E-Proceedings
of
The 5th International Conference on Highway Engineering 2024

"Future-Proofing Roads for Asia and Beyond"

4 - 6 September 2024
Bangkok International Trade & Exhibition Centre (BITEC)
Bangkok, Thailand

Organised by
Department of Highways
and
Roads Association of Thailand

Edited by
Ponlathep Lertworawanich, Ph.D.
Auckpath Sawangsuriya, Ph.D.
Harutus Phoban, Ph.D.

Published by
Department of Highways
Ministry of Transport

E-Proceedings
of
The 5th International Conference on Highway Engineering 2024

“Future-Proofing Roads for Asia and Beyond“

4 - 6 September 2024
Bangkok International Trade & Exhibition Centre (BITEC)
Bangkok, Thailand

Organised by
Department of Highways
and
Roads Association of Thailand

Edited by
Ponlathep Lertworawanich, Ph.D.
Auckpath Sawangsuriya, Ph.D.
Harutus Phoban, Ph.D.

Published by
Department of Highways
Ministry of Transport

Published by the Department of Highways, Ministry of Transport
2/486 Si Ayutthaya Road,
Thung Phaya Thai, Ratchathewi,
Bangkok, 10400, Thailand.

ISBN (e-book) 978-616-7754-87-1

Preface

This E-proceedings contains papers from the **International Conference on Highway Engineering 2024 (iCHE2024)** with the theme of **"Future-Proofing Roads for Asia and Beyond"**, held in Bangkok, Thailand, on September 4-6, 2024. This conference is the 5th International Conference on Highway Engineering in this conference series which serves as the platform for national and global leaders, policy-makers, industry professionals, specialists, research scholars, and academia to attend and exchange knowledge, experience, and insights on roads and road transport in these following topics: (a) efficient, inclusive, and safe road management, (b) smart mobility, digital technology, and innovation for roads, and (c) sustainability and resilience of road networks.

This E-proceedings includes a collection of papers from three main tracks of the iCHE2024. **Track A** involves efficient, inclusive, and safe road management. 37 papers were collected in this track. These papers described a wide range of road administrator tasks from planning for effective road networks to managing road infrastructure and transportation system for promoting economic growth and social development as well as improving the lives of citizens across the country. In addition, they also covered road safety management, which is essential towards achieving social equity in accessibility in rural areas.

Track B is associated with smart mobility, digital technology, and innovation for roads. 12 papers were collected in this track. These papers summarized mobility on roads, which was a crucial aspect of modern society, enabling people to access essential services, to engage in economic activities, and to connect with others. The papers addressed the application of new technologies and digital transformation concept in road transport, which optimized the road network operations and enhance the intelligent transportation system. In addition, they also discussed the implementation of emerging technologies and innovations that could help supporting road agencies during the design, construction, operation, and maintenance of roads.

Track C involves sustainability and resilience of road networks. 16 papers were collected in this track. These papers described sustainable and resilient road infrastructure and networks to ensure the serviceability and reliability of road transportation systems. The papers also demonstrated that an increasingly significant global impact of climate change and natural hazards such as floods, landslides, and earthquakes had a devastating impact on road networks, disrupted transportation, and caused significant economic and social harm. The papers presented the resilience of road networks, the development of robust disaster management plans, the investment in resilient infrastructure to mitigate the impact of these events as well as the ongoing monitoring and adaptation to changing conditions. In addition, they also discussed about the decarbonization efforts e.g. the use of low-carbon materials and the adoption of sustainable construction practices, making transportation systems more resilient to the impacts of climate change, and reducing the carbon footprint of road construction and maintenance.

The Editors,
Ponlathep Lertworawanich, Ph.D., Department of Highways
Auckpath Sawangsuriya, Ph.D., Department of Highways
Harutus Phoban, Ph.D., Department of Highways

Acknowledgments

Thanks are due to the authors, reviewers, technical committees, scientific committees, and chairman of iCHE2024 without whom this publication would not be possible.

iCHE2024 Chairman

Dr. Piyapong Jiwattanakulpaisarn
Deputy Director-General for Engineering, Department of Highways

iCHE2024 Scientific Committees

Prof. Dr. Agachai Sumalee, Chulalongkorn School of Integrated Innovation, Thailand
Prof. Dr. Kasem Choocharukul, Chulalongkorn University, Thailand
Prof. Dr. Kunnawee Kanitpong, A sian Institute of Technology, Thailand
Assoc. Prof. Dr. Sorawit Narupiti, Chulalongkorn University, Thailand
Prof. Dr. Saksith Chalermpong, Chulalongkorn University, Thailand
Prof. Dr. Vilas Nitivattananon, Asian Institute of Technology, Thailand
Prof. Dr. Somnuk Tangtermsirikul, Sirindhorn International Institute of Technology, Thailand
Dr. Ponlathep Lertworawanich, Department of Highways, Thailand
Dr. Auckpath Sawangsuriya, Department of Highways, Thailand
Dr. Sumet Ongkittikul, Thailand Development Research Institute (TDRI), Thailand
Prof. Dr. Hiroaki Nishiuchi, Kochi University of Technology, Japan
Dr. John Calvin Milton, Washington State Department of Transportation, United States
Mr. Christos S. Xenophontos, Rhode Island Department of Transportation, United States
Ms. Caroline Evans, PIARC Strategic Theme 1 Coordinator, Australia

iCHE2024 Technical Committees

Attasit Sawatparnich	Advisor
Ponlathep Lertworawanich	Chairman
Jakraphob Watcharamonthian	Member
Auckpath Sawangsuriya	Member
Songrit Chayanan	Member
Tunwin Svasdisant	Member
Sukit Yindeesuk	Member
Gridsana Pensomboon	Member
Patthanapong Tongsuk	Member
Poranic Jitareekul	Member
Chairat Supachawarote	Member
Jutha Sunitsakul	Member
Bhanitiz Aursudkij	Member
Pattharin Sarutipand	Member
Apivat Jotisankasa	Member
Wasin Rujikietgumjorn	Member
Puttipan Seraneeprakarn	Member
Pichanan Aursudkij	Member
Kavalin Wangsiripaisal	Member
Rajwanlop Kumpoopong	Member
Harit Prasertsin	Member
Waritz Rattanasiriphan	Member
Harutus Phoban	Member and Secretary
Warit Chookate	Member and Assistant Secretary
Siwakorn Dumduang	Member and Assistant Secretary
Patra Sukmoung	Member and Assistant Secretary
Nattapong Nusrikaew	Member and Assistant Secretary
Wuttichai Amornprasitpol	Member and Assistant Secretary
Punnawat Siripatthiti	Member and Assistant Secretary

Acknowledgments

Local Reviewers

Ackchai Sirikijpanichkul
Agachai Sumalee
Ampol Karoonsoontawong
Apichai Issariyanukula
Apivat Jotisankasa
Arkhom Termpittayapaisith
Auckpath Sawangsuriya
Boonchai Sangpetngam
Chaiyut Charoenphon
Chakrapan Tuakta
Chutipong Paraphantakul
Dussadee Satirasetthavee
Harit Prasertsin
Jetthawee Mark
Jittichai Rudjanakanoknad
Jumrus Pitaksringkarn
Kasem Choocharukul
Kazuya Tomiyama
Kerkritt Sriroongvikrai
Kriangkrai Pattanapakdee
Kunnawee Kanitpong
Ladda Tanwanichkul
Manop Kaewmoracharoen
Monchai Chumintarajak
Muanmas Wichiensin
Nongnuch Poolsawad
Nopadon Kronprasert
Noppakun Boongrapue
Nuksit Noomwongs
Pakpoom Limtong
Paramet Luathep
Parinya Sanguansat
Pasan Kulvanit
Pattara Kiatisevi
Peeranut Jeammaneeporn
Peerapol Sittivijan
Piya Chotickai
Pongsun Bunditsakulchai
Ponlathep Lertworawanich
Pornpote Nusen
Puttipan Seraneeprakarn
Rapee Tangsongsuwan
Rattaphol Pueboobpaphan
Rosalind Amornpitakpun
Sarawut Jansuwan
Saroch Boonsiripant
Siradech Surit
Sittinun Tongbai
Sivach Kaewcharoen
Somnuk Tangtermsirikul
Songpol Charuvisit
Songrit Chayanan
Sruangsaeng Chaikasetsin
Sukit Yindeesuk
Suladda Subsin
Tanakorn Chaiyarungyos
Teerapong Panboonyuen
Teerapun Phumrattanaprapin
Thaned Satiennam
Thapat Silalertruksa
Thavatchai Tayjasanant
Treerapot Siripirote
Tunwin Svasdisant
Usanee Raha
Warit Wipulanusat
Waritz Rattanasiriphan
Wasin Rujikietgumjorn
Watchara Sattayaprasert
Wichuda Satiennam
Winai Raksuntorn
Wirat Chinnan
Withit Pansuk

Table of Contents

Page

Track B: Smart Mobility, Digital Technology and Innovation for Roads **464**

ID 570 Zero-Shot Learning Based on Non-Motorized Vehicles Perception Based on a Surveillance System 465
Sruangsaeng Chaikasetsin, Mehrdad Nasri, Chenxi Liu and Yinhai Wang

ID 648 The study of the reflectance ratio on highway pavement surfaces by using High-Pressure Sodium (HPS) lamp type 486
Ratanachai Meeampol, Chainarong Benjapornthavee, Pattanapong Ngosorn and Pattarapon Sedokbuab

ID 756 Assessing BIM Implementation Success through Defined KPIs in Indonesia's Road and Bridge Infrastructure Projects 499
Heru Tri Saksena

ID 1136 Development of Machine Vision-Based Road Surface Maintenance Techniques for Climate Change 515
In Bae Kim

ID 1145 Simulation-Based Analysis of Variable Speed Limit for Traffic Performance and Safety During Incidents 527
S. Hemapan, A. Karoonsoontawong and K. Kanitpong

ID 1175 Optimizing Vehicle-to-Building Integration: A Sustainable, Adaptable Energy Management Framework 539
Yu-Shin Hu and I-Yun Lisa Hsieh

ID 1210 Multi-Objective Optimization in Energy Management using Vehicle-to-Building (V2B) Strategy and Energy Storage System 552
Jian Hern Yeoh, Kai-Yun Lo and I-Yun Lisa Hsieh

ID 1256 Exploring a Possible Mobility-as-a-Service Platform from the Perspective of Transport Policy Makers and Regulators in Thailand 564
Sirawit Chinvanichai and Sorawit Narupiti

ID 1260 A Automatic Recognition of Pavement Deterioration using Deep Learning: Using an AI-driven system for automatic pavement damage recognition, validating the identified damages and calculating the Pavement Condition Index (PCI) scores 577
SZU-HAN LU, Jyh-Dong Lin and Yi-Shian Chiou

ID 1276 Prediction and Variable Explanation of International Friction Index (IFI) using Deep learning: A Case Study in Thailand 588
Punnawat Siripatthiti and Ponlathep Lertworawanich

ID 1304 Real-time Traffic Data Collection at Mid-block and Intersection by UAVs and Computer Vision 601
Ponlathep Lertworawanich, Peerapol Sittivijan and Theerada Rungruangjarensuk

ID 1390 Evaluating Driver's Characteristics During Multiple Rear-end Collisions on Freeway Corridors 620
Ryuya Seki and Hironori Suzuki

Track B

Smart Mobility, Digital Technology and Innovation for Roads

ZERO-SHOT LEARNING BASED ON NON-MOTORIZED VEHICLES PERCEPTION BASED ON A SURVEILLANCE SYSTEM

Sruangsaeng Chaikasetsin[1*], Mehrdad Nasri[2], Chenxi Liu[3], Yinhai Wang[4]

[1]Department of Civil and Environmental Engineering/University of Washington, Seattle, USA
sruanc@uw.edu
[2]Department of Civil and Environmental Engineering/University of Washington, Seattle, USA
mehnasri@uw.edu
[3] Department of Civil and Environmental Engineering/University of Washington, Seattle, USA
lcx2017@uw.edu
[4] Department of Civil and Environmental Engineering/University of Washington, Seattle, USA
yinhai@uw.edu

ABSTRACT

Ensuring the safety of pedestrians and cyclists in active transportation is of paramount importance. However, current sensing technologies often struggle to detect and classify these vulnerable road users accurately. While existing computer vision models can identify people and bicycles, they encounter difficulties recognizing cyclists in transportation contexts. This study introduces an innovative solution using zero-shot learning techniques to enhance the detection of cyclists in images and videos. The approach builds upon the COCO dataset, which already identifies people and bicycles. It imposes five constraints on the computer vision model to train it to recognize a person riding a bicycle. The method is expanded to video detection and tracking using YOLO models and the ByteTrack algorithm. The research evaluates the efficacy of these zero-shot constraints compared to traditional machine learning methods, focusing on metrics such as Multiple Object Tracking Accuracy (MOTA), Identity F1 score (IDF1), and Higher Order Tracking Accuracy (HOTA). The results demonstrate significant improvements, with MOTA scores reaching 0.941, IDF1 up to 0.538, and HOTA as high as 0.924. This study's findings underscore the robustness of the proposed approach, contributing to improved cyclist detection and tracking, thereby enhancing traffic safety for active transportation users.

Keywords: zero-shot learning, transportation equity, traffic safety, traffic sensing

1. INTRODUCTION

The safety of non-motorized vehicles (NMVs), such as pedestrians and cyclists, has emerged as a critical concern for transportation systems, underscored by the alarming increase in fatalities among these vulnerable road users. Recent data from the Fatality Analysis Reporting System reveals a troubling trend: pedestrian fatalities have surged to their highest level in a decade, comprising 17% of all fatalities, while cyclist fatalities have risen to 2% (National Highway Traffic Safety Administration [NHTSA], 2023). Despite advancements in roadway infrastructure and technology, effectively safeguarding NMVs remains a significant challenge, exacerbated by the ongoing proliferation of these modes of transport.

In dense urban environments, monitoring vulnerable road users (VRUs) such as pedestrians and cyclists presents challenges due to their unpredictable movements and complex behaviors. The Federal Highway Administration's Traffic Monitoring Guide (2022) acknowledges the difficulty in identifying and tracking VRUs, particularly in dense scenarios. Thus, there is a clear need for advanced data collection and analysis methods. One promising solution is computer vision (CV) technology, which uses deep learning and sophisticated sensor systems (Elhenawy et al., 2020; Goldhammer et al., 2020). With CV, VRU movements can be accurately observed and analyzed in real-time, providing detailed insights into their behavior at urban intersections (Maurya & Choudhary, 2018). The latest algorithms and sensors used in CV technology can detect and predict pedestrian and cyclist movements on the road, thereby improving overall safety (Dimitrievski et al., 2020). Thus, the potential of CV applications to better understand and protect VRUs highlights the importance of further research and refinement in this area.

Sophisticated surveillance systems have the capability to supply up-to-the-minute information on non-motorized vehicles (NMVs) to improve road safety. Nevertheless, present systems encounter difficulty in precisely detecting and categorizing vulnerable road users (VRUs) like elderly persons, children, cyclists, and those with disabilities due to the inadequacy of diversity in existing datasets. The identification of VRUs presents a challenge for computer vision models, and even models trained on datasets like MS-COCO (Lin et al., 2014) struggle to detect cyclists due to the viewpoint problem—the appearance of cyclists changes based on their orientation (García-Venegas et al., 2021). This difficulty is mainly due to the limited variety of datasets available, such as MS COCO and Pascal VOC (Everingham et al., 2010), and the complexities involved in detecting cyclists (Zhang & Ling, 2017). To address this gap, our research proposes an innovative approach to optimize surveillance systems' capabilities in detecting and tracking VRUs, particularly for pedestrians and cyclists, by integrating Zero-Shot Learning (ZSL) techniques with computer vision (CV) detection and tracking algorithms. ZSL is a paradigm that allows recognition tasks to be tackled without examples of all classes during training (Xian et al., 2018). By leveraging high-level descriptions of categories and semantic similarities between classes, ZSL can bridge the gap between observed and unobserved classes, enhancing the detection and tracking of NMVs. Recent studies have shown that models like CLIP can achieve competitive performance in zero-shot tasks, outperforming traditional supervised models in terms of robustness and computational efficiency (Radford et al., 2021). Therefore, our objective is to enhance the detection and tracking of NMVs through the integration of ZSL techniques, particularly pedestrians and cyclists, who constitute the majority of vulnerable road users (VRUs). By leveraging the large amount of data collected by surveillance systems, coupled with advancements in CV, we aim to enhance safety measures for VRUs on roads, even in situations where there is no dedicated information available for all vulnerable groups. This approach has the potential to significantly improve road safety for NMVs without requiring additional investments in roadside infrastructure.

Our innovative ZSL approach focuses on identifying cyclists in transportation scenarios without requiring additional annotations. Five constraints are central to this approach, designed to enable ZSL to understand the concept of "a person riding a bicycle." These constraints include overlap, pedaling, height, width, and tracking, which are applied during the post-processing stage to match the features of detected persons and bicycles. This ensures accurate cyclist identification. Our method comprises three key modules: pre-processing, object detection, and post-processing. The pre-processing module prepares raw image or video data for detection, while the object detection module utilizes YOLOv5 (Jocher, 2020), YOLOv8 (Jocher et al., 2023), and YOLOv9 (Wang et al., 2024) detection models, each integrated with the ByteTrack (Zhang et al., 2022) tracking algorithm. These detection models are trained on the COCO (Lin et al., 2014) dataset, which includes both bicycle and pedestrian classes but notably excludes cyclists. Finally, the post-processing module introduces the ZSL method to match bicycle and pedestrian features detected by the models, facilitating precise cyclist identification. Moreover, our research evaluates the effectiveness of the zero-shot constraints compared to traditional machine learning methods. We also assess the performance of zero-shot learning models for object detection and tracking in traffic applications, using evaluation metrics such as Multiple Object Tracking Accuracy (MOTA) (Bernardin & Stiefelhagen, 2008), Identity F1 score (IDF1) (Ristani et al., 2016), and Higher Order Tracking Accuracy (HOTA) (Luiten et al., 2021).

The contributions of the research can be summarized into the following three points:

- Firstly, we propose a new approach to detect NMVs using ZSL techniques based on the surveillance systems while refining our detection models with human-derived parameters.
- Secondly, we utilize ZSL methods to label cyclists in pedestrian environments with minimal effort, extending the application of ZSL to more challenging semantic definitions. This includes combining two seen classes and defining actions.
- Lastly, we evaluate the effectiveness of ZSL in comparison to traditional methods in NMVs perception tasks, testing the feasibility of humans devising constraints in a ZSL framework.

Our efforts aim to advance the field of NMVs perception based on the surveillance system, offering innovative solutions for improved detection and tracking in transportation scenarios. This pursuit holds promise for enhancing road safety and mitigating risks for VRUs. Ultimately, our study aims to contribute to the development of more effective and reliable systems for safeguarding vulnerable individuals on our roads.

2. LITERATURE REVIEW

In this section, we review the current methods and advancements in cyclist detection and tracking within traffic systems. We first examine traditional detection-based methods, identifying the research gaps that our work aims to address. We then discuss the advantages of using computer vision techniques to solve these problems. Next, we explore the development of Zero-Shot Learning (ZSL), its applications, and its pros and cons. Following this, we present a review of the state-of-the-art object detection and tracking models, specifically focusing on why YOLO and ByteTrack were chosen for our approach. This comprehensive review sets the foundation for understanding the current landscape and highlights the need for innovative solutions in cyclist detection and tracking.

2.1 Relevant detection-based methods

Detection-based methods for identifying cyclists have traditionally relied on attributes such as helmet and pose detection. Past research by Hirota et al. (2017) and Mistry et al. (2017) extensively used helmet detection for identifying motorcyclists. However, these techniques primarily aim at enforcing helmet laws rather than distinguishing cyclists from pedestrians. Considering the more relaxed helmet regulations for cyclists and our goal to reduce manual data annotation, helmet detection is not suitable for our Zero-Shot Learning (ZSL) model.

Li et al. (2016) proposed the Tsinghua-Daimler Cyclist Benchmark (TDCB), classifying cyclists based on their orientation. While TDCB's approach of using height and width ratios is insightful, its direct application is limited in our project due to different focus areas.

Pose detection has also been explored. Fang et al. (2019) aimed to predict the intentions of vulnerable road users (VRUs) for autonomous vehicles. However, detecting cyclists from a distance using traffic cameras differs from predicting intentions, making pose detection less applicable for our needs. Abadi et al. (2022) proposed a method for detecting cyclists' crossing intentions based on body and head orientations, using segmentation and pose estimation neural networks. This approach highlighted the importance of head movement as a crossing signal, providing valuable insights.

Knura et al. (2021) differentiated moving from stationary bicycles in social media images using a three-step model, considering a bicycle moving if a person's and bicycle's Intersection Over Union (IOU) exceeded 0.10. While our model incorporates similar techniques, the IOU threshold will be adjusted to accurately capture cyclists.

Recent advancements have further enhanced detection capabilities and provided additional context for our approach. Abdel-Gawad et al. (2021) utilized deep learning for the detection and tracking of VRUs, emphasizing the need for high precision in various traffic scenarios. Lu et al. (2023) improved 3D VRU detection with point clouds, addressing the challenges of detecting small or distant objects with LiDAR sensors.

The research gap identified is the need for a method that accurately detects cyclists without extensive manual data annotation and can be applied effectively in real-time traffic scenarios.

2.2 Computer Vision Methods for Cyclist Detection

We use computer vision methods to solve our problem because they provide robust and scalable solutions for real-time detection and tracking. Computer vision techniques, leveraging advancements in deep learning, offer high accuracy in object recognition and classification, making them ideal for dynamic and complex environments such as traffic systems. Moreover, these methods can be integrated with existing surveillance infrastructure, enhancing their practicality and deployment feasibility.

2.3 Zero-Shot Learning for Cyclist Detection

Zero-Shot Learning (ZSL) is a technique that enables models to recognize objects without having seen explicit training examples of those objects. This approach leverages semantic embeddings, such as attributes or textual descriptions, to transfer knowledge from known classes to unseen classes. ZSL has gained traction due to its ability to address the challenge of limited annotated data, making it particularly useful for applications requiring scalability and adaptability.

Xian et al. (2019) provided a comprehensive evaluation of ZSL methods, identifying the strengths and weaknesses of various approaches. Antol et al. (2014) demonstrated the potential of ZSL in visual abstraction, where models could identify objects based on high-level descriptions. Romera-Paredes and Torr (2015) proposed an embarrassingly simple approach to ZSL, highlighting its efficiency and simplicity in implementation.

Applications of ZSL span various domains, including image and video recognition, natural language processing, and robotics. The primary advantage of ZSL is its ability to generalize knowledge, reducing the dependency on large, annotated datasets. However, a significant drawback is the potential for reduced accuracy in recognizing unseen classes compared to traditional supervised learning methods.

We chose ZSL for our problem to eliminate the need for additional annotated datasets specific to cyclists, thereby significantly reducing manual data annotation efforts. By leveraging pre-trained models and the COCO dataset, we apply constraints specific to cyclists to achieve accurate identification. This approach not only streamlines the training process but also enhances the model's adaptability to new classes without requiring extensive retraining.

2.4 State of The Art (SOTA) in Object Detection and Tracking Models

Recent advancements in object detection and tracking have produced highly efficient models. Abdel-Gawad et al. (2021) utilized deep learning for the detection and tracking of vulnerable road users (VRUs), emphasizing the need for high precision in various traffic scenarios. Lu et al. (2023) improved 3D VRU detection with point clouds, addressing the challenges of detecting small or distant objects with LiDAR sensors.

Among the state-of-the-art models, YOLO (You Only Look Once) and ByteTrack stand out for their performance. Jiang et al. (2022) provided a comprehensive review of YOLO developments, showcasing its advancements and adaptability for real-time applications. YOLO is renowned for its speed and accuracy in object detection tasks, making it suitable for real-time applications. It processes images in a single pass, significantly reducing computational time while maintaining high detection accuracy.

ByteTrack, as detailed by Zhang et al. (2022), is employed for tracking detected objects across frames, ensuring robust performance in dynamic environments. ByteTrack improves the reliability of tracking by using a simple yet effective method to associate objects across frames. It excels in maintaining high tracking performance, even under challenging conditions such as occlusions and rapid object movements.

3. METHODOLOGY

The structure of the proposed method is shown in Figure 1, with the detection phase shown in blue, and the post-processing phase is shown in yellow. In the beginning of the process, a detection model will detect the bounding boxes of person p and bicycle b, with annotations in the [x, y, width, height] format. From there, we determine a pool of potential cyclist candidates by using the intersection over union (IOU) method, which measures the overlapping area of the bounding boxes and divides it by their unified area (see Eq 1). Once the IOU is calculated for every possible person/bicycle combination in an image, each person is assigned to a bicycle using the Hungarian method (Kuhn, 1955). Using this method assigns each person to a bicycle, while also making sure that each person is assigned to only 1 unique bicycle, and vice versa. Any person/bicycle combination with an IOU value greater than zero is thrown into our pool of potential cyclist candidates.

$$IOU = \frac{Area\ of\ Overlap}{Area\ od\ Union} \tag{1}$$

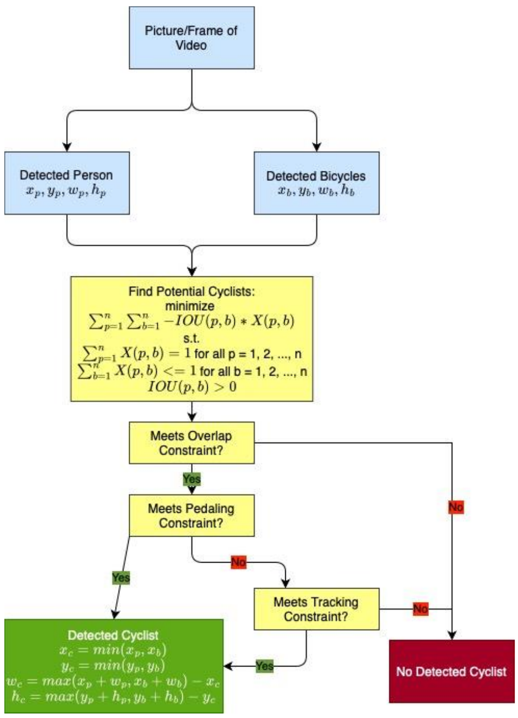
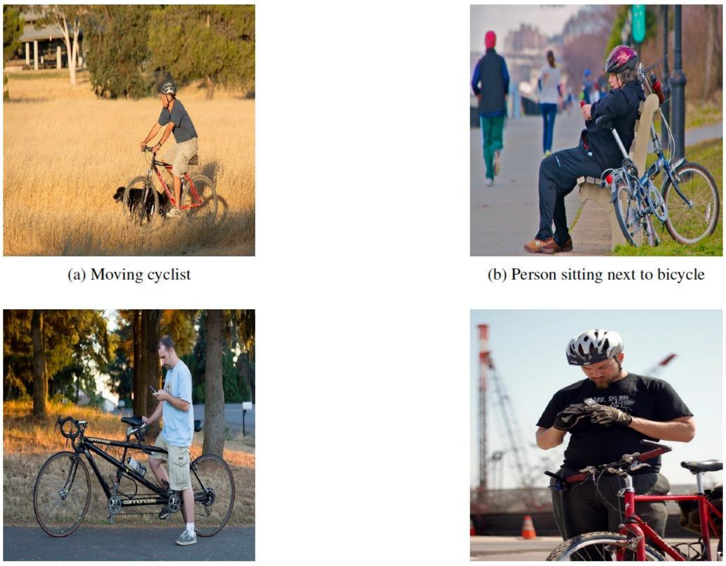


Figure 1 Architecture of our proposed ZSL model.

With our pool of potential cyclists created, a series of processes are devised to separate actual cyclists in the pool from overlapping person/bicycle combinations in an image that are not actually cyclists. To do this, a series of five constraints were tested to help us define what the act of "riding" is to a computer. The following sections detail how we define the act of "riding" on a ground truth level and explore how each of the metrics has been developed to detect that ground truth state.

3.1 Defining Riding on a Ground Truth Level

Before developing a method to define what "riding" is in our zero-shot learning model, we must first define what "riding" is at a ground truth level. From the perspective of VRU safety, the reason why we want to differentiate cyclists from pedestrians in the first place is their ability to reach higher top speeds. Therefore, for the purposes of this paper, the act of "riding" a bicycle is defined in our ground truth dataset as having a bicycle in between the legs of a person who is sitting, standing, and/or pedaling. Figure 2 (Lin et al., 2014) demonstrates how our definition works across 4 images in the COCO training dataset. Here, only 2a and 2c are considered cyclists. In the case of 2a, the person is pedaling the bike, so we expect that the person is capable of travelling faster than a pedestrian, and thus we choose to label to person and bicycle as a combined cyclist. We also expect 2c to be capable of travelling faster than a pedestrian. Even though the cyclist in 2c is not visibly travelling, they still have the potential to start riding at any moment and reach high speeds quickly. While one may assume that 2b and 2d would be considered cyclists since they are wearing helmets, their ability to travel faster than pedestrians depend on them getting on their bikes first. Since neither is pictured on their bikes, neither is considered a cyclist in the ground truth dataset. The goal of our zero-shot learning methodology is to find ways to detect instances like 2a and 3c as cyclists, while excluding images like 2b and 2d.



(a) Moving cyclist

(b) Person sitting next to bicycle

(c) Stationary cyclist

(d) Person standing behind bicycle

Figure 2 Examples of cyclists.
(Lin et al., 2014)

3.2 Defining Riding on a ZSL model

The following sections will discuss the 5 metrics that were devised in order to define what constitutes the act of "riding" is for our ZSL model to recognize. Of these 5 metrics, 4 were developed during the training phase, and 1 was developed for the testing phase. The training metrics were determined through trial and error by a human supervisor examining a subset of roughly 2500 images from the 2017 COCO training dataset that contained images of both people and bicycles. Each training metric developed is assigned a value of 1 if it was felt that this metric would be a defining characteristic of a cyclist and a value of 0 if not. It is worth noting that the metrics devised in the following sections were developed before ground truth data was created for the training dataset. This approach was taken in order to make sure that adjustments could not be made to the model with respect to the ground truth data before results were measured, which would've violated the ZSL assumptions as outlines by (Xian et al., 2018). Furthermore, annotating ground truth data of cyclists by hand before our ZSL model is implemented would've defeated the purpose of using a ZSL model in this project, so ground truth annotation will be avoided until our 5 metrics are set in stone.

Metric 1: Overlap Constraint

The first metric that was investigated as a potential constraint for our ZSL model was an overlap metric. This metric is designed to investigate the degree of overlap between the detected persons and bicycles in our pool of potential cyclists' image. The intersection over union (IOU) measurement is utilized to determine this degree of overlap. Here, the IOU will be utilized to measure the degree of overlap between a pedestrian and a bicycle in the model and determine the practical IOU threshold at which a cyclist could be detected. As noted earlier, Knura et al. (Knura et al., 2021) uses the IOU threshold of 0.1 to separate parked bicycles from bicycles being ridden by people, which was a very low threshold that worked for the purposes of their project, since they only wanted to count the number of bicycles in an image at the end of the day. With this in mind, we knew that our IOU constraint should lie somewhere above 0.1 if we wanted to detect cyclists accurately. Thus, IOU constraints were analyzed at intervals of 0.05 between the thresholds of 0.1 and 1, with the images of potential cyclists captured in our training dataset analyzed at each interval. Through this process, it was decided that the IOU threshold IOU of 0.25 seemed to perform very well at capturing cyclists, as higher thresholds missed many images of cyclists, while lower thresholds captured to many non-cyclists. Thus, the IOU threshold of 0.25 was added as our overlap constraint, as shown in Eq 2.

$$Overlap\ Constraint = \begin{cases} 1, & if\ IOU(p,b) \geq 0.25 \\ 0, & otherwise \end{cases} \tag{2}$$

Metric 2: Pedaling Constraint

With the lack of a ground truth dataset to test the effectiveness of our constraints, it was deemed necessary to develop our metrics in a way that would hopefully minimize any false negative errors when detecting objects. As mentioned earlier, one of the key differences between cyclists and pedestrians is the cyclist's ability to reach higher top speeds than pedestrians. From the perspective of VRU safety, a false positive error, where a bicycle or pedestrian is incorrectly detected as a cyclist, would likely result in an unnecessary level of caution in our transportation systems, as the CV model would recognize the person or bicycle of being capable of travelling much faster than they are able to. Conversely, a false negative error, where the model fails to recognize a person riding a bicycle as a cyclist, would result in a model that would believe that a cyclist would be moving much slower than they might actually be, and thus would exercise less caution than necessary, which could be dangerous for the cyclist. Because of this, a pedaling constraint is developed as one of the metrics in the model. While on the ground truth level we define a cyclist as anyone sitting, standing, or pedaling a bicycle, it is with the act of pedaling where a false negative error is considered particularly dangerous, since pedaling cyclists are the most likely to be travelling at a higher top speed than a pedestrian. In Knura et al.'s study (Knura et al., 2021), a pedaling constraint was implemented by checking if the centroid of a person's bounding box was above the centroid of a bicycle's bounding box. Our proposed model takes a different approach that will compare the position of a person's feet to the position of their bicycle. In the absence of pose detection, we will assume that a person's feet are located at the bottom of a person's bounding box, and the pedals of a bike are located somewhere between the top and the bottom of a bicycle's bounding box. Therefore, if a person is pedaling with their feet on the pedals of the bike, the bottom of a person's bounding box should be located somewhere in between the top and the bottom of an overlapping bicycle. This relationship is expressed mathematically in Eq 3.

$$Pedaling\ Constraint = \begin{cases} 1, & if\ y_b \leq y_p + h_p \leq y_b + h_b \\ 0, & otherwise \end{cases} \tag{3}$$

Metric 3: Height Constraint

A metric for height was developed because, in real life, the height of a bicycle has a deterministic relationship with the height of the person riding it. In other words, when we think of people riding bicycles, we imagine that the handlebars, seats, and pedals are in the same positions relative to a cyclist's body, regardless of the cyclist's height.

While the exact extent of this deterministic relationship is difficult to determine in the training dataset, due to all the varying angles from which the pictures of cyclists are taken, there were a few assumptions as to how the relationship between a person's height and the height of an overlapping cyclist could help us further refine our ZSL method. For example, if a person were riding a bicycle, it could be reasonable to assume that the bicycle should not be taller than the person riding said bike. Likewise, it could be reasonable to assume that the height of a bike should not be less than 1/4th the height of a person riding it. The only notable exception to the latter case would be if the bicycle were cut out of the picture, in which case a CV model may already struggle to correctly identify the bicycle. The constraints for height discussed above are expressed arithmetically in Eq 4.

$$Height\ Constraint = \begin{cases} 1, & if\ 0.25 \leq h_b/h_p \leq 1 \\ 0, & otherwise \end{cases} \tag{4}$$

Metric 4: Width Constraint

A width constraint was also developed in order to further refine the detection model. Handling a constraint for width was rather difficult, as the width of bicycles was far more variable in the training dataset than height, depending on the angle of the photo taken. An example of this situation can be shown in Figure 3 below, in two photos of cyclists taken from the COCO dataset. In 3a, the width of a bicycle in a photo of a cyclist taken from the side is wider than the width of the person riding the bike. In 3b, the photo of the cyclist taken from behind results in the bike's width being much narrower than the person's width. Cases similar to 3b made up many of the cyclist images that weren't successfully captured by the existing overlap constraint. Because of this, we want to make an effort to capture those cyclists in our dataset as well.

(a) Width ratio of cyclist - side view

(b) Width ratio of cyclist - front/back view

Figure 3 Examples of width ratio.
(Lin et al., 2014)

Since it would be difficult to come up with two separate width metrics that could account for both the side views and the front views of bikes in our images, an attempt was made to create a width metric that could encompass both scenarios shown in Figure 3. This unified metric is shown in Eq 5, which takes the ratio of the width of the bicycle to the width of the person, or the ratio of the width of the person to the width of the bicycle, depending on which of these two values is larger. From there, it was decided that a width ratio of 2.5 would be used as our width constraint, as it was thought that from a side view, a bike would never be over 2.5 times the width of the pedestrian, and from a front/back view, it would be rare for the width of the person to be over 2.5 times the width of the bicycle. Eq 6 shows this constraint mathematically.

$$Width\ Ratio\ W = max\ (\frac{w_b}{w_p}, \frac{w_p}{w_b}) \tag{5}$$

$$Height\ Constraint = \begin{cases} 1, & if\ W \leq 2.5 \\ 0, & otherwise \end{cases} \tag{6}$$

Metric 5: Tracking Constraint

The final metric devised for the ZSL model is a tracking constraint. This constraint would work during the testing phase, when the model would be used on video footage to detect a cyclist moving. Object tracking is different from object detection in that object tracking assigns unique IDs to each object detected in a model. The goal of object tracking is to track each unique object in a video, whereas the goal of object detection is to detect any object in a video without caring about differences between objects in the same class. Object tracking is an important part to include in our model, as we do want to be able to track an individual cyclist as it moves through a video. The ultimate goal of our tracking constraint is to recognize when a person and a bicycle in a video are moving in the same direction, indicating that the person may likely be a cyclist. The first step to finding an object's direction is to find out how the object's position has changed in the video. As shown in Eq 7 and 8, the change in position of an object i in video frame n is defined as the difference between the object's x or y position, and the object's x or y position from 5 frames ago. From here, we define the object's direction of travel Dx and Dy as positive, negative, or zero, as shown in Eq 9 and 10. Finally, our tracking metric will determine when a person and bicycle are moving as a cyclist when the person and bicycle are moving in the same x and y directions, as shown in Eq 11.

$$Change\ in\ x\ position = \delta_x = x_{i,n} - x_{i,n-5} \tag{7}$$

$$Change\ in\ y\ position = \delta_y = y_{i,n} - y_{i,n-5} \tag{8}$$

$$D_x = \begin{cases} 1, & if\ \delta_x > 0 \\ -1, & if\ \delta_x < 0 \\ 0, & otherwise \end{cases} \tag{9}$$

$$D_y = \begin{cases} 1, & if\ \delta_y > 0 \\ -1, & if\ \delta_y < 0 \\ 0, & otherwise \end{cases} \tag{10}$$

$$Tracking\ Constraint = \begin{cases} 1, & if\ D_{x,p} = D_{x,b}\ and\ D_{y,p} = D_{y,b} \\ 0, & otherwise \end{cases} \tag{11}$$

Since the training dataset consisted of still images only, it was difficult to determine where the tracking constraint should be implemented into our ZSL model without causing any drastic changes to the existing model. Ultimately, it was decided that the tracking constraint was measuring characteristics of cyclists that were similar to those measured by the pedaling constraint. Therefore, the tracking constraint was integrated into the pedaling constraint. In short, if the pedaling constraint wasn't met in a given frame of a video, but the tracking constraint determined that the person and bicycle were moving in the same direction, then it would be assumed that the pedaling constraint was met. Eq 12 shows how these constraints interact with one another during the testing phase.

$$Pedaling + Tracking\ Constraint = \begin{cases} 1, & if\ \ y_b \le y_p + h_p \le y_b + h_b \\ & \quad or\ D_{x,p} = D_{x,b}\ and\ D_{y,p} = D_{y,b} \\ 0, & otherwise \end{cases} \tag{11}$$

3.3 Determination

With our constraints determined, our tracking model was developed by using these metrics to train a decision tree classification. While the creation of the decision tree classifier required the ground truth dataset to be created first, it was assumed that the results of the decision tree classifier model would be simple enough that a human could've come up with it without any ground truth data. Furthermore, using the decision tree classifier allowed us to compare where we made splits in our model to where a machine-made splits in its model, as seen in Experiment 1. The decision tree was restricted to a maximum depth of 4, and further splits in the tree could not be made in the model if a potential leaf contained less than 1% of the training data. The final results of the decision tree classifier are shown in Figure 4. In these results, the height and width constraints that were developed did not prove to be significant factors in determining the presence of a cyclist. While this likely means that better height and width constraints probably could have been implemented, it could be argued that finding out that fewer variables were needed to detect a cyclist is a preferable finding, as it means that less work is needed to get a good ZSL model working to detect cyclists. However, this argument hinges on the ZSL model's ability to perform well, which will be explored more in the following section.

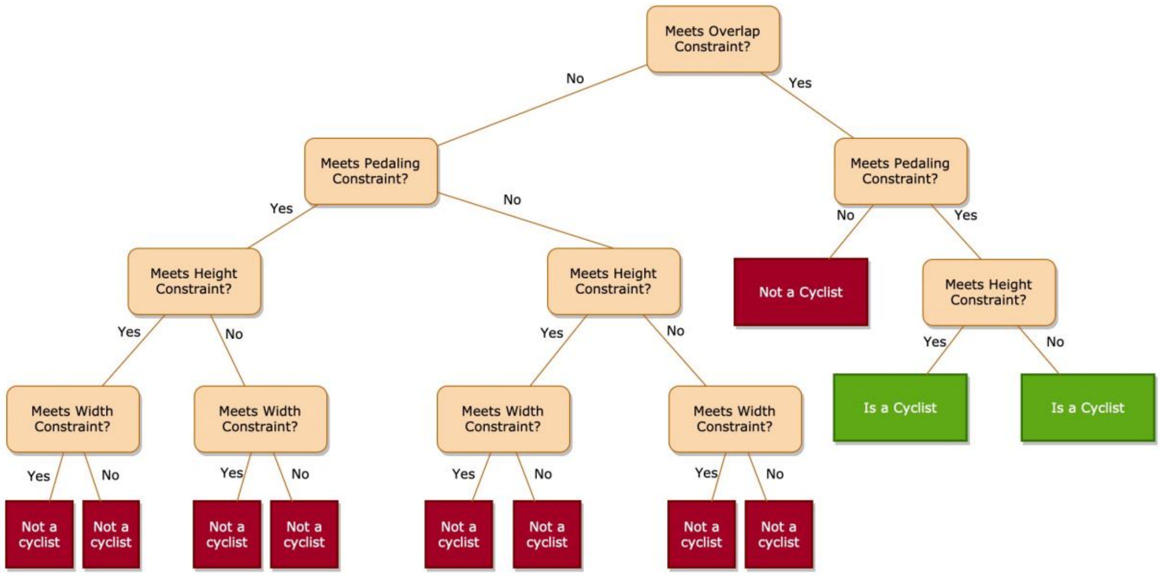


Figure 4 Decision Tree of our ZSL Model.

4. EXPERIMENT

During the evaluation process of this research, two experiments were run, one during the training phase and one during the testing phase. In all three experiments, a predicted cyclist is considered a correct detection if the IOU of a predicted cyclist's bounding box and a ground truth cyclist's bounding box is above 0.5.

4.1 Training Phase

During the training phase, the training dataset was created from images of people and bicycles was split into an 80% training / 20% testing split. The trained models are then used to predict the presence of cyclists in the testing images, and the model's precision, recall, and accuracy are extracted.

Experiment 1: Human Training Parameters vs. Standard Machine Learning

This first experiment aims to evaluate how well our human-trained ZSL model compares to machine learning models trained traditionally with access to ground truth data. To achieve this, two additional models were developed using our training data. The first model was a decision tree classifier, subjected to the same constraints as our ZSL model, as listed in the Determination section. However, this new model utilized the raw continuous values of our four training variables, allowing the machine to determine the optimal constraint thresholds. The second model was a logistic regression model trained on the continuous variables to develop the constraints. Unlike the decision tree classifier, the logistic regression model is not as easily interpretable by humans. The purpose of comparing the ZSL model to the logistic regression model is to evaluate how well our ZSL model performs against a sophisticated machine learning model. The results of this analysis can be observed in Table 1.

Table 1 Performance of Training Models.

Metric	ZSL Decision Tree	Machine-assisted Decision Tree	Logistic Regression
Precision	0.959	0.941	0.915
Accuracy	0.758	0.915	0.946
Recall	0.766	0.967	0.915

From these results, our ZSL model performed comparably well against the machine-learning models. Our ZSL model had high precision (0.959), indicating its effectiveness at ruling out false positives. However, it had a lower recall (0.766), meaning it missed more actual cyclists than the machine-assisted models.

In the machine-assisted decision tree model, the first split occurred at an IOU of about 0.17, which is very close to where we determined the first split to occur in our ZSL model. This similarity suggests that a human supervisor could effectively determine where the ZSL model should split the data to achieve good results. The machine-assisted decision tree showed a balance between precision (0.941) and recall (0.967), suggesting it effectively identified cyclists with fewer false negatives and false positives.

The logistic regression model, while achieving the highest accuracy (0.946) and recall (0.915), had lower precision (0.915), indicating it identified all cyclists but with a higher rate of false positives. The logistic regression model's coefficient values indicate that the overlap and pedaling constraints were the most significant predictors of cyclists, with the overlap constraint having the highest coefficient. This justifies our decision to use only our ZSL model's overlap and pedaling constraints.

When compared to the logistic regression model, our ZSL model outperformed it in terms of precision but did worse in terms of recall. This means our ZSL model effectively ruled out false positives but struggled more with false negatives. As mentioned in the Pedaling constraint section, our goal was to minimize the number of false negatives in our dataset, so our model's poorer performance in this area is a disappointing finding. However, since the COCO dataset is very broad, it is worth contextualizing these false negatives and false positives, as our testing phase has a narrower application of detecting cyclists from traffic cameras. Figure 5 shows multiple examples of false positive and false negative errors within our training data.

Figure 5 Examples of False Negative (FN) and False Positive (FP) Errors in Experiment 1.
(Lin et al., 2014)

While the images in Figure 5 do not encompass every error within our dataset, they do exemplify certain trends that were observed throughout the erroneous images. In cases like 5a, false negative errors occurred when cyclists were pictured from behind or from the front, which was an anticipated result if our width constraint proved to be ineffective. Since adjusting the width constraint at this point would violate our ZSL assumption, we hoped that in our testing footage, such scenarios would be rare in traffic camera settings. In cases like 5b, the ground truth cyclists are seen from very far away. During the testing phase, our selected detection models (YOLO) were noted to struggle with detecting objects that are far away, so we are not too concerned with the ramifications of this error. In cases like 5c, many cyclists present in an image at the same time would lead to a false negative error. This would be either due to a person or bicycle not being fully visible in the image (similar to the error seen in 5d), or the Hungarian method may have assigned a person to an incorrect overlapping bicycle. Scenarios where multiple cyclists were in close proximity to one another were not predicted to be a significant source of error during the testing phase of the project. In rare cases like 5e, you would have an image of a cyclist that does not neatly fit into any of the aforementioned error types but is still recognized as a false negative error as it did not meet either the overlap constraint or the pedaling constraint. Conversely, in this type of error, a majority of false positives were made up of images similar to 5f, where the overlap and pedaling constraints were both met in an image where a person was clearly not riding a bicycle. False positives in the dataset are not a big concern, given how effectively our ZSL model eliminates false positives. Therefore, when going into the testing phase of our dataset, we are most concerned about instances similar to 5a or 5e throwing off the results of our model.

While false positives were minimal and acceptable, false negatives were more concerning, particularly in images where cyclists were partially occluded or photographed from challenging angles. These cases highlight areas where the ZSL model might need further refinement or where real-world testing conditions differ significantly from the training dataset. Overall, these results show that our ZSL model, despite its simplicity, performs well and justifies using the overlap and pedaling constraints for our ZSL model. The findings emphasize the need for robust models capable of handling diverse real-world scenarios in cyclist detection.

4.2 Testing Phase

Testing footage for our methodology was taken from footage live-streamed from Ichibangai Street in the Kabukicho entertainment district located in Tokyo, Japan (Kabukicho et al. 2, 2024). This street experiences heavy pedestrian traffic, with occasional cyclists. We selected 15 seconds of footage, comprising 375 frames, and manually annotated the ground truth location of the cyclist in each frame to evaluate the performance of three models. During these 15 seconds, one cyclist is seen biking towards the camera amidst multiple pedestrians. Figure 6a shows an individual frame of the testing footage, with a bounding box around the ground truth cyclist.


(a) Ground Truth Cyclist Annotation


(b) Predicted Cyclist: YOLOv5


(c) Predicted Cyclist: YOLOv8


(d) Predicted Cyclist: YOLOv9

Figure 6 Ground Truth Annotation and Cyclist Predictions on Single Frame of Testing Footage.
(Kabukicho Live Channel 2, 2024)

For the testing phase, we selected the YOLO family of models for video detection and ByteTrack for tracking across all experiments. The YOLO (You Only Look Once) models are computer vision models trained on the COCO dataset, designed to quickly identify objects in an image, similar to how a human would glance at it (Redmon et al., 2016). While the YOLO algorithm may be less precise and accurate than other popular computer vision models, particularly when struggling with small objects, it compensates with its speed (Sanchez et al., 2020). At the time of writing this paper, ten versions of YOLO are publicly available. Figures 6b through 6d show the objects detected in a single frame by the three models used in the experiment. In all three models, the predicted cyclist matches the ground truth cyclist shown in 6a.

In our experiments, we evaluate the effectiveness of our detection and tracking models using three key metrics: Multiple Object Tracking Accuracy (MOTA), Identity F1 Score (IDF1), and Higher Order Tracking Accuracy (HOTA). These metrics are standard in the field of object detection and tracking and were specifically chosen because they were used in the development of the ByteTrack algorithm (Zhang et al., 2022).

Evaluation Metrics

To evaluate the effectiveness of our detection and tracking models, we use several performance metrics that are standard in the field of object detection and tracking:

- **Multiple Object Tracking Accuracy (MOTA):** This metric combines false positives, false negatives, and identity switches to provide an overall accuracy measure for tracking multiple objects. Higher MOTA values indicate better tracking performance (Bernardin & Stiefelhagen, 2008).
- **Identity F1 Score (IDF1):** This metric evaluates the accuracy of identity assignments by considering both precision and recall of correct identity matches. It is particularly useful for assessing the consistency of object tracking over time (Ristani et al., 2016).
- **Higher Order Tracking Accuracy (HOTA):** HOTA is a comprehensive metric that evaluates detection, association, and localization accuracy in tracking. It provides a balanced measure of tracking performance across different aspects of the task (Luiten et al., 2021).

These metrics provide a robust framework for evaluating and comparing the performance of different object detection and tracking models, ensuring a comprehensive assessment of our proposed method.

In the context of this project, the MOTA measures how effectively our model can detect the ground truth cyclist throughout the video. It accounts for false positives, false negatives, and identity switches, offering a comprehensive measure of tracking accuracy. The IDF1 assesses how well our model can associate a cyclist's unique ID with the ground truth ID, considering both precision and recall of correct identity matches to evaluate tracking consistency over time. The HOTA is an aggregate measure of detection accuracy and association accuracy, evaluated over a range of IOU values to determine matches with the ground truth dataset. It provides a balanced evaluation of tracking performance across different aspects. Since our models must detect both people and bicycles to identify a cyclist, detected cyclists do not initially have unique IDs. To address this, the ID of a cyclist is inherited from the ID of the detected person, and the ground truth ID of a given cyclist is defined as the ID of the cyclist appearing in the most frames of the video. The confidence threshold for detecting an object was set to 0.25 for all models. These metrics ensure a comprehensive assessment of our proposed method, allowing for effective comparison and evaluation of different object detection and tracking models.

Experiment 2: Model Performance on State-of-the-Art Computer Vision Models

This experiment aims to evaluate our model's performance using widely used computer vision models that can detect people and bicycles. The three models tested in this experiment were YOLOv5, YOLOv8, and the recently released YOLOv9. The performance of the three models can be seen in Table 2.

Table 2 ZSL Model Performance on YOLO models.

CV Model	MOTA	IDF1	HOTA
YOLOv5su	0.939	0.538	0.915
YOLOv8s	0.941	0.222	0.924
YOLOv9s	0.878	0.148	0.902

Based on the detailed results from the experiment, we observed distinct differences in performance among the YOLOv5, YOLOv8, and YOLOv9 models. YOLOv5 demonstrated balanced performance across multiple metrics, making it a reliable choice for our application. It achieved a Multiple Object Tracking Accuracy (MOTA) of 0.939, with 25 true positives (TP) out of 47 potential cyclists, 22 false positives (FP), and 21 false negatives (FN). The Intersection over Union (IoU) scores for YOLOv5 were consistently high, indicating precise bounding box predictions, as evidenced by IoU values of 0.931 for frame 262 and 0.965 for frame 265. Furthermore, YOLOv5 maintained better identity consistency over time, reflected in its Identity F1 Score (IDF1) of 0.538, demonstrating effective identity tracking with fewer switches.

In contrast, YOLOv8 exhibited excellent detection accuracy, as indicated by its highest MOTA of 0.941. However, it faced significant challenges with identity association, resulting in a low IDF1 score of 0.222. This model detected 58 true positives but produced 252 false positives and 155 false negatives out of 310 potential cyclists, indicating over-detection issues. Despite high IoU values, such as 0.915 for frame 262 and 0.923 for frame 261, the frequent identity switches and misidentifications undermined its tracking reliability.

YOLOv9 showed similar issues to YOLOv8, with frequent identity switches and misidentifications. It achieved a MOTA of 0.878 with 30 true positives, 189 false positives, and 156 false negatives out of 219 potential cyclists. The IDF1 score for YOLOv9 was the lowest at 0.148, reflecting substantial difficulties in maintaining consistent identity assignments. While its bounding box predictions were accurate, with IoU values such as 0.943 for frame 261 and 0.954 for frame 249, the high number of false positives highlighted the need for improvement in tracking consistency.

When diving deeper into these results, it was discovered that a series of missing frames from the beginning and the end of the footage were shared across the three models. These shared frames included the starting frames where the cyclist went behind a crowd of people. The person riding the bicycle remained visible, but the bicycle itself was not detectable at the set confidence threshold of 25 %, so classifying the cyclist during this time was impossible.

Figure 7 shows a sequence of frames from the testing footage. In the first four frames, the cyclist could be detected, as the bicycle (shown with a purple bounding box) could be detected. However, in the following frames, the bicycle cannot be detected, so no cyclist can be classified.

In conclusion, while YOLOv8 and YOLOv9 demonstrated high detection accuracy, their overall performance was compromised by poor identity-tracking capabilities. YOLOv5, with its balanced performance, fewer false positives, and better identity consistency, emerged as the most reliable model for our application. This comprehensive evaluation underscores the importance of robust identity association algorithms in enhancing the overall performance of object detection and tracking models in real-world scenarios. Future work should aim to address the identity tracking limitations observed in YOLOv8 and YOLOv9 to improve their applicability in complex detection tasks.

Figure 7 Detected and Missing Frames in the YOLOv5 Model.
(Kabukicho Live Channel 2, 2022)

5. CONCLUSIONS AND FUTURE WORK

Our research demonstrates that using a custom Zero-Shot Learning (ZSL) model, combined with pre-trained YOLO models and human-derived constraints, is effective for detecting cyclists as an unseen class. By applying constraints to pre-trained YOLOv5, YOLOv8, and YOLOv9 models, we successfully identified cyclists without requiring additional training data specific to cyclists. Overall, our ZSL approach performed well in most experiments, scoring above 70% in most testing metrics, except for some models where performance varied.

One notable area for improvement is the detection of cyclists based on the angle at which their picture is captured. Our approach attempted to create a width constraint that could accommodate both side and front/back views of cyclists rather than using angle-based classification. This method did not yield the desired results and contributed significantly to false negative errors. Future work could explore angle-based classification similar to the methodology used by Li et al. (2016) to improve detection accuracy.

Another area for enhancement is the tracking constraint. With limited knowledge of tracking performance prior to testing, our implementation was conservative. However, after observing the YOLOv8 model's tracking performance, it is evident that more sophisticated tracking algorithms could improve cyclist detection across multiple frames, especially when the bicycle is temporarily not visible.

Exploring different detection models for the video tracking phase could also improve the ZSL model. While the YOLO family of models was chosen to align with zero-shot learning objectives, alternative models with higher accuracy might be more suitable for VRU safety applications. These models could potentially offer better performance in detecting and tracking cyclists in complex environments.

In conclusion, while our ZSL model shows significant promise, particularly in ruling out false positives and maintaining identity consistency, there is ample room for enhancement in constraint formulation, tracking, and the choice of detection models. Our study contributes to the existing literature by demonstrating the feasibility of applying zero-shot learning techniques to cyclist detection, highlighting both the strengths and limitations of this approach. Future research should focus on refining these methods to improve overall detection and tracking performance, ultimately enhancing road safety for vulnerable road users.

6. ACKNOWLEDGMENTS

The authors extend their sincere appreciation to the UW STAR Lab for its invaluable resources and support, which made this research possible. The collaborative environment and access to state-of-the-art facilities provided by the UW STAR Lab were essential to the success of this work. The authors also acknowledge the University of Washington, Seattle for the resources that significantly contributed to the research's accomplishments.

7. REFERENCES

Abadi, A. D., Gu, Y., Goncharenko, I., & Kamijo, S. (2022, January). Detection of cyclists' crossing intentions for autonomous vehicles. *2022 IEEE International Conference on Consumer Electronics (ICCE)*, 1–6. https://doi.org/10.1109/ICCE53296.2022.9730559

Abdel-Gawad, A. H., Khamis, A., Said, L. A., & Radwan, A. G. (2021, December). Vulnerable road users detection and tracking using YOLOv4 and Deep SORT. 2021 9th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC), 140–145. https://doi.org/10.1109/JAC-ECC54461.2021.9691441

Bernardin, K., & Stiefelhagen, R. (2008). Evaluating multiple object tracking performance: The CLEAR MOT metrics. EURASIP Journal on Image and Video Processing, 2008(1), 246309. https://doi.org/10.1155/2008/246309

Dimitrievski, M., Van Hamme, D., Veelaert, P., & Philips, W. (2020). Cooperative multi-sensor tracking of vulnerable road users in the presence of missing detections. *Sensors*, *20*(17), 4817. https://doi.org/10.3390/s20174817

Elhenawy, M., Ashqar, H. I., Masoud, M., Almannaa, M. H., Rakotonirainy, A., & Rakha, H. A. (2020). Deep transfer learning for vulnerable road users detection using smartphone sensors data. *Remote Sensing*, *12*(21), 3508. https://doi.org/10.3390/rs12213508

Fang, Z., & López, A. M. (2019). Intention recognition of pedestrians and cyclists by 2D pose estimation. *IEEE Transactions on Intelligent Transportation Systems, 21*(11), 4773–4783. https://doi.org/10.1109/TITS.2019.2923319

Garcia-Venegas, M., Mercado-Ravell, D. A., Pinedo-Sánchez, L. A., & Carballo-Monsivais, C. A. (2021). On the safety of vulnerable road users by cyclist detection and tracking. *Machine Vision and Applications, 32*(5), 109. https://doi.org/10.1007/s00138-021-01231-4

Goldhammer, M., Köhler, S., Zernetsch, S., Doll, K., Sick, B., & Dietmayer, K. (2020). Intentions of vulnerable road users—Detection and forecasting by means of machine learning. *IEEE Transactions on Intelligent Transportation Systems, 21*(7), 3035–3045. https://doi.org/10.1109/TITS.2019.2923319

Hirota, A., Tiep, N. H., Van Khanh, L., & Oka, N. (2017). Classifying helmeted and non-helmeted motorcyclists. In F. Cong, A. Leung, & Q. Wei (Eds.), Advances in Neural Networks - ISNN 2017 (pp. 81–86). Springer International Publishing. https://doi.org/10.1007/978-3-319-59072-1_10

Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A review of YOLO algorithm developments. *Procedia Computer Science*, *199*, 1066–1073. https://doi.org/10.1016/j.procs.2022.01.135

Kabukicho Live Channel 2. (2024, April). *Tokyo Shinjuku Kabukicho Live 2 / Kabukicho Live Channel II* [YouTube video]. YouTube. https://www.youtube.com/watch?v=gFRtAAmiFbE (Note: Accessed on April 12, 2024)

Knura, M., Kluger, F., Zahtila, M., Schiewe, J., Rosenhahn, B., & Burghardt, D. (2021). Using object detection on social media images for urban bicycle infrastructure planning: A case study of Dresden. *ISPRS International Journal of Geo-Information*, *10*(11), 733. https://doi.org/10.3390/ijgi10110733

Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, *2*(1-2), 83–97. https://doi.org/10.1002/nav.3800020109

Li, X., Flohr, F., Yang, Y., Xiong, H., Braun, M., Pan, S., Li, K., & Gavrila, D. M. (2016, June). A new benchmark for vision-based cyclist detection. *2016 IEEE Intelligent Vehicles Symposium (IV)*, 1028–1033. https://doi.org/10.1109/IVS.2016.7535515

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), Computer Vision – ECCV 2014 (Vol. 8693, pp. 740–755). Springer International Publishing. https://doi.org/10.1007/978-3-319-10602-1_48

Luiten, J., Ošep, A., Dendorfer, P., Torr, P., Geiger, A., Leal-Taixé, L., & Leibe, B. (2021). HOTA: A higher order metric for evaluating multi-object tracking. *International*

Journal of Computer Vision, 129(2), 548–578. https://doi.org/10.1007/s11263-020-01375-2

Maurya, S. K., & Choudhary, A. (2018, September). Deep learning-based vulnerable road user detection and collision avoidance. *2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, 1–6. https://doi.org/10.1109/ICVES.2018.8519504

Mistry, J., Misraa, A. K., Agarwal, M., Vyas, A., Chudasama, V. M., & Upla, K. P. (2017, November). An automatic detection of helmeted and non-helmeted motorcyclist with license plate extraction using convolutional neural network. *2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 1–6. https://doi.org/10.1109/IPTA.2017.8310101

National Highway Traffic Safety Administration. (2023, April 7). Data visualization - Fatality analysis reporting system (FARS). https://www.nhtsa.gov/file-downloads

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021, February 26). Learning transferable visual models from natural language supervision. *arXiv*. https://doi.org/10.48550/arXiv.2103.00020

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). https://doi.org/10.1109/CVPR.2016.91

Ristani, E., Solera, F., Zou, R., Cucchiara, R., & Tomasi, C. (2016). Performance measures and a data set for multi-target, multi-camera tracking. In G. Hua & H. Jégou (Eds.), *Computer Vision – ECCV 2016 Workshops* (pp. 17–35). Springer International Publishing. https://doi.org/10.1007/978-3-319-48881-3_2

Romera-Paredes, B., & Torr, P. (2015). An embarrassingly simple approach to zero-shot learning. In *Proceedings of the 32nd International Conference on Machine Learning* (pp. 2152–2161). https://proceedings.mlr.press/v37/romera-paredes15.html

Sanchez, S. A., Romero, H. J., & Morales, A. D. (2020). A review: Comparison of performance metrics of pretrained models for object detection using the TensorFlow framework. *IOP Conference Series: Materials Science and Engineering, 844*(1), 012024. https://doi.org/10.1088/1757-899X/844/1/012024

Wang, C.-Y., Yeh, I.-H., & Liao, H.-Y. M. (2024, February 28). YOLOv9: Learning what you want to learn using programmable gradient information. *arXiv*. https://doi.org/10.48550/arXiv.2402.13616

Xian, Y., Lampert, C. H., Schiele, B., & Akata, Z. (2019). Zero-shot learning—a comprehensive evaluation of the good, the bad and the ugly. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 41*(9), 2251–2265. https://doi.org/10.1109/TPAMI.2018.2857768

Zhang, Y., Ling, Q. (2017, June 11). Bicycle detection based on multi-feature and multi-frame fusion in low-resolution traffic videos. *arXiv*. https://www.semanticscholar.org/paper/9f91fd3e9621b88769ecc330f362a591876f948f

Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W., & Wang, X. (2022, April 7). ByteTrack: Multi-object tracking by associating every detection box. *arXiv*. https://doi.org/10.48550/arXiv.2110.06864

The study of the Reflectance Ratio on Highway Pavement Surfaces by using High-Pressure Sodium (HPS) Lamp Type

Meeampol Ratanachai.[1], Ngosorn Pattanapong.[2], Benjapornthavee Chainarong.[3], Sedokbuab Pattarapon.[4]

[1] Bureau of Road Research and Development, Department of Highways, ratan.me@doh.go.th
[2] Bureau of Road Research and Development, Department of Highways, patta.ng@doh.go.th
[3] Bureau of Road Research and Development, Department of Highways, chain.be@doh.go.th
[4] Bureau of Road Research and Development, Department of Highways, s_pattarapon@hotmail.com

ABSTRACT

According to European standards, EN13201 standards, a report or system defines an appropriate lighting class in outdoor public areas in terms of parameters that later classify each ranking of pavement. Thailand has no specific specifications but mainly relies on European standards (EN) and the Illuminating Engineering Society (IES). This research aims to study the Reflectance Ratio of asphalt and concrete pavement. The research focuses on data from the field to analyze the Reflectance Ratio on the pavement during the night. Firstly, the field ranges from 0-1 year, 1-5 years, and more than 5 years for both asphalt concrete pavement and concrete pavement. Secondly, Electrical lighting equipment is a high-pressure sodium (HPS) type used for no more than 2 years. Lastly, the Luxmeter tool is used to collect Illuminance Data, and the Imaging Luminance measuring device (ILMD) is used to measure Luminance Data using a photometer imaging light technique. As a result, the reflection ratio of asphalt concrete pavement surfaces tends to develop Reflectance Ratio over time due to less intense skin colour, resulting in a better light reflection. However, the reflection ratio of concrete pavement surfaces tends to reduce over time due to the color of the concrete surface becoming darker due to the time passing by and the amount of use.

Keywords: Traffic Engineering, Transportation Engineering, Reflectance Ratio,

1. INTRODUCTION

Road safety consists of three main factors: driver readiness, vehicle readiness, and environmental conditions, especially during nighttime when ecological factors such as lighting significantly affect driving safety. If the lighting is insufficient, it can impair substantially visibility, making it difficult to see objects such as road edges or pedestrians crossing the road clearly while driving (Bureau of Road Research and Development, 2022). Moreover, Illuminance is currently used in the design and measurement of light intensity from lighting systems. Illuminance is the quantity of light falling on an object per unit area. However, there are standards for designing and measuring the intensity of light on roads using Luminance, which is the value of light that falls on an object and then reflects into the eyes of the driver (Bureau of Road Research and Development, 2019).

Upon closer examination, it is found that determining Luminance and designing street lighting poles and luminous electric lamps still rely on Illuminance principles, measured in lux per square meter or lux. Suppose Illuminance is used in the design process. In that case, users will perceive different road surfaces depending on the type of material used to make the road surface because Illuminance do not consider the coefficient of light reflection of the road surface or the ratio of light reflection on the surface. Furthermore, the engineering principles of luminous electric lighting design on highways are not explained. Using Illuminance in the design process involves considering the coefficient of light reflection of the road surface or the ratio of light reflection on the surface, allowing road users to see the road surface, regardless of the material used, and considering the appropriateness of the quality of light from various environmental factors such as traffic volume. This ensures that Illuminance is appropriate for the physical characteristics of different roads

2. LITERATURE REVIEW

2.1 Purpose of installing electrical lighting on the highway

The primary purpose of installing electrical lighting on highways is to enhance driving ability at night by providing accurate visibility of routes and roadside objects within a short time frame. Road users can maneuver effectively in emergencies and help reduce the likelihood of accidents. When designing the installation of electric lighting, the designer must consider the actual usage conditions of the road users during installation. Efficient and effective electric lighting provides road users with economic, social, and safety benefits [2]. The following objectives should be considered:

- To reduce road accidents at night and minimize the loss of life and property for highway users.
- To reduce crime and increase safety for highway users.
- To increase mobility and visibility for highway users.

2.2 High-Pressure Sodium (HPS) bulbs

HPS bulb is a light bulb that generates an electric charge from the evaporation of sodium in iron. It has a color temperature of light ranging from 2200 to 2400 Kelvin.

2.3 Factors that affect vision include

- Size and distance: Perceived size is contingent upon the Visual Angle. The Visual Angle broadens as the distance diminishes, resulting in the perception of larger objects.
- Illuminance: This quantifies the light falling on a specified surface area, typically measured in lux.
- Luminance: This photometric measure gauges the intensity of density of luminous intensity in a specified direction. It denotes the light emitted by an object and captured within a solid angle formed by the plane of the pupil or light reflection from a flat surface scattering light. Luminance delineates the Brightness perceived by our eyes when observing a surface from a particular angle, expressed in units of cd/m^2

2.4 Methods for measuring Illuminance and Luminance

The goal is to develop standards that can be used correctly and efficiently according to international principles, specifically EN 13201. Here are the methods for measuring Illuminance and Luminance with theoretical examples (Department of Alternative Energy Development and Efficiency, 2004).

- Illuminance: This measures the light incident on a surface per unit area. It may be called the Lighting Illuminance Level, indicating how much light an area receives. The unit is lumens per square meter or lux. Appropriate values for each area can be found in the recommendations from the TIEA - GD 003 standard of the Lighting Electrical Association of Thailand.
- Luminance: This measures the light reflected from any surface in a particular direction per unit area. Also called Brightness, it shows that the same amount of light falling on objects of different colors results in various amounts of reflected light, making objects appear to have different brightness levels. The unit is candela per square meter.
- The Reflectance Ratio on the pavement surface can be calculated using the Illuminance and Luminance equation.

$$\text{Reflectance Ratio } \left(\frac{\frac{cd}{lx}}{m^2}\right) = \frac{\text{Luminance } \left(\frac{cd}{m^2}\right)}{\text{Illuminance } (lx)}$$

2.5 Road classification according to IES RP-8-18 standards (Federal Highway Administration, 2018)

- Road Class R1: Reflectance Ratio of not less than 0.10 (Portland cement road/asphalt concrete road that reflects light in almost all directions).
- Road Class R2: Reflectance Ratio of not less than 0.07 (asphaltic concrete roads made from gravel-type aggregates/asphalt concrete roads made from synthetic aggregates that increase light reflection with mixed light reflection).
- Road Class R3: Reflectance Ratio of not less than 0.07 (asphaltic concrete roads made of black aggregates that reflect light directly).
- Road Class R4: Reflectance Ratio of not less than 0.08 (very smooth asphalt concrete road that reflects almost all direct light).

3. METHODOLOGY

3.1 Research preparation

1. Study related documents and research to determine the objectives and scope of the research.
2. Select routes consists of six areas which are divided into three age groups of pavement surfaces: 0 - 1 year, 1 - 5 years, and more than 5 years.
3. Conduct a study on the routes under the Department of Highways on asphalt concrete pavement and concrete surfaces.
4. Collect field data, analyze results, and summarize the data.

3.2 Research Tools

1. Luxmeter to collect Illuminance Data.
2. ILMD (Imaging Luminance Measuring Device) to collect Luminance Data.

3.3 Target Group/Population and Sample Group

The sample group includes routes under the responsibility of the Department of Highways:

- Pavement age of 0 - 1 year, consisting of one area of asphalt pavement and one concrete pavement.
- Pavement age of 1 - 5 years, consisting of one area of asphalt pavement and one concrete pavement.
- Pavement age of more than 5 years, consisting of one area of asphalt pavement and one area of concrete pavement.

4. RESULTS

4.1 The result on Asphalt Concrete Pavement Surfaces (IES RP-8-18 Standard, Road Class R3)

4.1.1 For Asphalt Concrete Pavement of 0 - 1 year, sample 1

Traffic volume was 24,031 vehicles per day. using single branch type light poles were installed in a parallel arrangement along the middle of Highway 347, Section 0200, from Bang Krasan to Bang Pahan, at km. 48+500. The average Illuminance is 24.98 Lux, the average Luminance is 1.50 cd/m², and the calculated average Reflectance Ratio is 0.068, which is slightly below the threshold value of 0.07, as shown in Figure 1.

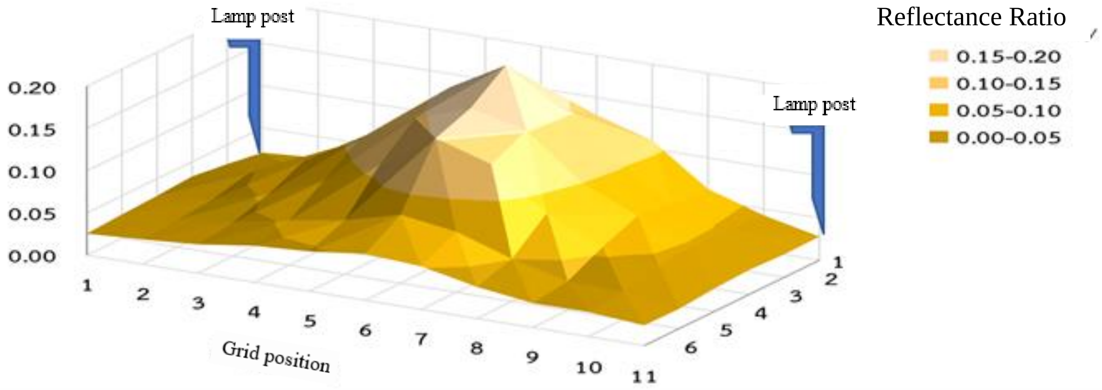


Figure 1 shows the Reflectance Ratio of streetlights on Highway No.347

4.1.2 For Asphalt Concrete Pavement of 0 - 1 year, sample 2

Traffic volume of 13,146 vehicles/day, using single branch type light poles parallelly installed on the left edge of Highway No.3121, section 0100, named Bang Kla - Long Yao, area km 5+700. The average Illuminance is 20.83 Lux, the average Luminance is 1.54 cd/m², and the calculated average Reflectance Ratio is 0.085, which is exceeding the threshold value of 0.07, as shown in Figure 2.

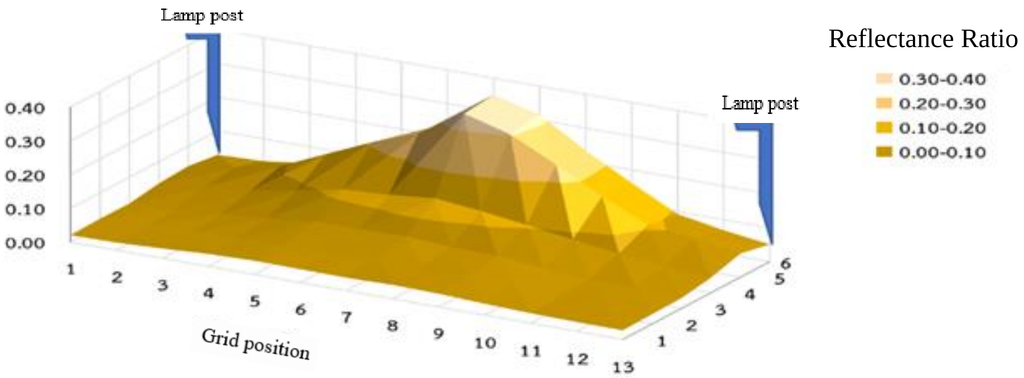


Figure 2 shows the Reflectance Ratio of streetlights on Highway No.3121

4.1.3 For Asphalt Concrete Pavement of 1 - 5 years, sample 1

Traffic volume of 8,792 vehicles per day, using single branch type light poles parallelly installed on the left edge on Highway No.3195, section 0201, named Lat Tan - Wiset Chai Chan, at kilometer 10+00, the average Illuminance is 46.32 Lux, the average Luminance is 2.88 cd/m2, and the calculated average Reflectance Ratio is 0.066, which is slightly below the threshold value of 0.07, as shown in Figure 3.

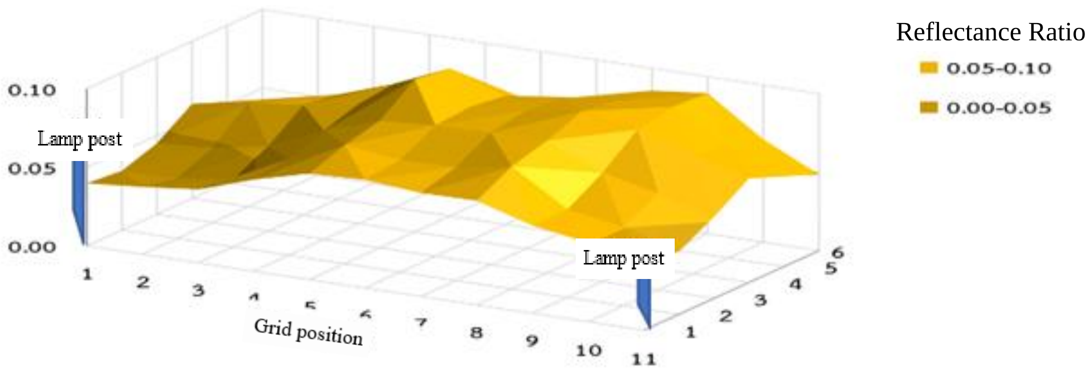


Figure 3 shows the Reflectance Ratio of streetlights on Highway No.3195

4.1.4 For Asphalt Concrete Pavement of 1 - 5 years, sample 2

Traffic volume of 23,871 vehicles/day, using single branch type light poles parallelly installed on the left edge on Highway No.352, Section 0100, named Thanyaburi - Khlong Rapeepat, at kilometer 9+100, the average Illuminance is 13.31 Lux, the average Luminance is 0.85 cd/m2, and the calculated average Reflectance Ratio is 0.082, which is exceeding the threshold value of 0.07, as shown in Figure 4.

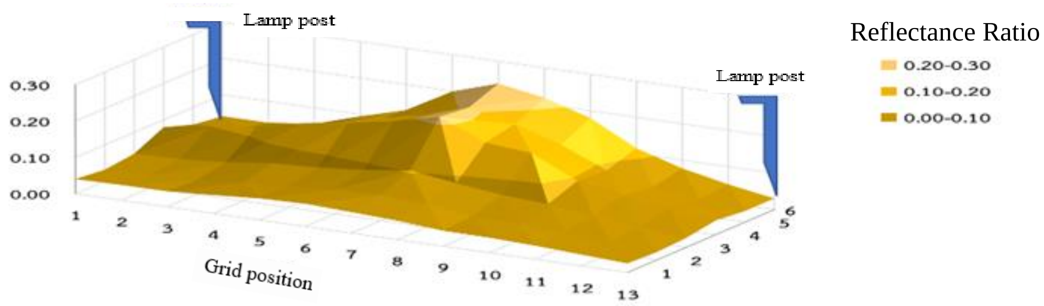


Figure 4 shows the Reflectance Ratio of streetlights on Highway No.352

4.1.5 For Asphalt Concrete Pavement of more than 5 years, sample 1

Traffic volume of 5,844 vehicles/day, using single branch type light poles parallelly installed on the right edge on Highway No.3603, section 0100, named Sam Ruean - Klang Wang Daeng, at km. 0+300, the average Illuminance is 27.21 Lux, the average Luminance is 1.75 cd/m², and the calculated average Reflectance Ratio is 0.073, which is exceeding the threshold value of 0.07, as shown in Figure 5.

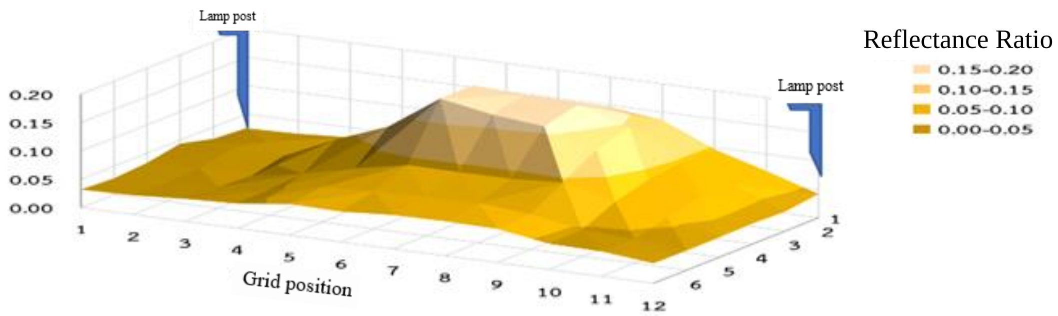


Figure 5 shows the Reflectance Ratio of streetlights on Highway No.3603

4.1.6 For Asphalt Concrete Pavement of more than 5 years, sample 2

Traffic volume of 6,713 vehicles/day, using single branch type light poles parallelly installed on the right edge on Highway No.3454, section 0100, named Na Khok - Sena, at km. 95+400, the average Illuminance is 11.35 Lux, the average Luminance is 1.17 cd/m2, and the calculated average Reflectance Ratio is 0.114, which is exceeding the threshold value of 0.07, as shown in Figure 6.

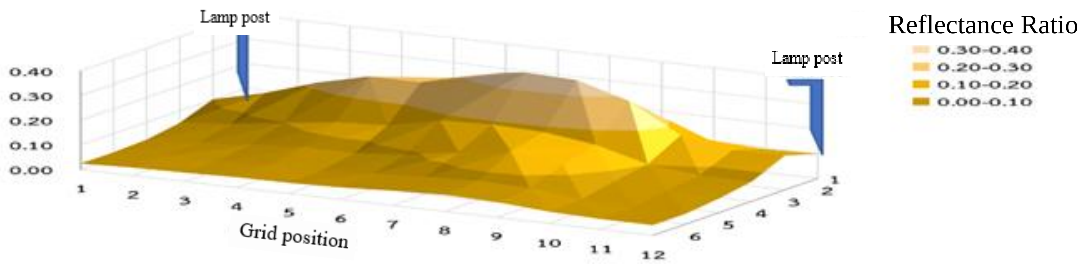


Figure 6 shows the Reflectance Ratio of streetlights on Highway No.3454

4.2 The result on Concrete Pavement Surfaces (IES RP-8-18 Standard, Road class R1)

4.2.1 For Concrete Pavement of 0 - 1 year, sample 1

Traffic volume of 24,552 vehicles/day, using single branch type light poles parallelly installed on the right edge on Highway 224, Section 0201, named Hua Thale - Chok Chai, at km. 11+600, the average Illuminance is 14.55 Lux, the average Luminance is 1.19 cd/m², and the average Reflectance Ratio is 0.095. which is slightly below the threshold value of 0.10, as shown in Figure 7.

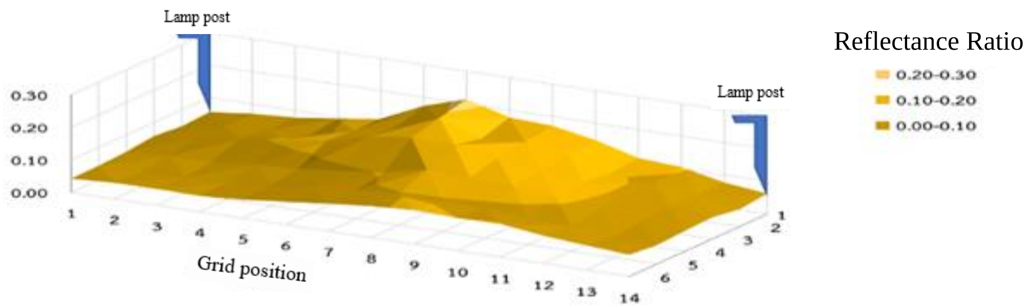


Figure 7 shows the Reflectance Ratio of streetlights on Highway No.224

4.2.2 For Concrete Pavement of 0 - 1 year, sample 2

Traffic volume of 14,938 vehicles/day, using single branch type light poles parallelly installed on the right edge on Highway 207, section 0100, named Ban Wat - Prathai, at km. 36+200, the average Illuminance is 23.43 Lux, and the average Luminance is 1.97 cd/m2, and the calculated average Reflectance Ratio is 0.095, which is slightly below the threshold value of 0.10, as shown in Figure 8.

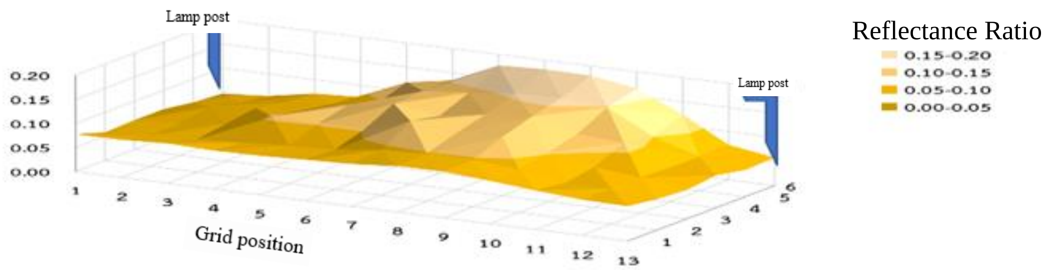


Figure 8 shows the Reflectance Ratio of streetlights on Highway No.207

4.2.3 For Concrete Pavement of 1- 5 years, sample 1

Traffic volume of 15,587 vehicles/day, using single branch type light poles parallelly installed on the left edge on Highway No.3267, Section 0100, named Chao Luek - Bang Khomt, at km. 13+600, the average Illuminance is 18.36 Lux, and the average Luminance is 2.21 cd/m2, and the calculated average Reflectance Ratio is 0.132, which is exceeding the threshold value of 0.10, as shown in Figure 9.

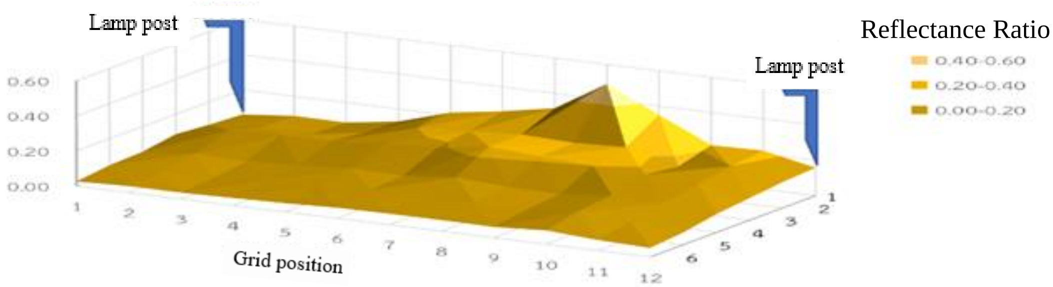


Figure 9 shows the Reflectance Ratio of streetlights on Highway No.3267

4.2.4 For Concrete Pavement of 1 - 5 years, sample 2

Traffic volume of 20,939 vehicles/day, using single branch type light poles parallelly installed on the left edge on Highway 346, section 0301, named Bang Len - Lam Luk Bua, at kilometer 48+800, the average Illuminance is 13.72 Lux, and the average Luminance is 1.08 cd/m², and the calculated Reflectance Ratio is 0.088, which is slightly below the threshold value of 0.10, as shown in Figure 10.

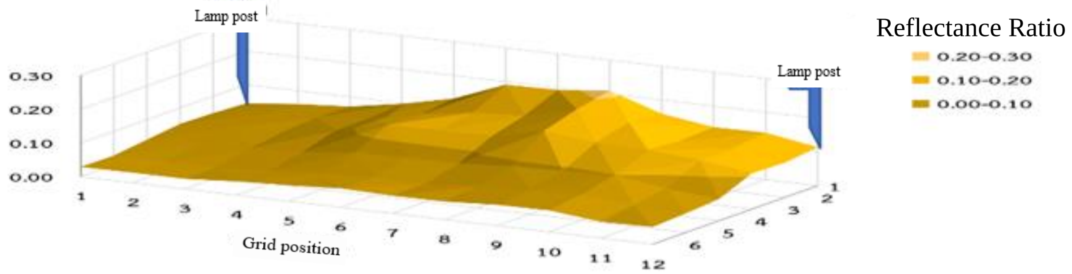


Figure 10 shows the Reflectance Ratio of streetlights on Highway No.346

4.2.5 For Concrete Pavement of more than 5 years, sample 1

Traffic volume of 5,883 vehicles/day, using single branch type light poles parallelly installed on the right edge on Highway No.3211, section 0100, named Hankha Entrance - Kabok Tia, in the area km. 2+300, the average Illuminance is 25.87 Lux, and the average Luminance is 1.81 cd/m², and the calculated Reflectance Ratio is 0.076, which is slightly below the threshold value of 0.10, as shown in Figure 11.

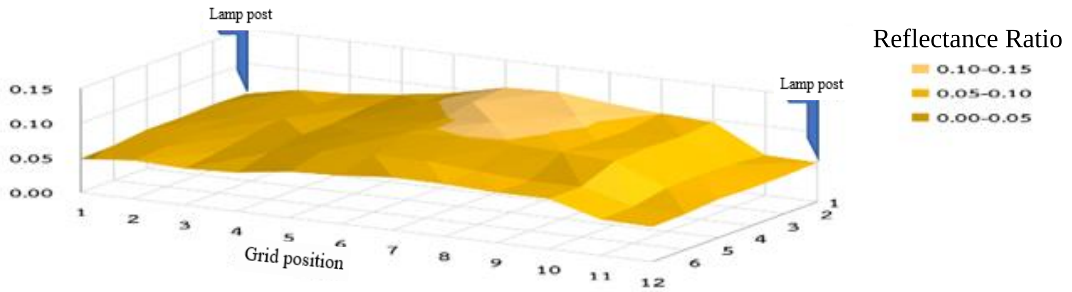


Figure 11 shows the Reflectance Ratio of streetlights on Highway No.3211

4.2.6 For Concrete Pavement of more than 5 years, sample 2

Traffic volume of 6,896 vehicles/day, using single branch type light poles parallelly installed on the right edge on Highway 340, section 0401, named Pak Nam - Sankhaburi Hospital, area km. 144+600, the average Illuminance is 32.34 Lux, the average Luminance is 1.60 cd/m², and the calculated Reflectance Ratio is 0.057, which is below the threshold value of 0.10, as shown in Figure 12.

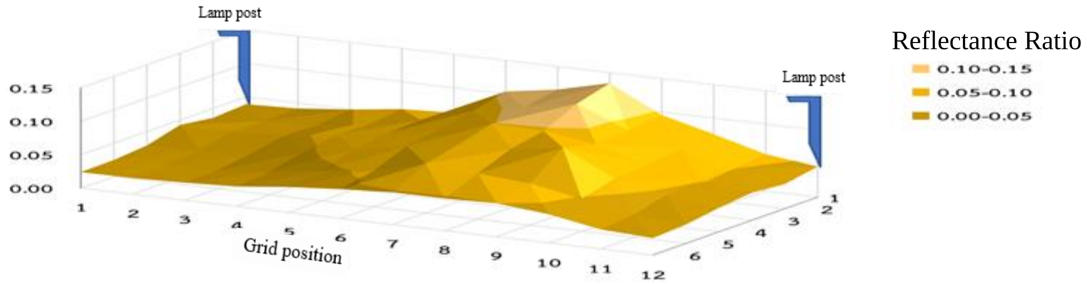


Figure 12 shows the Reflectance Ratio of streetlights on Highway No.340

4.3 Overall results of Reflectance Ratio of Asphalt Concrete Pavement and Concrete Pavement

The Overall results of Reflectance Ratio of Asphalt Concrete Pavement and Concrete Pavement as shown in Table 1

Table 1 The Overall results of Reflectance Ratio of Asphalt Concrete Pavement and Concrete Pavement.

Asphalt Concrete Pavement Surface (IES RP-8-18 Standard, Road Class R3)						
Pavement age (years)	0 - 1 year		1 - 5 years		More than 5 years	
Traffic volume (vehicles per day)	24,031	13,146	8,792	23,871	5,844	6,713
Average Illuminance (Lux)	24.98	20.83	46.32	13.31	27.21	11.35
Average Luminance (cd/m²)	1.5	1.54	2.88	0.85	1.75	1.17
Average Reflectance Ratio	0.068	0.085	0.066	0.082	0.073	0.114
IES RP-8-18 standards (Reflectance Ratio)	0.07					
Concrete Pavement Surface (IES RP-8-18 Standard, Road class R1)						
Pavement age (years)	0 - 1 year		1 - 5 years		More than 5 years	
Traffic volume (vehicles per day)	24,552	14,938	15,587	20,939	5,883	6,896
Average Illuminance (Lux)	14.55	23.43	18.36	13.72	25.87	32.34
Average Luminance (cd/m²)	1.19	1.97	2.21	1.08	1.81	1.6
Average Reflectance Ratio	0.095	0.095	0.132	0.088	0.076	0.057
IES RP-8-18 standards (Reflectance Ratio)	0.1					

5. CONCLUSIONS

The study results analyze the Reflectance Ratio of street lighting for asphalt concrete road surfaces according to the IES RP-8-18 standard, Roadway Class R3 and for concrete road surfaces according to the IES RP-8-18 standard, Roadway Class R1 can be shown in Table 2

Table 2 The Comparison of Reflectance Ratio of Asphalt Concrete Pavement and Concrete Pavement to the IES RP-8-18 standards.

Asphalt Concrete Pavement Surface (IES RP-8-18 Standard, Road Class R3)						
Pavement age (years)	0 - 1 year		1 - 5 years		More than 5 years	
Traffic volume (vehicles per day)	24,031	13,146	8,792	23,871	5,844	6,713
Average Illuminance (Lux)	24.98	20.83	46.32	13.31	27.21	11.35
Average Luminance (cd/m²)	1.5	1.54	2.88	0.85	1.75	1.17
Average Reflectance Ratio	0.068	0.085	0.066	0.082	0.073	0.114
IES RP-8-18 standards (Reflectance Ratio)	0.07					
	Lower	Higher	Lower	Higher	Higher	Higher
Concrete Pavement Surface (IES RP-8-18 Standard, Road class R1)						
Pavement age (years)	0 - 1 year		1 - 5 years		More than 5 years	
Traffic volume (vehicles per day)	24,552	14,938	15,587	20,939	5,883	6,896
Average Illuminance (Lux)	14.55	23.43	18.36	13.72	25.87	32.34
Average Luminance (cd/m²)	1.19	1.97	2.21	1.08	1.81	1.6
Average Reflectance Ratio	0.095	0.095	0.132	0.088	0.076	0.057
IES RP-8-18 standards (Reflectance Ratio)	0.1					
	Lower	Lower	Higher	Lower	Lower	Lower

For the age range of 0-1 year, with 2 test plots, it was found that plots with higher traffic volume had lower Reflectance Ratio compared to plots with lower traffic volume. Plots with higher traffic volume exhibited Reflectance Ratio that did not conform to the recommended values in international standards. This discrepancy may stem from the selected test plots experiencing surface abrasion due to heavy vehicular traffic, resulting in differences in the Reflectance Ratio from those specified by the standards.

For the age range of 1-5 years, with 2 test plots, it was found that plots with lower traffic volume had lower Reflectance Ratio compared to plots with higher traffic volume. Plots with lower traffic volume exhibited Reflectance Ratio that did not conform to the recommended values in international standards. This discrepancy may arise from the selected plots having been in use for a more extended period, with relatively lower traffic volume than plots with higher traffic volume, resulting in relatively minor light reflection. This may be attributed to the aggregate material's darker hue due to less usage.

For surfaces aged over 5 years, with 2 test plots, it was found that the Reflectance Ratio were in line with the recommended values in international standards. There was standardized light reflection. This may be attributed to both test plots having undergone surface abrasion due to heavy vehicular traffic and ageing, resulting in the aggregate material darkening and reducing overall light reflection to almost standardized levels.

The study results analyze the Reflectance Ratio of street lighting for concrete road surfaces according to the IES RP-8-18 standard, Roadway Class R1.

For the age range of 0-1 year, with 2 test plots, it was found that the Reflectance Ratio were in line with the recommended values in international standards. There was standardized light reflection. This may be attributed to both test plots having undergone surface abrasion due to heavy vehicular traffic, resulting in almost standardized light reflection for concrete road surfaces.

For the age range of 1-5 years, with 2 test plots, it was found that plots with higher traffic volume had lower Reflectance Ratio compared to plots with lower traffic volume. Plots with higher traffic volume exhibited Reflectance Ratio that did not conform to the recommended values in international standards. This discrepancy may arise from the selected test plots experiencing surface abrasion due to heavy vehicular traffic, resulting in differences in the Reflectance Ratio from those specified by the standards.

For surfaces aged over 5 years, with 2 test plots, it was found that the Reflectance Ratio did not conform to the recommended values in international standards. This may be attributed to both test plots having undergone surface abrasion due to heavy vehicular traffic and ageing, resulting in darkening aggregate material and reducing light reflection quality.

Thus, it can be concluded from the testing of both asphalt concrete and concrete surfaces that asphalt surfaces, as they age and are subject to usage, tend to exhibit Reflectance Ratio under the recommended standards. This may be due to the fading of surface color, resulting in improved light reflection. Conversely, as they age and are subject to usage, concrete surfaces tend to have lower values than the recommended standards. This may be attributed to the concrete surface becoming darker with age and usage.

6. ACKNOWLEDGEMENTS

The authors would like to thank the Director of the Bureau of Road Research and Development, Department of Highways, Mr. Ponlathep Lertworawanich for his encouragement in establishing this paper. Also, to thank our incredible mentor Senior Professional Level of the Bureau of Road Research and Development, Department of Highways, Mr. Pattarapon Sedokbuab for his unwavering support and insightful feedback which has been instrumental in refining the methodology, scope of work, literature review and manuscript preparation. His expertise added credibility to the findings of our manuscript.

7. REFERENCES

Bureau of Highway Safety. (2011). *Guide to Installing Safety Equipment and Facilities*. The Department of Highways.

Bureau of Highways Maintenance Management. (2006). *Road Maintenance Manual*. The Department of Highways.

Bureau of Road Research and Development. (2012). *Guide to Inspecting and Planning Maintenance of Pavement Surfaces to Ensure Appropriate Stiffness*. The Department of Highways.

Bureau of Road Research and Development. (2019). *Project to Study the Effects of Color Temperature and Brightness of Street lights on driving on Highways*. The Department of Highways.

Bureau of Road Research and Development. (2022). *Project to Study Innovative Technology for Measuring Light Intensity on Highways*. The Department of Highways.

Department of Alternative Energy Development and Efficiency. (2004). *Energy Efficient Design Guideline*. Ministry of Energy.

Federal Highway Administration. (2018). *FHWA Safety Program, Wed-Based Training for FHWA Roadway Lighting Workshop Module 3: Street and Roadway lighting Design*. U.S. Department of Transport.

Assessing BIM Implementation Success through Defined KPIs in Indonesia's Road and Bridge Infrastructure Projects

Heru Tri Saksena[1,*]

[1]Ministry of Public Works and Housing of Indonesia, saksena@pu.go.id

ABSTRACT

This research analyzes the implementation of Building Information Modeling (BIM) in nine road and bridge infrastructure projects in Indonesia during 2023. By defining specific Key Performance Indicators (KPIs) based on various parameters, the average BIM implementation KPI score for all projects was 78.44%. KPI-1: Completeness of BIM Documents and Support, achieved the highest compliance rate at 92.82%, while KPI-5: Cost and Time Efficiency through BIM Implementation, demonstrated the lowest compliance at 70.00%. Recommendations to enhance BIM effectiveness include concentrating efforts on improving KPI-3: Implementation of Preparation and Analysis of BIM-Based Design Drawings, further identifying opportunities for cost and time savings under KPI-5 and optimizing collaboration through the Directorate General of Highway's Common Data Environment as highlighted in KPI-6: Collaboration between Parties through The Directorate General of Highway's CDE. Key recommendations from this study include focusing on enhancing KPI-3 by refining BIM-based design processes, targeting cost and time efficiencies under KPI-5, and strengthening collaboration through optimized use of the Common Data Environment as outlined in KPI-6.

Keywords: Building Information Modeling (BIM), Digital Construction, Key Performance Indicators (KPIs)

1. INTRODUCTION

Building Information Modeling (BIM) has revolutionized the approach to design, construction, and facility management in infrastructure projects. Its significance lies in its ability to integrate various data and processes into a single digital model, which enhances collaboration, reduces errors, and improves decision-making throughout the project lifecycle. In infrastructure projects, particularly in road and bridge construction, BIM facilitates precise planning, real-time monitoring, and efficient resource management. This results in optimized project outcomes, including better cost control, time management, and sustainability. BIM's ability to simulate project scenarios, identify potential risks, and provide comprehensive visualizations enables stakeholders to make informed decisions that significantly reduce uncertainties and inefficiencies. BIM has increasingly been adopted in infrastructure projects, offering a comprehensive approach to design, construction, and facility management. Its integration into road and bridge projects is particularly crucial for several reasons. Firstly, BIM facilitates efficient asset and project management by enabling multiple applications simultaneously, including sustainability, lifecycle asset management, maintenance, condition monitoring, and real-time structural simulations (Kaewunruen et al., 2020). This multifaceted utility enhances overall infrastructure asset management. Despite the challenges of implementing BIM in infrastructure projects, it underscores the importance of considering the constructor perspective, which emphasizes a holistic view of BIM adoption (Bradley et al., 2016). Additionally, combining BIM with Geographic Information Science/Systems (GIS) is instrumental in managing asset information throughout the infrastructure's lifespan (Floros & Ellul, 2021). The interoperability between BIM and GIS provides a robust foundation for information exchange and management in these projects. The success of BIM in the building industry has led to its adoption in infrastructure design, including tunnels and large civil engineering projects, demonstrating its versatility and applicability (Barbieri et al., 2023). Specifically, in road and bridge infrastructure projects, BIM technology enhances planning accuracy, mitigates construction risks, and improves overall project lifecycle management (Rifai et al., 2022). Its application in transportation infrastructure projects shows great potential, from design to maintenance, indicating the broad spectrum of benefits BIM offers in this sector (Hao et al., 2021).

Before the adoption of BIM, infrastructure projects faced numerous challenges that often led to cost overruns, delays, and quality issues. Traditional project management methods relied heavily on fragmented and inconsistent information, resulting in poor communication and coordination among stakeholders. Design discrepancies and errors were common, as different teams often worked with outdated or incompatible versions of project documents. The lack of a unified data environment made it difficult to track project progress, manage changes, and ensure compliance with standards. Moreover, the manual handling of data increased the likelihood of human error, which further compounded the complexities of managing large-scale infrastructure projects. These challenges underscored the need for a more integrated and efficient approach to project management, which BIM effectively addresses.

BIM technology also assists in maintenance inspection, evaluation decision-making, and visual maintenance management of super-large bridges, highlighting its utility beyond the construction phase (Zhang & Dung, 2023). The integration of BIM into road and bridge projects provides various benefits, such as value control, decision support, reduction of design errors, identification of interface problems, enhancement of visual perception, and support for manufacturing, maintenance, and management of infrastructure assets (Vanjari & Kulkarni, 2021). The use of BIM models as a single point of truth during design and construction phases is recognized for improving efficiency, reducing rework, and minimizing extra costs in infrastructure projects (Nuttens et al., 2018).

In countries like Saudi Arabia and China, BIM implementation is assessed using Key Performance Indicators (KPIs) to evaluate effectiveness and maturity levels. For instance, a study in Saudi Arabia proposed a framework for BIM implementation based on the Project Management Institute's project management framework (Al-Yami & Sanni-Anibire, 2019). Similarly, in China, BIM practices were evaluated to understand their effectiveness across construction projects (Cao et al., 2015). Measuring BIM implementation with KPIs offers significant benefits. KPIs allow stakeholders to track progress, identify areas for improvement, and ensure effective utilization of BIM throughout the project lifecycle. They provide a quantitative and standardized way to evaluate BIM's impact on project outcomes, such as cost savings, improved collaboration, reduced errors, and enhanced visualization of project data (Sánchez et al., 2022). Moreover, KPIs help monitor BIM implementation performance, enabling informed decisions and adjustments to optimize project delivery (Lop et al., 2017). This approach enhances project efficiency, fosters stakeholder engagement, and improves overall project outcomes (Singh et al., 2021).

Indonesia's infrastructure sector, particularly in road and bridge projects, has been embracing digitalization and BIM implementation to enhance efficiency and effectiveness. Despite previous underinvestment, Indonesia recognizes the importance of digital transformation to address infrastructure deficits and spur economic growth (Ramces Hutapea & Natasaputra, 2021). The adoption of BIM in Indonesia's infrastructure projects is driven by various factors, including government requirements, belief in performance improvement, technological readiness, and top management support (Susanti et al., 2023). However, the lack of a policy and legal basis for BIM implementation across all infrastructure sectors is a noted limitation (Rifai et al., 2022). In road and bridge projects, the development of 4D BIM-based construction time planning standard operating procedures has been proposed to improve communication quality and serve as a guideline for BIM implementation (Jayadi & Riantini, 2019). Additionally, BIM integration in road infrastructure projects enhances asset management through effective information flow, accuracy, integrity, consistency, faster processing time, lower acquisition costs, and integrated decision-making (Guo et al., 2021). The digitalization of bridge inventory in Indonesia has also been explored through automated point cloud analysis for generating BIM models, addressing the challenges posed by complex bridge geometries (Hajdin, 2023). Moreover, integrating BIM with procedural modeling tools for road design is emphasized to improve the efficiency and accuracy of infrastructure projects (Biancardo et al., 2020).

2. LITERATURE REVIEW

2.1 Utilization of BIM and previous study

The successful adoption of Building Information Modeling (BIM) within organizations necessitates a focus on critical success factors and process management (Won et al., 2013). BIM functions as a digital innovation that not only encapsulates information about physical structures but also enhances collaboration and project management, thus becoming a crucial component of construction innovation (Koseoglu et al., 2019). The collaborative nature of BIM introduces advantages that improve project quality and delivery, underscoring its significance in enhancing project outcomes (Rahman, 2019). In the realm of infrastructure projects, BIM has been instrumental in refining road design, minimizing design errors, providing accurate and detailed information, enhancing the visualization of project data and environments, and fostering better coordination and collaboration among stakeholders (Sánchez et al., 2022). The application of BIM in infrastructure projects has been linked to benefits such as reducing time, costs, and waste, improving project performance, enhancing project simulation and visualization, increasing communication and collaboration, identifying conflicts, and reducing duplications (Samimpay & Saghatforoush, 2020).

BIM has been evolving towards the concept of a Digital Twin, which aims to optimize the performance of infrastructure systems by digitally enhancing asset value and leveraging asset data (Ammar et al., 2024). The integration of BIM with digital platforms and open standards facilitates interoperability and digital transformation in the construction sector, enabling stakeholders to make informed decisions and enhance project efficiency (Ciccone et al., 2022; Nübel et al., 2021). Globally, research has indicated that the implementation of BIM at the design stage may initially increase expenses, but the long-term benefits outweigh these costs, particularly during the construction phase (Lee et al., 2021). Critical success factors (CSFs) of BIM implementation have been identified, emphasizing the importance of factors such as knowledge management and demolition in successful BIM adoption in the construction industry (Yaakob et al., 2016).

Studies have also explored the challenges and opportunities related to BIM adoption in the construction industry, highlighting the need to address barriers to successful implementation (Dwi Hatmoko et al., 2019). In the context of road and bridge projects, the benefits of BIM implementation are evident in mitigating cost overrun factors, improving road design, reducing errors, enhancing visualization, and promoting collaboration among stakeholders (Sánchez et al., 2022). Challenges related to level of detail specifications, standards, and file-format issues in infrastructure projects for BIM Level Three have been addressed through case studies, emphasizing the importance of recognizing existing processes and challenges in implementing BIM in infrastructure projects (Hijazi & Omar, 2017).

In Indonesia, investigations into BIM adoption in the construction industry have been conducted to understand the challenges and opportunities associated with BIM implementation (Raza et al., 2017). Research has highlighted the benefits of BIM in green construction projects, emphasizing waste reduction, time optimization, and resource efficiency as key aims of BIM implementation in sustainable construction practices (Latupeirissa & Arrang, 2023). Additionally, studies have focused on BIM implementation readiness in building projects in Indonesia, defining readiness as the psychological willingness and preparedness for BIM implementation activities (Liao et al., 2019; Saksena & Sastrawiria, 2023).

In sum, previous studies on BIM implementation globally and in Indonesia have provided valuable insights into the challenges, benefits, critical success factors, and readiness for BIM adoption in road and bridge projects. These studies contribute to the understanding of BIM's impact on infrastructure development and underscore the importance of addressing barriers to successful BIM implementation for improved project outcomes. The consistent identification of challenges and critical success factors resonates across different contexts, underscoring the global significance of addressing these factors for optimized BIM implementation outcomes.

2.2 Key Performance Indicators (KPIs) for BIM success assessment

Evaluating the success of Building Information Modeling (BIM) implementation in construction projects relies heavily on Key Performance Indicators (KPIs). Numerous studies have pinpointed specific KPIs crucial for assessing the effectiveness and impact of BIM adoption. These metrics gauge various facets of BIM implementation and its influence on project outcomes. For instance, in public-private partnership projects, selecting performance objectives and KPIs is crucial for achieving value for money (Yuan et al., 2009). They stress the importance of aligning KPIs with project objectives to ensure successful partnership outcomes, emphasizing aspects like cost-effectiveness and efficiency. In Malaysian Private Finance Initiative projects, KPIs play a pivotal role in measuring performance, particularly in areas such as cost control, schedule adherence, quality management, and stakeholder satisfaction(Lop et al., 2017). Similarly, in managing scope creep in road projects, KPIs related to cost efficiency, scope adherence, schedule management, and quality control as essential for effective monitoring and management are highlighted (Jayalath & Somarathna, 2021). (Darwish et al., 2020) delve into critical success factors for BIM implementation, emphasizing KPIs associated with organizational readiness, stakeholder engagement, training effectiveness, and technology integration. Additionally, (Ershadi et al., 2022) underscore the importance of KPIs related to process management, implementation effectiveness, and BIM utilization in infrastructure construction projects. Furthermore, the widespread recognition and utilization of these diverse sets of KPIs across various construction projects underscore their universal relevance and effectiveness in evaluating BIM implementation. The consistency in their application highlights the consensus within the industry regarding the critical role KPIs play in gauging BIM's impact on project performance and overall success.

2.3 BIM regulation on Indonesia's road and bridge infrastructure projects

In Indonesia, the implementation of BIM is governed by various regulations, encompassing Government Regulation 2021, which pertains to the Implementing Regulations of Law 2002 focusing on Buildings. Additionally, Ministerial Instruction Number 2022 outlines strategies to prevent deviations in the Ministry's Goods/Services Procurement Process from 2022 to 2024, with one of the strategies is implemented BIM from early stages of construction. Directorate General of Highway provided Road and Bridge Infrastructure Guidelines 12th 2023 as comprehensive guidelines for the application of BIM within the road and bridge infrastructure. This guideline establishes the organization that implements BIM, the resource requirements, budgeting, minimum requirements, and technical provisions regarding modeling, simulation, data codification, and data management, as well as the digital asset handover process for Building Information Modeling (BIM) during the pre-planning, technical planning, land acquisition, construction, and building utilization phases.

The Directorate General of Highway plays a pivotal role in overseeing and ensuring the successful implementation of Building Information Modeling (BIM) across various infrastructure projects. To maintain a high standard of BIM execution, the Directorate conducts regular monitoring sessions online, held biweekly. These sessions are strategically designed to last between thirty minutes to an hour, providing a focused and efficient platform for project teams to discuss ongoing challenges and develop actionable solutions. During these online monitoring activities, project stakeholders, including engineers, BIM coordinators, and project managers, come together to review the progress of BIM implementation. The structured format of these sessions ensures that each project is thoroughly examined, with particular attention given to any obstacles that may hinder the optimal use of BIM. Common issues discussed include technical difficulties, data management challenges, and integration problems between different BIM tools and platforms. By identifying these issues early, the Directorate General of Highway can guide project teams in developing effective strategies to address them, thereby preventing minor problems from escalating into major setbacks.

3. METHODOLOGY

3.1 Key Performance Indicators (KPIs)

The integration of BIM into infrastructure projects is a dynamic process necessitating thorough monitoring and evaluation. Key Performance Indicators (KPIs) play a pivotal role, encompassing metrics such as BIM document completeness, design process quality, document preparation, training, cost and time efficiency, and collaboration facilitated by a Common Data Environment (CDE). The selection of KPIs is intricately linked to the principles delineated in ISO 19650, an internationally recognized standard governing information management across the lifecycle of constructed assets using BIM. By aligning KPIs with ISO 19650, organizations can methodically evaluate and enhance their BIM adoption efforts, ultimately improving project outcomes and adhering to international best practices in BIM management and delivery.

Table 1 Directorate General of Highway's KPIs and ISO 19650 criteria

KPIs	Indicators	ISO 19650 Clause(s) (BSI, 2018)	Description
1	Completeness of BIM Documents and Support	ISO 19650-2&3 - 5.1.1 (appointing party information management function)	This KPI ensures that BIM documents, including the BIM Execution Plan (BEP), folder formats, file naming conventions, and extensions, are complete and suitable for effective information management, aligning with the requirement to appoint an information management function according to ISO 19650.
1a	Completeness of BIM Execution Plan (BEP)		
1b	Completeness and Suitability of Folder Format, File/Activity Naming, and Extension		
1c	Completeness of Members in the Common Data Environment (CDE)		
2	Implementation of BIM Monitoring and Evaluation	ISO 19650-2 – 5.1.9 (activities for assessment and need) ISO 19650-3 – 5.1.5 (identify foreseeable trigger events for which information	This KPI focuses on the completeness of monitoring and evaluation sheets, the conduct of BIM coordination meetings, and the resolution of follow-up actions on issues, which are crucial for managing information effectively in response to foreseeable trigger events as per ISO 19650.
2a	Completeness of Monitoring and Evaluation Sheets (Monev)		
2b	Conduct of BIM Coordination Meetings		
2c	Resolution of Follow-Up Actions on Issues		

KPIs	Indicators	ISO 19650 Clause(s) (BSI, 2018)	Description
		shall be managed)	
3	Implementation of Preparation and Analysis of BIM-Based Design Drawings	ISO 19650-2 – 5.7 (information management process; information model delivery) ISO 19650-3 – 5.1.6 (asset information standard)	This KPI involves the preparation and analysis of BIM-based design drawings, including 3D BIM model-based drawings, scheduling simulations, conflict analyses, and financial quantifications based on BIM models, meeting the requirement to establish asset information standards outlined in ISO 19650.
3a	3D BIM Model-Based Design Drawings		
3b	Progress Scheduling Simulation Based on 4D BIM Model		
3c	Conflict Analysis Simulation Between Disciplines Based on 4D BIM Model		
3d	Quantification and Analysis of Financing Based on 5D BIM Model		
4	Implementation of Design Review and Approval through The Directorate General of Highway's CDE	ISO 19650-2 – 5.8.1 (archive the project information model) ISO 19650-3 – 5.1.9 (common data environment)	This KPI ensures the availability of workflows for design and document approvals, as well as issue management within The Directorate General of Highway's Common Data Environment (CDE), aligning with the establishment of a CDE as a platform for information exchange and collaboration as mandated by ISO 19650.
4a	Availability of Design/Working Drawing Approval Workflow		
4b	Availability of Other Project Document Approval Workflow		
4c	Issue Management on the CDE		
5	Cost and Time Efficiency through BIM Implementation	ISO 19650-3 – 5.2.4 (response requirements and evaluation criteria)	This KPI focuses on generating reports on cost and time efficiency achieved through BIM implementation, meeting the requirements to compile response requirements and evaluation criteria for effective project management and decision-making as per ISO 19650.
5a	Report on Cost and Time Efficiency through BIM Implementation		
6	Collaboration between Parties through The Directorate General of Highway's CDE	ISO 19650-2 – 5.8.1 (archive the project information model) ISO 19650-3 – 5.1.9 (common data environment)	This KPI emphasizes the role of The Directorate General of Highway's CDE as a platform for progress reporting, design approvals, document approvals, communication, and issue resolution among project parties, aligning with the establishment of a CDE for collaboration and information management as outlined in ISO 19650.
6a	CDE as a Platform for Progress Reporting and Weekly BIM Implementation Reports		
6b	CDE as a Platform for Design/Working Drawing Approval		
6c	CDE as a Platform for Approval of Other Project Documents		
6d	CDE as a Platform for Communication and Issue Resolution in Projects		

3.2 Projects sample and data collection method

The selection of the nine projects for this study was based on a combination of criteria, including geographic diversity, project type, and the level of BIM integration. These projects were chosen to provide a representative sample of Indonesia's road and bridge infrastructure sector. The selected projects ranged from large-scale toll road developments to smaller bridge rehabilitation efforts, reflecting the varied challenges and opportunities in BIM implementation across different contexts. The suitability of these projects was determined by their strategic importance, as well as the availability of detailed BIM implementation plans, which allowed for a thorough analysis of BIM's impact on project outcomes. These projects include the Flyover Sekip Ujung in South Sumatera, a bridge project aimed at easing traffic congestion; the Jalan Soe - Kefamenanu - Oelfaub in Nusa Tenggara Timur, a road rehabilitation project enhancing safety and accessibility; and three sections of the Jalan Tol Semarang Demak toll road project (Seksi 1A, 1B, and 1C) in Central Java, improving connectivity between Semarang and Demak. Additionally, the Jalan Tol Serang - Panimbang Seksi 3 in Banten focuses on regional connectivity and economic support, while the Jembatan Sambas in West Borneo constructs a vital bridge for better regional links. The KPBU Riau project in Riau involves road rehabilitation for improved infrastructure, and the Sofi Wayabula II project in North Maluku develops new road infrastructure for enhanced regional connectivity. BIM was employed in these projects to optimize design accuracy, manage construction processes, and ensure efficient project management.

Table 2 Project details.

No	Project Name	Location	Type
1	Flyover Sekip Ujung	South Sumatera	Bridge Construction
2	Jalan Soe - Kefamenanu - Oelfaub	East Nusa Tenggara	Road Rehabilitation
3	Jalan Tol Semarang Demak Seksi 1A	Central Java	Toll Road Development
4	Jalan Tol Semarang Demak Seksi 1B	Central Java	Toll Road Development
5	Jalan Tol Semarang Demak Seksi 1C	Central Java	Toll Road Development
6	Jalan Tol Serang - Panimbang Seksi 3	Banten	Toll Road Development
7	Jembatan Sambas	West Borneo	Bridge Construction
8	KPBU Riau	Riau	Road Rehabilitation
9	Sofi Wayabula II	North Maluku	Road Construction

Data collection in this study followed a structured, multi-phase approach designed to ensure both accuracy and relevance. For each project, data was directly collected from the project teams through biweekly submissions. These submissions comprised detailed reports on BIM implementation progress, supplemented by supporting evidence such as design documents, meeting minutes, and simulation results. The VDC/BIM Coordinator from the Directorate General of Highways verified the accuracy of this data by cross-referencing it with the predefined Key Performance Indicators (KPIs). Any discrepancies or missing information were flagged, requiring the project teams to rectify these issues in subsequent submissions, thus ensuring that the data accurately represented the project's status.

The biweekly online meetings played a crucial role in monitoring BIM implementation. Conducted via a secure video conferencing platform, these meetings typically lasted between 30 minutes to an hour. Each session was led by the VDC/BIM Coordinator, who facilitated discussions on the progress of BIM implementation, focusing particularly on any issues or challenges highlighted in the biweekly reports. The meetings followed a structured agenda, which included reviewing KPI scores, addressing discrepancies or challenges, and identifying action items for the coming week. Follow-up was managed through the Common Data Environment (CDE), where all meeting notes, action items, and deadlines were documented. Project teams were expected to update the status of each action item in the CDE before the next meeting, ensuring continuous progress and accountability.

KPIs submitted by the project teams were validated based on the work evidence provided for each project. This validation process ensured that the KPIs accurately reflected the project's progress and adherence to BIM standards. The final KPI score for each month was calculated as the average of the KPIs from the first and second biweekly meetings of that month, providing a balanced and comprehensive assessment of the project's BIM implementation performance. Feedback and improvement suggestions were communicated to each project through official memos from the Director of Road and Bridge Engineering, who oversees BIM implementation within the Directorate General of Highways. This structured approach facilitated thorough and continuous monitoring of BIM implementation across all projects, enabling early identification of issues and promoting effective project management and improvement strategies.

The calculation of KPI scores involved several key steps to ensure accuracy and a true reflection of project performance. Each KPI was scored monthly, with the scores derived from the data submitted by the project teams. This comprehensive process ensured the reliability and validity of the KPI scores, providing an accurate assessment of BIM implementation across the projects. The process included:

- Data Aggregation: Data related to each KPI was collected and aggregated from the biweekly reports submitted by the project teams.
- Normalization: The collected data was then normalized to ensure consistency across different projects and metrics.
- Weighting: Each KPI was assigned a weight based on its importance to the overall success of BIM implementation.
- Formula Application: The final KPI score was calculated using a weighted average formula, which combined the normalized data for each KPI with its assigned weight. This approach ensured that more critical KPIs had a greater impact on the overall score.
- Validation: The calculated scores were validated by comparing them with the previous month's scores to identify any anomalies or significant changes, which were then investigated further.

4. FINDINGS

The table below offers an in-depth analysis of BIM implementation success across nine different infrastructure projects in Indonesia, spanning a 12-month period throughout 2023. The analysis includes detailed monthly performance scores for each project, highlighting how they have progressed over time. Additionally, the table presents the average performance for the year, the highest and lowest scores achieved by each project, and the standard deviation, which measures the consistency of performance throughout the period.

This data provides a clear overview of how each project has performed in relation to key performance indicators (KPIs) that are critical to BIM implementation success. Importantly, these KPI achievements have been meticulously validated by the VDC/BIM Coordinator at the head office. This validation process is based on the monthly reports submitted by the project teams, ensuring that the data is accurate and reflects the true performance of each project. The overall aim of this analysis is to identify trends, highlight areas of excellence, and pinpoint opportunities for improvement in BIM practices across these projects.

Table 3 BIM Implementation through nine projects sample in 2023.

No	Projects/ Months	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Avg
1	Flyover Sekip Ujung	67.50	67.50	62.50	63.75	58.75	52.50	67.50	77.50	71.25	75.00	96.25	88.75	70.73
2	Jalan Soe - Kefamenanu - Oelfaub	35.00	35.00	60.00	60.00	67.50	85.00	78.75	85.00	77.50	77.50	77.50	75.00	67.81
3	Jalan Tol Semarang Demak Seksi 1A	82.50	82.50	87.50	90.00	97.50	97.50	91.25	100	100	97.50	100	97.50	93.65
4	Jalan Tol Semarang Demak Seksi 1B	42.50	42.50	47.50	85.00	100	97.50	97.50	100	100	97.50	96.25	97.50	83.65
5	Jalan Tol Semarang Demak Seksi 1C	42.50	42.50	60.00	65.00	85.00	100	88.75	97.50	100	100	97.50	95.00	81.15
6	Jalan Tol Serang - Panimbang Seksi 3	42.50	42.50	70.00	67.50	72.50	100	91.25	96.25	100	100	100	100	81.88
7	Jembatan Sambas	82.50	82.50	90.00	75.00	92.50	90.00	88.75	100	100	97.50	82.50	85.00	88.85
8	KPBU Riau	95.00	95.00	97.50	80.00	100	100	100	93.75	100	100	95.00	93.75	95.83
9	Sofi Wayabula II	12.50	12.50	8.75	15.00	50.00	45.00	53.75	55.00	48.75	67.50	72.50	67.50	42.40
Average		55.83	55.83	64.86	66.81	80.42	85.28	84.17	89.44	88.61	90.28	90.83	88.89	78.44

Table 4 KPIs details achievement score through nine projects sample in 2023.

No	Projects/ Months	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Avg
1	Completeness of BIM Documents and Support	55.93	74.07	87.96	93.52	98.15	100.00	89.81	100.00	97.22	99.07	100.00	100.00	91.31
2	Implementation of BIM Monitoring and Evaluation	68.52	68.52	75.00	73.15	87.96	87.04	87.96	96.30	94.44	95.37	96.30	93.52	85.34
3	Implementation of Preparation and Analysis of BIM-Based Design Drawings	40.00	40.00	54.44	64.44	75.56	84.44	70.00	81.11	83.33	83.33	83.33	80.00	70.00
4	Implementation of Design Review and Approval through The Directorate General of Highway's CDE	66.67	66.67	71.30	78.70	90.74	92.59	96.30	90.74	87.96	90.74	99.07	100.00	85.96
5	Cost and Time Efficiency through BIM Implementation	38.89	38.89	38.89	44.44	44.44	77.78	77.78	88.89	88.89	100.00	100.00	100.00	69.91

In Table 3, a meticulous breakdown of monthly performance scores for each project is provided, enabling a nuanced understanding of their trajectory and progress throughout the year. The average performance across all projects for the entirety of 2023 stands at 78.44%. This figure encapsulates the collective effort and efficacy of BIM implementation strategies across the sampled projects. However, the data reveals intriguing variations among individual projects. For instance, KPBU Riau emerges as a consistent front-runner, maintaining an impressive average score of 95.83%. This suggests a robust and well-executed BIM implementation framework within the project. Conversely, Sofi Wayabula II appears to face substantial challenges, registering a notably lower average score of 42.40%. This discrepancy underscores the diverse array of hurdles and opportunities encountered in different project contexts.

Table 4 furnishes a granular breakdown of KPI achievements across various dimensions of BIM implementation. Here, the achievements are delineated against specific KPIs, shedding light on the project's strengths and areas necessitating enhancement. For example, projects exhibited commendable performance in the completeness of BIM documents and support, with an average score of 91.31%. This underscores a concerted effort in ensuring comprehensive documentation and robust support structures, vital for the seamless execution of BIM methodologies. Conversely, the efficiency of cost and time through BIM implementation emerged as a notable challenge, with an average score of 69.91%. This indicates potential areas for streamlining processes and optimizing resource allocation to bolster overall project efficacy. The presented data not only offers a snapshot of the prevailing status of BIM implementation but also serves as a roadmap for future endeavors. By dissecting these performance metrics, stakeholders can discern overarching trends, identify areas of excellence, and pinpoint specific challenges requiring targeted interventions. This holistic approach fosters a culture of continuous improvement, propelling the trajectory of BIM implementation towards heightened efficiency and efficacy in future infrastructure projects.

4.1 KPI 1: Completeness of BIM Documents and Support

The average monthly KPI scores showed significant improvement from 55.93% in January to a consistent 100% from June onwards. KPBU Riau led with a perfect average score of 100% throughout the year, demonstrating exemplary adherence to BIM documentation standards. Other high performers included Jembatan Sambas and Jalan Tol Semarang Demak Seksi 1A, both achieving average scores above 90%. While most projects showed marked improvements and maintained high scores, Jalan Soe - Kefamenanu - Oelfaub and Jalan Tol Serang - Panimbang Seksi 3 started with lower scores but made substantial progress, ending the year with near-perfect scores. The data underscores the effectiveness of structured BIM implementation and monitoring in enhancing project documentation completeness over time.

4.2 KPI 2: Implementation of BIM Monitoring and Evaluation

Overall, the average performance across all projects was 85.34%, indicating a commendable level of success in implementing BIM monitoring and evaluation throughout the year. Projects like Flyover Sekip Ujung demonstrated a fluctuating trend, starting low in January but steadily improving and reaching peak performance by November and December. Conversely, Jalan Soe - Kefamenanu - Oelfaub exhibited consistent high performance throughout the year. Meanwhile, projects under Jalan Tol Semarang Demak Seksi 1A consistently maintained a perfect score. Challenges were evident in projects like Sofi Wayabula II, which showed lower initial scores but gradually improved over time.

4.3 KPI 3: Implementation of Preparation and Analysis of BIM-Based Design Drawings

The average KPI score across all projects was 70.00%, highlighting the overall performance trend and emphasizing the importance of continuous improvement and sharing best practices to enhance BIM implementation effectiveness. Projects like Flyover Sekip

Ujung and Jalan Tol Semarang Demak Seksi 1A consistently scored high, with average scores of 91.67% and 89.17% respectively, showcasing proficiency in utilizing 3D and 4D BIM models for design drawing preparation and analysis. Conversely, projects such as Jalan Soe - Kefamenanu - Oelfaub and Sofi Wayabula II struggled to maintain consistent performance, with average scores of 25.83% and 18.33% respectively, indicating potential challenges in leveraging BIM effectively for design drawing tasks.

4.4 KPI 4: Implementation of Design Review and Approval through The Directorate General of Highway's CDE

In the assessment of the Implementation of Design Review and Approval through The Directorate General of Highway's Common Data Environment (CDE), the projects exhibited varying degrees of performance throughout the year. While projects like Jalan Tol Semarang Demak Seksi 1A, Jalan Tol Serang - Panimbang Seksi 3, and Jembatan Sambas consistently maintained high scores, others, such as Flyover Sekip Ujung and Sofi Wayabula II, showed fluctuating performance trends. Overall, the average performance across all projects improved steadily from 66.67% in January to a peak of 100% in December, with an average of 85.96% over the year. This data underscores the importance of consistent monitoring and improvement efforts to ensure effective design review and approval processes in infrastructure projects.

4.5 KPI 5: Cost and Time Efficiency through BIM Implementation

The data reveals varying levels of success in achieving cost and time efficiency through BIM implementation across different projects. Projects like KPBU Riau consistently maintained maximum scores throughout the year, showcasing exemplary efficiency. Similarly, Jalan Tol Semarang Demak Seksi 1A and 1B demonstrated strong performance with consistent high scores. Conversely, projects like Flyover Sekip Ujung and Sofi Wayabula II struggled to achieve significant efficiency gains, with sporadic scores fluctuating throughout the year. Overall, while some projects maintained consistent efficiency levels, others faced challenges in maintaining momentum, resulting in an average efficiency rating of 69.91%. This highlights the need for continuous monitoring and improvement strategies to ensure sustained efficiency gains across all projects.

4.6 KPI 6: Collaboration through Common Data Environment

The data reflects the collaboration between parties through The Directorate General of Highway's Common Data Environment (CDE) across various projects over the course of a year. While some projects showed consistent or improving collaboration scores, others demonstrated fluctuations. For instance, Jalan Tol Semarang Demak Seksi 1A consistently maintained high collaboration scores throughout the year, indicating strong and sustained teamwork. Conversely, projects like Jembatan Sambas experienced fluctuating scores, suggesting challenges in maintaining consistent collaboration. Overall, the average collaboration score across all projects was 71.64%, indicating a moderate level of collaborative effectiveness through the CDE. These insights provide valuable feedback for optimizing collaboration strategies and enhancing project outcomes in future endeavors.

5. CONCLUSION

The comprehensive analysis of BIM implementation across nine infrastructure projects in Indonesia during 2023 reveals both successes and areas for improvement. The overall average score across all projects was 78.44%, indicating a generally positive trend in BIM implementation. KPI-specific insights reveal that KPI-1: Completeness of BIM Documents and Support saw high scores, averaging 91.31%, indicating strong documentation practices across projects. KPI-2: Implementation of BIM Monitoring and Evaluation averaged 85.34%, reflecting the importance of continuous monitoring and regular BIM coordination meetings. KPI-5: Cost and Time Efficiency through BIM Implementation varied widely, averaging 69.91%, highlighting the need for better process optimization and resource management. KPI-6: Collaboration between Parties through CDE had an average of 71.64%, with significant improvements in months where active collaboration platforms were used.

The successful implementation of Building Information Modeling (BIM) in infrastructure projects, as demonstrated by the majority of the nine projects studied, underscores the transformative potential of BIM in achieving significant cost and time savings. Projects such as KPBU Riau and Jalan Tol Semarang Demak Seksi 1A exemplified the efficiency and effectiveness that BIM can bring when fully integrated into project workflows. These projects achieved remarkable performance across multiple KPIs, particularly in terms of cost and time efficiency (KPI-5), collaboration (KPI-6), and documentation completeness (KPI-1). The ability of BIM to streamline processes, enhance communication, and provide real-time data has been pivotal in these successes, leading to improved project outcomes and reduced risks. However, the study also highlighted the challenges and obstacles that can impede BIM implementation. The Sofi Wayabula II project, with its low average score, revealed significant struggles in areas such as technical proficiency, data management, and stakeholder engagement. These challenges emphasize the need for comprehensive training, robust data management systems, and active stakeholder involvement to fully harness the benefits of BIM.

To address the technical challenges that some projects faced, such as those seen in the Sofi Wayabula II project, it is crucial to enhance training and capacity building within project teams. This can be achieved by investing in ongoing training programs that not only provide initial instruction on BIM tools but also offer continuous education and support. As BIM technologies and standards evolve, keeping the project team updated will ensure they can effectively utilize these tools to their full potential, thereby improving overall project performance. Another key area for improvement is strengthening data management practices. Effective data management is fundamental to the success of BIM implementation. Projects should adopt standardized data management protocols to ensure that information remains consistent, accurate, and accessible throughout all phases of the project. The Flyover Sekip Ujung project highlighted the issues that can arise from poor data management, such as data loss, duplication, and inconsistencies. By implementing robust data management systems, projects can mitigate these risks and maintain a high level of operational efficiency.

Promoting stakeholder engagement is also essential for the success of BIM projects. Active and consistent engagement from all stakeholders ensures alignment and commitment to the project's goals. This can be facilitated by regular meetings, clear communication channels, and the effective use of the Common Data Environment (CDE). Encouraging collaboration and transparency will help overcome resistance to change, ensuring that all stakeholders are fully on board with BIM processes, which is vital for the smooth execution of the project. Additionally, implementing change management strategies is necessary to address resistance to adopting new technologies and processes. Effective change management involves clear communication of the benefits of BIM, involving key stakeholders in decision-making processes, and providing support systems to help teams transition smoothly to new ways of working. By fostering a culture that embraces innovation, projects can mitigate the barriers that often hinder BIM implementation.

Lastly, a strong focus on continuous monitoring and evaluation is recommended. Regularly assessing BIM implementation through predefined Key Performance Indicators (KPIs) should become standard practice. This approach allows for the early identification of issues and enables the timely implementation of corrective measures. The success of projects like Jembatan Sambas illustrates the importance of maintaining consistent monitoring and evaluation efforts to sustain project momentum and achieve desired outcomes. By adhering to these practices, future infrastructure projects can fully realize the benefits of BIM and achieve higher levels of efficiency, quality, and sustainability.

6. REFERENCES

Al-Yami, A., & Sanni-Anibire, M. O. (2019). BIM in the Saudi Arabian Construction Industry: State of the Art, Benefit and Barriers. *International Journal of Building Pathology and Adaptation*. https://doi.org/10.1108/ijbpa-08-2018-0065

Ammar, A., Nassereddine, H., & Dadi, G. B. (2024). *Perspective Chapter: Roadmap to a Holistic Highway Digital Twin – A Why, How, and Why Framework*. https://doi.org/10.5772/intechopen.108546

Barbieri, G., Giani, M., Panicis, E. De, Biagi, A., Femina, D. Della, & Bella, G. Di. (2023). *BIM and Tunnelling – A Norwegian Application: The Sotra Link Project*. https://doi.org/10.11159/icgre23.123

Biancardo, S. A., Capano, A., de Oliveira, S. G., & Tibaut, A. (2020). Integration of BIM and Procedural Modeling Tools for Road Design. *Infrastructures*. https://doi.org/10.3390/infrastructures5040037

Bradley, A., Li, H., Lark, R. J., & Dunn, S. (2016). BIM for Infrastructure: An Overall Review and Constructor Perspective. *Automation in Construction*. https://doi.org/10.1016/j.autcon.2016.08.019

BSI. (2018). BS EN ISO 19650 Organization of information about construction works — Information management using building information modelling. *The International Organization for Standardization (ISO), 2018*.

Cao, D., Wang, G., Li, H., Skitmore, M., Huang, T.-C., & Zhang, W. (2015). Practices and Effectiveness of Building Information Modelling in Construction Projects in China. *Automation in Construction*. https://doi.org/10.1016/j.autcon.2014.10.014

Ciccone, A., Stasio, S. Di, Asprone, D., Salzano, A., & Nicolella, M. (2022). Application of openBIM for the Management of Existing Railway Infrastructure: Case Study of the Cancello–Benevento Railway Line. *Sustainability*. https://doi.org/10.3390/su14042283

Darwish, A. M., Tantawy, M. M., & Elbeltagi, E. (2020). Critical Success Factors for BIM

Implementation in Construction Projects. *Saudi Journal of Civil Engineering*. https://doi.org/10.36348/sjce.2020.v04i09.006
Dwi Hatmoko, J. U., Fundra, Y., Wibowo, M. A., & Zhabrinna. (2019). Investigating Building Information Modelling (BIM) Adoption in Indonesia Construction Industry. *Matec Web of Conferences*. https://doi.org/10.1051/matecconf/201925802006
Ershadi, M., Jefferies, M., Davis, P. G., & Mojtahedi, M. (2022). Implementation of Building Information Modelling in Infrastructure Construction Projects: A Study of Dimensions and Strategies. *International Journal of Information Systems and Project Management*. https://doi.org/10.12821/ijispm090403
Floros, G. S., & Ellul, C. (2021). LOSS OF INFORMATION DURING DESIGN &Amp; CONSTRUCTION FOR HIGHWAYS ASSET MANAGEMENT: A GEOBIM PERSPECTIVE. *Isprs Annals of the Photogrammetry Remote Sensing and Spatial Information Sciences*. https://doi.org/10.5194/isprs-annals-viii-4-w2-2021-83-2021
Guo, X., Tian, C., Chen, Y., & Zhang, J. (2021). Case Study of Building Information Modeling Implementation in Infrastructure Projects. *Transportation Research Record Journal of the Transportation Research Board*. https://doi.org/10.1177/03611981211045060
Hajdin, R. (2023). Digitalization of Bridge Inventory via Automated Analysis of Point Clouds for Generation of BIM Models. *Ce/Papers*. https://doi.org/10.1002/cepa.1995
Hao, Z., Zhang, W., & Zhao, Y. (2021). Integrated BIM and VR to Implement IPD Mode in Transportation Infrastructure Projects: System Design and Case Application. *Plos One*. https://doi.org/10.1371/journal.pone.0259046
Hijazi, A., & Omar, H. A. (2017). *Level of Detail Specifications, Standards and File-Format Challenges in Infrastructure Projects for Bim Level Three*. https://doi.org/10.2495/bim170141
Jayadi, A., & Riantini, L. S. (2019). The Development of 4D BIM Based Infrastructure Construction Time Planning Standard Operating Procedure to Enhance Communication Quality From Infrastructure Contractor Perspective. *Journal of International Conference Proceedings*. https://doi.org/10.32535/jicp.v2i1.393
Jayalath, C., & Somarathna, K. K. G. P. (2021). *Key Performance Indicators in Upholding Scope Creep Management in Road Projects*. https://doi.org/10.31705/wcs.2021.33
Kaewunruen, S., Sresakoolchai, J., & Zhou, Z. (2020). Sustainability-Based Lifecycle Management for Bridge Infrastructure Using 6D BIM. *Sustainability*. https://doi.org/10.3390/su12062436
Koseoglu, O., Keskin, B., & Ozorhon, B. (2019). Challenges and Enablers in BIM-Enabled Digital Transformation in Mega Projects: The Istanbul New Airport Project Case Study. *Buildings*. https://doi.org/10.3390/buildings9050115
Latupeirissa, J. E., & Arrang, H. (2023). *Sustainability Factors of Building Information Modeling (BIM) Implementing for Successful Construction Project Managemet Life Cycle in Indonesia*. https://doi.org/10.21203/rs.3.rs-2316649/v2
Lee, S. Y., Chung, S.-W., Kwon, S.-W., Cho, C.-S., & Lee, K. (2021). Assessment of BIM Competencies and Correlation Analysis Between Competencies and Career Characteristics of FAB Construction Project Participants. *Applied Sciences*. https://doi.org/10.3390/app11188468
Liao, L., Ai Lin, E. T., & Low, S. P. (2019). Assessing Building Information Modeling Implementation Readiness in Building Projects in Singapore. *Engineering Construction & Architectural Management*. https://doi.org/10.1108/ecam-01-2019-0028
Lop, N. S., Ismail, K., & Isa, H. M. (2017). The Implementation of Key Performance Indicators in the Malaysian Private Finance Initiative Projects. *Environment-Behaviour Proceedings Journal*. https://doi.org/10.21834/e-bpj.v2i5.686
Nübel, K., Bühler, M. M., & Jelinek, T. (2021). Federated Digital Platforms: Value Chain Integration for Sustainable Infrastructure Planning and Delivery. *Sustainability*. https://doi.org/10.3390/su13168996
Nuttens, T., Breuck, V. De, Cattoor, R., Decock, K., & Hemeryck, I. (2018). Using BIM Models for the Design of Large Rail Infrastructure Projects: Key Factors for a Successful Implementation. *International Journal of Sustainable Development and Planning*. https://doi.org/10.2495/sdp-v13-n1-73-83
Rahman, I. A. (2019). Awareness and Challenges of Building Information Modelling (BIM)

Implementation in the Yemen Construction Industry. *Journal of Engineering Design and Technology*. https://doi.org/10.1108/jedt-03-2019-0063
Ramces Hutapea, B. S., & Natasaputra, S. (2021). Affecting Factors That Improve Performance Project With Line of Balance and Critical Path Method at Pier Works in Toll Road Project at Jakarta Indonesia. *International Journal of Engineering Research and Advanced Technology*. https://doi.org/10.31695/ijerat.2021.3739
Raza, H., Tanoli, W. A., Lee, S., & Seo, J. (2017). *Flexible Earthwork BIM Module Framework for Road Project*. https://doi.org/10.22260/isarc2017/0056
Rifai, A. I., Thalib, H., Isradi, M., & Prasetijo, J. (2022). Implementation of Building Information Modelling for Road Rehabilitation and Reconstruction Project: Liquefaction Disaster of Palu Area. *Ijebd (International Journal of Entrepreneurship and Business Development)*. https://doi.org/10.29138/ijebd.v5i4.1914
Saksena, H. T., & Sastrawiria, R. P. P. (2023). Implementation of Laser Scanning Technology for Digital Asset Management of Bridges in Indonesia: A Case Study of Pulau Balang Bridge in East Kalimantan. *Asia Australia Road Conference*. https://doi.org/https://doi.org/10.58674/phpji.v16i1.384
Samimpay, R., & Saghatforoush, E. (2020). Benefits of Implementing Building Information Modeling (BIM) in Infrastructure Projects. *Journal of Engineering Project and Production Management*. https://doi.org/10.2478/jeppm-2020-0015
Sánchez, O., Castañeda, K., Herrera, R., & Pellicer, E. (2022). *Benefits of Building Information Modeling in Road Projects for Cost Overrun Factors Mitigation*. https://doi.org/10.1061/9780784483978.049
Singh, S., Chinyio, E., & Suresh, S. (2021). The Key Enablers, Techniques and Benefits of Managing Stakeholders Within BIM Supported Projects. *Journal of Engineering Design and Technology*. https://doi.org/10.1108/jedt-04-2021-0203
Susanti, R., Fauziyah, S., & Aziz, S. (2023). Factor and Barrier Implementing Building Information Modelling (BIM) in Green Construction in Indonesia. *Iop Conference Series Earth and Environmental Science*. https://doi.org/10.1088/1755-1315/1203/1/012018
Vanjari, P. B., & Kulkarni, S. S. (2021). Building Information Modeling (BIM) as Tool to Develop Solution for Bridge Rehabilitation. *Asean Journal of Science and Engineering*. https://doi.org/10.17509/ajse.v2i1.37715
Won, J. S., Lee, G., Dossick, C. S., & Messner, J. (2013). Where to Focus for Successful Adoption of Building Information Modeling Within Organization. *Journal of Construction Engineering and Management*. https://doi.org/10.1061/(asce)co.1943-7862.0000731
Yaakob, M., Athirah Ali, W. N., & Radzuan, K. (2016). *Identifying Critical Success Factors (CSFs) of Implementing Building Information Modeling (BIM) in Malaysian Construction Industry*. https://doi.org/10.1063/1.4960945
Yuan, J., Zeng, A. Y., Skibniewski, M. J., & Li, Q. (2009). Selection of Performance Objectives and Key Performance Indicators in Public–private Partnership Projects to Achieve Value for Money. *Construction Management and Economics*. https://doi.org/10.1080/01446190902748705
Zhang, Y., & Dung, V. P. (2023). Monitoring and Maintenance of Highway Super-Large Bridge Based on BIM Technology. *Wireless Communications and Mobile Computing*. https://doi.org/10.1155/2023/5495549

Development of Machine Vision-Based Road Surface Maintenance Techniques for Climate Change

In Bae Kim[1,*]

[1] R&D Division, Korea Expressway Corporation, Hwaseong, South Korea, crafstman74@ex.co.kr
[*] Correspondence

ABSTRACT

Recently, the lack of expansion joint gaps in highway bridges has been frequently observed due to ground shifting and concrete pavement expansion caused by factors such as alkaline silica reaction. Additionally, with the increase in the number of heatwave days, narrowing of expansion joints leads to issues like zero-gap and pavement blow-up, posing threats to traffic and structural safety. Consequently, we developed an inspection technique capable of monitoring the expansion joint device gap through image analysis while driving at a high speed of 100 km/h, replacing the current manual inspection method. To address error factors encountered during image analysis in the trial application, we employed machine learning to enhance the measurement accuracy of the expansion joint device gap. The proposed method aims to empower maintenance personnel to conduct preventive maintenance and take preemptive measures before problems arise, thereby enhancing traffic safety and reducing costs.

Keywords: Bridge, Expansion Joint, Joint-gap, Bridge Maintenance, Sensors, Structural Health Monitoring, Line-scan Camera, Image Detection, Machine Learning, Machine Vision

1. INTRODUCTION

The Korea Expressway Corporation (KEC) is responsible for constructing and maintaining expressways in Korea. As of 2020, approximately 9,800 highway bridges were in use across the country. Recently, there has been an increase in bridge expansion joint zero-gap failures due to concrete pavement expansion and ground shifting (see Figure 1). These accidents can have adverse effects on the structural integrity of bridges and pose risks to traffic safety. Korea experiences distinct seasons, with significant temperature variations between cold and hot periods. These fluctuations impact bridge expansion joint devices, emphasizing the importance of effective maintenance.

After the completion of bridge construction and a period during which the effects of superstructure creep and shrinkage stabilize, temperature emerges as the primary factor influencing bridge expansion and contraction behaviors. To verify the recent rise in heatwave frequency, we compared average annual temperature, maximum temperature, and heatwave days data from the Korea Meteorological Administration for the 8-year period from 2011 to 2018 with corresponding data from the 30-year period spanning 1981 to 2010 (see Table 1 and Figure 2). The results clearly indicate that the average temperature during the 8-year period exceeded that of the 30-year period. Notably, the number of heatwave days increased by approximately 45%, highlighting the impact of climate change. This trend is expected to continue due to unusual weather phenomena like the heat dome effect.

(a) (b) (c) (d)

Figure 1 Examples of failures at road bridge expansion joints: (a) Blow-up in road pavement, (b) Lack of joint gap in expansion joint, (c) Failure to wing wall of bridge abutment, and (d) Failures by excessive bridge support movement.

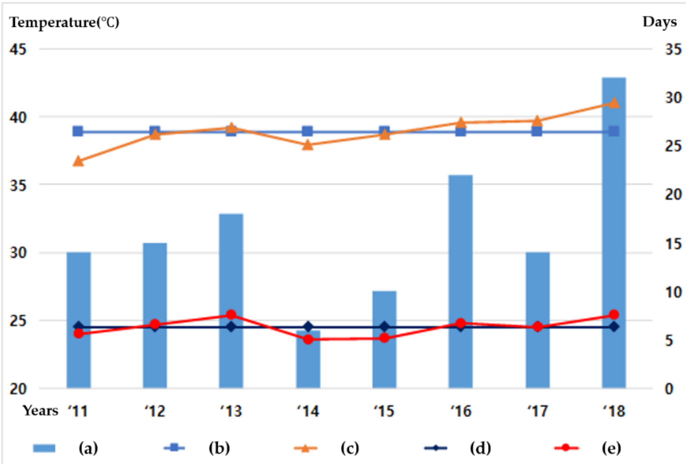


Figure 2 Average annual temperature, average annual maximum temperature, and number of heatwave days (KOREA): (a) Number of heatwave days (days), (b) Recent average maximum temperature (°C), (c) Maximum temperature (°C), (d) Recent average temperature (°C), and (e) Average temperature (°C).

In 2018, the highest number of heatwave days since 2013 was recorded, leading to a rapid increase in cases of zero-gap and pavement blow-up damage. KEC conducted a comprehensive survey of bridges under its management, identifying zero-gap in 276 bridges (2.96% of the total 9,334 bridges). Table 2 outlines the main causes of zero-gap based on onsite investigations.

Table 1 Average annual temperature, average annual maximum temperature, and number of heatwave days (KOREA).

Year	Recent Average (2011–2018)	Older Average (1980–2010)	Difference	Ratio
Average Temperature (°C)	24.5	23.6	+0.8	3.8%
Maximum Temperature (°C)	38.9	37.5	+1.4	3.7%
Number of Heatwave (days)	14.2	9.8	+4.4	45%

Bridge zero-gap damage is prohibitively expensive to repair, and restoration is nearly impossible unless the bridge undergoes significant remodeling. To address these challenges, a novel inspection method was developed to detect issues early and minimize damage in bridge expansion joints. This method relies on automated maintenance inspections, replacing the traditional manual approach. Specifically, we employed a vehicle-type high-speed inspection method using a line scan camera and artificial intelligence (AI) techniques. Through image identification, we accurately determined the expansion joint gap.

Table 2 Analysis of causes of insufficient expansion joint gap (KEC, 2018).

Causes of insufficient expansion joint gap	Number of Bridges	Main causes	
		Expansion of pavement	Deformation of backfill
Sum(① + ②)	276	166	110
① Lack of Abutment + Expansion Joint	96	55	41
② Lack of Abutment	180	111	69

2. LIMITATIONS OF STRUCTURE MANAGEMENT CONDITIONS (considering structural inspectors)

Recently, the Korean government has actively encouraged the adoption of cutting-edge structural safety inspection technologies by revising various laws and regulations. For instance, in March 2020, Annex 10 of the Enforcement Decree of the Special Act on the Safety and Maintenance of Facilities was amended to include provisions for "appearance investigation and image analysis using new technology or inspection robots." In essence, this allows the application of new technology for structural inspections. Additionally, the revision of Article 167 of the Occupational Safety and Health Act emphasized safety management for internal employees by imposing stricter punishment standards on employers (with a maximum fine of $100,000 USD) in cases of workplace accidents resulting from negligence in worker safety management.

As a consequence, the structure management requirements within KEC have become increasingly stringent. Following the revision of the Special Act on Safety and Maintenance of Facilities, the establishment of facility and performance evaluations necessitated evaluating 2,611 Type 3 facilities (bridges) and 4,626 structures. Consequently, structural inspectors conducted 25,219 regular safety inspections, 2,205 detailed safety inspections, and 258 precise safety diagnoses in 2020, both internally and externally. The excessive workload has led to a shortage of management personnel (see Table 3).

Table 3 Number of structural inspections (KEC, 2020).

Safety management	Sum	Bridges	Tunnels	Box culverts
Sum (EA)	27,682	15,636	2,118	9,928
Regular safety inspection[1]	25,219	13,648	1,643	9,928
Precision safety inspection[2]	2,205	1,783	422	-
Precision safety diagnosis	258	205	53	-

[1] Regular safety inspection: An average of approximately 450 EA/yr/branches performed by structural staff
[2] Precision safety inspection (2020): Self-manpower performance (684EA), Performing external services (1,521EA)

Given that these stricter structure management conditions place greater demands on manpower, it is crucial to promote inspection methods and smart technologies that can replace human resources.

3. DEVELOPMENT OF MONITORING SYSTEM (using line scan cameras)

KEC analyzed historical safety inspection data and technical advice, revealing that 78.3% of expansion joint devices requiring replacement lacked proper gaps. Since 2016, we have been researching monitoring methods to proactively maintain suitable gaps between expansion joints.

To achieve this goal, we integrated high-speed line scan cameras, commonly used in fields such as semiconductor inspection and road pavement investigation. We combined image processing technology, which relies on AI-based (machine learning) image sensing during the analysis process, with automatic control technology. The result was the development of the Nonstop Bridge Expansion Joint-Gap Measuring Utility System (NEXUS). This system measures the distance between expansion joints with a 1.0-mm resolution while driving at a high speed of 100 km/h (see Figure 3).

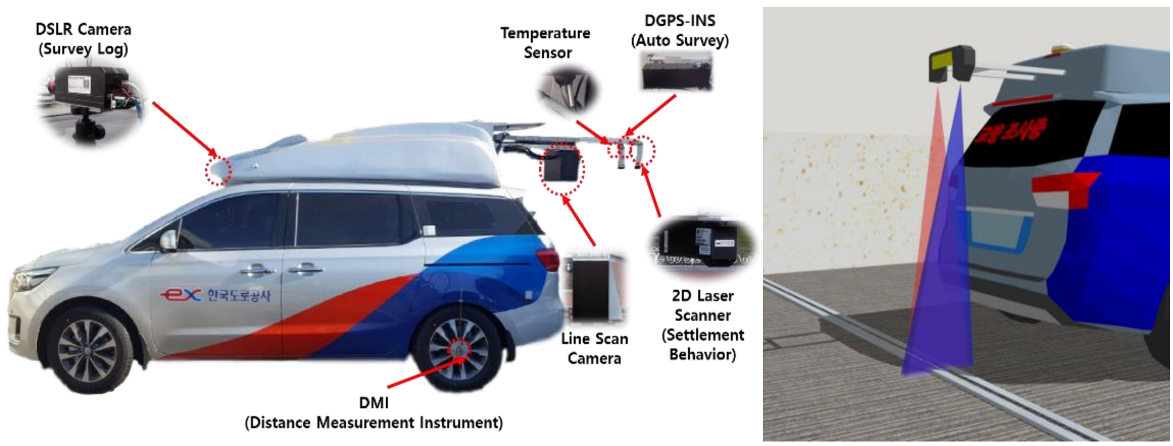


Figure 3 NEXUS system and its configurations.

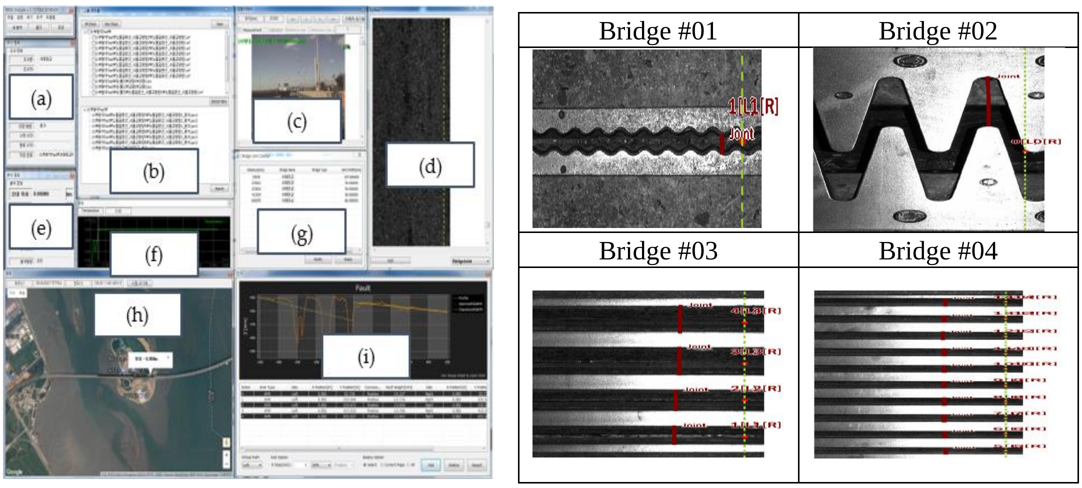


Figure 4 Component items in analysis program and analysis results: (a) survey information, (b) group control, (c) status, (d) road surface image, (e) analysis information, (f) temperature, (g) bridge information, (h) survey trajectory, and (i) flush.

Unlike existing methods that manually measure gaps in highway bridge expansion joints through traffic control and partial blockage, our proposed method utilizes a high-speed line scan camera mounted on a vehicle traveling at 100 km/h. It captures an image across a 40-cm-wide car lane. By integrating geographical information system data (using the longitude coordinate system) from the bridge, the system performs automatic measurements in conjunction with the mounted GPS device. This approach creates a database of expansion gap measurements based on accurate survey images without disrupting traffic flow. In our study, we conducted a test survey on approximately 5,000 bridges along the highway (see Figure 4). The analysis results were then used for big-data-based machine learning algorithms to accurately determine the length of the expansion joint gap based on their type and site conditions.

4. IMPROVEMENT OF IMAGE DISCRIMINATION (using machine learning)

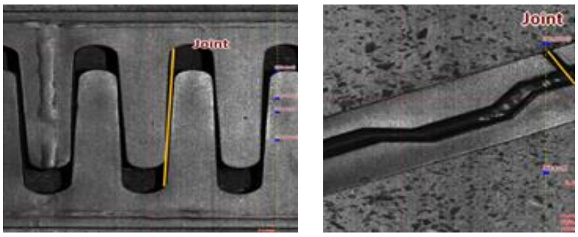


Figure 5 False-detection of expansion joint-gaps.

We aimed to overcome the limitations of the traditional gap identification algorithm. Machine learning can be leveraged to learn from previously investigated images, treating them as big data. This approach allows us to identify the expansion joint device from an input image more accurately and determine the gap value. The input image obtained using the NEXUS equipment was recorded as multiple 10,000 × 1,024-pixel images, depending on the irradiation distance. The image analysis system consisted of cascaded AI vision modules, where the work process analyzed high-resolution input images to facilitate effective calculations. The AI process involves the following steps: original image input → expansion joint area extraction AI network (classification) → expansion joint area image cropping input → gap extraction AI (segmentation) → gap measurement (algorithm) (see Figure 6).

To measure the gap value, the position of the expansion joint device must be specified in the image. Achieving high measurement accuracy requires recognizing the expansion joint device at the pixel level. This can be accomplished using image segmentation methods commonly employed in computer vision. Additionally, a deep learning model trained on extensive data is used to address challenges related to large variations in illuminance, angle, image shape, and noise generation, such as those encountered in road environments.

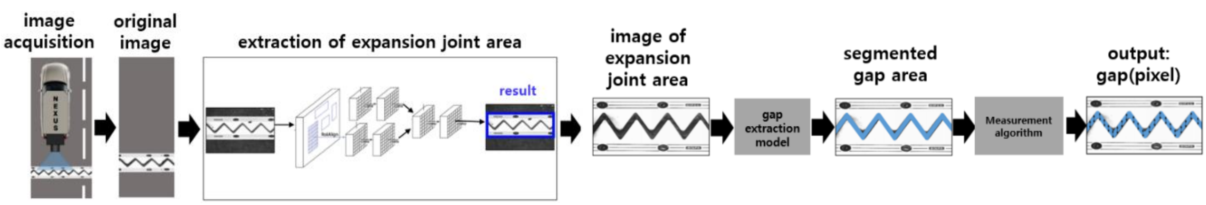


Figure 6 Expansion joint device analysis process for AI method (i.e., machine learning).

The AI machine learning model for discovering expansion joint devices was implemented to solve the classification problem. We used ResNet as a feature extractor and trained the model in the Tensorflow2 environment. The original image, with a size of 10,000 × 1,024 pixels, was divided into image patches measuring 1,000 × 1,024 pixels. These patches were input to the AI model with minimal distortion. Subsequently, 19 image patches were generated for each line scan image, and classification was performed to identify the patches containing the expansion joint device (see Figure 7).

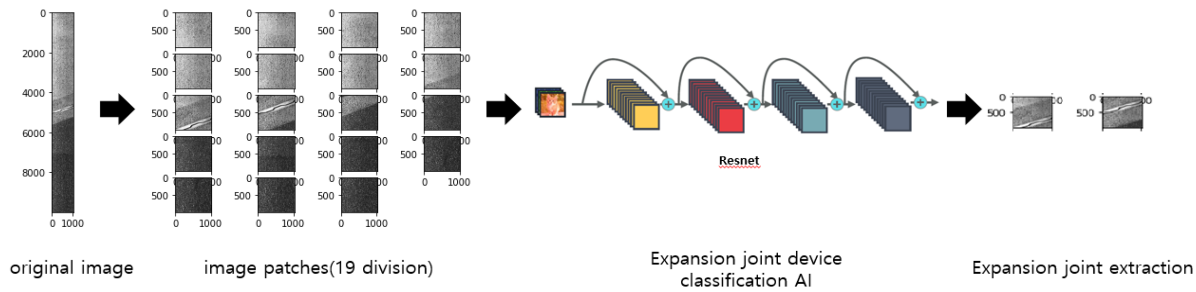


Figure 7 Image segmentation and classification to search for expansion joint device.

To pinpoint the point with the minimum gap distance, precise segmentation of the gap region from the image patch where the expansion joint device appears is necessary. For this purpose, we employed the U-Net model, a representative segmentation model, which was trained in the Tensorflow2 environment. The first AI model generated and learned training data by labeling pixels in the corresponding area to extract metal parts with characteristic textures from the line scan images of the expansion joint device. The second AI model was trained to find the gap region between the extracted joint devices (see Figure 8).

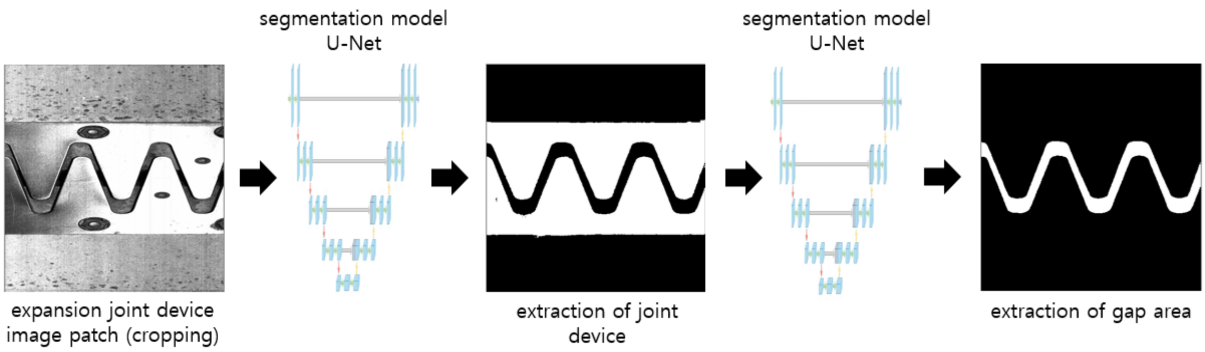


Figure 8 Segmentation process for extracting gap area.

After obtaining a binarized image with the gap region and the background, we identified the point with the minimum gap distance. The minimum gap distance in pixels was then converted to the final gap value in millimeters by multiplying it by the information in millimeters per pixel.

4.1 AI-based Image Analysis

Line-scan images obtained using NEXUS were converted into high-resolution digital images with dimensions of 10,000 × 1,024 pixels. The unit length (mm) of the pixels can be calculated by considering the vehicle speed at the time of capture. If the expansion joint device can be measured in pixels, the actual gap distance value can be inferred within the error range. Therefore, accurately recognizing the position of the expansion joint device in the line scan image using a program and measuring the gap size in pixels are critical for achieving high measurement accuracy.

To perform highly accurate analyses, we trained a large number of line scan images using deep learning. Cascade-type analysis procedures were configured to efficiently analyze high-resolution input images. The analysis involved the following steps: (1) ultrafast line scan image input, (2) extraction of small image patches in a square matrix, (3) recognition of the expansion joint device among the image patches, (4) expansion gap division, and (5) gap distance analysis and actual value calculation (see Figure 9).

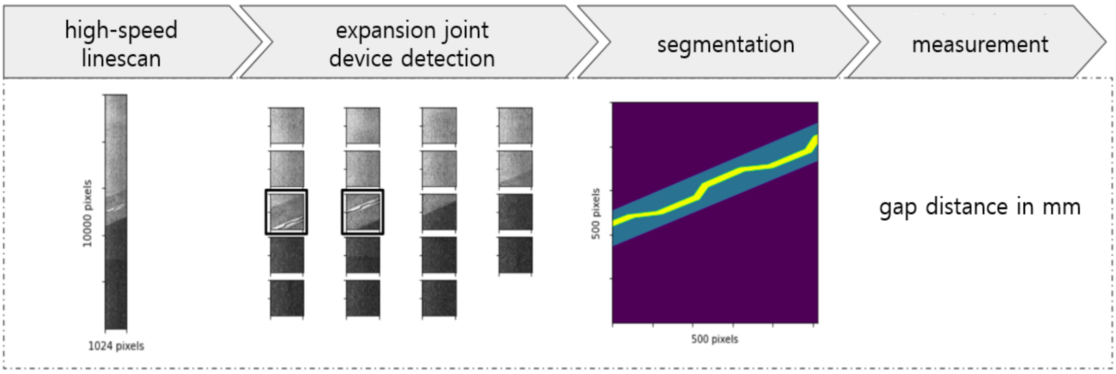


Figure 9 Spacing measurement process of expansion joint device.

4.2 Expansion Joint Device Recognition

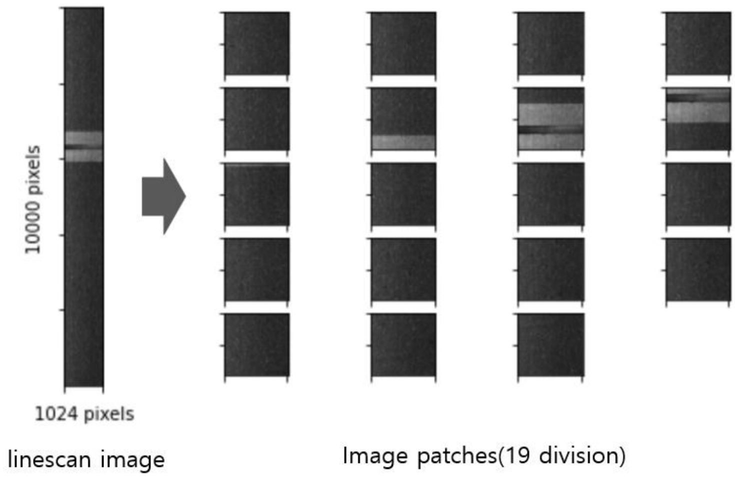


Figure 10 Line scan image segmentation to search for expansion joint devices.

Line-scan images have a high resolution of more than 10,000,000 pixels and require significant memory and computational resources when analyzed using deep neural networks. Consequently, it is necessary to divide the image into smaller patches for efficient computations. Using the window sliding method, we extracted images every 1000 pixels in the longitudinal direction, resulting in a total of 19 image patches with dimensions of 1000 × 1024 (see Figure 10). Each image patch was then resized to 224 × 224 pixels and subjected to binary classification.

We employed a convolutional neural network (CNN), an effective architecture for image analysis. We labeled patches containing the expansion joint device as y = 1 and patches without the expansion joint device as y = 0, classifying them using logistic regression. The CNN inferred the probability P that an expansion joint exists in the image patch. The prediction error for the image patch was expressed using the binary cross-entropy (BCE) method as follows:

$$L = -ylogP - (1-y)log(1-P) \tag{1}$$

Due to class imbalance (a ratio of 1:9), we assigned weights to each class:

$$L_{weighted} = -\epsilon \times ylogP - \mu \times (1-y)log(1-P) \qquad \text{where } \epsilon = 9.0\,, \mu = 1.0. \tag{2}$$

The cost function (J) was calculated by averaging errors for (N) training data and adding L2 regularization to reduce model overfitting:

$$J = \frac{1}{N}\sum_{j=1}^{N} L_{weighted}^{j} + \lambda|w|^2 \qquad \text{where } \lambda = 10^{-4} \tag{3}$$

The gradient descent method updated the model parameters to minimize the cost function (J):

$$w \Leftarrow w - \alpha \cdot g \qquad \text{where } \alpha = 10^{-4},\ g = \frac{\partial J}{\partial w} \tag{4}$$

We tested ResNet, MobileNet, and EfficientNet—well-known CNN architectures (see Fig. 11). Through testing, the EfficientNet model demonstrated the highest performance, achieving a recognition accuracy of 97.57% for the expansion joint device. Notably, starting learning from randomized initial parameters outperformed transfer learning from ImageNet. This difference arises because the analysis image exhibits distinct characteristics compared to the general features found in ImageNet.

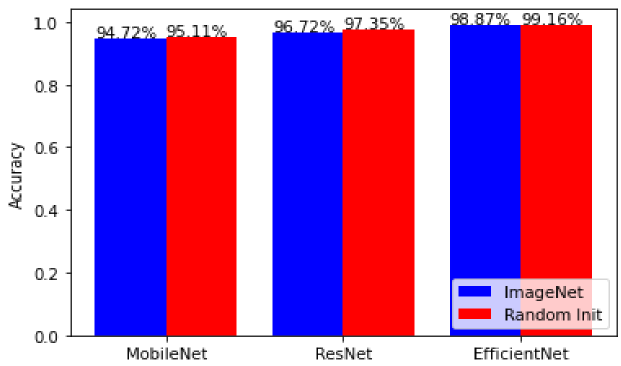


Figure 11 Recognition accuracy test results for each CNN structure.

4.3 Expansion Joint Gap Segmentation

Even after extracting the image patch of the expansion joint from the line scan image, image segmentation is necessary within the detected image to accurately measure the expansion gap. Image segmentation involves pixel-level classification, determining the class to which each pixel belongs (i.e., expansion joint device or background). An image mask representing the pixels corresponding to the expansion joint device can be generated using an image segmentation algorithm.

Figure 12 illustrates the image patch (left) and the correct mask (right) of the expansion joint device. The neural network structure takes image patches as input and performs binary classification on individual pixels. The deep learning model learns to predict a mask that closely aligns with the correct answer. In this case, pixel classification should consider both global and local characteristics of the image rather than individual pixel values.

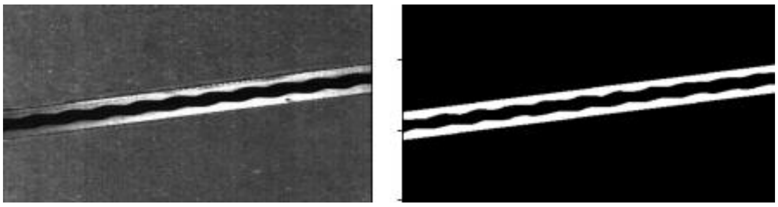


Figure 12 Image of flexible expansion joint and mask of correct answer.

For this study, we employed the U-Net model, originally developed for sophisticated analyses of organ lesions in biomedical science. U-Net is well-suited for precisely detecting the shape of the analysis target by simultaneously learning global and local information from the image. We modified the U-Net structure to create EfficientNet, a neural network optimized for extracting image features.

U-Net infers the probability that an arbitrary pixel is an expansion joint device and learns the correct answer for each pixel from the correct answer masking. Therefore, the prediction error for all pixels (M) of the image patch expressed as BCE is:

$$L_{patch} = \sum L_{pixel} = \sum_{i=1}^{M}(-y_i \log P_i - (1 - y_i)\log(1 - P_i)) \qquad i = 1,2,3,\dots,N \tag{5}$$

The final cost function (J) is expressed by summing the mean error for (N) image patches and the L2 regularization term:

$$J = \frac{1}{N}\sum_{j=1}^{N}\left(L_{patch}^{j} + \lambda|w|^2\right) \qquad \text{where } \lambda = 10^{-4} \tag{6}$$

The gradient descent method updates model parameters in the direction of reducing the cost function (J) as follows:

$$w \Leftarrow w - \alpha \cdot g \qquad \text{where } \alpha = 10^{-4}, g = \frac{\partial J}{\partial w} \tag{7}$$

We compared the performance between U-Net's masking image and the correct masking image in the test dataset. Table 4 shows the pixel-level classification performance for the test set: the pixel precision of the expansion joint device was 96.61%, recall rate was 94.38%, and f1-score was 95.49%. The f1-score of the expansion joint device pixel detection was within 5%. In the post-processing process, by correcting the error of the predicted masking image of U-Net, the minimum gap point was detected and the distance was measured.

Table 4 Gap identification accuracy by type of expansion joint (2017–2019, KEC).

Accuracy by type of expansion joint (%)	Precision	Recall	F1-Score
Positive (pixels of expansion joints)	96.61	94.38	95.49
Negative (other pixels)	99.23	99.55	99.39

4.4 Gap Distance Analysis Algorithm

The texture of the pixels in the expansion gap area to be measured appears somewhat irregular depending on the gap, foreign matter, and type of expansion joint device. Texture irregularity is a factor that makes it difficult to distinguish pixels in the expansion gap area.

The metal surface constituting the expansion joint device has a consistent texture compared to the gap region. This means that it is easier to extract the expansion joint device than the gap area. Therefore, to analyze the gap distance, the expansion joint device is extracted first, and the gap area is extracted again from the resulting image. Image segmentation using U-Net was applied to both area extraction processes (see Figure 13).

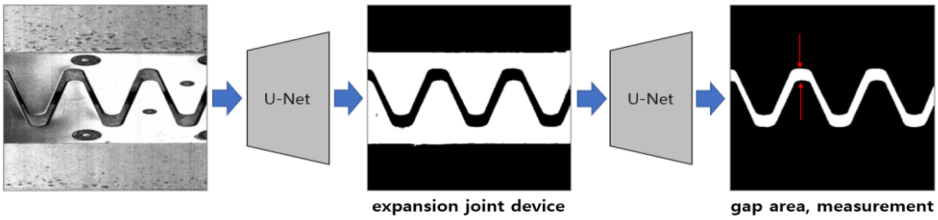


Figure 13 An approach for elaborate gap area extraction.

Because the U-Net output represents the probability that each pixel is a gap area, binarization is performed by applying a threshold value of 0.5. When defining a binarization function as a response, the formula to search the x-coordinate of the minimum gap is as given below. Note that the same applies to rail-type expansion joint devices:

$$\underset{1 \le x \le 512}{\arg\min} \sum_{y=1}^{512} response(x, y) \tag{8}$$

The input and output of the U-Net model have a size of 512 × 512. By binarizing the final output probability map, we search for the x-coordinate with the smallest number of pixels with a value of 1. The number of pixels at the coordinates is the gap distance in pixels, and the actual gap distance is obtained by multiplying the distance value per pixel:

$$(distance) = (\#\, of\ pixel) * (\frac{mm}{pix}) \tag{9}$$

Figure 15 shows the pseudocode for obtaining the minimum gap distance from the output probability map of U-Net.

```
response = threshold(output of U-Net, 0.5)
for x in range(0, width of response)
    distList[x] = # of response which value is 1
    if distList[x] is minimum
        minIdx = x
minimum distance = distList[minIdx] * mm/pixel
```

Figure 14 Pseudocode to obtain minimum gap distance from U-Net output.

4.5 Gap Identification Verification

Based on the abovementioned results of AI-based gap identification, we randomly selected 10,526 places among 12,825 expansion joint device big data images obtained previously to determine the expansion joint device-discrimination of expansion joint device gap. The results are presented below for each expansion joint device type. The position where the minimum spacing was measured is indicated by a red line. For rail-type expansion joints in which several gaps appear at once, the starting and end gaps of the part with the smallest actual gap value are indicated by red lines (see Figure 16).

After comparing the identification rate results from the existing traditional algorithm, we found that the Finger type, which initially had the lowest discrimination accuracy, improved the most by 41% to achieve an accuracy of approximately 92%. Overall, the average identification accuracy improved by 27.5% to reach 95%. Despite the significant improvement in identification accuracy, certain scenarios still pose challenges. For instance, if a foreign substance exists inside the expansion gap, if a sealing agent is applied to the surface to prevent water leakage, or if the body of the expansion joint device is damaged, manual correction becomes necessary (see Table 8).

Table 1 Accuracy of expansion gap identification using machine learning algorithm.

Accuracy (%)	Average	Mono-cell	Finger	Rail	etc.
Conventional algorithm	67.5	71	51	87	61
AI algorithm (machine learning)	95	98	92	99	91
Improvement rate	↑27.5	↑27	↑41	↑12	↑30

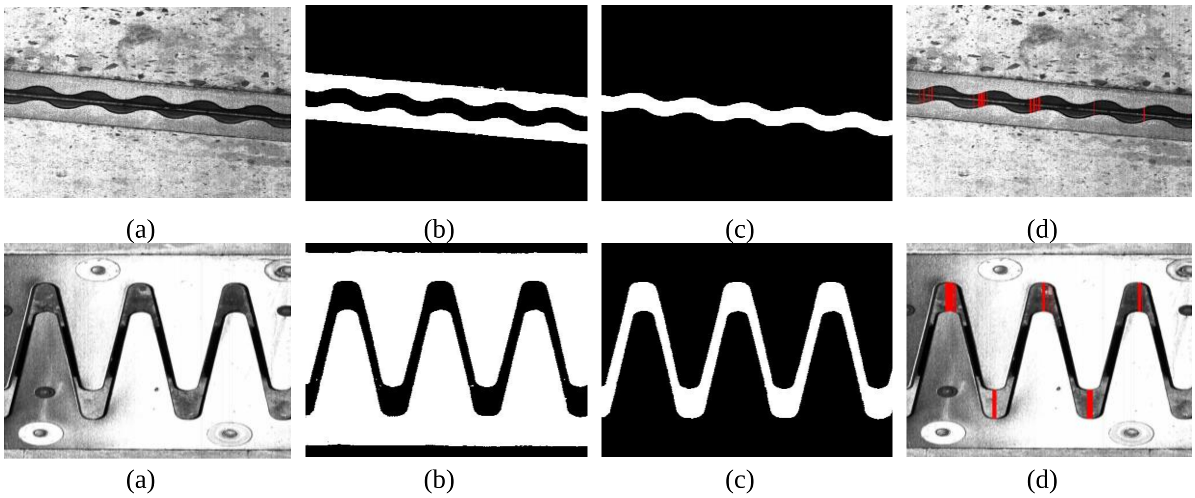


Figure 15 Machine-learning-based verification results of expansion joint device gap identification: (a) original image, (b) extraction of expansion joint area image, (c) extraction of gap area image, and (d) final measurement result image.

5. CONCLUSIONS

The following conclusions were derived from this study:

1. Temperature Trends and Gap Monitoring:

Based on data from the Korea Meteorological Administration, we analyzed the average annual temperature, average maximum temperature, and annual heatwave days for the 8-year period from 2011 to 2018. We compared these findings with data from the 30-year period spanning 1981 to 2010. Notably, during the 8-year period, the average temperature was significantly higher, and the number of heatwave days increased by approximately 45%. These results underscore the importance of monitoring the bridge expansion joint device gap, especially in areas where temperature-related expansion behavior dominates.

2. Development of the NEXUS Device:

To monitor the highway bridge expansion joint device gap, we developed the NEXUS device. This device investigates road surface images with a 1.0-mm resolution, determining the expansion gap through image sensing. Equipped with a high-speed line scan camera, GPS, and distance measurement instrument sensors, the NEXUS system operates while driving at 100 km/h.

3. Improving Gap Identification Accuracy:

The machine learning model for identifying expansion joint devices in survey images utilized ResNet as a feature extractor, while the representative segmentation model for detecting gap areas employed U-Net. These models effectively addressed the classification problem. Testing with a random selection of 10,526 previously acquired big data images demonstrated an improved expansion gap identification accuracy of 27.5%, increasing from 67.5% to 95.0%.

4. Efficiency and Safety Benefits:

The use of the NEXUS device for automatic surveys of bridge expansion joint device gaps while driving at high speed significantly reduces survey time. The average time per bridge decreased from approximately 1 hour (using the existing manual inspection method) to approximately 3 minutes per bridge opening. Additionally, proactive maintenance based on accumulated gap data can prevent gap narrowing (lack of joint gap) and mitigate risk factors associated with future temperature-related expansion behavior of bridges. This contributes to both traffic safety and cost reductions.

6. REFERENCES

Elgaaly, M. (1988). Thermal gradients in beams, walls, and slabs. Structural Journal, 85(1), 76-81.

Fu, H. C., Ng, S. F., & Cheung, M. S. (1990). Thermal behavior of composite bridges. Journal of Structural Engineering, 116(12), 3302-3323.

Ikeuchi, K., & Kanade, T. (1988). Automatic generation of object recognition programs. Proceedings of the IEEE, 76(8), 1016-1035.

Kwak, H. G., Kwon, S. H., & Ha, S. H. (2011). Temperature Distribution and It's Contribution to Self-Equilibrium Thermal Stress in Bridge. Journal of the Computational Structural Engineering Institute of Korea, 24(5), 531-542.

Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. Artificial intelligence, 31(3), 355-395.

Park, J. C. (2015). Evaluation of thermal movements of a Cable-Stayed Bridge using temperatures and displacements data. KSCE Journal of Civil and Environmental Engineering Research, 35(4), 779-789.

Shin, D. W., Kim, K. N., Choi, C. H., & Lee, S. H. (2013). Evaluation of Effective Temperature for Estimate Design Thermal Loads in Steel Deck of Steel Box Girder Bridges. Journal of the korea Institute for Structural Maintenance and Inspection, 17(6), 77-87.

Vecchio, F. J. (2000). Disturbed stress field model for reinforced concrete: formulation. Journal of structural engineering, 126(9), 1070-1077.

SIMULATION-BASED ANALYSIS OF VARIABLE SPEED LIMIT FOR TRAFFIC PERFORMANCE AND SAFETY DURING INCIDENTS

S. Hemapan [1,*], A. Karoonsoontawong [2] and K. Kanitpong [3]

[1] Transportation Engineering/School of Engineering and Technology, Asian Institute of Technology, 58 Moo 9, Km. 42, Paholyothin Rd, Klong Luang, Pathum Thani /Email surapat.he@gmail.com
[2] Department of Civil Engineering/Faculty of Engineering, King Mongkut's University of Technology Thonburi, 126 Pracha Uthit Rd, Bangmod, Thung Kru, Bangkok/Email ampol.kar@kmutt.ac.th
[3] Transportation Engineering/School of Engineering and Technology, Asian Institute of Technology, 58 Moo 9, Km. 42, Paholyothin Rd, Klong Luang, Pathum Thani /Email kanitpon@ait.asia

ABSTRACT

Traffic congestion poses substantial issues to the environment, society, and the economy, with traffic incidents playing a crucial role. Intelligent Transportation Systems, particularly active traffic management tactics such as Variable Speed Limit (VSL) systems, offer options for reducing congestion and increasing safety. VSL systems typically include fixed sensors to assess traffic conditions and dynamic variable message signs (VMS) positioned at predefined places to implement adjusted speed limits based on the average, volume, and percentage occupancy. A simulation model was developed using the VISSIM COM interface with MATLAB to evaluate the effect of VSL strategies on traffic safety and performance during incidents. This study compared the effectiveness of VSL with VMS spacings at 1 km and 2 km intervals. The study area encompasses Motorway Number 7, which spans from KM 19+000 to KM 21+500. The results show that VSL systems significantly improve traffic flow, particularly in mild to moderate congestion situations. However, during periods of high traffic congestion, the effectiveness of VSL system decreases because speed limit signs may not reduce vehicle speeds effectively before reaching incident areas. To optimize safety and performance benefits, it is recommended to install VMS signs at intervals of 1 kilometer. The analysis of VSL with VMS spacing at 1 kilometer at different flow rates (0.25, 0.50, 1.00, and 1.25 times as much as the peak traffic volume) reveals up to 48%, 18%, 15%, 89%, 88%, 79% and 84% improvement in average delay per vehicle, average speed, total travel time, average number of stops per vehicle, rear-end conflicts, lane change conflicts, and total number of conflicts, respectively.

Keywords: Traffic Simulation, Variable Speed Limit, Variable Message Sign, Traffic Incident

1. INTRODUCTION

Accidents on Thai highways frequently occur due to the high speeds at which most people drive. From 2019 to 2024, there were 11,232 recorded accidents on highways 7 and 9 (Sangakong, 2567). Traffic accidents, Vehicle breakdowns, road debris, roadside distractions, and cargo spills are among the scenarios included in the word "incident". The impact of incident-induced congestion is critical, as accidents make the capacity of the road drop which leads to traffic breakdowns (Mfinanga & Fungo, 2013). In high-density metropolitan areas, traffic congestion creates significant environmental, social, and economic problems.

To solve traffic congestion problems, Nowadays Intelligent Transportation Systems (ITS) are used in modern transportation, especially Active traffic management (ATM) solutions, such as variable speed limits, ramp metering, and dynamic lane use control, which seek to minimize both recurring and non-recurring congestion by altering traffic conditions in real-time. These solutions improve traffic flow and help to create more efficient and sustainable transportation networks (Boyles et al. 2018). There was a study found that exceeding the speed limit is a major contributor to fatal crashes. Roadways and segments with higher design speeds result in fewer crashes than those with lower design speeds. Thus, variable Speed Limit (VSL) systems, use dynamic VSL signs and detectors to continuously monitor traffic conditions. These VSL signs can show appropriate speed limits, allowing transportation managers to set speed limits based on current traffic, weather, or other factors such as lane closures, limited visibility, work zones, or incidents in real time. VSL systems are seen as tools for smoother traffic flow, preventing congestion, enhancing safety, reducing driver stress, and improving travel time (Allaby et al., 2007; Sadat & Celikoglu, 2017; Farrag et al., 2020; Srisurin & Chalermpong, 2021; Guerra et al., 2022).

This study simulated VSL by using the VISSIM COM interface with MATLAB, with two primary goals: 1) to assess and improve VSL strategies for improving traffic performance and safety during incidents, and 2) to compare and discuss the benefits and drawbacks of various VSL strategies and existing scenarios. While previous research has largely supported the usefulness of VSLs in improving traffic safety and efficiency, their performance is impacted by a variety of circumstances. This project attempts to test these tactics on Thailand's roads and provide insights into their usefulness and effectiveness in practical applications.

2. LITERATURE REVIEW

The beginning of Variable Speed Limit (VSL) systems started in the United Kingdom in the 1960s. After that, Germany and the Netherlands established their systems in the 1970s and 1980s, respectively, and have since spread globally. VSL evolved from incident detection to be the tools used to improve safety, prevent congestion, and improve traffic flow and environmental sustainability over time passed (Lu & Shaladover, 2014; Grumert et al., 2018; Soriguera et al., 2017; Hoogen & Smulders, 1994; Hegyi et al., 2005; Khondaker & Kattan, 2015; Sadat & Celikoglu, 2017; Farrag et al., 2020; Zegeye et al., 2011). In the year 2018, Grumert et al. (2018) categorized VSL algorithms into four categories: Rule-Based, Fuzzy

Logic-Based, Analytical, and Control Theory-Based Algorithms, each with a distinct goal and varying complexity. In this paper, the rule-based model is used because it is easy to understand and most used practically. An interesting algorithm is rule-based and has been used in several research studies as follows. Allaby et al. (2007) developed rule-based algorithms, which use data from loop detectors to estimate speed limitations based on characteristics such as volume, occupancy, and average speed by using real-time data from downstream detectors to change speed restrictions. While these systems are efficient in improving safety, but increase travel times. Building on this basis, Sadat and Celikoglu (2017) used the rule-based approach to optimize VSL systems, focusing on the importance of driver compliance and demonstrating that VSL systems successfully reduce congestion and enhance environmental conditions when driver compliance is high. Boyles et al. (2018) employed this rule-based VSL system and used it in specific corridors with recurring and non-recurring congestion situations. The results showed traffic advantages, including fewer delays and faster travel times, demonstrating the effectiveness of rule-based algorithms in optimizing VSL tactics for better traffic flow and safety.

This study adopted this algorithm in different locations which is a motorway in Thailand during the incident. To create the microsimulation model, PTV VISSIM which is the major microscopic simulation program, provides the necessary realism and precision to examine various traffic situations before deployment is used for this study (MANUAL PTV Vissim, 2023). But VISSIM does not have an incident-create function, so the incident is represented using incident simulation approaches such as the 'Parking Lot' approach (Xie et al., 2022; Karaer et al., 2020; Farrag et al., 2021). The result from VISSIM will provide the performance aspect and the trajectory data which is exported to the FHWA's Surrogate Safety Assessment Model (SSAM) to evaluate the conflict frequency related to the safety aspect (Kurker et al., 2014)

3. METHODOLOGY

3.1 Study Area

This study focuses on a specific road section of Thailand's Motorway Number 7, which starts from km 19+000 to km 21+500 inbound, covering around 2.5 kilometers. VISSIM is used to create a road section that begins with 5 lanes and merges into 4 lanes, each lane is 3.6 meters wide. The study area is shown in Figure 1.

Figure 1 Study area

3.2 Data Collection

First, the geometric data is collected from construction drawings and aerial photographs which show the road's shape and design. Second, the demand and speed data, the traffic flow, vehicle speed, and type were collected from 15:30 to 18:00 in 5-minute intervals across 12 typical workdays, with a warm-up period from 15:30 to 16:00 and analysis beginning at 16:00. The vehicle composition was cars (93%) and trucks (7%). Lastly, Incident data is collected from the previous five years including the incident duration and the number of lanes closed for each accident. All data were acquired from the Department of Highways, Inter-City Motorway Division.

3.3 Calibration Process

For the calibration process the study area is divided into three sections which are upstream (5 lanes), merging (5 to 4 lanes), and downstream (4 lanes) as shown in Figure 2. The Wiedemann 99 car-following model was used for calibration, with three parameters which are standstill distance (CC0), time headway (CC1), and following variation (CC2). These variables are adjusted separately in each section to reflect capacity, flow, and average speed in real life by ensuring capacity differences were within 5% of observed data (Dowling et al. 2004). For the average speed, using the RMSE approach (Hollander & Liu, 2007), the aim was to keep the average speed deviation within 10 km/h. Flow calibration employed the GEH statistic, with values less than 5 indicating a good match between simulated and real traffic flows. CC0, CC1, and CC2 settings for best-fit calibration results in each section are as follows. Upstream: CC0 = 1.5 m, CC1 = 1.5 s, CC2 = 11 m; Merging: CC0 = 2.2 m, CC1 = 0.9 s, CC2 = 13 m; Downstream: CC0 = 1.5 m, CC1 = 1.5 s, CC2 = 12 m.

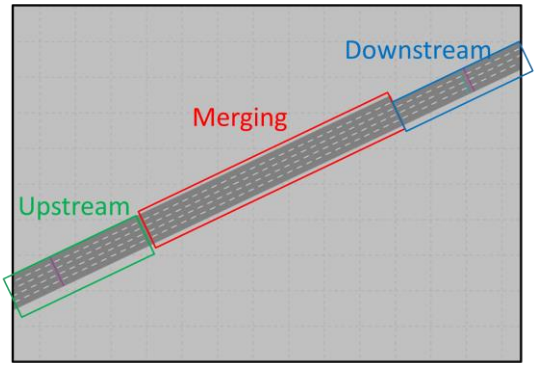


Figure 2 Road Calibration Sector

To reduce the impact of stochastic from VISSIM the simulation was run 8 times, and the result was as follows: The model's capacity was calculated using a speed-flow graph (Figure 3) (Srisurin, 2020), with flow rates that vary from 2500 to 8000 veh/h. The capacity from the model, based on the point of maximum flow before a speed drop, was determined to be 6414 veh/h/ln, representing a 4.2% variance from field data which is calculated from the capacity formula (Krisingson S. & Methawaranthon A., 2019). The RMSE for speed was 3.00 km/h, and all GEH values were less than 5.

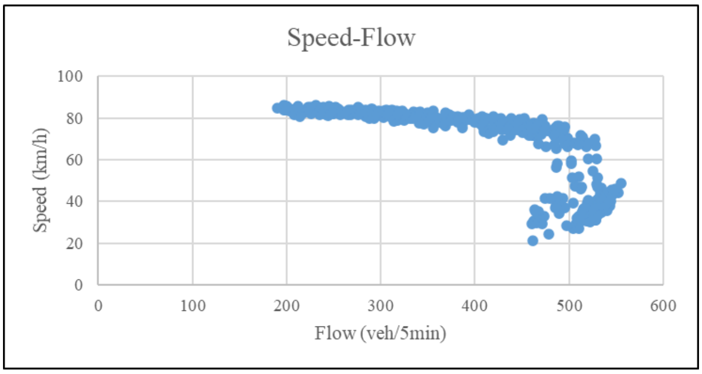


Figure 3 Modeling Speed-Flow graph.

3.4 Incident Simulation

The "parking lot" technique is used to create the incident in VISSIM illustrated in Figure 4. The incident duration is represented by park duration and the number of lane closures is represented by the number of parking lots.

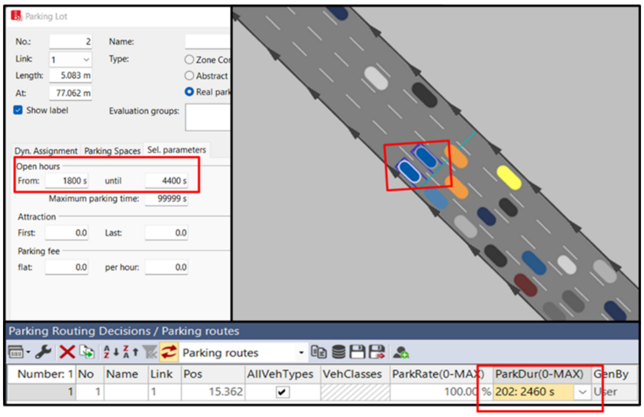


Figure 4 Incident simulation by using the parking lot.

3.5 Variable Speed Limit Simulation

VSL simulation was performed using VISSIM and MATLAB via a COM interface. Variable message signs show speed limits, while detectors measure volume, average speed, and occupancy. These data were sent to MATLAB, which computed new speed limits every 60 seconds based on Sadat and Celikoglu (2017), shown in Figure 5.

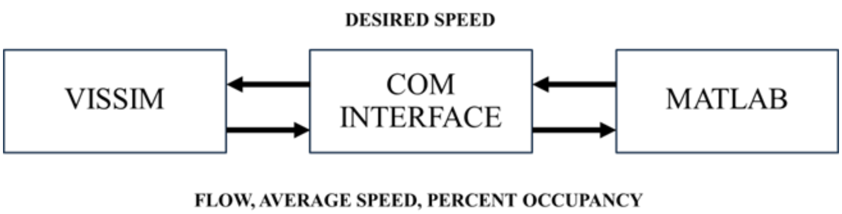


Figure 5 VISSIM works with MATLAB via COM INTERFACE.

There is one restriction for variable speed limits. In the case that there are 2 variable message signs in a row, the upstream speed limit must not be higher than the downstream speed limit by more than 20 km/h to preserve smooth traffic flow and avoid congestion risk. (Allaby et al. 2007).

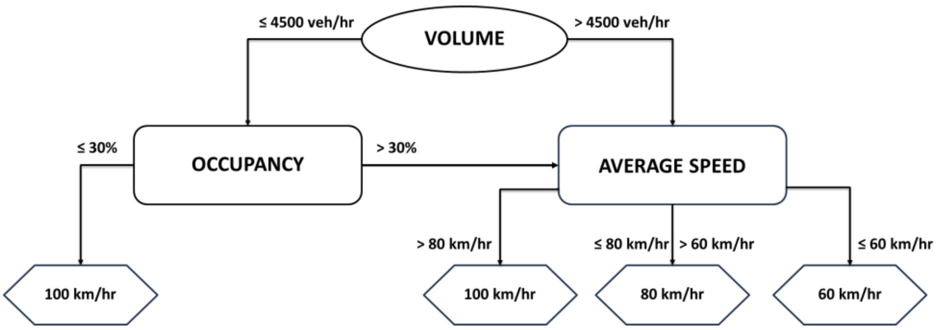


Figure 6 Variable Speed Limit Algorithm for car.

Figure 6 shows 3 thresholds adapted from Allaby et al. (2007). First is volume threshold 4500 veh/hr is vehicle flow at the level of service C of this road section. Second is percent occupancy, thirty percent is the percent occupancy from the data plot that represents the critical occupancy before the average speed of the vehicle breaks down from the free flow speed. Third is the appropriate average speed that was set as 100, 80, and 60 km/hr. The average speed assigned to cars is 20 km/h higher than that for trucks for the safety aspect.

3.6 Alternative Scenarios

Three alternatives as shown in Figure 7 were evaluated:

1. Existing Situation: Current conditions without VSL.
2. 1 km VSL Coverage: Speed limit signs covering 1 km.
3. 2 km VSL Coverage: Speed limit signs covering 2 km.

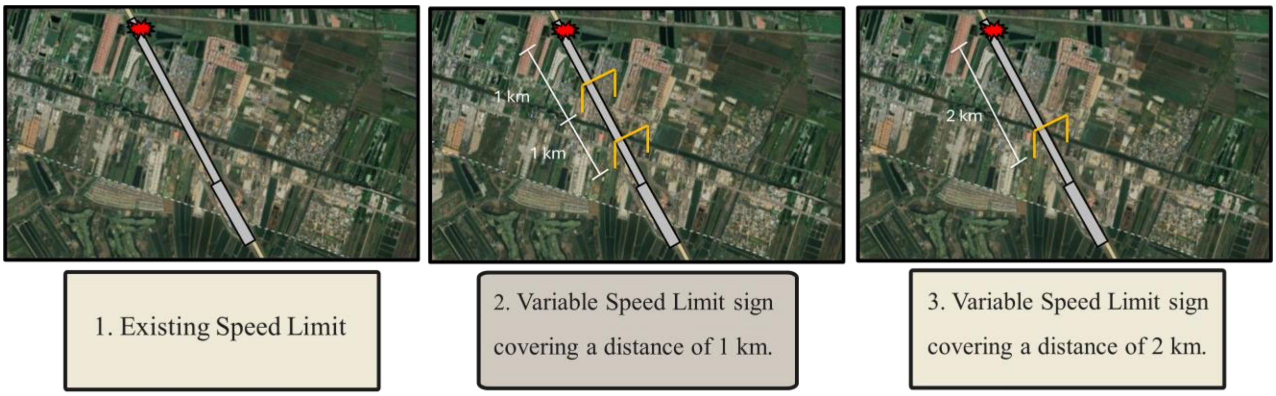
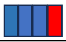
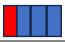
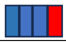
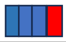


Figure 7 Alternative Scenarios.

The objective was to assess the impact of without and with VSL in different coverage distances on traffic performance and safety. Each alternative was visualized to illustrate the differences and expected outcomes.

***Table 1 Summarize all scenarios*.**

Alternatives	Number of lanes closed (lanes)	Close patterns	Incident duration (min)	Number of scenarios
Existing situation	1		41	2
			87	
			41	2
			87	
	2		41	2
			87	
			41	2
			87	
VSL that cover 1 km	1		41	2
			87	
			41	2
			87	
	2		41	2
			87	
			41	2
			87	
VSL that cover 2 km	1		41	2
			87	
			41	2
			87	
	2		41	2
			87	
			41	2
			87	
Total				24

In Table 1, three alternatives are outlined: the existing condition, a VSL covering 1 km, and a VSL covering 2 km. Each alternative considers scenarios with one or two lanes closed. For single-lane closures focus on leftmost and rightmost lanes, which historically have the highest accident rates (60% and 26%, respectively). Two-lane closure scenarios include closing the two leftmost or rightmost lanes because incidents in the leftmost or rightmost lanes might lead to closures in adjacent lanes due to their severity. Incident durations of 41 minutes (mean) and 87 minutes (90th percentile) were used to represent typical and severe incidents, respectively, resulting in 24 analyzed scenarios.

3.7 Performance Measure

This study's performance measures are divided into two categories: traffic performance and traffic safety. VISSIM's traffic performance outputs include Average Delay per Vehicle, Average Speed, Total Travel Time, and Vehicle Throughput, which reflect cars that arrived at their destinations. Traffic safety measures produced by VISSIM, and the Surrogate Safety Assessment Model (SSAM) include the average number of stops per vehicle, which indicates the possibility of secondary accidents (Farrag et al. 2020), as well as Rear-End Conflicts, Lane Change Conflicts, and Total Conflicts (Rear-End + Lane Change Conflicts).

4. RESULTS AND DISCUSSIONS

This chapter shows the simulation results and analyzes how the Variable Speed Limit (VSL) approach affects traffic performance and safety. The paper focuses on the key parameters that impact traffic performance and safety which are average delay per vehicle and total conflict to represent the traffic performance and traffic safety respectively. The graph consists of the dot which shows the amount of that performance measure and the error bar which is 95 confidence intervals of each scenario from 8 runs, used to show the difference between the Existing, VSL covering 1 and 2 km, which is differently significant or not. If the error lines in each alternative collab each other means they are not significantly different. The following charts mark the lanes as follows: Lane 1 is the most left lane, and Lanes 2, 3, and 4 follow orderly to the right. The abbreviations VSL1 and VSL2 refer to variable speed limit sign that covers one and two kilometers, respectively.

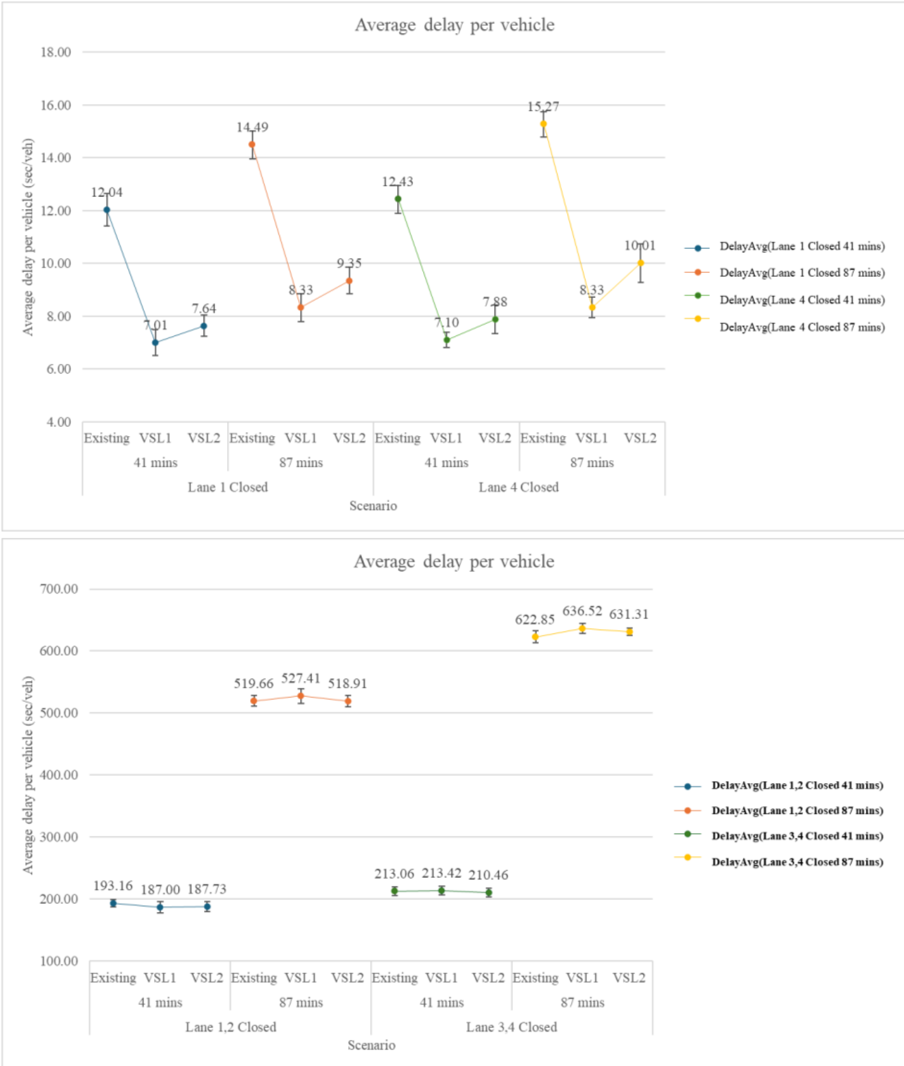


Figure 8 Average delay per vehicle results

Figure 8 shows that when one lane (lane 1 or lane 4) is closed, the average delay per vehicle of VSL covering 1 km and 2 km is significantly better than the existing scenario. When two lanes are closed (lanes 1 and 2 or lanes 3 and 4), the results from the existing scenario and VSL scenario are not significantly different. Furthermore, when comparing VSL cover 1 and 2 km itself, there is no scenario that these two are significantly different.

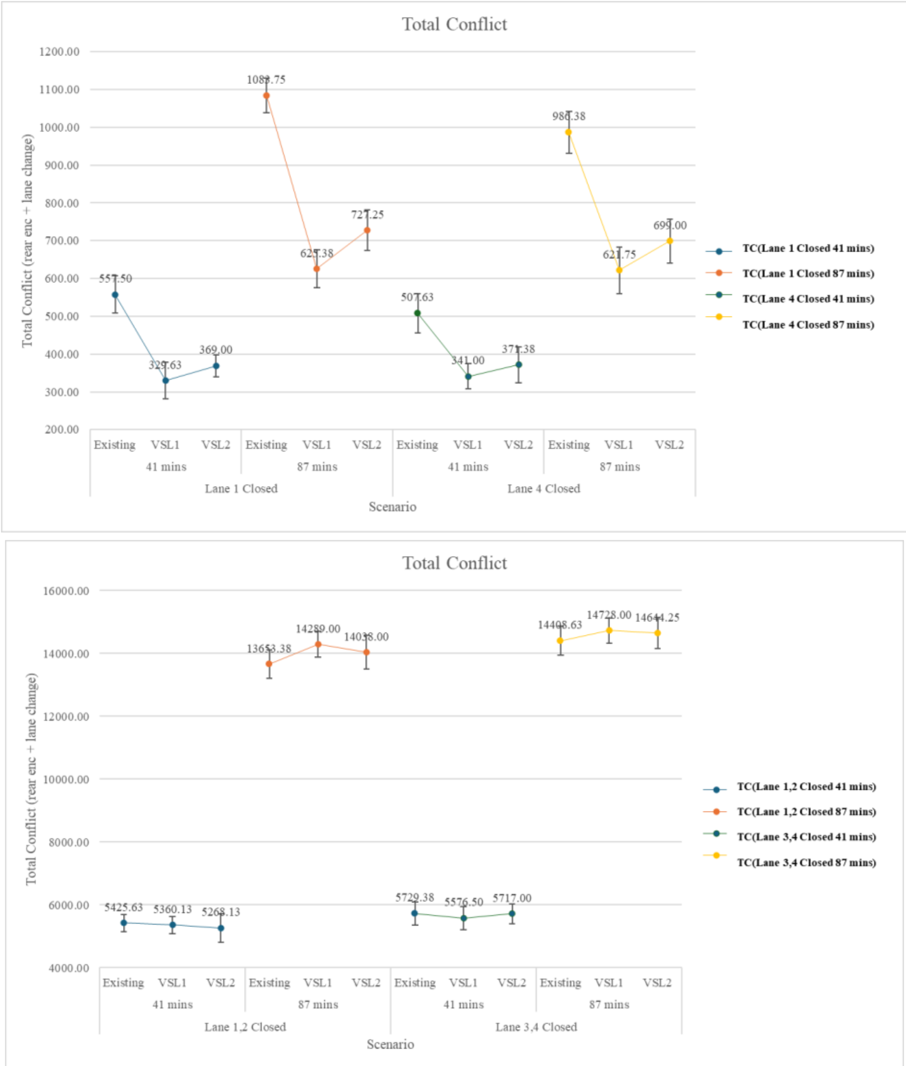


Figure 9 Total conflict results

Figure 9 shows that when one lane (lane 1 or lane 4) is closed, the total conflict of VSL covering 1 km and 2 km is significantly better than the existing scenario. When two lanes are closed (lanes 1 and 2 or lanes 3 and 4), the results from existing scenarios and VSL scenarios are not significantly different. Furthermore, when comparing VSL cover 1 and 2 km itself there is no scenario that these two are significantly different.

Figures 8 and 9, when comparing VSL and existing conditions the results show that if only one lane (lane 1 or lane 4) is closed, both VSL1 and VSL2 show significantly better than the existing alternative. The VSL1 shows a percent improvement of up to a 45% reduction in average delay per vehicle and total conflict dropped by up to 42%. VSL2 improves average vehicle delay by up to 37% and total conflict improved by decreasing up to 34%. When two lanes are closed (lanes 1 and 2 or lanes 3 and 4), there was no significant difference in either average delay per vehicle and total conflict. When comparing VSL1 and VSL2 in every scenario there are no scenarios that average delay per vehicle and total conflict between VSL1 and VSL2 that are significantly different.

The sensitivity analysis, adjusting calibrated flow rates to 0.5, 0.75, 1, and 1.25, explores the impact of traffic fluctuations on outcomes. It reveals that while Variable Speed Limit (VSL) systems are beneficial during lighter traffic, their effectiveness diminishes, and they may exacerbate congestion during heavier periods. Moreover, the comparison between VSL1 and VSL2 performance under varying traffic conditions shows no significant difference.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The study investigated the usefulness of Variable Speed Limit (VSL) systems in enhancing traffic performance and safety by comparing scenarios with and without VSL installation. The findings revealed that VSL implementation can greatly improve traffic performance and safety in mild to moderate congestion scenarios. However, in circumstances of high congestion, where vehicle lines stretch to the VSL signs, the efficiency of VSL in controlling traffic flow and reducing congestion is reduced. This happens because the VSL system may not have sufficient time to assign and show new speed limits to vehicles before they enter the congested zone. As a result, whereas VSL devices provide significant benefits under normal traffic circumstances, their effectiveness may be limited during periods of high traffic congestion.

The comparison of VSL covering 1 km and 2 km produced interesting results. VSL with a 1 km spacing performed better in some variables in scenarios where it collected flow from the field. However, sensitivity analysis with calibrated flows of 0.5, 0.75, and 1.25 flow revealed no significant difference between VSL spacings of 1 km and 2 km. These findings indicate that the spacing of VSL may not have a significant influence on traffic performance under varying flow circumstances.

5.2 Limitations

This study focused on just one section of Thailand's Highway Number 7. Furthermore, the study only looked at one incident's location and only two different durations, offering insights into specific circumstances. The findings may not be relevant to other places or event durations, highlighting the need for caution in interpreting the findings. In the real situation, all the drivers may not follow the speed limit that shows on the sign which may make the results different.

5.3 Recommendations for Future Study

Future research should look at alternative Intelligent Transportation Systems (ITS) or adjust the placement of Variable Message Signs (VMS) to solve the problem of long vehicle lineups inhibiting the efficiency of VSL systems under heavy traffic, especially when lanes are closed. Investigating the impact of various techniques on traffic performance and safety might provide useful information for improving traffic management systems in densely populated areas.

Furthermore, future research should look at a variety of event locations, durations, and times of the day to have a more comprehensive understanding of VSL efficiency under various scenarios. This deeper investigation will improve the usefulness and adaptability of VSL systems for maintaining traffic flow and increasing road safety in many kinds of situations.

6. ACKNOWLEDGEMENTS

I am extremely thankful to Prof. Kunnawee Kanitpong, Assoc. Prof. Ampol Karoonsoontawon, and Asst. Prof. Punyaanek Srisurin for their important guidance and assistance during the completion of my thesis.

Furthermore, I am grateful to all of the professors and instructors at the Asian Institute of Technology for providing valuable knowledge and help during my academic career, which has considerably aided my growth and development.

Finally, I'd like to thank the Royal Thai Government for giving a major scholarship that enabled me to complete my master's degree. The financial assistance has been important in furthering my academic career, allowing me to devote myself entirely to my studies and thesis.

REFERENCES

Kurker, M., Fournier, C., Zhao, Q., Hakimi, S., Qi Y., Tang, S., Liu R., Juri, N. R., Duthie, J., & Machemehl, R. (2014). *Minimizing User Delay and Crash Potential through* . Austin: Center for Transportation Research The University of Texas at Austin .

Allaby, P., Hellinga, B., & Bullock, M. (2007). Variable Speed Limits: Safety and Operational Impacts of a Candidate Control Strategy for Freeway Applications. *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 8*, 4.

Boyles, S. D.,Walton, C. M., Duthie, J., Jafari, E., Jiang, N., Khani A., Li J., Osorio, J., Pandey, V., Rambha, T., & Yahia, C. (2018). *A Planning Tool for Active.* Austin: The University of Texas At Austin Center For Transportation Research.

Dowling, R., Skabardonis, A., & Alexiadis, V. . (2004). *Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software.* Washington, DC: U.S. Department of Transportation Federal Highway Administration.

Ellen F. Grumert, Andreas Tapani, and Xiaoliang Ma. (2018). Characteristics of variable speed limit. *European Transport Research Review*.

Farrag, S. G., Outay, F., Yasar A. U., Janssens D., Kochan B., & Jabeur N. (2021). Toward the improvement of traffic incident management systems using Car2X technologies. *Personal and Ubiquitous Computing*, 163-167.

Farrag, S., El-Hansali, M. Y., Yasar, A., & Shakshuki, E. M. (2020). Simulation-based Evaluation of Using Variable Speed Limit in Traffic Incidents. *Procedia Computer Science 175*, 340-348.

Guerra, A., Gadhiya, V., & Srisurin, P. (2022). CRASH PREDICTION ON ROAD SEGMENTS USING MACHINE LEARNING METHODS. *ASEAN Engineering Journal*, 12(3), 27-37.

Hegyi, A., Schutter, B. D., & Hellendoorn, J. (2005). Optimal Coordination of Variable Speed Limits to Suppress Shock Waves. *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 6*, 1.

Hollander, Y., & Liu, R. (2007). The principles of calibrating traffic microsimulation models. *Springer Science+Business Media, LLC.*

Hoogen, E. V. D., & Smulders, S. (1994). Control by variable speed signs: results of the Dutch experiment. *IET*, 145-149.

Karaer, A., Ulak M. B., Ozguven E. E., & Sando T. (2020). Reducing the Non-Recurrent Freeway Congestion with Detour Operations: Case Study in Florida. *Transportation*

Engineering 2.
Khondaker, B., & Kattan L. (2015). Variable speed limit: An overview. *Transportation Letters The International Journal of Transportation Research*.
Krisingson S. & Methawaranthon A. (2019). *Traffic Congestion Index and Density Analysis Report In 2018.* Bangkok: Department of Highway.
Lu, X. Y., & Shaladover, D. E. (2014). Review of Variable Speed Limits/Advisories – Theory, Algorithms and Practice. *93rd TRB Annual Meeting.* Washington, D.C.
MANUAL PTV Vissim 2023. (2023). Karlsruhe, Germany: PTV Planung Transport Verkehr GmbH.
Martin, P. T., Chaudhuri, P., Tasic, I., & Zlatkovic, M. (2011). *FREEWAY INCIDENTS: SIMULATION AND ANALYSIS.* Salt Lake City: Department of Transportation, University Transportation Centers Program.
Mfinanga, D. A. & Fungo, E. (2013). Impact of Incidents on Traffic Congestion in Dar es Salaam City. *International Journal of Transportation Science and Technology · vol. 2 · no. 2*, 95-107.
Sadat, M., & Celikoglu, H. B. . (2017). Simulation-based Variable Speed Limit Systems Modelling: An Overview and A Case Study on Istanbul Freeways. *Transportation Research Procedia 22*, 607-614.
Sangakong, J. (2567, January 24). *ข้อมูลอุบัติเหตุในเส้นทาง บนทางหลวงพิเศษระหว่างเมือง หมายเลข 7 และ หมายเลข 9.* Retrieved from Motorway: https://www.motorway.go.th/%e0%b8%82%e0%b9%89%e0%b8%ad%e0%b8%a1%e0%b8%b9%e0%b8%a5%e0%b8%ad%e0%b8%b8%e0%b8%9a%e0%b8%b1%e0%b8%95%e0%b8%b4%e0%b9%80%e0%b8%ab%e0%b8%95%e0%b8%b8%e0%b9%83%e0%b8%99%e0%b9%80%e0%b8%aa%e0%b9%89%e0%b8%99/
Soriguera, F., Martínez, I., Sala, M., & Menéndez, M. (2017). Effects of low speed limits on freeway traffic flow. *Transportation Research Part C 77*, 257-274.
Srisurin, P. (2020). *Microscopic Simulation Model for Mixed Traffic of Connected Automated Vehicles and Conventional Vehicles on Freeways .* Honolulu.
Srisurin, P., & Chalermpong, S. (2021). Analyzing Human, Roadway, Vehicular and Environmental Factors Contributing to Fatal Road Traffic Crashes in Thailand. *Engineering Journal*, 25(10), 27-38.
Xie, J., Yang, Z. , Liu, Y., Yang, X. B.,Teng, T., & Tham, C. K. (2022). Deep reinforcement learning for dynamic incident-responsive traffic information dissemination. *Transportation Research Part E*.
Zegeye, S. K.,Schutter, B. D., Hellendoorn, J., & Breunesse, E. A. (2011). Reduction of Area-Wide Emissions Using an Efficient Model-Based Traffic Control Strategy. *IEEE Forum on Integrated and Sustainable Transportation Systems*.

Optimizing Vehicle-to-Building Integration: A Sustainable, Adaptable Energy Management Framework

Yu-Shin Hu[a] and I-Yun Lisa Hsieh[a,b,*]

[a] Department of Civil Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei, 106319, Taiwan

[b] Department of Chemical Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei, 106319, Taiwan

[*] iyhsieh@ntu.edu.tw

ABSTRACT

Achieving nearly zero-energy buildings (NZEB) is a pivotal objective in sustainable development. Electric vehicles (EVs), with their inherent battery storage capacity, present a novel solution to energy management challenges through Vehicle-to-Building (V2B) strategies. These strategies harness the potential of EVs during extended parking times for peak energy shaving and reducing reliance on the electrical grid. Current research, however, often relies on idealized scenarios that neglect the real-world variability of EV user schedules, leading to participant inconvenience and limited application. Furthermore, many studies overlook significant factors such as associated costs and environmental implications, which are vital for realizing truly sustainable systems. Addressing these shortcomings, this study introduces a two-stage optimization framework designed to enhance the feasibility and sustainability of V2B operations. This framework is divided into a day-ahead optimization (DAO) phase and a real-time optimization (RTO) phase. The DAO phase employs multi-objective optimization, considering peak load shaving, cost minimization, and carbon reduction concurrently. To promote EV owner participation, we introduce strategic compensation for battery degradation and participation incentives. The RTO phase is developed to accommodate unpredictable EV usage patterns and to fine-tune day-ahead forecasts, thereby reducing user inconvenience. This dynamic system integrates grid and renewable energy sources, catering to the dual demands of building energy consumption and EV charging. Simulation outcomes indicate a substantial reduction in costs by 14.20% and a decrease in external carbon burdens by 14.37%. Analyses validate the model's resilience, displaying its capacity to adapt to uncertainties and variable conditions. This research underscores the role of V2B systems in advancing smart, sustainable energy management critical for the development of connected road networks, aligning with the push towards digitalization in modern transportation infrastructure.

Keywords: Vehicle-to-Building Integration, Nearly Zero-Energy Buildings, Smart Energy Management, Multi-Objective Optimization, Uncertainty Resilience

1. INTRODUCTION

The global pursuit of achieving net-zero emissions (NZE), in line with the Paris Climate Accords, underscores the urgent need for decarbonization across various sectors. The road transportation and building sectors represent major contributors to global carbon emissions, accounting for 13% (Albuquerque et al., 2020) and 25% (UN Environment Programme, 2022), respectively. Additionally, the transition to intermittent renewable energy sources has intensified the demand for flexibility services in power systems. In this context, the proliferation of electric vehicles (EVs) offers a sustainable alternative to traditional internal combustion engine vehicles, providing enhanced energy efficiency, lower operating costs, and reduced environmental impact (IEA, 2024; National Development Council, 2022).

Beyond these contributions, EVs hold further potential through Vehicle-to-Building (V2B) technology. This technology leverages the idle time of EVs (Ravi & Aziz, 2022), utilizing them as mobile storage units with bidirectional charging and discharging capabilities to mitigate fluctuations in power supply and demand within both buildings and power grids (Gschwendtner & Krauss, 2022). Given the inherent interconnection among the transportation, building, and energy sectors, V2B technology not only enables peak shaving and cost reduction for buildings, but also facilitates the integration of renewable energy into power systems, thereby supporting energy transitions and carbon emission reduction. Moreover, it enhances the reliability and efficiency of the overall system through frequency regulation, spinning reserve, and congestion management services (Banol Arias et al., 2019).

However, despite extensive research on V2B's potential in peak electricity balancing and renewable energy integration (Pearre & Ribberink, 2019), comprehensive strategies considering economic and environmental implications remain limited. Additionally, existing studies often prioritize building operators' interests, with limited attention to EV owners' perspectives and incentives (Luo, Cao, et al., 2024); prevalent idealized scenarios neglect prediction errors and variations in EV schedules, potentially undermining participant willingness (Chai et al., 2023). To bridge the identified research gaps, this study introduces a two-stage optimization framework that incorporates the considerations of both buildings and EV owners. This multifaceted approach is aimed at enhancing the adaptability and flexibility of V2B strategies in achieving a more feasible and sustainable energy system.

The rest of the paper is structured as follows: Section 2 delineates the methodology, encompassing the system delineation, data collection, and the proposed two-stage optimization model. Section 3 demonstrates the designed model's effectiveness in handling multifaceted optimization and adapting to uncertainty. The last section concludes the principal findings and suggests possible future research directions to extend this work.

2. MATERIAL AND METHODOLOGY

This study introduces a two-stage planning method, incorporating day-ahead (DAO) and real-time optimization (RTO), to address uncertainties in V2B strategy implementation within real-world applications. Figure 1 presents the research framework, showing the inputs, outputs, and analyses of the two-stage process.

In the day-ahead (DA) stage, the multi-objective optimization plan is formulated based on day-ahead prediction (DAP) data, aiming to anticipate and prepare for expected conditions. Subsequently, in the real-time (RT) stage, adjustments to the DA plans are made by considering real-time prediction (RTP) data. This approach enhances system flexibility and adaptability by allowing advance planning and responsive adjustments to RT conditions, effectively mitigating uncertainties and increasing feasibility. Additionally, this study encompasses a multidimensional scope by simultaneously addressing peak shaving, cost reduction, and carbon emissions mitigation, thereby enhancing overall energy sustainability.

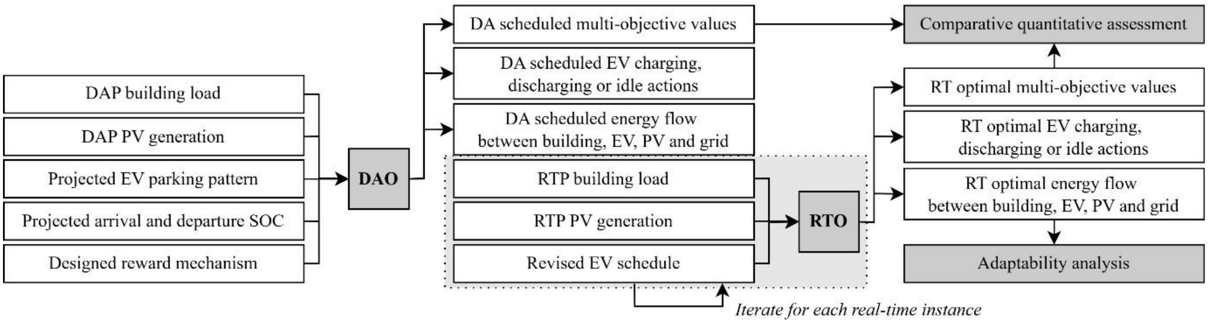


Figure 1 Proposed two-stage optimization model framework.

2.1 System and data

The research subject's system focuses on a compact campus system, which comprises a building (abbreviated as B), rooftop solar panels (R for renewable energy), electric vehicles (V), and the power grid system (G). The energy flow within the system is as follows: The grid powers the building (G2B) and electric vehicles (G2V), while solar panels support the building (R2B) and electric vehicles (R2V). Additionally, vehicles supply electricity to the building (V2B) when this mechanism is considered.

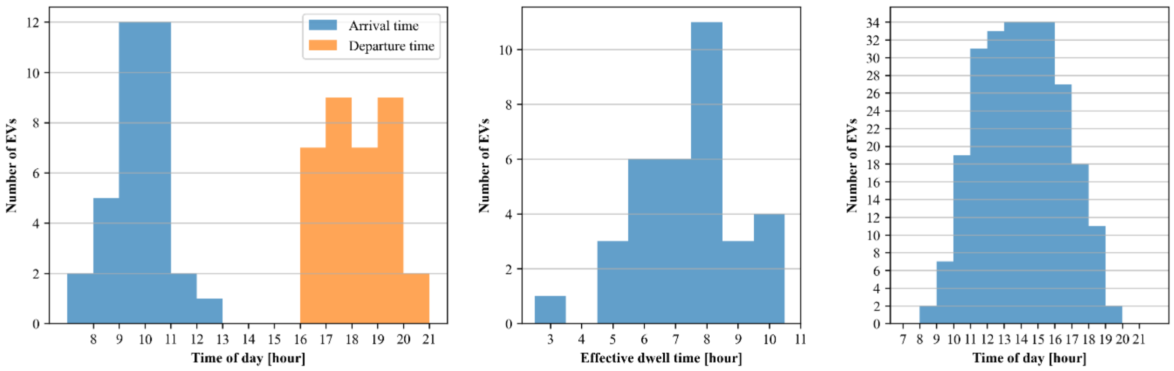
2.1.1 Building

In this study, we conduct a case study on the Civil Engineering Research Building at National Taiwan University in Taiwan. The building has a roof area of 322.14 m² available for solar panels and includes 34 underground parking spaces. Hourly building energy consumption data, measured in kilowatt-hours (kWh), is sourced from the NTU Digital Electricity Meter Monitoring System (National Taiwan University, 2024).

2.1.2 Electric vehicles

Assuming 34 EVs will park at the building, the EV parking process follows a random pattern, with variables constrained to a specified range by a truncated normal distribution (Ahmadi et al., 2020). Equation (1) gives the probability density function (PDF) for random variables x, where φ and ϕ represent PDF and cumulative distribution function (CDF) of standard normal distribution function, a and b are the upper and lower bounds for x, and μ and σ denote mean and standard deviation of x, respectively. EV arrival times span from 07:00 to 13:00 with an average of 10:00 and a 1-hour standard deviation, while departures occur between 16:00 to 22:00, averaging 18:00 with a 1.5-hour standard deviation. The binary availability status is set to 1 when the vehicle is at the building and 0 otherwise. Figure 2 illustrates the simulated distributions of parking patterns for a weekday.

$$f(x;\mu,\sigma,a,b) = \frac{1}{\sigma}\frac{\varphi(\frac{x-\mu}{\sigma})}{\phi\left(\frac{b-\mu}{\sigma}\right)-\phi(\frac{a-\mu}{\sigma})} \tag{1}$$



(a) **(b)** **(c)**

Figure 2 Simulated EV parking patterns for a weekday: (a) Distribution of arrival and departure patterns; (b) Distribution of effective dwell times; (c) Distribution of the parked and effective numbers.

The arrival state of charge (SOC) is generated with a uniform distribution between 30% and 60%, while the required departure SOC is between 65% and 90%, ensuring the required SOC is higher than the arrival SOC. The SOC is maintained between 20% and 80% to safeguard the health of lithium-ion batteries during high-power charging (Wang et al., 2023). Furthermore, this study references electric vehicle models in the Chinese market and adopts a battery storage capacity of 67.5 kWh (Luo, Yuan, et al., 2024). EV battery life reaches 2000 to 3000 cycles to reflect a margin of safety for real-world statistical distributions (Thunder Said Energy, 2024), with an average of 2500 cycles considered. Also, the global average price of lithium-ion battery packs are expected to decrease from 156 to 80 USD/kWh from 2022 to 2026, with this study employing an average of 118 USD/kWh, equivalent to 3797 NTD/kWh (Goldman Sachs, 2024). The charging piles are uniformly configured with a maximum charging and discharging rate of 7 kW, consistent with Level 2 charging station specifications (Lo et al., 2023). The charging and discharging efficiency are established at 90% considering circuit losses (Liu et al., 2024). Table 1 presents the parameters for EVs and charging piles.

Table 1 Parameters related to EVs and charging piles.

Parameter	Notation	Value and source
Arriving and departing time of vthEV (h)	t_v^a, t_v^d	Simulated with a truncated normal distribution (Ahmadi et al., 2020)
Availability status for vth EV at timeslot t	$a_{v,t}$	
Arrival and departure SOC for vth EV	SOC_v^a, SOC_v^d	Simulated with a uniform distribution
Battery energy capacity (kWh)	π^{Cap}	67.5 (Luo, Yuan, et al., 2024)
Battery lifetime in terms of cycle life (cycles)	π^{CL}	2500 (Thunder Said Energy, 2024)
Battery cost per kilowatt-hour (NTD/kWh)	c^{Batt}	3797 (Goldman Sachs, 2024)
Minimum and maximum SOC for battery health; Depth of discharge	SOC^{min}, SOC^{max}; π^{DOD}	20%, 80% (Wang et al., 2023); 60% (PowerTech, 2019)

Maximum charging and discharging rate (kW)	$r^{ch,max}, r^{dch,max}$	7 (Lo et al., 2023)
Charging and discharging efficiency	η^{ch}, η^{dch}	90% (Liu et al., 2024)

2.1.3 Solar panels

The solar panels on the target building's rooftop have dimensions of 1.002 meters by 1.667 meters (United Renewable Energy, 2023). Given the available rooftop area, an estimated 192 panels can be accommodated. Hourly solar energy data is collected. Figure 3 illustrates the solar power generation and building load, for a specific weekday in both summer and non-summer months.

2.1.4 Grid system

The time-of-use (TOU) electricity prices for the building and EV are sourced from Taiwan Power Company (Taiwan Power Company (Taipower), 2023b, 2023a), in New Taiwan Dollars per kilowatt-hour (NTD/kWh). These rates vary by season, day of the week, and time of day, with the highest prices charged during peak hours on weekdays in the summer.

Additionally, the hourly greenhouse gas emissions intensity (GEI) of the grid in Taiwan is quantified using the electricity emission intensity evaluation framework (Tseng & Hsieh, 2023), measured in kilograms of carbon dioxide equivalent per kilowatt-hour (kgCO_2eq/kWh). To assess the environmental impact of grid electricity, the greenhouse gas (GHG) emissions of grid load are calculated by multiplying the GEI by the energy consumed.

Figure 4 illustrates the TOU electricity prices for buildings and EVs, along with the GEI from the grid, for a specific weekday in both summer and non-summer months. It indicates elevated electricity prices during peak hours on summer weekdays, with staggered peak usage periods for the building and EVs. Due to increased renewable energy generation, the GEI is lower at midday in summer. Table 2 summarizes the parameters relevant to the building, solar power generation, and the grid system.

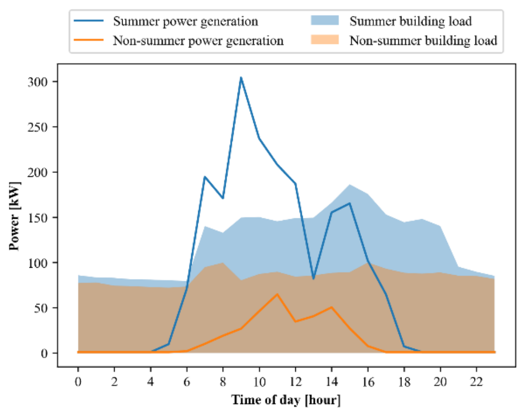


Figure 3 Solar generation and building load on a specific weekday in summer and non-summer months.

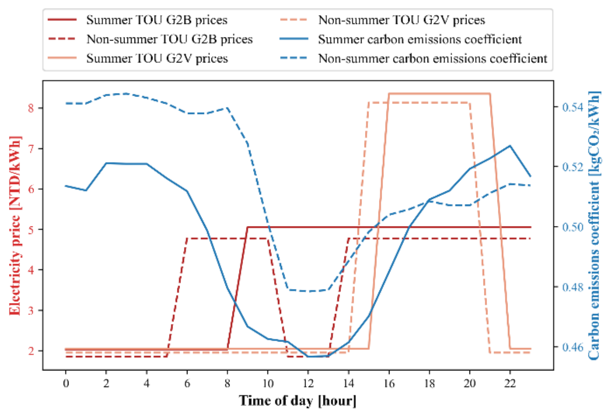


Figure 4 Grid TOU electricity prices and GEI on a specific weekday in summer and non-summer months.

Table 2 Parameters related to building, solar power generation, and grid system.

Parameter	Notation	Source
Energy demand for building at timeslot t (kWh)	D_t^B	Taipower
Solar power generation (supply) at timeslot t (kWh)	S_t^R	Own collected data
TOU electricity prices for building and EV at timeslot t (NTD/kWh)	c_t^{G2B}, c_t^{G2V}	Taipower
Carbon emissions intensity of the grid ($kgCO_2eq/kWh$)	π_t^{Cn}	(Tseng & Hsieh, 2023)

2.2 Two-stage optimization model

The two-stage optimization model comprises DAO and RTO. It assumes no detailed consideration of electric grid intricacies, transmission capacity constraints, or losses, given the centralized setup of charging stations directly connected to the high-voltage transmission grid. Operating on an hourly basis, the model assumes consistent charging or discharging rates throughout each interval. For model validation, we employ the Gurobi Optimizer, a commonly used solver for linear, integer, and mixed-integer programming, which aligns with our problem requirements.

2.2.1 Day-ahead optimization (DAO)

DAO aims to utilize DA forecasting for generate a 24-hour optimized schema, grounded in holistic considerations. The decision variables encompass the hourly energy levels and states (charging, discharging, or idle) of EVs, along with the intra-system energy flow, as outlined in Table 3.

Table 3 Decision variables within DAO.

Decision variables	Notation
SOC value for vth EV at time point t	$SOC_{v,t}$
Charging, discharging and idle state for vth EV at timeslot t	$x_{v,t}^{ch}, x_{v,t}^{dch}, x_{v,t}^{idle}$
G2V, R2V, V2B, G2B, R2B energy flow at timeslot t (kWh), with lowercase indicating energy flow to vth EV	$EF_t^{G2V}, EF_t^{R2V}, EF_t^{V2B}, EF_t^{G2B}, EF_t^{R2B}, ef_{v,t}^{G2V}, ef_{v,t}^{R2V}, ef_{v,t}^{V2B}$

Moreover, to enhance EV owner participation and the strategy's feasibility, it is imperative to provide incentives without imposing additional charging costs. Consequently, reward mechanism is designed to operate under V2B conditions based on the duration of stay, with pertinent parameters and calculation formulas detailed in Equation (2) and Table 4.

$$o_{v,t}^{Rwd} = a_{v,t} \cdot \pi^{Rwd} \cdot \frac{\sum_{t_1 \in T^{excl}}\left(\left(EF_{t_1}^{G2B} - {}^{*}EF_{t_1}^{G2B}\right)\cdot c_{t_1}^{G2B} + \left(EF_{t_1}^{V2B} - {}^{*}EF_{t_1}^{V2B}\right)\cdot c_{t_1}^{V2B}\right)}{\sum_{v_1 \in V}\left(t_{v_1}^{d} - t_{v_1}^{a}\right)} \tag{2}$$

Table 4 Parameters related to the designed reward mechanism.

Parameter	Notation	Value
Reward allocation ratio for EV engaged in V2B mechanism	π_t^{Rwd}	50%

In this phase, optimal solutions are identified from a building perspective using multi-objective optimization. It enables comprehensive considerations across various dimensions, incorporating peak shaving, cost efficiency, and environmental sustainability aspects. For addressing the challenge of multi-objective optimization, the weighted sum method is utilized. We minimize the objective function represented by Equation (3). Objectives' details and mathematical expressions can be found in Table 5, with lowercase indicating objectives at timeslot t. Here, we posit a consistent effect of unit electricity on battery degradation.

$$\min[w^{Pk} \cdot (O^{Pk}) + w^{Cts} \cdot (O^{Elec} + O^{Deg} + O^{Rwd}) + w^{Env} \cdot (O^{Cn} + O^{Wst})] \tag{3}$$

Table 5 Objectives within DAO.

Aspect	Objective	Notation	Calculation
Peak shaving	Grid loads	O^{Pk}, o_t^{Pk}	$o_t^{Pk} = EF_t^{G2B} + EF^{G2V}$
Cost Efficiency	Building electricity costs	O^{Elec}, o_t^{Elec}	$o_t^{Elec} = EF_t^{G2B} \cdot c_t^{G2B} + EF_t^{G2V} \cdot c_t^{G2V}$
	Battery degradation costs	$O^{Deg}, o_{v,t}^{Deg}$	$o_{v,t}^{Deg} = -c^{Batt} \cdot \frac{\pi}{\square}(\pi^{CL} \cdot \pi^{DOD}) \left(SOC_{v,t+1} - SOC_{v,t}\right) \cdot \left(x_{v,t}^{ch} - x_{v,t}^{dch}\right)$
	Rewards paid to EV owners	$O^{Rwd}, o_{v,t}^{Rwd}$	Design-specified
Environmental Sustainability	Carbon emissions	O^{Cn}, o_t^{Cn}	$o_t^{Cn} = (EF_t^{G2B} + EF_t^{G2V}) \cdot \pi_t^{Cn}$
	Unutilized solar power	O^{Wst}, o_t^{Wst}	$o_t^{Wst} = S_t^R - EF_t^{R2V} - EF_t^{R2B}$

Equation (4) to (13) delineates constraints for decision variables and their relationships. The constraints on energy flow include lower bounds and energy balance within the system components (building, PV, and EVs), as described in Equation (4) to (8).

$$D_t^B = EF_t^{V2B} + EF_t^{G2B} + EF_t^{R2B} \tag{4}$$

$$S_t^R - EF_t^{R2V} - EF_t^{R2B} \geq 0 \tag{5}$$

$$\left(SOC_{v,t+1} - SOC_{v,t}\right) \cdot x_{v,t}^{ch} \cdot \frac{\pi^{Cap}}{\eta^{ch}} = ef_{v,t}^{G2V} + ef_{v,t}^{R2V} \tag{6}$$

$$\left(SOC_{v,t+1} - SOC_{v,t}\right) \cdot x_{v,t}^{dch} \cdot \pi^{Cap} \cdot \eta^{dch} = -ef_{v,t}^{V2B} \tag{7}$$

$$EF_t^{G2V}, EF_t^{R2V}, EF_t^{V2B}, EF_t^{G2B}, EF_t^{R2B}, ef_{v,t}^{G2V}, ef_{v,t}^{R2V}, ef_{v,t}^{V2B} \geq 0 \tag{8}$$

EV-related constraints include SOC requirements, SOC and rate range limitations, along with state, availability, and energy flow relationships, presented in Equations (9) to (13).

$$SOC_{v,t_v^a} = SOC_v^a \wedge SOC_{v,t_v^d} = SOC_v^d \tag{9}$$

$$SOC_{max} \geq SOC_{v,t} \geq SOC_{min} \tag{10}$$

$$\begin{cases} SOC_{v,t+1} - SOC_{v,t} = 0, & t \geq t_v^d \ or \ t < t_v^a \\ \frac{r^{ch,max}}{\pi^{Cap}} \cdot \eta^{ch} \geq SOC_{v,t+1} - SOC_{v,t} \geq \frac{r^{dch,max}}{\pi^{Cap} \cdot \eta^{dch}}, & \text{elsewhere} \end{cases} \tag{11}$$

$$x_{v,t}^{ch} + x_{v,t}^{dch} + x_{v,t}^{idle} \leq a_{v,t} \tag{12}$$

$$\begin{cases} ef_{v,t}^{G2V} + ef_{v,t}^{R2V} > 0 \text{ if } x_{v,t}^{ch} = 1 \\ ef_{v,t}^{G2V} + ef_{v,t}^{R2V} < 0 \text{ if } x_{v,t}^{dch} = 1 \end{cases} \tag{13}$$

Furthermore, under V2B conditions, rewards should outweigh battery degradation and charging fees for EV owners, thereby encouraging their participation, as Equation (14).

$$O^{Rwd} \geq O^{G2V} + O^{Deg} \tag{14}$$

2.2.2 Real-time optimization (RTO)

The RTO endeavors to enhance system flexibility and cater to the requirements of EV owners. It utilizes RT and RTP data to refine the DA plan, adapting to parameter uncertainties, including forecast data and the demand for unplanned trips. Over the course of a day, RTO iterates 24 times to generate optimized patterns reflecting actual conditions. Each iteration exclusively accounts for parameters at the current and subsequent time points.

The DAO-planned schedule is incorporated into the parameters, with DAP and RTP data distinguished by the superscripts $\hat{X}$ and $\tilde{X}$, respectively. Additionally, the concept of 'energy anxiety' is introduced as a binary parameter to represent unplanned trips by EV owners, with a value of 1 indicating that energy anxiety is activated and there is an urgent need to leave. Table 6 details the parameters utilized in the RTO process.

Table 6 Parameters within RTO.

Parameter	Notation
DAP and RTP energy demand for building for the ongoing timeslot	$\widehat{D}^B, \widetilde{D}^B$
DAP and RTP solar power generation for the ongoing timeslot	$\hat{S}^R, \tilde{S}^R$
DAO SOC value for vth EV at this and next time point	$\widehat{SOC}_v^{next}$
True SOC value for vth EV at this time point	SOC_v^{now}
Binary energy anxiety for vth EV leaving earlier than planned and as quickly as possible at this time point	λ_v

The decision variables comprise intra-system energy flow at the current timeslot and the resulting SOC values of EVs at the subsequent time point, elaborated in Table 7.

Table 7 Decision variables within RTO.

Decision variable	Notation
SOC value for vth EV at next time point	SOC_v^{next}
Charging, discharging and idle state for vth EV at next time point	$x_v^{ch}, x_v^{dch}, x_v^{idle}$
G2V, R2V, V2B, G2B, R2B energy flow at this timeslot (kWh), with lowercase indicating energy flow to vth EV	$EF^{G2V}, EF^{R2V}, EF^{V2B}, EF^{G2B}, EF^{R2B}$, $ef_v^{G2V}, ef_v^{R2V}, ef_v^{V2B}$

In RT stage, the optimization objective is to address forecast errors and trip demand changes through slight adjustments, thereby avoiding significant alterations. Detailed formulations are provided in Equation (15) and Table 8.

$$\min\left[w^{Diff} \cdot \left(O^{Diff}\right) + w^{EarLv} \cdot \left(O^{EarLv}\right)\right] \tag{15}$$

Table 8 Objectives within RTO.

Purpose	Objective	Notation	Calculation
Ensure adjustments to the schedule are minimal	Difference between DA and RT planned schedules	O^{Diff}, o_v^{Diff}	$o_v^{Diff} = \left\|SOC_v^{next} - \widehat{SOC}_v^{next}\right\|$
Accommodate early departure needs	EV energy deviation during anxiety activation	O^{EarLv}, o_t^{EarLv}	$SOC_v^{next} \geq \lambda_v \cdot \left(\widehat{SOC}_v^{next} + o_v^{EarLv}\right)$

Model constraints are outlined in Equation (16) to (24). Equation (16) to (20) describe the constraints on RTP energy flow for system components, namely building, PV, and EVs.

$$\widetilde{D}^B = EF^{V2B} + EF^{G2B} + EF^{R2B} \tag{16}$$

$$\tilde{S}^R - EF^{R2V} - EF^{R2B} \geq 0 \tag{17}$$

$$\left(SOC_v^{next} - SOC_v^{now}\right) \cdot x_v^{ch} \cdot \frac{\pi^{Cap}}{\eta^{ch}} = ef_v^{G2V} + ef_v^{R2V} \tag{18}$$

$$\left(SOC_v^{next} - SOC_v^{now}\right) \cdot x_v^{dch} \cdot \pi^{Cap} \cdot \eta^{dch} = -ef_v^{V2B} \tag{19}$$

$$EF^{G2V}, EF^{R2V}, EF^{V2B}, EF^{G2B}, EF^{R2B}, ef_v^{G2V}, ef_v^{R2V}, ef_v^{V2B} \geq 0 \tag{20}$$

Equation (21) to (24) provide constraints on the expected SOC at the next time point. Constraints ensure the attainment of the required SOC for EV departure. Additionally, if energy anxiety is triggered, the EV becomes unassignable upon reaching full charge.

$$SOC_v^{next} \geq SOC_v^d, \forall v \in V \text{ if at } t_v^d \quad (21)$$

$$SOC_{max} \geq SOC_v^{next} \geq SOC_{min} \quad (22)$$

$$\begin{cases} SOC_v^{next} - SOC_v^{now} = 0, & t \geq t_v^d \text{ or } t < t_v^a \\ \frac{r^{ch,max}}{\pi^{Cap}} \cdot \eta^{ch} \geq SOC_v^{next} - SOC_v^{now} \geq \frac{r^{dch,max}}{\pi^{Cap} \cdot \eta^{dch}}, & \text{elsewhere} \end{cases} \quad (23)$$

$$SOC_v^{next} = SOC_v^{now}, \text{if } \lambda_v = 1 \text{ and } SOC_v^{now} = SOC_v^d \quad (24)$$

3. RESULTS AND DISCUSSION

In the results section, we offer analyses and comparisons for the proposed mechanisms to evaluate the model's efficacy in multifaceted optimization and uncertainty adaptation.

3.1 Day-ahead optimization (DAO)

The analysis of DAO results encompasses a dissection of objectives and energy flow details, alongside a comparison between scenarios with and without the V2B mechanism.

3.1.1 V2B mechanism quantitative assessment

To conduct multi-objective optimization, we normalize each objective by its optimal value to eliminate unit-based bias, thereby ensuring that no objective is disproportionately emphasized or overlooked. Figure 5 illustrates the DAO objective values over time. The energy flow composition of the overall system and the energy flows of its individual components are depicted in Figures 6 and 7, respectively.

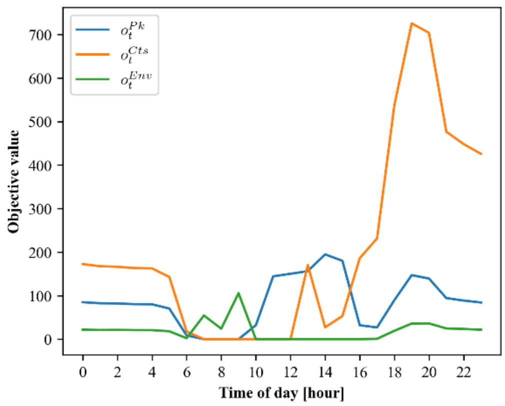


Figure 5 DAO objective values over time.

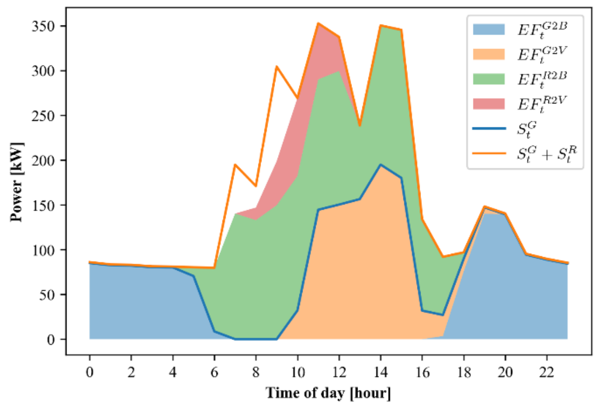
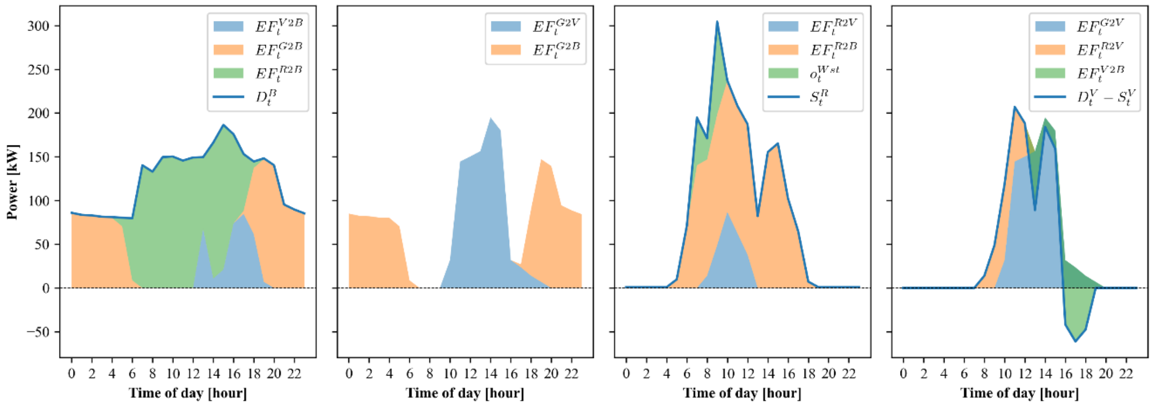


Figure 6 Energy flow composition of the system.

The environmental-related objective peaks around 8 AM due to wasted solar energy. This is attributed to the fact that most EVs have not yet arrived and building energy consumption is relatively low, rendering surplus energy storage unused, thereby leading to unavoidable wastage. The economic objective becomes more pronounced after 6 PM as buildings increase their power consumption, drawing from both EVs and the grid, to compensate for the reduced solar energy contribution. Additionally, the peak shaving objective, directly correlated with grid supply, peaks around midday and after 6 PM owing to heightened EV charging demand and insufficient solar energy, respectively.

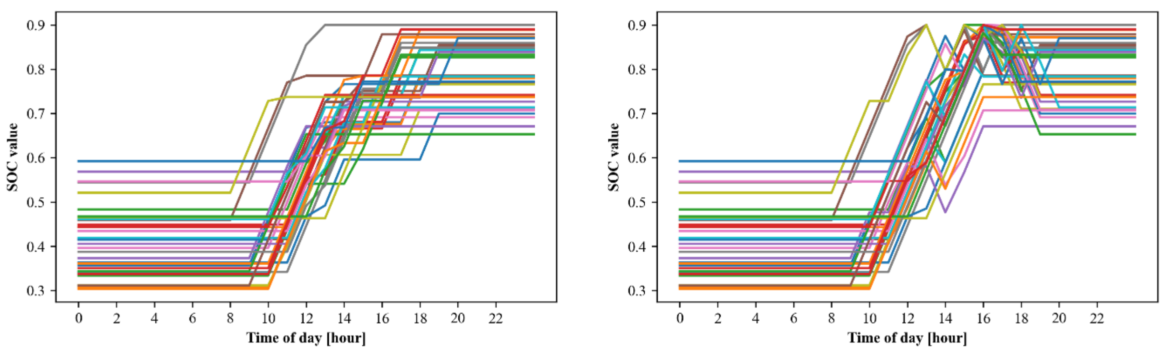


(a) **(b)** **(c)** **(d)**

Figure 7. Energy flow composition of system components: (a) Building energy flow composition; (b) Grid energy flow composition; (c) PV energy flow composition; (d) EV energy flow composition.

3.1.2 Comparative analysis of with and without V2B mechanism

Figure 8 depicts the temporal variations of SOC for each EV, reflecting their charging and discharging behaviors. It indicates that EVs tend to discharge energy to buildings in the afternoon in scenarios integrating V2B mechanisms, thus alleviating grid stress induced by buildings. By comparing objective values with and without the V2B mechanism, a reduction in daily costs of approximately 824 NTD is observed, reflecting a 14.20% decrease in cost-related objectives. Also, environmental-related objectives decrease by 14.37%, suggesting a shift in grid load towards periods of lower CEI and a reduction of wasted renewable energy.



(a) **(b)**

Figure 8 Temporal SOC variations of 34 EVs, each represented by a different colored line: (a) Without V2B mechanism; (b) With V2B mechanism.

3.2 Real-time optimization (RTO)

RTO results analysis entails examining the model's adaptability to uncertainty, including its ability to accommodate variations in EV schedules and prediction errors.

The RT phase provides EV owners with flexibility for early departures through the energy anxiety mechanism. Upon activation, the EV must rapidly charge to the required level to accommodate the owner's adjusted schedule. It also allows for adjustments in the model to account for prediction errors in DAP and RTP. Here, we assume a larger error in DAP, whereas RTP is deemed to be closer to the actual values.

Figure 9 illustrates the accelerated charging behavior of an EV upon the activation of the energy anxiety mechanism at 12 PM, with a DAP error of 10%. Compared to the DAO-scheduled results, this EV can depart at 2:30 PM, two hours earlier than the previously planned 4:30 PM departure. Additionally, RTO adjustments enable comprehensive fine-tuning of the system plan, including the charging and discharging behaviors of EVs, effectively managing the overall impacts.

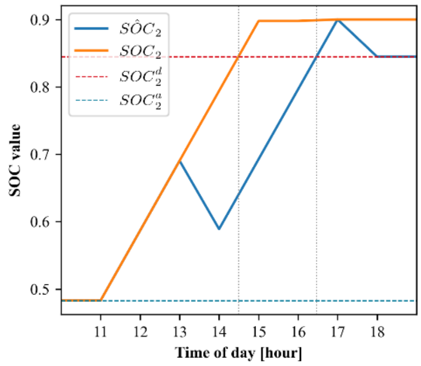


Figure 9 Temporal SOC variations of an EV (v=2) requiring early departure since 12 PM, with a DAP error of 10%.

4. CONCLUSIONS

This study introduces a two-stage model that integrates renewable energy sources to enhance the sustainability and feasibility of V2B strategies. The results highlight 14.20% and 14.37% reduction in costs and negative environmental impacts respectively, confirming the model's efficacy. Additionally, analyses validate its adaptability to uncertainty. Future research could delve deeper into refining the proposed framework and analysis, encompassing the model's adaptability to varying weather conditions, assessment of designed incentive mechanisms, integration of advanced forecasting techniques, and sensitivity analysis between scenarios. In conclusion, this study lays a solid foundation for advancing smart and sustainable energy solutions and addressing broader goals for balancing environmental and economic considerations, vital for the development of interconnected transportation networks.

5. ACKNOWLEDGMENTS

This study was supported by the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan (grant no. NTU- 113L893705).

6. REFERENCES

Ahmadi, S., Arabani, H. P., Haghighi, D. A., Guerrero, J. M., Ashgevari, Y., & Akbarimajd, A. (2020). Optimal use of vehicle-to-grid technology to modify the load profile of the distribution system. *Journal of Energy Storage*, *31*, 101627. https://doi.org/10.1016/j.est.2020.101627

Albuquerque, F. D. B., Maraqa, M. A., Chowdhury, R., Mauga, T., & Alzard, M. (2020). Greenhouse gas emissions associated with road transport projects: Current status, benchmarking, and assessment tools. *Transportation Research Procedia*, *48*, 2018–2030.

https://doi.org/10.1016/j.trpro.2020.08.261

Banol Arias, N., Hashemi, S., Andersen, P. B., Traeholt, C., & Romero, R. (2019). Distribution System Services Provided by Electric Vehicles: Recent Status, Challenges, and Future Prospects. *IEEE Transactions on Intelligent Transportation Systems*, *20*(12), 4277–4296. https://doi.org/10.1109/TITS.2018.2889439

Chai, Y. T., Che, H. S., Tan, C., Tan, W.-N., Yip, S.-C., & Gan, M.-T. (2023). A two-stage optimization method for Vehicle to Grid coordination considering building and Electric Vehicle user expectations. *International Journal of Electrical Power & Energy Systems*, *148*, 108984. https://doi.org/10.1016/j.ijepes.2023.108984

Goldman Sachs. (2024, February). *Lower battery prices are expected to eventually boost EV demand*. https://www.goldmansachs.com/insights/articles/even-as-ev-sales-slow-lower-battery-prices-expect

Gschwendtner, C., & Krauss, K. (2022). Coupling transport and electricity: How can vehicle-to-grid boost the attractiveness of carsharing? *Transportation Research Part D: Transport and Environment*, *106*, 103261. https://doi.org/10.1016/j.trd.2022.103261

IEA. (2024, April). *Global EV Outlook 2024*.

Liu, Q., Li, Y., Xu, T., Qian, F., Meng, H., Yao, Y., & Ruan, Y. (2024). Peak shaving potential and its economic feasibility analysis of V2B mode. *Journal of Building Engineering*, *90*, 109271. https://doi.org/10.1016/j.jobe.2024.109271

Lo, K.-Y., Yeoh, J. H., & Hsieh, I.-Y. L. (2023). Towards Nearly Zero-Energy Buildings: Smart Energy Management of Vehicle-to-Building (V2B) Strategy and Renewable Energy Sources. *Sustainable Cities and Society*, *99*, 104941. https://doi.org/10.1016/j.scs.2023.104941

Luo, J., Cao, X., & Yuan, Y. (2024). Comprehensive techno-economic performance assessment of PV-building-EV integrated energy system concerning V2B impacts on both building energy consumers and EV owners. *Journal of Building Engineering*, *87*, 109075. https://doi.org/10.1016/j.jobe.2024.109075

Luo, J., Yuan, Y., Joybari, M. M., & Cao, X. (2024). Development of a prediction-based scheduling control strategy with V2B mode for PV-building-EV integrated systems. *Renewable Energy*, *224*, 120237. https://doi.org/10.1016/j.renene.2024.120237

National Development Council. (2022). *Taiwan's Pathway to Net-Zero Emissions in 2050*.

National Taiwan University. (2024, May). *National Taiwan University Campus Digital Electricity Meter Monitoring System*. https://epower.ga.ntu.edu.tw/?fbclid=IwAR1_crXmTrEojnqGZCh6z2hesnkZ1Bsd7YBEnyAyzEyHOoIvr-xjA8sBAqo

Pearre, N. S., & Ribberink, H. (2019). Review of research on V2X technologies, strategies, and operations. *Renewable and Sustainable Energy Reviews*, *105*, 61–70. https://doi.org/10.1016/j.rser.2019.01.047

PowerTech. (2019, July 4). *New high-performance lithium LiFePO4 cells*. https://www.powertechsystems.eu/new-high-performance-lithium-lifepo4-cells/

Ravi, S. S., & Aziz, M. (2022). Utilization of Electric Vehicles for Vehicle-to-Grid Services: Progress and Perspectives. *Energies*, *15*(2), 589. https://doi.org/10.3390/en15020589

Taiwan Power Company (Taipower). (2023a). *Taiwan Power Company Electric Vehicle Charging and Switching Facility Electricity Prices (台灣電力公司電動車充換電設施電價)*.

Taiwan Power Company (Taipower). (2023b). *Taiwan Power Company Electricity Pricing Regulations (Implemented from April 1, 2023) (台灣電力公司電價表 (自112年4月1日起實施))*.

Thunder Said Energy. (2024). *Electric vehicle: Battery life?* https://thundersaidenergy.com/downloads/electric-vehicle-battery-life/

Tseng, W. C., & Hsieh, I. Y. L. (2023). Impacts of electric fleet charging patterns under different solar power penetration levels: Hourly grid variations and operating emissions. *Transportation Research Part D: Transport and Environment*, *122*. https://doi.org/10.1016/j.trd.2023.103848

UN Environment Programme. (2022). *2022 Global Status Report for Buildings and Construction.*

United Renewable Energy. (2023). *D1K_H3A monocrystalline solar module product specifications* *document*.

Wang, H., Wang, C., Jiang, C., Zhou, J., Liu, W., Ji, Z., Meng, G., & Zhao, F. (2023). High-power charging strategy within key SOC ranges based on heat generation of lithium-ion traction battery. *Journal of Energy Storage*, *72*, 108125. https://doi.org/10.1016/j.est.2023.108125

Multi-Objective Optimization in Energy Management using Vehicle-to-Building (V2B) Strategy and Energy Storage System

Jian Hern Yeoh[1], Kai-Yun Lo[2] and I-Yun Lisa Hsieh[3, *]

[1] Department of Civil Engineering, National Taiwan University, Taiwan, r12521626@ntu.edu.tw
[2] Department of Civil Engineering, National Taiwan University, Taiwan, r10521604@ntu.edu.tw
[3] Department of Civil Engineering, Department of Chemical Engineering, National Taiwan University, Taipei, Taiwan, *iyhsieh@ntu.edu.tw

ABSTRACT

The growing use of electric vehicles (EVs) has highlighted their influence on electrical grids, especially in the context of building energy management. While optimized EV charging schedules and Energy Storage Systems (ESS) have been proposed to alleviate grid load, the potential of Vehicle-to-Building (V2B) interactions has yet to be fully harnessed. V2B conceptualizes EVs as mobile energy storage units that can level peak demands and enhance grid stability. This study introduces a multi-objective optimization model focused on reducing peak loads, lowering energy costs, and minimizing emissions. We construct a simulated setup based on certain facilities at National Taiwan University, complete with four buildings, solar photovoltaics (PV), bidirectional EV chargers, and an extensive ESS. The impacts of these components are evaluated across four scenarios to assess the benefits of PV, V2B, and ESS both individually and in combination. Our findings highlight a 46.79% annual reduction in total grid energy consumption through the holistic adoption of these technologies. Additionally, substantial declines in key indicators are recorded: daily peak load decreased by 63.49%, electricity costs by 51.15%, and carbon emissions by 45.99%. This research contributes to the development of smarter energy solutions in transportation infrastructure, reflecting a move towards more intelligent and resource-efficient road networks.

Keywords: Vehicle-to-Building (V2B) Optimization, Peak Load Shaving, Energy Cost Reduction, Carbon Emissions Mitigation, Sustainable Energy Management.

1. INTRODUCTION

In recent years, governments around the world have shifted switching focus to transportation electrification to mitigate environmental impact such as reducing greenhouse gases emissions. Despite economic or political factors which caused the uneven distribution of global electric vehicles (EVs), the rapid growth is documented (Tan et al., 2023). The surge in EVs sales posed significant challenges such as load profile distortion and grid instability, which are caused by an increased in energy demand (Mahmud et al., 2023). In addition, the mass installation of EVs chargers within buildings has introduced significant challenges, yet presented opportunities for energy management within the building sector. According to (OECD, 2023), the building sector has been contributing one-third of global energy demand influenced by the increasing amount of electric appliances. This suggests that a smart charging system is crucial to address the downside of sustainable transportation.

Within the realm of smart energy management systems (SEMS), extensive research has been conducted on scheduling strategies for EVs and Energy Storage Systems (ESS) (Achour et al., 2021; Leone et al., 2022; Leonori et al., 2020; Z. Wang et al., 2020). Various optimization models and algorithms have been proposed, including rolling window approaches for EV charging optimization, strategies for ultra-fast charging, and predictive control methods aimed at reducing peak demand within specific environments.

Moreover, the integration of Vehicle-to-Building (V2B) strategies into energy infrastructure has garnered attention as a means to further optimize energy usage (Dagdougui et al., 2019; Odkhuu et al., 2018; Ouammi, 2021). These strategies focus on leveraging EVs not only for transportation but also as distributed energy resources capable of discharging energy back into buildings during peak demand periods. Incorporating renewable energy sources and predictive control strategies, such as those involving photovoltaic (PV) generation and ESS, can further enhance grid stability and reduce peak load while ensuring EV charging requirements are met.

The primary contribution of our study focuses on developing a day-level multi-objective optimization model, showcasing the potential benefits of the V2B strategy. This model integrates peak load management, cost savings, and carbon emissions reduction within a defined energy boundary, advancing SEMS design by aligning environmental and economic goals dynamically. We introduce a flexible algorithm adaptable to various energy situations, enabling resilient EMS responses to evolving energy demands, balancing environmental responsibility with cost-effectiveness. It enhances grid efficiency and promotes sustainability in energy management practices, serving as an essential decision-support tool.

2. MATERIAL AND METHODS

2.1 Research Framework

Our study focuses on a specific area within National Taiwan University (NTU) in Taipei, Taiwan, serving as a boundary for our case study. This zone features multiple buildings interconnected by a shared power line infrastructure. We have conceptualized a SEMS for this area, incorporating rooftop solar panels, a 1,500-kWh ESS, and bidirectional EV chargers in the parking lots. The batteries of EVs in the system are treated as portable energy storage system supported by 56 bidirectional charging stations.

Figure 1 shows the structure of our framework, beginning with the collection of data on the building initial electricity use, grid electricity cost and grid emissions. We also simulate models for the solar energy generation patterns and vehicle parking behavior. An optimization function is then designed to optimize three sub-objective functions, mainly minimizing peak load, electricity cost and carbon emissions. We can obtain the hourly optimal action for both V2B strategy and ESS smart charging from the solution of the function. This comprehensive approach quantifies both the environmental and economic impacts of our energy management strategies, ultimately supporting a smart, ecologically responsive urban energy ecosystem.

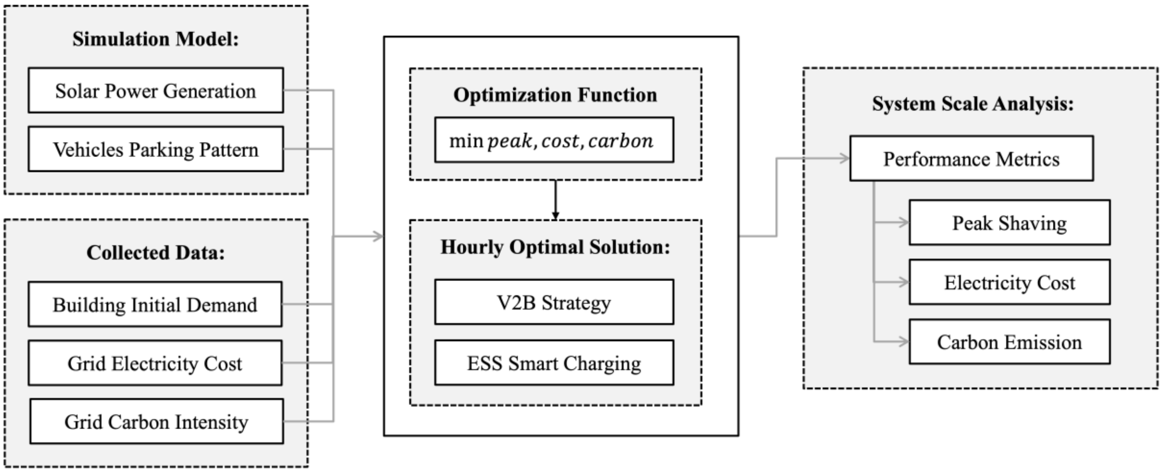


Figure 1 Study Framework

2.2 Simulation Model

Solar Power Generation

Solar energy has become an increasingly attractive alternative to traditional power sources. To capitalize on this, our study aims to use the available rooftop space across several buildings within the NTU campus for solar panel installation. Drawing on the work of (Wu et al., 2018), we adopt a formula to estimate the power generated by rooftop PV installations.

$$P_{PV} = \eta_{PV} \times S \times I \times \left(1 - 0.005 \times (T_a - 25)\right) \tag{1}$$

P_{PV} represents the output power of the solar panels in kilowatts (kW), η_{PV} denotes the conversion efficiency of PV cells set at 20% (Center for Sustainable Systems, University of Michigan., 2023), S is the available rooftop area for PV panel deployment in square meters (m²), I is the solar radiation in kilowatts per square meter (kW/m²), and T_a is the ambient temperature in degrees Celsius (°C). The weather data is obtained from a smart meter installed within NTU and the roof area is estimated from NTU digital map.

Vehicle Parking Pattern

Within our study area at NTU, 56 parking slots are equipped with bidirectional charger. This setup enhances energy management flexibility and underscores the role of EVs in the V2B strategy. Accurately modeling EV parking patterns is crucial for simulating their contribution to the V2B ecosystem.

We adapt a distribution model from (Lo et al., 2023) to simulate parking patterns, incorporating various vehicle models prevalent in Taiwan as of 2022. Key details such as battery capacity and energy efficiency are included to ensure accuracy. To capture the variability in EV parking behaviors, our model integrates multiple probability distributions, following the stochastic methodology recommended by (Thomas et al., 2018). The parameters defining these distributions are detailed in Table 1. Campus data indicates that a chi-squared distribution best models the skewed arrival times of vehicles, a method endorsed by (D. Wang et al., 2016) for home-based trips. The mean vehicle departure time follows a normal distribution centered at 6 pm with a standard deviation of 1.5 hours. Arrival and departure states of charge (SoC) are modeled using uniform distributions: $U\ (0.3, 0.75)$ for arrival SoC and $U\ (0.75, 0.9)$ for departure SoC.

Table 1 EVs Stochastic behavior patterns.

EVs Functions	Method	Parameters
Arrival time	Chi-squared distribution	$\chi^2\ (4)$ with *loc* = 6, *scale* = 0.5
Departure time	Normal Distribution	$N\ (18, 1.5^2)$
Arriving SOC	Uniform Distribution	$U\ (0.3, 0.75)$
Departing SOC	Uniform Distribution	$U\ (0.75, 0.9)$

(Thomas et al., 2018)

2.3 Data Collection

Building Initial Demand

The data for each building's energy usage were obtained from NTU's smart metering system. A total of 4 buildings within NTU campus are considered in this study, namely Agricultural Exhibition Hall, College of Art, Department of Civil Engineering, and Department of Chemical Engineering. Figure 2 shows the monthly overview for the energy within the system, i.e. total energy consumption and PV generation, along with the average monthly temperature in degrees Celsius (°C) on a second y-axis. This illustrates the parallel trends between energy consumption, PV generation, and temperature variations.

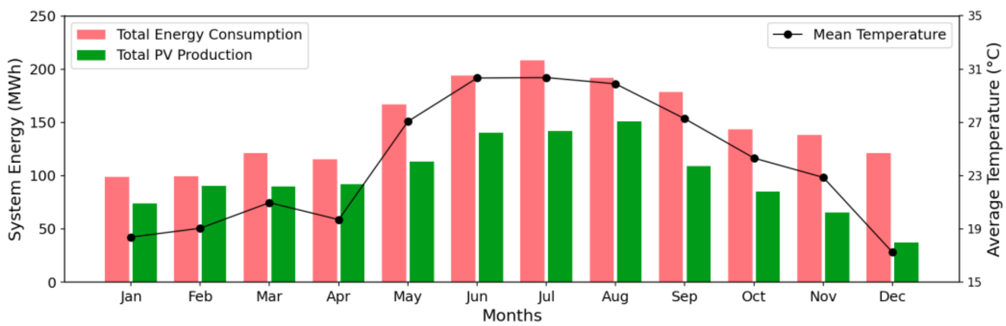


Figure 2 System Monthly Energy Overview

Grid Electricity Cost

In our study, we assumed that electricity obtained from the grid to fulfill system needs is supplied by Taiwan Power Company. Consequently, data for calculations, such as electricity rates and carbon intensity, are based on information from the Taipower website. For electricity cost calculations, we adapted the process to fit two stages. During the optimization process, we focused on scheduling the best charging slot, considering only the hourly electricity rate for 2020. Table 2 shows the breakdown of energy charges based on seasons, days of the week, and times of day within our research horizon.

Table 2 Energy Charges Rates Breakdown by Taipower Company.

Day	Period	Time	Cost (NTD/kWh)	
			Summer	Non-Summer
Monday To Friday	Peak Period	07:30~22:30	3.42	3.33
	Off-Peak Period	00:00~07:30 22:30~24:00	1.46	1.39
Saturday	Partial-Peak Period	07:30~22:30	2.14	2.06
	Off-Peak Period	00:00~07:30 22:30~24:00	1.46	1.39
Sunday & Off-Peak Day	Off-Peak Period	00:00~24:00	1.46	1.39

(Taipower Company, 2020) Note: Off-Peak days include national holidays in Taiwan.

Grid Carbon Intensity

Electricity carbon emissions are a major contributor to global emissions, especially in Taiwan, where fossil fuel power generation leads to significant CO_2 output and environmental degradation. The integration of EVs can increase building electricity demand, potentially negating their environmental benefits. To mitigate these impacts, it is crucial to quantify carbon emissions from electricity consumption. We calculate the hourly carbon intensity of Taiwan's electricity grid using (Tseng & Hsieh, 2023) methodology, measuring in $kgCO_2eq/kWh$. This approach, which considers consumption patterns and grid emissions, reveals a carbon intensity range of 0.3916 to 0.5861 $kgCO_2eq/kWh$, providing a clear picture of the environmental footprint of electricity use.

2.4 Proposed Multi-objective Optimization Function

Objective Functions

The roles of ESS and EVs are central to this study, aiming to determine the most effective hourly actions within specified constraints to alleviate grid pressure, reduce electricity costs, and lower carbon emissions from electricity generation. This strategy facilitates peak load reduction and off-peak energy storage, decreasing total electricity expenditures. Our model, using the *scipy.optimize* library and the basin hopping technique, searches for the global minimum with *scipy.optimize.minimize* as a local minimizer.

We developed three objective functions to assess daily energy performance, focusing on grid demand, electricity costs, and carbon emissions. The optimization goal is to minimize these functions, promoting balanced energy consumption. The peak score is calculated by squaring the hourly grid load, creating a quadratic relationship with demand. The cost score is derived by multiplying grid demand by the current electricity rate, following Taipower's tariff structure. The carbon score is obtained by multiplying grid demand by hourly carbon intensity. We normalize these scores using min-max scaling to ensure fair comparison of grid performance, cost efficiency, and carbon footprint.

$$\min F(\boldsymbol{x}) = \sum_{t=0}^{23} \gamma_{peak} \cdot F_{peak}(x_t) + \gamma_{price} \cdot F_{cost}(x_t) + \gamma_{carbon} \cdot F_{carbon}(x_t) \tag{2}$$

The objective function $F(\boldsymbol{x})$ is designed to minimize three key aspects of energy management: peak load reduction, electricity cost reduction, and carbon emissions reduction. In this function, γ_{peak}, γ_{price} and γ_{carbon} represents the weight assigned to peak load reduction, reducing electricity costs, and lowering carbon emissions respectively. The variable $\boldsymbol{x}$ denotes the optimal set of actions in kWh, which is a varying variable of the optimization function for the output. $\boldsymbol{x}$ is returned as a list with length 24 over a 24-hour period, capturing the timing and intensity of these actions to achieve the desired balance among the three objectives affected by factors such as such as arrival and departure times, charging rates, and energy demand.

$$F_i(x_t) = \frac{f_i(x_t) - \min(f_i)}{\max(f_i) - \min(f_i)} \quad for\ i \in \{peak,\ cost,\ carbon\} \tag{3}$$

$$f_{peak}(x_t) = (\ Grid_t\ + x_t\ )^2 \tag{4}$$

$$f_{cost}(x_t) = (\ Grid_t\ + x_t\ )\ \times\ cost_t \tag{5}$$

$$f_{carbon}(x_t) = (Grid_t\ + x_t)\ \times\ CI_t \tag{6}$$

$$\sum_{i}^{N} \gamma_i = 1 \quad for\ i \in \{peak,\ cost,\ carbon\} \tag{7}$$

In the model, f_{peak} is the objective function focused on minimizing peak grid demand by performing a squared on the total demand, f_{cost} targets the reduction of electricity costs, and f_{carbon} aims to decrease carbon emissions. The variable $Grid_t$ represents the grid demand at hour *t* before the current iteration of optimization is applied, while $cost_t$ denotes the electricity cost per kilowatt-hour at hour *t*. CI_t stands for the carbon intensity per kilowatt-hour at hour *t*. Variables including $Grid_t$, $cost_t$ and CI_t are external input based on historical data highlighted in Section 2.3.

Model Constraints

To promote the longevity of EV batteries and avoid frequent replacements, we set a SoC range for each EV with a lower threshold and an upper limit of 90%. The maximum power for EVs and ESS charging/discharging is set at 7kW and 500kW respectively. The constraints equations are written as eqs8-11. The equation to calculate the SoC is shown in eq12, which take into consideration of the charging efficiency of 90% (Schram et al., 2020).

$$30\% \leq SoC_{EV} \leq 90\% \tag{8}$$

$$20\% \leq SoC_{ESS} \leq 90\% \tag{9}$$

$$P_{max,EV} \leq 7kW \tag{10}$$

$$P_{max,ESS} \leq 500kW \tag{11}$$

$$SoC_t = \begin{cases} SoC_{t-1} + x_t \cdot \eta_{ch} & if\ x_t > 0 \\ SoC_{t-1} + x_t \cdot \dfrac{1}{\eta_{disch}} & if\ x_t < 0 \end{cases} \tag{12}$$

Key parameters include SoC_{EV} and SoC_{ESS}, which indicates the SoC for EVs and ESS. The maximum power $P_{max,EV}$ and $P_{max,ESS}$ draws the power limit for EVs and ESS. SoC_t is the SoC at time t, calculated by charging and discharging efficiencies, η_{ch} and η_{disch}.

EVs Scheduling Priority

To effectively schedule EVs based on their priorities, we have developed a priority equation. This scheduling approach considers the parking patterns of EVs to optimize slot usage for all vehicles. The priority score, representing the average energy required per hour for each vehicle, is calculated using eq13.

$$P_n = \frac{E_{EV\_req,n}}{t_{dep,n} - t_{arr,n}} \tag{13}$$

$E_{EV\_req,n}$ describes the energy required for EV_n charging in kWh, while $t_{dep,n}$ and $t_{arr,n}$ represents departure time and arrival time for EV_n.

2.5 Scenario Analysis

To evaluate the impacts of different energy management strategies, we constructed three distinct scenarios in addition to the baseline case, providing a framework for comparison as shown in Table 3.

The baseline scenario, referred to as *Scenario 0*, serves as our control group, where EVs are charged immediately upon arrival with no new installations, reflecting standard building demand. *Scenario 1* explores the use of solar panels alone, potentially leading to the underutilization of excess energy produced by the PV systems. *Scenario 2* introduces the V2B strategy, with EVs parked in slots equipped for bidirectional charging. *Scenario 3* is the most comprehensive, combining the V2B strategy with an ESS to maximize the use of surplus PV-generated energy.

Table 3 Definition of Scenarios for Comparative Analysis

Scenario	PV	V2B	ESS	Description
Scenario 0: Control	-	-	-	Standard operation: immediate EV charging upon arrival, no new installations.
Scenario 1: PV	✓	-	-	Utilizes only solar panels, with potential excess energy not fully utilized.
Scenario 2: PV + V2B	✓	✓	-	Implements V2B strategy with custom EV charging/discharging to optimize load and costs.
Scenario 3: PV + V2B + ESS	✓	✓	✓	Integrates ESS, combining V2B and ESS strategies for maximal energy utilization.

2.6 Optimization Algorithm

Algorithm 1 outlines the step-by-step algorithm used for optimizing the charging and discharging schedules of EVs and ESS within our integrated energy framework. The algorithm processes daily data for the entire year (2020), adapting the energy management system dynamically to the daily variations in energy consumption, PV generation, and EV parking patterns. The arrival, departure and stochastic characteristics of EVs are simulated as presented in Section 2.2.

Algorithm 1: Optimization of EV and ESS Charging/Discharging

Input: Initial System Energy Consumption
PV Power Generation
EVs Parking Pattern and Characteristics
Objective functions, Constraints, Charging parameters

Output: Charging/discharging schedules for EVs and ESS

```
FOR each day in year ( day ∈ {1,2, … … ,366} ) DO
    GET Energy Consumption, PV generation, EVs info for single day
    COMPUTE scheduling priorities for all vehicles
    FOR each vehicle after sorting, starting with highest P (Eq.13) DO
        IF required energy per hour > P_max,EV THEN
            Skip vehicle, not eligible for V2B
        ELSE
            COMPUTE optimization problem for vehicle
            GET optimal V2B charging schedule for vehicle
            UPDATE Energy consumption with EV charging
        ENDIF
    ENDFOR
```

13	**COMPUTE** reserve amount for excess RES generated
14	**COMPUTE** optimization problem for ESS
15	**GET** optimal charging/discharging schedule for ESS
16	**ENDFOR**

3. RESULTS AND DISCUSSIONS

3.1 Variability in Model Response

In this section, we visualize the model's response to each scenario with different settings as shown in Figure 3. The figure illustrates the optimization results for EVs and ESS in a chosen exemplary week, showcasing the flexibility and contribution towards energy flow within the system.

Exemplary day with limited solar generation

During low solar generation, e.g. on January 1st, EV charging is strategically scheduled to utilize available solar energy, flattening demand peaks. This contrasts with immediate charging, which further straining on the grid on a day with low renewable power. The ESS further enhances flexibility by charging during off-peak hours and providing power during peak demand, achieving a peak shaving, and optimizes electricity cost and carbon emissions.

Exemplary day with surplus solar generation

On January 2nd to 4th, a high solar generation with surplus solar energy is observed. This optimization strategy allows synchronization of EV discharge upon arrival, then charge during high solar generation hours. ESS can then be used to store excess energy for later use during lower generation periods.

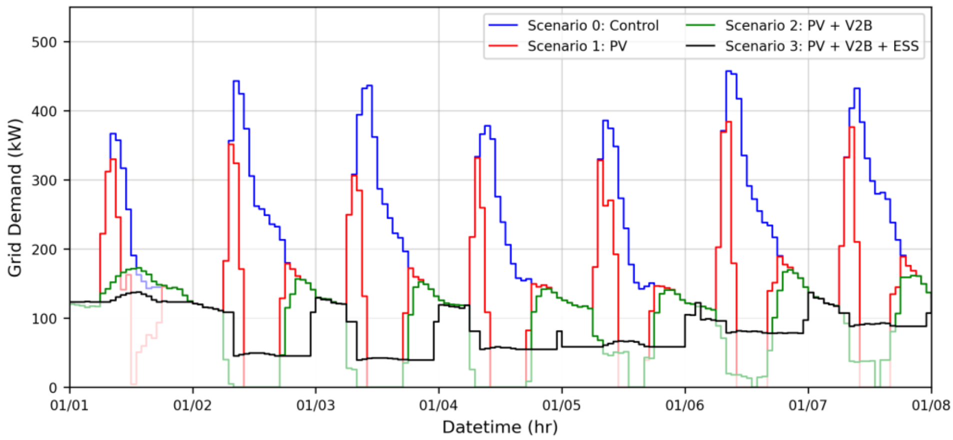


Figure 3 Hourly Energy profile of Exemplary Week

3.2 Performance Metrics across subobjectives

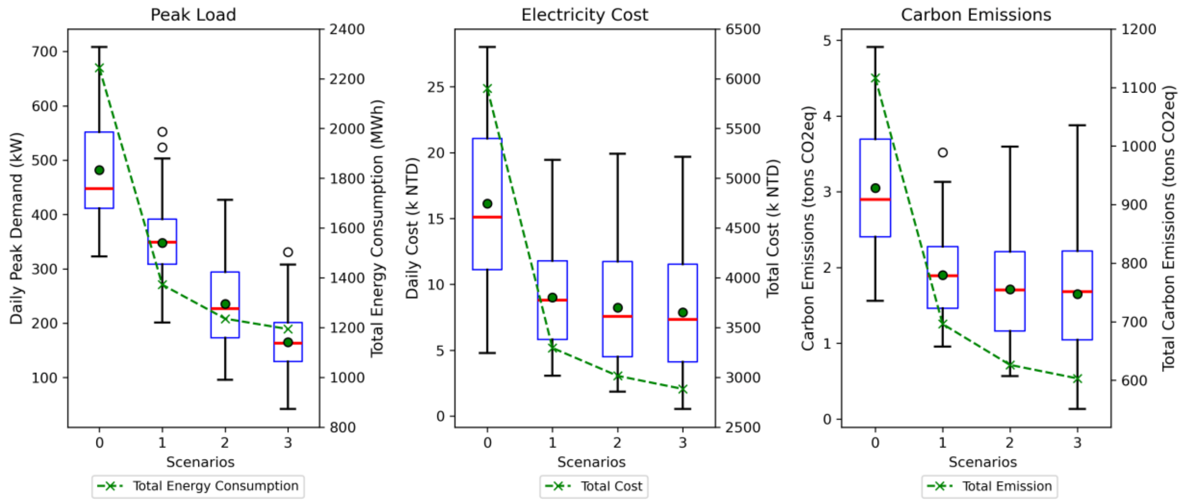
The model's performance is assessed based on its ability to minimize daily peak load, electricity costs, and carbon emissions, which aligns with our objective functions. Figure 4 presents a series of boxplots and accompanying statistical data comparing the performance metrics under various energy management scenarios. These scenarios illustrate the individual and combined effects of different strategies, highlighting their respective impacts on cost efficiency, demand management, and environmental sustainability. The visualization and statistics together convey the efficacy of each scenario in optimizing energy use.

Peak Load Reduction

The daily peak demand, a measure of the system's efficiency in smoothing electricity demand spikes, shows a significant decline across scenarios. Figure 4(a) displays the median peak load, which drops from 447 kW in *Scenario 0: Control* to 163 kW in the most integrated *Scenario 3: PV+V2B+ESS*, marking a 63.49% reduction. The total energy consumption marks a 46.79% drop from 2,243MWh to 1,194kWh. This substantial decrease underscores the robustness of the optimization model in managing peak loads effectively.

Electricity Cost Reduction

The electricity costs reveal a reduction in both median daily cost and total electricity cost, indicating effective cost optimization. The median cost drops from 15.1 kNTD to 7.32 kNTD from *Scenario 0: Control* to *Scenario 3: PV+V2B+ESS*, representing a 51.51% reduction. Similarly, the total electricity cost sees a 51.15% decrease in Scenario 3. Figure 4(b) highlights this downward trend, emphasizing the model's adeptness at cost management. However, a slight increase in upper cost ranges is noted in Scenario 2, suggesting operational challenges in aligning vehicle charging schedules with optimal electricity pricing within this scenario.



(a) (b) (c)

Figure 4 Comparative Analysis of Performance Metrics (a) Peak Load Reduction (b) Electricity Cost Reduction (c) Carbon Emissions Reduction

(NOTE: The blue box shows the interquartile range (IQR) and black whiskers extend to the minimum and maximum values within 1.5 times of the IQR. The red line represents the median, and the green dot indicates the mean.)

Carbon Emissions Reduction

Carbon emissions, reflecting the environmental impact, also show a notable decline across scenarios. Scenario 3 achieves a 41.96% reduction in median daily emissions and a 45.99% reduction in total emissions compared to Scenario 0. Figure 4(c) illustrates this trend, demonstrating the model's effectiveness in reducing emissions associated with electricity generation. However, there is an increasing spread in distribution for carbon emissions as scenarios transition, indicating the model's dynamic strategy to balance immediate grid demands with long-term emission reductions, showcasing its flexibility in achieving sustainable energy management outcomes.

4. CONCLUSIONS

Along with the global shift towards EVs adoption and net-zero emissions, this study offers an V2B strategy optimization modeling assessment, examining the interplay between mobility trends and energy infrastructure. Focusing on the building sector's significant energy consumption and carbon footprint, we introduce a multi-objective optimization model adaptable to different requirements, bridging gaps in opportunities for the rising demand for EV charging.

The model's flexibility is evident through adjustable weights, allowing prioritization of specific objectives. A convincing result is obtained with equalized weight shown in Section 3. For peak reduction, V2B and ESS strategies lead to a 63.49% decrease in median peak power requirements, validating the model's capacity to alleviate grid stress. Electricity cost and carbon emission shows a decrease median value, with a slight rise at higher percentiles, indicating the model's strategic adaptability in balancing grid stability and environmental objectives, highlighting its ability to manage complex energy trade-offs.

5. ACKNOWLEDGMENTS

This study was supported by the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan (grant no. NTU- 113L893705).

6. REFERENCES

Achour, Y., Ouammi, A., & Zejli, D. (2021). Model Predictive Control Based Demand Response Scheme for Peak Demand Reduction in a Smart Campus Integrated Microgrid. *IEEE Access*, *9*, 162765–162778. https://doi.org/10.1109/ACCESS.2021.3132895

Center for Sustainable Systems, University of Michigan. (2023). *Photovoltaic Energy Factsheet* (Pub. No. CSS07-08.).

Dagdougui, H., Ouammi, A., & Dessaint, L. A. (2019). Peak Load Reduction in a Smart Building Integrating Microgrid and V2B-Based Demand Response Scheme. *IEEE Systems Journal*, *13*(3), 3274–3282. https://doi.org/10.1109/JSYST.2018.2880864

Leone, C., Longo, M., Fernandez-Ramirez, L. M., & Garcia-Trivino, P. (2022). Multi-Objective Optimization of PV and Energy Storage Systems for Ultra-Fast Charging Stations. *IEEE Access*, *10*, 14208–14224. https://doi.org/10.1109/ACCESS.2022.3147672

Leonori, S., Martino, A., Frattale Mascioli, F. M., & Rizzi, A. (2020). Microgrid Energy Management Systems Design by Computational Intelligence Techniques. *Applied Energy*, *277*, 115524. https://doi.org/10.1016/j.apenergy.2020.115524

Lo, K.-Y., Yeoh, J. H., & Hsieh, I.-Y. L. (2023). Towards Nearly Zero-Energy Buildings: Smart Energy Management of Vehicle-to-Building (V2B) Strategy and Renewable Energy Sources. *Sustainable Cities and Society*, *99*, 104941. https://doi.org/10.1016/j.scs.2023.104941

Mahmud, I., Medha, M. B., & Hasanuzzaman, M. (2023). Global challenges of electric vehicle charging systems and its future prospects: A review. *Research in Transportation Business & Management*, *49*, 101011. https://doi.org/10.1016/j.rtbm.2023.101011

Odkhuu, N., Lee, K.-B., A. Ahmed, M., & Kim, Y.-C. (2018). Optimal Energy Management of V2B with RES and ESS for Peak Load Minimization. *Applied Sciences*, *8*(11), 2125. https://doi.org/10.3390/app8112125

OECD. (2023). *World Energy Outlook 2023*. OECD. https://doi.org/10.1787/827374a6-en

Ouammi, A. (2021). Peak load reduction with a solar PV-based smart microgrid and vehicle-to-building (V2B) concept. *Sustainable Energy Technologies and Assessments*, *44*, 101027. https://doi.org/10.1016/j.seta.2021.101027

Schram, W., Brinkel, N., Smink, G., Van Wijk, T., & Van Sark, W. (2020). Empirical Evaluation of V2G Round-trip Efficiency. *2020 International Conference on Smart Energy Systems and Technologies (SEST)*, 1–6. https://doi.org/10.1109/SEST48500.2020.9203459

Taipower Company. (2020). *Taipower Electricity Rate—Starting 1st April 2018 (台灣電力公司電價表中華民國107年4月1日起實施)*. Taiwan Power Company. https://www.taipower.com.tw/upload/1478/2019092713462323288.pdf

Tan, K. M., Yong, J. Y., Ramachandaramurthy, V. K., Mansor, M., Teh, J., & Guerrero, J. M. (2023). Factors influencing global transportation electrification: Comparative analysis of electric and internal combustion engine vehicles. *Renewable and Sustainable Energy Reviews*, *184*, 113582. https://doi.org/10.1016/j.rser.2023.113582

Thomas, D., Deblecker, O., & Ioakimidis, C. S. (2018). Optimal operation of an energy management system for a grid-connected smart building considering photovoltaics' uncertainty and stochastic electric vehicles' driving schedule. *Applied Energy*, *210*, 1188–1206. https://doi.org/10.1016/j.apenergy.2017.07.035

Tseng, W.-C., & Hsieh, I.-Y. L. (2023). Impacts of electric fleet charging patterns under different solar power penetration levels: Hourly grid variations and operating emissions. *Transportation Research Part D: Transport and Environment*, *122*, 103848. https://doi.org/10.1016/j.trd.2023.103848

Wang, D., Guan, X., Wu, J., Li, P., Zan, P., & Xu, H. (2016). Integrated Energy Exchange Scheduling for Multimicrogrid System With Electric Vehicles. *IEEE Transactions on Smart Grid*, *7*(4), 1762–1774. https://doi.org/10.1109/TSG.2015.2438852

Wang, Z., Jochem, P., & Fichtner, W. (2020). A scenario-based stochastic optimization model for charging scheduling of electric vehicles under uncertainties of vehicle availability and charging demand. *Journal of Cleaner Production*, *254*, 119886. https://doi.org/10.1016/j.jclepro.2019.119886

Wu, Y., Ravey, A., Chrenko, D., & Miraoui, A. (2018). A Real Time Energy Management for EV Charging Station Integrated with Local Generations and Energy Storage System. *2018 IEEE Transportation Electrification Conference and Expo (ITEC)*, 1–6. https://doi.org/10.1109/ITEC.2018.8450235

Exploring a Possible Mobility-as-a-Service Platform from the Perspective of Transport Policy Makers and Regulators in Thailand

Sirawit Chinvanichai1,*, Sorawit Narupiti2

[1,2] Department of Civil engineering, Chulalongkorn University,
Bangkok, Thailand sirawit41489@gmail.com

ABSTRACT

Mobility as a Service (MaaS) is an innovative approach to smart mobility, offering integrated mobility solutions through a digital platform. This study presents a prototype MaaS platform designed for Bangkok, Thailand. To create and better understand MaaS platform components, a focus group meeting was held with Transport Policy makers and Transport regulatory agencies. The findings highlight the importance of updating regulations to support the evolving MaaS landscape. Additionally, the study suggests that the success of MaaS platforms depends on factors such as investor engagement and revenue models, which require further investigation. Furthermore, the study examines the advantages, disadvantages, and implementation possibilities of various platform configurations, including ownership, openness, data sharing incentives, and revenue streams. These findings provide valuable guidance for future MaaS research and contribute to a better understanding among stakeholders interested in MaaS platforms. This research sheds light on the complexities of implementing MaaS in Bangkok, Thailand, and lays the foundation for future MaaS initiatives in the study area.

Keywords: Mobility as a Service Platform, Transport policy maker perspectives, Transport Regulator perspectives

1. INTRODUCTION

Office of the Permanent Secretary, Ministry of Transport mentioned that, Under the 20-year policy and planning framework (2018 to 2027) for digital development in the national economy and society, the government of Thailand launched the "Digital Thailand" initiative. This initiative aims to digitalize Thai society, requiring related organizations and ministries to develop digital protocols as part of a digital-driven economy. In response, the Department of ICT, Office of the Permanent Secretary Ministry of Transportation (OPSMOT), developed a 5-year protocol (2023 to 2027) called the "Digital Transport Plan 2027" as a framework for driving digital transformation in transportation to enhance economic value. This plan also promotes the National Multimodal Transport Integration Center (NMTIC) and its networks.

The concept of reforming the transportation system using Mobility as a Service (MaaS) was raised during the "Transportation Trends & Innovation in 2024" event, organized by OPSMOT. The event aimed to create a session for idea exchange among executives and personnel from the Ministry of Transport (MOT) and related organizations. Participants included the Land Transport Federation of Thailand, Thai Transport and Logistics Association, Thai Shipowners' Association, Thai International Freight Forwarders Association (TIFFA), and Thai Airfreight Forwarders Association (TAFA).

This study was conducted during the "Smart Transportation Infrastructure & Facilities" session, with a focus on MaaS. The study aims to explore the possibility of establishing MaaS services in Thailand. The research question is: "What should be considered for establishing a Mobility as a Service system in Thailand?" The following questions were asked to the panel:

- What factors should be considered when developing a travel application with features like trip planning, booking, payment, and enroute assistance?

- Evaluate the advantages and disadvantages, identify potential impacts, and determine the most suitable options for implementing MaaS in Thailand, focusing on the following areas: MaaS Service Provider, Openness of the MaaS System, Data Sharing Policy and System's Fee Collection

The first question highlighted the importance of considering the components of MaaS, including the regulator (prospect MaaS regulator), regulations, investors, and business models. The second question explored potential app and platform operators, platform openness, solutions for enabling data sharing, and possible service agreements between the platform and service providers (SPs). The findings were collected and analyzed using thematic analysis.

Section 2 provides an understanding of the roles and responsibilities of stakeholders in the MaaS system, considerations for the mobility platform, and service agreements between MaaS Operators (MO) and SPs. Section 3 focuses on data collection and analysis in this study. Section 4 presents the preliminary results of the study. Conclusions and recommendations are proposed in Section 5.

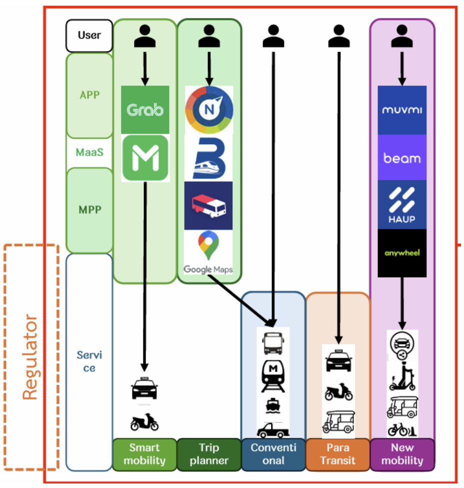


Figure 1 Current operations of Bangkok's transport system

2. LITERATURE REVIEW

Current transport situation

Figure 1 presents an overview of the current landscape where numerous public transport providers operate their own services, with some integrated with others and many functioning independently. According to Narupiti, S. (2019) Transport Service Providers (TSPs) can be categorized into five groups: Smart Mobility Providers (SM), Trip Planner Service Providers (TP), Conventional Service Providers (C), Paratransit Providers (P), and New Mobility Providers (N).

Smart Mobility Providers (SM, such as Grab and Lineman, offer mobility services through applications that feature trip planning, booking, payment, and subscription options for transport modes like taxis, motorcycle taxis, and car sharing. These services are provided under partnership programs and regulated by the Digital Platform Service Law (DPS) of the Electronic Transactions Development Agency (ETDA).

Trip Planner Service Providers (TP), including Google Maps, Namtang, BKK Rail, and Via Bus, offer trip planning services through their applications. They gather information from conventional transport modes using public data and GPS tracking.

Conventional Service Providers (C), such as buses, MRT, BTS, riverboats, canal boats, and Songthaews, offer fixed-route services at roadside stops without the use of applications or digital access. These services operate under route concession contracts with the Ministry of Transport of Thailand.

Paratransit Providers (P), such as taxis, tuk-tuks, and motorcycle taxis, provide variable-route services at roadside stops, with some partnering with SMs' applications. They also operate under route concession contracts with the Ministry of Transport of Thailand.

New Mobility Providers (N), such as Muvmi, Bean, Anywheel, GCOO, and HAUP, offer services through their applications, similar to Smart Mobility Providers. However, these providers rely on their own mobility assets. For example, Muvmi owns a fleet of E-Tuk Tuks, and GCOO operates a combined fleet of E-scooters and E-bikes. New Mobility Providers' application services are regulated by the DPS law, while their mobility assets are governed by various regulations established by the Ministry of Transport of Thailand. In contrast, Smart Mobility Providers are only regulated by the DPS law since their transport services are provided through partnership programs.

a. Feature of MaaS

Narupiti, S. (2019) identified five essential functions of prospective MaaS in Bangkok, Thailand: information, trip planning, booking (ticketing), payment, and enroute assistance. Additionally, Sochor, J. (2021) classified MaaS service integration into five levels, ranging from Level 0 to Level 4. These levels are defined as follows: Level 0 involves no integration across modes; Level 1 includes information integration (enabling trip planning features); Level 2 incorporates booking and payment integration; Level 3 offers integrated services like transport bundles and subscription services; and Level 4 integrates societal goals such as policies and incentives through MaaS. Therefore, the features of the MaaS application discussed in this paper are as follows:

- Trip Planning: This feature relies on the integration of information.
- Booking: Developed from integrated data but not yet including payment.
- Payment: Relies on the integration of payment systems.
- Enroute Assistance: Provides optimal suggestions for users in case of incidents during the trip, based on transport data and user preferences.

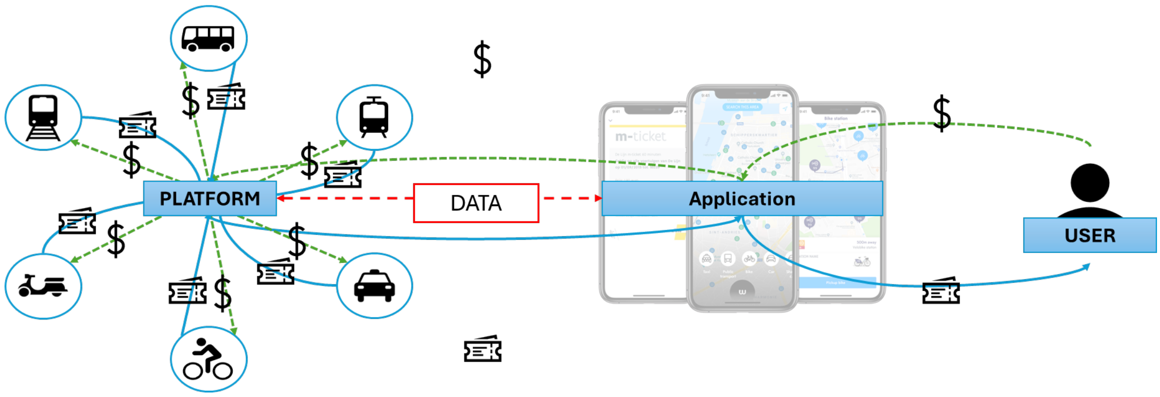


Figure 2 Mobility as a Service system

Figure 2 illustrates the Mobility as a Service (MaaS) system, divided into three layers: the Platform layer, Application layer, and User layer. On the left-hand side of Figure 2, the Platform layer in the MaaS system is responsible for gathering transport information such as fees, locations, availability, travel times, and waiting times, which support trip planning and enroute assistance features. The platform also acts as a data exchange intermediary between users and Transport Service Providers (TSPs), facilitating tasks like matching availability with user demand for booking features and serving as a fee distributor or clearinghouse for payment features.

In the middle layer of Figure 2, the Application(s) collect information from the platform to process user demands and payments through four account features and respond to these demands with transport information and/or tickets from TSPs. These two layers (Platform and Application) are thus identified as the MaaS Provider (MP) due to their intermediary-like role between TSPs and users.

Stake holder in MaaS

Figures 1 and 2 show the potential stakeholders in Bangkok's MaaS system. Users are on the right side of Figure 2, followed by MaaS application providers and platform providers, together called MP, as defined in Section b. Figure 1 categorizes possible Transport Service Providers (TSPs) into five groups, showing their roles and positions within the MaaS system, especially in the Platform layer on the left side of Figure 2. Transport regulators and their areas of control are shown in an orange rounded rectangle at the bottom-left of Figure 1. According to Kamargianni, M., & Matyas, M. (2017), Sochor, J. (2021), Narupiti, S. (2019), Mulley, C., & Nelson, J. (2020) and ERTICO – ITS Europe (Ed.), MPs play a key role in MaaS by offering mobility services through an application and bringing together services from TSPs through a platform. Narupiti, S. (2019) suggests three possible setups for how the Platform and Application layers of an MP could be aligned: (a) MP from a public transport provider; (b) MP from a private transport provider; and (c) MP from a third party or a public-private partnership (PPP). Mulley, C., & Nelson, J. (2020) and ERTICO – ITS Europe (Ed.) further explain that how MPs are organized in the mobility market depends on the ecosystem they belong to. MPs might independently offer their own MaaS services in a free market, or they could be publicly run, which would be called public-operated MaaS. MPs could also offer their applications on a regulated platform, which might be managed by a public organization, a group of TSPs, or a third-party application provider working with TSPs.

MaaS Provider (MP)

MP acts as an intermediary that offers services such as trip planning, booking, payment, and enroute assistance to users through an application. Additionally, MP provides integration services to Transport Service Providers (TSPs) via a platform, including demand management, data pooling, data analysis, payment processing, and clearinghouse services. As mentioned in Mulley, C., & Nelson, J. (2020) and ERTICO – ITS Europe (Ed.) the organization and limitations of MP operations can vary depending on the environment and platform scenarios discussed in.

MaaS in free market: MP operates in a marketplace environment where both MP and TSPs manage their agreements. Competition among MP is common, with no regulation of usage fees by transport authorities or transport regulator. The development of application (part of MP) focuses on customer-oriented solutions. Data sharing is limited to public authorities, restricting public transport service development and data planning. There is also a higher risk of biased transport option presentations.

Public MaaS: A MaaS system operated by a public transport authority acting as the MP, including selected mobility services. Regulations require other services to open their APIs. This model aims for sustainable mobility and social inclusion but may be less customer-oriented and innovative.

PPP MaaS: Established by a public authority with set rules, this MP serves as public infrastructure, enabling various actors to build MaaS solutions using platform-provided data. Competition focuses on front-end services, aiming to be customer-oriented, innovative, and impartial. Local mobility providers may join, but financial support remains an issue.

Transport Service Provider (TSP)

There are five categories of Mobility Service Providers (MSPs) in Bangkok, Thailand, as mentioned in Section a. Smart Mobility Providers (SM) operate private MaaS businesses, offering customers up to Level 3 MaaS services, including subscription and clearinghouse systems across various transport modes. SMs also establish partnerships with mobility suppliers, such as private car-sharing, motorcycle-sharing, and taxi services.

Investor

As highlighted in Kamargianni, M., & Matyas, M. (2017), investors hold a pivotal role in the development of the Mobility as a Service (MaaS) market. Both private investors and public funds are essential contributors to this market. Public investors, such as government authorities, may redirect funds from travel schemes and public transport subsidies to MaaS providers, potentially enhancing the efficiency of transport services and optimizing the use of public resources. Additionally, private investors have the opportunity to explore crowdfunding as an alternative means of raising capital. Nevertheless, further research is required to determine the most effective strategies for implementing these funding mechanisms.

Transport Regulator [show different between TSPs regulator, MP regulator]

Kamargianni, M., & Matyas, M. (2017) and Mulley, C., & Nelson, J. (2020)underscore the vital role of transport regulators and policymakers in shaping the Mobility as a Service (MaaS) market. Transport regulatory bodies, such as the Ministry of Transport and Communications (MTC) in Finland and the state government in New South Wales (NSW), Australia, are essential in establishing standards and regulations that ensure the interoperability and sustainability of MaaS. These transport regulators provide policy frameworks that address various aspects of MaaS, including fair competition, passenger rights, and sustainable development. The differing approaches of Finland and NSW highlight the importance of context-specific regulatory strategies in fostering MaaS innovation and growth. Overall, the literature emphasizes the need for regulators to work closely with industry stakeholders and international organizations to create a comprehensive regulatory framework that promotes innovation, efficiency, and sustainability within the MaaS ecosystem.

b. MaaS platform and business model

Digital mobility platform

As discussed in Sá, E., Carvalho, A., Silva, J., & Rezazadeh, A. (2022), experts evaluated key aspects of stakeholder collaboration on the platform. **Platform ownership and management** were reviewed in four scenarios: public initiatives offer trust but may lack leadership and business focus; ecosystem member ownership fosters inclusivity but faces cost and transparency issues; consortium platforms encourage innovation but struggle with stability; private ventures drive innovation but may prioritize profit and raise privacy concerns. For **platform openness**, two scenarios were considered: an open platform encourages broad participation but risks fairness, while a closed platform ensures quality and security but limits inclusiveness. In terms of **data sharing**, approaches include mandatory participation for transparency and commitment, a digital data marketplace for fair and incentivized sharing, voluntary sharing for privacy and wide participation, and rewards-based sharing to encourage contribution.

Service-Based Business Model

This model directly offers services to users through a digital platform, such as trip planning, booking, and payments. Revenue is generated from service fees, subscriptions, and transaction commissions. It can attract many users by bundling convenient services, but its success depends on how effectively the platform can monetize these services. This paper will explore the service-based business model as a potential approach.

Data-Based Business Model

This model relies on the data generated by platform usage. The platform provides analytics and insights to stakeholders, like cities and service providers, who can use this data to improve their operations. Revenue is earned by selling data access, analytics, and subscriptions to these insights. This model can create more long-term value by helping with better decision-making and planning, but it requires a strong system for managing data and protecting privacy.

Service agreement

König, D., Eckhardt, J., Aapaoja, A., Sochor, J., & Karlsson, M. (2016) indicates that the types of agreements between MaaS Providers (MP) and transport providers can vary depending on the model used. A single MaaS service may involve multiple types of agreements. The primary types identified are re-sold services and negotiated services. Re-sold services typically use list fares or fixed reduction percentages, as seen in examples like Telia Company and VR (Finnish Railways) or SNCF iDAVIS. Negotiated services are based on bilateral agreements, customized to the specific needs of the involved parties, such as Telia Company's Sonera Reissu with taxis or SNCF's Pack mobilité. These agreements are crucial for defining the terms of collaboration and ensuring smooth integration of services within the MaaS platform.

3. MATERIAL AND METHODS

a. Data collection

This study's data collection was conducted using a focus group method during a specific session of the event. The research team divided a panel of 23 experts, as shown in Table 1, into four smaller panels to ensure a diverse range of expertise and experience.

Table 1List of Expert participate in the Focus group

Organization	Number or participant	% of participant
Office of the Permanent Secretary Ministry of Transportation	1	4%
Office of Transport and Traffic Policy and Planning	6	26%
Department of Highways	3	13%
Department of Rural roads	3	13%
Expressway Authority of Thailand	5	22%
Department of Land Transport	4	17%
Bangkok Mass Transit Authority	1	4%
Total	23	

a brief lecture session to establish a common understanding among participants, focusing on a predetermined set of questions for the panels. Following this, each panel was asked to present their ideas on the possibilities, enablers, barriers, and issues related to different scenarios of Mobility as a Service (MaaS) based on the following questions:

- What factors should be considered when developing a travel application with features like trip planning, booking, payment, and enroute assistance?
- Evaluate the advantages and disadvantages, identify potential impacts, and determine the most suitable options and conditions for implementing MaaS in Thailand, focusing on the following areas with deterministic choices:
    - Platform Ownership and Management:
        - Public Initiative: Trustworthy and inclusive but may lack leadership and business focus.
        - Ecosystem Member Ownership: Inclusive but faces challenges with cost, bias, and transparency.
        - Consortium: Encourages innovation and collaboration but struggles with long-term stability.
        - Private Venture: Promotes independence and innovation but raises privacy concerns and may prioritize profit.
    - Openness of the MaaS System:
        - Open Platform: Allows broad participation and growth but risks fairness and control.
        - Closed Platform: Ensures quality and security but limits reach and inclusiveness.
    - Data Sharing Policy:
        - Mandatory: Ensures success with committed members and transparency.
        - Digital Data Marketplace: Encourages data sharing with benefits and fair access.

    - Voluntary: Supports privacy and broad participation without obligation.
    - Rewards-Based: Incentivizes data sharing by offering access to additional services.
  - System's Fee Collection:
    - List Fare
    - Fixed Reduction (%)
    - Negotiated Service

b. Analysis method

In this research, several analytical methods were used to thoroughly examine the data collected from the focus group sessions.

Data Collection and Preparation:
During the pitching sessions, discussions were recorded, transcribed, and then cross-validated with visual observations to ensure the accuracy and validity of the responses.

Thematic Analysis:
A thematic analysis was conducted to identify key themes and sub-themes, ensuring that participants' responses were consistent and accurately reflected the research questions. This helped filter out any misunderstandings.

Content Analysis:
Content analysis was used to extract important concepts and patterns related to MaaS implementation, providing deeper insights into the qualitative data.

Frequency Analysis:
Frequency analysis identified the most common words and phrases, highlighting the main concerns and ideas expressed by the participants.

These methods provided a comprehensive understanding of the possibilities, enablers, barriers, and issues related to MaaS. They also supported the evaluation of options for platform ownership, system openness, data sharing, and fee collection in the context of MaaS implementation in Thailand.

4. RESULTS

Table 2 Thematic analysis result

Theme	Sub-theme	Quotes	Key Points
Data Integration	Digital Data	"DLT have not access to Rail transport data (BTS, MRT)", "Enablers: Digital timetable Data validation"	Lack of centralized digital data repository; validation of timetable data is crucial.
	Data Openness	"Lack of data openness: Government need to initiate open data (opened by default)"	Governmental support is needed to promote open data policies.
	Static to Dynamic Schedule	"No Dynamic schedule: need to move from static to dynamic schedule"	Transitioning to dynamic scheduling is essential for real-time data accuracy and enhancing user experience.
Planning and Regulation	Regulation and Regulatory Bodies	"Create Proper Regulator is as important as identify the most suitable MaaS regulator", "Find and assign the proper Data owner"	Identifying proper regulators and managing data ownership are critical
	Business Model and Investor Involvement	"Investors are more important to consider than Business model"	Clear business models and investor involvement are necessary after establishing regulatory frameworks.

a. Thematic analysis result

Table 2 Indicate the result of thematic analysis, three main themes were identified: 1. Data Integration, 2. Planning and Regulation, and 3. Service Customization. In Data Integration, key challenges include the lack of a central digital data system and the need to move from static to dynamic scheduling for better real-time accuracy. The Planning and Regulation theme highlighted the need for clear regulations, proper regulatory bodies, and a solid business model supported by investor involvement. Service Customization showed that offering personalized services can increase user satisfaction, but the current lack of customization is a major obstacle. These findings will help shape recommendations for the future development of MaaS systems.

b. Content analysis result

Digital Data Integration: There is a significant focus on the lack of centralized digital data, particularly static data like timetables, and the challenges this creates for dynamic scheduling. The document highlights the need for government intervention to promote open data and encourage the development of dynamic schedules.

Planning and Regulation: The discussion centers around identifying appropriate regulatory bodies to oversee the MaaS system. The importance of a top-down regulatory approach and the need for a clear business model to attract investors are emphasized.

Service Models: The document explores different service models, including public, private, and consortium approaches. Each model has its advantages and disadvantages, particularly concerning sustainability, control, and profitability.

Data Sharing: Various approaches to data sharing are discussed, including mandatory participation, voluntary sharing, and the creation of a digital data marketplace. The focus is on balancing the benefits of data openness with the need for privacy and control.

c. Frequency analysis result

Most Frequently Mentioned Concepts:

Data (Integration, Sharing, Digital): The document frequently discusses data in terms of integration, openness, and the need for dynamic systems.

Regulation/Regulators: The importance of having proper regulatory bodies and frameworks is a recurring theme.

Service Models: Various service models (public, private, consortium) are discussed in detail, emphasizing their impact on the MaaS system.

d. Implication

Data Integration:

The development of a MaaS system using the current resources from "Namtang" will encounter significant challenges related to data integration. The primary issue is the lack of sufficient static data, such as service schedules and timetables provided by Service Providers (SPs). While some SPs have attempted to create and share timetables, these efforts often suffer from data validation issues, with many datasets not being up to date. The absence of reliable static data presents a major barrier, forcing firms or organizations interested in initiating trip planning services to generate their own data. This includes dynamic data, such as GPS-based vehicle location tracking, which is costly and could deter potential stakeholders from investing. To encourage investment in the trip planner business, the government must implement regulations that require SPs to initiate, validate, and share their data with the overseeing regulator. Additionally, promoting the development of dynamic schedules will make SPs more appealing to trip planner providers by reducing the cost burden. Finally, the public sector needs to establish policies that create a win-win situation for SPs, particularly those who are hesitant to share their data.

Development Permutations:

The prevailing suggestions point toward a top-down approach, where the government must identify a regulator capable of overseeing all SPs within the MaaS system. For example, the Department of Land Transport can regulate buses, taxis, and motorcycle taxis, but not rail systems, which fall under the Department of Rail Transport. The next step involves addressing the current regulatory environment, which is not conducive to MaaS development. It has been proposed that regulatory improvements could be deferred during the early stages of MaaS implementation, placing greater importance on developing a viable business model and attracting investors. This approach hinges on the establishment of pilot licensing, allowing regulators to address complications and refine regulations based on lessons learned from the pilot stage. The interplay between business models and investor interests is crucial; some argue that potential investors should drive business model development, as they are likely to create models that maximize their advantage within the MaaS system. However, regulators should ensure that private sector gains from public infrastructure are balanced with sufficient incentives to attract investment while safeguarding public interests.

e. Recommendation

Application (Platform) Ownership and Management:

The findings suggest that a MaaS application owned and managed by the public sector would enhance sustainability and provide benefits through the development of a more accessible system for various user groups, including those with disabilities or low income. However, regulators must ensure a legal framework that supports the continuous operation of such a system. Conversely, allowing private service providers (SPs) to own and manage the MaaS platform may prioritize profit, potentially complicating the regulatory environment. Despite this, the rapid adaptability of private firms could be advantageous for developing demand-responsive applications. For private sector involvement to be viable, the MaaS ecosystem must be sufficiently profitable to attract investment.

A consortium-based approach, where both the public sector and SPs collaborate in managing the platform, could offer broader service coverage. This model allows SPs to share or co-own resources, reducing costs for all parties involved. Additionally, it enables SPs to leverage shared resources to generate marginal benefits from the demand synergy within the platform. Regulators may also have opportunities to apply regulations or provide incentives to support or subsidize SPs as needed.

Platform Openness:

The panel recommends that an open platform, particularly for data, during the early stages of MaaS implementation, will make the system more attractive to high-value players. Allowing more SPs to join and benefit from the platform will expand service coverage and market size, making the platform easier to scale up.

Data Sharing Conditions:

Creating a "Data Marketplace" is strongly recommended as the most effective and beneficial way to encourage SPs to share their data across the platform. Since generating and owning data involves costs, enabling SPs to add value to their data is a way to increase profits without raising fares. A data platform would also streamline data management, reducing costs for SPs. Given that data is a valuable asset, limitations, conditions, and standard protocols for data sharing should be established. For instance, sensitive data could be restricted from being shared or sold in the marketplace. As a regulator, the government can enhance this data platform by creating an initial market and implementing regulations to ensure fair play as the marketplace scales up.

Service Agreements:

Service agreements between platform providers and SPs should operate in a free-market environment, meaning that government intervention in service agreements should be minimal. The public sector should focus on establishing service agreement protocols for various SP sectors and creating a fair environment for new SPs who wish to join the MaaS platform in the future.

5. CONCLUSIONS

In developing a Mobility as a Service (MaaS) system in Thailand, data integration emerges as the most critical factor, with the need for a robust Public-Private Partnership (PPP) platform to drive collaboration and innovation. An open platform during the early stages will attract key service providers, while transitioning to a partially regulated market can ensure sustainability and inclusiveness. Future research should focus on overcoming data-related challenges and exploring the mutual benefits of collaboration among stakeholders, ensuring the successful implementation and growth of MaaS in Thailand.

Data is the First to Consider

The thematic analysis and content analysis highlight that data integration is the primary challenge in developing a MaaS system in Thailand. The lack of centralized, validated static data, such as timetables, and the absence of dynamic scheduling are significant barriers. To overcome this, public and private sector service providers (SPs) must begin generating and sharing both static and dynamic data. Government intervention is crucial to enforce data validation and sharing protocols, ensuring that the data platform supports the successful implementation of MaaS. The establishment of a data marketplace could incentivize SPs to share data, adding value to the ecosystem and reducing operational complexities.

PPP Platform

A Public-Private Partnership (PPP) platform is recommended as the most advantageous ownership and management model for MaaS. This model allows the public sector to guide the platform's operations toward maximizing social benefits, while private partners focus on innovation, collaboration, and service development. Depending on the public partner's decision, the platform can be established either through a public-led approach, prioritizing regulatory structures, or through a private-led approach, emphasizing business models and value creation. The PPP model balances the strengths of both sectors, fostering a sustainable and socially inclusive MaaS system.

Partial Regulated Market

During the initial stages of MaaS implementation, an open platform is advisable to attract high-value service providers and build a robust customer base. This openness can facilitate platform growth, making it easier to scale up. However, as the platform evolves, transitioning to a partially regulated market may be necessary, with specific conditions and criteria applied to new service providers. This approach ensures that the platform remains inclusive and socially beneficial while maintaining the quality and security of the services offered. The establishment of a data marketplace within this framework can further enhance collaboration and innovation among SPs.

Future Research

Future research should focus on identifying the barriers and enablers of data generation and sharing, which are critical for the success of the MaaS data platform. Studies should explore the perspectives of different stakeholders, including application developers, platform providers, and transport operators, to gain a comprehensive understanding of the ecosystem. Additionally, research on the collaboration between platform providers and SPs, particularly regarding the mutual benefits and pay-offs, could be enriched using frameworks like game theory. This will provide valuable insights for optimizing the MaaS platform and ensuring its long-term success.

6. ACKNOWLEDGMENTS

We would like to thank you Asst. Prof. Pongsun Bunditsakulchai, Ph. D. as one of the contributors of this paper, the opportunity given from him to lets me enroll in "Transportation Trends & Innovation in 2024" as his research assistant is a crucial part of the contribution of this study. Furthermore, I would like to thank you to whom that contributed and are a part of this seminar for allowing me to conduct this study during the seminar. And finally, I would like to gratefully thank my advisor and co-author for wisdom and experience through the process we have discussed and advising.

7. REFERENCES

สำนักงานปลัดกระทรวงคมนาคม.(2018).ข่าวสาร.กระทรวงคมนาคม. https://motapplication.mot.go.th/mot/20-news-web/default.html?P=S20WS000_1&ROW_ID_NEWS_M_GROUP=3&DATA_TOPIC=&submit=submit

MaaS Alliance. (2023, May 29). Mobility as a feature (MaaF): Rethinking the focus of the second generation of MaaS. MaaS Alliance. https://maas-alliance.eu/2023/05/29/mobility-as-a-feature-maaf-rethinking-the-focus-of-the-second-generation-of-maas

Kamargianni, M., & Matyas, M. (2017). The business ecosystem of Mobility-as-a-Service.

Sochor, J. (2021). Piecing together the puzzle: Mobility as a Service from the user and service design perspectives.

Narupiti, S. (2019). Exploring the possibility of MaaS service in Thailand, implications from the existing conditions and experts' opinions on "Who should be the MaaS provider in Bangkok?". IATSS Research, 43(4), 226-234. https://doi.org/10.1016/j.iatssr.2019.10.001

Mulley, C., & Nelson, J. (2020). How Mobility as a Service impacts public transport business models. International Transport Forum Discussion Papers (No. 2020/17). OECD Publishing. https://doi.org/10.1787/4d8237e9-en

ERTICO – ITS Europe (Ed.). (n.d.). *Mobility as a Service (MaaS) and Sustainable Urban Mobility Planning*.

Sá, E., Carvalho, A., Silva, J., & Rezazadeh, A. (2022). A Delphi study of business models for cycling urban mobility platforms. Research in Transportation Business & Management, 45, 100907. https://doi.org/10.1016/j.rtbm.2022.100907

König, D., Eckhardt, J., Aapaoja, A., Sochor, J., & Karlsson, M. (2016). Deliverable 3:

Automatic Recognition of Pavement Deterioration using Deep Learning

Szu-Han, Lu[1,*], Jyh-Dong, Lin[2], Yi-Shian Chiou[3]

[1]Civil Engineering/ National Central University, Taiwan/ vul3au33@gmail.com
[2]Civil Engineering/ National Central University, Taiwan / jyhdongl@ncu.edu.tw
[3]Civil Engineering/ National Central University, Taiwan / Lettermail64@gmail.com

ABSTRACT

In recent years, there has been an increasing variety of methods for using deep learning to identify pavement damage. However, there is a lack of unified standards and norms for recognition outcomes. This study focuses on two main objectives. Firstly, it aims to determine whether to identify damage on urban roads in Taiwan according to the Flexible Pavement Distress Survey Manual formulated by the domestic Construction and Planning Agency or based on the American Society for Testing and Materials (ASTM) standard D 6433-23. After identifying the types of damage, this research utilizes deep learning to recognize various pavement damages. It compares the recognition results of different models to explore the precision and accuracy of identifying each type of damage. Recommendations are then provided regarding the preferred model for using deep learning to recognize pavement damage. Additionally, the study compares automatic recognition with manual recognition using the established model to observe differences in the Pavement Condition Index (PCI) derived from both methods, aiming to determine whether deep learning can achieve results comparable to manual recognition.

Keywords: Pavement Condition Index, deep learning, image recognition, convolutional neural networks

1. INTRODUCTION

In modern transportation infrastructure, the accurate assessment of the Pavement Condition Index (PCI) is crucial for road maintenance. However, the lack of unified standards and certification mechanisms in previous PCI evaluations may result in inconsistencies and reduced reliability of inspection results. This introduction will explore this issue and use ASTM D6433-23 as a unified inspection standard to establish a clear inspection framework, thereby improving the consistency and reliability of PCI assessments.

This study aims to achieve three primary objectives. Firstly, based on the ASTM D6433-23 standard, our team will establish a comprehensive PCI inspection method, including indicators for 20 types of flexible pavement distress, to evaluate pavement conditions more thoroughly and meticulously. Secondly, we will analyze the accuracy and precision of data obtained by different inspectors, explore its impact on PCI scores, and use the coefficient of variation to verify the consistency of inspection results. Finally, through education and training, we aim to enhance the inspectors' ability to identify pavement distress, reduce subjective errors, and establish a reliable manual inspection and certification mechanism. These objectives will contribute to improving the standardization level of PCI inspections, thus providing more reliable assessment tools for road maintenance and management.

2. LITERATURE REVIEW AND DISCUSSION

2.1 Pavement Distress Survey Manual

The Pavement Distress Survey Manual, promulgated by the Construction and Planning Agency of the Ministry of the Interior in Taiwan in 2002, serves as an essential guideline in the field of road engineering in Taiwan. It aims to assist engineers in evaluating the condition and extent of damage to flexible pavements. This manual references the ASTM D6433-99 (1999) testing methods and covers various forms of distress such as cracking, potholes, and manhole issues. This comprehensive approach enhances the accuracy and consistency of damage assessments. The manual is specifically designed considering Taiwan's climate, road design, and traffic conditions. The types of distress are categorized into four main types: "Surface Cracks," "Surface Deformations," "Surface Damage," and "Others." These four types are further subdivided into 13 specific forms of distress, which can be detailed into a total of 16 specific distress types, as shown in Table 1.

Table 1 Comparison of Distress Types between Flexible Pavement Survey Manual and ASTM D6433-23

Item	Pavement Distress Survey Manual			ASTM D6433-23	
	Distress Type	Code	Distress Type	Code	Distress Type
1	Surface Cracks	1	Allgator Cracking	1	Allgator Cracking
2		2-1	Longitudinal Cracking	2	Bleeding
3		2-2	Transverse Cracking	3	Block Cracking
4		3	Block Cracking	4	Bumps and Sags
5	Surface Deformations	5	Rutting	5	Corrugation
6		7	Shoving	6	Depression
7		8	Bumps and Sags	7	Edge Cracking
8		10	Corrugation	8	Joint Reflection Cracking
9		11	Lane/Shoulder Drop-off	9	Lane/Shoulder Drop-off
10	Surface Damage	4-1	Potholes	10	Longitudinal/Transverse Cracking
11		4-2	Cover	11	Patching and Utility Cut Patching
12		4-3	Delamination	12	Polished Aggregate
13		12	Slippage Cracking	13	Potholes
14		13	Raveling	14	Railroad Crossing
15	Others	6	Patching	15	Rutting
16		9	Bleeding	16	Shoving
17	-	-	-	17	Slippage Cracking
18	-	-	-	18	Swell
19	-	-	-	19	Raveling
20	-	-	-	20	Weathering (surface wear)

2.2 Deep Learning

Deep learning is primarily implemented through neural networks (NN). Neural networks consist of countless neurons that operate similarly to human neurons. However, instead of transmitting bioelectric signals, they handle variables, weights, and biases. Each neuron acts as a small function, and data is gradually transformed within the neural network, ultimately forming a predictive model. The main computational models in deep learning are divided into two categories.

The first category is feedforward neural networks (FNN), where artificial neurons can respond to surrounding units within their coverage area, performing exceptionally well in large-scale image processing. The second category is recurrent neural networks (RNN), which excel in describing dynamic temporal behaviors and are used primarily for tasks like predicting future outcomes such as stock analysis and weather forecasting. Since this study focuses on image recognition, convolutional neural networks (CNN) are used as the primary method for identifying pavement damage.

3. RESEARCH METHODOLOGY

This study installs a simple image recognition device on inspection vehicles. To enhance the recognition rate of the convolutional neural network, high-resolution industrial digital cameras are used to capture images of all pavement damages. The Global Navigation Satellite System (GNSS) is used to confirm the location of the captured images, and a Distance Measuring Instrument (DMI) is used to ensure that each photo is taken at fixed intervals. This ensures that the photos are equidistant, enabling accurate calculation of the Pavement Condition Index (PCI).

3.1 Detection Equipment

The equipment used in this study is divided into five main modules for pavement damage recognition: the main module, camera module, positioning module, distance module, and power module. The detailed descriptions are as follows:

- Main Module: Processes all detection equipment operations, captures pavement images, positioning information, and driving distance information, and transmits the detected data in real-time to the cloud, where pavement damage is automatically recognized.
- Camera Module: This module captures various pavement images using different focal lengths, apertures, and angles of view, storing them in the main module.
- Positioning Module: This module uses the Global Positioning System (GPS) to determine the real-time location of the vehicle and corrects the positioning errors of the damage locations based on the camera module, accurately locating the damage.
- Distance Module: This module measures the distance traveled by the vehicle. It captures relevant pavement images at different distances and converts the capture frequency and images into the detection area, facilitating the calculation of the Pavement Condition Index (PCI).

3.2 Automatic Pavement Damage Recognition

This study primarily uses deep learning methods to recognize pavement damage. In computer vision (CV) image analysis algorithms, image analysis can be divided into four categories: Image Classification, which typically identifies only one type of damage object in a single image; Object Detection, which can identify multiple damage objects in a single image, of the same or different categories, without limitations on size or quantity; Semantic Segmentation, which can outline the correct area of objects but groups all pixels of the same category together when multiple overlapping objects of the same damage category are present, making it difficult to determine the size, position, number, and boundaries of the damage objects; and Instance Segmentation, which, unlike semantic segmentation, can distinguish between multiple objects of the same damage category and can accurately outline the irregular areas of the damage. This study evaluates the effectiveness of these four commonly used algorithm models and determines which is most suitable for this research. The models used are:

- Image Classification: Inception v4, a model in the GoogLeNet series that uses smaller network layers while achieving high accuracy.
- Object Detection: YOLOv8, currently the fastest and most accurate system for object detection.
- Semantic Segmentation: FCN, which can accept input images of any size, avoiding issues with pixel duplication and calculation.
- Instance Segmentation: YOLACT++, the first model to achieve real-time recognition in instance segmentation, mainly by adding a mask branch directly to the one-stage object recognition algorithm.

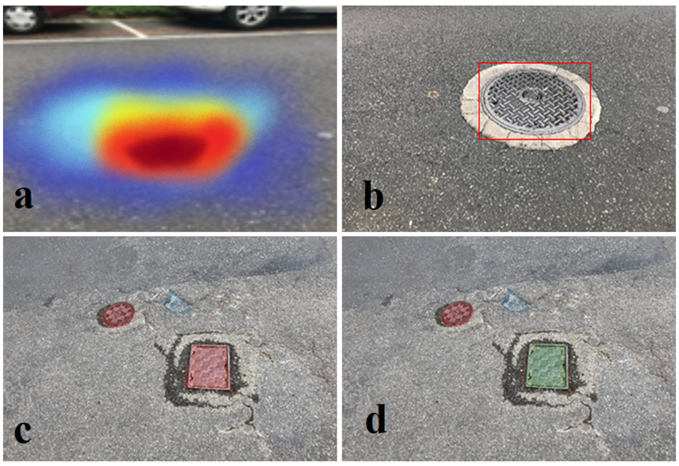


Figure 1 Recognition Results of Different Models
a. Image Classification; b. Object Detection; c. Semantic Segmentation; d. Instance Segmentation.

(Szu-Han LU ,2023)

Since the primary goal of this study is to accurately calculate the Pavement Condition Index (PCI), it is essential to correctly determine the damage area and the number of damage instances. Therefore, instance segmentation will be used as the image recognition method for this study. The recognition results of these models are illustrated in Figure 1

4. ANALYSIS AND DISCUSSION OF AUTOMATIC RECOGNITION RESULTS

The models for automatic pavement damage recognition are trained by computers, primarily identifying damage types and categories from image data. Currently, the study follows the 13 damage types from the domestic Flexible Pavement Distress Survey Manual, subdividing them into 16 specific types of pavement damage. For instance, longitudinal and transverse cracks are separated into distinct categories. The 16 types of damage include alligator cracking, longitudinal cracking, transverse cracking, block cracking, rutting, corrugation, heave and settlement, shoving, lane-shoulder separation, potholes, manhole cover difference, delamination, slippage cracking, raveling, patching, and bleeding.

After defining these 16 damage types, the collected damage data is annotated accordingly. Once all data is annotated, an instance segmentation model is selected for use. After selecting the model, deep learning training is conducted. During the training process, the model's automatic validation results are continuously evaluated to see if they exceed the expected recognition rate. If the expected recognition rate is not met, the relevant recognition parameters are adjusted. This process is repeated until the best model prediction results are achieved. The workflow is illustrated in Figure 2.

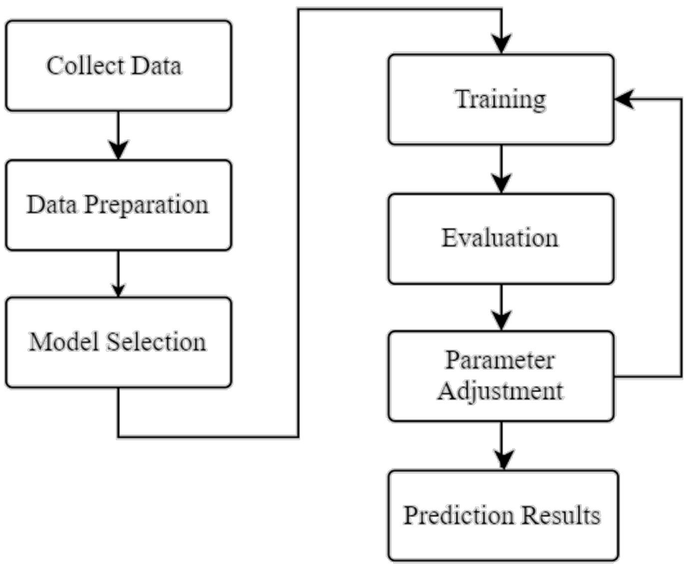


Figure 2 Training Model Workflow

(Szu-Han LU ,2023)

By implementing this structured approach, the study aims to develop an effective and reliable model for automatically recognizing pavement damage, ultimately improving the accuracy of the Pavement Condition Index (PCI) calculations

4.1 Training the Model

Based on the Flexible Pavement Distress Survey Manual and ASTM D6433 standards, this study identifies the most severe types of damage on roads as: alligator cracking, longitudinal cracking, transverse cracking, block cracking, potholes, manholes, and patching, totaling seven types of damage. Approximately 18,000 pavement damage photos were collected on-site. The details are shown in Table 2. The photos were taken on-site, and the difference between effective and ineffective photos lies in whether the damage can be clearly identified from the photos. As shown in Figure 3, the left image clearly shows alligator cracking, while the right image does not, thus it is classified as an ineffective photo.

Table 2 Number of Pavement Damage Photos

Damage Type	Number of Photos	Effective Photos	Effective Ratio
Alligator Cracking	4,675	3,646	78.0%
Potholes	1,358	1,208	89.0%
Cover	7,245	5,361	74.0%
Patching	5,218	3,913	75.0%
Total	18,496	14,128	76.4%

Figure 3 Difference Between Effective and Ineffective Photos in Images

(Szu-Han LU ,2023)

The concept of model training involves continuously adjusting the parameters within the model to increase its accuracy. This process aims to minimize the loss function, where the loss represents the residual between actual values and predicted values. In machine learning, the most common approach is to learn from the data and build a model, a process known as model training, which enables identification and prediction from the data.

In this study, the training data is divided into three groups: Training Dataset, Validation Dataset, and Test Dataset. The data is split with 70% for the training set, 20% for the validation set, and 10% for the test set. The trained data is then compared, with the quantities detailed in **Table 3**.

Table 3 Proportion of Training Photos

Damage Type	Total Photos	Training Set	Validation Set	Test Set
Alligator Cracking	4,675	3,272	935	468
Potholes	1,358	951	272	135
Cover	7,245	5,072	1,449	724
Patching	5,218	3,653	1,044	522
Total	18,496	12,948	3,700	1,849

The training set is primarily used during the machine training phase to fit the model, directly participating in the parameter adjustment process. The validation set is used during training to assess the model's initial performance and guide hyperparameter tuning, utilizing cross-validation methods to reduce issues such as overfitting or selection bias. Common cross-validation methods include holdout CV, leave-one-out CV, and k-fold CV. By adjusting the proportions of different photo quantities during parameter tuning, better recognition performance can be achieved. The test set is used to evaluate the model's recognition capability after training. To accurately assess the model's true performance, the test set is not involved in parameter adjustments or feature selection. The recognition results after completing the training are shown in Figure 4.

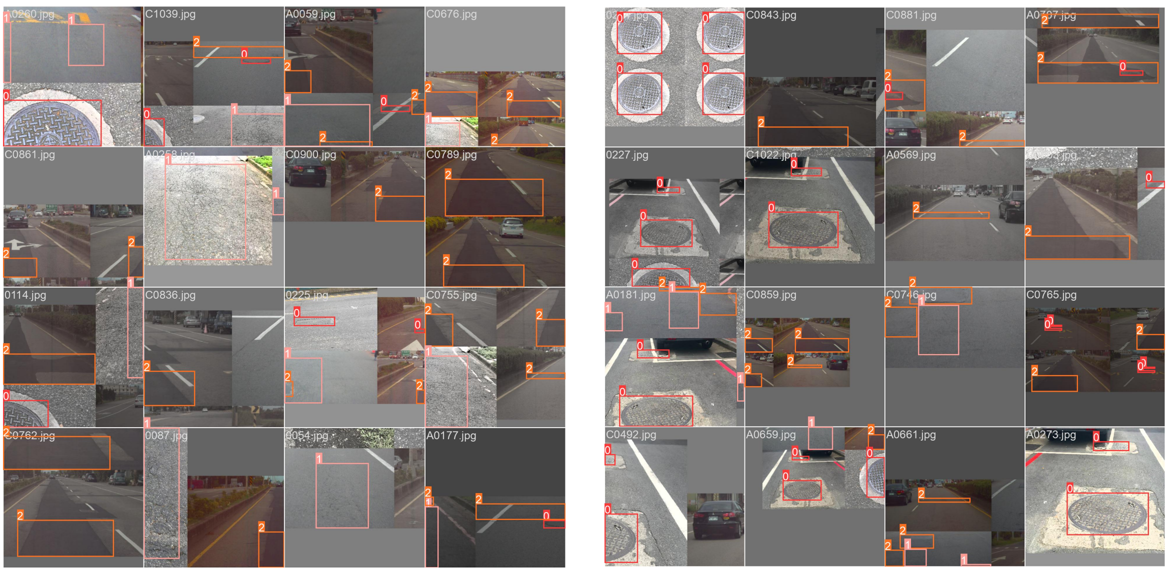


Figure 4 Recognition Results

4.2 Evaluating the Model

In machine learning, the goal is often to maximize or minimize a function or metric, commonly referred to as the objective function. In deep learning, this objective function is typically a loss function. The quality of a trained model largely depends on the design of the loss function, as shown in Equation 1.

$$loss = y - \hat{y} \tag{1}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3}$$

In machine learning statistics, the most commonly used loss function for regression is the Mean Square Error (MSE), as shown in Equation 2. It calculates the mean of the squares of the differences between the predicted and actual values. A higher MSE indicates a poorer model. To address negative values, the Root Mean Square Error (RMSE) is used, as shown in Equation 3. It provides a more accurate measure by taking the square root of MSE.

During the training process, it is essential to consider multiple iterations (epochs). Passing the entire dataset through the neural network once is insufficient; it must be passed multiple times. As the number of iterations increases, the frequency of weight updates in the neural network also increases, which can cause the curve to transition from underfitting to overfitting. This relationship between the loss function and the number of iterations is illustrated in Figure 5.

Figure 5 presents various metrics used to evaluate the performance of the deep learning model during training and validation. The graphs are divided into two rows: the top row shows metrics for the training set, and the bottom row shows metrics for the validation set.

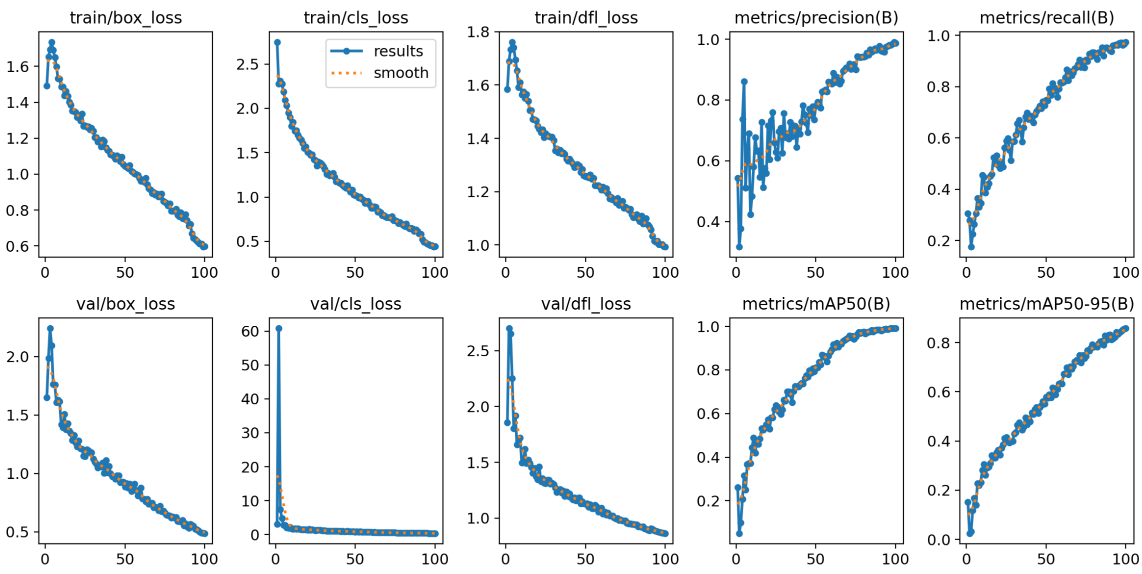


Figure 5 Evaluation Metrics of Deep Learning Model During Training and Validation

(Szu-Han LU ,2023)

4.3 Model Results

Since this study focuses on various types of pavement damage, it specifically examines the training results for each type of pavement damage and applies the trained models for testing. The following figures present the evaluation metrics of the model in recognizing pavement damage.

- ***Precision Curve***

The precision-confidence curve shows the precision for different damage categories at varying confidence levels. Precision is plotted against the confidence threshold, illustrating how precision changes with varying confidence levels. The blue line represents the overall performance across all categories, achieving a precision of 1.00 at a confidence threshold of 0.804. This indicates that at higher confidence thresholds, the model can accurately identify most instances of damage.

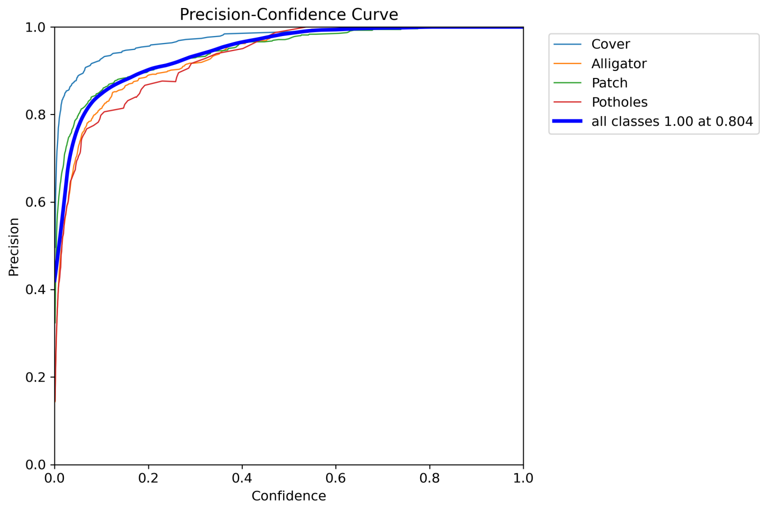


Figure 6 Precision-Confidence Curve

(Szu-Han LU ,2023)

- ***Recall Curve***

The recall-confidence curve shows the relationship between recall and confidence levels for different damage categories. Recall is plotted against the confidence threshold, illustrating how recall changes with varying confidence levels. The blue line represents the overall performance across all categories. In Figure 6, the model achieves a recall of 1.00 at a confidence threshold of 0.000. This indicates that at lower confidence thresholds, the model can identify all instances of damage, albeit potentially at the expense of precision.

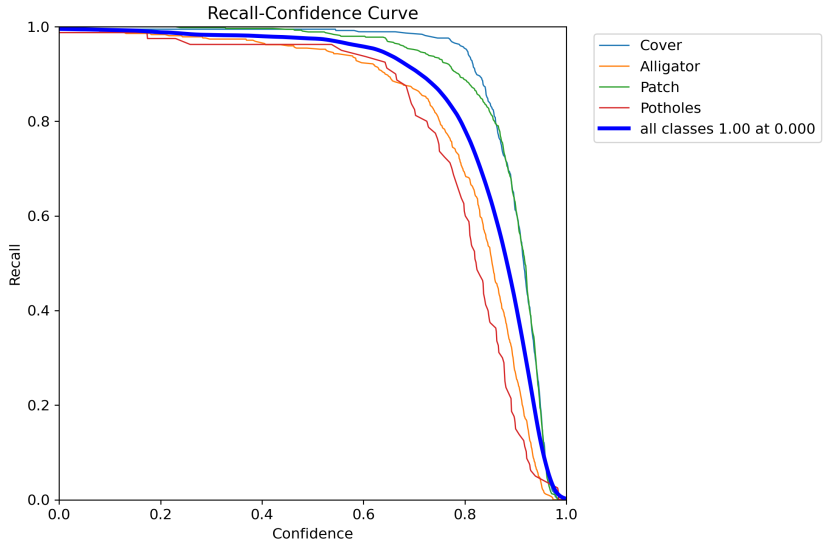


Figure 7 Recall-Confidence Curve

(Szu-Han LU ,2023)

- ***Precision-Recall Curve***

The precision-recall curve shows the precision and recall for different damage categories. Precision measures the proportion of true positive predictions among all positive predictions, while recall measures the proportion of true positive predictions among all actual positives. In Figure 7, the blue line represents the overall performance across all categories. The model demonstrates high precision and recall for different damage categories, with a combined mean average precision (mAP) of 0.991 at an IoU threshold of 0.5.

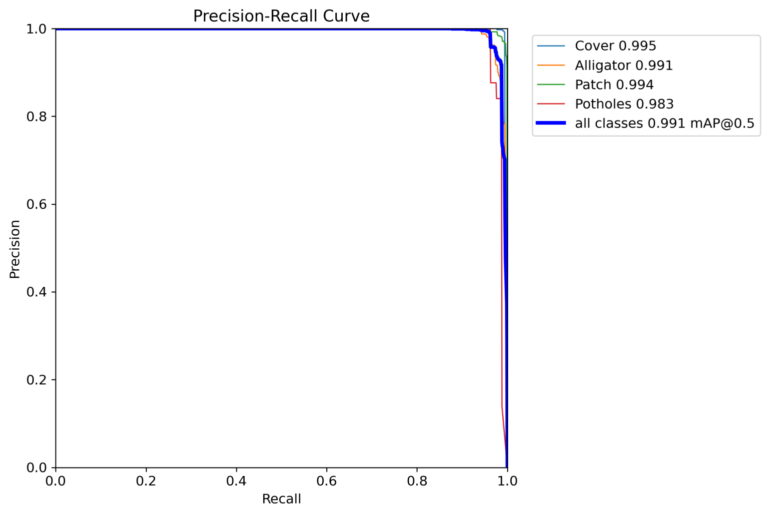


Figure 8 Precision-Recall Curve

(Szu-Han LU ,2023)

The evaluation metrics presented in the figures illustrate that the YOLO model performs well in recognizing different types of pavement damage. The model achieves high precision and recall, indicating its effectiveness in accurately identifying and classifying pavement damage. By adjusting the confidence thresholds, the model can optimize its performance to balance precision and recall as needed.

5. CONCLUSION AND RECOMMENDATIONS

This study applied deep learning to develop a model capable of automatically recognizing common types of pavement damage on urban roads. Upon validation, it was found that models trained with larger datasets achieved higher recognition rates. For instance, the model achieved a recognition rate of 93.18% for manhole covers. However, some categories exhibited lower recognition accuracy. After iterative parameter adjustments through training and testing, it was determined that the primary cause of this issue was an insufficient number of training images for certain damage types.

To address this, future efforts should focus on collecting and categorizing more images of the less frequent damage types. Additionally, it is recommended that damage categories be further subdivided into more detailed classifications. For example, the category of patching, which includes various forms of damage, could be divided into subcategories such as pothole patching, side ditch patching, excavation patching, and temporary patching. By expanding the categories in this manner, the overall recognition accuracy can be improved, resulting in more precise identification outcomes.

6. REFERENCES

Construction and Planning Agency, Ministry of the Interior. (2002). 柔性鋪面損壞調查手冊.

ASTM International. (2011). Standard practice for roads and parking lots pavement condition index surveys (D6433-11).

Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017). Understanding of a convolutional neural network. In 2017 international conference on engineering and technology (ICET) (pp. 1-6). IEEE. https://doi.org/10.1109/ICEngTechnol.2017.8308186

Jozefowicz, R., Zaremba, W., & Sutskever, I. (2015). An empirical exploration of recurrent network architectures. In International conference on machine learning (pp. 2342-2350). PMLR.

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In Thirty-first AAAI conference on artificial intelligence.

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431-3440).

Bolya, D., Zhou, C., Xiao, F., & Lee, Y. J. (2020). Yolact++: Better real-time instance segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(4), 1134-1141. https://doi.org/10.1109/TPAMI.2019.2956511

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In Proceedings of the 14th International Joint Conference on Artificial Intelligence (pp. 1137-1143). Montreal, Canada.

Prediction and Variable Explanation of International Friction Index (IFI) using Machine Learning: A Case Study in Thailand

Punnawat Siripatthiti[1], Ponlathep Lertworawanich[2]

[1]Bereau of Road Research and Development, Department of Highways, Thailand, punna.si@doh.go.th
[2]Bereau of Road Research and Development, Department of Highways, Thailand, ponla.le@doh.go.th

ABSTRACT

Highways are an essential part of land transport and can incubate large traffic to rural areas. One way to assess pavement condition is to use a composite index called the International Friction Index (IFI). If the index value is lower than certain thresholds, it will trigger an investigation or intervention. However, the IFI is sensitive to measurement variability, based on the assumption of linear behavior, and dependence on look-up tables. Several studies have attempted to develop mathematical models to predict road surface conditions while addressing those limitations using linear regression and machine learning techniques. Those techniques could not yield high-class accuracy, i.e. they can only predict unhealthy road conditions. This study explores varieties of regression techniques ranging from regression models i.e. Multivariate Regression to machine learning i.e. Random Forest, Gradient Boosting, and Feedforward Neural Networks. The Feedforward Neural Network gives the best results with the R-squared value of 0.76 and the F1 score of 0.957 for the investigatory level. Then, the study illustrates how a model understands the importance of variables by using SHAP. It shows that the increasing number of AADT and 10-wheel trucks has a distinctive negative impact, conversely, MPD has the most positive impact on the IFI values. While the impact of most choices in categorical variables is not distinctive, the choice of using AC 60-70 seems to have a negative impact on the IFI values.

Keywords: Machine Learning, Deep Learning, Highway Engineering, International Friction Index, Pavement Management System

1. INTRODUCTION

Highways are an essential part of land transportation and can incubate large traffic in rural areas. The highway's pavement surface is expected to withstand repetitive traffic and accumulated fatigue; hence, it needs to be monitored frequently to provide and sustain pavement in a condition acceptable with safety and reliability to the road users at the least life cycle cost (AASHTO, 1990). One way to assess pavement condition is to use a composite index such as the International Friction Index (IFI). The IFI was developed by Wambold et al., 1995. It aims to harmonize friction measurements with different equipment to a common calibrated index and can be used for comparisons internationally. The IFI combines two parameters (1) macrotexture, and (2) measured friction (FRS) on wet pavement. Measurement of the pavement macrotexture is used to estimate the speed constant (Sp). The measurement is represented by Mean Profile Depth (MPD) per ASTM Practice E 1845, i.e. high-speed laser measurement; or MTD per Test Method E 965, i.e. sand patch, if the first method is not available. Then the measured friction at some speed (S) is used with the speed constant of the pavement (Sp) to calculate friction at 60 km/h (FR60) then the calibrated friction (F60) is calculated from a linear transformation. The coefficient and constant in both parameters' calculations are dependent on the empirically determined device-specific coefficients which are predetermined in look-up tables to interpret the values (ASTM E1960, 1998). If the IF is lower than certain thresholds, it will trigger an investigation or intervention. The thresholds are the results of research by the Department of Highway, Thailand, and are estimated by identifying the relationship between accident rates per kilometer and IFI values. It should be noted that the thresholds are different from country to country due to different innate traffic characteristics (*คู่มือแนะนำ และวางแผนบำรุงรักษาผิวทางให้มีความเสียดทานผิวทาง (u) ที่เหมาะสม*, 2016).

The calculation poses some challenges such as its assumption based on the linear correlation that existed between measurements converted to the standard slip speed of 60 km/h (FR60) acquired from the devices that operate at 100% slip condition and the dynamic friction tester, however, the relation would be nonlinear if the devices operated at 10-20% slip (Fuentes & Gunaratne, 2011). Moreover, it is dependent on generalized and predetermined look-up tables, and it is sensitive to measurement variability i.e. different countries and different measuring devices present different values (Flintsch et al., 2009). Hence, the IFI exhibits some degree of subjectivity, which may impact its consistency across different contexts.

Another challenge of obtaining the IFI value is its measurement procedures that require the road section to be shut down, and it is time- and workforce-consuming. This challenge restraints the feasibility of the measurement procedures, thus, the procedures usually commence after the highway pavement survey has been reported of a certain number of road incidents or road damages. Consequently, the IFI dataset could be biased toward one type of road conditions (Fülöp et al., 2000).

Thailand has been regularly collecting road friction data, yet the dataset has not been used for developing a predictive model. This study aims to develop the best predictive model for the Thailand friction dataset with a focus on balancing performance on both healthy class and unhealthy class prediction. The prediction models will help road agencies to reduce time and cost for the pavement management system while direct measurements should be reserved for those with critical IFI values. Additionally, the study explores an explainable tool to understand how each attribute impacts the model prediction which could help road agencies understand the deterioration of the road surface.

The remainder of the paper is organized as follows: in Sec. 2, the related work presents existing methodologies and their limitations. In Sec. 3, the methodology that is used in this study. The results are reported in Sec. 4 and conclusions are drawn in Sec. 5.

2. Related Work

As a part of the analysis scheme in the pavement management system (PMS), there are several attempts to develop algorithms used to interpret data in meaningful ways. With the rise of computer-based optimization algorithms, researchers can develop more sophisticated predictive models (Katorimzadeh & Shoghli, 2020). There are several studies that attempt to develop mathematical models to predict road surface conditions. In Marcelino et al., (2017), linear regression is used to predict friction number or skid number from variables like average monthly temperature, total monthly precipitation, accumulated traffic, international roughness index, and initial friction number from Long-Term Pavement Performance (LTPP) open dataset. The best parameters yield the R^2 of 0.583. Fülöp et al., 2000 proposed a probabilistic model like Markov chains rather than a linear equation that takes similar parameters as described in the previous section. The data was collected on the Hungarian asphalt-concrete motorways system. However, these models couldn't yield a satisfactory result, so the study is not using these models.

2.1 Model Selection

This study will use four predictive models to achieve the best IFI prediction results. Those models are the multivariant regression model, ensemble decision tree; Random Forest and Gradient Boosting decision tree, and Feed Forward Neural Network (FNN). The study will try to optimize each model to its best performance (except the multivariant regression).

Ensemble decision tree methods are machine learning algorithms that combine multiple decision trees to create a more powerful predictive model. The main idea behind ensemble methods is that a group of weak learners can form a strong learner. Examples of ensemble decision tree methods include Random Forest and Gradient Boosting. Random Forest creates a set of uncorrelated decision trees and averages their predictions, which helps to reduce the variance and prevent overfitting. Each tree in the ensemble is built from a sample drawn with replacement from the training set. Gradient Boosting is also a powerful machine learning technique that builds an ensemble of decision trees in a sequential manner. Each tree in the sequence is designed to correct the mistakes of its predecessor. The algorithm uses the gradient of the loss function to guide the construction of each new tree so that each additional tree takes a step in the direction that minimizes the loss (Bashar & Torres-Machi, 2021).

Feedforward Neural Networks (FNNs) are the simplest type of artificial neural network (ANN). In an FNN, information moves in only one direction, i.e. forward, from the input layer, through any hidden layers, to the output layer. Each neuron in a layer is connected to all neurons in the previous layer and all neurons in the next layer, forming a fully connected network. The weights of all nodes are calibrated when the model is being trained in order to minimize the loss between the predicted values and the actual values (Basnet et al., 2023). Furthermore, the architecture of neural networks allows researchers to expand their architecture to be deeper (Haddad et al., 2022) and more complex like autoencoder (Liu et al., 2022).

2.2 Interpretability

The final part of this study is going to explore the explainability of each model. For the first model: multivariate regression, the model only shows coefficient values which only express how many degrees each attribute influences the model prediction. The same goes for the random forest model which has a function to inspect the feature importance of each attribute but doesn't which direction the attribute influences. Thus, this study requires an explainability tool that has additive explanations, i.e. offering a straightforward interpretation of which direction that individual feature impacts, appealing visualizations, and consistency across models, i.e. can be used on top of different model architecture trained on the same dataset. This study uses Shapley Additive Explanations, (SHAP) (Lundberg & Lee, 2017).

SHAP is an explanation method for machine learning models that aims to assign responsibility for the model's predictions to individual input features. It uses concepts from cooperative game theory, specifically Shapley values, to provide interpretable explanations of model outputs. To simplify, features are matched with "players" in a game, while the model function is matched with the game rules. Then SHAP values represent how much each feature contributes to the overall prediction. The core concept is to calculate the expected model output when a feature is included versus excluded and compare those expectations to determine the features' contributions.

3. Methodology

This section will be divided into 4 parts: data exploration, data preparation, model selection, and training setup. The first part is exploring data attribution and its distributions, the second part is explaining data preparation, and finally, six models will be optimized to find the best predictive model.

3.1 Data Exploration

The dataset used in this study was acquired from field data across Thailand in a total of 42,346 data points. It is important to note that the dataset is considered tabular data, and it is not a time-series data. The distribution can be seen in Figure 1. The overall distribution is skewed right, the median is 0.27 and the maximum is 0.73. The skewed distribution of the IFI came from the challenge of obtaining it as mentioned before. Moreover, the percentage of data points below the intervention threshold is 85% while only 2% passes the investigatory threshold and would be considered a healthy road condition. The IFI thresholds are different from country to country. This is due to distinctive innate traffic characteristics. This research employs the plot between accident rate per kilometer and IFI value to identify the thresholds. To mitigate the problem of an imbalanced dataset, only the intervention threshold is used to classify unhealthy and healthy road conditions in this study.

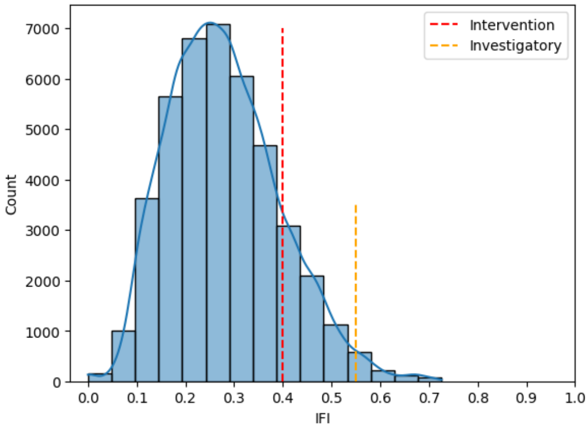


Figure 1 Dataset distribution.

The attributes in the friction database are listed in Table 1which is in accordance with (ASTM E1960, 1998). The attributes consist of two categorical variables; the road surface type and the road pavement aggregate type, and the other nine variables are numerical variables which represent the mean profile depth of the road (MPD), the age of the road, the number of lanes, the number of annual average daily traffic (AADT), and the number of trucks of each type per day. The acquiring method of MPD and AADT are in (ASTM E1845, 2015; FHWA, 2018) respectively.

Table 1 Attributes in pavement friction database.

Attribute	Type	Unit	Description
MPD	Numerical	Millimeter	Mean Profile Depth
Surface Type	Categorical	-	PMA, AC 60/70, Para AC, Para Slurry Seal, Slurry Seal
Age	Numerical	Days	Age of pavement in the days (at the time of observation)
No_lane	Numerical	Lanes	Number of pavement section (both directions)
Aggregate Type	Categorical	-	Basalt, Limestone, Andesite
AADT	Numerical	Vehicles/day	Annual Average Daily Traffic
Truck 4WH	Numerical	Trucks/day	Number of 4-wheel truck per day
Truck 6WH	Numerical	Trucks/day	Number of 6-wheel truck per day
Truck 10WH	Numerical	Trucks/day	Number of 10-wheel truck per day
Truck semi-trailer	Numerical	Trucks/day	Number of semi-trailer truck per day
Truck Trailer	Numerical	Trucks/day	Number of trailer truck per day
IFI	**Numerical**	**-**	**International Friction Index**

Polymer-modified asphalt (PMA), asphalt concrete (AC_60_70), and para-asphalt concrete (Para_AC) are used either in overlayed asphalt in road construction or in pavement rehabilitation. The slurry seal and para slurry seal are used to repair damage on service pavement surfaces as either a preparatory treatment for other maintenance treatments or as a wearing course. The majority of roads are using AC 60-70 as surface, as it is the most economical; and limestone as aggregate, because it is the most accessible (*คู่มือแนะนำ และวางแผนบำรุงรักษาผิวทางให้มีความเสียดทานผิวทาง (μ) ที่เหมาะสม*, 2016).

The data attribute selection is based on data availability. Thus, there are a few attributes that were used in other literature but not in this study such as pavement thickness, ESAL, surface temperature, precipitation rate, and humidity (Santos et al., 2021; Ziari et al., 2016). This study doesn't have access to these attributes, or the acquired data is not complete compared to the other attributes. Hence, this study dismisses the incomplete feature and remains with attributes as shown in Table 1.

When exploring how each surface type and aggregate type can influence the unhealthy class (intervention) in Figure 2, it appears that the slurry seal is the only surface type that has the most significant influence. This is because the slurry seal is used to maintain or improve the road surface then after this implementation, the IFI values are collected causing the number of IFI values to be higher than the other types.

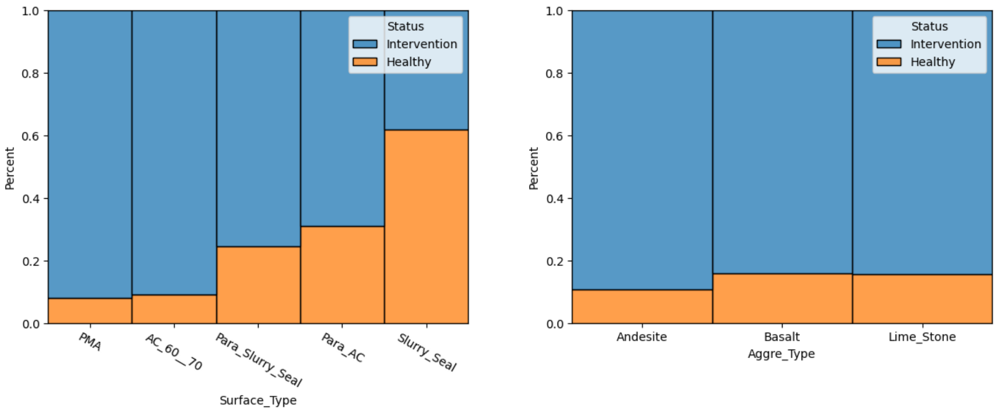


Figure 2 Distribution plots comparing between unhealthy and health road in each surface type and aggregate type.

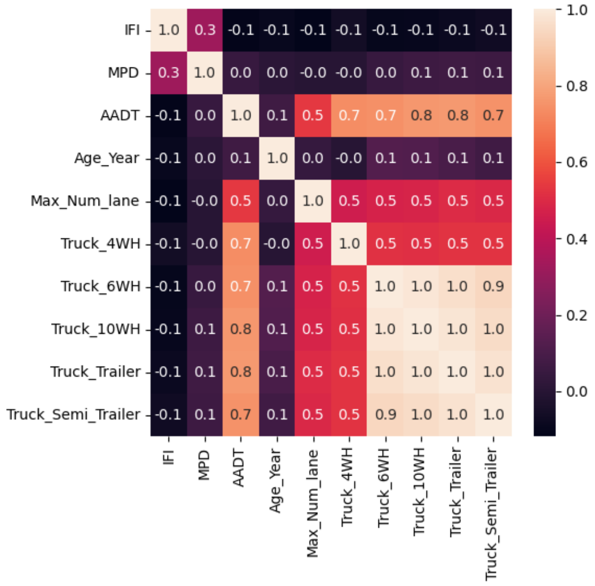


Figure 3 Correlation matrix heat map of continuous variables; IFI, MPD, ADDT, Age_Year, Max_Num_lane, and Truck_* variables

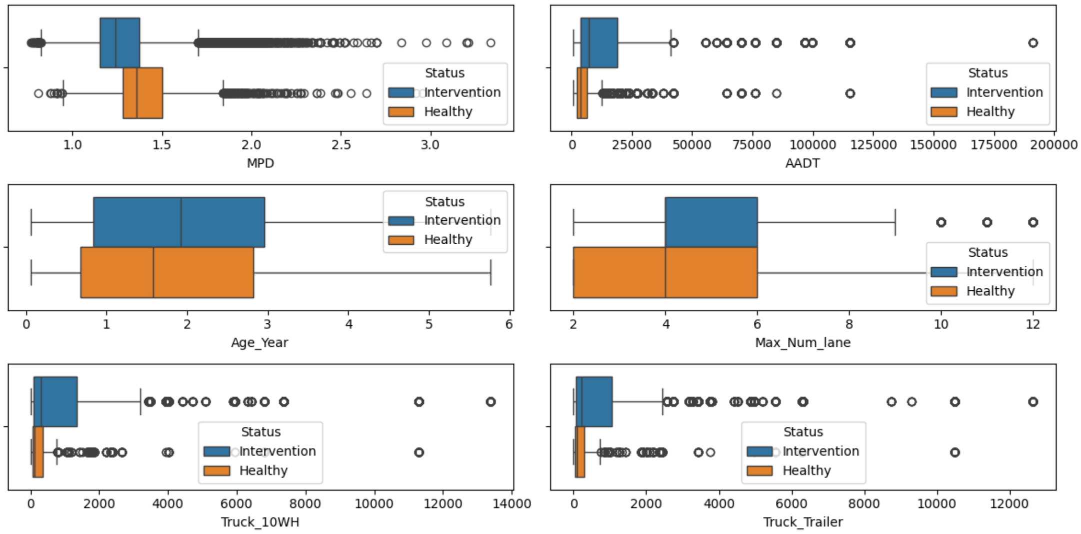


Figure 4 Box plots of continuous variables; IFI, MPD, ADDT, Age_Year, Max_Num_lane, Truck_10WH, and Truck_Trailer

The correlation matrix in Figure 3 shows that each variable does not have a strong linear relationship with the IFI value. Except for ADDT which implies traffic in a similar way as variables with the prefix "Truck_" and Max_Num_lane. Moreover, box plots in

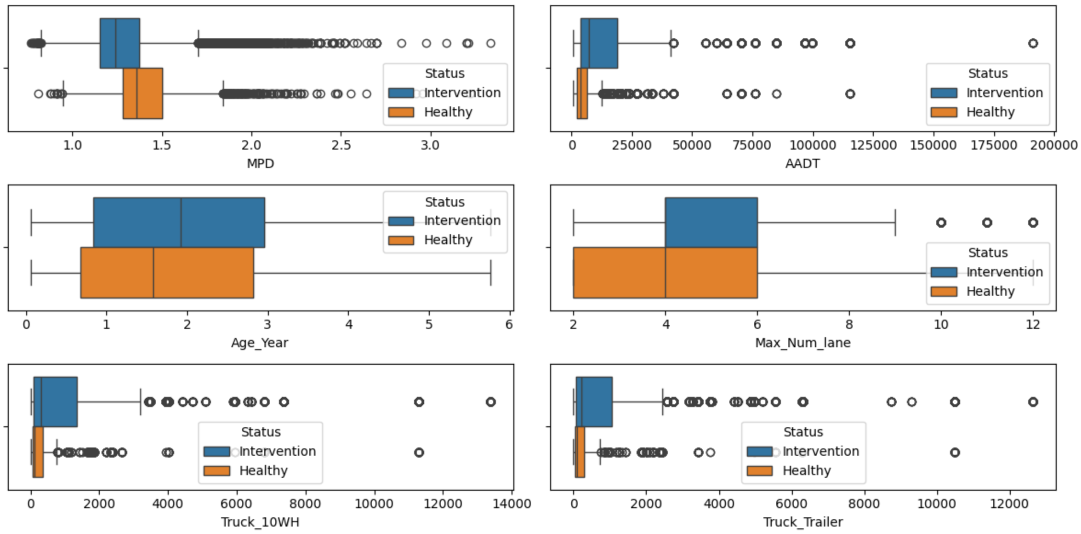


Figure 4 show a statistically insignificant difference between the healthy class and the intervention class of each continuous variable. Hence, these figures confirm that a regression model that based on linear assumption might not be suitable for predicting the IFI value.

3.2 Data Preparation

In the data preparation process, the raw dataset is cleaned by removing null data or duplicated data. The categorical variables are transformed into dummy variables, so in total there are 15 variables. After that, the numerical variables are scaled by subtracting with mean and dividing by the standard deviation. Next, the dataset is partitioned into train and test datasets by 80:20 proportion. Importantly, each partition is stratified to ensure that it always maintains the proportion between the unhealthy class and the healthy class. In each training epoch, the models are trained on the train dataset and then evaluated on the validate dataset. Finally, the model is run on the test dataset to collect the results.

3.3 Training Setup

The study trains the models on the Google Colab using Python. Random Forest, Gradient Boosting, FNN, and FNN + ResBlock have optimized their hyperparameters on the train dataset with stratified 5-fold cross-validation to find the lowest MSE Loss.

The predictive model will predict the IFI value then they are converted to each class responding to the intervention threshold. The reason why the models do not directly predict healthy class and unhealthy class because the intervention threshold is arbitrary and often influenced by the fiscal budget. This approach will allow a more flexible interpretation.

3.4 SHAP Setup

After training the Gradient Boosting model and FNN model, both models are validated with SHAP. Unfortunately, the SHAP library in Python does not support the random forest model. The training dataset is used in the SHAP explanatory function. However, due to the space complexity, only 1.5% of training dataset can be used in the FNN model, but the whole the train dataset can be used in the Gradient Boosting model. Hence, the study is going to demonstrate the usage of SHAP only from the trained Gradient Boosting model.

4. Results

The results in

Table 2 show experiment results. The best performance model is the feed forward neural network (FNN) model with an accuracy of the healthy class of 69.2% and R^2 of 0.76. Followed by the Random Forest model with an accuracy of the healthy class of 66.6% while there is no significant difference in R^2.

The worst performance model is the multivariate regression model with an accuracy of the healthy class of 12.2% and R^2 of 0.23 although the model almost has the most accuracy on intervention class. This implies that the model is biased toward unhealthy road conditions even though the dataset is not severely imbalanced. The results from multivariate regression agree with the assumption that the dataset is unfit for the linear assumption

Table 2 Experiment results of each predictive model on test dataset.

Method	**Accuracy on Healthy**	**Accuracy on Unhealthy**	**F1 score**	**R^2**
Multivariate Regression	12.2%	<u>98.6%</u>	0.92	0.23
Random Forest	66.6%	97.0%	<u>0.96</u>	0.75
Gradient Boosting	63.2%	97.3%	0.95	0.74
FNN	**69.2%**	96.8%	<u>0.96</u>	**0.76**

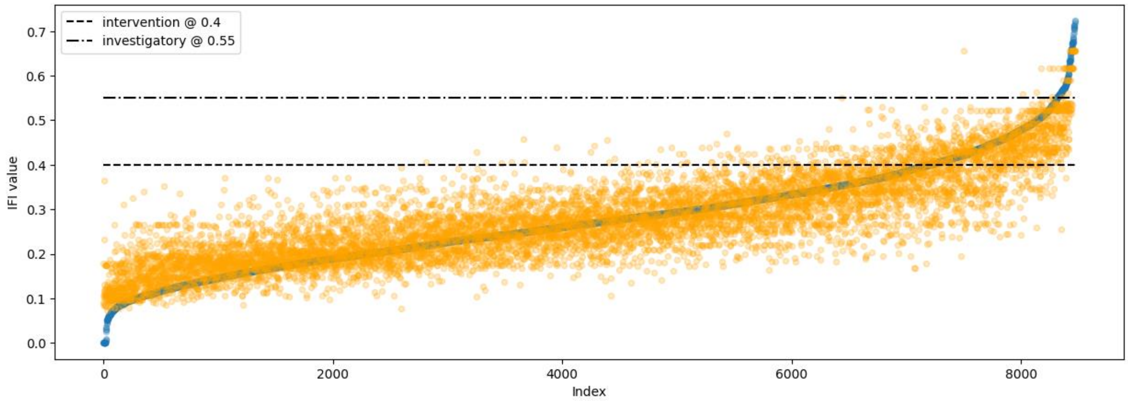


Figure 5 comparison between true values and prediction values from FNN on test dataset.

Next, the study arranges actual IFI values (blue dots) of the test dataset ascendingly, then plots their corresponding predicted values (orange dots) to visualize the prediction performance. The result will be like in Figure 5. Most predicted values are clustering around the blue dot line. Notice that the predictions of the lower index data points (below 40,000) are overpredictions and the upper index data points are vice versa. Additionally, the over-predicting values mostly not exceed the intervention threshold line.

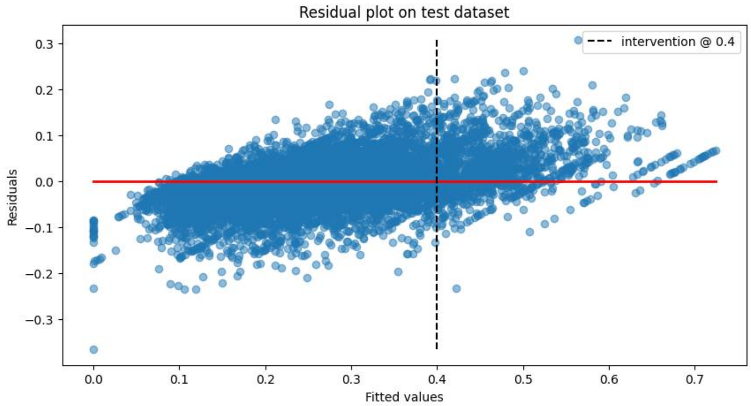


Figure 6 residual plot from FNN on test dataset.

Figure 6 shows the residual plot of the results from the FNN model on the test dataset. The residual is calculated from the difference between the actual value and the predicted value. The overall distribution of the residuals shows a normal distribution. However, when dividing the residual plot into four tiles based on the intervention threshold and no residual point. The top tiles show that the model is under-predicting while the bottom tiles are over-predicting. The overprediction of intervention values, bottom left tile, could compromise the integrity of the road surface. Nonetheless, the cluster of over-predicting intervention values is not bigger than the others.

The expected model output is a sum of a base value, the expected model output without considering any attribute, and SHAP values from all attributes. To illustrate, in Figure 7 shows a force plot of summation from significant SHAP values from attributes. The base value is 0.28 plused positive SHAP values (red bars), then the expected model output is 0.67 (the actual IFI value is 0.66). Another example is in Figure 8 where the same base value plus with positive SHAP values (red bars) and negative SHAP values (blue bars) yields the expected model output as 0.26 (the actual IFI value is 0.26). One significant difference between these two examples is the rocketing increase in AADT and Truck_10WH in the later example.

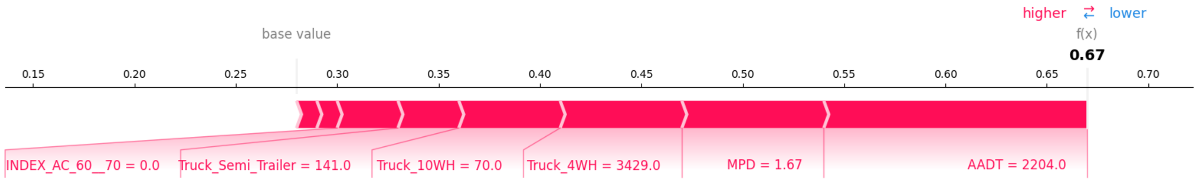


Figure 7 force plot from SHAP values on the prediction of actual IFI value 0.66.

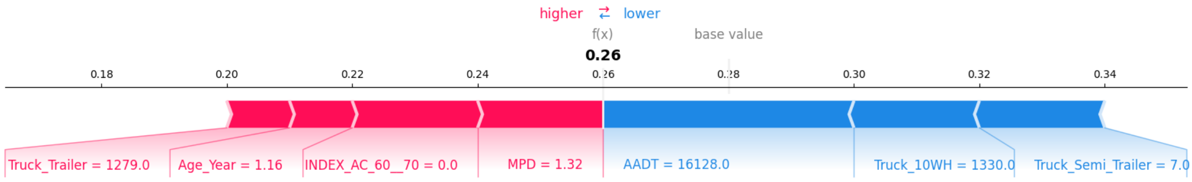


Figure 8 force plot from SHAP values on the prediction of actual IFI value 0.26.

The combination of all SHAP values from every predicted data point is shown in Figure 9. The right color bar shows the feature value of each attribute meaning that a red dot at AADT row implies a higher value of AADT than blue dots. There are three highlights in the figure. First, the three most important attributes are AADT, MPD, and Truck_10WH. Both AADT and Truck_10WH are similar where clusters of red dots (the data points with a high value) have negative SHAP values while blue dots dissipate toward positive SHAP values. MPD, on the other hand, has vivid bi-polarised clusters where the higher the MPD values show the higher the SHAP values. This is agreeable with the correlation between MPD and IFI shown in Figure 3.

The same pattern also appears on INDEX_AC_60_70. A red dot on its row implies a choice of using asphalt concrete 60/70 for the surface type and it gives a negative SHAP value. The last highlight is the other categorical variables. It seems that the choices of choosing any other surface type and aggregate type do not matter much compared to the choice of using asphalt concrete 60/70 as the surface type. This insight explanation is beyond what Figure 2 illustrates.

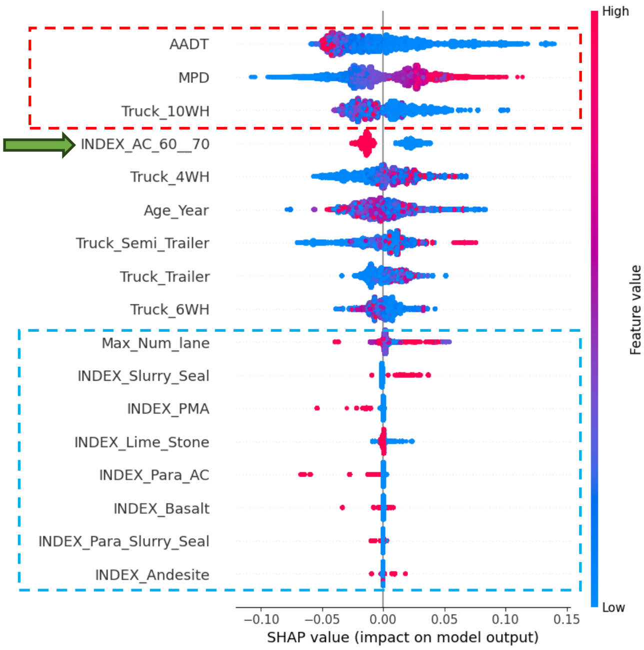


Figure 9 summary plots of SHAP values on gradient boosting.

5. Conclusion

In conclusion, this study has demonstrated the potential of machine learning models in predicting the International Friction Index (IFI) for highway pavements in Thailand. The Feedforward Neural Network (FNN) model outperformed other models, including Random Forest, Gradient Boosting, and Multivariate Regression, in terms of accuracy for the healthy class and R^2 score. The FNN model achieved an accuracy of 69.2% for the healthy class and an R^2 score of 0.76, indicating a strong correlation between the predicted and actual values.

The reason why

Table 2 shows the accuracy of the healthy road condition is that the best predictive model would be used for giving decision support. If the model produces a large number of false positive predictions, then the road authority must spend time and resources to conduct a maintenance scheme. Moreover, maintenance is often required to close the road section, which will propagate traffic delays to the whole road network.

However, it's important to note that while the FNN model performed well overall, it tended to overpredict for lower index data points and underpredict for higher index data points. This discrepancy could potentially compromise the integrity of the road surface if not properly accounted for. Despite this, the FNN model's performance suggests that it could be a valuable tool for predicting pavement conditions and informing maintenance decisions.

Finally, SHAP is a powerful tool to illustrate how important each attribute is to the model prediction. It shows that Annual Average Daily Traffic (AADT), Mean Profile Depth (MPD), and number of trucks with 10-wheel (Truck_10WH) have the most importance to the model prediction. Additionally, SHAP also gives an insight beyond descriptive statistic capabilities. It shows that the choice of using asphalt concrete 60/70 (INDEX_AC_60_70) has a negative impact on the model prediction.

6. Discussion

This study uses the IFI dataset that was acquired only one year and lacks a broader representation. The future study should include attributes like pavement thickness, surface temperature, precipitation rate, etc. to effectively utilize SHAP. Furthermore, the dataset should expand to at least three years to perform future prediction models.

The results of this study highlight the importance of selecting the right model for the right task. The IFI is a complex index that depends on a variety of factors, and the models used in this study may not fully capture these complexities although the performance is promising. Future research could explore other machine learning models or feature engineering techniques to improve prediction accuracy.

Although the best performer is FNN, the time complexity is exponentially longer than the RF and GB while training. In addition, the SHAP function makes the model less feasible in explainability tasks. The future study should consider another explainable AI tool that allows the FNN model to be understood.

Finally, this study contributes to the ongoing efforts to automate and improve the monitoring of highway pavements. By developing a predictive model for the IFI, this study provides a tool that could potentially save time, reduce costs, and improve road safety. However, further work is needed to validate and refine the model before it can be used in practice.

7. REFERENCES

AASHTO. (1990). *AASHTO Guidelines for Pavement Management Systems*. American Association of State Highway and Transportation Officials. https://books.google.co.th/books?id=TRT9GwAACAAJ

ASTM E1845. (2015). *Standard Practice for Calculating Pavement Macrotexture Mean Profile Depth*. https://www.astm.org/e1845-15.html

ASTM E1960. (1998). *Standard Practice for Calculating International Friction Index of a Pavement Surface*. https://doi.org/10.1520/E1960-98

Bashar, M. Z., & Torres-Machi, C. (2021). Performance of Machine Learning Algorithms in Predicting the Pavement International Roughness Index. *Transportation Research Record*, *2675*(5), 226–237. https://doi.org/10.1177/0361198120986171

Basnet, K. S., Shrestha, J. K., & Shrestha, R. N. (2023). Pavement performance model for road maintenance and repair planning: A review of predictive techniques. *Digital Transportation and Safety*, *2*(4), 253–267. https://doi.org/10.48130/DTS-2023-0021

FHWA. (2018). *Traffic Data Computatoin Method Pocket Guide*. Federal Highway Admisnistration.

Flintsch, G., Izeppi, E., Mcghee, K., & Roa, J. (2009). Evaluation of International Friction Index Coefficients for Various Devices. *Transportation Research Record*, *2094*, 136–143. https://doi.org/10.3141/2094-15

Fuentes, L., & Gunaratne, M. (2011). Revised Methodology for Computing International Friction Index. *Transportation Research Record: Journal of the Transportation Research Board*, *2227*, 129–137. https://doi.org/10.3141/2227-14

Fülöp, I. A., Bogárdi, I., Gulyás, A., & Csicsely-Tarpay, M. (2000). Use of Friction and Texture in Pavement Performance Modeling. *Journal of Transportation Engineering*, *126*(3), 243–248. https://doi.org/10.1061/(ASCE)0733-947X(2000)126:3(243)

Haddad, A. J., Chehab, G. R., & Saad, G. A. (2022). The use of deep neural networks for developing generic pavement rutting predictive models. *International Journal of Pavement Engineering*, *23*(12), 4260–4276. https://doi.org/10.1080/10298436.2021.1942466

Karimzadeh, A., & Shoghli, O. (2020). Predictive Analytics for Roadway Maintenance: A Review of Current Models, Challenges, and Opportunities. *Civil Engineering Journal*, *6*, 602–625. https://doi.org/10.28991/cej-2020-03091495

Liu, J., Liu, F., Zheng, C., Fanijo, E. O., & Wang, L. (2022). Improving asphalt mix design considering international roughness index of asphalt pavement predicted using autoencoders and machine learning. *Construction and Building Materials*, *360*, 129439. https://doi.org/10.1016/j.conbuildmat.2022.129439

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

Marcelino, P., de Lurdes Antunes, M., Fortunato, E., & Gomes, M. C. (2017). Machine Learning for Pavement Friction Prediction Using Scikit-Learn. In E. Oliveira, J. Gama, Z. Vale, & H. Lopes Cardoso (Eds.), *Progress in Artificial Intelligence* (pp. 331–342). Springer International Publishing. https://doi.org/10.1007/978-3-319-65340-2_28

Santos, A., Freitas, E. F., Faria, S., Oliveira, J. R. M., & Rocha, A. M. A. C. (2021). Prediction of Friction Degradation in Highways with Linear Mixed Models. *Coatings*, *11*(2), Article 2. https://doi.org/10.3390/coatings11020187

Wambold, J. C., Antle, C. E., Henry, J. J., & Rado, Z. (1995). *International PIARC Experiment to Compare and Harmonize Texture and Skid Resistance Measurements*. PIARC. https://urn.kb.se/resolve?urn=urn:nbn:se:vti:diva-18200

Ziari, H., Sobhani, J., Ayoubinejad, J., & Hartmann, T. (2016). Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods. *International Journal of Pavement Engineering*, *17*(9), 776–788. https://doi.org/10.1080/10298436.2015.1019498

คู่มือแนะนำ และวางแผนบำรุงรักษาผิวทางให้มีความเสียดทานผิวทาง (น) ที่เหมาะสม. (2016). [Manual guide]. Bureau of road research and development, Department of Highways, Ministry of Transport.

Real-time Traffic Data Collection at Mid-block and Intersection by UAVs and Computer Vision

Ponlathep Lertworawanich [1,*], Peerapol Sittivijan [2] and Theerada Rungruangjarensuk [1]

[1] Department of Highways, Bangkok, Thailand, theer.ru@doh.go.th
[2] Burapha University, Chonburi, Thailand, peerapol@go.buu.ac.th

ABSTRACT

Given the recent rapid urban expansion, there is a growing need to focus on traffic surveillance and monitoring to provide authorities with better insights and be able to efficiently respond to traffic situations and incidents in near real-time manner. In order to real time traffic management, various kinds of sensors are needed to be install to collect traffic data. However, installation of these sensors is costly and, currently, there are areas in Bangkok not covered by these sensors yet. UAVs can address the currently lacking monitoring infrastructure system in these areas. In this study, we illustrated that by deploying Unmanned Aerial Vehicles (UAVs) coupled with computer vision techniques, it is possible to efficiently collect traffic data required for traffic management at both mid-blocks and intersections areas. We proposed We proposed a practical framework for urban traffic surveillance in real-world environment. From the experiments, parameters such as the height of UAV, camera angle and resolution can largely affect the accuracy of detection and tracking algorithms which can vary from 30% to 95%. The best parameter setting from the experiment enabling around 95% and 90% of accuracy at mid-blocks and intersections is provided.

Keywords: UAV, Traffic Monitoring, Computer Vision, Traffic Video Processing

1. INTRODUCTION

Unmanned aerial vehicles (UAVs) have proven to be valuable for automation and data collection. In terms of data collection, the installed onboard sensors and cameras outperform conventional methods in terms of efficiency and cost saving. UAVs have a diverse range of application such as in agriculture, package delivery, aerial photography, land surveying, and mapping. The UAV is pivotal in assisting in these fields.

In civil engineering, apart from land surveying and constructing digital twins, data acquired using UAVs is used to support decision-making processes (Kaewunruen et al., 2023). Transportation and traffic issues remain at the forefront of challenges, given their vital role in societal development and economic growth. That said, Bangkok, a mega-city with rapid urban expansion in the past decade, requires more attention to traffic surveillance and monitoring in response to the significant increase in demand for goods and transportation. Therefore, there is an urgent need to develop a reliable traffic monitoring system. However, given the lack of infrastructure in some areas, the main challenge is that, in some areas, infrastructure has not been able to keep up with demand, leading to drivers over speeding in unmonitored area (Bashir et al., 2022). The highway network has limited pre-installed loop detectors; even the CCTV camera monitoring system is not installed in some areas.

The Department of Highways aims to deploy UAVs for traffic monitoring, enabling the collection of real-time data across highways and local roads during the day. This approach also proves advantageous in real dynamic traffic scenarios, which are challenging to capture even with modern traffic simulation software. This will ultimately lead to more informed and accurate decision making to support traffic control and transportation authorities. This data could help mitigate traffic congestion, facilitate rerouting plans in case of incidents, expedite emergency responses, positively impact travel behavior, and increase highway-network efficiency (Kanistras et al., 2013; Khan et al., 2020).

Advances in computer vision, convolutional neural networks, and advanced deep learning, have enabled major state-of-the-art solutions to engineering problems. Object detection and tracking algorithms are especially significant in this context, as they enable the monitoring and analysis of moving objects within the footage captured by UAVs. For example, accurately extracting vehicle speed (Guido et al., 2016), analyze confliction zone (Chen et al., 2020), conducting behavioral studies (Barmpounakis et al., 2020). Relevant algorithms contribute to the accurate extraction of crucial traffic parameters and characteristics (Ke et al., 2015).

While numerous studies focused on improving the capabilities of deep-learning models in vehicle detection and tracking, UAVs are still vulnerable to some environmental conditions such as high wind speed, poor lighting, loss of GPS signal, and environmental hazards in certain areas.

This study comprehensively focused on both the practical framework as well as implementation of an object detection and tracking model. The designated study area presents a complex scenario characterized by multiple types of vehicles mixed in traffic; detectors in previous studies proved inefficient in capturing non-homogeneous traffic flows.

The remainder of this paper is structured as follows: Section 2 briefly discusses the evolution of traffic parameter extraction from a mathematical model to non-conventional methods such as UAV-assisted methods. Previous studies that use different approaches to machine learning as well as deep learning are explored as well. Section 3 describes how this study is set up and conducted. Explaining the new dataset, UAV configuration, as well as YOLOv8 object detector and BoT-SORT tracker. Section 4 presents the evaluation performance of the method proposed in this study and discusses the results and notable findings. Section 5 concludes the paper and addresses the limitations of this approach.

2. Related work

The research effort prior to UAV-assisted traffic monitoring and control mainly involved mathematical models. The formulated models, such as Greenshields, Underwoods, Greenbergs, Daganzos, are used to determine traffic characteristics such as density, flow, and speed. Nevertheless, the accuracy of these models varies depending on specific road conditions, traffic flow pattern, as well as saturation level. To satisfy these model requirements, large datasets and continuous data collection such as pre-installed detectors to monitor at least two of the three variables are needed, which is expensive in both data collection and analysis.

Salvo et al. (2014) found that some of the traditional mathematical models have significantly large error ranges. On the contrary, Greenshields can accurately represent traffic conditions with an error rate of 0.50%. It is noteworthy that this study was conducted in the city of Palermo, different study areas have different traffic patterns and driver behaviors.

Several studies explored the potential use of small quadcopters and UAVs in traffic and transportation. Outay et al. (2020) conducted a comprehensive literature review and classified the potential use case into three main scenarios: road safety, traffic monitoring and management, and highway infrastructure management. Butila and Boboc (2022) surveyed the scientific contributions specifically related to traffic monitoring. They found a growing trend in attempts to develop recognition and tracking of moving vehicles, which remains a challenge. Most studies either focus on traffic monitoring such as vehicle detection and tracking or traffic analysis such as congestion analysis and traffic parameter extraction.

Ke et al. (2017) proposed a framework for extracting traffic parameters from aerial video footage. The proposed system identifies the directions of traffic streams and extracts the flow for each stream using four steps: Kanade–Lucas–Tomasi (KLT) tracker, k-mean clustering, connected graphs, and traffic flow theory. While the first two components are used for interest-point-based motion analysis, four constraints were proposed to determine the connectivity of points belonging to one traffic stream cluster. Then the traffic parameters can be calculated based on the previous steps. Khan et al. (2017) proposed an extensive and systematic framework for optimal application of UAVs that streamlined the extraction of vehicle trajectories of multiple vehicles on a specific road corridor from UAV-acquired data. The framework consists of five main modules. After obtaining UAV footage, the preprocessing module is initiated to perform trimming, image rectification, and masking region of interest (ROI), then stabilization, georegistration, vehicle detection and tracking using optical flow approach, background subtraction and blob analysis, and lastly, managing trajectories in storage.

Other studies focused on developing various computer vision algorithms as well as using existing models to classify and track vehicles. Xu et al. (2017) deployed a Faster-RCNN for detection tasks from a low-altitude UAV, achieving acceptable accuracy along with high detection speed. Niu et al. (2018), Ke et al., (2018), Ke et al.(2020), and Kyrkou et al.(2018) also presented hybrid models combining the Haar Cascade algorithm and convolutional neural network (CNN) to improve vehicle detection robustness and extracted traffic parameters such as count and speed. Zhu et al. (2018) proposed an advanced deep-neural-network model that demonstrated the superiority of deep-learning approaches over traditional vision-based algorithms as well as the effect of ultra-high-resolution video in enhancing the accuracy of vehicle analysis.

Zhao et al. (2018) and Barmpounakis (2019) assessed the limitations and challenges of UAVs such as instability, data transmission, and flight altitude. Their work underlies having a robust data collection technique such as camera stabilization and geo-referencing to improve the reliability of the collected data.

Zhang et al. (2019), Liu et al. (2019), Zhu et al. (2018), and Ahmed et al. (2020) used advanced deep-learning models and R-CNN models for detection and tracking tasks. The presence of multiple types of vehicles resulted in high errors in vehicle counts.

Adaimi et al. (2023) used UAVs to monitor high-resolution aerial images. While this condition is still a challenging task for object-detector models, they developed the Butterfly detector, an anchor-free method that requires a composite field and describes spatial features as well as the scales of detected objects. The algorithm uses a voting mechanism to locate the object center. It was tested on two public datasets UAVDT and VisDrone2019 and a new roundabout dataset with desirable accuracy.

2.1 Contribution

This study addresses several gaps in the field of traffic surveillance and parameter extraction using UAVs by building upon the existing research work. The main contributions of this study can be summarized as follows.

1. Enhanced detection and tracking accuracy with variation in vehicle types: While most previous studies have mainly focused on low-altitude video footage and passenger cars, this study recognizes the importance of a more practical and comprehensive perspective by extending the scope to collect a wider range of vehicles, including motorcycles and heavy vehicles, which are the main challenge to previous research.

2. Comprehensive framework for practical implementation: This study introduces a novel framework that includes deploying drones with longer flight times, capable of capturing high-resolution footage as well as deployment of a more robust algorithm in vehicle detection and tracking, and integrated traffic analysis to analyze possible congestion and parameter extraction.

3. Optimized altitude and camera angle: This study examines the influence of the UAV altitude and different camera angles on the efficiency of vehicle detection and tracking models. Different combinations of altitudes and angles are systematically investigated to identify the optimal settings for suitable data collection.

3. MATERIAL AND METHODS

The materials and methods or experimental section provides specific details of how the research was conducted.

3.1 Overview

The general framework of the UAV-assisted monitoring system that is proposed in this study is illustrated in Figure 1. The framework can be categorized into three different layers. The first layer of the framework covers data processing from raw UAV footage, with a built-in stabilizer *(gimbal)* to help with video processing (Salvo et al., 2017) through the trained detection model to classify and count vehicles. In the second layer, after the vehicle composition has been determined, tracking is conducted to collect data of each vehicle in any frame instance to collect travel time when the vehicle enters the predetermined regions of interest (from upstream to downstream) and extract the space-mean speed. The final layer incorporates traffic flow theory, using the time-space diagram and cumulative diagram to estimate queue length and delay (Khan et al., 2018). The detailed process is discussed in this section.

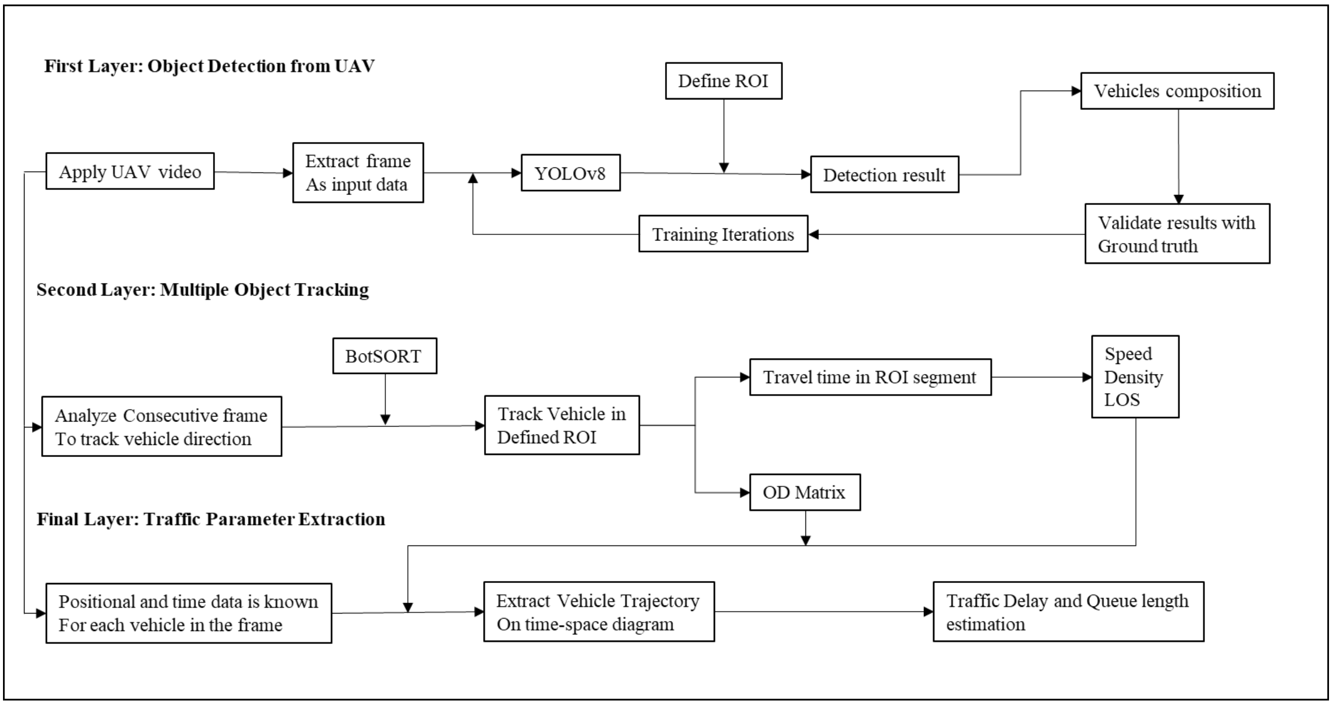


Figure 1. Core 3-layer framework for UAV-assisted traffic monitoring.

3.2 Data Collection

This study selected 31 distinct study sites, each chosen based on a combination of factors such as service level, highway functionality, surrounding geography, and built environment. These sites represent a variety of transportation infrastructure types, including highway corridors, freeways, intersections, and multi-level interchanges.

The data collection process involved capturing turning movement counts (TMC) and mid-block classified counts (MBC). Supplementary field equipment (ground video camera, speed laser gun) was deployed to serve as the benchmark test for evaluating the efficacy of the UAV-assisted approach.

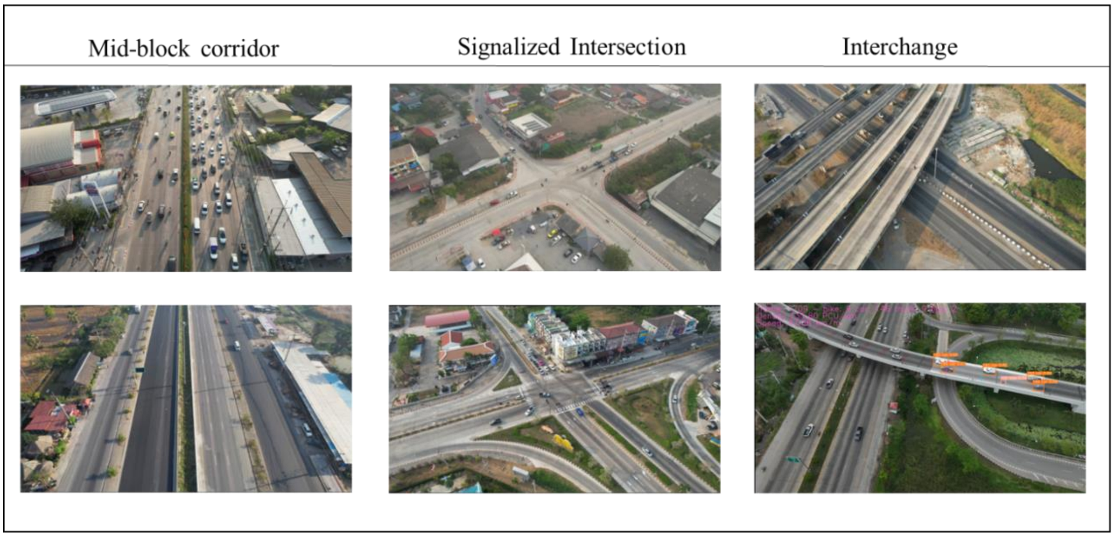


Figure 2. Designated study area.

It is noteworthy that the study concentrated on three predominant vehicle categories that are both commonly encountered in road networks of Thailand and suitable for analysis within the context of a CNN model. Specifically, the study focused on passenger cars, motorcycles, and heavy vehicles (trucks and buses). By narrowing the scope to these three major vehicle types, the study aimed to optimize the accuracy and efficiency of the model in detecting and classifying traffic flow patterns within the chosen study sites.

Table 1. Vehicle detection with ground data distribution

	Vehicle Count			Total	Weight Distribution		
	Passenger Car	Motorcycle	Heavy Vehicle		Passenger Car	Motorcycle	Heavy Vehicle
Multi-lane	102,785	27,230	6,654	136,669	0.752	0.199	0.048
Single-lane	1,605	252	61	1,918	0.837	0.131	0.032
Interchange	19,942	4,425	2,852	27,219	0.733	0.163	0.105
Intersection	3,335	1,899	143	5,377	0.620	0.353	0.026

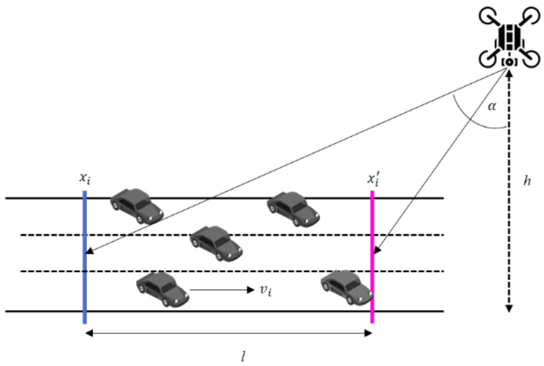


Figure 3. UAV-assisted traffic monitoring. l is effective corridor length, h is the aerial vehicle height from ground level, alpha is the camera angle perpendicular to ground level, xi is the designated upstream location, and xi' is the designated downstream location.

Table 2. Drone specification

UAV	Phantom 4 Pro	DJI Air2S
Flight time	30 min	31 min
Distance	7 km	12 km
Sensor	CMOS 1"	CMOS 1"
	4K/60 fps video	5.4K/30 fps
Battery	5,870 mAh	3,500 mAh
Video Resolution	3840 x 2160	5472 x 3078

To obtain the traffic surveillance footage from the UAVs, the data collection process conducted in this study involved 360 minutes of video footage for each designated site. For each flight, a minimum flight duration of 15 minutes was ensured before recharging the UAV. Additionally, all flights were within permissible airspace and clear of potential obstacles within the operational area.

Data collection was conducted during peak hours, specifically 8:00 AM to 10:00 AM, 11:00 AM to 1:00 PM, and 4:00 PM to 6:00 PM. This approach aimed to capture a comprehensive range of traffic scenarios and patterns during periods of peak demand.

During flight, the UAVs also employed stabilization mechanisms. A built-in gimbal system ensured steady and smooth footage acquisition under various conditions such as high wind speed in certain areas. Variability was introduced by conducting flights at three different altitudes: 50, 70, and 90 meters, as well as adjustments in the camera angle, with settings at 30, 45, 70, and 90 degrees. This approach allowed the acquisition of data from diverse perspectives, contributing to a more comprehensive understanding of the UAV and object detection and tracking model capabilities (Wang et al., 2016) as shown in Figures 4 and 5.

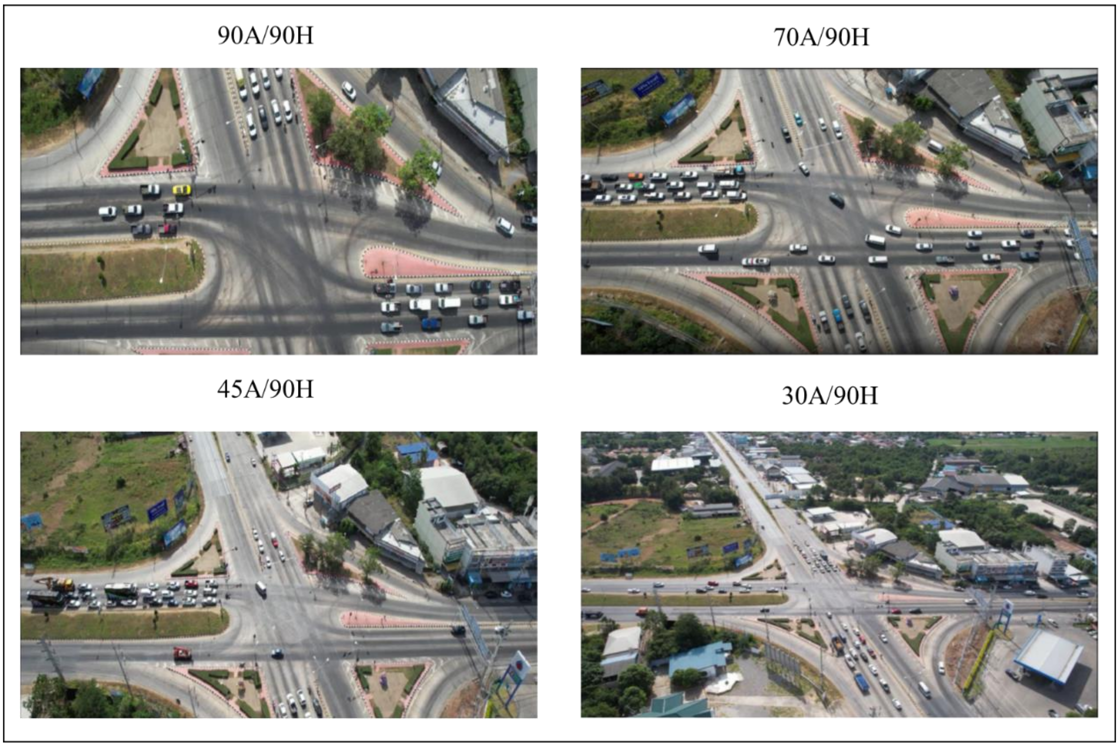


Figure 4. Different UAV camera angles at a 90-meter altitude hovering above signalized intersection for turning movement count, TMC

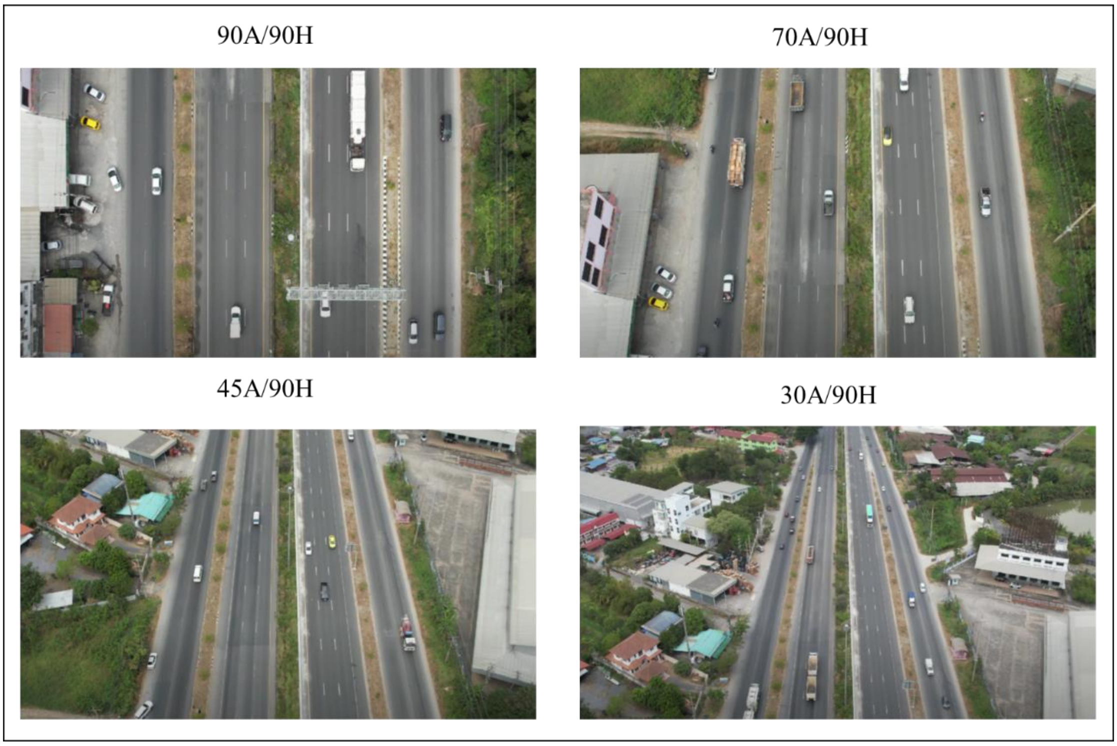


Figure 5. Different UAV camera angles at a 90-meter altitude hovering above a corridor section for mid-block classified counts, MBC

3.3 Object Detection (YOLOv8)

The You Only Look Once (YOLO) object detector was first developed by Joseph Redmon and Ali Fardahi in 2015. Originally, the YOLO model was based on a pyramidal hierarchy of deep convolutional neural networks, with a hierarchical top-down pyramid architecture designed to enhance the performance of CNN in handling more complex features and improve the overall model efficiency. Since then, there have been many improvements and iterations over its successors including YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv6, and YOLOv7 (Benjdira et al., 2019; Bozcan et al., 2020; Shizari et al., 2020).

In standard CNNs, features are typically captured at a single scale, which can significantly limit the capacity to recognize objects of varying sizes. Notably, YOLOv7 introduced the incorporation of the 'Focal Loss' function, which is designed to address class imbalance problem that proved to be challenge for previous study that is lacking different vehicle type in the data set. The focal loss function assigns greater weight to more complex examples and reduces the influence of the class that is predominant in vehicle type distribution. This approach aligns well with our detection tasks, given the imbalanced nature of our collected data in each category represented in Table 1

YOLOv8 represents the latest iteration of the well-established and accurate object detection model. The model backbone is based on *CSPDarknet53* architecture, which leverages cross-stage feature aggregation to amplify feature representation. The network is structured in a way that captures information at different levels of abstraction, from lower-level to higher-level features (Lin et al., 2017). A notable enhancement in YOLOv8 is the incorporation of anchor-free detection. This means the model predicts directly onto the center of the object, without relying on offsets from anchor boxes as seen in its predecessor. Anchor boxes are predetermined boxes with specific height and width ratios that frame the objects, their selection is determined by object sizes in the training data. Consequently, this leads to models that have greater flexibility and efficiency in detection tasks.

Therefore, YOLOv8 was fine-tuned to our collected vehicle datasets to increase its accuracy and efficiency. The proposed model serves as the pivot foundational framework for traffic surveillance detection and parameter extraction.

3.4 Object Tracking (BoT-SORT)

BoT-SORT is a state-of-the-art algorithm for object tracking. The goal of multi-object tracking is to detect and track all the objects in a scene while keeping a unique identifier for each tracked object. The BoT-SORT algorithm combines the advantages of motion and appearance information as well as a more accurate Kalman filter state vector (Aharon et al., 2022).

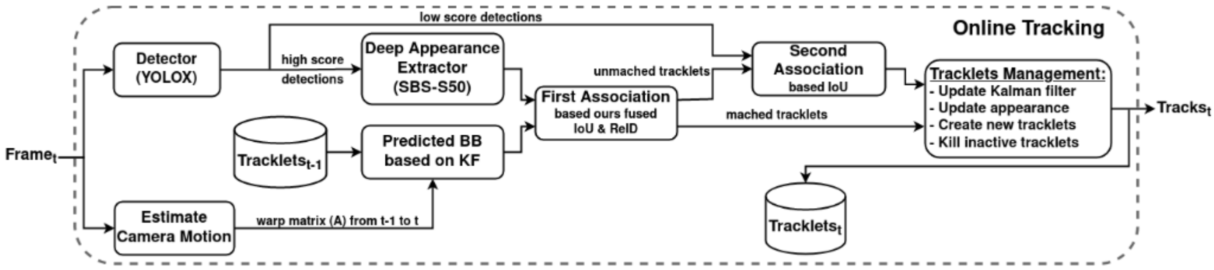


Figure 6. Overview of BoT-SORT-ReID tracker pipeline (Aharon et al., 2022).

Figure 6 shows the main pipeline of BoT-SORT object tracker that analyzes consecutive frames and produces a tracklet, which is a unique ID for each of the approaching vehicles in the ROI. In the use case of this study, the challenge is to keep a consistent ID on different motions of the target vehicle (i.e., straight movement as in the mid-block classifier and TMC on the intersection area). Keeping a unique tracklet for TMC is more challenging since the vehicle has different angles towards the UAV, causing the tracks to be eliminated owing to sudden changes in appearance. The tracker model then requires more training data to achieve the fine-tuned state when this scenario occurs.

3.5 Traffic Parameter Extraction

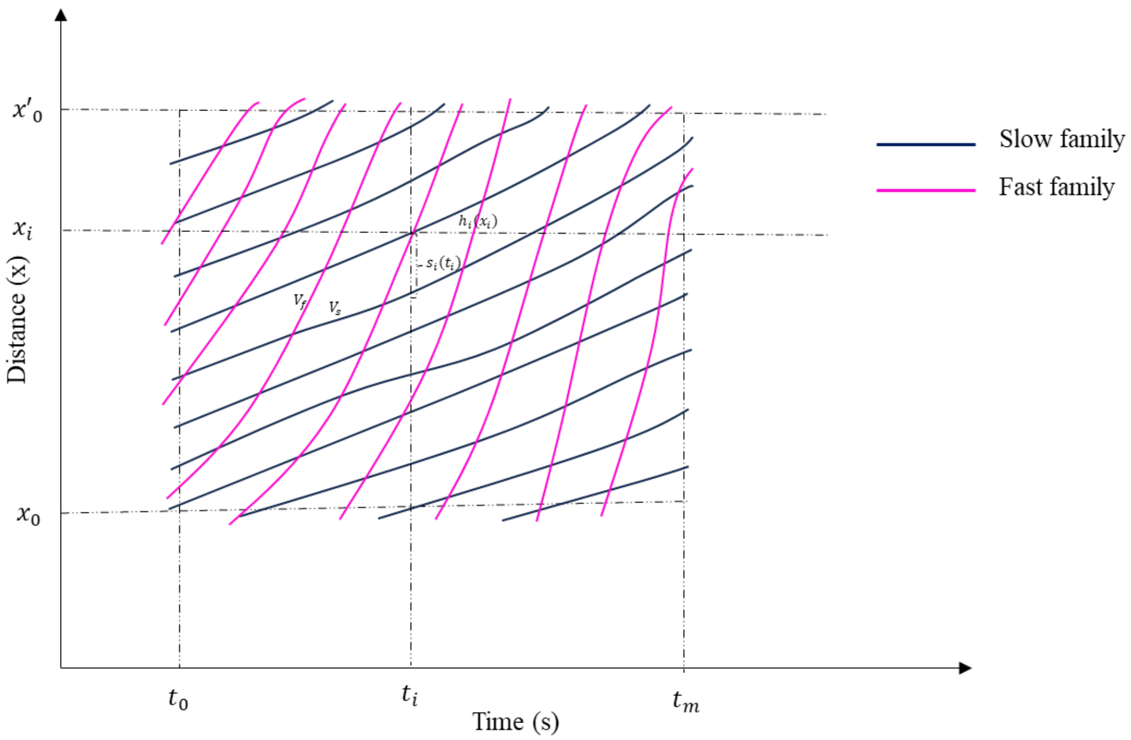


Figure 7 Generalized flow and density definition

From Figure 7, since this study is the implementation of aerial photography, we can draw generalized definitions from traffic flow theory. Figure 3 depicts two different families of vehicles (i.e., motorbikes and passenger cars as fast and slow families). Here flow is denoted as q(T,x) where $T = \sum h_i(x_i)$

Therefore, Flow = $q = \frac{m}{T}$ (1)

Density is denoted as the number of vehicles that enter segment L at any time $t_i = \frac{n}{L}$

Then any parameters are expressed as a space-mean function as shown in Equation 2

$$\alpha(L, x_i) = \frac{1}{n}\sum_{j=1}^{n} a_j(t_i) \tag{2}$$

Additionally, to define the captured vehicle speed, which consists of different speed profiles for each vehicle type, the space-mean speed must be determined. The flow and density properties of each vehicle profile are allowed to be additive functions to represent the overall traffic condition in the focused corridor segment. The flow for different vehicle families can then be extracted using Equation 3

$$\sum q_l = \sum_l k_l v_l \cdot \frac{k}{k} = k(\sum_l \frac{k_l}{k} v_l) \tag{3}$$

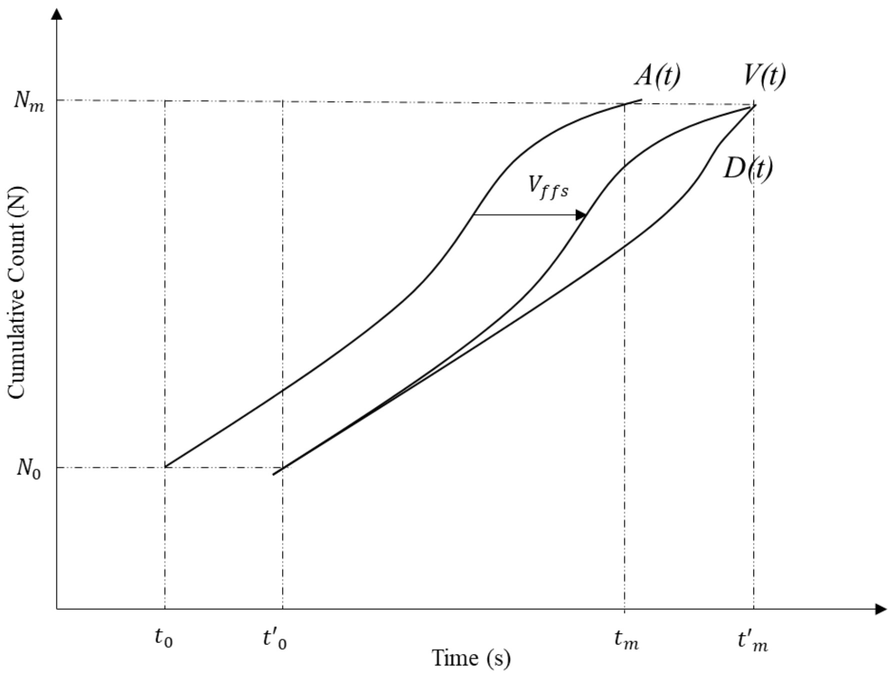


Figure 8. Cumulative diagram for delays extraction

Once the generalized parameters have been defined and the vehicle trajectory has been extracted from the vehicle data, we can then construct a cumulative diagram to obtain the excess cumulation of traffic at any particular timestamp, as well as the systematic delay for that particular segment.

With the robustness of data collected by UAVs, we can construct three curves that represent vehicle counts and status for each cell in each timeframe. The A(t) curve denotes the arrival curve of the vehicle that was the first to enter the cell entrance (upstream). The D(t) curve is the departure curve; it shows when the vehicle that has been tracked by the object tracker entered the cell entrance and went through the cell exit (downstream) when the queue in the segment has dissipated. The area between the A(t) and D(t) curves then represents the time the vehicle spent in each cell as well as the quantity of vehicles. We can then determine the overall time spent by all vehicles by calculating the area under the curve.

Finally, the virtual departure curve, V(t), is what the D(t) curve should look like when no incidents or systematic delays are in the system. We can assume that all vehicles continue to travel with the free flow speed, therefore, shifting the A(t) curve in the direction of the x-axis with the time it would take each vehicle with free flow speed to leave the cell. The system delay can then be determined by subtracting the area under the V(t) curve and D(t) curve to obtain the excess time spent by vehicles when there are delays. This section draws on Little's formula to express the rate of service to determine the systematic delay expressed in Equations 4 and 5. This application is useful when estimating the delay of signalized intersections to determine the quality of services (Ashquer et al., 2023).

$$\lambda = \frac{\bar{Q}}{\bar{w}} \tag{4}$$

where λ is the rate of service, $\bar{Q}$ is the number of excess vehicles accumulated, and $\bar{w}$ denotes the average excess time.

4. RESULTS AND DISCUSSION

4.1 Object Detection and Tracking Evaluation

The concept of confusion matrix is used to evaluate the performance of both object detection and object tracking. Therefore, it is essential to evaluate the results from the deployed model to ground truth.

The confusion matrix consists of four fundamental components: True positives (TP), which refers to vehicles being correctly classified to their respective class; True negatives (TN), which refers to vehicles not being classified into a particular class and those vehicles are not in that particular class; False Positives (FP), which refers to vehicles being misclassified into another vehicle class; False Negatives (FN), which refers to vehicles being wrongly classified as negative.

Multiple objects tracking accuracy (MOTA) is a widely used metric to evaluate tracking performance. Our goal was to follow the trajectories of vehicles in the ROI as they move and their interaction within the footage. In addition to the confusion matrix, additional factors are employed to evaluate the performance of object tracking. IDSW refers to a tracklet (ID) switch to a different vehicle, usually owing to confusion between vehicles. Missed targets refers to the number of times the tracking system loses track of a vehicle that is in the ROI, usually due to perceptive angle change owing to the effect of turning movement. GT represents the number of ground truth vehicles within the ROI in the footage.

With the concepts of TP, TN, TP, and FN defined, the accuracy of the classification model can be measured and its performance can be evaluated as shown in Table 3

Table 3. Evaluation criteria

Benchmark	Formulation	Description
Accuracy	$A = \frac{TP + TN}{TP + TN + FP + FN}$	Proportion of correctly classified to all samples
Precision	$P = \frac{TP}{TP + FP}$	Ratio of TP classified to all samples predicted as positive
Recall	$R = \frac{TP}{TP + FN}$	Ratio of TP to total samples that are actually positive
F1-Score	$F1 = \frac{2 \cdot P \cdot R}{P + R}$	Harmonic mean of precision and recall
MOTA	$MOTA = 1 - \frac{\sum_t (FN_t + FP_t + IDSW_t)}{\sum_t GT_t}$	Accuracy of tracking multiple objects over time in sequences

The results of model performance are shown in Table 4. The dataset used in this study had different physical geometric designs resembling common highways and corridors, located in downtown areas, central business district, interstate highways, etc.

The configuration of YOLOv8 model training and hyperparameters were as follows: batch size 64 using SGD optimizer with momentum 0.937, learning rate set to 0.01, weight decay of 0.0005 with focal loss gamma set to 0.0; to increase the performance of the model, the training data were augmented by modifying hue, saturation, scaling, and image flipping. For BoT-SORT tracker, the hyperparameters were as follows: appearance threshold: 0.481, sparseOptFlow method, confidence threshold 0.3501, frame rate: 30, lambda: 0.9896, matching threshold: 0.2273, new track threshold: 0.2114.

Table 4. Overall performance evaluation

	Section	MOTA
MBC	Corridor	
	Front facing, Aggregated	96.8
	Backward facing, Aggregated	96.6
	Front facing, single lane	76.0
	Backward facing, single lane	88.3
Hybrid	Interchange	
	Altitude – Low	92.1
	Altitude – Medium	93.5
	Altitude – High	70.7
TMC	Traffic Count on Intersection	62.3

Note: MBC = mid-block classifier, TMC = turning movement classifier, Hybrid = mix between MBC and TMC depending on camera angle and interchange layout

From Table 4, the MBC (in the corridor section) shows that the results achieved promising accuracy when compared to TMC. The testing was performed on 51 video footages of 60 minutes each. The accuracy in classifying and tracking vehicles in an ideal setup was more than 95 percent. Three different vehicle classes of interest (passenger car, motorcycle, heavy vehicle) along with the vehicle speed, traffic density, as well as delay were extractable from the collected UAV footage. We found hybrid movement type—a mix of MBC with different levels and mixed flow of TMC—to be common in interchanges of designated areas.

The model performed exceptionally well on lower and medium flight altitudes, achieving more than 90 percent accuracy. However, at an operating height of 90 meters, the model performance decreased significantly to 70.7 percent, trading off between model performance and more coverage area.

Given the trade-off, there is an optimal height that covers as much ground without diminishing the quality of video footage and model efficiency. To achieve maximum accuracy, we found that the optimal height and angle should be 50 meters and 45 degrees, respectively, for MBC movement. For hybrid capturing, data should be collected from between 60-70 meters to capture different interchange elevations while maintaining optimal model performance. The results of vehicle detection and tracking can be seen in Figures 9-11.

Figure 9. Model performance on multi-lane corridor ROI and single-lane corridor ROI. (ROI: region of interest).

Figure 10. Model performance on signalized intersection with different movement direction.

Figure 11. Model performance on multi-level interchange with different movement direction.

Table 5. Detailed performance for different vehicle types.

		Ground Truth	TP	FP	FN	A	P	R	Class Weight
MBC	Passenger Car	144116	136507	854	6755	0.9472	0.9937	0.9528	0.7528
	Motorcycle	34383	31277	-	3106	0.9097	1.0000	0.9097	0.1796
	Heavy Vehicle	12917	11998	269	650	0.9288	0.9781	0.9486	0.0675
TMC	Passenger Car	20700	16903	20	3777	0.8166	0.9988	0.8173	0.6772
	Motorcycle	7156	2495	-	4661	0.3486	1.0000	0.3486	0.2341
	Heavy Vehicle	2709	2439	-	270	0.9003	1.0000	0.9003	0.0886

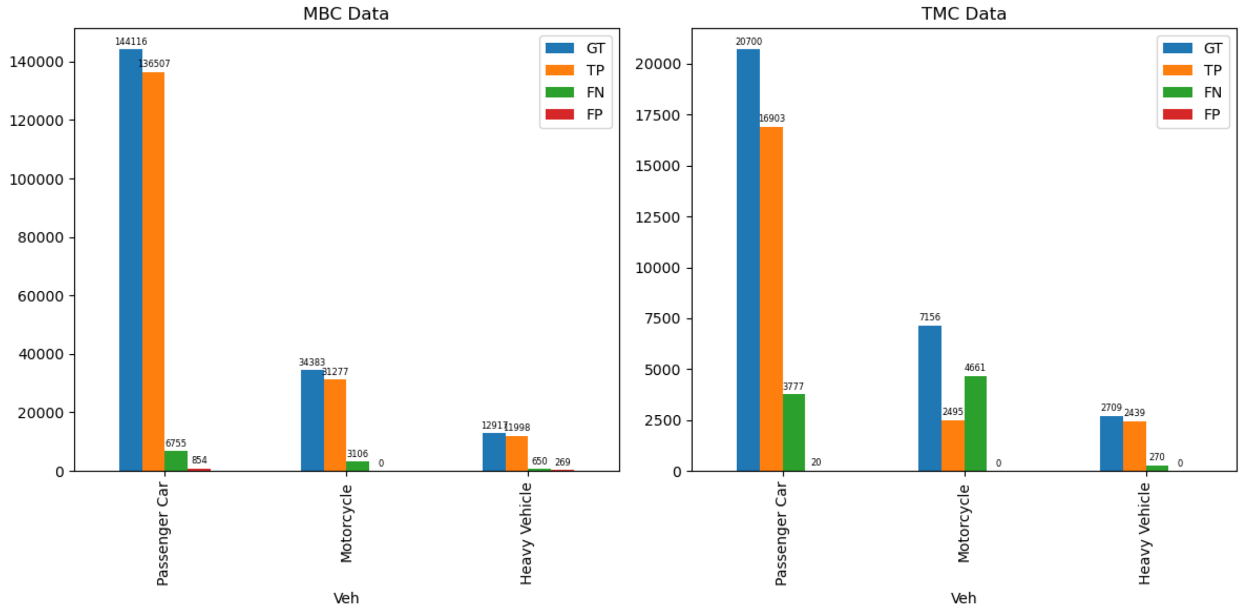


Figure 12. Vehicle distribution and detection results for MBC and TMC

Figure 12 shows the distribution and detailed model performance results on the testing dataset. From Table 5, it is noteworthy that, despite the presence of class imbalance within both the training and testing datasets (0.75/0.17/0.06), our mid-block classifier still demonstrates reliable performance. This outcome would not be possible if the training were conducted on the previous anchor-based model, where the existence of class imbalances usually led to generally worse model performance. The anchor-free method directly predicts the object location without relying on the offset of bounding boxes.

It is worth noting that footage analysis on TMC is still a challenge. Segregating each designated lane involves the change of vehicle appearance in the video footage, which confuses the object tracker. This causes redundant counting of the same vehicle, which the tracker registers as a different vehicle, or occlusion, where the UAV loses the line of sight to the vehicle. At 70 meters, elevated positioning is required to cover larger areas effectively. However, motorcycles are significantly smaller in size compared to standard passenger cars and heavy vehicles, which greatly diminishes the efficiency of the model in tracking vehicles, achieving only 30 percent accuracy. The intersection area is more sensitive and prone to the aforementioned problems. However, at an operation height of 50 meters, the accuracy significantly improved to 85 percent. However, the limitation is that only two legs of the intersection can be fully observed; there is need to capture longer legs of intersections to determine queue length as well as delay estimation considering that the goal is to improve traffic signals at intersections.

Intuitively, we might lower the operating height to 50 meters to maximize accuracy, such as shown in the corridor study section for MBC count. However, this approach will need multiple UAVs flying simultaneously side-by-side in the same area to capture all traffic directions.

5. RESULTS AND DISCUSSION

In this paper, we introduced a three-layer framework for traffic surveillance and extraction of crucial traffic parameters in urban areas, focusing on rush hour scenarios. Our framework leveraged robust computer vision algorithms in conjunction with data collected from high-resolution UAV footage. We employed the YOLOv8 algorithm for object classification and the BoT-SORT algorithm for vehicle tracking. Our approach achieved outstanding performance: a maximum accuracy rate of 96.8 percent in a highway corridor setting.

Additionally, we conducted a systematic investigation of the impact of UAV altitude and various camera angles on the overall model efficiency. This study provides valuable insights into the trade-off between model performance and the extent of the coverage area. Our findings suggest that to attain optimal results, it is advisable to operate the UAV at elevations ranging from 50 to 70 meters with a camera angle of 45 degrees.

For future research endeavors, there is an opportunity to enhance the efficiency of our model in detecting smaller and less common vehicle types beyond passenger cars and the tracking model robustness when monitoring vehicles undergoing changes in direction, such as during turns. These potential advances would further enhance the applicability and accuracy of the traffic surveillance system.

6. Acknowledgments

This work was supported by the Bureau of Road Research and Development, Department of Highways, Thailand

7. References

Adaimi, George, Kreiss, Sven, Alahi, Alexandre, 2023. Traffic perception from aerial images using butterfly fields. Transp. Res. C 153 104181

Aharon, Nir, Orfaig, Roy, Bobrovsky, Ben-Zion, 2022. BoT-SORT: Robust Associations Multi-Pedestrian Tracking. arXiv:2206 14651

Ahmed, Outay, Zaidi, Adnan, Ngoduy, 2020. Examining queue-jumping phenomenon in heterogeneous traffic stream at signalized intersection using UAV-based data. Pers. Ubiquit. Comput.

Ashqer, Ashqar, Elhenawy, Almannaa, Aljamal, Rakha, Bikdash, 2023. Evaluating a Signalized Intersection Performance Using Unmanned Aerial Data.

Ashutosh, Manohara, Ujjwal, Radhika, 2020. Efficient Vehicle Counting by Eliminating Identical Vehicles in UAV aerial videos. IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), pp. 246-251.

Barmpounakis, Vlahhogianni, Golias, Babinec, 2019. How accurate are small drones for measuring microscopic traffic parameters? Transportation Letters. Vol. 11, issue 6, pp. 332-340.

Barmpounakis, Sauvin, Geroliminis, 2020. Lane Detection and Lane-Changing Identification with High-Resolution Data from a Swarm of Drones. Transportation Research Record: Journal of the Transportation Research Board. Vol. 2674, issue 7, pp. 1-15.

Bashir, Boudjit, Zeadally, 2022. A closed-loop control architecture of UAV and WSN for traffic surveillance on highways. Computer Communications. Vol. 190, pp. 78-86.

Benjdira, Khursheed, Koubaa, Ammar, Ouni, 2019. Car Detection using Unmanned Aerial Vehicles: Comparison between Faster R-CNN and YOLOv3. IEEE 1st International Conference on Unmanned Vehicle Systems-Oman (UVS), pp 1-6.

Bozcan, Kayacan, 2020. AU-AIR: A Multi-modal Unmanned Aerial Vehicle Dataset for Low Altitude Traffic Surveillance. IEEE International Conference on Robotics and Automation (ICRA), pp 9504-8510.

Butilă, Boboc, 2022. Urban Traffic Monitoring and Analysis Using Unmanned Aerial Vehicles (UAVs): A Systematic Literature Review. Remote Sensing. Vol. 14, issue 3, pp. 620.

Chen, Chiu, Hsieh, Lin, Angah, 2020. Conflict analytics through the vehicle safety space in mixed traffic flows using UAV image sequences. Transp. Res. C 119 102744.

Guido, Gallelli, Rogano, Vitale, 2016. Evaluating the accuracy of vehicle tracking data obtained from Unmanned Aerial Vehicles. International Journal of Transportation Science and Technology. Vol. 5, Issue 3, pp. 136-151.

Kaewunruen, AbdelHadi, Kongpuang, Pansuk, Remennikov, 2023. Digital Twins for Managing Railway Bridge Maintenance, Resilience, and Climate Change Adaptation. Sensors, article 252.

Kanistras, Martins, Rutherford, Valavanis, 2013. A survey of unmanned aerial vehicles (UAVs) for traffic monitoring. 2013 International Conference on Unmanned Aircraft

Systems (ICUAS), pp. 221-234.

Ke, Li, Tang, Pan, Wang, 2019. Real-Time Traffic Flow Parameter Estimation From UAV Video Based on Ensemble Classifier and Optical Flow. IEEE Trans. Intell. Transport. Syst. Vol 20, issue 1, pp.54-64.

Ke, Li, Kim, Ash, Cui, Wang, 2017. Real-Time Bidirectional Traffic Flow Parameter Estimation From Aerial Videos. IEEE Transactions on Intelligent Transportation Systems. Vol. 18, issue 4, pp. 890-901.

Ke, Li, Tang, Pan, Wang, 2018. Real-time traffic flow parameter estimation from UAV video based on ensemble classifier and optical flow. IEEE Trans. Intell. Transp. Syst., pp. 54-64.

Ke, Feng, Cui, Wang, 2020 Advanced framework for microscopic and lane-level macroscopic traffic parameters estimation from UAV video, IET Intell. Transp. Syst., pp. 724-734.

Khan, Ectors, Bellemans, Janssens, Wets, 2017. UAV-Based Traffic Analysis: A Universal Guiding Framework Based on Literature Survey. Transportation Research Procedia. Vol. 22, pp. 251-550.

Khan, Ectors, Bellemans, Janssens, Wets, 2018. Unmanned Aerial Vehicle-Based Traffic Analysis: A Case Study for Shockwave Identification and Flow Parameters Estimation at Signalized Intersections. Remote Sensing. Vol. 10, issue 3, pp. 458.

Khan, Jhanjhi, Brohi, Usmani, Nayyar, 2020. Smart traffic monitoring system using Unmanned Aerial Vehicles (UAVs). Computer Communications. Vol. 157, pp. 434-443.

Kyrkou, Timotheou, Kolios, Theocharides, Panayiotou, 2018. Optimized vision-directed deployment of UAVs for rapid traffic monitoring. IEEE International Conference on Consumer Electronics, ICCE 2018, 2018-Janua, pp. 1-6.

Lin, Dollár, Girshick, He, Hariharan, Belongie, 2017. Feature Pyramid Networks for Object Detection. arXiv:1612 03144

Liu, Wang, Shi, Liu, Li, Mao, 2019. Vehicle tracking by detection in UAV aerial video. Science China Informat. Sci. 62, 24101.

Niu, Gonzalez-Prelcic, Heath, 2018. A UAV-based traffic monitoring system-invited paper. IEEE 87th Vehicular Technology Conference (VTC Spring), pp. 1-5.

Outay, Mengash, Adnan, 2020. Applications of unmanned aerial vehicle (UAV) in road safety, traffic and highway infrastructure management: Recent advances and challenges. Trans. Res. Vol.141, pp.116-129.

Salvo, Caruso, Scordo, 2014. Urban Traffic Analysis through an UAV. Procedia – Social and Behavioral Sciences. Vol. 111, pp. 1083-1091.

Salvo, Caruso, Scordo, Guido, Vitale, 2017. Traffic data acquirement by unmanned aerial vehicle. European Journal of Remote Sensing. Vol. 50, issue 1, pp. 343-351.

Shirazi, Patooghy, Shisheie, Haque, 2020. Application of Unmanned Aerial Vehicles in Smart Cities using Computer Vision Techniques. IEEE International Smart Cities Conference (ISC2), pp1-7.

Wang, Chen, Yin, 2016. Detecting and tracking vehicles in traffic by unmanned aerial vehicles. Automation in Construction. Vol. 72, pp. 294-308.

Xu, Yu, Wang, Wu, Ma, 2017. Car Detection from Low-Altitude UAV Imagery with the Faster R-CNN. Journal of Advanced Transportation. pp. 1-10.

Zhu, Sun, Jia, Li, Hou, Lin, Liu, Qiu, 2018. Urban Traffic Density Estimation Based on Ultrahigh-Resolution UAV Video and Deep Neural Network. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. Vol. 11, issue 12, pp. 4968-4981.

Zhao, Zhang, Brooks, Banach, Aden, 2018. A comprehensive overview of improving traffic flow observability using UAVs as mobile sensors. 97th Annual Meeting, Transportation Research Board, Washington DC, USA. Paper No. 18-04949.

Zhang, Liptrott, Bessis, Cheng, 2019. Real-time traffic analysis using deep learning techniques and UAV based video. In: 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)., pp. 1-5.

Evaluating Driver's Characteristics During Multiple Rear-end Collisions on Freeway Corridors

Ryuya Seki[1,*] and Hironori Suzuki[2]

[1] Department of Mechanical Engineering, Toyo University, Saitama, Japan, s36a02400169@toyo.jp
[2] Department of Mechanical Engineering, Toyo University, Saitama, Japan, suzuki054@toyo.jp

ABSTRACT

The prevention of multiple rear-end collisions on highways is crucial due to the potentially severe consequences compared to single-vehicle rear-end accidents. While some characteristics of multiple rear-end collisions have been studied, the exact mechanisms leading to these accidents remain incompletely understood. Therefore, it is essential to clarify the states of drivers leading up to multiple rear-end collisions. This study aims to elucidate the driving characteristics of drivers who were able to avoid multiple rear-end collisions and those who were not. We conducted driving simulator experiments with five passenger cars following each other and replicated accidents in which an automated control vehicle collided with a stopped vehicle at the end of a traffic jam in a tunnel section. The differences in characteristics between drivers who could avoid multiple rear-end collisions and those who could not are then revealed by estimating five parameters of the Intelligent Driver Model (IDM), which accurately reproduces measured driver acceleration, using a dual particle filter (DPF).The numerical analyses showed that the colliding vehicles primarily operated in a speed control mode for the whole period, focusing solely on their own speed, with little awareness of the leading vehicle. Conversely, drivers who avoided accidents mostly operated in a speed control mode as well, with reduced awareness of the leading vehicle, but frequently controlled their vehicle by a "distance control mode", where desired and actual headway distance against the preceding vehicle are completely synchronized. This indicates frequent awareness of the leading vehicle, suggesting that frequent attention to the leading vehicle may be one characteristic of drivers who can avoid multiple rear-end collisions. Additionally, ANOVA revealed significant statistical differences in most of the IDM parameters between participants who successfully avoided the collision and those who did not. It can be concluded that the proposed approach to evaluating the characteristics of drivers who collided and those who did not could be elucidated through the estimation of IDM parameters using the DPF.

Keywords: Intelligent Driver Model, Particle filter, Driving Control Mode, Multiple rear-end collisions, Driving Characteristics

1. INTRODUCTION

In the 2021 traffic accident statistics in Japan, rear-end collisions accounted for 30.5 % of all incidents, representing the highest proportion. Notably, on highways, rear-end collisions have the highest occurrence rate, and the proportion of fatal accidents among these incidents is 2.7 %, more than three times higher than the 0.8 % observed on general roads (Japan Cabinet Office, n.d.). The significance of preventing multi-vehicle rear-end collisions on highways, where vehicles travel at high speeds and densities, is considerable due to the greater severity of such accidents compared to single-vehicle rear-end collisions.

Ivan et al. modeled the impact of various independent variables on the occurrence of single-vehicle accidents and multi-vehicle collisions by analyzing traffic data (Ivan et al., 1999). Their results confirmed that the high traffic volume (poor Level of service (LOS)) and geometric characteristics of the road significantly influence the occurrence of accidents. Additionally, it was shown that site conditions are not a critical factor in single-vehicle accidents. However, their analysis only revealed characteristics of road environmental factors such as shoulder width, LOS, and distance between signals, without clarifying driver characteristics. Hirata et al. conducted a simulator experiment in a Virtual Reality (VR) environment that replicated an underground urban road, where participants simultaneously operated two vehicles (Hirata et al., 2007). This experiment elucidated the characteristics of multi-vehicle collisions and the relationship between driving behavior and accident types. Nevertheless, the mechanism of multi-vehicle collisions is not yet fully understood, and accident prevention measures have not been adequately examined. To elucidate the mechanisms leading to multi-vehicle rear-end collisions, it is necessary to clarify the state of the driver leading up to such incidents.

As long as the driver's behavior is accurately described by the driver model, the driver's characteristics can be reflected in the model's parameters. Therefore, estimating these parameters is crucial for identifying the driver's state during incidents. Monteil et al. (2015) and Buyer et al. (2019) attempted to dynamically estimate the parameters of the intelligent driver model (IDM), which is well-known and widely used for modeling longitudinal driver behavior (Kesting et al., 2010). Kim and Mahmassani (2010) investigated the correlations among these parameters, while Punzo et al. (2014) conducted a detailed analysis and concluded that only two out of the five parameters are sensitive and need to be estimated. While their research focused on estimating or analyzing the IDM parameters, there has been no investigation into the driver's characteristics derived from these estimated values.

In this study, we will first recreate multi-vehicle rear-end collisions based on the analysis of actual accidents that occurred on highways within a driving simulator (DS). A DS experiment will be conducted, where subjects drive one of five passenger cars following each other in a sequence. Using data obtained from the DS experiment, we aim to identify the driver characteristics that differentiate those who avoided collisions from those who did not. This will be achieved by employing the driver state estimation method proposed by Suzuki and Seki (2024), which dynamically estimates the IDM parameters using a dual particle filter as a Bayesian estimator.

2. ONLINE ESTIMATION OF DRIVER MODEL PARAMETERS AND EVALUATION OF DRIVER CHARACTERISTICS

In this study, we aim to elucidate the differences in characteristics between drivers who were able to avoid multi-vehicle rear-end collisions and those who were not, using the parameters of the Intelligent Driver Model (IDM). The parameters of the IDM will be estimated online using a dual particle filter (DPF). The DPF will estimate the five IDM parameters necessary to accurately reproduce the measured acceleration. From these parameters, we will identify the driver characteristics that distinguish subjects who avoided collisions from those who did not.

2.1 Intelligent driver model (IDM)

The Intelligent Driver Model (IDM) is one of the most widely used driver models for calculating the acceleration of a following vehicle relative to a leading vehicle (Treiber et al., 2000). In the IDM, the acceleration of vehicle i at time k denoted as $acc_i(k)$, is calculated using the vehicle's own speed $v_i(k)$ the speed of the preceding vehicle $v_{i-1}(k)$, and the headway distance $d_i(k)$, as shown in Equation (1). The desired distance gap, $D(k)$, which represents the driver's preferred following distance, is calculated using Equation (2).

$$acc_i(k) = a\left[1 - \left(\frac{v_i(k)}{V}\right)^4 - \left(\frac{D(k)}{d_i(k)}\right)^2\right] \tag{1}$$

$$D(k) = s + v_i(k) \cdot T - \frac{v_i(k)\big(v_{i-1}(k) - v_i(k)\big)}{2\sqrt{ab}} \tag{2}$$

The five model parameters in Equations (1) and (2) are as follows, with the domain set to $0 < a \leq 10\,, 0 < b \leq 10\,, 0 < V \leq 40\,, 0 < s \leq 10\,, 0 \leq T$:

a : Maximum acceleration (m/s^2)

b : Comfortable deceleration (m/s^2)

V : Desired speed (m/s)

s : Minimum gap at standstill (m)

T : Desired time headway THW(s)

2.2 Dual particle filter (DPF)

The dual particle filter (DPF) is an online estimation system that applies the concept of dual filtering, where state and measurement variables are mutually estimated and their results reflected upon each other, using particle filters. The five IDM model parameters to be estimated by the DPF are state variables $\theta_k = [a(k),\ b(k),\ V(k),\ s(k),\ T(k)]^T$. The observation variable is $y_k = acc_i(k)$, and the input variables are $x_k = [v_i(k),\ v_{i-1}(k),\ d_i(k)]^T$. The state-space model is defined as follows, with g representing the IDM equations (1) and (2), r_k the system error at time k, and n_k the observation error at time k.

$$\theta_k = \theta_{k-1} + r_{k-1} \tag{3}$$

$$y_k = g(\theta_{k-1}, x_{k-T}) + n_k \tag{4}$$

First, the five model parameters to be estimated are decomposed as shown in Equation (5) below.

$$\begin{aligned} a_k &= a_{k-1} + r_{a,k} \\ b_k &= b_{k-1} + r_{b,k} \\ V_k &= V_{k-1} + r_{V,k} \\ s_k &= s_{k-1} + r_{s,k} \\ T_k &= T_{k-1} + r_{T,k} \end{aligned} \tag{5}$$

Next, individual particle filters, PF_a, PF_b, PF_V, PF_s, PF_T, are assigned to each of the five parameters. Subsequently, each particle filter executes the procedure depicted in Figure 1. Here, we explain the estimation of a_k at time k using PF_a as an example. Table 1 provides a description of each step in this process.

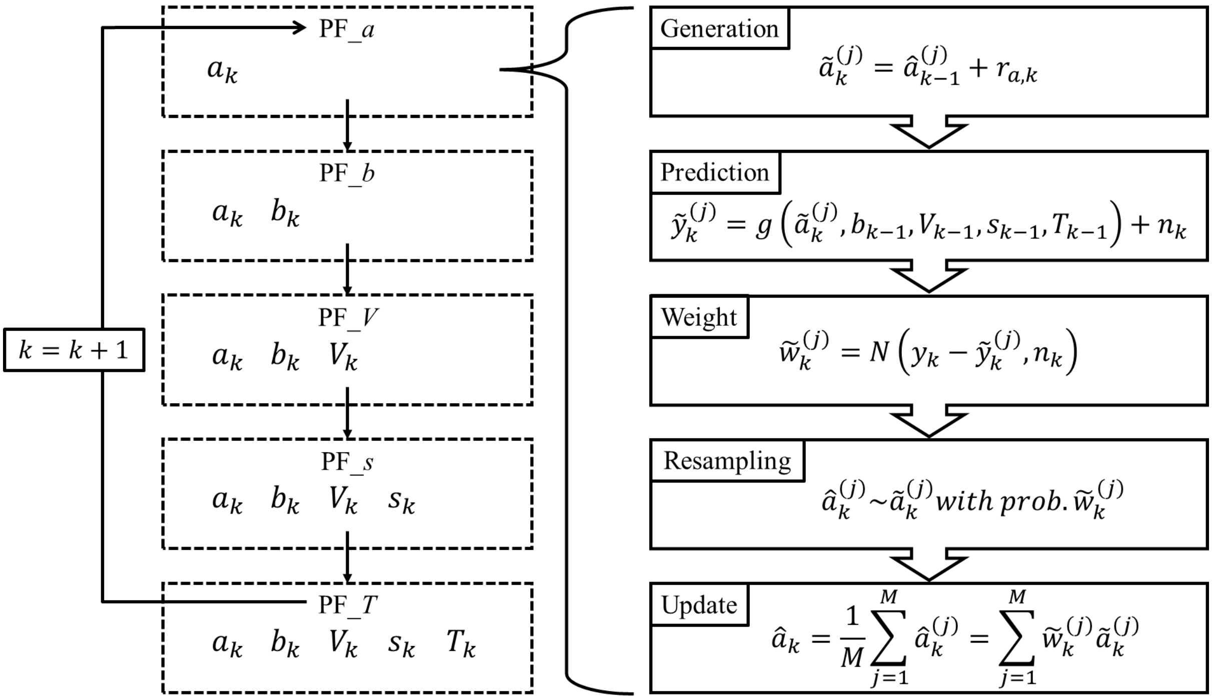


Figure 1 Process of the dual particle filter.

Table 1 Description of each step in the dual particle filter.

Step	Description
1. Generation	From Equation (5), generate particles $\tilde{a}_k^{(j)}$ using the system model. j denotes the particle number $j = 1, \cdots, M$
2. Prediction	Input each generated particle $\tilde{a}_k^{(j)}$ into the function g in Equation (4) to obtain the one-step-ahead prediction of the measurement variable $\tilde{y}_k^{(j)}$. At this point, use the latest estimated values for parameters other than the one being estimated. Specifically, in PF_a, the parameters other than a use their estimates from time $k-1$.
3. Weight	Compare the one-step-ahead predictions $\tilde{y}_k^{(j)}$ with the actual observation y_k. Assign weights $\widetilde{w}_k^{(j)}$ based on a Gaussian distribution: particles with predictions close to the observation receive higher weights, while those with predictions far from the observation receive lower weights.
4. Resampling	Exclude particles with low weights and replicate particles with high weights as needed. Resample particles based on the weights $\widetilde{w}_k^{(j)}$ to select only valid particles.
5. Update	Using all resampled valid particles, compute the expected value $\hat{a}_k$ and update the estimate.

3. METHODOLOGY

3.1 Driving simulator experiment

Hirata et al. (2017) collected data from newspaper articles published between 2003 and 2005, yielding information on a total of 467 traffic accidents. From this data, traffic accidents involving three or more vehicles were defined as multi-vehicle rear-end collisions and were investigated. The results showed that the number of vehicles involved typically ranged from three to five, and accidents involving more than ten vehicles were often caused by adverse weather conditions (heavy rain, dense fog, mist). The most common cause of multi-vehicle rear-end collisions was single-vehicle accidents, and large vehicles frequently played a role.

Using these findings as a reference, we conducted an experiment using a fixed-base driving simulator (hereafter, DS) to replicate a multi-vehicle rear-end collision, based on the analysis of an actual highway incident. The experiment involved subjects driving five passenger cars in a continuous following scenario. This DS is a system that allows the simultaneous operation of five independent vehicles in the same virtual space.

The experimental road was a 16 km, two-lane highway. The section from approximately 1.5 km to 14.5 km was a tunnel. Traffic volume was simulated at approximately 600 vehicles per hour (VPH) in the overtaking lane and approximately 300 vehicles per hour in the driving lane. An overview of the experimental road is shown in Figure 2.

The collision scenario was replicated with an AI-controlled large truck (hereafter, AI vehicle) crashing into a stationary vehicle at the end of a traffic jam inside the tunnel. Subject ID 01 followed the AI vehicle, while subjects ID 02 to ID 05 followed the vehicle driven by the preceding subject. Subjects were instructed to drive as they normally would, ensuring they did not collide with the vehicle ahead. Data on the driving behavior of those who avoided the AI vehicle's collision and those who did not were collected.

The experiment involved ten participants, all holding a driver's license, with ages ranging from their 20 s to 50 s. The average driving experience of the participants in their 20 s was 2 years, while that of the participants in their 50 s was 31 years. Figure 4 shows the screen of the DS experiment.

The experiment was non-invasive, participation was voluntary, subjects could withdraw at any time, and their data was anonymized and only accessible to the researchers. These conditions were thoroughly explained to the subjects, and consent was obtained before proceeding.

3.2 Calculation conditions

In the Dual Particle Filter (DPF), the covariance of the system error and observation error were set between 0.1 to 0.5 and 0.3, respectively. The number of particles was set to 500. Parameters were estimated every 0.1 seconds from the data obtained in the DS experiment.

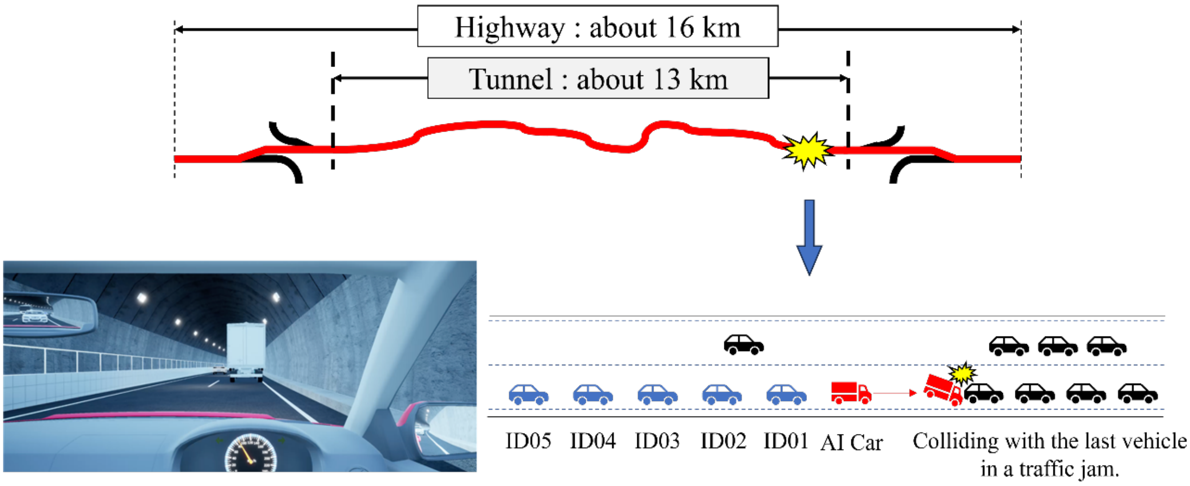


Figure 2 Overview of the driving simulator experiment.

3.3 Detection method of driver characteristics

Driver characteristics were classified into three control modes: "Speed Control Mode," "Distance Control Mode," and "Feedforward Control Mode." Figures 3 and 4 illustrate the graphs for each control mode, with the orange line representing the actual values and the blue line representing the IDM parameters estimated by the DPF.

First, the "Speed Control Mode" occurs when the driver's desired speed V matches the actual speed, as estimated by the IDM parameters through the DPF. In this mode, it can be inferred that the driver is regulating acceleration based on feedback from the vehicle's speed only.

Next, the "Distance Control Mode" is defined when the driver's desired headway distance D, estimated from the IDM parameters via the DPF, matches the actual headway distance. In this mode, the driver is adjusting acceleration based on feedback from the distance to the preceding vehicle.

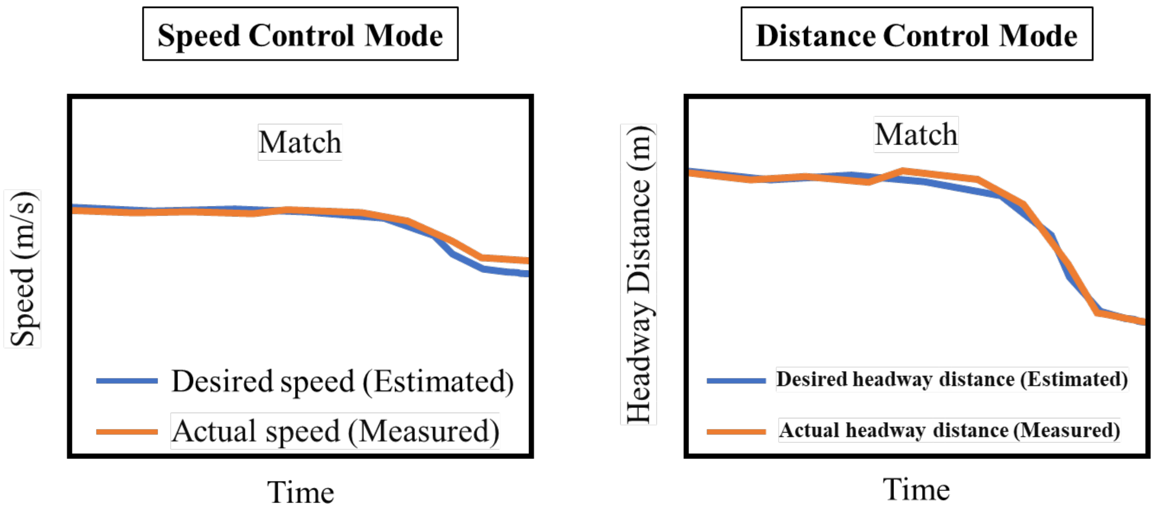


Figure 3 Speed Control Mode and Distance Control Mode.

Finally, the "Feedforward Control Mode" is characterized by neither "Speed Control Mode" nor "Distance Control Mode". In other words, neither the speed nor the headway distance matches the IDM parameters estimated by the DPF. In this mode, the driver is not using feedback from either the vehicle speed or the headway distance, indicating a driving behavior that does not incorporate feedback from the surrounding vehicles or environment, which can be risky in certain situations.

A multi-vehicle rear-end collision typically occurs when drivers are not aware of the preceding vehicle despite a reduced headway distance following the initial collision. Therefore, drivers need to be in "Distance Control Mode" among the three control modes. Any other control mode suggests a lack of awareness of the preceding vehicle. This study investigated the differences in control modes based on the number of vehicles from the initiating collision vehicle at the time of the collision. Additionally, we hypothesized that there would be differences in control modes during normal driving before the accident between drivers who could avoid the multi-vehicle rear-end collision and those who could not. This hypothesis was examined and verified.

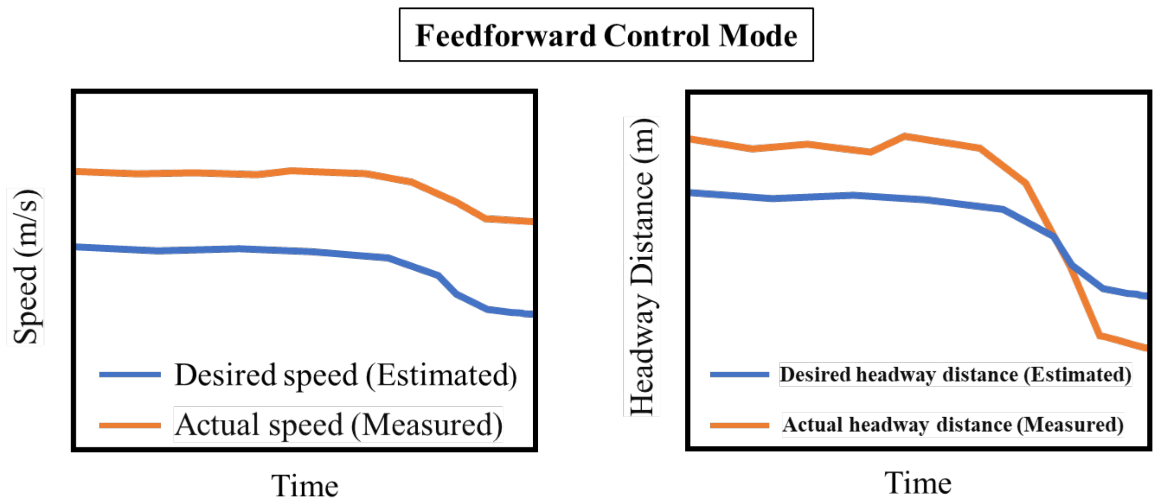


Figure 4 Feedforward Control Mode.

4. EXPERIMENTAL RESULTS AND DISCUSSION

For example, Figure 5 illustrates the results for ID 01, while Figure 6 presents those for ID 04. The graphs depict (a) acceleration, (b) Time Headway (THW), (c) velocity, and (d) headway distance from top to bottom. The blue lines represent the values estimated by the DPF, indicating the desired values of the drivers, while the orange lines represent the measured values.

4.1 Drivers involved in rear-end collisions

Figure 5 shows the results of ID 01, who failed to avoid the rear-end collision and exhibited a lack of awareness of the preceding vehicle for most of the time. This conclusion is substantiated by the following observations:

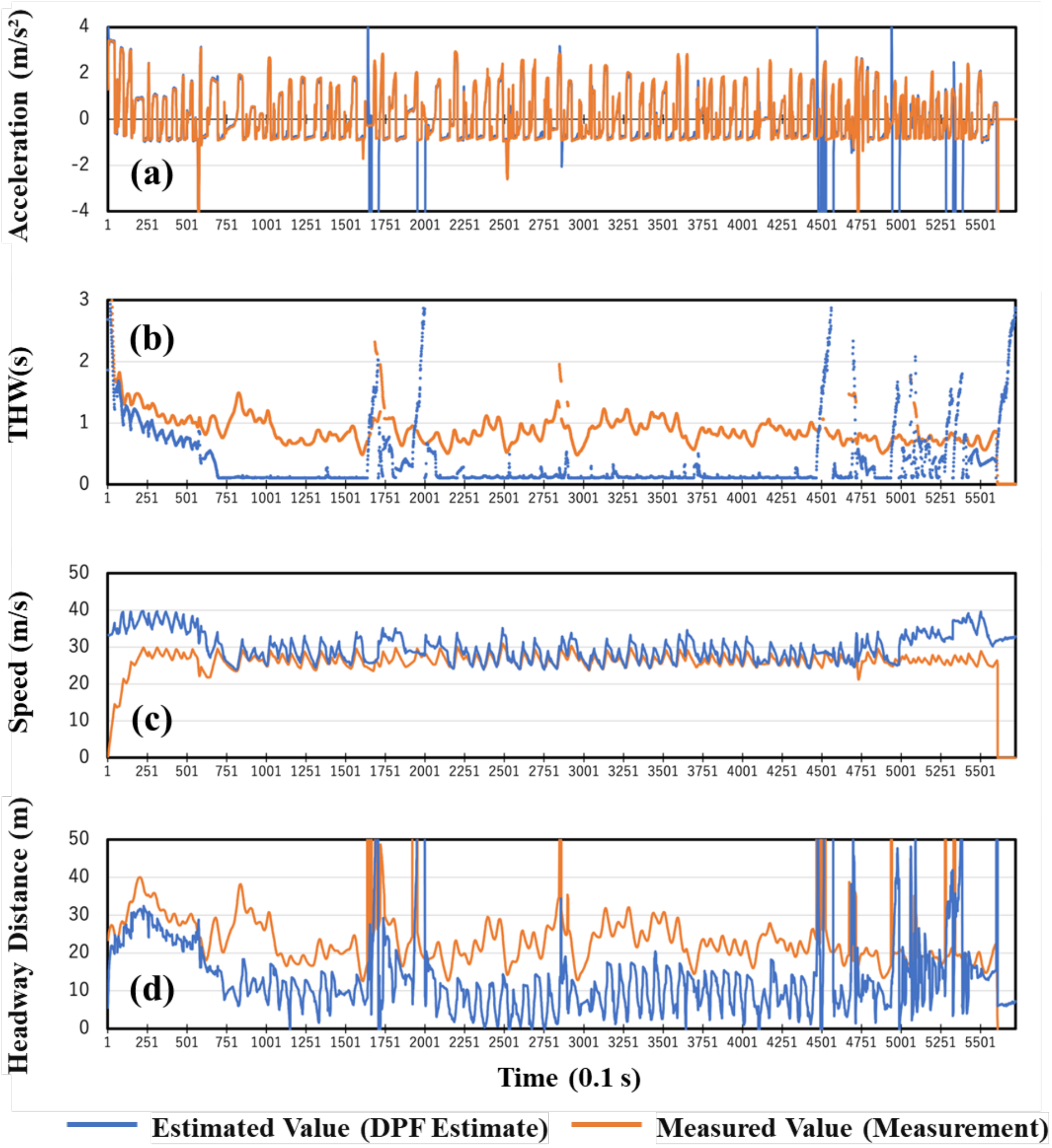


Figure 5 Experimental results for ID 01.

1) The estimated desired speed closely matches the actual speed of the driver (Fig. 5 (c)), indicating that ID 01 was primarily focused on their own vehicle's speed rather than the vehicle ahead.

2) The estimated desired headway distance significantly diverges from the actual headway distance as shown in Fig. 5 (d). This discrepancy suggests that the driver did not incorporating feedback regarding the position of the preceding vehicle.

Although ID 01 was following the preceding vehicle in an extremely risky situation where the THW was below 1 second (Fig. 5 (b)), the driver was not aware of the preceding vehicle for most of the time. This lack of awareness is one of the reasons why the rear-end collision occurred.

Similar results were observed for both ID 02 and ID 03. Thus, a lack of attention to the preceding vehicle likely led to the rear-end collision.

The analyses of the three participants mentioned above suggest that a lack of awareness of the preceding vehicle is a characteristic of those who could not avoid a rear-end collision. As demonstrated here, our proposed approach has the significant advantage of estimating a driver's lack of attention to the preceding vehicle using only numerical data, without relying on cameras for face recognition.

4.2 Drivers who avoided rear-end collisions

Figure 6 illustrates the results of ID 04, who successfully avoided the rear-end collision, indicating frequent awareness of the preceding vehicle during the experiment. This frequent awareness is represented by the green zone in Figs. 6 (c) and (d), where the desired headway distance consistently aligns with the actual headway distance. This conclusion is drawn from the following observations:

When focusing on the headway distance in Fig. 6 (d), there are frequent instances where the estimated desired headway distance aligns with the actual headway distance, in contrast to the driver involved in the rear-end collision. This alignment suggests that the driver was incorporating feedback regarding the position of the preceding vehicle.

Examining THW, ID 04 was following the preceding vehicle in a potentially hazardous situation where THW is below 1 second, but demonstrated awareness of the preceding vehicle very frequently during the experiment (Fig. 6 (b)). It is believed that this frequent awareness of the preceding vehicle contributed to the avoidance of the collision. These characteristics are similarly observed in ID 05, indicating a potential driver characteristic of those who could avoid a multi-vehicle rear-end collision.

4.3 Discussion

Drivers like ID 01, who could not avoid the collision, typically drove in a speed control mode, determining their acceleration based on feedback from their own vehicle speed. On the other hand, drivers like ID 04, who avoided the collision, also primarily drove in speed control mode. However, only the drivers who could avoided the collision frequently switched to headway distance control mode, where they determined acceleration based on feedback from the distance to the vehicle ahead. This suggests that drivers who avoided the collision were more sensitive to the headway distance before the occurrence of the multi-vehicle rear-end collision, enabling them to respond appropriately when the collision occurred. This sensitivity to the headway distance is potentially one of the driver characteristics that contributed to avoiding the collision.

4.4 Statistical test for IDM parameters

ANOVA was conducted for each of the five IDM parameters to ensure that there is a statistically significant difference between participants who successfully avoided the multiple rear-end collision and those who did not. The results revealed statistically significant differences in four out of the five parameters, with p−values less than 0.001: p = 5.38E−155 for "*a*", 2.57E−132 for "*b*", 0.00 for "*s*", and 5.48E−249 for "*T*". According to the study by Punzo et al. (2014), the most sensitive IDM parameters are *a* and *T*. Our statistical analysis showed that participants with higher maximum acceleration (a = 7.53 m/s^2) were more successful in avoiding the collision, compared to those who were less successful (a = 6.48 m/s^2). In contrast, individuals who avoided the collision had a shorter desired safe time headway (T = 0.27 s) than those who failed to avoid it (T = 0.53 s). The drivers who successfully avoided the collision closely followed the preceding vehicle with a shorter headway (*T*), while ensuring safety through a more sensitive adjustment of the parameter *a*.

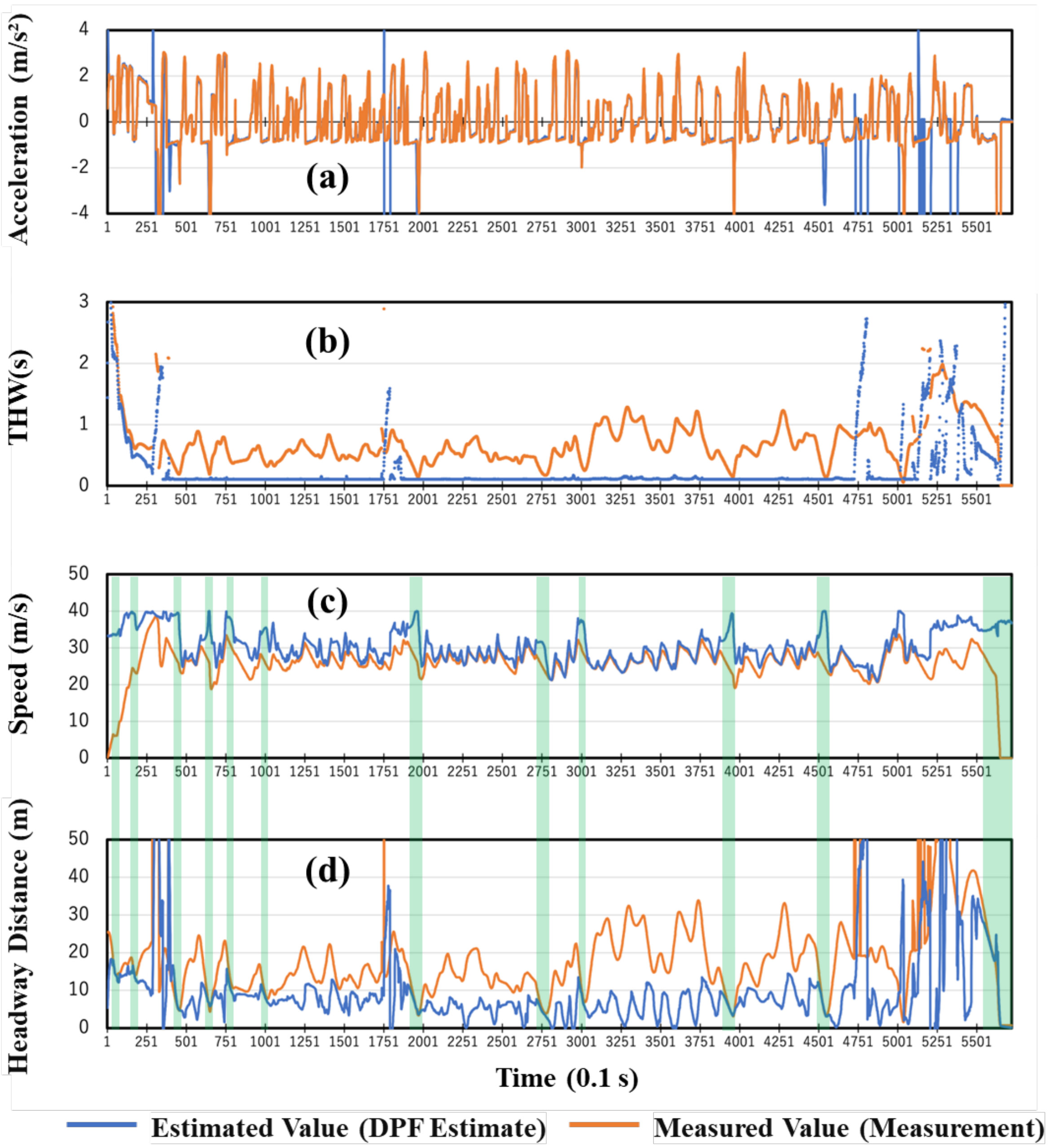


Figure 6 Experimental results for ID 04.

5. CONCLUSION

This study aimed to elucidate the driver characteristics that differentiate those who avoided a multi-vehicle rear-end collision from those who did not. The experiment replicated a real-world multi-vehicle rear-end collision on a DS, involving participants driving five consecutive passenger cars. We utilized a DPF to estimate the five parameters of the IDM that accurately reproduced the measured accelerations from the DS experiment. Additionally, we analyzed three control modes such as speed control, headway distance control, and feedforward control modes to identify the differences in driver behavior between those who avoided the collision and those who did not.

The findings revealed that in highly risky situations where the THW was less than one second, drivers who failed to avoid the collision predominantly operated in the speed control mode, wherein they determined acceleration based on feedback from their own vehicle speed, indicating a lack of awareness of the preceding vehicle.

Conversely, drivers who successfully avoided collisions often operated in headway distance control mode, adjusting their acceleration based on feedback from the distance to the vehicle ahead. In other words, they frequently paid close attention to the vehicle directly in front of them. These driver characteristics were identified solely from numerical data without relying on face recognition cameras, highlighting differences in the drivers' awareness of the preceding vehicle during normal driving conditions, not just immediately before the collision.

Additionally, ANOVA revealed significant statistical differences in four out of five parameters between participants who avoided the collision and those who did not. This indicates that drivers who successfully avoided the collision closely followed the preceding vehicle with lower desired time headway "T" while maintaining safety through a more sensitive adjustment of the parameter "a".

The identified driver characteristics have significant implications for the development of traffic control, vehicle control, and autonomous driving systems that aim to prevent not only single-vehicle rear-end collisions but also multi-vehicle rear-end collisions, without the need for face recognition cameras.

Despite these findings, several challenges remain. The current experiment's limitations, such as the extremely short headway times and the small number of participants, make it difficult to conclusively determine the utility of these findings in elucidating the mechanisms behind multi-vehicle rear-end collisions. Future research should focus on collecting driving data from a larger number of participants under more typical headway conditions to further investigate and validate the detection of driver characteristics.

6. REFERENCES

Buyer1, J., Waldenmayer1, D., Sußmann1, N., Z¨ollner, R., & Z¨ollner, J. M. (2019). Interaction Aware Approach for Online Parameter Estimation of a Multi-lane Intelligent Driver Model. Proc. of IEEE ITSC, 3967-3973. https://doi.org/10.1109/ITSC.2019.8917257

Hirata, T., Yamaguchi, K. & Yai, T. (2007). VR Analysis of multi-vehicle accidents in underground urban expressway. Journal of the Eastern Asia Society for Transportation Studies, 7, 2655-2664. https://doi.org/10.11175/easts.7.2655

Ivan, J. N., Pasupathy, R. K., & Ossenbruggen, P. J. (1999). Differences in causality factors for single and multi-vehicle crashes on two-lane roads. Accident Analysis and Prevention, 31(6), 695-704. https://doi.org/10.1016/S0001-4575(99)00030-5

Japan Cabinet Office. (n.d.). Road traffic accident situation in 2021. White Paper on Traffic Safety in Japan. (in Japanese) https://www8.cao.go.jp/koutu/taisaku/r04kou_haku/zenbun/genkyo/h1/h1b1s12.html

Kesting, A., Treiber, M. & Helbing, D. (2010). Enhanced Intelligent Driver Model to Access the Impact of Driving Strategies on Traffic Capacity. Philosophical Transactions of the Royal Society, 368(1928), 4585-4605. https://doi.org/10.1098/rsta.2010.0084

Kim, J., & Mahmassani H. S. (2011). Correlated Parameters in Driving Behavior Models - Car-Following Example and Implications for Traffic Microsimulation. TRR 2249.

Monteil, J., O'Hara, N., Cahill, V., & Bouroche, M. (2015). Real-Time Estimation of Drivers' Behaviour. Proc. of IEEE ITSC, 2046-2052. https://doi.org/10.1109/ITSC.2015.331

Punzo, V., Cliuffo, B., & Montanino, M. (2014). Do We Really Need to Calibrate All the Parameters? Variance-Based Sensitivity Analysis to Simplify Microscopic Traffic Flow Model. IEEE Transactions on Intelligent Transport Systems, 16(1), 184-193. https://doi.org/10.1109/TITS.2014.2331453

Suzuki, H. & Seki, R. (2024). Dual particle filtering of intelligent driver model parameters for indirect detection of driving anomalies. Human Interaction and Emerging Technologies, 120, 336-343. https://doi.org/10.54941/ahfe1004584

Treiber, M., Hennecke, A. & Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. Physical Review E, 62, 1805-1824. https://doi.org/10.1103/PhysRevE.62.1805

iCHE
BANGKOK 2024

iCHE
BANGKOK 2024