hard to detect it. In fact, it does not prohibit people from placing something on the top of their car. In some situation, the sign might be might faultily detected, causing an incorrect prediction.

Another main difference is the print text on the vehicle's body – the word 'TAXI' must be printed on the side of a taxi. It is a strong indicator for identifying taxi, even without any relevant characteristics. In this case, searching for the text on the taxi's body is a solid solution.

Figure 3.26: Side-view of a taxi

Similarly, for the public light bus, the words 'PUBLIC LIGHT BUS' should be marked on both sides, as shown in Figure 3.27, according to Cap. 374A Road Traffic (Construction and Maintenance of Vehicles) (https://www.elegislation.gov.hk/hk/cap374A) Regulations no. 50.

Additional markings on public light buses

Every public light bus shall be plainly marked—

(a) in block letters and Chinese characters, of uniform size not less than 100 millimetres in height, on the outside of the vehicle on both the near and off sides, with the words "Public Light Bus" and the characters "公共小型巴士";

For more information, the model of the public light bus mainly used in 2020 constitutes: GMI Gemini, Golden Dragon XML6701J18, Mitsubishi Fuso Rosa, Optare Solo SR and Toyota Coaster.

Figure 3.27: Side-view of a mini-bus

In this study, Tesseract will be used for word detection. Tesseract is an optical character recognition engine, developed and sponsored by Google since 2006. One of the famous applications of this technology is license plate recognition, which is similar to the application of this research. There are several page segmentation modes (PSM) for Tesseract, and PSM 11 is set for the proposed system. In this PSM, Tesseract will try to find as much text as possible as in the frame. As a result, a list of words is acquired. If the profiling system detects the word 'TAXI' or 'PUBLIC' + 'LIGHT' + 'BUS', then it is able to classify a taxi and a public light bus respectively.

3.8.3 Dimensional analysis

The dimension of a vehicle is one of its key features. The profiling system should familiarise the typical size and restriction size of most vehicle types. Table 3.5 summaries the dimensional data of different main vehicles used in Hong Kong in 2020.

Table 3.5: The main models of single-decked buses used in 2020

Model	Length	Width	Height
Young Man JNP6122UC **	12	2.54	3.84
Young Man JNP6122G	12	2.5	3.2
Scania K230UB	10.6-12	2.527	3.15
Young Man JNP6120GR	11.6	2.54	3.17
Young Man JNP6105GR	10.5	2.54	3.17
Volvo B7R	12	25	3.5
Yixing SDL6105／SDL6120	11.6-10.5	2.5	-
MAN NL273F	12	2.5	-
Isuzu FRR	8.8	2.26	2.89
ISUZU LV434R	11	2.4	3.5
MAN A91	12	2.5	3.5
ISUZU LV423R	12	2.4	3.4
Daewoo BH117L	11	-	3.5-3.7
Enviro200 Dart	10.4-11.3	2.432	3.05
BYD K9R	11.6-12	2.5	3.2
BYD K9A	12	2.55	3.2
Size range	8.8-12	2.26-2.55	2.89-3.84

** This model is a capacitive bus, with the maximum height of 3.84m, but it includes an electric charging port. Its appearance is shown in Figure 3.28.

Figure 3.28: Young Man JNP6122UC

Table 3.6: The main models of double-decked buses used in 2020

Model	**Length**	**Width**	**Height**
Alexander Dennis Enviro400	10.45	2.55	4.4
Alexander Dennis Enviro500	11.3-12	2.51	4.4
Dennis Trident 3	10.3-12	2.5	4.17-4.4
MAN A95	12-12.8	2.5-2.55	4.2-4.43
Scania K280UD	12	2.5	4.4
Scania K310UD	12	2.5	4.4
VDL DB300	10	2.55	4.38
Volvo B8L	12-12.8	2.55	4.35-4.4
Size range	10-12.8	2.5-2.55	4.17-4.43

Table 3.7: The main models of light goods vehicle (van) used in 2020

Model	Length	Width	Height
TOYOTA HIACE	4.695-5.38	1.695-1.88	1.98-2.105
NISSAN URVAN	4.695-5.08	1.695-1.88	1.99-2.285
Volkswagen Transporter	4.892-5.292	1.904-1.959	1.935-2.476
FORD TRANSIT***	4.972-5.339	1.986-2.29	1.969-2.022
PEUGEOT EXPERT	4.609-5.309	1.905-1.94	2.204
HYUNDAI H1	5.15-5.125	1.920	1.925
M-BENZ VITO	4.748-5.223	1.906	1.875
LAND ROVER DISCOVERY	4.838-4.970	2.022-2.22	1.888-1.909
Size range	4.695-5.38	1.695-2.22	1.875-2.476

*** There are many sub-models for FORD TRANSIT. The dimension stated in the table is the size of the ones commonly used in Hong Kong.

The model of private light buses can also be used as public light buses, and vice versa. In fact, most of the models of private light bus are the same as those of public light buses in Hong Kong. Therefore, their size can indicate the same model.

Table 3.8: The main models of light bus used in 2020

Model	Length	Width	Height
GMI Gemini	6.99	2.04	2.865
Golden Dragon XML6701J18	6.97	2.05	2.65
Mitsubishi Fuso Rosa	6.245-7.73	2.01	2.63-2.735
Optare Solo SR	7.87	2.34	2.885
Toyota Coaster	6.255-7.725	2.035	2.585-2.6
Size range	6.245-7.73	2.01-2.34	2.585-2.88

Table 3.9: The main models of taxi used in 2020

Model	**Length**	**Width**	**Height**
Toyota Crown Comfort	4.695	1.695	1.525
Toyota Comfort Hybrid (JPN)	4.4	1.695	1.88
Toyota Prius	4.48	1.745	1.49
Toyota Noah	4.71	1.73	1.825
Nissan Cedric	4.69	1.695	1.445
Nissan NV200	4.4	1.695	1.850
Ford Transit Connect	4.825	1.810	1.93
Size range	4.4-4.825	1.695-1.81	1.445-1.93

According to Tables 3.5, 3.6, 3.7, 3.8 and 3.9, light bus, double-decked bus, single-decked bus and light goods vehicle can be classified based on the analysis of dimension. The length of a double-decked bus and single-decked bus is about 12m normally, while that of a light bus and light good vehicle is shorter than 8m. The classification line of these two groups, therefore, can be set to 9m. The length of the light bus is within 6.245–7.73m while that of light goods vehicle is within 4.695–5.38m. The classification line between them can hence be set to 6m. The height of a double-decked bus is 4.17–4.43m while that of a single-decked bus is 2.89–3.7m. The classification line between them can therefore be set to 3.9m. The height of private car is assumed to be similar to the taxi; a taxi is 1.445–1.93m while that light goods vehicle is 1.875–2.476m. The classification line between them can therefore be set to 1.9m, but it is not correct, further improvement should be considered to fix this issue.

3.8.4 Role of deep learning

As mentioned above, the deep learning technique is commonly used in the object detection nowadays. However, before it can have a high accuracy rate, several problems need consideration. First, greater resources are required in order to obtain enough a data set large enough to make the training process of deep learning. Second, the annotation process of data sets needs huge human resources. Finally, in order to speed up the training process, a high computing power and high standard Graphic Processing Unit (GPU) is needed. To design a novel and cost-effective profiling system, the role of deep learning should be optimised.

As the dimensions of the double-decked bus, medium goods vehicle, and heavy goods vehicle are similar, their classification by dimensional analysis is not valid. In this case, appearance analysis through image deep learning technology is a better solution. The appearance of goods vehicles usually differs from that of other types of vehicles, while the classification of other types of vehicles are analysed by other features.

3.8.5 Trained Dataset for YOLO

The performance of YOLO detection is directly influenced by the trained data set. This data should include all types of vehicles that are going to be classified. The source training of a dataset can be varied, any photos that contain the classified class are suitable, and it can be taken by a camera or a mobile phone by any person anywhere, as well as on the internet.

3.8.6 Image labelling

In order to enhance the performance of object detection, the dataset is customized so that-more images are going to be trained. Before the training process, the images need to be annotated. Each image contains a lot of information, including ROI and background. The ROI should be labelled with a certain category, which is a step called annotation. After collecting lots of images and point cloud, all data should be processed with annotation manually one by one. There is no auto-annotation program available which is good enough to perform annotation. Though the task is tedious, but it is a very important step for the profiling system. Since a wrong annotation label will affect the performance of the profiling system. Below is an example of one image from the dataset.

Figure 3.29: A goods vehicle
(https://www.hgvtraders.com/vehicles/scania-r440-44-tonne-curtain-side-for-sale).

In Figure 3.29, the photograph contains a goods vehicle, a traffic road and lots of plants. The region of interest (ROI) is a vehicle, which is the object to be classified. Therefore, the area of the vehicle should be marked and labelled as shown in Figure 3.30.

Figure 3.30: Labelling of an image inside the data set

An .xml file is created after the above step. This file contains information on the bounding area and labelling name in text format, as shown as Figure 3.31.

```
<annotation>
    <folder>JPEGImages</folder>
    <filename>01_75.jpg</filename>
    <path>C:\Users\Shek\Desktop\car_detect\yolo\VOCdevkit\JPEGImages\01_75.jpg</path>
    <source>
        <database>Unknown</database>
    </source>
    <size>
        <width>1017</width>
        <height>678</height>
        <depth>3</depth>
    </size>
    <segmented>0</segmented>
    <object>
        <name>good_vehicle</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <difficult>0</difficult>
        <bndbox>
            <xmin>101</xmin>
            <ymin>166</ymin>
            <xmax>944</xmax>
            <ymax>564</ymax>
        </bndbox>
    </object>
</annotation>
```

Figure 3.31: Information in .xml after labelling

Each image matches with its particular .xml file. Together with the images, the training process will train on ROI only and hence it can increase the efficiency of the training and exhibit better profiling performance.

3.8.7 Counting of axles

Medium goods vehicles and heavy goods vehicles are similar in appearance with different size, as shown as Figure 3.32, so their tunnel fee is different in charge, and hence, the profiling system should distinguish between them.

Figure 3.32: A medium goods vehicle: Isuzu FSR90 (the left one) and a heavy goods vehicle: Isuzu CYH52 (the right one), the Length/Width/Height are 7.875m/2.20m/2.53m and 10.66m/2.49m/2.98m respectively (https://www.isuzu.com.hk/index-cn)

In addition, heavy goods vehicles can be divided into two types, one is rigid and the other is articulated. The latter is usually connected with a trailer. Sometimes, there might not be any containers on the trailer, and hence the length measurement might fail. Moreover, according to the Code of Practice for the Loading of Vehicles from the Transport Department, the goods container can be placed on the middle part of the vehicle or trailer. Figure 3.33 shows one of the possible placements of a container. The length dimension measured may be divided into two parts: the vehicle head and the container. This may increase the complexity of how to define the length of the vehicle. The variation of the heavy goods vehicle is too large, and hence, it is not recommended to identify it by dimension or magnetic field.

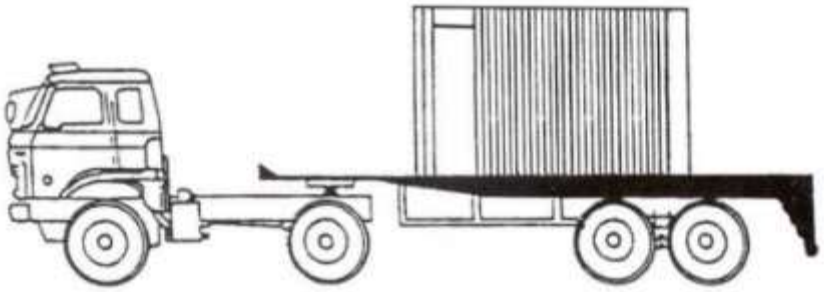


Figure 3.33: Possible placement of container for a heavy goods vehicle

In Hong Kong, most of the heavy goods vehicle is 30 tonnes or above, and they most frequently have 4 or more axles. On the other hand, light goods vehicle and medium goods vehicle must have 2 to 3 axles. Therefore, counting the number of axles can help classification it if it is a heavy goods vehicle.

Table 3.10: Weight, length and axles for goods vehicle

Type of goods vehicle	Total weight (Tonnes)	Description
Light goods vehicle	≤5.5	Vehicle length is shorter than 6m normally 2 axles
Medium goods vehicle	>5.5 to 24	Vehicle length is longer than 6m normally 2 to 3 axles
Heavy goods vehicle	>24	Vehicle length is longer than 7m normally 4 or more than 4 axles

The number of axles can give a hint in classifying the vehicle type: larger sized vehicles usually needs more axles. In fact, the vehicle type classified in the Electronic Toll Collection (ETC) system, referring to Figure 3.34, in Hong Kong also concerns the number of axles. Therefore, the counting of axles is essential in the profiling system.

Figure 3.34: List of vehicle types defined by the ETC system in Hong Kong

Most of the vehicles have 2 axles. Only heavy vehicles need 3 or more axles to support the weight of the vehicle and its loading. In fact, besides goods vehicles, double-decked bus is another common vehicle that might require 3 axles. However, not all of them have 3 axles as shown in Figure 3.35. Therefore, the number of axles is not a critical characteristic for identifying a double-decked bus.

Figure 3.35: Model Enviro400 Euro V

According to Hong Kong Transportation Department, if a vehicle has more than two axles, then an additional charge is needed for each additional axle in most tunnels (https://www.td.gov.hk/en/transport_in_hong_kong/tunnels_and_bridges/toll_rates_of_road_tunnels_and_lantau_link/index.html). Therefore, the counting of axles is a essential function of a profiling system.

There are two proposed methods to identify wheels in this research. The first one uses the deep learning approach, which can be also done by YOLO, or other deep learning algorithm for object detection. This is similar to classifying the goods vehicles as described in the previous section. On the other hand, it requires lots of wheel images for training purpose, so the Hough transform algorithm method is proposed in this project, it can detect any circle in the image.

The image will first transform into an edge picture by Canny edge detector, as shown in Figure 3. 36, to give the shape of all obvious objects. The Hough transform method then detects any line or shape of the edge picture. For instance, all the wheels must be in shape of a circle. Therefore, all circles inside the edge image will be detected.

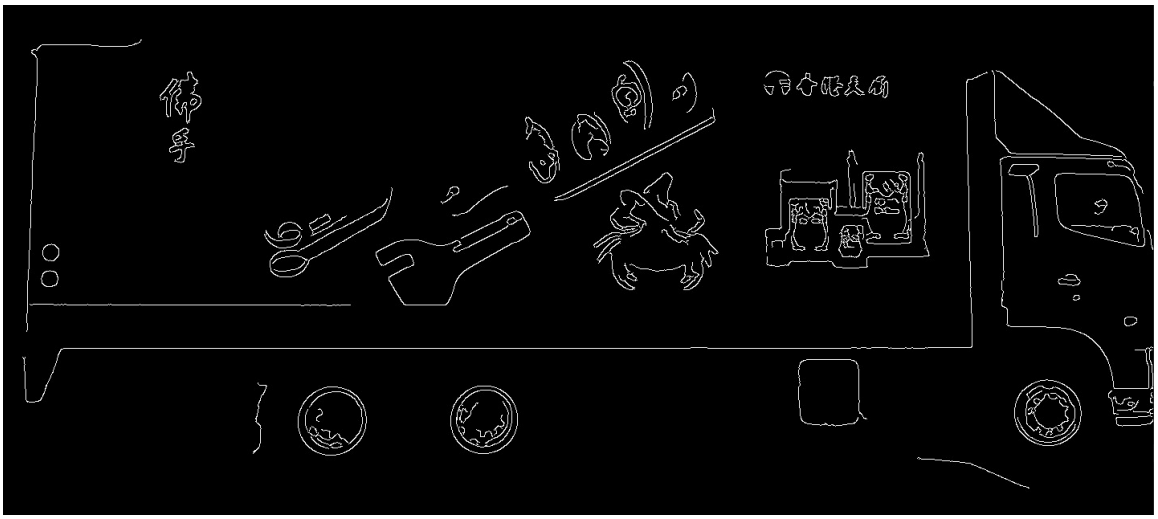


Figure 3.36: Edge picture of an RGB image captured by camera

In some scenarios, there are other circles on the vehicle's body. This will affect the correct rate of axle counting. To prevent fault detection, there are three rules to be considered while applying the Hough transform method. First, the radius of the circle is defined within a valid range and this value should be adjusted as different sites might have a different position of the camera. Second, the circle must be aligned with almost the same vertical level, to make sure there is no mix up with another circle from the surroundings or the vehicle's body. Third, only the lowest set of circles is counted, to prevent any circles above the wheel to be counted. If there are two or more circle which obey the above rules, these circles are counted as wheels as shown in Figure 3.37.

Figure 3.37: Wheel counting after applying the Hough transform algorithm

3.8.8 Multiple-criteria decision analysis (MCDM)

After acquiring the final classification result mentioned in the previous sections, it is going to calculate the confidence score of the final classification result $S^{(final)}$.

$$S^{(final)} = \alpha_0 KP^{(yolo)} + \alpha_1 S^{(text\_T)} + \alpha_2 S^{(text\_P)} + \alpha_3 S^{(axles)} + \alpha_4 S^{(leng\ )} + \alpha_5 S^{(heigh\ )} \quad (3.27)$$

As mentioned in the previous sections, some types of vehicles can be simply classified by YOLO detection with a well-trained data. This can be used to check the validation of the profiling system. The COCO data set is used in current stage, it can be classifying the vehicle

into three categories only, including motorcycle, bus and goods vehicle. In the equation 3.27, the value of $P^{(yolo)}$ can be obtained by YOLO directly during operation. It is a confidence percentage of its classification result, and its range is 0 to 100. K is the scaling coefficient which make the value of $P^{(yolo)}$consistent with S, and its value is set to be 0.02 in current stage. α_0 - α_5 are the coefficients of each of the criteria, and the sum of them is equal to 1. Their value should be investigated with more testing result. In fact, the value can be varied in different environments, such as day or night, sunny or rainy, etc. Their value is set to be 0.2 in current stage. $S^{(text\_T)}, S^{(text\_P)}, S^{(axles)}, S^{(length)} and\ S^{(height)}$ are the confidence score of different criteria, and the value of it is defined in the tables 3.11 to 3.15.

Table 3.11: Confidence score for the word "TAXI" scanned

	Taxi	Others
Four letters of "TAXI"	2	0
Three letters of "TAXI"	1	1
Two or less letters of "TAXI"	0	2

Table 3.12: Confidence score for the word "PUBLIC LIGHT BUS" scanned

	Public light bus	Others
Ten or more letters of "PUBLIC LIGHT BUS"	2	0
Five to nine letters of "PUBLIC LIGHT BUS"	1	1
Four or less letters of "PUBLIC LIGHT BUS"	0	2

Table 3.13: Confidence score for the number of axles count

	Heavy goods vehicle	Medium goods vehicle	Double-deck bus	Others
Four or more axles	2	0	0	0
Three axles	0	2	2	0
Two axles	0	0	2	2

Table 3.14: Confidence score for the length acquire

	Heavy goods vehicle, Single / double-decked bus	Medium goods vehicle	Private / public light bus	Others
Length>10.5	2	0	0	0
10.5>Length > 9.5	1	1	0	0
9.5>Length>8.5	0	2	1	0
8.5>Length>7.5	0	2	2	0
7.5>Length>6.5	0	2	2	0
6.5>Length>5.5	0	1	1	1
Length <5.5	0	0	0	2

Table 3.15: Confidence score for the height acquired

	Heavy / Medium goods vehicle, Single / double-decked bus	Light goods vehicle	Private / public light bus	Others
Height > 4	2	0	0	0
4> Height >3	1	0	1	0
3> Height >2	0	2	2	1
Height <2	0	2	0	2

The range of scores is set to be 0 to 2 at current stage. The range and the score tables should be revised afterward according to the testing results. If the testing data is large enough, it may apply a numerical machine learning / AI to find out the most suitable value in the future.

In addition, if the final confidence score is too low, i.e. smaller than 1, the system decides a failed case of the classification. Otherwise, it can determine to be a successful case and show the classification result to the user.

Chapter 4 Result and Analysis

4.1 System prototype and data collection

The prototype has three main components: A LiDAR camera Intel® RealSense™ L515, an SBC IEI NANO-BT-i1-J19001-R11 and a geomagnetic sensor, referring to Figure 4.1, 4.2 and 4.5 respectively. The simulation test is carried out by combining all the data collected by the sensors.

A long-range and wide FoV LiDAR is expensive. Therefore, a several metres range LiDAR with model cars is used in the prototype to simulate depth measuring. Intel® RealSense™ L515 has both a camera and a LiDAR inside. The horizontal FoV and vertical FoV are 70° and 55°respectively for LiDAR while 70°and 43° respectively for camera.

Figure 4.1: LiDAR camera – Intel® RealSense™ L515

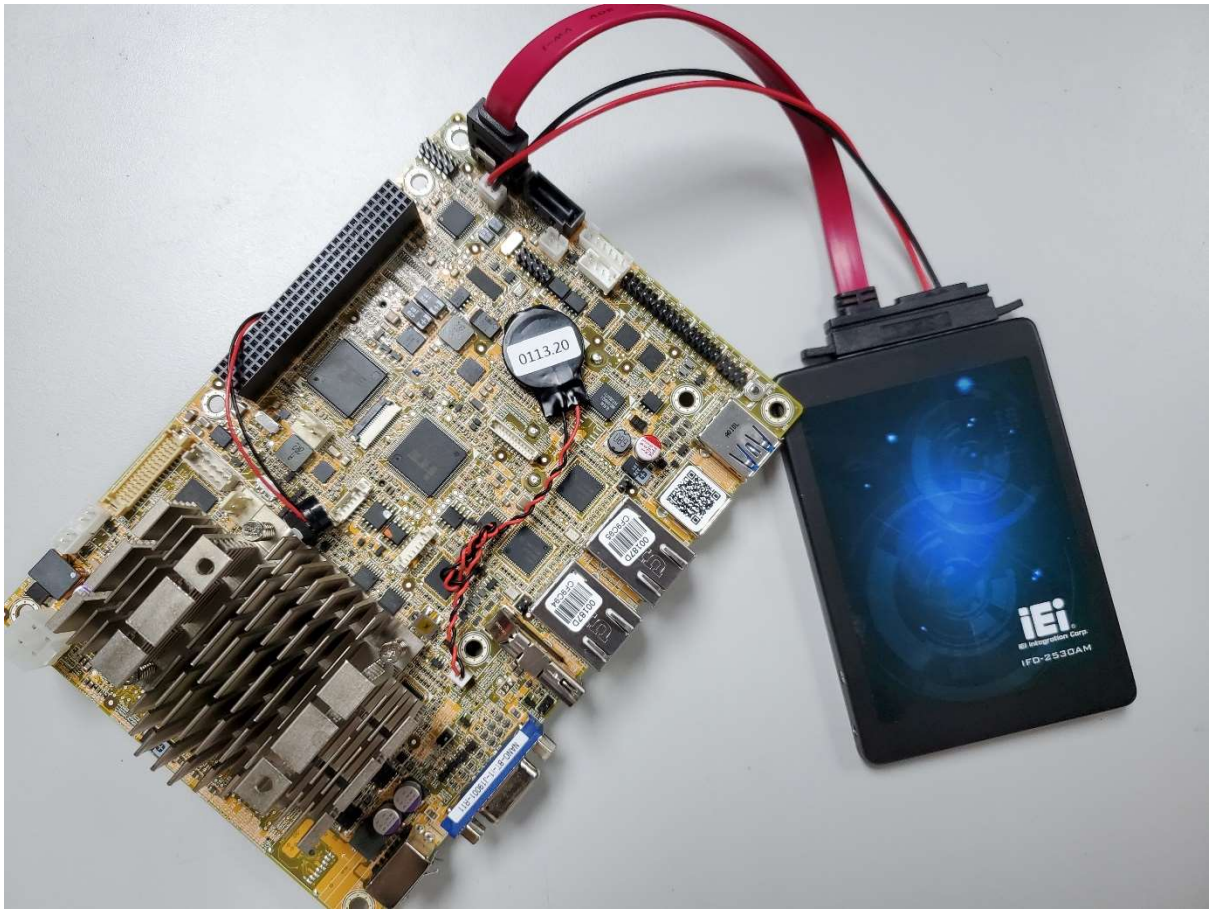
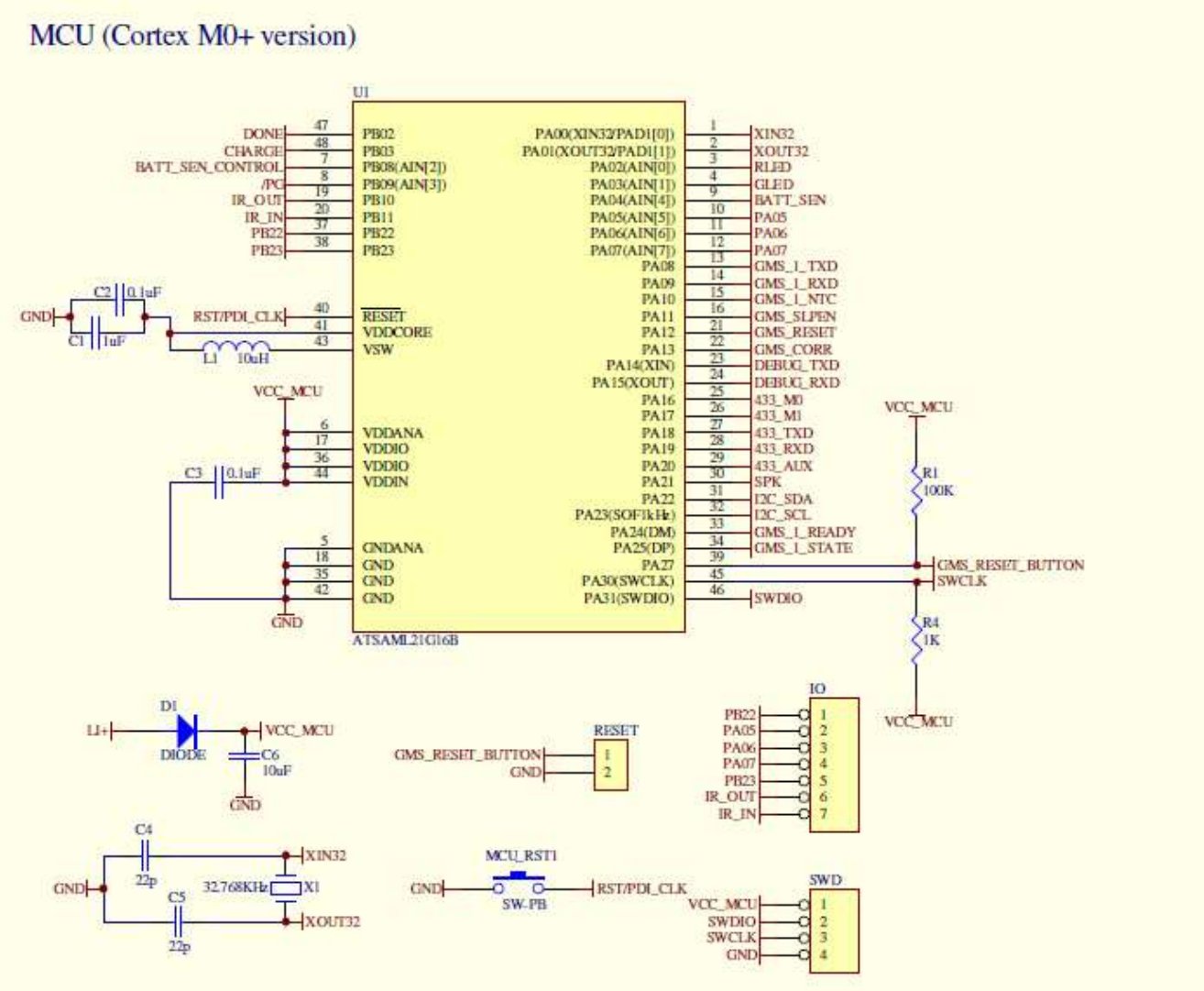


Figure 4.2: An SBC IEI NANO-BT-i1-J19001-R11

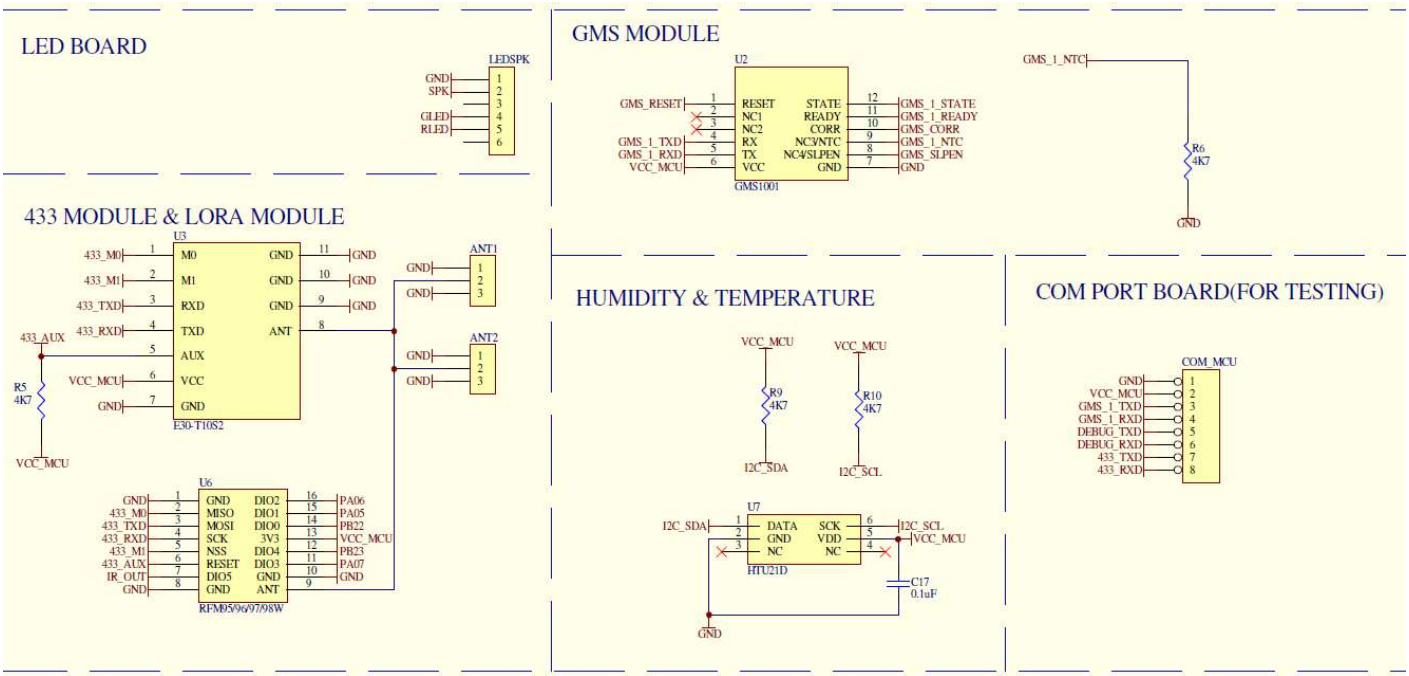
COCO (Common Objects in COntext) trained dataset is to be used for YOLO. It is an open-source data set with around 90 categories, and only 4 of them are extracted and applied in this project, i.e. car, truck, motorcycle, and bus. With an increasing number of images being collected, the original COCO data set can be further trained by the user so that a customised dataset can be used. A COCO dataset with version “2017 Train/Val/Test” is implemented; the public download link is https://cocodataset.org/#download.

A geomagnetic sensor is developed to achieve a low-cost and energy-efficient solution. The energy consumption is a very important factor, as the sensor may dig into the ground, and hence, it is hard to take it out once it is installed. This geomagnetic sensor is designed to communicate both with wires and wirelessly. The IC used in the geomagnetic sensor module is LSM303D produced by STMicroelectronics. Its circuit schematic is presented in Figure 4.3 and 4.4.



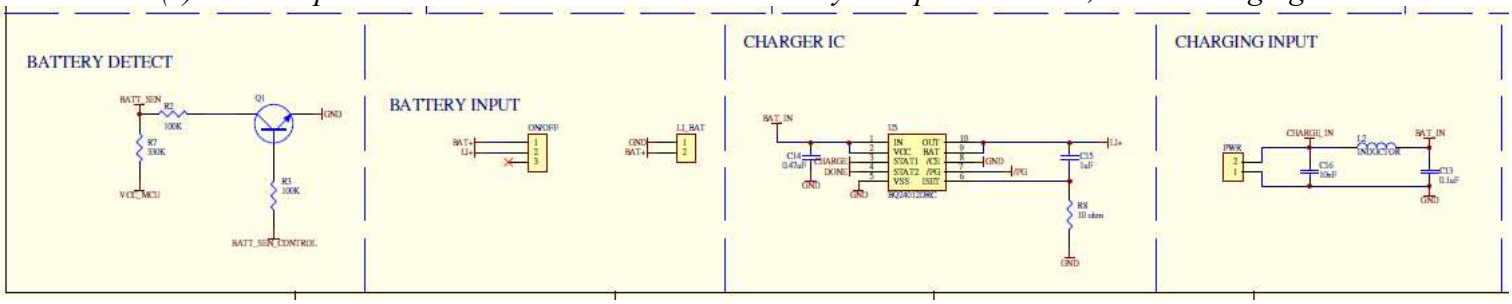
(a) MCU part: a firmware is embedded and controls the whole device, connecting all the other parts to perform its functions

Figure 4.3: Circuit schematic of the geomagnetic sensor being developed, (a) MCU part



(b) Sensor and communication part: a GMS module is a magnetic field sensor which sends an MFI data by SERCOM. An IC sensor measures humidity and temperature, which sends a raw data directly by I2C. Two LEDs and one buzzer are connected to MCU for debugging purpose., 433 and LORA module is a wireless communication interface, and COM port is a wire communication with RS232 TTL interface

(c) Power part: connected to a 3.6V Lithium battery as a power source, with a charging IC



which can charge the battery by 3.6~4V charging input. A voltage detection circuit is also implemented to detect the voltage of the battery

Figure 4.4: Circuit schematic of the geomagnetic sensor being developed (con't), (b) sensor and communication part, (c) power part.

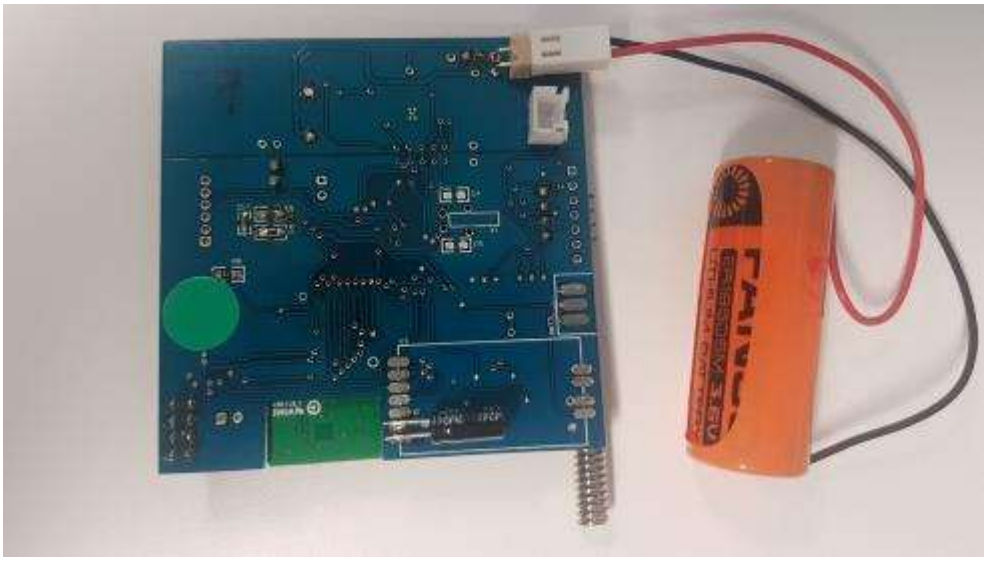


Figure 4.5: A geomagnetic sensor with a battery

Figure 4.6: Testing environment of the geomagnetic sensor

Figure 4.7: Geomagnetic sensor placed into a mounting case, view without a cover (Left) and view with a cover (Right)

A geomagnetic sensor is placed under the ground as shown in Figure 4.7, and it sends raw data via wireless communication. A vehicle is entering into the detection area as shown in Figure 4.8 and 4.9. The magnetic field flux density has greatly changed while the vehicle passes across. Graph plots in Figure 4.10 show the variation in the magnetic field against time.

Figure 4.8: Private car for the testing geomagnetic sensor

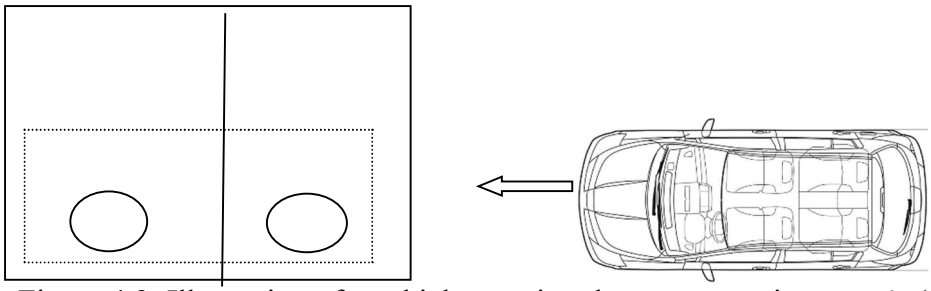


Figure 4.9: Illustration of a vehicle entering the geomagnetic sensor's (circle) detection area

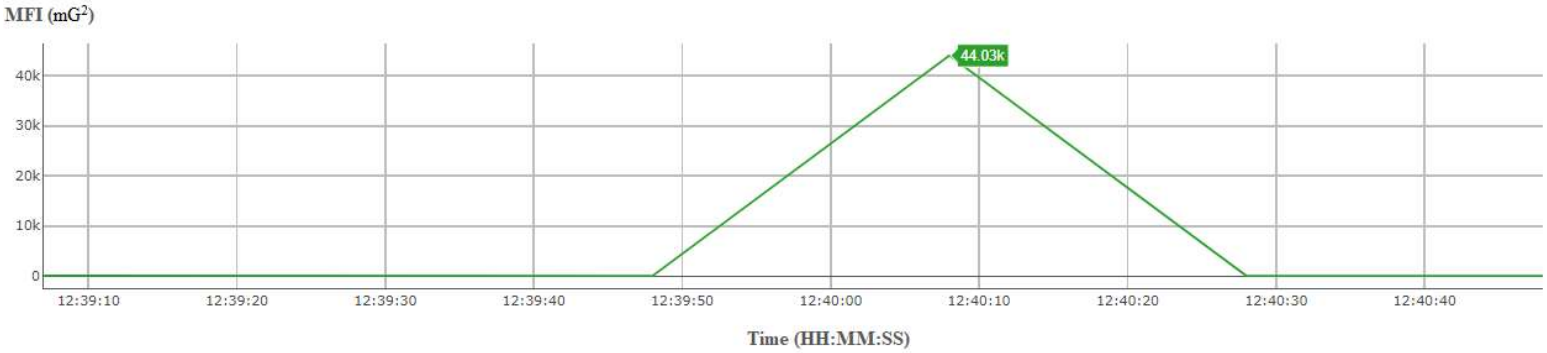


Figure 4.10: Variation of magnetic field while a private car enters and exits the detection area

Other types of vehicles, i.e., medium goods vehicles, referring to Figure 4.11, and motorcycles, referring to Figure 4.12, were tested. However, as this test was carried in a private area, i.e., the Hong Kong Institute of Vocational Education (Sha Tin), only approved vehicles were allowed to enter into the site; hence, not all vehicle types could be tested.

Figure 4.11: Medium goods vehicle for geomagnetic sensor testing

Figure 4.12: Motorcycle for geomagnetic sensor testing

4.2 Data consolidation and result simulation

Several executions are performing in slightly different situations, such as different positions on the ground, different angles of moving direction and different moving speeds. The model of each type of vehicles is the same for each unit tests.

4.2.1 Data summary

Table 4.1: MFI (in mG^2) measured from the geomagnetic sensor

Type of vehicles	No. of execution	Min.	Max.	Avg.	Std. Dev.	Median
Motorcycles	30	1356	19866	11063	3926	10683
Private cars	30	6779	57266	32406	15837	28696
Medium goods vehicles	30	63432	65535	65351	558	65535

*The maximum measuring value of geomagnetic sensor is 65535.

Table 4.2: Length (in m) deduced by data measured from the camera and the LiDAR

Type of vehicles	No. of execution	Min.	Max.	Avg.	Std. Dev.	Median
Private cars	30	4.4	5.4	4.69	0.207	4.7
Taxis	30	4.1	4.9	4.68	0.149	4.7
Public light buses	30	6.5	7.3	6.89	0.205	6.9
Private light buses	30	6.5	7.4	6.9	0.213	6.9
Light goods vehicles (Van)	30	4.7	5.2	4.89	0.169	4.9
Medium goods vehicles	30	5.7	6.5	5.99	0.168	6
Heavy goods vehicles	30	10	11.2	10.85	0.236	10.9
Single-decked buses	30	11.7	12.8	11.99	0.234	12
Double-decked buses	30	11.5	12.6	11.96	0.258	12

Table 4.3: Height (in m) deduced by data measured from the camera and the LiDAR

Type of vehicles	No. of execution	Min.	Max.	Avg.	Std. Dev.	Median
Private cars	30	1.5	1.8	1.59	0.081	1.6
Taxis	30	1.5	1.8	1.61	0.085	1.6
Public light buses	30	2.5	2.8	2.68	0.083	2.7
Private light buses	30	2.6	2.8	2.69	0.068	2.7
Light goods vehicles (Van)	30	1.9	2.1	2	0.071	2
Medium goods vehicles	30	2.5	2.9	2.69	0.098	2.7
Heavy goods vehicles	30	4	4.2	4.08	0.075	4.1
Single-decked buses	30	3.1	3.3	3.2	0.068	3.2
Double-decked buses	30	4.2	4.5	4.39	0.08	4.4

Table 4.4: Performance of a vision task by data from the camera

Function	Item description	No. of execution	No. of correct	Correct rate (%)
Word identification	'TAXI'	30	24	80
	'PUBLIC' or 'LIGHT'	30	26	86.7
	No words / other words	240	240	100
YOLO object detection	Goods vehicles (excluding vans)	60	59	98.3
	Motorcycles	30	30	100
	Buses*	120	117	97.5
	Other vehicles	90	60	66.7
Wheel detection	Detected at least 2 but not more than 3 axles for goods vehicles	30	19	63.3
	Detected at least 4 axles for goods vehicles	30	11	36.7

*Including public light buses, private light bus, single-decked buses and double-decked buses

4.2.2 Overall simulated classification result

The above profiling data is connected separately according to each sensor. After recoding all the data, it is entered into a classification program and processed according to the workflow mentioned in Section 3.8. Table 4.5 shows the classification result.

Table 4.5: Overall classification performance

Classification type	No. of execution	No. of correct	Correct rate (%)	Overall correct rate (%)
1) Motorcycles	30	30	100	86
2) Private cars	30	30	100	
3) Taxis	30	24	80	
4) Public light buses	30	26	86.7	
5) Private light buses	30	30	100	
6) Light goods vehicles (Van)	30	28	93.3	
7) Medium goods vehicles	30	19	63.3	
8) Heavy goods vehicles	30	11	36.7	
9) Single-decked buses	30	30	100	
10) Double-decked buses	30	30	100	

4.3 Result analysis and limitation

4.3.1 Different vehicle types with the same vehicular model

The main limitation of the proposed profiling system is to classify some vehicles of the same model but different vehicle types. This problem occurs between taxi and private cars as mentioned in the previous sections, but the body of the taxi has the word 'TAXI' marked on it for identification. However, the same problem can also occur between a van (LGV) and a private car. According to Cap. 374A Road Traffic (Construction and Maintenance of Vehicles) (https://www.elegislation.gov.hk/hk/cap374A) Regulations 81A, a van should be followed below regulations.

Protective partitions inside vans

(1) There shall—

(a) be in every van, permanently, between the goods compartment and the seating accommodation, a partition, which shall be of such height, width, strength, design and construction as to be capable of—

(i) withstanding or bearing the weight of any load the partition is likely to encounter or the van is likely to carry; and

(ii) acting as a protective barrier against goods in the goods compartment shifting into the seating accommodation; and

(b) not be any side windows in the goods compartment of any van.

(2) In this regulation, "van" means a light goods vehicle constructed with a fully enclosed body which is an integral part of the vehicle.

(3) This regulation shall apply—

(a) to a van registered on or after 1 July 1990, as from the date of such registration; and

(b) to a van registered before 1 July 1990, as from 1 July 1991.

Figure 4.13: A private car with model Toyota HiAce

Figure 4.14: A van with model Toyota HiAce

Figures 4.13 and 4.14 show a private car and a van with the same model – Toyota HiAce. Their shapes and dimensions are both similar or even identical. The main difference is the existence or no existence of the side window. In this case, the proposed system is not able to distinguish between them. Further improvement should be focused on identifying the existence of the side windows.

4.3.2 Different types with similar features

In fact, there is a light goods vehicle that is very similar to a medium goods vehicle – 5.5-tonne goods vehicle and 9-tonne goods vehicle, respectively. Their features like size, appearance, number of axles, etc. are nearly the same, and hence it is very difficult to

distinguish between them. According to Road Traffic (Construction and Maintenance of Vehicles) (https://www.elegislation.gov.hk/hk/cap374A) Regulations 40A.

Sideguards

(1) Subject to paragraph (2), this regulation shall apply to—

(a) a goods vehicle which has a permitted gross vehicle weight exceeding 5.5 tonnes and the distance between the centres of any 2 consecutive axles exceeds 3 metres;

(b) ...

(2) ...

(3) Every vehicle to which this regulation applies shall be fitted with sideguards constructed and fitted to meet the following requirements —

(a) ...

(4) The requirements in paragraph (3) shall apply only as far as is practicable in the case of a vehicle fitted with a tank for carrying fluid which is provided with valves and hose or pipe connections for loading and unloading and in the case of a vehicle with extending stabilisers required for stability during loading, unloading or while used for operations for which it is designed or adapted.

(5) Every sideguard fitted to a vehicle in pursuance of this regulation shall at all times while the vehicle is used on a road be maintained free from any defect which might in any way affect its effectiveness.

Figure 4.15: Sideguard installed in a goods vehicle

All 9 tonne goods vehicles must install sideguards, as shown in Figure 4.15. This can be one main feature that must exist for medium goods vehicles. On the other hand, drivers of 5.5 tonne goods vehicle can also install sideguards although it is not mandatory. In practice, besides the safety issue, some drivers of 5.5 tonne goods vehicle want to pretend that their vehicles are 5.5 tonne goods vehicles in order to escape the prosecution of overload. In this case, the detection of safeguards is not a valid distinguishing characteristic. Therefore, the distinguishing between light goods vehicles and medium goods vehicles cannot be done by observing the appearance. The division of goods vehicle is based on its maximum weight including loading, not the dimension.

4.3.3 Improvement and exceptional case in axles counting

The result of axles counting shown in Chapter 4 is not very good. It is because the edge finding has less tolerance to handle different situations. For example, the image of wheel is interfered by shadow and the whole wheel is in black colour as shown as Figure 4.16. It sometimes fails to indicate the wheel as a circle. To improve this, more filters are needed or the deep learning algorithm should be used instead. Besides, some mechanical sensors can be

used for axles counting, such as piezoelectric detectors and load cell. Further investigation on this topic is needed.

Figure 4.16: Black wheels with shadow

In Hong Kong tunnel service, besides the type of vehicle, an additional charge is needed for each axle if the vehicle has more than two axles. In fact, some goods vehicles are raising their last axle, and the corresponding wheels do not touch the ground if they are not carrying any goods. In practice, the raising axles are not counted as valid axles, and hence, there is no need to pay an additional charge, although no law or instruction about this practice is mentioned.

For the proposed system, as the raising axle is not aligned with other axles in horizontal, its wheel will be discerned as a circle object but not a wheel. An improvement can be made to identify the raising axles if needed.

4.3.4 Selection and application of LiDAR

Before embarking on an in-depth investigation about the practicality of the profiling system and finding an authorized place to test the entire system, a cheaper LiDAR with a shorter detecting range is used for concept development in this project. In fact, a long-range LiDAR is required for vehicle profiling, as the dimension of vehicles is in metres. The FoV of LiDAR, as well as camera, should be considered with the possible installation position, which can check by equation 3.5 and 3.6. Based on the above calculation, a possible LiDAR is the Velodyne Veladome, in which HFoV and VFoV are both 180°, is one of the possible LiDAR for future development and testing, and its pricing is around US$4,000 per unit.

In addition, a further development can be also focusing on detecting the speed of a vehicle by using the LiDAR, and the dimension of the vehicle can be calculated by direct measuring and speed of vehicle. In this case, the HFoV of the LiDAR does not need to cover the whole vehicle, and hence the LiDAR can place closer to the lane and it can enhance the flexibility of installation and the selection of the LiDAR.

4.4 Qualitative comparison with other methods

In this section, the proposed method is compared with methods of those mentioned in Chapter 2. The 'SICK method' is the product sold by the SICK company (https://www.sick.com/ag/en/system-solutions/profiling-systems/free-flow-profiler/c/g442952.)

Number of classification types

There are many types of vehicles that a good profiling system should be able to classify based on the profiling data. If only a few types can be classified, the market value will be much lower, and the system may not be viable.

Table 4. 6: Description of the number of classification types for different methods

Method	Description
Proposed	Together with point cloud, magnetic field and images, ten of the vehicles can be classified, and potentially for all vehicle types.
SICK	The three dimensions are going to be detected, and hence it can deduce most vehicle. However, it fails to classify different vehicle types with same model.
Axle-based	It has a limited classification performance since it uses the axle information and most vehicles are two-axles.
Sound-based	It has a limited classification performance since it uses audio information and the variation of sound signal is too larger even for the same vehicle type with a different model or producer.
Magnetic-based	The detection of magnetic field mainly acquires the size of the vehicle, it is hard to distinguish types of vehicles with similar sizes.
Video-based	Traditional algorithm is limiting in classification performance; it is hard to define all the common features of all the vehicle types. With the new deep learning algorithm, there is a potential to classify all the vehicle types if the training data set is huge enough in the future.

Classification accuracy

The classification accuracy is different with profiling accuracy. Besides the preciseness of the profiling data, the type of data also affects the classification accuracy.

Table 4. 7: Description of classification accuracy of different methods

Method	Description
Proposed	Although the profiling accuracy is not the main concern, it can profile several data, including dimension, number of axles, change in magnetic flux density and text on the vehicle, which can increase the classification accuracy.
SICK	A relative precise three dimension is detected, and hence it can result in a high classification accuracy.
Axle-based	As it is a simple counting method, the classification accuracy is very high.
Sound-based	The noise of data is significant; it affects the classification result a lot.
Magnetic-based	In the normal situation, the change in the magnetic field is directly proportional to its size, so the classification is solid.
Video-based	The image contains many pieces of useless information such as background and noise, it affects the classification performance.

Manufacturing cost

The manufacturing cost includes all the expenditure used in assembling the whole system and installing it in the working position. Sometimes, even though the system hardware cost is low, the mounting cost may be high and hence the total cost becomes higher. In addition, the maintenance cost also counts as a potential manufacturing cost. A cost-effective system should also consider its lifecycle.

Table 4. 8: Description of manufacturing cost for different methods

Method	Description
Proposed	A LiDAR camera and a geomagnetic sensor with speed bump are used.
SICK	Three 2D LiDARs are used. However, a gantry is needed so that the sensors can be mounted on a suitable place to work properly.
Axle-based	Several sensors can be used for detecting axle, such as piezoelectric detectors, load cell, capacitance mat or optical fiber. Actually, they are all mechanical sensors, and hence the lifecycle may vary with the working environment.
Sound-based	The main component used is only several directional microphones, which is relative cheap components.
Magnetic-based	Only geomagnetic sensors are used. These sensors can be placed either under the lane or inside the speed bump, and hence its expenditure is not high.
Video-based	A high-quality camera is needed; the system is based on frames recoded to perform the classification. Nowadays, the pricing of camera is not high even it has a good quality.

Flexibility of installation

Not all locations are suitable to install the profiling system due to the environment or human aspects. The size and complexities and working performance of the profiling system are also of concern for installation.

Table 4. 9: Description of flexibility of installation for different methods

Method	Description
Proposed	A speed bump is needed, but some users are resisting it. The LiDAR camera should be placed on one side of the lane with a certain distance.
SICK	There should be enough space and affordable infrastructure to install a gantry; otherwise, the sensors cannot be mounted onto suitable regions for profiling.
Axle-based	A flat sensor is placed on the road, so that its installation not complicated.
Sound-based	The installation requires finding a suitable location for sensors; otherwise, it fails to collect valid data.
Magnetic-based	The geomagnetic sensors should be placed into speed bumps or dug into the ground.
Video-based	Only a camera is placed in a particular region, so its installation is very flexible.

System processing time

The system processing time is the time required for analysing the collected data. Normally, a big size of point cloud data requires a longer processing time.

Table 4. 10: Description of system processing time for different methods

Method	Description
Proposed	It needs to process point cloud, images and digital data, so the processing time is relatively higher.
SICK	Although only point cloud data needs to be handled, its size is very big as three LiDAR are used. The processing time is still significant.
Axle-based	A simple digital count and classification mapping occurs in this method, so the processing time is very fast.
Sound-based	Only audio signal is processed. It is relatively fast.
Magnetic-based	Only magnetic field signal is processed. It is relatively fast.
Video-based	Only the image is processed. It is relatively fast.

Vehicle speed allowed

A user-friendly system should allow the user to employ the system in any condition. The vehicle speed is one of the main concerns.

Table 4. 11: Description of the vehicle speed allowed for different methods

Method	Description
Proposed	One of the main profiling is text detection. The vehicle should not be too fast; otherwise, the camera cannot capture a clear image to identify the words due to technological limitations.
SICK	The working principle is based on dimension measuring and speed detecting. A relatively high speed is allowed.
Axle-based	The sensor detection is a simple physical touch, so it allows a high-speed vehicle.
Sound-based	The speed of the vehicle also affects the sound made by the vehicle. To reduce this variation, the difference of speed between the vehicles should not be too large.
Magnetic-based	The speed allowed depends on the sensor's sampling speed. A good quality sensor can be used for high-speed detection.
Video-based	The image should be clear enough to find out the desired features, and it depends on the requirement of preciseness. Rolling shutter might occur if the vehicle is moving too fast. This causes distortion and hence affects the accuracy.

Table 5.7 shows Pugh's selection matrix (Pugh, 1981) being used to compare the quality of the different methods. As this research aims to develop a marketable system, its performance plays a critical part in it. Therefore, the number of classification types and the classification accuracy are the most concerning criteria. Next, the cost of manufacturing and installation should be considered. The processing time and the vehicle speed allowed are less significant. Based on the above criteria weights, it can be observed from the table that the proposed method in this research is the better option compared to the others.

Table 4. 12: Pugh's selection matrix for different methods

Criteria	Weight	Method					
		Proposed	SICK	Axle-based	sound-based	Magnetic-based	Video-based
Number of classification types	3	1	1	-1	-1	-1	0
Classification accuracy	3	1	1	1	-1	1	0
Manufacturing cost	2	0	-1	0	1	1	1
Flexibility of installation	2	0	-1	0	0	0	1
System processing time	1	-1	-1	1	1	1	0
Vehicle speed allowed	1	-1	0	1	-1	0	-1
	Total	4	1	3	-4	3	3

Chapter 5 Conclusion

5 Conclusion

A novel vehicle profiling system with multiple sensors is proposed in this project, which can reduce the reliance on any one of the sensors, so each sensor can compensate to make the system more functional and robust. Besides, to explore the potential application of the profiling system, it is better to profile several characteristics of the vehicles, such as change in magnetic flux density, dimension, number of axles and text on vehicles' body. For future development, colour, sound pattern, temperature, shape can be also considered.

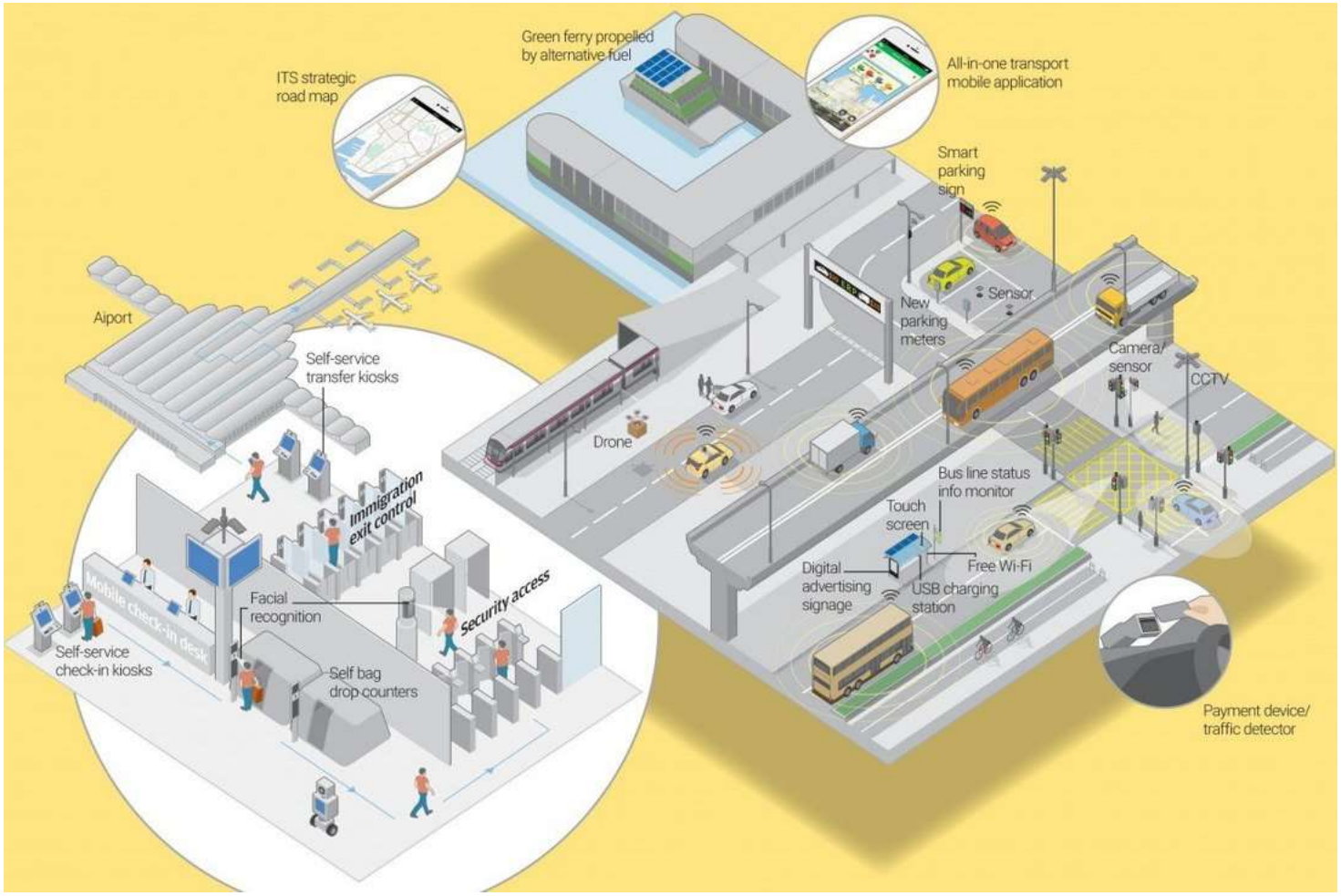


Figure 5.1: The smart mobility segment of Hong Kong's smart city initiative includes the development of an intelligent transport system
(Deng & Perez, 2020)

Together with other systems, an intelligent transport system can be assembled as shown in Figure 5.1. As the profiling system can classify the type of vehicle, the classification result can further send information to other sub-systems via a cloud server for usage other than the

tunnel toll auto-charging service. For example, whether or not there is a taxi or other public transport in a particular area can be checked, and to see what types of passengers are coming across certain areas so that a corresponding service can be designed. Businesses can use it to track consumer patterns or learn what advertisements to display according to the type of passenger. To give precise insights that result in safer cars, smarter homes, and optimized enterprises, an IoT strategy requires a connection to many other data-gathering devices, and hence, an intelligent transport system is a big part of the AI industry. An all-round profiling system can collect much more data and store it as big data to do serve above purpose.

References

Abdelkader, A., Saeed, A., Harras, K. A., & Mohamed, A. (2015, August). The Inapproximability of Illuminating Polygons by α-Floodlights. In *CCCG*. http://cccg.ca/proceedings/2015/25.pdf

Alkhatib, A. A., Hnaif, A., & Kanan, T. (2019). Proposed simple system for road traffic counting. *International Journal of Sensors Wireless Communications and Control*, *9*(2), 269-277. doi : 10.2174/2210327908666181107110441

Alkheder, S. A., Sabouni, R., El Naggar, H., & Sabouni, A. R. (2013). Driver and vehicle type parameters' contribution to traffic safety in UAE. *Journal of Transport Literature*, *7*, 403-430. doi: 10.1590/S2238-10312013000200021

Altshuller, Shulyak, L., & Rodman, S. (2005). *40 principles : TRIZ keys to technical innovation* (Extended ed..). Technical Innovation Center.

Altshuller, G. S. (1999). *The innovation algorithm: TRIZ, systematic innovation and technical creativity*. Technical innovation center, Inc.

Angelopoulou, E., & Wright Jr, J. R. (1999). Laser scanner technology. University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-99-16. https://repository.upenn.edu/cis_reports/74/

Benjdira, B., Khursheed, T., Koubaa, A., Ammar, A., & Ouni, K. (2019, February). Car detection using unmanned aerial vehicles: Comparison between faster r-cnn and yolov3. In *2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS)* (pp. 1-6). IEEE. 10.1109/UVS.2019.8658300

Bradski, G. (2000). The OpenCV library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, *25*(11), 120-125.

Brownlee. J, (2020, August 14). *What is Deep Learning?* Machine Learning Mastery. Retrieved on June 6, 2021 from https://machinelearningmastery.com/what-is-deep-learning/

Bugdol, M., Segiet, Z., Kręcichwost, M., & Kasperek, P. (2014). Vehicle detection system using magnetic sensors. *Transport problems*, *9*(1). 49-60. http://transportproblems.polsl.pl/pl/Archiwum/2014/zeszyt1/2014t9z1_06.pdf

Chvátal, V. (1975). A combinatorial theorem in plane geometry. *Journal of Combinatorial Theory, Series B*, *18*(1), 39-41. https://www.sciencedirect.com/science/article/pii/0095895675900611

Coifman, B. A., & Lee, H. (2011). LiDAR based vehicle classification. https://www.purdue.edu/discoverypark/nextrans/assets/pdfs/123OSUY2.1_Summary%20and%20Final%20Technical%20Report.pdf

Deng, I. & Perez, B. (2020, June 8). *How Hong Kong's smart mobility ambitions will be boosted by 5G roll-out.* South China Morning Post. https://www.scmp.com/tech/innovation/article/3087927/how-hong-kongs-smart-mobility-ambitions-will-be-boosted-roll-out-5g

Florez, S. A. R. (2010). *Contributions by vision systems to multi-sensor object localization*

and tracking for intelligent vehicles (Doctoral dissertation, Université de Technologie de Compiègne).
https://tel.archives-ouvertes.fr/file/index/docid/635330/filename/thesis.pdf

Garg, R., Bg, V. K., Carneiro, G., & Reid, I. (2016, October). Unsupervised cnn for single view depth estimation: Geometry to the rescue. In *European conference on computer vision* (pp. 740-756). Springer, Cham.
https://arxiv.org/pdf/1603.04992.pdf

Godard, C., Mac Aodha, O., & Brostow, G. J. (2017). Unsupervised monocular depth estimation with left-right consistency. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 270-279).
https://openaccess.thecvf.com/content_cvpr_2017/papers/Godard_Unsupervised_Monocular_Depth_CVPR_2017_paper.pdf

Goldman, M., Hassner, T., & Avidan, S. (2019). Learn stereo, infer mono: Siamese networks for self-supervised, monocular, depth estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 0-0).
https://arxiv.org/pdf/1905.00401.pdf

Haferkamp, M., Al-Askary, M., Dorn, D., Sliwa, B., Habel, L., Schreckenberg, M., & Wietfeld, C. (2017, June). Radio-based traffic flow detection and vehicle classification for future smart cities. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)* (pp. 1-5). IEEE, doi: 10.1109/VTCSpring.2017.8108633

Hamilton, A., Waterson, B., Cherrett, T., Robinson, A., & Snell, I. (2013). The evolution of urban traffic control: changing policy and technology. *Transportation planning and technology*, *36*(1), 24-43.
http://dx.doi.org/10.1080/03081060.2012.745318

Hirschmuller, H. (2007). Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on pattern analysis and machine intelligence*, *30*(2), 328-341. DOI: 10.1109/TPAMI.2007.1166

Ho, G. T. S., Tsang, Y. P., Wu, C. H., Wong, W. H., & Choy, K. L. (2019). A computer vision-based roadside occupation surveillance system for intelligent transport in smart cities. *Sensors*, *19*(8), 1796.
https://doi.org/10.3390/s19081796

Hu, P., Tan, J., Yang, H., Zhao, X., & Liu, S. (2011, July). Phase-shift laser range finder based on high speed and high precision phase-measuring techniques. In *The 10th International Symposium of Measurement Technology and Intelligent Instruments* (Vol. 29).
http://www.ismtii2011.org/article/xml/sub/file_download.kin?main_no=92&mode=pdf

Hughes, C., Glavin, M., Jones, E., & Denny, P. (2009). Wide-angle camera technology for automotive applications: a review. *IET Intelligent Transport Systems*, *3*(1), 19-31.
https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.140.4995&rep=rep1&type=pdf

Kaewkamnerd, S., Chinrungrueng, J., Pongthornseri, R., & Dumnin, S. (2010, June). Vehicle classification based on magnetic sensor signal. In The 2010 IEEE International Conference on Information and Automation (pp. 935-939). IEEE.
10.1109/ICINFA.2010.5512140

Keung, K. L., Lee, C. K. M., Ng, K. K. H., & Yeung, C. K. (2018, December). Smart city application and analysis: Real-time urban drainage monitoring by iot sensors: A case study of Hong Kong. In *2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)* (pp. 521-525). IEEE. https://ieeexplore.ieee.org/document/8607303

Klein, L. R. (2002). *Sensor Technologies for ITS* [PowerPoint slides]. Transportation Research Board Freeway Operations and Signal Systems Mid-Year Committee Meeting. http://www.rf114.com/lib/download.php?code=tbl_board&seq_name=bseq&seq=754

Mann, D. L. (2003). Better technology forecasting using systematic innovation methods. *Technological Forecasting and Social Change*, *70*(8), 779-795, doi: https://doi.org/10.1016/S0040-1625(02)00357-8

Mann, D. (2009). *Matrix 2010: Re-updating the TRIZ contradiction matrix*. Bidefor, UK: Ifr Press.

Martin, P. T., Feng, Y., & Wang, X. (2003). *Detector technology evaluation* (No. MPC Report No. 03-154). Fargo, ND: Mountain-Plains Consortium. https://www.ugpti.org/resources/reports/downloads/mpc03-154.pdf

Masino, J., Thumm, J., Frey, M., & Gauterin, F. (2017). Learning from the crowd: Road infrastructure monitoring system. *Journal of Traffic and Transportation Engineering (English Edition)*, *4*(5), 451-463, doi: 10.1016/j.jtte.2017.06.003

Mayer, N., Ilg, E., Hausser, P., Fischer, P., Cremers, D., Dosovitskiy, A., & Brox, T. (2016). A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4040-4048). doi: 10.1109/CVPR.2016.438

Meta, S., & Cinsdikici, M. G. (2010). Vehicle-classification algorithm based on component analysis for single-loop inductive detector. *IEEE Transactions on Vehicular Technology*, *59*(6), 2795-2805. https://ieeexplore.ieee.org/document/5460952

Nellore, K., & Hancke, G. P. (2016). A survey on urban traffic management system using wireless sensor networks. *Sensors*, *16*(2), 157. https://doi.org/10.3390/s16020157

Nooralahiyan, A. Y., Dougherty, M., McKeown, D., & Kirby, H. R. (1997). A field trial of acoustic signature analysis for vehicle classification. *Transportation Research Part C: Emerging Technologies*, *5*(3-4), 165-177. https://www.sciencedirect.com/science/article/pii/S0968090X97000119

Object detection. (2021, June 6). In *Wikipedia*. https://en.wikipedia.org/wiki/Object_detection

Pfeifer, N., & Briese, C. (2007, April). Laser scanning–principles and applications. In *GeoSiberia 2007-International Exhibition and Scientific Congress* (pp. cp-59). European Association of Geoscientists & Engineers. 10.3997/2214-4609.201403279

Pugh, S. (1981, March). Concept selection: a method that works. In *Proceedings of the*

International conference on Engineering Design (pp. 497-506).

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
10.1109/CVPR.2016.91

Sato, T., Aoki, Y., & Takebayashi, Y. (2014). Vehicle Axle Counting using Two LIDARs for Toll Collection Systems. In *ITS World Congress* (Vol. 12249).
https://www.researchgate.net/publication/290573783

Simony, M., Milzy, S., Amendey, K., & Gross, H. M. (2018). Complex-yolo: An euler-region-proposal for real-time 3d object detection on point clouds. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops* (pp. 197-209). https://rd.springer.com/book/10.1007/978-3-030-11009-3

Souchkov, V. (2015). *A brief history of TRIZ*.
http://www.xtriz.com/BriefHistoryOfTRIZ.pdf

Sun, Y., Xu, H., Wu, J., Zheng, J., & Dietrich, K. M. (2018). 3-D data processing to extract vehicle trajectories from roadside LiDAR data. *Transportation research record*, *2672*(45), 14-22.
https://journals.sagepub.com/doi/pdf/10.1177/0361198118775839

Suryatali, A., & Dharmadhikari, V. B. (2015, March). Computer vision based vehicle detection for toll collection system using embedded Linux. In *2015 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2015]* (pp. 1-7). IEEE. https://ieeexplore.ieee.org/document/7159412

Tóth, C. D. (2002). Art galleries with guards of uniform range of vision. *Computational Geometry*, *21*(3), 185-192. https://www.sciencedirect.com/science/article/pii/S0925772101000244

Transport department of HKSAR (2022, January 25). *TOLL RATES OF ROAD TUNNELS*. Retrieved on January 25, 2022 from https://www.td.gov.hk/en/transport_in_hong_kong/tunnels_and_bridges_n/toll_matters/toll_rates_of_road_tunnels_and_lantau_link/index.html

Ueda, T., Takemura, K., Ogata, S., & Yamashita, T. (1997, November). Optical axle-counting equipment. In *Proceedings of Conference on Intelligent Transportation Systems* (pp. 326-331). IEEE. https://doi.org/10.1109/ITSC.1997.660496

Urazghildiiev, I., Ragnarsson, R., Ridderstrom, P., Rydberg, A., Ojefors, E., Wallin, K., ... & Lofqvist, G. (2007). Vehicle classification based on the radar measurement of height profiles. *IEEE Transactions on intelligent transportation systems*, *8*(2), 245-253. https://ieeexplore.ieee.org/document/4220668

Werner, M. (2014). *Indoor location-based services: Prerequisites and foundations* (pp. 1-233)

Wulf, O., & Wagner, B. (2003, July). Fast 3D scanning methods for laser measurement systems.

In *International conference on control systems and computer science (CSCS14)* (pp. 2-5). https://pdfs.semanticscholar.org/633d/42113180b581e64ac04b846cb55c6b9bffd2.pdf

Xiang, W., Otto, C., & Wen, P. (2008). Automated vehicle classification system using advanced noise reduction technology. In *Proceedings of the 1st International Conference on Signal Processing and Communication Systems (ICSPCS 2007)* (pp. 413-417). DSP for Communication Systems. https://www.dspcs-witsp.com/icspcs_2007/Proceedings_ICSPCS2007.pdf

Zbontar, J., & LeCun, Y. (2016). Stereo matching by training a convolutional neural network to compare image patches. *J. Mach. Learn. Res.*, *17*(65):1-32, 2016. https://arxiv.org/pdf/1510.05970.pdf

Zhang, G., Avery, R. P., & Wang, Y. (2007). Video-based vehicle detection and classification system for real-time traffic data collection using uncalibrated video cameras. *Transportation research record*, *1993*(1), 138-147. https://journals.sagepub.com/doi/10.3141/1993-19

Zhou, Y., & Tuzel, O. (2018). Voxelnet: End-to-end learning for point cloud based 3d object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4490-4499). 10.1109/CVPR.2018.00472

Lovas, T., Toth, C. K., & Barsi, A. (2004). Model-based vehicle detection from lidar data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 134-138.

https://www.researchgate.net/publication/228963569_Model-based_vehicle_detection_from_lidar_data

Platzman, M. M. (1974). *Automatic vehicle classification and ticket issuing system* (U. S. Patent No. 3794966). U. S. Patent and Trademark Office. https://www.freepatentsonline.com/3794966.html

Bennett, N. (1988). *Computerized vehicle classification system* (U. S. Patent No. 4,789,941). U. S. Patent and Trademark Office. https://www.freepatentsonline.com/4789941.html

Kuwagaki, H. (1995). *Vehicle classification system using profile* (U. S. Patent No. 5,392,034). U. S. Patent and Trademark Office. https://www.freepatentsonline.com/5392034.html

Becker, G., Borsken, N., Gudesen, A., Klemp, J., Tummoscheit, G., Vilmar, G. (1995). *Method for classifying vehicles passing a predetermined waypoint* (U. S. Patent No. 5,446,291). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/5446291.html

Brady, M. J., Hamerly, M. E. (1997). *Vehicle classification system using a passive audio input to a neural network* (U. S. Patent No. 5,619,616). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/5619616.html

Hasselbring, R. E. (1998). *Doppler-radar based automatic vehicle-classification system* (U. S. Patent No. 5,717,390). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/5717390.html

Kohnen, K., Eaton, W. W. (2000). *Passive vehicle classification using low frequency electro-magnetic emanations* (U. S. Patent No. 6,014,447). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/6014447.html

Wangler, R. J., Myers, J. T., Gustavson, R. L., McConnell, R. E. II (2002). *Vehicle classification and axle counting sensor system and method* (U. S. Patent No. *20020140924*). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/y2002/0140924.html

Owen, D. K., Klamer, D. M. (2004). *System and method for classifying vehicles* (U. S. Patent No. 6,828,920). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/6828920.html

Bertrand, J., Dicko, M. (2005). *Method and device for classifying vehicles* (U. S. Patent No. 6,865,518). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/6865518.html

Dingwall, R., Milnick, P., Emerick, J., Jameison, G. (2005). *Systems and methods for classifying vehicles* (U. S. Patent No. 6,894,233). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/6894233.html

Devdhar, P. P. (2005). *Method for vehicle classification* (U. S. Patent No. *20050267657*). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/y2005/0267657.html

Chung, J. (2012). *Vehicle classification by image processing with laser range finder* (U. S. Patent No. 8,311,343). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/8311343.html

Lameyre, B., Jouannais, J. (2015). *Automatic classification system for motor vehicles* (U. S. Patent No. *20150269444*). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/y2015/0269444.html

Bhanu, B., Thakoor, N. (2016). *Method and system for vehicle classification* (U. S. Patent No. 9,239,955). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/9239955.html

Neuman, J. W. (2016). *Vehicle classification system and method* (U. S. Patent No. 9,361,798). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/9361798.html

Bhanu, B., Kafai, M. (2016). *Dynamic Bayesian Networks for vehicle classification in video* (U. S. Patent No. 9,466,000). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/9466000.html

Khan, S. M., Cheng, H., Matthies, D. L., Sawhney, H. S., Jung, S. H., Broaddus, C., Matei, B. C. M., Divakaran, A. (2018). *3-D model based method for detecting and classifying vehicles in aerial imagery* (U. S. Patent No. 9,977,972). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/9977972.html

Kwon, Y. J., Kim, D. H., Lee, S. J. (2018). *Vehicle classification system and method* (U. S. Patent No. 9,984,704). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/9984704.html

Salti, S., Sambo, F., Taccari, L., Bravi, L., Simoncini, M., Lori, A. (2019). *Vehicle classification using a recurrent neural network (RNN)* (U. S. Patent No. 5,392,034). U. S. Patent and Trademark Office. http://www.freepatentsonline.com/10345449.html