Journal of Informatics and Web Engineering

Vol. 5 No. 1 (February 2026) eISSN: 2821-370X

Vehicle Re-identification System using Residual Network with Instance-Batch Normalization

Wei-Jie Low[1], Kah-Ong Michael Goh[2*], Check-Yee Law[3], Connie Tee[4], Yong-Wee Sek[5] , Md Ismail Hossen[6]**

[1]Faculty of Information Science and Technology, Multimedia University, Jalan Ayer Keroh Lama, 75450 Melaka, Malaysia.

[2,4]Center for Image and Vision Computing, COE for Artificial Intelligence, Faculty of Information Science and Technology (FIST), Multimedia University, Jalan Ayer Keroh Lama, 75450 Bukit Beruang, Melaka, Malaysia.

[3]Centre for Intelligent Cloud Computing, COE for Advanced Cloud, Faculty of Information Science and Technology (FIST), Multimedia University, Jalan Ayer Keroh Lama, 75450 Bukit Beruang, Melaka, Malaysia.

[5]Faculty of Artificial Intelligence and Cyber Security, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.

[6]School of Information and Communication Technology, Griffith University, Nathan, QLD, Australia.

**corresponding author: (michael.goh@mmu.edu.my; ORCiD: 0000-0002-9217-6390)*

***corresponding author: (m.hossen@griffith.edu.au; ORCiD: 0000-0003-3632-937X)*

Abstract - Vehicle Re-identification (Re-ID) has become extremely important due to the increasing number of vehicles on the road and its potential to address traffic-related challenges. As a result, there is also a growing need for efficient methods to track and identify vehicles across multiple traffic cameras. One of the biggest challenges of this task is the variations in vehicle appearances across different camera angles. This is because vehicles can appear significantly different when captured from various camera angles and viewpoints. Furthermore, the current vehicle Re-ID solutions typically require extensive coding knowledge, making it inaccessible to many potential users. Therefore, we focus on developing a user-friendly software application that simplifies the entire Re-ID workflow. This includes tasks like dataset preparation and data preprocessing using YOLO, model training with ResNet-ibn, performance evaluation, and visualization of results. The application provides a comprehensive pipeline that enables users to perform vehicle Re-ID tasks without requiring advanced programming skills. The experiment results shown that ResNet-IBN model achieved the highest results on custom dataset MMUVD_1500 with mAP of 87.63% and Rank@1 of 84.68% respectively. For instance, users would be able to input query vehicle images and receive matched gallery images from different camera viewpoints through the application interface. Thus, this makes it easier for users to track vehicles across multiple locations, enhance the usability and broaden the accessibility of vehicle Re-ID tasks. The final outcome is a complete software solution with a user-friendly interface that allows users to perform vehicle Re-ID tasks effortlessly.

Keywords—Vehicle Re-identification, Vehicle Re-ID, Residual Network, ResNet-IBN, Deep Learning, Traffic.

Received: 10 July 2025; Accepted: 12 October 2025; Published: 16 February 2026

Journal of Informatics and Web Engineering
https://doi.org/10.33093/jiwe.2026.5.1.17

Published by MMU Press. URL: https://journals.mmupress.com/jiwe

1. INTRODUCTION

Re-ID is a computer vision task that involves identifying the same vehicle across multiple camera viewpoints with different viewing angles. This task is applicable for traffic impact assessment systems as it can be used in urban areas to improve the traffic flow and management. Vehicle Re-ID is based on extracting the features of a vehicle and comparing the features obtained in the different camera views [1], [2] in order to determine whether the same vehicle has been observed in the other camera locations. This would be done using deep learning algorithms to enhance the general accuracy and performance in recognition and classification of various types of vehicles in camera based in ever changing environment and situations [3].

However, with the different angles of the camera, motion blur and lighting differences on the moving vehicle, the same vehicle may look like different vehicles. The same vehicle would have the same characteristics when they appeared in other camera angles [4]. This makes it more challenging to the system to recognize the vehicle in case of multiple and varied camera positions and viewing angles. This is known as intra-class variability. Therefore, the objective of this research work is to identify unique and distinguishable features from the same vehicles that appear in different viewing angles and conditions. In the following sections, section 2 presents an overview of existing works on vehicle Re-ID. This is followed by research methodology used to propose our vehicle Re-ID method and findings in Section 3. Section 4 mainly shows the experiment and visual results obtained. At last, best results and future plans are concluded in Section 5.

2. LITERATURE REVIEW

2.1 Overview

Vehicle Re-ID has developed and grown in recent years due to the advancement of transportation and heavy traffic around the world. The fast-growing number of vehicles, including the autonomous driving vehicle, smart vehicles and electric vehicles, it is no doubt that all these vehicles contribute to the increasing number of vehicles on the road to this date [3]. Therefore, vehicle Re-ID is a crucial task to help improve the management of traffic and traffic flow analysis. Vehicle Re-ID models can be labelled into two types of learning, supervised and unsupervised learning [3]. Most of the vehicle Re-ID systems are supervised learning, which the model uses labelled training data for model training, in this case the data would be annotated to show the vehicles in each frame of camera footage from multiple viewing angles. On the other hand, unsupervised learning would use unlabelled data to perform model training.

2.2. Supervised Learning

A plethora of research related to supervised learning has been conducted to develop highly accurate vehicle Re-ID models by leveraging labelled datasets to learn discriminative and robust feature representations across different camera views and environmental conditions. Initially, Franco et al. [5] created a lightweight vehicle Re-ID system that integrates approximate nearest neighbour search with a sample mining strategy to boost feature extraction quality using CPU hardware only. It achieved a mAP of 86.20% on VeRi dataset with their Coupled Loss model. Moral, Garcia-Martin and Martinez [6] introduced an ensemble of orientation and appearance DL features combined with multi-loss optimization in both image and video-based settings. It reached 36.26% mAP on the CityFlow-ReID dataset. Zhou et al. [7] designed an Attention-Aware Network by fusing channel and spatial attention modules under a multi-loss joint training regime to address sample imbalance. The results shown 75.14% mAP on VeRi-776 dataset. Zhu et al. [8] presented VOC-ReID, that reframed shape and background similarity as orientation and camera Re-ID and merged distance matrices across modalities to achieve 78.10% mAP on CityFlow and 79.70% on VeRi-776.

Furthermore, Dilshad and Song [9] developed a dual-stream Siamese network with dilated convolutions for simultaneous shape and license-plate feature extraction, achieving 95.70% accuracy in their two-stream strategy. Kamenou et al. [10] introduced a multi-level Re-ID network utilizing a structured metric learning loss adapted for vehicle Re-ID, improving mAP by 6% to 12% across image-to-image and tracklet benchmarks on CityFlow-ReID and VeRi-776. Truong and Mei [11] addressed domain variability by building a MegaVeriVehicle dataset through multi-domain learning on selected external data, achieving up to 85.23% mAP on the combined dataset. Zhao et al. [12] applied a novel Pompeiu-Hausdorff distance learning method on their VVeRI-901 video dataset to enhance video-to-video matching, attaining mAP of 47.20%. Zheng et al. [13] created VehicleNet by combining four public datasets with a two-stage progressive approach, securing mAP of 86.07%.

In addition, Chen et al. [14] proposed the Viewpoint-Aware Loss function with a multi-decision boundary mechanism to learn viewpoint invariant representations, and obtained mAP of 81.36% on VeRi-776 dataset. Gao et al. [15] embed a multi-dimensional attention network combining height, channel and width modules to refine feature extraction, reaching 78.33% mAP on VeRi-776 and up to 84.17% on VeRi-Wild. Kumar et al. [16] released the CARLA Re-ID dataset and utilized YOLOv7, SORT and OSNet with trajectory metrics to achieve 98.70% mAP and 99.50% Top-1 accuracy. Martinel et al. [17] introduced a Hanoi pooling layer to suppress orientation-induced variance without extra labels, obtaining up to 99.02% HIT@5 on VeRi-776. Shi et al. [18] proposed LG-CoT, a transformer framework with a local-attention-guided post-optimization module, achieving 79.70% mAP on VeRi-776 and 90.50% on VehicleID Test800. Spagnolo et al. [19] leveraged a CNN to learn key vehicle characteristics beyond license plates, attaining Rank-10 of 97.32% on VeRi-776 dataset. Sun et al. [20] presented SCAN, a collaborative attention network with refined local and edge modules, reaching 83.30% mAP on VeRi-776 and 98.10% Rank-5 on VehicleID. Yang et al. [21] developed the TIMS framework combining multi-level features and spatial-temporal camera graph inference for network-level traffic sensing, recording an F1 score of 75.34% on their LHTV dataset. Zhang et al. [22] applied vehicle Re-ID to travel time distribution estimation in lane-level, achieving 5.43s of mean error and 0.63s of standard deviation.

Moreover, Almeida et al. [23] proposed a lightweight multi-branch deep architecture with Loss-Branch-Split and grouped convolutions, yielding 84.72% mAP on VeRi-776 and up to 93.46% on VeRi-Wild. Du et al. [24] integrated side information embeddings and visual prompt tuning into a ViT-based framework to reduce variance and obtained mAP of 82.10% on VeRi-776 dataset. Zhao et al. [25] created DSTNet, a dynamic-static transformer network with high-resolution constraint inputs to harmonize HR and super-resolved images, achieving 78.00% mAP on VeRi-776.

Additionally, Jiao et al. [26] introduced the VRAI dataset for aerial imagery and proposed an Orientation Adaptive and Salience Attentive network, obtaining 85.10% mAP on VRAI-Image. Li et al. [27] tackled day and night cross-domain Re-ID with the DNDM framework and the DN-Wild dataset, reporting up to 92.60% Rank-5 in night-to-day scenarios. Lian et al. [28] presented MED, a multi-branch enhanced discriminative network using spatial sub-maps for fine-grained features, achieving 83.40% mAP on VeRi-776. Sheng et al. [29] developed the Co-occurrence Attention Net with global and aware branches plus a partition-reunion loss, reaching 83.43% mAP on VeRi-776. Lastly, Wang et al. [30] proposed BCV-ReID, a blockchain-based system combining viewpoint-identity queries and a secure VehicleChain, achieving 83.74% mAP on VeRi-776 and up to 87.48% on VeRi-Wild.

2.3. Unsupervised Learning

Besides that, unsupervised vehicle Re-ID is another research topic that uses unlabelled data to develop the Re-ID system, compared to supervised which utilize the labelled datasets to perform model training. Wang et al. [31] introduced an unsupervised vehicle Re-ID framework based on FDN and hard triplet centre loss that clustered unlabelled samples across camera viewpoints, achieving up to mAP of 36.90% on VeRi-776 and mAP of 61.60% on VehicleID. He et al. [32] proposed a multi-level progressive learning method with multi-branch feature extraction, density-based pseudo-label clustering, and dynamic contrast learning, which yielded mAP of 45.10% on VeRi-776 and 65.30% on VehicleID respectively.

Tao et al. [33] developed a Semantic Camera Self-Aware Contrastive Learning framework that leveraged spatial transformer–derived semantic attributes and a camera-aware contrastive loss for cross-camera clustering, reaching 43.00% mAP and 63.90% mAP on VeRi-776 and VehicleID respectively. Finally, Yu et al. [34] presented an approach using automatically available camera and tracklet IDs to generate pseudo labels, reporting 59.0% mAP on VeRi-776, up to Rank-1 of 83.0% on VeRi-Wild, and 32.34% mAP on VVeRI-901.

3. RESEARCH METHODOLOGY

One of the most significant challenges in completing vehicle Re-ID tasks is the lack of accessible and user-friendly tools to prepare high-quality datasets specifically for this task. To address this challenge, we propose a comprehensive solution, which is a user-friendly, Python-based application designed specifically to streamline the entire pipeline of vehicle Re-ID tasks as shown in Figure 1.

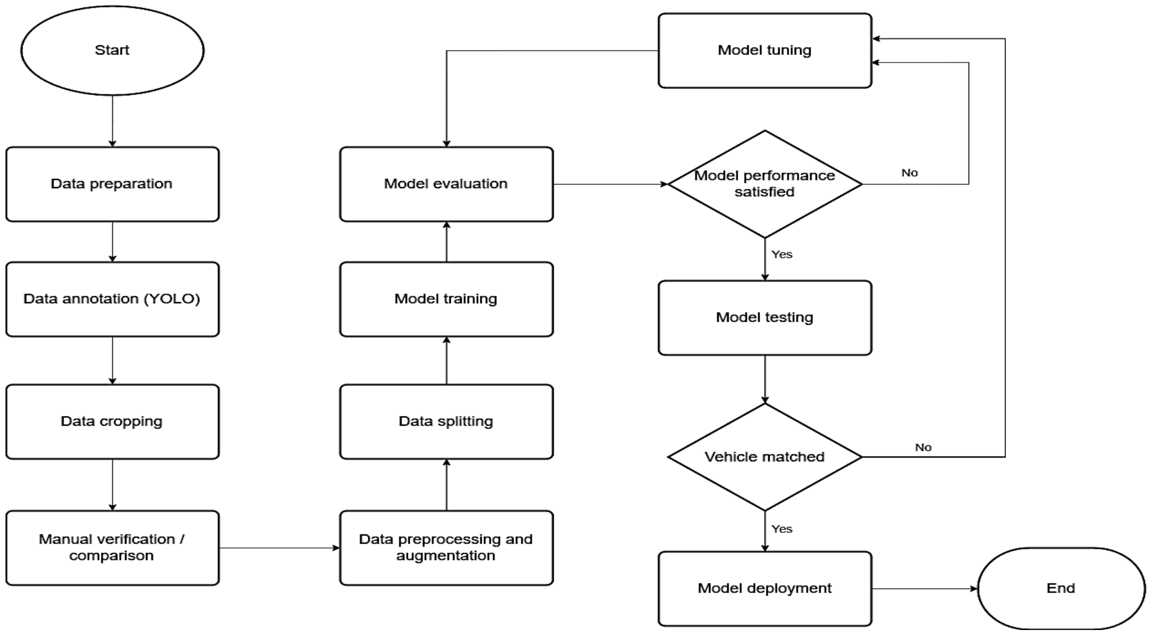

Figure 1. System Flowchart

The research work starts with data preparation. Data preparation involves collecting, processing, and organizing raw data into a format suitable for training and testing Re-ID models. This may be a time-consuming and tedious process in traditional workflow. This usually needs manual labelling or manual annotation of data. In this case though, the object detection model, in the form of YOLO [35], is used to simplify the task. This pretrained model assists in automation of detection and tracking of vehicle images within video footage. Additionally, the tracking mechanism in YOLO is provided by BoT-SORT [36], which is an advanced tracking algorithm capable of providing each vehicle under tracking a unique identifier (ID). This unique ID enables the system to maintain consistent identification of vehicles across frames, even when they move or change positions within the video.

Subsequently, by passing an object detection algorithm like YOLO or other algorithms to detect vehicles, every vehicle detected is given an individual identifier depending on how it looks and moves in one video. This ID is however not persistent between videos or camera feeds. Thus, the feature comparing images in the application is presented to overcome this problem. This aspect enables the users to compare the images of the vehicles side-by-side manually to confirm and decide whether they are of the same vehicle. Besides that, the application also provided the augmentation process that is powered by Albumentations [37], a highly efficient and versatile Python library specifically designed for image augmentation in deep learning and computer vision tasks.

In vehicle Re-ID, the dataset must be split in a way that reflects real-world challenges to effectively evaluate the capability of the model. Simply dividing the data randomly as in standard classification tasks would result in overlapping information between the training and testing sets, making the evaluation overly optimistic and unreliable. For this reason, vehicle Re-ID datasets are carefully organized into three main subsets, which are the training set, validation set and test set that are further divided into query set and gallery set. Each subset serves a different purpose, and understanding their roles is essential for ensuring a robust and meaningful evaluation.

Model training is one of the core processes in the vehicle Re-ID project, where the application utilizes state-of-the-art methods and algorithms commonly used in both person and vehicle Re-ID research. The training script used in the application is an open-source repository developed for person Re-ID tasks, but it works for vehicle Re-ID with slight modification of the code [38]. This task transforms the prepared dataset into a robust model capable of identifying and matching vehicles across different camera views.

The training script used in the application is a deep learning-based architecture specifically designed for Re-ID tasks. The script supports multiple backbone architectures, including ResNet-IBN ResNet with Instance Batch Normalization), a variant of ResNet [39] that incorporates Instance Batch Normalization layers to improve generalization across domains [40]. Furthermore, DenseNet [41], Swin Transformer [42], NASNet [43], HRNet [44] and EfficientNet [45] are other popular architectures that can be selected based on the user's preference. These backbones extract high-dimensional feature embeddings from input images, capturing discriminative information about the vehicles. On top of the backbone, a classifier head is added to map the feature embeddings to class labels, which are the vehicle unique IDs. The script allows customization of the embedding dimension and the final classification layer.

Additionally, the script also supports multiple loss functions such as circle loss that focuses on optimizing similarity scores between positive and negative pairs, triplet loss that encourages the model to make the distance between positive pairs as minimal as possible while maximizing the distance between negative pairs, and contrastive loss which is similar to triplet loss but operates on pairs of samples. These loss functions are combined with mining strategies to select hard-positive and hard-negative pairs during training.

Apart from this, the script also allows data augmentation like colour jittering, random horizontal flipping and random erasing in order to diversify the training data. Other than that, the model is optimized with SGD with momentum and decay of weights. The scheduling of the learning rate is done in the form of step decay or cosine annealing based on the preference of the user. In addition, gradient clipping is used to avoid divergence during training.

4. RESULTS AND DISCUSSIONS

4.1 mAP

mAP evaluates the entire ranked list for each query by calculating the average precision (AP) and then averaging over all queries. The equation (1) of the mAP is defined as:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{1}$$

where N is the total number of queries. mAP provides a broader measure of performance across all ranks and balancing both precision and recall. A higher mAP specified better overall ranking performance of the Re-ID model by balancing both the number of retrieved items that are applicable, known as precision, and the number of applicable items that are retrieved, known as recall. This metric is widely used in object detection and information retrieval as it encompasses the precision and recall trade off relationship, including both FP and FN [46].

4.2 Rank@K (Rank@1, Rank@5, Rank@10)

For a retrieval-based Re-ID task, the Rank@K metric is defined as the proportion of queries whose first correct match appears at rank K or in the equation (2) of:

$$Rank@K = \frac{1}{Q}\sum_{i=1}^{Q} 1(rank_i \leq K) \tag{2}$$

where Q showed the total number of query images, $rank_i$ denoted as the position of the first correctly matched gallery image for query i and $1(\cdot)$ suggested as the indicator function, yielding 1 if its argument is true and 0 otherwise. By varying K, each $Rank@K$ represent the fraction of queries with a correct match within the top 1, 5 and 10 retrievals respectively. These metrics emphasize how well the model could find relevant results early in the ranked list and they are commonly used in person and vehicle Re-ID benchmarks [47].

4.3 Datasets

The experiment leverages two different types of datasets, which are the custom dataset generated through the application provided task and public benchmark datasets created for vehicle Re-ID tasks. This dual-dataset approach enables a comprehensive evaluation of the vehicle Re-ID model's performance under both controlled experimental conditions and real-world scenarios while conforming to broader research standards. To validate the vehicle Re-ID model against established research standards, two widely recognized public vehicle Re-ID benchmarks are used, which are VeRi-776 [48] and VRIC [49] datasets respectively.

In this experiment, the custom dataset is generated from four synchronized cameras positioned at a four-way junction, each capturing traffic flow from one direction, which is north, south, east and west respectively. Both custom datasets, MMUVD_0800 and MMUVD_1500 [50], [51] consist of wide variety of vehicle types, where each vehicle is captured during sunny daytime with good lighting condition and minor occlusion cases where the vehicle is blocked out of view from large size vehicle. The custom datasets are split with train, validation and test split ratio of 60%, 20% and 20% respectively. These sets are also labelled with a name format of MMUVD_0800 where MMUVD is the location name and 0800 is the time recorded in 24-hour time format. Therefore, these subsets further test the model's robustness to illumination changes and sensor noise.

Additionally, the custom datasets created may encounter class imbalance challenges where certain vehicles appear more often in each different camera. As a solution to this, data augmentation can be done as re-sampling technique to increase the number of vehicle images to reduce model bias.

Before conducting the experiment, both public datasets are split with a train-validation ratio of 75% and 25% respectively. Note that the test subset for both datasets was originally available which consists of the gallery and query subsets. Thus, only the train subset is further split into train and validation set for a more balanced and fair comparison results.

4.4 Experiment Results

Table 1 shows the overall summary of the comparison results between each different types of model and metric losses used for public benchmark datasets and the custom datasets generated for this experiment. The model ResNet50-IBN has the best performance compared to other model types, with even higher results if metric losses are applied. The results shown that the custom dataset MMUVD_1500 achieved the highest mAP of 87.63%, while MMUVD_0800 obtained the highest Rank@1 of 85.53% with circle and contrastive metric losses respectively.

Table 1. Summary of Experiment Results

Models	Metric Losses	Public Datasets				Custom Datasets			
		VeRi-776		VRIC		MMUVD_0800		MMUVD_1500	
		mAP	Rank@1	mAP	Rank@1	mAP	Rank@1	mAP	Rank@1
ResNet50-ibn (Proposed method)	-	67.46	61.50	72.50	68.40	80.24	76.97	81.91	79.03
	Triplet	68.81	63.50	75.90	71.90	85.05	82.89	83.70	80.65
	Contrast	**72.11**	**66.50**	76.20	72.30	86.37	84.21	85.09	82.26
	Circle	70.93	65.50	76.20	72.10	85.70	83.55	85.81	83.06
	Circle+Contrast	71.17	66.00	**77.10**	**73.10**	**87.28**	**85.53**	**87.63**	**84.68**
	Triplet+Contrast	69.00	63.00	68.32	63.35	84.44	81.58	81.11	77.42
	Triple+Circle	71.19	65.50	67.79	63.02	83.73	80.92	86.29	83.87
ResNet50	-	59.56	52.50	69.20	64.60	76.30	72.37	82.36	79.84
DenseNet121	-	65.22	58.50	70.20	65.80	78.68	75.00	84.94	82.26
HRNet	-	69.12	63.50	74.20	70.40	81.19	78.29	86.35	83.87
EfficientNet-b0	-	59.72	53.50	68.20	63.70	80.24	76.97	79.04	75.00

Table 2 shows the complete experiment results conducted on the custom dataset MMUVD_0800, including metrics such as mAP, Rank@1, Rank@5 and Rank@10 respectively. The dataset achieved the highest mAP of 87.28% and

Rank@1 of 85.53% using ResNet50-ibn model type with circle and contrastive metric losses. On the other hand, the dataset obtained the highest Rank@5 and Rank@10 of 94.08% and 96.05% with contrastive metric loss respectively.

Table 2. Experiment Results of Custom Dataset MMUVD_0800

Models	Metric Losses	mAP	Rank@1	Rank@5	Rank@10
ResNet50-ibn (Proposed method)	-	80.24	76.97	91.45	93.42
	Triplet	85.05	82.89	90.79	92.76
	Contrast	86.37	84.21	**94.08**	**96.05**
	Circle	85.70	83.55	93.42	94.08
	Circle+Contrast	**87.28**	**85.53**	92.76	94.08
	Triplet+Contrast	84.44	81.58	92.76	94.74
	Triple+Circle	83.73	80.92	92.76	94.74
ResNet50	-	76.30	72.37	88.82	93.42
DenseNet121	-	78.68	75.00	90.13	94.08
HRNet	-	81.19	78.29	90.79	92.76
EfficientNet-b0	-	76.59	73.03	88.82	90.13

Table 3 shows the complete experiment results conducted on the custom dataset MMUVD_1500. The dataset achieved the highest mAP of 87.63%, Rank@1 of 84.68%, Rank@5 of 96.77% and Rank@10 of 97.58% using ResNet50-ibn model type with circle and contrastive metric losses respectively. This dataset obtained the best results compared to other datasets in terms of all metrics.

Table 3. Experiment Results of Custom Dataset MMUVD_1500

Models	Metric Losses	mAP	Rank@1	Rank@5	Rank@10
ResNet50-ibn (Proposed method)	-	81.91	79.03	92.74	95.97
	Triplet	83.70	80.65	93.55	95.97
	Contrast	85.09	82.26	95.16	96.77
	Circle	85.81	83.06	96.77	97.58
	Circle+Contrast	**87.63**	**84.68**	**96.77**	**97.58**
	Triplet+Contrast	81.11	77.42	94.35	96.77
	Triple+Circle	86.29	83.87	94.35	97.58
ResNet50	-	82.36	79.84	90.32	93.55
DenseNet121	-	84.94	82.26	93.55	96.77
HRNet	-	86.35	83.87	95.16	96.77
EfficientNet-b0	-	79.04	75.00	93.55	95.16

Figures 2 and 3 illustrate the visual sample of the vehicle Re-ID model trained with custom datasets MMUVD_0800 and MMUVD_1500.

Figure 2. Custom Dataset MMUVD_0800 Visual Result

Figure 3. Custom Dataset MMUVD_1500 Visual Result

The model variation used in this visualization is ResNet50-IBN model type with circle and contrastive metric losses. The image with black border represents the query image, and the rest of the images are gallery images. Correctly matched gallery images are shown with green borders whereas mismatched gallery images are shown with red borders. Each of the vehicle images are labelled with their corresponding similarity score and vehicle image filename that consists of useful metadata such as its location name, camera ID, frame ID and unique vehicle ID as shown in Figure 4.

Figure 4. Vehicle Metadata with Similarity Score

Figure 5 and 6 illustrates the visual sample of the vehicle Re-ID model trained with public datasets. The model variation used in this visualization is ResNet50-IBN model type with contrast metric loss for VeRi-776 dataset and circle, contrastive metric losses for VRIC dataset. The similarity scores for both datasets are consistently high across every vehicle tested.

Figure 5. Public Dataset VeRi-776 Visual Result

Figure 6. Public Dataset VRIC Visual Result

Figure 7 illustrates the design of the software solution developed for users to perform vehicle Re-ID tasks. This application integrates multiple functionalities, starting from dataset preparation and data augmentation to model training, testing, and visualization. Furthermore, vehicle image comparison is one of the key highlights of this application. This aspect enables the user to compare the images of vehicles taken at different camera angles in a side-by-side manner on the application so that the final dataset is a true reflection of the real-life diversity that may be encountered in various vehicle Re-ID situations. The application also automates the process of saving these comparisons, enabling users to effortlessly generate a high-quality dataset that suits their specific needs. The application also automates the task of saving these comparisons so that users can easily create a good quality dataset that fits their exact requirements.

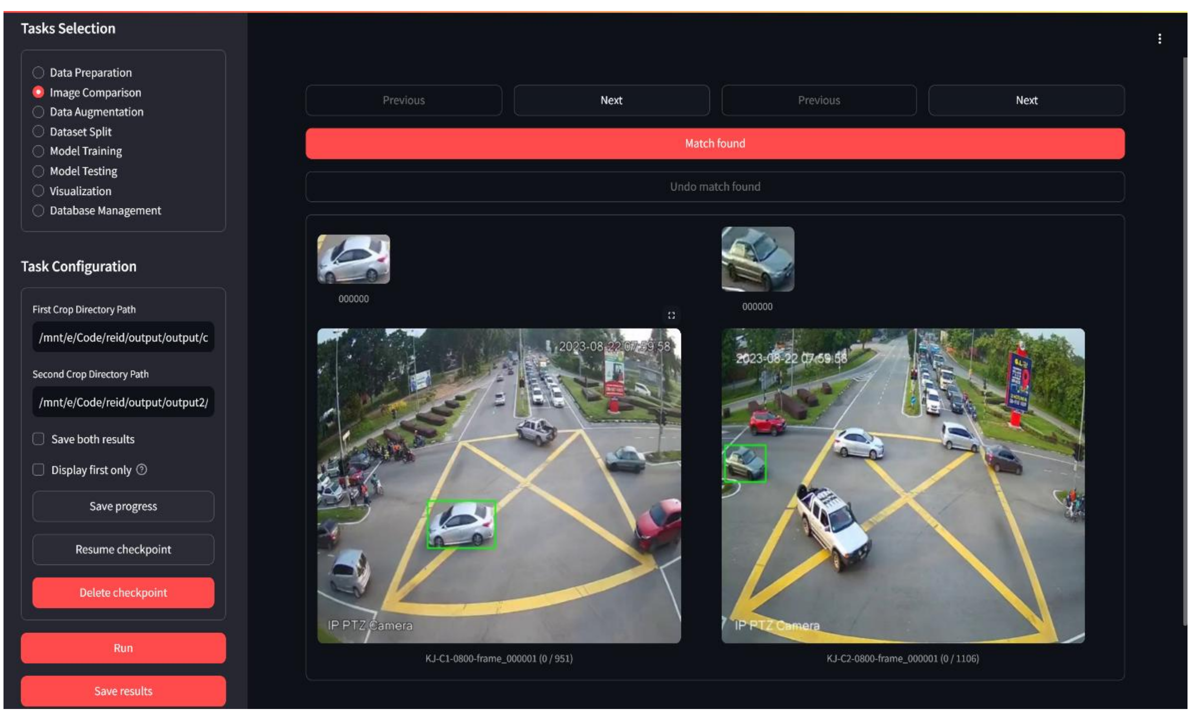

Figure 7. Application User Interface

5. CONCLUSION

In conclusion, we successfully developed an application for Re-ID tasks using deep learning, addressing the need for a more accessible solution compared to traditional coding approaches. The proposed Re-ID model, specifically ResNet-IBN with metric losses demonstrated strong performance across all project objectives, effectively identifying and matching vehicles under various conditions. Final experiment results shown that the custom dataset MMUVD_1500 model outperformed public datasets with mAP of 87.63%, Rank@1 of 84.68%, Rank@5 of 96.77% and Rank@10 of 97.58% respectively.

Future works could include integrating real-time processing capabilities for instant vehicle Re-ID in traffic monitoring systems. This would allow for immediate vehicle tracking across multiple camera feeds, making it more suitable for live traffic monitoring. Besides that, extending the application's functionality to include traffic flow analysis and anomaly detection could make it a comprehensive tool for traffic impact assessment.

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers for the suggestions to improve the paper.

FUNDING STATEMENT

The authors gratefully acknowledge the support provided by TM R&D and Multimedia University (MMU), Malaysia for this project, funded under Grant ID: MMUE/250015.

AUTHOR CONTRIBUTIONS

Wei-Jie Low: Conceptualization, Data Curation, Methodology, Validation, Writing – Original Draft Preparation.
Kah-Ong Michael Goh: Supervision, Ideation, Acquire and Editing.
Check-Yee Law: Facilitate, Review & Editing.
Connie Tee: Result Verification & Review.
Yong-Wee Sek: Data Validation & Review.
Md Ismail Hossen: Review & Editing.

CONFLICT OF INTERESTS

No conflict of interests were disclosed.

ETHICS STATEMENTS

Our publication ethics follow The Committee of Publication Ethics (COPE) guideline. https://publicationethics.org/.

DATA AVAILABILITY

- The data that support the findings of this study are openly available in VeRi-776 at http://doi.org/10.1109/ICME.2016.7553002. These data were derived from sources in the public domain https://github.com/VehicleReId/VeRi.
- The data that support the findings of this study are openly available in VRIC at http://doi.org/10.48550/arXiv.1809.09409. These data were derived from sources in the public domain https://qmul-vric.github.io/.
- The data that support the findings of this study are openly available in MMUVD at http://doi.org/10.62527/joiv.8.1.2313. These data were derived from sources in the public domain https://www.kaggle.com/datasets/michaelgoh/road-vehicle.

REFERENCES

[1] S. Teng, and T. Dong, "Unsupervised vehicle re-identification via raw UAV videos," in *Image and Graphics*, H. Lu, W. Ouyang, H. Huang, J. Lu, R. Liu, J. Dong, and M. Xu, Eds., Cham: Springer Nature Switzerland, pp. 360–372, 2023, doi: 10.1007/978-3-031-46308-2_30.

[2] H. Taleb, "Vehicle re-identification using deep learning methods with a focus on challenges and recent solutions," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 4, Art. no. 4, Jun 2024.

[3] A. Amiri, A. Kaya, and A. S. Keceli, "A comprehensive survey on deep-learning-based vehicle re-identification: models, data sets and challenges," *arXiv*: arXiv:2401.10643, Jan 2024, doi: 10.48550/arXiv.2401.10643.

[4] H. Li *et al.*, "Attributes guided feature learning for vehicle re-identification," *arXiv*: arXiv:1905.08997, Nov. 2021, doi: 10.48550/arXiv.1905.08997.

[5] A. O. R. Franco *et al.*, "Vehicle re-identification by deep feature embedding and approximate nearest neighbors," in 2020 *International Joint Conference on Neural Networks (IJCNN),* Glasgow, United Kingdom: IEEE, pp. 1–8, Jul. 2020, doi: 10.1109/IJCNN48605.2020.9207150.

[6] P. Moral, A. Garcia-Martin, and J. M. Martinez, "Vehicle re-identification in multi-camera scenarios based on ensembling deep learning features," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, WA, USA: IEEE, pp. 2574–2580, Jun. 2020, doi: 10.1109/CVPRW50498.2020.00310.

[7] H. Zhou, C. Li, L. Zhang, and W. Song, "Attention-aware network and multi-loss joint training method for vehicle re-identification," in *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Chongqing, China: IEEE, pp. 1330–1334, Jun. 2020, doi: 10.1109/ITNEC48623.2020.9085181.

[8] X. Zhu, Z. Luo, P. Fu, and X. Ji, "VOC-RelD: Vehicle re-identification based on vehicle-orientation-camera," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, WA, USA: IEEE, pp. 2566–2573, Jun. 2020, doi: 10.1109/CVPRW50498.2020.00309.

[9] N. Dilshad, and J. Song, "Dual-stream siamese network for vehicle re-identification via dilated convolutional layers," in *2021 IEEE International Conference on Smart Internet of Things (SmartIoT)*, Jeju, Korea, Republic of: IEEE, pp. 350–352, Aug. 2021, doi: 10.1109/SmartIoT52359.2021.00065.

[10] E. Kamenou, J. M. Del Rincon, P. Miller, and P. Devlin-Hill, "Multi-level deep learning vehicle re-identification using ranked-based loss functions," in *2020 25th International Conference on Pattern Recognition (ICPR)*, Milan, Italy: IEEE, pp. 9099–9106, Jan. 2021, doi: 10.1109/ICPR48806.2021.9412415.

[11] Q. Truong, and B. Mei, "Not all data matters: An efficient approach to multi-domain learning in vehicle re-identification," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, Indianapolis, IN, USA: IEEE, pp. 1278–1285, Sep. 2021, doi: 10.1109/ITSC48978.2021.9564945.

[12] J. Zhao, F. Qi, G. Ren, and L. Xu, "PhD Learning: Learning with Pompeiu-hausdorff Distances for video-based vehicle re-identification," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA: IEEE, pp. 2225–2235, Jun. 2021, doi: 10.1109/CVPR46437.2021.00226.

[13] Z. Zheng, T. Ruan, Y. Wei, Y. Yang, and T. Mei, "VehicleNet: Learning robust visual representation for vehicle re-identification," *IEEE Transactions on Multimedia*, vol. 23, pp. 2683–2693, 2021, doi: 10.1109/TMM.2020.3014488.

[14] H. Chen, Y. Liu, Y. Huang, W. Ke, and H. Sheng, "Partition and reunion: A viewpoint-aware loss for vehicle re-identification," in *2022 IEEE International Conference on Image Processing (ICIP)*, Bordeaux, France: IEEE, pp. 2246–2250, Oct. 2022, doi: 10.1109/ICIP46576.2022.9897541.

[15] Y. Gao, L. Lu, B. Xu, and D. Chu, "Multi-dimensional attention network for vehicle re-identification," in *2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, Nanjing, China: IEEE, pp. 1–5, Oct. 2022, doi: 10.1109/CVCI56766.2022.9965183.

[16] A. Kumar, T. Kashiyama, H. Maeda, F. Zhang, H. Omata, and Y. Sekimoto, "Vehicle re-identification and trajectory reconstruction using multiple moving cameras in the CARLA driving simulator," in *2022 IEEE International Conference on Big Data (Big Data)*, Osaka, Japan: IEEE, pp. 1858–1865, Dec. 2022, doi: 10.1109/BigData55660.2022.10020814.

[17] N. Martinel, M. Dunnhofer, R. Pucci, G. L. Foresti, and C. Micheloni, "Lord of the Rings: Hanoi pooling and self-knowledge distillation for fast and accurate vehicle reidentification," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 1, pp. 87–96, Jan. 2022, doi: 10.1109/TII.2021.3068927.

[18] Y. Shi, X. Zhang, and X. Tan, "Local-guided global collaborative learning transformer for vehicle reidentification," in *2022 IEEE 34th International Conference on Tools with Artificial Intelligence (ICTAI)*, Macao, China: IEEE, pp. 793–798, Oct. 2022, doi: 10.1109/ICTAI56018.2022.00122.

[19] P. Spagnolo, P. L. Mazzeo, F. Buccoliero, P. L. Carcagni, and C. Distante, "A deep learning approach for vehicle re-identification," in *2022 7th International Conference on Smart and Sustainable Technologies (SpliTech)*, Split /Bol, Croatia: IEEE, pp. 1–6, Jul. 2022, doi: 10.23919/SpliTech55088.2022.9854225.

[20] Y. Sun *et al.*, "SCAN: Stronger Collaborative Attention Network for vehicle re-identification," in *2022 International Joint Conference on Neural Networks (IJCNN)*, Padua, Italy: IEEE, pp. 1–8, Jul. 2022, doi: 10.1109/IJCNN55064.2022.9892171.

[21] H. Yang, J. Cai, M. Zhu, C. Liu, and Y. Wang, "Traffic-Informed Multi-Camera Sensing (TIMS) system based on vehicle re-identification," *IEEE Transactions on Intelligent Transportation System*, vol. 23, no. 10, pp. 17189–17200, Oct. 2022, doi: 10.1109/TITS.2022.3154368.

[22] C. Zhang *et al.*, "Vehicle re-identification for lane-level travel time estimations on congested urban road networks using video images," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 12877–12893, Aug. 2022, doi: 10.1109/TITS.2021.3118206.

[23] E. Almeida, B. Silva, and J. Batista, "Strength in diversity: Multi-branch representation learning for vehicle re-identification," in 2*023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, Bilbao, Spain: IEEE, pp. 4690–4696, Sep. 2023, doi: 10.1109/ITSC57777.2023.10422175.

[24] L. Du, K. Huang, and H. Yan, "ViT-ReID: A vehicle re-identification method using visual transformer," in *2023 3rd International Conference on Neural Networks, Information and Communication Engineering (NNICE)*, Guangzhou, China: IEEE, pp. 287–290, Feb. 2023, doi: 10.1109/NNICE58320.2023.10105738.

[25] C. Zhao, M. Wang, N. Su, Y. Yan, and S. Feng, "DSTNet: Dynamic-Static Transformer Style Network for cross-resolution vehicle reidentification," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1–5, 2023, doi: 10.1109/LGRS.2023.3270883.

[26] B. Jiao, L. Yang, L. Gao, P. Wang, S. Zhang, and Y. Zhang, "Vehicle re-identification in aerial images and videos: Dataset and approach," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 3, pp. 1586–1603, Mar. 2024, doi: 10.1109/TCSVT.2023.3298788.

[27] H. Li, J. Chen, A. Zheng, Y. Wu and Y. Luo, "Day-night cross-domain vehicle re-identification," *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 12626-12635, 2024, doi: 10.1109/CVPR52733.2024.01200.

[28] J. Lian, D.-H. Wang, Y. Wu, and S. Zhu, "Multi-Branch enhanced discriminative network for vehicle re-identification," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 2, pp. 1263–1274, Feb. 2024, doi: 10.1109/TITS.2023.3316068.

[29] H. Sheng *et al.*, "Discriminative feature learning with co-occurrence attention network for vehicle ReID," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 5, pp. 3510–3522, May 2024, doi: 10.1109/TCSVT.2023.3326375.

[30] S. Wang, D. Yang, H. Sheng, J. Shen, Y. Zhang, and W. Ke, "A blockchain-enabled distributed system for trustworthy and collaborative intelligent vehicle re-identification," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 2, pp. 3271–3282, Feb. 2024, doi: 10.1109/TIV.2023.3347267.

[31] H. Wang, J. Peng, G. Jiang, and X. Fu, "Learning multiple semantic knowledge for cross-domain unsupervised vehicle re-identification," in *2021 IEEE International Conference on Multimedia and Expo (ICME)*, Shenzhen, China: IEEE, pp. 1–6, Jul. 2021, doi: 10.1109/ICME51207.2021.9428440.

[32] Z. He, H. Zhao, J. Wang, and W. Feng, "Multi-level progressive learning for unsupervised vehicle re-identification," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 4, pp. 4357–4371, Apr. 2023, doi: 10.1109/TVT.2022.3228127.

[33] X. Tao, J. Kong, M. Jiang, and X. Luo, "Semantic camera self-aware contrastive learning for unsupervised vehicle re-identification," *IEEE Signal Processing Letters*, vol. 31, pp. 2175–2179, 2024, doi: 10.1109/LSP.2024.3449233.

[34] J. Yu, H. Oh, M. Kim, and J. Kim, "Weakly supervised contrastive learning for unsupervised vehicle reidentification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 11, pp. 15543–15553, Nov. 2024, doi: 10.1109/TNNLS.2023.3288139.

[35] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, real-time object detection," in *2016 IEEE conference on computer vision and pattern recognition (CVPR)*, Las Vegas, NV, USA: IEEE, pp. 779–788, Jun. 2016, doi: 10.1109/CVPR.2016.91.

[36] N. Aharon, R. Orfaig, and B.-Z. Bobrovsky, "BoT-SORT: Robust Associations multi-pedestrian tracking," *arXiv*: arXiv:2206.14651, Jul. 07, 2022, doi: 10.48550/arXiv.2206.14651.

[37] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, "Albumentations: Fast and flexible image augmentations," *Information*, vol. 11, no. 2, Art. no. 2, Feb. 2020, doi: 10.3390/info11020125.

[38] Z. Zheng, X. Yang, Z. Yu, L. Zheng, Y. Yang, and J. Kautz, "Joint discriminative and generative learning for person re-identification," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA: IEEE, pp. 2133–2142, Jun. 2019, doi: 10.1109/CVPR.2019.00224.

[39] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *arXiv*: arXiv:1512.03385, Dec. 10, 2015, doi: 10.48550/arXiv.1512.03385.

[40] X. Pan, P. Luo, J. Shi, and X. Tang, "Two at once: Enhancing Learning and Generalization Capacities via IBN-Net," *arXiv*: arXiv:1807.09441, Mar. 23, 2020, doi: 10.48550/arXiv.1807.09441.

[41] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," *arXiv*: arXiv:1608.06993, Jan. 28, 2018, doi: 10.48550/arXiv.1608.06993.

[42] Z. Liu *et al.*, "Swin transformer: Hierarchical vision transformer using shifted windows," *arXiv*: arXiv:2103.14030, Aug. 17, 2021, doi: 10.48550/arXiv.2103.14030.

[43] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," *arXiv*: arXiv:1707.07012, Apr. 11, 2018, doi: 10.48550/arXiv.1707.07012.

[44] J. Wang *et al.*, "Deep high-resolution representation learning for visual recognition," *arXiv*: arXiv:1908.07919, Mar. 13, 2020, doi: 10.48550/arXiv.1908.07919.

[45] M. Tan, and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," *arXiv*: arXiv:1905.11946, Sep. 11, 2020, doi: 10.48550/arXiv.1905.11946.

[46] V7, "Mean Average Precision (mAP) explained: Everything you need to know," V7labs, Mar. 7, 2022. [Online]. Available: https://www.v7labs.com/blog/mean-average-precision [Accessed: Apr. 23, 2025].

[47] D. Xu *et al.*, "Vehicle re-identification system based on appearance features," *Security and Communication Networks*, vol. 2022, Jan. 2022, doi: 10.1155/2022/1833362.

[48] X. Liu, W. Liu, H. Ma, and H. Fu, "Large-scale vehicle re-identification in urban surveillance videos," in *2016 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, Jul. 2016, doi: 10.1109/ICME.2016.7553002.

[49] A. Kanaci, X. Zhu, and S. Gong, "Vehicle re-identification in context, " *arXiv*: arXiv:1809.09409, Sep. 2018, doi: 10.48550/arXiv.1809.09409.

[50] S. X. Tan, J. Y. Ong, K. O. M. Goh, and C. Tee, "Boosting vehicle classification with augmentation techniques across multiple YOLO versions", *International Journal on Informatics Visualization (JOIV)*, vol. 8, No. 1, 2024, doi: 10.62527/joiv.8.1.2313.

[51] C. H. Lim, C. Tee., T. S. Ong *et al.*, "Visual-based vehicle detection with adaptive oversampling," *International Journal of Information Technology*, vol. 16, no.8, pp. 4767–4777, 2024, doi: 10.1007/s41870-024-01977-w.

BIOGRAPHIES OF AUTHORS

	Wei-Jie Low is a student at Multimedia University pursuing a Bachelor's degree in Computer Science, majoring in Artificial Intelligence. He is passionate about the advancement of software and AI technologies. He is dedicated to build user-friendly and efficient software applications that can positively impact the society and community. His academic and personal interests include computer vision, machine learning, deep learning, and software development. He can be contacted at 1211101082@student.mmu.edu.my.
	Kah-Ong Michael Goh is an Associate Professor at Multimedia University, specializing in AI, biometrics, machine learning, bioinformatics, and cybersecurity. His research has appeared in top-tier journals such as IEEE Access, Applied Soft Computing, and Expert Systems with Applications. He has received prestigious awards including the ITEX Gold Medal, RICES Gold Medal, and iNVENTX Award. His recent work focuses on advancing intelligent systems in healthcare, security, and computer vision through high-impact, interdisciplinary research and innovation. He can be contacted at michael.goh@mmu.edu.my.
	Check-Yee Law is a lecturer at Faculty of Information Science and Technology (FIST), Multimedia University (MMU), Melaka, Malaysia. Her research interests span the fields of teaching and learning to development of systems, software, and mobile applications. Topics of interest include but are not limited to educational technology, information systems, user-centered design, visual analytics, Internet of Things (IoT), and mobile computing. She can be contacted at cylaw@mmu.edu.my.

	Connie Tee received both the MSc (IT) and PhD (IT) degrees from Multimedia University in 2005 and 2015, respectively. She is a Professor in the Faculty of Information Science and Technology Multimedia University since 2021. She is currently holding the position of the Dean of the Institute for Postgraduate Studies. Her research interests include computer vision, machine learning, deep learning and image processing. She is a senior member of the IEEE. She can be contacted at tee.connie@mmu.edu.my.
	Yong-Wee Sek is a researcher at the Faculty of Artificial Intelligence and Cyber Security (FAIX), Universiti Teknikal Malaysia Melaka. His research interests include technology adoption, Internet of Things (IoT), mobile computing, smart farming, and supply chain innovation. He has published in high-impact journals such as IEEE Access, IET Computer Vision, and AI Open. He actively contributes as a reviewer for journals including IEEE Access, F1000, and Brain Informatics. He can be contacted at ywsek@utem.edu.my.
	Md Ismail Hossen is currently pursuing the Ph.D. degree in computer science with Griffith University, Australia. His expertise lies in image processing and big data analysis, with a particular focus on agricultural applications. In addition, he serves as a Lecturer with American International University, Bangladesh. Outside of academia, he cherishes spending quality time with family and friends and enjoys traveling to enrich his understanding of diverse cultures and perspectives. He can be contacted at m.hossen@griffith.edu.au.