AERIAL BASED TRAFFIC MONITORING AND VEHICLE COUNT DETECTION USING BACKGROUND SUBTRACTION

MUHAMAD ZULHILMI BIN MUHAMAD

SCHOOL OF AEROSPACE ENGINEERING
UNIVERSITI SAINS MALAYSIA

2021

AERIAL BASED TRAFFIC MONITORING AND VEHICLE COUNT DETECTION USING BACKGROUND SUBTRACTION

by

MUHAMAD ZULHILMI BIN MUHAMAD

Thesis submitted in fulfilment of the requirements for the Bachelor degree of Engineering (Honours) (Aerospace Engineering)

July 2021

ENDORSEMENT

I, Muhamad Zulhilmi Bin Muhamad hereby declare that all corrections and comments made by the supervisor and examiner have been taken consideration and rectified accordingly.

(Signature of Student)

Date: 10 July 2021

Assoc. Prof. Dr. ELMI BIN ABU BAKAR
Lecturer & Researcher
Room 1.25, School of Aerospace Engineering,
Engineering Campus, Universiti Sains Malaysia,
14300 Nibong Tebal, Penang
Tel : +604-599 5954 / 5933 Fax : +604 599 6911

(Signature of Supervisor)

Name: Assoc. Prof. Dr Elmi Bin Abu Bakar

Date: 10 July 2021

DR. HO HANN WOEI
Pensyarah
Pusat Pengajian Kejuruteraan Aeroangkasa
Kampus Kejuruteraan
Universiti Sains Malaysia
14300 Nibong Tebal, Penang
Tel: +604 599 5870 Fax: +604 599 6911

(Signature of Examiner)

Name: Dr Ho Hann Woei

Date: 10 July 2021

DECLARATION

This thesis is the result of my own investigation, except where otherwise stated and has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any other degree.

(Signature of Student)

Date: 10 July 2021

ACKNOWLEDGEMENT

First and foremost, I am using this opportunity to express my deepest gratitude to my thesis advisor, Associate Professor Dr Elmi Abu Bakar and Dr Muhammad Nishat Akhtar, who has supported me with their continuous patience and immense knowledge throughout this study. Without their constant encouragement and valuable guidance this thesis would not have been completed or written. I am also indebted to my examiner, Dr Ho Hann Woei for his comments and suggestions which were very precious for this research work.

I would like to thank to my family members who always play a role in giving a support for me to finish this thesis. Also, I would like to thank all my friends that has been continuously support throughout this project. Without each one supports I would not be able to do my project alone. Their advice is really helpful.

AERIAL BASED TRAFFIC MONITORING AND VEHICLE COUNT DETECTION USING BACKGROUND SUBTRACTION

ABSTRACT

Increasing population in an area of the world increasingly increases the density of the area. This happened by an increase in vehicle volume resulting in congestion. This project uses Python as its programming language and OpenCV as an open-source library for programming, and Raspberry Pi. The objective of this study was to develop a vision-based system for road vehicle counting and tracking. The system will be able to achieve counting with very good accuracy even in difficult scenarios related to occlusions or the presence of shadows. The principle of the system is to install a camera on the pedestrian bridges and track the vehicular traffic congestion by incorporating a unique ID. Moving objects were tracked using simple background subtraction and moving object monitoring was conducted using the MOSSE (Minimum Output Sum of Squared Error) tracker. The video processing model is combined with a motion detection procedure, which correctly allows the positioning of moving vehicles depending on the space and time when the experiment was conducted. More trials need to be carried out comprising of peak periods and different vehicle types, and occlusions need to be observed between close moving vehicles and between cars and heavy vehicles. Using the proposed method, the identification of severe shadows based on solidity can be calculated through the nature of the shape and this classification allows its accuracy to be estimated.

AERIAL BASED TRAFFIC MONITORING AND VEHICLE COUNT DETECTION USING BACKGROUND SUBTRACTION

ABSTRAK

Pertambahan penduduk diserata dunia yang semakin meningkat secara langsung berhubungkait kepadatan di sesuatu kawasan. Malah, fenomena ini berlaku menyumbang kepada peningkatan jumlah kenderaan yang mengakibatkan kesesakan berlaku di jalanraya. Kajian projek ini menggunakan *Python* sebagai Bahasa pengaturcaraannya dan *OpenCV* sebagai sumber perpustakaan terbuka untuk pengaturcaraan, dan *Raspberry Pi*. Objektif kajian ini untuk mengembangkan sistem berasaskan penglihatan bagi mengira dan mengesan kenderaan di jalan raya. Sistem ini dapat mencapai kiraan dengan ketepatan yang sangat baik walaupun dalam senario sukar berkaitan dengan kehadiran bayang – bayang. Prinsip asas sistem ini adalah memasang kamera di jejantas pejalan kaki dan mengesan kesibukan lalulintas kenderaan melalui penggunaan ID unik. Objek bergerak dapat di jejaki dengan menggunakan teknik pengurangan latar belakang yang sederhana dan pemantauan objek bergerak dilakukan menggunakan penjejak *MOSSE* (Jumlah Output Minimum Ralat Kuadrat). Model pemprosesan video digabungkan dengan prosedur pengesanan gerakan, dengan betul memungkinkan kedudukan posisi kenderaan bergerak bergantung kepada ruang dan waktu ketika ujikaji dijalankan. Lebih banyak percubaan perlu dilakukan yang terdiri dari masa puncak dan jenis kenderaan yang berlainan, dan bayang-bayang perlu diperhatikan antara kenderaan bergerak dekat dan antara kereta dan kenderaan berat. Dengan menggunakan kaedah yang dicadangkan, pengenalan bayang-bayang teruk berdasarkan kekukuhan dapat dikira melalui sifat bentuk dan pengkelasan ini membolehkan ketepatannya dapat dianggarkan.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS

k	Equivalent to the number of clusters
2^n	Dimension of pixels frame (n=1,2,3,…)

LIST OF ABBREVIATIONS

OpenCV	Open-source Computer Vision
ID	Identity Document
MOSSE	Minimum Output Sum of Squared Error
KNN	K Nearest Neighbor
MOG	Mixtures Of Gaussian
MOG2	Mixtures Of Gaussian 2
GMG	Geometric Multi-Grid
UAV	Unmanned Aerial Vehicle
RTM	Road Traffic Monitoring
HD	High Definition
CVF	Camera’s View Field
LTS	Long-Term Support
PSR	Peak-to-Sidelobe Ratio
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transformation
PM	Post Meridiem
FPS	Frames Per Second
SIFT	Scale Invariant Feature Transform
SURF	Speeded Up Robust Features
CNN	Convolutional Networks
SUV	Sport Utility Vehicle
3D	3-Dimensional

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Research Background

Object recognition is a computer technology that deals with identifying instances of semantic objects of a certain class in digital images and videos. It is related to computer vision and image processing. Face detection and pedestrian detection are two well-studied object detection domains. Object recognition can be used in a variety of computer vision applications, including image extraction and video surveillance (Bhaskar and Yong, 2014). Detecting the actual movement of an object in a given location or area is known as moving object detection. The motion of moving objects could be monitored and analysed later by using segmentation between moving objects and stationary areas or regions. Consider a video as a structure made up of single frames, moving object detection is the process of locating foreground moving target(s) in each video frame or only when the moving target makes its first appearance in the video (Huttenlocher, 2004).

Computer vision aims to automate functions that can be performed by the human visual system. It entails the development of a theoretical and algorithmic foundation for automatic visual comprehension. Video loops, multiple camera views, or multi-dimensional data from a medical scanner are all examples of image data. Computer vision is a scientific discipline that studies the principle behind artificial systems that extract knowledge from images. Computer vision, as a technological discipline, aims to apply its theories and models to the creation of computer vision systems (Rosebrock, 2017).

OpenCV was developed to provide computing resources for computer vision applications and to help commercial products incorporate machine perception more quickly. The library is widely used by companies, academic organisations, and government agencies (Uke and Thool, 2013). As a part of computer vision applications, it is built to provide a common infrastructure to accelerate in the commercial products of the use of machine perception. The library's optimized algorithms comprise more than 2500 and includes a comprehensive collection of computer vision and machine learning algorithms, both classic and state-of-the-art. The used of these algorithms is to detect and recognize vehicles, identify objects and many more.

Unmanned Aerial Vehicles (UAVs), also known as drones, are used in an increasing number of civil and commercial applications. Among these applications, Road Traffic Monitoring (RTM) systems constitute a domain where the use of UAVs is receiving significant interest. Unmanned Aerial Vehicle (UAVs) are becoming an attractive solution for road traffic monitoring because of their mobility, low, cost and broad view range (Khan *et al.*, 2020). Up to now, existing traffic monitoring systems based on UAVs only use UAV with a fixed trajectory to extract information about vehicles.

Aerial imagery should be a tool for topographical mapping and interpreting places, objects, and behaviour. Traditional asset inspection methods need time-consuming and costly operations. The utilisation of aerial imagery has made remote asset monitoring possible. High-resolution aerial imagery enhances infrastructure management while closely monitoring assets over time, from public utilities to private facilities. Although drones have the ability to conduct traffic surveillance more efficiently, the aerial image through a static standstill system is considered in this work.

1.2 Problem Statement

Rising traffic flow is a significant problem in cities and efficient traffic congestion is an urgent issue. Traffic monitoring with the fixed camera is getting increasingly inefficient as they cannot identify issues beyond their immediate location (Kamble and Kounte, 2020).

Urban planners need to analyse traffic density, road capacity and traffic flow to draw strategies to reduce urban congestion. This will optimize traffic flow, cut fuel use, and may help to tackle urban environmental issues.

Traffic monitoring via aerial can overcome the limitations of traditional methods of monitoring due to its simplicity, mobility and ability to cover large areas. High-resolution real-time videos from aerial can be relayed to the command and control centre to assist on-ground personnel in road monitoring, traffic guidance, traffic activity analysis, identify and track individual vehicles, read the license plate and more. Aerial analysis can be equipped with different type of payloads like HD camera and thermal camera for day and night surveillance (Gleason *et al.*, 2011).

The aerial analysis can provide on-ground situational awareness in emergencies like road accidents, oil leaks and also collect evidence for the same. The data collected by aerial can be analysed to improve traffic flow and road safety.

1.3 Research Objective

A research study designed to implement and assess the performance of the proposed technique involves the following research objectives:

I. To develop a unique algorithm for vehicle detection using selected background subtraction method (KNN, MOG, MOG2 and GMG)

II. To address the issue of detecting vehicle from video frames and compare the algorithm used performance in a specific location, environment and time

1.4 Research Scope

The current research aims to develop an automatic vehicle counting system that can process videos recorded from stationary cameras installed over roads, such as cameras installed on pedestrian bridges or near traffic intersections, and count the number of vehicles passing a spot in a specific time for further vehicle data collection. To solve the problem, a basic solution based on background subtraction approaches was used (KNN, MOG, MOG2 and GMG). A narrow region, a line, in the video frames that has been developed as an indicator for the vehicle that passes this narrow region will be counted and counted based on a unique ID.

CHAPTER 2

LITERATURE REVIEW

2.1 Generalities of Aerial-Based Vehicle Monitoring

2.1.1 Challenges of Aerial-Based Vehicle Monitoring

There are numerous difficulties with vision-based vehicle monitoring. The process of monitoring may be influenced by the sensor's quality. First and foremost, the sensor's quality (noise, vibration, video format and resolution, occlusion, colour representation, processing power available) may have an impact on the monitoring process. There are significant context-based initialization and stages which is camera calibration, picture region of interest determination, initial maps (Datondji *et al.*, 2016).

2.1.2 Sensing Technologies for Monitoring

Cameras are commonly applied at bridges to accurately detect and give detailed visual data. Visible light cameras can be utilised for daytime activities. However, at night or in inclement weather, a visible light camera is unlikely to meet long-term performance requirements (Koutsia *et al.*, 2008). Because the temperature of the vehicle's tyres is higher than the isotherm level at night, by use of infrared cameras can be a viable alternative for night-time vision.

Camera networks can provide numerous benefits, including obtaining better data, resolving occlusion difficulties, and providing redundancy. However, while installing a network, essential factors like mobility, power consumption, and mandatory spatial and temporal calibration must be considered. Camera networks are seldom utilised in the setting of junctions (Datondji *et al.*, 2016).

2.2 Existing Traffic Count Available Today

2.2.1 Manual Counts

People manually counting vehicles is a basic yet accurate technique of traffic counting. A person either uses an electronic hand-held counter or a tally sheet to record data. They may stand by the roadside or, more often, watch a video of the road and count from that. During testing, manual vehicle counting was 99% accurate over the course of the counting period. Manual counts gather a short sample of data, usually in less than a day, and the findings are extrapolated for the remainder of the year or season. This is when mistakes enter the picture because a small sample is rarely perfectly indicative of the full year (Zheng and Mike, 2012).

2.2.2 Vehicle Detection using Computer Vision

There are now systems available that will automatically analyse video images as vehicles pass by cameras, identifying vehicles with the same precision as individuals watching the footage (Pal'o *et al.*, 2019). This technique of vehicles counting has numerous benefits over other automated systems. It is cost-effective since it can count in many directions at the same time; just one camera is required for multiple lanes or exits at a junction. From an computer, it is simple to add or change the zones through which vehicles are tallied. Counts may be readily confirmed by watching the video and comparing it to automatic counts from their web browsers.

2.2.3 Pneumatic Road Tube Counting

A pair of tubes attached to a data logger may be extended over many lanes of traffic. When a set of wheels collides with the tube, the air pressure in the squashed tube triggers the data logger, which records the event's time (Puan *et al.*, 2019). Vendors claim 99% accuracy however, tests reveal that the absolute inaccuracy of a normal 15-minute count was closer to 10%.

2.2.4 Piezoelectric Sensor

Piezoelectric sensors gather information by turning mechanical energy into electrical energy (Trindade, 2008). When an automobile drives over a piezoelectric sensor, it squeezes it, causing an electric potential to be created. The magnitude of the signal varies with the degree of distortion. When the vehicle drives away, the voltage reverses and may be used to identify and count vehicles. Data may be gathered locally on a laptop through an Ethernet or RS232 connection, or it can be sent wirelessly.

2.2.5 Inductive Loop

An inductive loop is a square of wire that is placed in or beneath the road. The loop works on the concept that introducing a magnetic field near an electrical conductor induces an electrical current (Oliveira *et al.*, 2010). A huge metal truck serves as the magnetic field in the case of traffic monitoring. The produced signals are recorded by a device at the roadside.

2.2.6 Magnetic Sensor

This detector detects cars by monitoring changes in the earth's magnetic field as the vehicles pass over it (Lenz, 1990). The sensor is either buried in the road or housed in a box at the roadside. If two cars are following each other very closely, the magnetic detector may struggle to distinguish between them.

2.2.7 Acoustic Detector

This detects vehicles based on the sound they make as they pass. The sensor is placed on a pole and is aimed down towards the traffic. It can gather data for one or more traffic lanes. Some can send their counts wirelessly (Swerdlow *et al.*, 2009).

2.2.8 Passive Infrared

Passive infrared detectors detect cars by measuring the amount of infrared energy emitted by the detection zone (Student, 2014). When a car passes by, the energy emitted changes and the count rises. Slow temperature variations on the road surface caused by changing weather conditions are disregarded. Lane coverage ranges from one to two lanes.

2.2.9 Doppler and Radar Microwave Sensors

Doppler microwave detection systems send a continuous signal of low-energy microwave radiation to a specific region and then evaluate the reflected signal. When the microwave source and the vehicle move relative to one another, the detector detects a change in wave frequency. As a result, the gadget can identify moving vehicles (Balageas, Fritzen and Güemes, 2010).

Radar can detect distant objects and determine their location and speed of movement. A device focuses on high-frequency radio waves at the roadway to measure the time delay of the return signal, allowing the distance to the detected vehicle to be calculated (Balageas, Fritzen and Güemes, 2010).

2.3 Detection Occurs during the Day and at Night

Low-level image analysis modules perform vehicle extraction under the two main different illumination conditions: daylight and night. Target extraction is based on Spatio-temporal segmentation. The system is adaptive to luminance variations in time and space. The operator filters isolated spots due to the small movements of sensors and avoid de-localization of extracted points. A morphological closure is performed by following the points with a high gradient to obtain the moving object contour. All metric parameters used for vehicle extraction are scaled linearly along the main traffic direction. A moving object is classified and labelled as a vehicle if its size (in pixels) is by an initial scene calibration (Cucchiara and Piccardi, 1999).

The only salient visual features are headlights and their beams, street lamps, and street lamps. The main goal is to identify vehicles in terms of pairs of headlights, since vehicles with single headlights, such as motorbikes, can be present. The final verification is based on the correlation of headlights from the same pair, which accomplished by correlating brightness levels along the normal to the major traffic direction (Pavan Kumar and Bharathi, 2019).

2.4 Background Subtraction Method

Background subtraction is one tool for detecting a moving target. Background subtraction, also known as foreground identification, is one of the methods used in the area of image processing and computer vision to detect foreground (objects) from the background for more processing. Background subtraction is a tool widely used to track moving objects from static cameras to images (Mandellos, Keramitsoglou and Kiranoudis, 2011). The mechanism of detecting a moving target using the background subtraction approach is based on the discrepancy between the reference background and the frame. OpenCV has a BackgroundSubtractor class that may be used to do foreground and background segmentation. BackgroundSubtractor is a fully-fledged class with lots of methods that perform background subtraction and improve background detection in time through machine learning and save the classifier to a file (Piccardi, 2004).

2.4.1 Nonparametric Classification Methods

The classification method is often separated into two phases of development: training and detection. When the training period is sufficiently long, nonparametric approaches are efficient.

2.4.1(a) Principal of K-Nearest Neighbor (KNN)

The KNN method is frequently used for classification in pattern recognition and data mining. The principle of the algorithm is that if the majority of the k most similar samples to a query point in the feature space belong to a certain category, then a verdict can be made that the query point falls in this category. Similarity can be measured by the distance in the feature space, so this algorithm is called the K-Nearest Neighbor algorithm. The KNN method is easy to implement and can handle high-dimensional data sets. However, if the test set, train set, and data dimension are all bigger than predicted, the computational complexity and operation time will be enormous (Yu *et al.*, 2009).

2.4.2 Parametric Classification Methods

The most majority of moving object extraction algorithms rely on the temporal development of each pixel in the image. The detection technique entails categorising each pixel in the object or background classes individually.

2.4.2(a) Principal of Mixture of Gaussian (MOG)

Based on this, a Gaussian mixture, or MOG, was first suggested (KaewTraKulPong and Bowden, 2002) & (Stauffer and Grimson, 1999). Each backdrop pixel is modelled by a blend of k Gaussian distributions, with k values between 3 and 5. The authors presume that various distributions indicate various colour in the background and foreground. On the model, the weight of each of the utilised distributions is related to the amount of time each colour remains on that pixel. As a result, when the weight of a pixel distribution is low, that pixel is labelled as foreground.

MOG has a low rate of compatibility, complexity, and memory consumption and the ability to detect objects in an outside area. In the background subtraction method, this technique is much more adaptable and resilient, and it can handle multi-modal distributions (Mohamed, Tahir and Adnan, 2010).

2.4.2(b) Principal of Mixture of Gaussian 2 (MOG2)

The MOG2 approach was built on the works to overcome one of MOG's limitations: the fixed number of usable distributions (Zivkovic, 2004) & (Zivkovic and Van Der Heijden, 2006). MOG2 achieves a better depiction of the complexity of colours in each frame by employing a configurable number of Gaussian distributions that are mapped pixel by pixel.

MOG2, a new and updated version of MOG, utilizes the same concept as the original MOG but contains some additional capabilities. The most convenient number of Gaussian distributions is chosen for each pixel on its own. There is also the option to choose whether or not to detect shadows. It adapts well to varied settings due to variations in lighting (Zivkovic and Van Der Heijden, 2006).

2.4.2(c) Principal of Geometric Multi-Grid (GMG)

The proposed GMG method models the background using a mixture of Bayesian Inference and Kalman Filters (Godbehere and Goldberg, 2014). The first stage of the approach gathers weighted values for each pixel based on how long a colour remains in that place. New observations are added to the model for each frame, changing these values. Background colours are those that remain steady for an extended period of time. To minimise noise from the first stage, the second step filters pixels in the foreground.

2.5 Video Frame Difference Method

Using the video frame difference method, the variance is calculated according to the pixel values of two or three consecutive video frames. Moreover, the moving foreground region is separated by the threshold. By using this method and suppressing noise, the stopping of the vehicle can also be detected. When the background image in the video is fixed, the background information is used to establish the background model. Then, each frame image is compared with the background model and the moving object can also be segmented.

2.6 Optical Flow Method

The method of using optical flow can detect the motion region in the video. The generated optical flow field represents each pixel's direction of motion and pixel speed. Vehicle detection methods using vehicle features, such as the Scale Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SURF) methods, have been widely used. For example, 3D models have been used to complete vehicle detection and classification tasks. Using the correlation curves of 3D ridges on the outer surface of the vehicle, the vehicles are divided into three categories: cars, SUVs and minibuses.

2.7 Complex Deep Learning Method

The use of deep convolutional networks (CNNs) has achieved amazing success in the field of vehicle object detection. CNNs have a strong ability to learn image features and can perform multiple related tasks, such as classification and bounding box regression (Zhao *et al.*, 2019). The detection method can be generally divided into two categories. The two-stage method generates a candidate box of the object via various algorithms and then classifies the object by a convolutional neural network. The one-stage method does not generate a candidate box but directly converts the positioning problem of the object bounding box into a regression problem for processing. In the two-stage method Region-CNN (R-CNN) uses selective region search in the image. The image input to the convolutional network must be fixed-size, and the deeper structure of the network requires a long training time and consumes a large amount of storage memory. Drawing on the idea of spatial pyramid matching, SPP NET allows the network to input images of various sizes and to have fixed outputs.

2.8 Object Tracking vs Object Detection

Tracking algorithms outperform detection algorithms in terms of speed. A successful tracking algorithm would use all available information about the object up to that point. Tracking algorithms have a better understanding of the particular instance of the class they are tracking (Yin, Makris and Velastin, 2007). Tracking algorithms are prone to accumulating errors. To solve these problems with tracking algorithms, a detection algorithm is run regularly to deal with the problems with the tracking algorithm.

The tracking data can be used to predict the position of the object in the next frame. It can also lose track of an object if it is obstructed for an extended period of time or moves so quickly that the tracking algorithms cannot keep up. A good tracking algorithm will handle some level of occlusion (Cheong and Chew, 2018).

Object detection produces an array of rectangles containing the object, but there is no identity associated with the object. A detector that detects red dots outputs rectangles for all of the dots it detects in a frame. Tracking, on the other hand, allows one to simply connect the dots (Tian *et al.*, 2014).

2.9 Summary

This project, therefore, proposes designing and implementing vehicle detection using computer vision and the background subtraction method (KNN, MOG, MOG2 and GMG) will be used in this project.

CHAPTER 3

METHODOLOGY

3.1 Overview

The proposed automatic vehicle counting system uses video data received from a stationary traffic camera to estimate the number of cars present in a scene by conducting causal mathematical operations over a series of frames derived from the video. The KNN, MOG, MOG2, and GMG separate the objects in motion from the background in each frame by monitoring identified objects within a narrow region, a line, in the video frames and then counting.

3.2 Getting Started

3.2.1 Object Features

Vehicles can be classified according to their size, colour or form. The selection of an appropriate and robust vehicle feature representation about the application is a crucial question during start-up. In general, the tracking technology is defined by vehicle representation. The capability to keep vehicle tracks alive in the scene for as long as feasible is the most essential aim. Vehicle tracking must be accurate to perform posterior trajectory analysis and behaviour identification (Datondji *et al.*, 2016).

Vehicle occlusion, changes in vehicle perception, and abrupt vehicle motion are all major problems. Some approaches need the vehicle to be detected first and then tracked, whilst others use vehicle tracks as detection signals (Cheong and Chew, 2018). Other difficulties make a vehicle detection and tracking difficult. The study identified three key challenges: vehicle-vehicle, infrastructure-carrier, and shadow identification and removal.

3.2.2 Camera Set-up

In this study, the device used to record video throughout the analysis of traffic conditions was a mobile smartphone camera. The images are taken from a mobile smartphone camera, Mi Note 10 that is mounted on an octopus mini tripod stand mount for mobile smartphone. The camera has an overall resolution of 1920 x 1080 pixels with a focal length of 44 mm, and a resolution of 30 fps. The resolution and rate of fps are selected to provide sufficient detail in the image to identify individual vehicles and to capture sequential images rapidly enough so that individual vehicles can be tracked between images without shifting more than one vehicle length between images.

The camera is then wrapped around a bridge pole with a viewing angle facing straight into the road. The smartphone is fixed in a horizontal position for the CVF (Camera's View Field) angle. The standard height of a pedestrian bridge is 5.5 m and the height of the camera from the road approximately 7.0 m. Figure 3.1 shows how the camera setup at the bridge.

Figure 3.1 Octopus Mini Tripod Stand Mount for mobile smartphone

3.2.3 Datasets for Traffic Analysis

Over the last decade, there has been an increase in interest in traffic monitoring at pedestrian bridges, with an emphasis on environment modelling and vehicle behaviour analysis. The co-creative effort benefits in the exchange of data and the advancement of research. This has resulted in the creation of difficult datasets for evaluation and benchmarking (Kamble and Kounte, 2020). The recording was taken from around 1.00 pm to 4.00 pm on the same day. All videos were recorded for 2 minutes at 30 fps.

NTebal1 (Figure 3.1): the traffic dataset from front view of pedestrian bridge in front of SK SAUJANA INDAH, is for research on activity analysis of low traffic scenes.

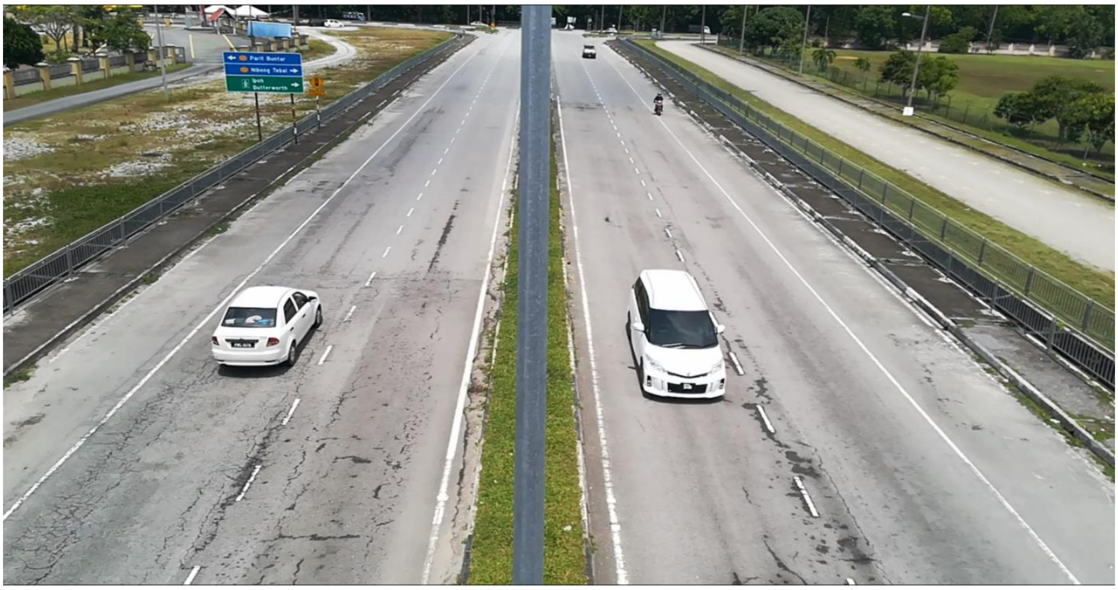

Figure 3.2 Frame taken from a video (NTebal1)

NTebal2 (Figure 3.2): the traffic dataset from back view of pedestrian bridge in front of SK SAUJANA INDAH, is for research on activity analysis of low-medium traffic scenes.

Figure 3.3 Frame taken from a video (NTebal2)

Jawi1 (Figure 3.3): the traffic dataset from front view of pedestrian bridge in front of SK JAWI, is specifically intended for activity analysis and behaviour understanding. It contains a big shadow come from the tree in the left frame of the video which gives a noise for detection and tracking activities.

Figure 3.4 Frame taken from a video (Jawi1)

Jawi2 (Figure 3.4): the traffic dataset from front back view of pedestrian bridge in front of SK JAWI, provides a single view of a heavy traffic situations.

Figure 3.5 Frame taken from a video (Jawi2)

SgBakap1 (Figure 3.5): the traffic dataset from front view of pedestrian bridge in front of SK SUNGAI BAKAP, is meant to generate a comprehensive dynamic model of the ongoing traffic.

Figure 3.6 Frame taken from a video (SgBakap1)

SgBakap2 (Figure 3.6): the traffic dataset from back view of pedestrian bridge in front of SK SUNGAI BAKAP, provides data for bench-marking unsupervised trajectory-based activity analysis algorithms.

Figure 3.7 Frame taken from a video (SgBakap2)

3.2.4 Software Setup

The first developed of the program are with the latest OS image – Ubuntu 20.04 LTS. Linux-Ubuntu 20.04 LTS was installed using Oracle VM VirtualBox. Oracle VM VirtualBox is cross-platform virtualization software that enables users to expand their existing machine to run several operating systems concurrently. To jumpstart the development process, the useful computer vision from all other libraries such as OpenCV and Python were enabled. Python 3.8 was used with the latest version for program the coding.

3.3 Flowchart

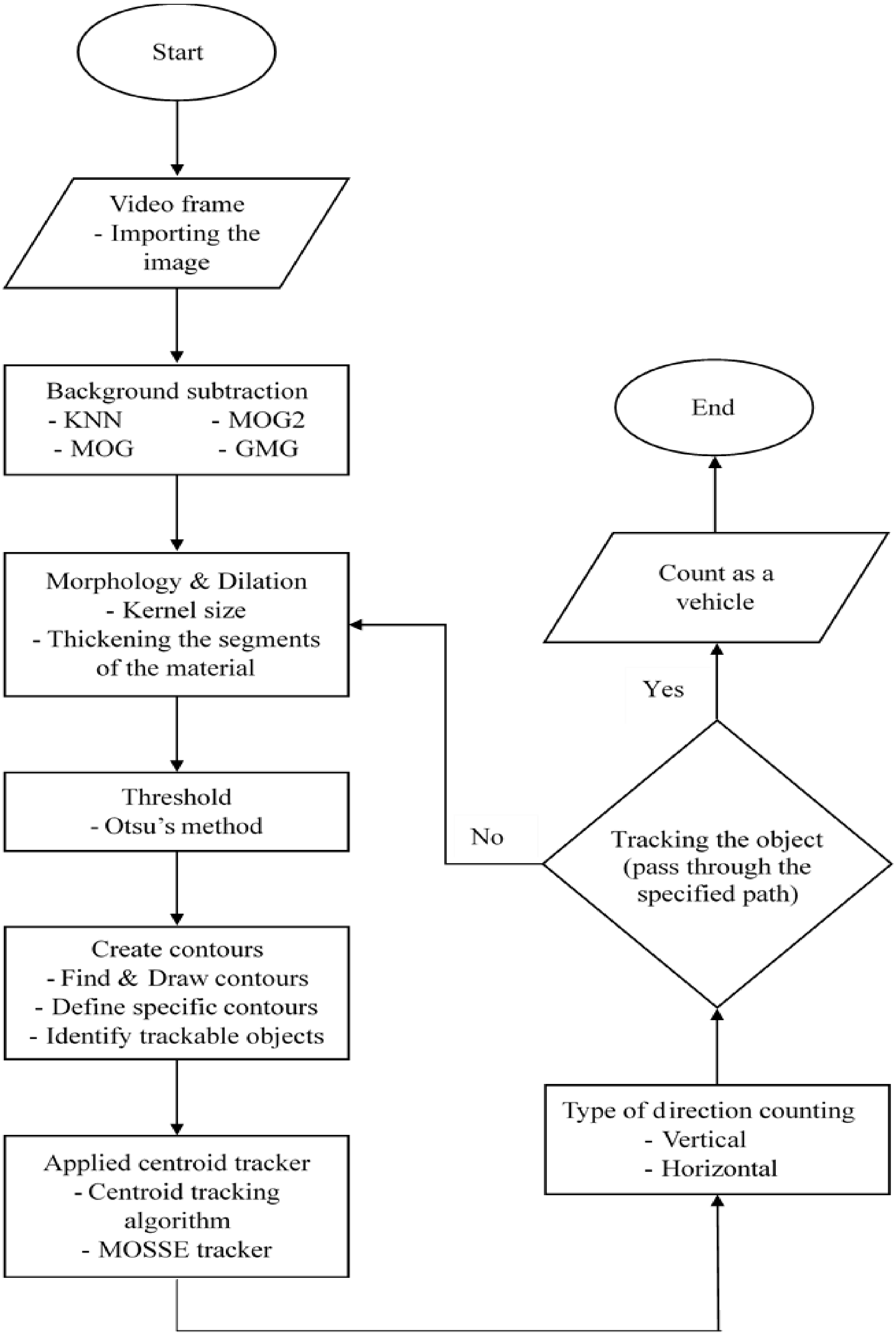

Figure 3.8 Flowchart of the system

Figure 3.8 shows the main process for this system. This algorithm would grab a reference to the video file, if a video path was not supplied, the algorithm would grab a reference to the webcam. Next, the frame will be converted to the grayscale image and then blur it. The selected background subtraction model applied which returns a mask. When the frame in grayscale image, the dilation kernel will initialize and define. The next important step to counting vehicles is to apply thresholds to the image to allow better isolation. Otsu's method will be used because the value of threshold would be automatically defined. After that, the binary image will be used to create contours around the vehicles. The centroid tracker will be instantiated and initialize a dictionary to map each unique object ID to a trackable object. Then initialize the direction info variable used to store information such as up/down vehicles count. A horizontal line drew in the frame, once an object crosses this line will be determine whether the vehicles were moving up or down and the object will be count as a vehicle after cross the line.

3.4 Importing the Image

OpenCV (Open-Source Computer Vision) is a computer vision library that contains various functions to perform operations on images or videos. OpenCV library can be used to perform multiple operations on videos. To capture a video, cv2.VideoCapture function was used. This function has the device index or the name of a video file. The device index file is just the number to specify which camera. If we pass 0 then it is for the first camera, 1 for the second camera so on. The video was capture frame by frame. Note: Video file should in the same directory where the program is executed.

cv2.VideoCapture(0) (3.1)

cv2.VideoCapture(1) (3.2)

cv2.VideoCapture('file name.mp4') (3.3)

The function shows that number '0' means first camera or webcam, number '1' means second camera or webcam and number 'file name.mp4' means video file.

3.4.1 Background Subtraction using OpenCV

Background subtraction has several use cases in everyday life, it is being used for object segmentation, security enhancement, pedestrian tracking, counting the number of visitors, number of vehicles in traffic and the rest. It can learn and identify the foreground mask (Sobral and Vacavant, 2014).

As the name suggests, it can subtract or eliminate the background portion in an image. Its output is a binary segmented image that essentially gives information about the non-stationary objects in the image. There lies a problem in this concept of finding a non-stationary portion, as the shadow of the moving object can be moving and sometimes being classified in the foreground. The popular background subtraction algorithms.

cv2.createBackgroundSubtractorKNN() (3.4)

cv2.bgsegm.createBackgroundSubtractorMOG() (3.5)

cv2.createBackgroundSubtractorMOG2() (3.6)

cv2.bgsegm.createBackgroundSubtractorGMG() (3.7)

3.4.2 Morphology and Dilation

Morphology is an image processing technique focused on an image fragment that attempts to enhance segmentation results. Morphological methods are typically used with binary images or in some cases, may also be extended to grey images or