NEAR-MISS TRAFFIC TRAJECTORY DETECTION BASED ON DEEP LEARNING

YANG LU

UNIVERSITI SAINS MALAYSIA

2025

NEAR-MISS TRAFFIC TRAJECTORY DETECTION BASED ON DEEP LEARNING

by

YANG LU

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

February 2025

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to **ASSOCIATE PROFESSOR DR. AHMAD SUFRIL AZLAN MOHAMED** and **DR. MAJID KHAN MAJAHAR ALI** for their unwavering guidance, invaluable insights, and continuous support throughout my Ph.D. journey. Their mentorship and encouragement have been instrumental in shaping my research and personal growth. I am deeply thankful to USM for providing me with the resources, facilities, and research opportunities that facilitated the completion of this thesis. I would like to extend my appreciation to my fellow Ph.D. candidates and colleagues who have been an incredible source of inspiration and encouragement. Their camaraderie and intellectual discussions have enriched my research experience. I am indebted to Fujian Province Education and Research Foundation for their financial support, which enabled me to pursue my research aspirations. I am profoundly grateful to my family for their unwavering love, encouragement, and understanding throughout this challenging journey. Lastly, I wish to acknowledge the support and assistance provided by Jimei University Chengyi College, whose efforts have been indispensable to the success of my research. Thank you to everyone who has contributed to my Ph.D. experience and made this achievement possible.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

LIST OF SYMBOLS

$a_{sim,t}^{ij}$	The Similarity of the Two Vehicles i and j at the moment t
v_t^i, v_t^j	The Speed of the Vehicles i and j at the moment t
TP	The Number of Detection Boxes with IoU > 0.5
FP	The Number of Detection Boxes with IoU <= 0.5
FN	The Number of Ground-Truth Boxes without Detected
$F1$	A Weighted Average of the Accuracy and Recall of the Classification Model
IoU	The Overlapping Area Between the Ground-Truth Box and the Prediction Box
AP_i	An Indicator of the Effectiveness of Object Detection
mAP	The Mean AP Value for Each Real Target Category Separately
MOTA	An Indicator of the Effectiveness of a Multi-Target Tracker
IDSW	The ID Switch
$silu$	A Variant of the Swish Activation Function
$swish$	A Type of the Activation Function
p_0^{HW*C}	The Interaction Order
$\emptyset_{in}, \emptyset_{out}$	The Linear Projection Layers Performing Channel Mixing
$\mathbb{R}^{HW*2C}$	The Input Characteristic
$\odot$	The Hadamard Product (Component-Wise Multiplication for Matrices)
$\otimes$	The Multiplication by One Element
F	The Intermediate State
M_c, M_s	The Channel Attention Map and Spatial Attention Map
β_{ij}^t	A Scalar Indicating the Influence of Vehicle j on Vehicle i
s_{ij}^t	The Weight of Distance Between the Vehicle to the Target Vehicle
A	The Target State Transfer Matrix Representing the Transfer Relationship Between the States of the System From Moment t-1 to Moment t
H	The Target Observation Matrix

w_t, v_t	The Process Noise and Observation Noise Respectively Satisfying the Normal Distributions
θ	The Displacement Direction of the Target
x_i, x_j	The Displacement Coordinates of the Target in Two Frames
k	The Number of Road Users
S_x^{0j}, S_y^{0j}	The Longitudinal and Transverse Distance Between the Two Road Users
w_{ij}	A Weight for Near-Miss Process
X_j	The Relative Longitudinal and Transverse Distance Between the Subject Road User
L_j , L_0	The Length of the Target
W_j, W_0	The Width of the Target
$MADR_x^0$	The Braking Capacity of the Road User
α	The Frames Per Second
β	The Paranoia Constant Derived from the Inverse Perspective Mapping

LIST OF ABBREVIATIONS

CCTV	Closed-Circuit Television
AP	Average Precision
mAP	mean Average Precision
GPS	Global Positioning System
DREAM	Driving Reliability and Error Analysis Method
MTO	Man, Technology, Organization
NIDB	Near-Miss Incident Database
TDD	Trajectory-pooled Deep-convolutional Descriptors
CNN	Convolutional Neural Network
R-CNN	Region-based Convolutional Neural Network
Fast R-CNN	Fast Region-based Convolutional Neural Network
Faster R-CNN	Faster Region-based Convolutional Neural Network
Mask R-CNN	Region-based Convolutional Neural Network with Mask
FPN	Feature Pyramid Network
YOLO	You Only Look Once
FCOS	Fully Convolutional One-stage Object Detection
SSD	Single Shot Detector
COCO	Common Objects in Context dataset
DBL	Darknet Bottleneck Layer
GTSDB	German Traffic Sign Detection Benchmark
STS	Swedish Traffic Sign
DET2	Facebook Detectron2
ACmix	Self-Attention and Convolution Mixed Model
NWD	Normalized Wasserstein Distance
SIOU	Semantic Intersection over Union
CARAFE	Content-Aware Reassembly of Features
SPD-Conv	Space-to-Depth Convolution
NAM	Normalization-Based Attention Module
EMOD	Efficient Moving Object Detection Pipeline
HOG	Histogram of Oriented Gradients
SVM	Support Vector Machine

ADAS	Advanced Driver Assistance Systems
HOG-SVM	Histogram of Oriented Gradients with Support Vector Machine
LP	License Plate
ANPR	Automatic Number Plate Recognition
YOLACT	You Only Look At CoefficienTs
GPU	Graphics Processing Unit
ResNets	Residual Neural Network
SOT	Single Object Tracking
CF	Correlation Filter
MOOSE	Minimum Output Sum of Squared Error
CSK	Circulant Structure of Tracking-by-detection with Kernels
KCF	Kernelized Correlation Filters
CN	Color Names
HSV	Hue, Saturation, Value
MOT	Multiple Object Tracking
MTT	Multiple Target Tracking
ID	Identifier
SORT	Simple Online and Real-time Tracker
IOU	Intersection over Union
DeepSORT	Deep Simple Online and Real-time Tracker
DBT	Detection-Based Tracking
N-YOLO	Not Pipelining You Only Look Once
SFFSORT	Shallow Feature Fusion Algorithm Based on the Simple Online and Realtime Tracking
XGBoost	Extreme Gradient Boost
VGG	Visual Geometry Group
Social-STGCNN	Social Spatio-Temporal Graph Convolutional Neural Network
SAT	Self-Adversarial Training
IPM	Inverse Perspective Mapping
OGM	Occupancy Grid Map
MSE	Mean Square Error
BBox	Bounding Box
StrongSORT	Strong Simple Online and Real-time Tracker
FPS	Frames Per Second
MBPP	Majlis Bandaraya Pulau Pinang

mIOU	Mean Intersection over Union
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
MOTA	Multi-Object Tracking Accuracy
FDE	Final Displacement Error
ADE	Average Displacement Error
CBS	Convolutional layer, a Batch Normalization layer, and a SiLU activation layer
SiLU	Swish activation function
ELAN	Efficient Network Architecture in the Backbone Feature Extraction Network
MP	Multilayer Perceptron
BN	Batch Norm
REP	Repvgg Block
FFN	Feed-Forward Network
HorNet	High Order Spatial Attention
SA-NET	Shuffle Attention for Deep Convolutional Neural Networks
SimAM	A Simple, Parameter-Free Attention Module for Convolutional Neural Networks
CC-Attention	Cross-Correlated Attention
S2-MLPv2	Spatial-Shift MLP Architecture for Vision
SKAttention	Selective Kernel-based Attention
NAMAttention	Normalization-based Attention Module
BAM	Block Attention Module
CBAM	Convolutional Block Attention Module
SEAttention	Squeeze-and-Excitation-based Attention Module
ECA-Net	Efficient Channel Attention for Deep Convolutional Neural Networks
GAM	Global Attention Module
CACSA	Channel Attention and Convolutional Space Attention
MLP	Multilayer Perceptron
NAM	Normalization-Based Attention Module
ConvNeXt	convolutional neural networks and the Fully-Convolutional Nature
BoTNet	Bottleneck Transformers for Visual Recognition Network
CotNet	Contextual Transformer Network
DN	Distance Neighbors

PICUD	Index of Collision Urgent Deceleration
PSD	Proportional Stopping Distance
ITS	Intelligent Transportation Systems
KF	Kalman Filter
ViTs	Vision Transformers
G3HN	g^n Conv *3 with HorNet
SPPCSPC	Spatial Pyramid Pooling with Cross Stage Partial Concatenation

PENGESANAN TRAJEKTORI TRAFIK HAMPIR TERLANGGAR BERASASKAN PEMBELAJARAN MENDALAM

ABSTRAK

Kaedah berasaskan penglihatan komputer memang telah digunakan secara meluas untuk memantau keadaan lalu lintas jalan raya. Keselamatan jalan raya merupakan kebimbangan penting di persekitaran bandar, dengan peristiwa hampir terlanggar berfungsi sebagai penunjuk berharga kemalangan yang berpotensi. Dalam kajian ini, satu rangka kerja inovatif dicadangkan yang menggabungkan YOLOv7 dengan struktur berasaskan transformer dan teknik segmentasi untuk pengesanan objek yang kukuh, penjejakan, dan analisis peristiwa hampir terlanggar dalam senario lalu lintas. Dengan menggunakan keupayaan pengesanan objek masa nyata YOLOv7, ia dipertingkatkan melalui integrasi seni bina transformer. Peningkatan ini membolehkan penangkapan kebergantungan jarak jauh dan maklumat kontekstual, seterusnya meningkatkan ketepatan dalam pengecaman dan penyetempatan objek. Selain itu, kaedah segmentasi digunakan untuk menggariskan objek dalam tempat kejadian, memperhalusi kotak pengesanan untuk lebih sesuai dengan objek sasaran. Dengan mengintegrasikan algoritma penjejakan objek, trajektori dan interaksi entiti yang dikesan dipantau dari semasa ke semasa, memudahkan pengenalpastian dan analisis peristiwa hampir terlanggar dengan ketepatan yang lebih tinggi. Pendekatan komprehensif bukan sahaja mengesan dan menjejaki objek secara berkesan tetapi juga memberikan wawasan tentang dinamik ruang dan masa yang mendasari insiden hampir terlanggar. Eksperimen adalah berdasarkan dataset (CCTV) POL37 dari Pulau Pinang, Malaysia. Keputusan eksperimen menunjukkan bahawa algoritma YOLOv7 yang dipertingkatkan yang dicadangkan berkesan mengenal pasti sasaran kecil semasa

fasa pengesanan objek, mencapai ketepatan pengesanan melebihi 94%. Ketepatan tinggi ini dicapai semasa menggunakan model yang lebih kecil dan mencapai kelajuan latihan yang lebih pantas. Selain itu, algoritma berjaya memenuhi tugas pengawasan untuk mengesan kejadian nyaris dalam adegan pengawasan CCTV. Penilaian percubaan menunjukkan keberkesanan dan kecekapan rangka kerja YOLOv7 yang dipertingkatkan ini dalam mengendalikan persekitaran trafik yang kompleks, memajukan pemahaman tentang kejadian hampir terlepas dan menyumbang kepada peningkatan langkah keselamatan jalan raya.

NEAR-MISS TRAFFIC TRAJECTORY DETECTION BASED ON DEEP LEARNING

ABSTRACT

Computer vision-based methods have indeed been widely employed for monitoring road traffic conditions. Traffic safety is a critical concern in urban environments, with near-miss events serving as valuable indicators of potential accidents. In this research, an innovative framework is proposed that combines YOLOv7 with transformer-based structures and segmentation techniques for robust object detection, tracking, and near-miss event analysis in traffic scenarios. Utilizing the real-time object detection capabilities of YOLOv7, it is augmented through the integration of transformer architectures. This enhancement enables the capture of long-range dependencies and contextual information, thereby improving accuracy in object recognition and localization. Additionally, segmentation methods are employed to delineate objects within the scene, further refining the detection box to better fit the target object. By integrating object tracking algorithms, the trajectories and interactions of detected entities are monitored over time, facilitating the identification and analysis of near-miss events with heightened precision. The comprehensive approach not only detects and tracks objects effectively but also provides insights into the spatial and temporal dynamics underlying near-miss incidents. The experiments are based on the POL37 Closed-Circuit Television (CCTV) dataset from Penang, Malaysia. The experimental results indicate that the proposed improved YOLOv7 algorithm effectively identifies small targets during the object detection phase, achieving a detection accuracy exceeding 94%. This high accuracy is attained while utilizing smaller models and achieving faster training speeds. Additionally, the

algorithm successfully fulfils the surveillance task of detecting near-miss events in CCTV surveillance scenes. The experimental evaluations demonstrate the efficacy and efficiency of this enhanced YOLOv7 framework in handling complex traffic environments, advancing the understanding of near-miss events and contributing to the enhancement of road safety measures.

CHAPTER 1

INTRODUCTION

1.1 Background

Rapid economic expansion and urban development in Asian cities, including major hubs like Beijing, Shanghai, Kuala Lumpur, and Bangkok, have led to a notable rise in traffic congestion, presenting a significant challenge (Loo & Huang, 2021). Effective monitoring of road traffic is essential for managing urban congestion and improving safety. However, as urbanization continues and population density increases, traffic-related challenges are becoming more pronounced. Furthermore, commuters are increasingly turning to personal vehicles for transportation due to long wait times, unpredictable schedules, and overcrowding on public transportation systems. This situation exacerbates traffic congestion, leading to commuters spending significant amounts of time stuck in traffic jams throughout the day. In such scenarios, motorcycles, bicycles, and electric motorcycles offer an alternative means of transportation that can help alleviate congestion and improve commute times.

Nevertheless, it is crucial to recognize that the growing prevalence of motorcycles and bicycles also correlates with a rise in traffic accidents involving these modes of transportation. Studies have shown that motorcyclist fatalities constitute a significant proportion of traffic accidents (Kitamura et al., 2018). Hence, there is an urgent necessity for comprehensive analysis and investigation of road incidents concerning vehicles, motorcycles and bicycles to minimize risks and bolster road safety measures.

On the other hand, it is important to note that research on traffic accidents extends beyond actual crashes to include near misses and unsafe situations. Over the years, various researchers have referred to safety-related incidents using different terms such as near misses, near-crashes, traffic conflicts, safety-critical incidents, traffic interactions, and traffic encounters (Polders, 2018). Unlike collisions, which entail evident indications of contact and quantifiable damages, traffic conflicts represent incidents where no physical interaction occurs between road users, despite the potential for significant harm (Zheng et al., 2021).

By leveraging advanced technologies such as AI-powered traffic monitoring systems, smart traffic lights, and predictive analytics, cities can enhance their ability to manage traffic and respond to evolving urban transportation needs. This research aims to employ deep learning-based image processing techniques to efficiently identify near-miss events and classify them according to risk levels (safe or near-miss risk) and occurrence frequencies. The research encompasses both four-wheeled vehicles and motorcycles. By including motorcycles in the analysis, the research seeks to offer a thorough grasp of safety-related incidents and near misses across different vehicle types, contributing to the evolution of effective strategies for improving road safety for all road users and the evolution of proactive measures for enhancing road safety.

1.1.1 Near-Miss

A near-miss accident refers to any event that may cause personal injury or property or environmental damage but has no actual consequences. Only a lucky break in a series of incidents can prevent injury, death, or damage. In other words, this was a very thrilling event. Heinrich's law (Stone, 1931) demonstrates that for each accident leading to a severe injury, there are roughly 29 incidents resulting in minor injuries and 300 incidents causing no injuries. This underscores a direct relationship between real casualties and near-miss occurrences. By mitigating the probability of danger, the likelihood of casualties can also be decreased. Therefore, it is essential to heed the lessons from both accidents and near misses (Terum & Svartdal, 2019).

Cho et al., (2023) has established the relevance of Heinrich's law in studying traffic accident within traffic contexts and leveraged media descriptions of road accidents and conducted analyses of accident causes as the foundation for developing a safety framework for accidents involving motor vehicles. In the context of traffic incidents, Heinrich's Law implies that studying minor traffic incidents and near-miss events can provide valuable insights into identifying potential accident hotspots and improving overall road safety. By leveraging digital tachograph data, which records information about vehicle movements and driver behaviour, researchers can analyse patterns and trends in near-misses to pinpoint areas where accidents are more likely to occur. The occurrence of near-miss events contains the fundamental elements of accidents, underscoring their significance in the realms of traffic and industrial safety. Continuous investigation and analysis of potential hazards and vulnerabilities are essential to implementing effective preventive measures and ensuring overall traffic safety.

1.1.2 Surveillance Tracking

As modern management practices continue to evolve, the role of closed-circuit television (CCTV) systems is also transforming. Kurniawan et al., (2018) has emerged as a widely utilized tool for monitoring and managing various environments. This technology enables real-time monitoring and communication of on-site situations through images and text, facilitating swift and effective responses from management, maintenance, and safety personnel. By capturing and recording critical data, CCTV systems provide valuable insights and evidence for incident analysis and accident response. Leveraging the capabilities of CCTV technology, urban traffic management has also adopted CCTV monitoring networks extensively. These systems are deployed across major cities worldwide, serving as vital sources of information for tracking vehicles and pedestrians within the traffic network and identifying potential incidents or anomalies.

1.1.3 Advancements in Computer Vision for Urban Traffic Management and Near-Miss Detection

As deep learning technology advances and computing power increases, coupled with the abundance of visual data, computer vision has made significant strides in the realm of urban intelligent traffic management. Among them, object detection (Zhao et al., 2019) and object tracking (Porikli & Yilmaz, 2012) are the main research directions in this field. Object detection identifies various entities such as vehicles, pedestrians and cyclists within a given scene, enabling traffic management systems to analyse and respond effectively to dynamic situations. Meanwhile, object tracking enables the continuous monitoring of these entities as they move through the environment, providing valuable insights into traffic flow patterns and potential congestion points. Moving on from these fundamental components, the detection of near misses is a critical extension.

Near-miss incidents are typically recognized based on spatial proximity or maneuvers to avoid collisions (Tang et al., 2024). Evasive actions such as braking and swerving are commonly employed to prevent accidents. Analysing near-miss events entails considering various factors, including the types of vehicles or road users involved and the characteristics of the road infrastructure. Both spatial distance and the behaviour of road users, including their speeds, are critical parameters in traffic monitoring and management. These factors contribute significantly to overall road safety and traffic flow.

First of all, it is necessary to determine which target is in the surveillance video (this study refers to all kinds of vehicles and pedestrians) to solve the problem existing in the target. Secondly, determining the specific location of a target object in a surveillance video involves object localization. It is necessary to lock the same object in different frames and continuously acquire the position and trajectory of the target. Ultimately, the position of each detected target is analysed to assess the possibility of a near-miss event, and predictions are made regarding the trajectory of these targets in the immediate future to anticipate potential near-miss situations.

1.1.4 Application of Near-Miss

Near-miss events analysis in traffic road scenarios finds diverse applications, primarily categorized into three types: analysis of driving behaviours and patterns (Jomnonkwao et al., 2023), safety analysis using offline data, and real-time monitoring systems.

Driving behaviours and habits, encompassing factors like speed, turning angles, and trip duration, are primary contributors to near-miss events, with analysing these behaviours crucial for identifying accident patterns and enhancing driver awareness of road conditions. While some studies rely on textual reports (Ahadh et al., 2021) for analysis, offering valuable insights despite lacking real-time data, others emphasize the significance of real-time monitoring systems in enhancing traffic safety. These systems utilize video image data to train models based on historical data, subsequently deployed to detect and alert for potential near-miss events in real-time using CCTV camera footage. Therefore, deploying such real-time monitoring systems on roads is imperative, as they leverage video data to construct models trained on historical datasets, enabling the detection and identification of near-miss events as they occur.

1.2 Problem Statement

With the current road traffic situation becoming more and more complex, the number of motorcycle drivers and car drivers accounts for a high proportion of traffic deaths (Chowdhury & Chakraborty, 2017). The purpose of this research is to reduce motorcycle accidents and solve traffic-related problems. Unfortunately, near misses cannot be collected solely by police reports, hospital records, or other official statistics. On the other hand, most existing work on near misses is written work based on a small number of investigations. Therefore, the main purpose of this research is to develop a video surveillance-based system through image and video processing technology combined with deep learning to improve road safety.

As people increasingly realize that traffic safety is a very important problem that needs to be solved, the following are three main challenges.

- Image-based object detection algorithms are currently less accurate in detecting small targets to judge near-miss events, especially the motorcycles in the CCTV data.

Most methods are context-based (Girotto et al., 2016) and (Kwayu et al., 2021), such as police reports and hospital records. On the other hand, many studies only pay attention to the accidents/incidents reported in official sources in European countries.

Context detection methods based on reports mainly focus on text data mining and rule judgment. Therefore, it is difficult to realize real-time detection, and it is difficult to get intuitive detection results. In addition, although many image datasets on traffic are available to researchers today, there is no systematic annotation for traffic near-miss incidents. Based on the traffic situation in Asia, the research and dataset of near-miss events about motorcycles, bicycles, or electric motorcycles are still missing (Kataoka et al., 2018). Research in this area has a poor effect on the real-time detection of motorcycles and bicycles. Key challenges include real-time/small objects detection and the absence of systematic annotation for near-miss incidents, posing significant hurdles to comprehensive analysis and effective mitigation strategies.

Inaccurate detection frames remain a common issue in object detection algorithms. For example, detection frames may fail to align precisely with the target object, resulting in excessive empty space surrounding the object or parts of the object extending beyond the frame boundaries.

- A significant challenge arises from the decreased accuracy when tracking small targets, exacerbated by hurdles like missing and overlapping instances. Furthermore, less study has explored using trajectory prediction for near-miss events and challenges with target missing and overlapping in tracking.

At firstly, small objects have limited visual features, making it challenging for detection algorithms to distinguish them from the background or other objects. Secondly, small targets are more prone to being completely or partially occluded by other objects in the scene, leading to difficulties in continuous tracking. Thirdly, in crowded scenes, small objects can easily be overlooked or confused with surrounding elements, leading to tracking errors and potential target losses (ID-Switch Problem). Fourthly, small and fast-moving targets can experience motion blur in video frames, making it challenging to accurately localize and track them over time. Finally, imbalanced datasets where small objects are underrepresented, can result in models that are biased towards larger objects, leading to poorer performance on small targets (Yu Wang et al., 2023).

- Most existing research primarily addresses accidents, lacking near-miss events, which are crucial for preventing future incidents, under-explored in the visual data analysis.

Limited scope of some studies which is designed and trained primarily for accident detection, which limits its ability to effectively identify and analyse narrowly escaping or near-miss events. Near-miss events are likely to be higher frequent than accidents, therefore, it is not possible to annotate directly in the dataset. This can affect the model's ability to learn and generalize well for the near-miss scenarios. Models trained specifically for accident detection may lack the flexibility to generalize well to diverse near-miss scenarios. Near-miss events can involve different dynamics, distances, and interactions compared to accidents. Relying on a single technology or type of model may limit the system's overall effectiveness. Defining what constitutes a near-miss event can be complex and subjective (Yamamoto et al., 2020). Different situations and cultures may have varied interpretations, making it challenging to establish a universal criterion for near-miss detection. Finally, the evaluation metrics used for accident detection may not be suitable for assessing the performance of near-miss detection.

Therefore, the main problem of this research is to investigate near-miss events without official reports and context descriptions. This is an accident that has not caused material or personal damage. However, it is likely that damage has already been caused.

1.3 Research Questions

Research questions related to traffic near-miss events can encompass various aspects of road safety, transportation engineering, human factors, and data analysis. The potential research questions in this area are shown as below:

- Frequency and Distribution

What is the frequency and distribution of near-miss events in different types of road environments (e.g., different times, different streets)? How do near-miss events vary across different times of day, days of the week, and seasons?

- Predictive Modelling and Forecasting

Can predictive models be developed to forecast the likelihood of near-miss events based on real-time traffic data, historical incident data, and environmental factors? How can deep learning algorithms be leveraged to identify patterns and trends in near-miss event data and improve predictive accuracy?

- Data Collection and Analysis

What are the best practices for collecting, processing, and analysing near-miss event data from various sources, including traffic surveillance cameras, vehicle telematics, and crowd-sourced incident reports? How can advanced data analytics techniques, such as spatial analysis, temporal analysis, and network analysis, be applied to gain insights into near-miss event patterns and trends?

- Evaluation of Near-Miss Reporting Systems

How effective are existing near-miss reporting systems in capturing and documenting near-miss events?

1.4 Objectives

This study primarily aims to automate the process of detecting of motorcycle near-miss incidents and precisely categorize precisely categorize that it is near-miss event or not. Literature reviews and surveys indicate a scarcity of reported near-misses in the single-vehicle category, while a significant number are observed in the Intersection and Overtaking/Lane Change categories. The main scenes are intersections, encounters, and lane changes. The focus of this research is to find these near-miss situations and scenes from the video data, and analyse them accurately and rapidly. The three key objectives of this study are outlined, each corresponding to one of the problem statements in Section 1.2. The first objective addresses the first problem statement, the second objective addresses the second, and the third objective resolves the final problem statement.

i. To design a vision-based method for accurate, real-time traffic object detection with minimal false positives in complex traffic environments.

Compare the differences between the context-based and vision-based methods in tracking other four-wheeled vehicles. Using data expansion method and geometric transformation to optimize the vision-based method to track 2-4-wheeled vehicles. Designing a object detection algorithm for small targets (motorcycles).

ii. To develop a tracking method for real-time, accurate trajectory prediction of traffic objects in dynamic environments.

The study data is CCTV video data, so a tracking algorithm needed to be designed to determine the continuity of the target. The target angle is adjusted using the IPM algorithm to ensure the precision of the tracking and detection data. Incorporating the acceleration of the object as a parameter can further improve the accuracy of the outcomes.

iii. To create a data-driven method for analysing traffic near-miss events, improving detection accuracy and risk assessment.

Utilizing both the proposed vehicles and motorcycles dataset, the vision-based method and parameter-net structure proposed in this study are employed to estimate near-miss incidents. This approach involves leveraging the rich data available in the vehicle's dataset, which comprises diverse scenarios involving various types of vehicles, to train the model. Additionally, curated to focus on road users-related incidents, provides targeted and specialized information crucial for accurate near-miss estimation in road users-centric contexts. By integrating these datasets and utilizing the proposed method and parameter-net structure, the aim is to enhance the model's ability to identify and predict near-miss incidents effectively, particularly those involving motorcycles, vehicles thus contributing to improved safety measures on the roadways.

1.5 Motivation

The motivation behind near-miss traffic analysis is based on several key factors, in particular that this research addresses a critical gap in current traffic management and safety efforts. The other motivations are outlined below.

- Near-miss incidents provide valuable insights into potential hazards and risks on the road. By analysing near-miss events, road safety measures can be improved to prevent future accidents.
- Understanding near-miss events makes it possible to identify patterns and behaviours that can lead to accidents. This insight enables the implementation of preventative measures to reduce risk and improve road safety.
- While there is a wealth of data on road accidents, near-miss events are often under-reported or overlooked. Analysing near-miss data fills this gap and provides a more comprehensive understanding of road safety issues.
- By identifying areas of high near-miss frequency, resources can be allocated more effectively to address potential safety concerns and prioritise interventions where they are most needed.
- By focusing on near-miss incidents and specifically addressing the challenges related to motorcycle safety.

Overall, by focusing on near-miss incidents and directly addressing the challenges associated with car/motorcycle safety, the research aims to make a significant contribution to wider traffic management and safety efforts. It not only fills critical gaps in data collection, but also uses emerging technologies to improve emergency response and overall road safety.

1.6 Contributions of Thesis

This research presents an approach to identify and forecast near-miss incidents in urban intelligent traffic management using object detection and tracking techniques. The primary objective is to enhance the accuracy of detection while minimizing false positives. The achievements of this study can be outlined as follows.

- The enhanced YOLOv7 algorithm by introducing tailored image segmentation and transformer-based adjustments designed to handle the unique data characteristics of urban traffic systems is proposed. This enhanced approach incorporates image segmentation and transformer structure tailored to the specific data characteristics of the research context. By optimizing weight values and bounding box size, the algorithm aims to achieve superior classification performance compared to traditional methods.

- Refining the object tracking method is achieved by integrating angle and acceleration data of the target object. This approach employs perspective and geometric transformations to enhance accuracy. By leveraging geometric properties and perspective, the method effectively refines object tracking, especially in scenarios where traditional methods may be less accurate.

- To ascertain the potential for a near collision between target objects, an exploration and investigation of braking force and distance metrics is conducted using various evaluation methods. Leveraging the actual characteristics observed, a predictive algorithm is employed to assess the likelihood of a near-miss event occurring in the upcoming frames.

In summary, the first novelty of the proposed methods in this study lies in the enhancement and integration of existing algorithms. Specifically, this work builds upon the YOLOv7 algorithm designed to handle the unique data characteristics of urban traffic systems. These modifications include optimizing weight values and adjusting bounding box sizes, which lead to more accurate classification performance compared to standard implementations. Secondly, the tracking method is enhanced by incorporating contextual data, such as angle and acceleration of the target object, alongside geometric and perspective transformations. This approach refines the tracking process by leveraging spatial and directional cues, particularly in complex scenarios where traditional tracking might be less precise. Finally, a novel predictive model is developed to assess near-miss events by evaluating braking force and distance metrics, allowing for improved near-collision predictions in the context of urban traffic safety.

1.7 Scope of Research

The motorcycle near-miss event dataset can be sourced from vehicle driving recorders, CCTV footage with hazard level predictions, and datasets provided by other researchers. These datasets typically include a significant number of scenes depicting near-miss events involving private vehicles. Each video is labelled with a risk rate, while each frame is annotated with motorcycle bounding boxes in different colours corresponding to the near-miss severity. However, this study specifically utilized

CCTV surveillance data from Majlis Bandaraya Pulau Pinang (MBPP). The dataset used in this research comprises videos captured from various vehicles, locations, times, and weather conditions, providing diverse and comprehensive traffic scenarios for analysis.

The scope of this research includes the development and evaluation of a comprehensive framework for the detection of traffic manoeuvres. This includes the implementation of object detection and tracking algorithms to accurately identify and monitor vehicles and other road users. The research also focuses on the definition and evaluation of near-miss events to improve road safety analysis. In addition, the research involves the meticulous cleaning and labelling of the dataset to ensure the accuracy and reliability of the detection and tracking processes.

1.8 Challenge and Limitation

The study also has certain limitations. First, the probability of using motorcycles to travel in the Asian region is much greater than that of bicycles, so that the number of bicycles in the data set is small. When two models are similar but differ significantly in size, the detection accuracy of the smaller model tends to decrease. Second, the study focuses on the use of the CCTV dataset. In the road section and time with the small pedestrian flow, the number of pedestrians photographed is small, and the accident rate is extremely low, so the behaviour of predicting the probability of their near-miss is trivial.

A clear research gap in traffic near-miss judgment lies in the lack of standardized, real-time detection and evaluation methods that can consistently identify and classify near-miss events across diverse traffic environments. While existing systems can detect near-miss events, many struggle with variability in traffic conditions, environmental factors (e.g., lighting, weather), and the precise distinction between actual collisions and near-misses. Furthermore, most methods rely heavily on post-event analysis, limiting the potential for real-time intervention. Addressing these gaps by developing a robust, adaptable, and real-time near-miss detection framework is crucial for enhancing traffic safety.

1.9 Summary

The thesis consists of 7 chapters presented in sequential order: introduction, literature review, research methodology, results utilizing object detection technology, results utilizing trajectory prediction technology, results utilizing near-miss judgment technology and discussion. The structure of the thesis is delineated as follows:

Chapter 1 sets the stage by providing an overview of Asian traffic conditions and the context of traffic accidents. It outlines the research objectives, problem statement, and underscores the significance of the study. Additionally, it defines the scope of the research, establishes the theoretical framework, and enumerates the research contributions.

Chapter 2 reviews the existing literature and relevant works pertinent to the study are critically reviewed and analysed.

Chapter 3 delineates the research methods employed, offering comprehensive insights into the procedural steps followed throughout the study.

Chapter 4 demonstrates the experimental process and results of the improved object detection method. Chapter 5 focuses on the experimental process and results of object tracking and object trajectory prediction. In Chapter 6, the determination process and result analysis of near-miss events are illustrated. The evaluation of each stage includes methods and results, along with comparisons with popular benchmark algorithms.

Chapter 7 would discuss the contribution, limitation, gap, and future work of the proposed model.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In recent years, considerable research effort has been directed towards vision-based approaches for detecting traffic accidents. This chapter provides a comprehensive review of prior studies relevant to the research conducted in this thesis. It delves into existing algorithms, methodologies, and applications aimed at mitigating traffic near-miss incidents. The overview diagram of the literature review is shown as Figure 2.1. It encompasses various components related to near-miss events, traffic surveillance algorithms, deep learning algorithms for vehicle detection, calculation of near-miss events, applications of traffic accident detection models, and traffic datasets and metrics.

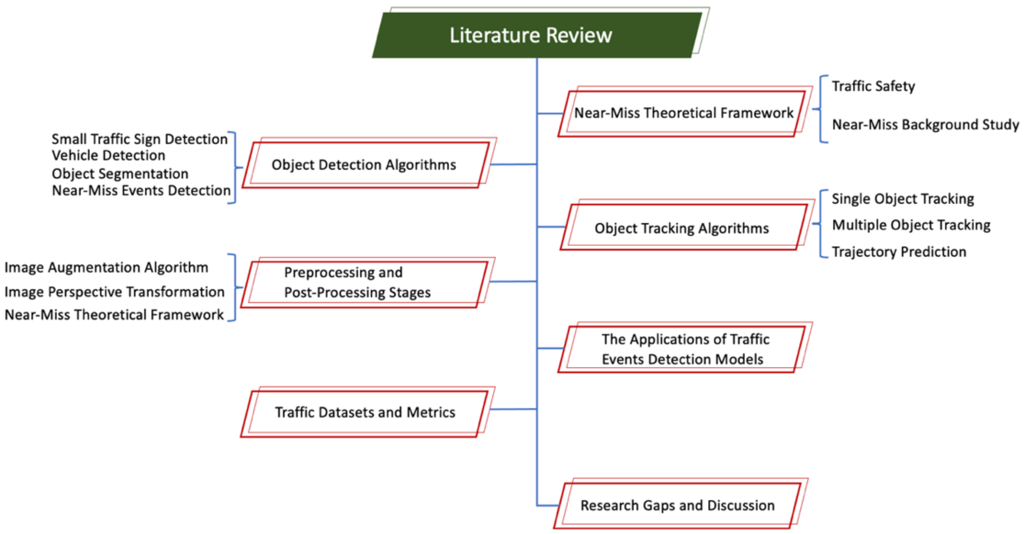

Figure 2.1 Overview Diagram of the Literature Review

As Figure 2.1 shown, the chapter is divided into seven main sections. Section 2.2 provides an overview of traffic safety and near-miss events in traffic scenarios, emphasizing their importance as precursors to potential accidents. It discusses the definition, classification, and significance of near-miss events in road safety research. In Section 2.3, the focus is on providing an overview of the current popular object detection algorithms. In Section 2.4, the focus shifts to providing a brief overview of current popular object tracking algorithms. The Section 2.5 aims to provide a comprehensive overview of the pre-processing and post-processing methods commonly employed in object detection and tracking pipelines. Post-processing techniques include the analysis and interpretation of the results obtained from object detection and tracking algorithms to determine and analyse traffic near-miss events. Section 2.6 examines the practical applications of traffic accident detection models derived from surveillance data. It highlights the use of machine learning and deep learning techniques for instantaneous detection accident detection, emergency response optimization, and traffic management. Section 2.7 discusses the availability and characteristics of traffic datasets used for training and evaluating surveillance and detection algorithms. It encompasses commonly utilized metrics for evaluating the effectiveness of detection models, such as accuracy, mAP, and F1-score.

2.2 Theoretical Framework for Near-Miss Analysis in Traffic Safety

This section aims to analyze the relationship between traffic safety and near-miss events and provides a systematic understanding of the algorithms utilized for defining and analyzing near-miss events in recent years. At first, clarify the concept of near-miss events in the context of traffic safety research, emphasizing their significance as indicators of potential safety hazards and vulnerabilities in traffic systems. Secondly, explore the connection between near-miss events and traffic safety, highlighting the importance of near-miss analysis in identifying pre-collision scenarios, understanding contributing factors, and preventing accidents. Finally, review the algorithms and methodologies employed for defining and analyzing near-miss events, encompassing both traditional statistical techniques and advanced machine learning methods, including the application of deep learning techniques, spatiotemporal modeling, and data fusion approaches.

2.2.1 Traffic Safety and Risk Assessment

Traffic safety refers to the measures and practices put in place to prevent accidents, injuries, and fatalities on roadways. It encompasses a wide range of strategies, regulations, and technologies aimed at ensuring the safe movement of vehicles, pedestrians, and cyclists (Chu & Sun, 2021). In order to ensure traffic safety, people are now beginning to use a variety of technological tools, including smart traffic management systems, traffic monitoring cameras, and vehicle-to-vehicle communication, can enhance safety. Furthermore, regular analysis of accident data helps identify trends, high-risk areas, and common causes of accidents, which in turn informs targeted safety interventions. By addressing these various technologies, communities can work towards reducing accidents and ensuring safer roadways for everyone.

With Asia's large population and the growth of courier and takeaway industries in many countries, there has been a significant increase in the number of road users, consequently raising the risk of traffic accidents, particularly involving motorcycles (Kitamura et al., 2018). The courier and food delivery population often work in fast-paced, high-demand environments. They face unique challenges related to traffic safety, including tight delivery schedules and exposure to various weather conditions. In their quest for speed, they usually use motorbikes as a means of transport. With the increasing use of motorbikes for deliveries, there is a growing need to address the associated risks. Motorbikes can reach high speeds and are often used in dense urban environments, which can lead to a higher likelihood of accidents. Addressing the safety of this occupational group also has broader implications for their long-term health and well-being. Preventing accidents and injuries can lead to a higher quality of life for these workers. (Z. Wang et al., 2021) underscores the importance of recognizing and addressing the specific challenges faced by courier and food delivery workers, particularly those using e-bikes. It calls for a multi-faceted approach involving policy, education, technology, and stakeholder collaboration to improve safety in this occupational group.

An innovative smartphone application (Kontaxi et al., 2021) is used to monitor and address the speeding behaviour of motorcyclists is a forward-thinking and technology-driven strategy to improve road safety. The application can provide real-time data on the speed of motorcyclists. This allows for immediate feedback and intervention in case of excessive speeding. The application can utilize GPS technology to track the location and speed of motorcyclists. This information can be used to identify high-risk areas or routes associated with speeding.

The application can send automated alerts or notifications to motorcyclists when they exceed a predefined speed limit. These alerts can serve as immediate reminders to adhere to speed limits. The application can allow users to set their own speed limits based on road conditions, personal preferences, or legal requirements. This customization encourages a sense of ownership over safe driving habits. The application can provide feedback to motorcyclists about their driving behaviour, including statistics on speeding incidents, trends over time, and comparisons to safe driving norms. By harnessing the capabilities of smartphone technology, this innovative application has the potential to significantly impact road safety, particularly among motorcyclists. It combines real-time monitoring, personalized feedback, educational resources, and community engagement to promote safer driving habits.

Through thorough investigation and exploration of the road infrastructure itself, (Shahi et al., 2023) aim to uncover the root causes of potential dangers and identify areas prone to hazards. By pinpointing these hazardous zones to determine which road policies require revision to enhance overall road safety. A comprehensive approach for systematically evaluating different sections of the road network have been outlined to identify potential hazards. This involves analysing risk factors, identifying and documenting situations where accidents were narrowly avoided. The factors mentioned, including sharp curves, blind spots, poor lighting, and high traffic density, are critical elements that contribute to the overall safety and risk profile of a road network. Based on the analysis and on-site assessments, provide recommendations for improvements in identified high-risk areas. Document the findings from the data analysis and recommendations in a comprehensive report. In addition, government departments can implement technology solutions like traffic monitoring cameras, sensors, or GPS data to gather real-time information about traffic flow, speed patterns, and accident occurrences. Finally, by combining thorough investigations, expert input, stakeholder collaboration, and data-driven analysis, targeted interventions can be developed to address specific dangers and improve overall road safety. This is a good approach which would help create a safer environment for all road users.

Merely detecting an accident is insufficient. Predicting and preventing dangerous situations on apparently safe roads, even when road users are compliant with traffic rules, is an important and challenging research direction. This proactive approach aims to identify potential risks before they escalate into accidents or near-miss incidents.

2.2.2 Background and Evolution of Near-Miss Analysis

A "near-miss" refers to a situation in which an accident almost occurs, but is narrowly avoided (Cho et al., 2023). The study of near-miss incidents is crucial for understanding potential hazards and improving safety measures.

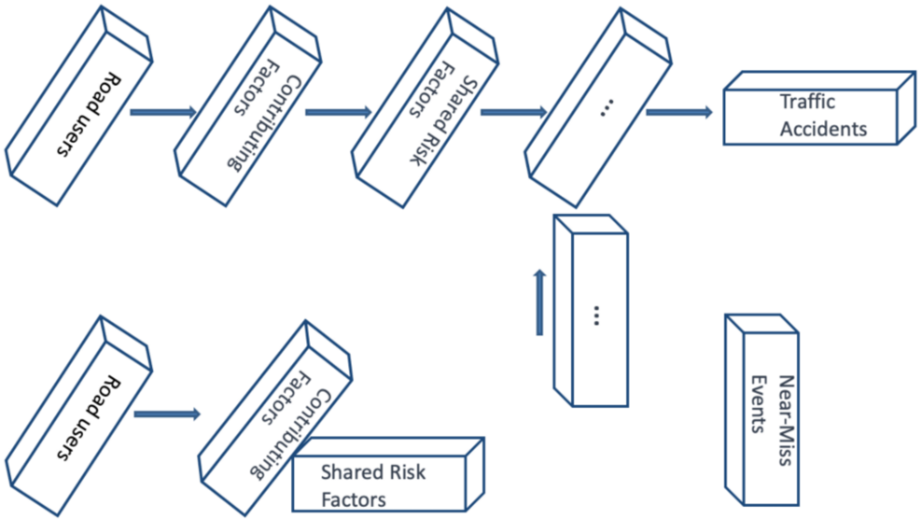

Figure 2.2 The Connection Between Accidents and Near-Misses

Near-misses and traffic accidents exhibit a significant positive correlation, wherein near-misses represent instances where the sequence of events leading to a collision is interrupted, thereby preventing the accident from materializing. Accidents, on the other hand, emerge as the culmination of various contributing factors, including but not limited to reckless driving, adverse weather conditions, and inadequate road infrastructure. When these contributing factors converge, they act as catalysts, triggering a chain reaction that culminates in an accident, akin to falling dominoes. However, accidents can be prevented if any of these contributing factors are eliminated or disrupted. This action breaks the chain, resulting in a near-miss event. The conceptual framework elucidates the intricate relationship between near-misses and accidents. It is visually illustrated in Figure 2.2, providing a clear depiction of how accidents transpire when all contributing factors align. Conversely, near-misses occur when the sequence of events is interrupted.

The study of near-miss incidents has evolved significantly over time (Al Shaaili et al., 2023). The concept of near-misses began to gain attention in the early to mid-20th century as safety became a focal point in various industries, particularly aviation and engineering. Researchers and practitioners recognized that incidents narrowly averted could provide critical insights into safety vulnerabilities (Beukenkamp, 2016). The aviation industry was one of the pioneers in near-miss studies (Rains, 2023). The introduction of flight data recorders (black boxes) in the 1940s allowed for the systematic collection of data on incidents, including those that did not result in accidents. This marked a significant leap in understanding the importance of near-miss analysis. In the late 20th century, the healthcare industry began to focus on near-miss incidents in patient safety. Healthcare professionals recognized that incidents that narrowly avoided harm to patients could offer crucial insights into system weaknesses and opportunities for improvement (Sheikhtaheri, 2014). The nuclear industry also became a key area for near-miss studies. The inherently high stakes and potential catastrophic consequences of nuclear incidents emphasized the importance of understanding incidents that stopped short of a major disaster. The study of near-miss incidents became increasingly integrated with human factors engineering and risk management disciplines (M G Gnoni et al., 2022). Researchers and practitioners sought to not only identify near-misses but also to analyse the underlying causes and develop strategies for prevention. The advent of advanced data collection technologies, including digital recording systems, sensors, and advanced analytics, greatly facilitated the study of near-miss incidents. These technologies allowed for more comprehensive data collection and analysis.

Today, the concept of near-miss is applied across a wide range of industries including transportation, manufacturing, energy, healthcare, and more (M G Gnoni et