The following publication Zhang, C., Chen, B. Y., Lam, W. H., Ho, H. W., Shi, X., Yang, X., ... & Chow, A. H. (2021). Vehicle re-identification for lane-level travel time estimations on congested urban road networks using video images. IEEE Transactions on Intelligent Transportation Systems, 23(8), 12877-12893 is Transportation Systems is available at https://doi.org/10.1109/TITS.2021.3118206.

Vehicle Re-identification for Lane-level Travel Time Estimations on Congested Urban Road Networks Using Video Images

Cheng Zhang, Bi Yu Chen, *Member, IEEE*, William H. K. Lam, H. W. Ho, Xiaomeng Shi, Xiaoguang Yang, Wei Ma, *Member, IEEE*, S. C. Wong, and Andy H. F. Chow *Member, IEEE*

***Abstract*—The provision of lane-level travel time information can enable accurate traffic control and route guidance in urban roads with distinctive traffic conditions among lanes. However, few studies in the literature have been conducted to estimate lane-level travel time distributions. This study proposes a new vehicle re-identification (V-ReID) method for estimating lane-level travel time distributions using video images from widely deployed surveillance cameras. In the proposed method, a lane-based bipartite graph matching is introduced to obtain optimal matches between upstream and downstream vehicles by considering lane-level traffic conditions and vehicles' lane changing behaviors and visual features. A lane-based travel time estimation technique is introduced to real-time estimate full spectrum of lane-level distribution parameters, including not only the mean but also the standard deviation and the distribution type. A comprehensive case study is carried out on a congested urban road in Hong Kong. Results of case study show that the proposed method outperforms the state-of-the-art link-based V-ReID method and is capable for providing accurate lane-level travel time distribution information on congested urban roads.**

***Index Terms*—Vehicle re-identification, lane-level travel time, lane changing behaviors, video images.**

I. Introduction

Accurate and updated travel time information is a crucial requirement for many traffic monitoring, traffic control and route guidance applications. The provided travel time information not only can enable travelers to make informed route choice decisions to avoid problematic roads, but also can allow transport operators to identify bottlenecks for proactively deploying effective controls [1-3].

The estimation of travel time information in congested urban road networks are not trivial. Firstly, travel times in congested urban road networks are highly stochastic, due to demand fluctuation, traffic controls, incidents and adverse weather, etc. This stochasticity inherent to congested urban networks has led to an increasing focus on estimating travel time distributions rather than a single measure of travel times (e.g., mean travel time) [1-3]. Secondly, the provision of lane-based travel time information in congested urban road networks is of great importance for traffic management, since different lanes on urban roads may have distinctive traffic conditions and different lane-changing movements [4]. Lane-based travel time information enables urban traffic control and management in a more detailed and fine-grained manner. Therefore, it is necessary to investigate methods for estimating lane-based travel time distributions on congested urban roads.

In the literature, various sensor technologies have been used to estimate travel times on congested urban roads. Commonly used traffic sensors could be roughly classified into three types. The first one is probe vehicles equipped with GPS (Global Positioning Systems) devices or cell phones [5]. Trajectories of probe vehicles travelling in the network are collected to

Manuscript received October 14, 2020; revised July 9, 2021; revised October 3, 2021; accepted October 1, 2021. This work was supported in part by the Research Grants Council of Hong Kong Special Administrative Region, China under Project PolyU R5029-18, in part by the Research Institute for Sustainable Urban Development of The Hong Kong Polytechnic University under Project 5-ZJM5, in part by the National Natural Science Foundation of China under Grant 52072264, in part by the National Natural Science Foundation of Hubei Province under Grant 2020CFA054, in part by the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS) Special Research Funding, and in part by the Open Foundation of Key Laboratory of Advanced Public Transportation Science, Ministry of Transport, China. The Associate Editor for this article was S. S. Nedevschi. (Corresponding author: chen.biyu@whu.edu.cn)

Cheng Zhang is with the Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong. He is now with the Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai, 201804, China (email: zhangchengvalue@163.com).

Bi Yu Chen is with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China, also with the Key Laboratory of Advanced Public Transportation Science, Ministry of Transport, Beijing 100029, China, and also with the Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong (email: chen.biyu@whu.edu.cn)

William H. K. Lam, H. W. Ho and Wei Ma are with the Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong (email: william.lam@polyu.edu.hk; hwai.ho@gmail.com; wei.w.ma@polyu.edu.hk).

Xiaomeng Shi is with the School of Transportation, Southeast University, Nanjing, 210096, China (email: shixiaomeng@seu.edu.cn).

Xiaoguang Yang is with the Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai, 201804, China (email: yangxg@tongji.edu.cn).

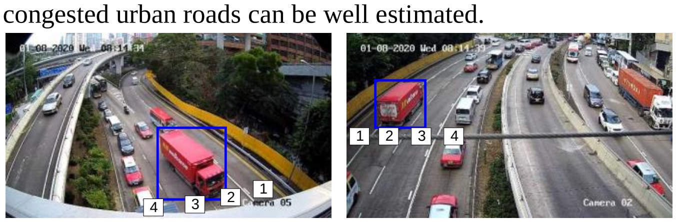
S. C. Wong is with the Department of Civil Engineering, The University of Hong Kong, Hong Kong (email: hhecwsc@hku.hk)

Andy H. F. Chow is with Department of Advanced Design and Systems Engineering, City University of Hong Kong, Hong Kong (email: andychow@cityu.edu.hk)

estimate only link-based travel times, since GPS positional accuracy usually cannot reach the lane level [6]. The second one is point detectors installed at fixed locations, such as loop detectors [7], video cameras [8], thermal cameras [9], weigh-in-motion sensors [10], Radar sensors [11] and LiDAR [12]. These detectors are capable for collecting lane-level spot speeds at a specific location in a road section. Thereby, lane-based travel time can be estimated [13]. However, the travel time estimations do not perform well under congestion, since spot speeds vary significantly at different locations along the road section [14]. The third one is interval detectors, which consist of a pair of sensors deployed at two fixed locations in the network. This type of sensors is capable for directly estimating travel times between the device pair using the time difference of a vehicle passing through two sensor locations [15].

Several types of sensors have been used for interval detectors, such as loop detectors, video cameras and wireless magnetic sensors. In recent years, video camera has become a popular means of interval detectors, due to its widespread availability in urban areas for various applications. To identify a specific vehicle in video images, license plate recognition is a commonly used technique [8]. However, it may not be applicable for video images with low resolution and/or frequent occlusions. More importantly, due to the privacy issue, license plate numbers are not allowed to use in many cities, such as Hong Kong.

As an anonymous way, vehicle re-identification (V-ReID) technique recently attracts much attention [10, 16-21]. Fig. 1 illustrates this V-ReID technique using two video images captured at upstream and downstream locations, respectively. It is to assign a random ID for each vehicle in video images of an upstream location; and to re-identify this vehicle in video images of the downstream location. Using experienced travel times of all re-identified vehicles, travel time information on congested urban roads can be well estimated.



(a) (b)

Fig. 1. Video images from (a) upstream and (b) downstream camera views. The red trucks in the blue bounding boxes are to be matched as the same vehicle.

This study aims to develop an effective method for accurately estimating lane-based travel time distributions by using V-ReID technique. The remainder of this paper is organized as follows. Section II reviews existing V-ReID methods and summarizes the contributions of the current study. Section III briefly describes a state-of-the-art link-based V-ReID method to provide necessary research background. Section IV presents the proposed lane-based V-ReID method. Section V reports the case study on a congested urban road of Hong Kong. Finally, Section VI gives the conclusions and recommendations for further study.

II. Related work

In the literature, many V-ReID methods have been proposed to estimate link-level travel times. As shown in Fig. 1, the identification of the same vehicles in video images captured at the upstream and downstream locations is a critical issue. Most existing V-ReID methods relied on the calculation of similarity among vehicles at two sensor locations based on their visual features, e.g. vehicle color, type and length [22]. Two types of methods have been developed for measuring similarity of visual features: distance-based and probabilistic-based methods. Distance-based methods evaluate vehicle similarity using deterministic distance measures. Sun et al. [18] calculated the distance of each vehicle feature, including inductive signatures, color, and spot speed; then fused these distances by weights to get an overall distance metric between two vehicles; and finally matched the vehicle pair with minimum distance metric. Oliveira-Neto et al. [23] introduced an edit distance to measure the similarity between two strings of license plate numbers. Tang et al. [24] applied deep convolutional neural networks techniques to learn vehicle features with the objective of minimizing their distance measures. Bai et al. [25] adopted group-sensitive triplet embedding to minimize the distances between the same vehicles at the same time keep different vehicles apart away. To further incorporate the noise of vehicle features into the similarity evaluation, an advanced probabilistic method was developed to calculate a matching probability rather than a deterministic similarity metric. Using Bayesian models, Kwong et al. [19] established statistical models to depict the distribution of distance measures between correct matches (the same vehicle) and wrong matches (different vehicles).

To further improve the matching accuracy, several V-ReID methods have introduced a time window constraint to filter out infeasible matches of vehicles. Given a vehicle at the downstream location, a time window can be established to delimit feasible time period in which this vehicle should be observed at the upstream location. The size of time window should be determined carefully, neither too large to include too many candidates for matching, nor too small to filter out correct matches. Sumalee et al. [22] proposed a fixed time window strategy based on the vehicle's historical link-based travel time distribution. Wang et al. [26] further introduced an adaptive time window strategy based on real-time estimated link-based travel time distributions.

As summarized in Table 1, existing V-ReID methods focused on the estimation of link-based travel time distributions, but failed to distinguish significant differences among lanes. However, as illustrated in Fig. 2, travel time distributions vary by lanes. Aggregating travel times of all lanes into a single link-based distribution can lead to inappropriate time windows for different lanes and cause the degradation of V-ReID performance.

It also can be seen from Table 1 that most existing V-ReID methods mainly focused on the estimation of mean link travel times. These V-ReID methods usually assumed that link travel times follow normal distributions, and associated standard deviations or coefficient of variations (the ratio of standard deviation to mean) are fixed and pre-given. However, as illustrated in Fig. 2, travel times in congested urban roads are complicated and highly dynamic, and could not satisfy these over-simplified assumptions. Therefore, full spectrum of travel time distribution parameters should be carefully estimated on a real-time basis, including not only the mean but also the standard deviation and the distribution type.

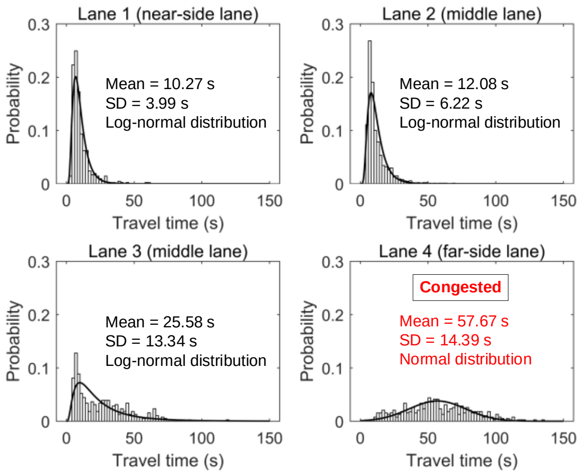


Fig. 2. Travel time distributions by vehicles' lane location at the downstream section in the peak hour. There are significant differences of travel times among lanes. Lane 4 has longer and more spread-out travel time distribution than other lanes.

To fill the gaps, this study develops a new lane-based V-ReID method for estimating lane-level travel time distributions on congested urban roads. It contributes to existing link-based V-ReID studies in following aspects.

(1) A lane-based vehicle re-identification technique is proposed. Lane-based time windows are well constructed to capture significant differences of traffic conditions among lanes. Vehicles' lane changing behaviors are also extracted to estimate their lane location probability at the downstream section. Then, a lane-based bipartite graph matching is developed to determine the optimal matches between vehicles at upstream and downstream sections. Therefore, by explicitly considering the detailed lane-level characteristics, this proposed technique can significantly improve the vehicle matching accuracy.

(2) A lane-based travel time estimation technique is developed. A weighting scheme is introduced to mitigate the effects of wrong matching by considering both distinctness of vehicle features and consequence of wrong matches. Then, a full spectrum of distribution parameters is estimated for each lane including not only the mean but also the standard deviation and the distribution type. An intra-period adjusting process is further developed to iteratively improve travel time estimation and vehicle matching accuracy. Therefore, by real-time estimating the full spectrum of lane-based travel time distributions, this developed technique can significantly enhance the travel time estimation accuracy.

(3) A comprehensive case study using real data collected on a congested urban road in Hong Kong is carried out. Results of case study demonstrate that the developed V-ReID method is capable for accurately estimating lane-based travel time distributions during peak and non-peak hours. The developed V-ReID method outperforms the state-of-the-art link-based V-ReID method [26] with respect to both vehicle matching accuracy and travel time estimation accuracy.

TABLE I
COMPARISONS OF EXISTING VEHICLE RE-IDENTIFICATION METHODS

Vehicle re-identification method	Detector/sensor	Similarity metric	Time window constraint	Vehicle matching	Travel time estimation: Mean	Travel time estimation: Standard deviation	Travel time estimation: Distribution type
Sun et al. (2004) [18]	Video camera and loop detector	Distance-based method	Link-based	Link-based	None	None	None
Oliveira-Neto et al. (2012) [23]	License-plate recognition machine				Link-based	Link-based	
Tang et al. (2018) [24]	Video camera				Link-based	None	
Hyun et al. (2017) [10]	WIM sensor and loop detector	Probabilistic method			None	None	
Kwong et al. (2009) [19]	Wireless magnetic sensor				Link-based	Link-based	
Sumalee et al. (2012) [22]	Video camera				Link-based	Link-based	
Wang et al. (2014) [26]	Video camera				Link-based	Link-based	
The proposed method	Video camera		Lane-based	Lane-based	Lane-based	Lane-based	Lane-based

III. STATE-OF-THE-ART LINK-BASED METHOD

As illustrated in Fig. 1, there are two video cameras established at upstream and downstream sections along a network link. The link-based V-ReID problem is to identify vehicles passing the upstream section during time period t, then re-identify these vehicles at the downstream section, and finally estimate link travel time distribution T_t during time period t based on the experienced travel times of re-identified vehicles.

Fig. 3 illustrates the overall framework of a state-of-the-art link-based V-ReID method [26], which consists of following four major steps.

1) Step 1: Link-based Vehicle Feature Extraction

Using video image processing techniques, this step is to identify all vehicles passing through upstream section and downstream section, and extract their vehicle visual features and mobility statuses. The extracted visual features include vehicle color, vehicle type and vehicle length; while the extracted mobility statuses include arrival time τ and spot speed v. Consequently, a set of vehicles at the upstream section, denoted by U, can be identified, and each upstream vehicle, $i \in U$, is represented by $(color_i^U, type_i^U, length_i^U, \tau_i^U, v_i^U)$. Similarly, the set of vehicles at the downstream section, denoted by D, can be identified, and each downstream vehicle $j \in D$ is represented by $(color_j^D, type_j^D, length_j^D, \tau_j^D, v_j^D)$.

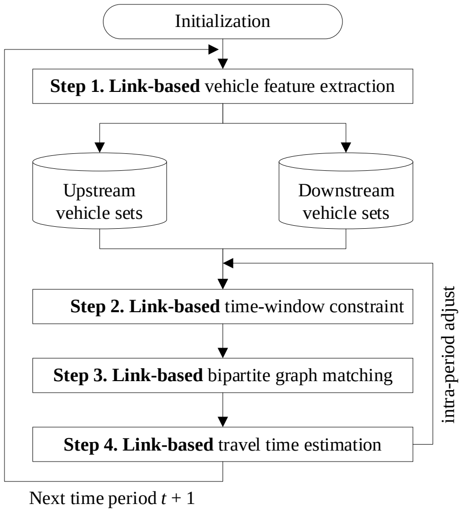


Fig. 3. Flowchart of the state-of-the-art link-based V-ReID method, which implied the assumption that different lanes have the same travel time estimation because differences of travel times in different lanes were not considered.

2) Step 2. Link-based Travel Time Window Construction

This step is to construct a link-based time window to delimit feasible matches between upstream and downstream vehicles.

Without loss of generality, supposes that the link travel time distribution during the last time period $t-1$, denoted by $T_{t-1}(\mu_{t-1}, \sigma_{t-1})$, was estimated and available, where μ_{t-1} is its mean and σ_{t-1} is its standard deviation. It is assumed that $\tilde{T}_t$ follows normal distribution and the coefficient of variation, denoted by ϕ, is given and can be calibrated based on the historical data [26].

Using T_{t-1}, a link-based inter-period adjusting process [26] is introduced to roughly predict the link time distribution $\tilde{T}_t(\tilde{\mu}_t, \tilde{\sigma}_t)$ during current time period t, where $\tilde{\mu}_t$ and $\tilde{\sigma}_t$ are its mean and standard deviation respectively. In (1), the mean travel time $\tilde{\mu}_t$ is predicted by the smoothing technique as:

$$\tilde{\mu}_t = \mu_{t-1} + \beta_\mu(\mu_{t-1} - \tilde{\mu}_{t-1}) \tag{1}$$

where $\tilde{\mu}_{t-1}$ and μ_{t-1} respectively are predicted and estimated mean travel time at the last time period, and β_μ is the smoothing parameter with respect to the prediction error, i.e., $(\mu_{t-1} - \tilde{\mu}_{t-1})$. In (2), this $\tilde{\mu}_t$ is further adjusted to take account the change of spot speeds during these two periods as:

$$\tilde{\mu}_t = \frac{v_{t-1}^U + v_{t-1}^D}{v_t^U + v_t^D} \tilde{\mu}_t \tag{2}$$

where v_{t-1}^U and v_{t-1}^D respectively are average spot speeds for all vehicles at the upstream and downstream sections during the last time period $(t-1)$, and v_t^U and v_t^D are the corresponding values during the current time period t. Then, in (3) the standard deviation $\tilde{\sigma}_t$ is calculated as:

$$\tilde{\sigma}_t = \phi \tilde{\mu}_t \tag{3}$$

With the predicted $\tilde{T}_t$, its α confidence interval of $\tilde{T}_t$, denoted by $[Lb_t, Ub_t]$, can be formulated by $\left[\Phi_{\tilde{T}_t}^{-1}(\frac{1-\alpha}{2}), \Phi_{\tilde{T}_t}^{-1}(\frac{1+\alpha}{2})\right]$, where $\Phi_{\tilde{T}_t}^{-1}()$ is the inverse of cumulative distribution function (CDF) of $\tilde{T}_t$ at a confidence level. Under the normal distribution case, they can be calculated as:

$$[Lb_t, Ub_t] = \left[\tilde{\mu}_t - \mathrm{z}(\frac{1-\alpha}{2})\tilde{\sigma}_t, \tilde{\mu}_t + \mathrm{z}(\frac{1+\alpha}{2})\tilde{\sigma}_t\right] \tag{4}$$

where z() is the inverse of CDF of the standard normal distribution at a confidence level and it can be obtained from the standard normal table or calculated by numerical approximation.

3) Step 3. Link-based Bipartite Graph Matching

This step is to determine the optimum matches between upstream and downstream vehicles using a bipartite graph matching technique.

The bipartite graph contains two parts of vertices and a set of edges between them. To be specific, the vertices are divided into two disjoint sets, i.e. upstream vehicle set U and downstream vehicle set D. For each upstream vehicle $i \in U$, there has several edges connecting to its feasible matching downstream vehicles. Given arrival time of vehicle i at upstream section τ_i^U, the feasible arrival time window of vehicle i at downstream section can be calculated by $[\tau_i^U + Lb_t, \tau_i^U + Ub_t]$. Then, the set of feasible downstream matches, $S(i)$, can be determined as described in (5).

$$S(i) = \left\{\forall j \in D \middle| \tau_j^D \in [\tau_i^U + Lb_t, \tau_i^U + Ub_t]\right\} \tag{5}$$

where τ_j^D is arrival time of vehicle j at the downstream section.

For an edge (i, j) connecting $i \in U$ and $j \in S(i)$, its weight represents the matching probability P_{ij}, which is calculated by similarity between visual features (color, type and length) of vehicles i and j using following Bayesian formulas:

$$P_{ij} = P\left(\delta_{ij} = 1 \middle| d_{color}, d_{type}, d_{length}\right) \tag{6}$$

where $\delta_{ij} = 1$ indicates an edge exist between vehicles i and j, and d_{color}, d_{type}, and d_{length} are difference (or distance) between vehicles' color, type, and length features, respectively. A training dataset that contains a number of pairs of correctly matched vehicles is utilized to estimate P_{ij} for all edges. Consequently, a weighted bipartite graph can be constructed.

The optimal matches between U and D can be obtained by solving the bipartite graph matching problem [27].

4) Step 4. Link-based Travel Time Estimation

This step is to estimate the link travel time T_t based on the experienced travel times of re-identified vehicles, i.e., matched vehicle pairs.

A thresholding process is used to filter out unreliable matches in those scenarios, in which several candidate downstream vehicles have similar matching probability. The thresholding process is to calculate a distinctiveness value $\mathcal{D}_{ij}$ for each vehicle match (i, j) obtained in Step 3. Let $i^{(2)}$ be the vehicle with the second largest matching probability in $S(i)$. The distinctiveness index $\mathcal{D}_{ij}$ is expressed in (7).

$$\mathcal{D}_{ij} = \frac{P_{ij}}{P_{i^{(2)}}} \tag{7}$$

Then, unreliable vehicle matches with $\mathcal{D}_{ij}$ less than a threshold parameter ε_D (e.g., $\varepsilon_D \geq 1$) are removed without consideration. Other reliable matches form a valid set of vehicle matches, denoted by M, for travel time estimation.

It is also assumed T_t following the normal distribution, $N(\mu_t, \sigma_t)$. Accordingly, mean travel time μ_t is calculated by (8).

$$\mu_t = \frac{1}{|M|} \sum_{\forall (i,j) \in M} T_{ij} \tag{8}$$

where T_{ij} is the travel time experienced by a valid match $(i, j) \in M$, and $|M|$ is the total number of matches in M. Using the assumption that the coefficient of variation ϕ of the link travel time is given, the standard deviation σ_t is determined by (9).

$$\sigma_t = \phi \mu_t \tag{9}$$

Since time windows constructed in Step 2 have significant impacts on the V-ReID performance, a link-based intra-period adjusting process is introduced to utilize estimated T_t instead of predicted $\tilde{T}_t$ for constructing more accurate time windows. With newly constructed time windows, Steps 3 and 4 are also re-performed to determine a better estimation of T_t. This iterative process will continue until the relative change in the estimated μ_t is smaller than a threshold parameter, ε_μ.

5) *Discussion on the state-of-the-art link-based method*

The existing link-based V-ReID method [26] can provide reasonable estimations of link mean travel times in highways. However, this existing method may not be applicable to congested urban roads due to following two reasons. Firstly, this V-ReID method only estimated link travel time distributions but failed to distinguish significant differences among lanes of the same road section. Secondly, this V-ReID method assumed travel times following normal distributions with fixed and pre-given the coefficient of variation. Nevertheless, travel times in congested urban roads are complicated and highly dynamic, and could not satisfy these over-simplified assumptions. To address these two issues, a new lane-based V-ReID method is proposed and presented in the next section.

IV. Lane-based Vehicle Re-identification

The flowchart of the proposed lane-based V-ReID method is given in Fig. 4. It consists of four major steps, which are described in detail as following sub-sections.

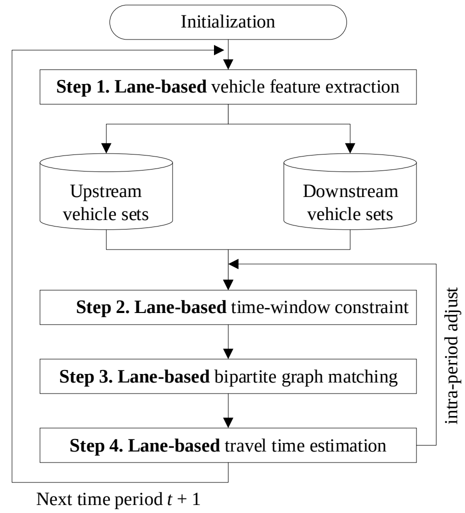


Fig. 4. Flow chart of the proposed lane-based vehicle re-identification method. Appreciable differences of travel times in different lanes are explicitly considered.

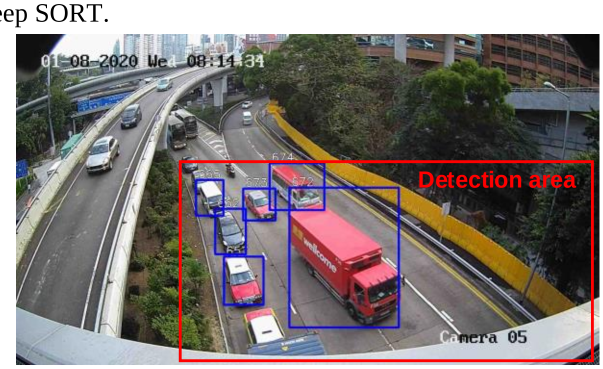
A. *Step 1. Lane-based Vehicle Feature Extraction*

This step is to identify vehicles passing through upstream and downstream sections and exact vehicle information using the video image processing technique. Following the previous link-based V-ReID methods [22], we extract several robust visual features for every vehicle. To fulfill the lane-based V-ReID, we also extract the vehicle's lane-based mobility status, such as lane location and lane changing pattern.

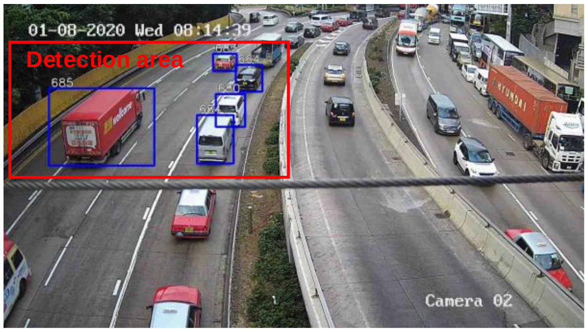
This step can be further divided into three sub-steps, which respectively exact vehicle trajectories, lane-based mobility status and visual features. Their detailed procedures are described as below.

1) *Sub-step 1.1. Trajectory Extraction*

This sub-step is to extract the trajectory for every vehicle in both upstream and downstream videos. We use a well-developed tool, YOLOv3 (You Only Look Once version 3) [28], for vehicle detection and a powerful tool, Deep SORT (Simple Online and Realtime Tracking) [29], for multiple vehicle tracking. As illustrated in Fig. 5, vehicle images in blue boxes are clipped from the original video frames based on the results of vehicle detection using YOLOv3. A random ID is assigned to each vehicle for tracking its trajectory in the video using Deep SORT.



(a)

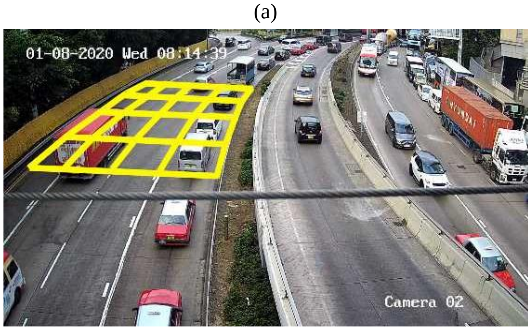


(b)

Fig. 5. Detected vehicles in blue bounding boxes: (a) upstream; (b) downstream. The areas of interest in this study are bounded by red boxes as detection areas.

Given a vehicle at the upstream section $i \in U$, there are a sequent of n vehicle frames recorded at different timestamps. Let $I_{i,k}^{U}$ be vehicle frame recorded at k^{th} timestamp, denoted by $\tau_{i,k}^{U}$. The vehicle's image coordinates at $\tau_{i,k}^{U}$ can be determined as the center of the corresponding vehicle images. By projecting these image coordinates into virtual grids [24], we can obtain the vehicle's world coordinates $(x_{i,k}^{U}, y_{i,k}^{U})$, where $x_{i,k}^{U}$ and $y_{i,k}^{U}$ represent the vehicles' lateral and longitudinal locations respectively (illustrated in Fig. 6). Thereby, a record of $(I_{i,k}^{U}, \tau_{i,k}^{U}, x_{i,k}^{U}, y_{i,k}^{U})$ can be determined for each timestamp $\tau_{i,k}^{U}$. The vehicle's trajectory in the upstream video can be obtained by corresponding sequent of n records. Similarly, the trajectory of a downstream vehicle $j \in D$ can be obtained by a series of records, $(I_{j,q}^{D}, \tau_{j,q}^{D}, x_{j,q}^{D}, y_{j,q}^{D})$, where q denotes the q^{th} timestamp.

(a)



(b)

Fig. 6. Camera calibration by using projected virtual grids: (a) upstream; (b) downstream. Image coordinates can be transformed to world coordinates by homography matrixes.

2) Sub-step 1.2. Lane-based Mobility Status Extraction

This sub-step is to extract the lane-based mobility statuses for all vehicles. The extracted information includes the vehicle's lane location, arrival time, spot speed and lane changing pattern.

The lane location of upstream vehicle i at the pre-given section $y_{i,k}^{U}$ is denoted by l_{i}^{U}. It can be represented in the form of lane number, which is in the order from the near-side lane to the far-side lane. This l_{i}^{U} parameter can be determined by its lateral location $x_{i,k}^{U}$ and the lane width (e.g., 3.0 m in Hong Kong). For example, a vehicle's lateral location $x_{i,k}^{U}$ equals 4.4 m, this vehicle is within the range of Lane 2 from 3.0 to 6.0 m, and thus l_{i}^{U} is Lane 2. After all vehicles' lane locations at the upstream section have been identified, the set of vehicles on each upstream lane l, denoted by U_l, can be obtained.

The arrival time of upstream vehicle i at the pre-given section $y_{i,k}^{U}$ is denoted by τ_{i}^{U}. It can be simply determined as the timestamp $\tau_{i,k}^{U}$ when the vehicle passed the section $y_{i,k}^{U}$.

The spot speed of upstream vehicle i is denoted by v_{i}^{U}. It can be calculated by the distance of its longitudinal movement in the camera view and the time it takes. The equation is given by (10).

$$v_{i}^{U} = \frac{y_{i,n}^{U} - y_{i,1}^{U}}{\tau_{i,n}^{U} - \tau_{i,1}^{U}} \tag{10}$$

where $y_{i,n}^{U}$ and $y_{i,1}^{U}$ are the longitudinal locations of the last and the first record, $\tau_{i,n}^{U}$ and $\tau_{i,1}^{U}$ are the timestamps of the last and the first record respectively.

The lane changing pattern of upstream vehicle i is denoted by LC_{i}^{U}. It is extracted based on its beginning and final lane locations, $l_{i,1}^{U}$ and $l_{i,n}^{U}$, in the upstream video as:

$$LC_{i}^{U} = \begin{cases} 0, l_{i,1}^{U} = l_{i,n}^{U} \\ 1, l_{i,1}^{U} < l_{i,n}^{U} \\ -1, l_{i,1}^{U} > l_{i,n}^{U} \end{cases} \tag{11}$$

where $LC_{i}^{U} = 0$, 1 and -1 respectively represents vehicle i staying on the same lane, changing to far-side lanes and changing to near-side lanes.

Using the same approach as the upstream case, we can also determine the lane location l_{j}^{D}, arrival time τ_{j}^{D} and spot speed v_{j}^{D} for each downstream vehicle j. Consequently, we can obtain the set of vehicles on each downstream lane l, denoted by D_l.

3) Sub-step 1.3. Visual Feature Extraction

This sub-step follows previous methods [22] to extract the vehicle's visual features, including $color$, $type$ and $length$. We firstly extract the color feature for upstream vehicle i, denoted by $color_{i}^{U}$. It is depicted by a three-dimensional (i.e. red, green and blue) color histogram [24] and all three channels are divided into 8 bins. Accordingly, $color_{i}^{U}$ is a vector of 8×8×8=512 elements, representing the color distribution of the vehicle image. The technique of background subtraction is adopted to exclude the color information of the background [30].

We then extract the type feature of upstream vehicle i, denoted by $type_{i}^{U}$. Following the work of [31], the template matching method is adopted by calculating the similarity of the vehicle's image to template images of reprehensive vehicle types in the study area. In this method, the vehicle image $I_{i,k}^{U}$ is firstly converted into a gray-scale image I_{gray} to remove color information but preserve vehicle shape and size. Then, a similarity score between the vehicle's gray-scale image I_{gray} and a gray-scale template image of the κ-th vehicle type, Tem_κ, is calculated by (12).

$$type(\kappa) = 1 - \frac{\sum_{a=1}^{A}\sum_{b=1}^{B}\left|I_{gray}(a,b) - Tem_\kappa(a,b)\right|^2}{255 \times AB} \quad (12)$$

where A and B are width and height of the template image respectively, and (a,b) are the pixel coordinates in images. The similarity score $type(\kappa)$ ranges from 0 to 1. Larger $type(\kappa)$ means more similarity to the κ^{th} vehicle type.

It is noticed that the appearance of the same vehicle varies from lane to lane, especially when vehicles are captured from a side angle. Thus, the template images, Tem_κ in (12), are prepared for each vehicle type, each camera view and each lane.

In this study, six vehicle types are adopted including sedan, taxi, van, minibus, truck and bus. Accordingly, the type feature $type_i^U$ is a vector of six elements in terms of similarity metrics. Takes an upstream vehicle in Fig. 1(a) as an example, and its type feature (i.e. similarity-score vector) is provided in Table II.

TABLE II
SIMILARITY SCORE FOR EACH TEMPLATE

Template image	 Sedan	 Taxi	 Van	 Mini-bus	 Truck	 Bus
 Input	0.5815	0.6193	0.6670	0.6867	**0.7066**	0.6238

Finally, we extract the length feature of upstream vehicle i, denoted by $length_i^U$. Following the work of [22], it is estimated by the normalized height of the vehicle image as:

$$length_i^U = height_i^U \times \eta_l \quad (13)$$

where $height_i^U$ is the image height and η_l is the normalizing factor that needs calibration for each lane location l and each video camera. In this study, a training dataset is prepared for η_l calibration, see Table III as an example of the training dataset. The ground truth of vehicle length $\widehat{length}_i^U$ is determined by vehicle type while the image height $height_i^U$ is measured from the vehicle images. Using an ordinary least square technique, η_l parameter can be calibrated by (14-15).

$$\min \sum_i \left(\widehat{length}_i^U - length_i^U\right)^2 \quad (14)$$

$$\eta_l = \sum_i length_i^U \Big/ \sum_i height_i^U \quad (15)$$

Using the same approach, color feature $color_j^D$, type feature $type_j^D$, and length feature $length_j^D$ can be extracted for every downstream vehicle j.

TABLE III
GROUND TRUTH OF VEHICLE LENGTH AND IMAGE HEIGHT

Vehicle type	 Sedan	 Taxi	 Van	 Mini-bus	 Truck	 Bus
Vehicle length(m)	4.8	4.8	6.0	8.0	10.0	12.0
Image height (pixel)	189	186	229	281	327	499

B. Step 2. Lane-based Travel Time Window Construction

This step is to construct lane-based travel time windows for each downstream lane to delimit feasible matches between upstream and downstream vehicles. It extends the previous link-based travel time window construction step [26] by explicitly considering distinctive traffic conditions at different lanes.

Without loss of generality, supposes that travel time distribution of lane l during the last time period $t-1$, denoted by $T_{l,t-1}(\mu_{l,t-1}, \sigma_{l,t-1}, \pi_{l,t-1})$, is estimated and available, where $\mu_{l,t-1}$ is its mean, $\sigma_{l,t-1}$ is its standard deviation, and $\pi_{l,t-1}$ is its distribution type. In this study, it was assumed that $T_{l,t-1}$ follows either normal or lognormal distribution according to the previous empirical studies [3, 32]. It should be noted that this assumption can be relaxed by using other distribution types, e.g., Gamma or Weibull.

To determine lane-based time windows, a lane-based inter-period adjusting process is introduced to determine a rough prediction of travel time distribution of lane l during the current time period t, denoted by $\tilde{T}_{l,t}(\tilde{\mu}_{l,t}, \tilde{\sigma}_{l,t}, \tilde{\pi}_{l,t})$. Following the previous work [26], the mean travel time $\tilde{\mu}_{l,t}$ is adjusted using a smoothing technique as:

$$\tilde{\mu}_{l,t} = \mu_{l,t-1} + \beta_\mu(\mu_{l,t-1} - \tilde{\mu}_{l,t-1}) \quad (16)$$

where $\tilde{\mu}_{l,t-1}$ and $\mu_{l,t-1}$ respectively are predicted and estimated means at the last time period $t-1$, and β_μ is the smoothing parameter with respect to prediction error of mean travel time. This predicted mean is further adjusted according to the changes of spot speed as:

$$\tilde{\mu}_{l,t} = \frac{v_{l,t-1}^U + v_{l,t-1}^D}{v_{l,t}^U + v_{l,t}^D}\tilde{\mu}_{l,t} \quad (17)$$

where $v_{l,t-1}^U$ and $v_{l,t-1}^D$ respectively are average spot speeds for all vehicles on lane l of the upstream and downstream sections during the last time period $(t-1)$, and $v_{l,t}^U$ and $v_{l,t}^D$ are the corresponding values during the current time period t. In this study, the standard deviation is also real-time estimated by the smoothing technique as:

$$\tilde{\sigma}_{l,t} = \sigma_{l,t-1} + \beta_\sigma(\sigma_{l,t-1} - \tilde{\sigma}_{l,t-1}) \quad (18)$$

where $\tilde{\sigma}_{l,t-1}$ and $\sigma_{l,t-1}$ respectively are predicted and estimated standard deviations at the last time period $t-1$, and β_σ is the corresponding smoothing parameter. However, it is assumed $\tilde{T}_{l,t}$ having the same distribution type as $T_{l,t-1}$, i.e., $\tilde{\pi}_{l,t} = \pi_{l,t-1}$. It should be noted that all three distribution parameters will be further updated in Step 4.

With the predicted $\tilde{T}_{l,t}$, its α confidence interval $\left[\widetilde{Lb}_{l,t}, \widetilde{Ub}_{l,t}\right]$ can be formulated by $\left[\Phi_{\tilde{T}_{l,t}}^{-1}(\frac{1-\alpha}{2}), \Phi_{\tilde{T}_{l,t}}^{-1}(\frac{1+\alpha}{2})\right]$, where $\Phi_{\tilde{T}_{l,t}}^{-1}()$ is the inverse of CDF of $\tilde{T}_{l,t}$. When $\tilde{T}_{l,t}$ follows normal distribution, it can be calculated by (19).

$$\left[\widetilde{Lb}_{l,t}, \widetilde{Ub}_{l,t}\right] = \left[\tilde{\mu}_{l,t} - \mathrm{z}(\frac{1-\alpha}{2})\tilde{\sigma}_{l,t}, \tilde{\mu}_{l,t} + \mathrm{z}(\frac{1+\alpha}{2})\tilde{\sigma}_{l,t}\right] \quad (19)$$

Otherwise, $\tilde{T}_{l,t}$ follows a lognormal distribution and its α confidence interval can be calculated by (20).

$$\left[\widetilde{Lb}_{l,t}, \widetilde{Ub}_{l,t}\right] = \left[e^{\tilde{\mu}_{l,t}^{\log} - \mathrm{z}\left(\frac{1-\alpha}{2}\right)\tilde{\sigma}_{l,t}^{\log}}, e^{\tilde{\mu}_{l,t}^{\log} + \mathrm{z}\left(\frac{1-\alpha}{2}\right)\tilde{\sigma}_{l,t}^{\log}}\right] \quad (20)$$

where z() is the z-score, i.e. the inverse of CDF of the standard normal distribution at a confidence level. $\tilde{\mu}_{l,t}^{\log}$ and $\tilde{\sigma}_{l,t}^{\log}$ are respectively mean and standard deviation of the logarithm of $\tilde{T}_{l,t}$ and they can be expressed as:

$$\tilde{\mu}_{l,t}^{\log} = \log(\tilde{\mu}_{l,t}) - \frac{1}{2}\log\left(1 + \left(\frac{\tilde{\sigma}_{l,t}}{\tilde{\mu}_{l,t}}\right)^2\right) \quad (21)$$

$$\tilde{\sigma}_{l,t}^{\log} = \sqrt{\log\left(1 + \left(\frac{\tilde{\sigma}_{l,t}}{\tilde{\mu}_{l,t}}\right)^2\right)} \quad (22)$$

C. Step 3. Lane-based Bipartite Graph Matching

This step is to determine optimal matches between upstream and downstream vehicles using a lane-based bipartite graph technique. It extends the previous link-based matching step by explicitly considering detailed lane-level traffic conditions and vehicles' lane changing behaviors. This step can be further divided into three sub-steps, including lane-based bipartite graph construction, matching probability calculation, and optimal matching. Their detailed procedures are described as below.

1) Sub-step 3.1. Lane-based Bipartite Graph Construction

This sub-step is to construct a bipartite graph to determine optimal matches between upstream and downstream vehicles. Fig.7 illustrates the concept of lane-based bipartite graph. This graph contains two disjoint parts of vertices representing upstream vehicle set U and downstream vehicle set D, which are further divided into corresponding lane-based sub-sets, $(\dots, U_l, \dots)$ and $(\dots, D_l, \dots)$.

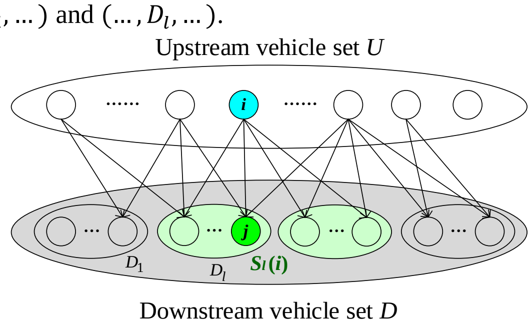


Fig. 7. Lane-based bipartite graph representation. For an upstream vehicle i, its feasible matches of downstream vehicles j are reduced from the link-based search space (grey color) to the lane-based search space (green color).

For each upstream vehicle $i \in U$, there has several edges connecting to its candidate matches in each downstream lane set D_l. Given arrival time of vehicle i at upstream section τ_i^U, its feasible arrival time window through the downstream lane l is determined as $\left[\tau_i^U + Lb_{l,t}, \tau_i^U + Ub_{l,t}\right]$. Accordingly, the set of feasible matches on downstream lane l, denoted by $S_l(i)$, can be determined by (23).

$$S_l(i) = \left\{\forall j \in D_l \middle| \tau_j^D \in \left[\tau_i^U + Lb_{l,t}, \tau_i^U + Ub_{l,t}\right]\right\} \quad (23)$$

where τ_j^D is arrival time of vehicle j at the downstream section. The whole set of feasible matches at the downstream section, $S(i)$, can be obtained as the union of all $S_l(i)$ as:

$$S(i) = \bigcup_{\forall l} S_l(i) \quad (24)$$

For an edge (i, j) connecting $i \in U$ and $j \in S_l(i)$, its weight represents the matching probability P_{ij}.

2) Sub-step 3.2. Matching Probability Calculation

This sub-step is to calculate matching probability P_{ij} for every edge, i.e., feasible vehicle match (i, j). Following the previous work [26], the matching probability P_{ij} is formulated in this study using the statistical matching method. Nevertheless, this study extends the previous method [26] by considering not only vehicles' visual features but also lane changing behaviors and arrival time probability.

According to Bayesian approach, the matching probability P_{ij} is formulated as:

$$P_{ij} = P\left(\delta_{ij} = 1 \middle| d_{ij}\right) \quad (25)$$

where δ_{ij} denote whether vehicles i and j are matched, i.e. 1 means matched and 0 otherwise; and d_{ij} is the dis-similarity measures of these two vehicles' visual features. According to Bayes' theorem, we have following posterior probability:

$$P\left(\delta_{ij} = 1 \middle| d_{ij}\right) = \frac{P\left(d_{ij} \middle| \delta_{ij} = 1\right) P\left(\delta_{ij} = 1\right)}{P(d_{ij})} \quad (26)$$

where $P\left(d_{ij} \middle| \delta_{ij} = 1\right)$ is the likelihood function, and $P(\delta_{ij} = 1)$ is the prior probability without considering the vehicle feature information. The vehicle feature dis-similarity distribution $P(d_{ij})$ can be further expressed as:

$$P(d_{ij}) = P\left(d_{ij} \middle| \delta_{ij} = 1\right) P\left(\delta_{ij} = 1\right) + P\left(d_{ij} \middle| \delta_{ij} = 0\right) P\left(\delta_{ij} = 0\right) \quad (27)$$

$$P\left(\delta_{ij} = 0\right) = 1 - P\left(\delta_{ij} = 1\right) \quad (28)$$

Therefore, the calculation of P_{ij} depends on likelihood functions $P\left(d_{ij} \middle| \delta_{ij} = 1\right)$ and $P\left(d_{ij} \middle| \delta_{ij} = 0\right)$, and prior probability function $P(\delta_{ij} = 1)$.

In this study, the prior probability $P(\delta_{ij} = 1)$ is formulated by explicitly considering vehicle's lane changing behaviors and arrival time probability. Based on the historical data, we can observe some patterns of lane changing behaviors as follows: (1) vehicles changing lanes at upstream are less likely to change lanes at downstream; (2) if vehicles change to near-side lanes at upstream, it is almost impossible for them to change to far-side lanes at downstream, and vice versa; (3) lane changing patterns vary among vehicles of different sizes. It is easier for small vehicles to change lanes when it comes to congestion. Vehicles are divided into two categories (i.e. small and large vehicles) based on vehicle lengths. The threshold of 7.2 m is determined for classification by a clustering method. Therefore, the probability of upstream vehicle i on downstream lane , denoted by $P(l_j^D)$, depends on its lane location l_i^U, lane changing behavior at the upstream section LC_i^U and its vehicle length $length_i^U$ as:

$$P(l_j^D) = P\left(l_j^D \middle| l_i^U, LC_i^U, length_i^U\right) = \frac{P(l_i^U, LC_i^U, length_i^U, l_j^D)}{\sum_{l_j^D} P(l_i^U, LC_i^U, length_i^U, l_j^D)} \quad (29)$$

where $P(l_i^U, LC_i^U, length_i^U, l_j^D)$ is the joint probability density estimated with training dataset in the same time duration.

The probability of vehicle i arriving at the downstream section on τ_i^D, denoted by $P(\tau_j^D)$, depends on travel time distribution of downstream lane l as:

$$P(\tau_j^D) = f_{\tilde{T}_{l,t}}(\tau_j^D - \tau_i^U) \cdot \gamma_{time} \quad (30)$$

where $f_{\tilde{T}_{l,t}}(\cdot)$ is the lane-based travel time distribution estimated in the last time period, and γ_{time} is a normalizing factor which could be the class interval of the histogram related to the travel time distribution.

By fusing $P(l_j^D)$ and $P(\tau_j^D)$ based on the logarithmic opinion pool approach, the prior probability $P(\delta_{ij} = 1)$ determined by (31).

$$P(\delta_{ij} = 1) = \frac{1}{\gamma_{LT}} P\left(l_j^D \middle| L_i^U, l_i^U, LC_i^U\right)^{\theta_{lane}} P(\tau_j^D)^{\theta_{time}} \quad (31)$$

where γ_{LT} is the normalizing constant, and θ_{lane} and θ_{time} are corresponding weighting parameters.

Following the previous work [26], the likelihood functions, $P(d_{ij}|\delta_{ij} = 1)$ and $P(d_{ij}|\delta_{ij} = 0)$, are calculated by visual features including color, length and type. Let $d_{color}(i,j)$, $d_{type}(i,j)$ and $d_{length}(i,j)$ be the distance measure between vehicles i and j for color, length and type features respectively. The color distance measure, $d_{color}(i,j)$, is calculated by the Bhattacharyya distance as:

$$d_{color}(i,j) = \left[1 - \sum_{r=1}^{512} \sqrt{color_i^U(r) \cdot color_j^D(r)}\right]^{1/2} \quad (32)$$

where r means the r-th element of the color feature vector. The type distance measure, $d_{type}(i,j)$, is calculated by the Manhattan distance as:

$$d_{type}(i,j) = \sum_{\kappa} \left|type_i^U(\kappa) - type_j^D(\kappa)\right| \quad (33)$$

where κ represents the κ-th element of the type feature vector. The length distance measure, $d_{length}(i,j)$, is calculated by the absolute difference as:

$$d_{length}(i,j) = \left|length_i^U - length_j^D\right| \quad (34)$$

Let $P(d_{color}(i,j)|\delta_{ij} = 1)$ and $P(d_{color}(i,j)|\delta_{ij} = 0)$ be likelihood functions of color feature; $P(d_{type}(i,j)|\delta_{ij} = 1)$ and $P(d_{type}(i,j)|\delta_{ij} = 0)$ be likelihood function of type feature; $P(d_{length}(i,j)|\delta_{ij} = 1)$ and $P(d_{length}(i,j)|\delta_{ij} = 0)$ be likelihood function of length feature. All these likelihood functions of each visual feature can be well estimated by the training dataset. By fusing them based on the logarithmic opinion pool approach, the overall likelihood functions, $P(d_{ij}|\delta_{ij} = 1)$ and $P(d_{ij}|\delta_{ij} = 0)$, can be determined as:

$$P(d_{ij}|\delta_{ij} = 1) = \frac{1}{\gamma_{CTL}} P(d_{color}|\delta_{ij} = 1)^{\theta_{color}} P(d_{type}|\delta_{ij} = 1)^{\theta_{type}} P(d_{length}|\delta_{ij} = 1)^{\theta_{length}} \quad (35)$$

$$P(d_{ij}|\delta_{ij} = 0) = \frac{1}{\gamma_{CTL}} P(d_{color}|\delta_{ij} = 0)^{\theta_{color}} P(d_{type}|\delta_{ij} = 0)^{\theta_{type}} P(d_{length}|\delta_{ij} = 0)^{\theta_{length}} \quad (36)$$

where γ_{CTL} is a normalizing constant, and θ_{color}, θ_{type} and θ_{length} are corresponding weighing parameters that can be estimated by training dataset to maximize the matching accuracy.

3) *Sub-step 3.3. Optimum Matching*

This sub-step is to determine optimal matches using the lane-based bipartite graph constructed in Sub-steps 3.1 and 3.2. The optimal matches can be obtained by solving the following maximization problem in the constructed bipartite graph:

$$\max \sum_{\forall i \in U} \sum_{\forall j \in D} P_{ij}\delta_{ij} \quad (37)$$

$$\text{s.t.}\ \delta_{ij} \in \{0,1\}, \forall i \in U, \forall j \in S(i) \quad (38)$$

$$\sum_{\forall j \in D} \delta_{ij} \le 1, \forall i \in U \quad (39)$$

$$\sum_{\forall i \in U} \delta_{ij} \le 1, \forall j \in D \quad (40)$$

(37) is the objective function that maximizes the total matching probability. (38) indicates the decision variables as binary integers. (39) ensures that any upstream vehicle i can only be matched with at most one downstream vehicle, while (40) ensures that any downstream vehicle j can only be matched with at most one upstream vehicle. This bipartite graph matching problem can be solved by a well-developed algorithm proposed by Galil [27].

D. *Step 4. Lane-based Travel Time Estimation*

This step is to estimate the lane-level travel time distributions on real-time basis. It extends the previous study [26] to estimate the full spectrum of lane-level travel time distributions, including not only the mean but also the standard deviation and the distribution type. This step can be further divided into two sub-steps as below.

1) *Sub-step 4.1. Travel Time Distribution Estimation*

This sub-step is estimate travel time distribution for each lane based on the results of vehicle matches.

In this study, a weighting scheme is introduced to mitigate the effects of wrong matching. The weight of a vehicle match (i,j), denoted by w_{ij}, is defined in (41).

$$w_{ij} = \left(\frac{P_{ij}}{P_{i^{(2)}}}\right) \Big/ \left(1 + \left|\frac{T_{ij} - T_{i^{(1)}}}{T_{i^{(1)}}}\right|\right) \quad (41)$$

where P_{ij} is the result of graph match for vehicle i and T_{ij} is its associated experienced travel time, $P_{i^{(2)}}$ is the second largest matching probability among $S(i)$, and $T_{i^{(1)}}$ is the experienced travel time of the vehicle match with the largest matching probability among $S(i)$. The numerator of w_{ij} is a distinctive index [26] representing the uniqueness of vehicle match (i,j), while the denominator of w_{ij} represents the consequence of wrong matching. Therefore, the more distinctive of vehicle match the larger weighting, and the more serious consequence of wrong match the lower weighting.

Using the introduced weighting scheme, travel time distribution $T_{l,t}(\mu_{l,t},\sigma_{l,t},\pi_{l,t})$ is estimated for each lane l during the current time period t. The mean travel time, $\mu_{l,t}$, can be calculated by (42).

$$\mu_{l,t} = \frac{\sum_{j\in D_l}\delta_{ij}T_{ij}w_{ij}}{\sum_{j\in D_l}\delta_{ij}w_{ij}} \tag{42}$$

The standard deviation, $\sigma_{l,t}$, is calculated by (43).

$$\sigma_{l,t} = \sqrt{\frac{\sum_{j\in D_l}\delta_{ij}w_{ij}\left(T_{ij}-\mu_{l,t}\right)^2}{\sum_{j\in D_l}\delta_{ij}w_{ij}}} \tag{43}$$

The distribution type $\pi_{l,t}$, either normal or lognormal, is examined by using Kolmogorov-Smirnov (K-S) hypothesis tests [33] for the data collected in the case study. The p-value of K-S tests is used to determine the goodness of fit for both normal and lognormal distributions. Larger p-value means better fitting.

In real-time operations, the sample size during off-peak periods may not sufficient to determine a robust estimation of lane-level distributions, especially $\sigma_{l,t}$ and $\pi_{l,t}$ parameters. To address this issue, we try to utilize the matched samples on the same lane l obtained at the previous time periods, i.e., $t-1$ and $t-2$. In this case, a discount factors of w_{t-1} and w_{t-2} are applied to all weighting of these matched samples collected on $t-1$ and $t-2$ respectively.

2) *Sub-step 4.2. Intra-period Adjusting*

This sub-step introduces a lane-level intra-period adjusting process to iteratively improve the travel time estimation and vehicle matching accuracy. This sub-step extends the previous link-based intra-period adjusting process [26] by updating lane-level distribution parameters.

Since lane-based travel time windows constructed in Step 2 have significant impacts on the V-ReID performance, the introduced intra-period adjusting process is to utilize estimated $T_{l,t}$ instead of predicted $\tilde{T}_{l,t}$ for constructing more accurate time windows according to (19-22). With updated lane-level travel time windows, Steps 3 and 4 are also re-performed to determine a better estimation of $T_{l,t}$. This process continues to iteratively improve the travel time estimation and vehicle matching accuracy. Let $T_{l,t}^{(n)}(\mu_{l,t}^{(n)},\sigma_{l,t}^{(n)},\pi_{l,t}^{(n)})$ and $[Lb_{l,t}^{(n)},Ub_{l,t}^{(n)}]$ respectively denote the lane-based travel time distribution and time window obtained at the n^{th} iteration. Such process can be terminated when the travel time windows are convergent as:

$$\sum_{\forall l}\left|\frac{Lb_{l,t}^{(n)}-Lb_{l,t}^{(n-1)}}{Lb_{l,t}^{(n-1)}}\right|+\left|\frac{Ub_{l,t}^{(n)}-Ub_{l,t}^{(n-1)}}{Ub_{l,t}^{(n-1)}}\right|\le\varepsilon_S \tag{44}$$

where ε_S is the termination threshold.

V. CASE STUDY

To verify the effectiveness of the proposed lane-based V-ReID method, a comprehensive case study was conducted by using real data collected on a congested urban road in Hong Kong.

A. *Testing Site and Parameter Calibration*

As shown in Figs. 8 and 9, the testing site was on a four-lane urban road in Hong Kong (Chatham Road South, Westbound). Two video cameras were installed on a footbridge. Fig. 1 illustrates the video images at upstream and downstream from different angles. The footbridge locates over the testing road and connects Block Z building and Innovation Tower of The Hong Kong Polytechnic University.

The installed video cameras had a resolution of 1920×1080. However, as the view of cameras captured the whole section of testing road, the image size of each extracted vehicle was much smaller, say around 140×140. As a result, the license plate number of the extracted vehicles cannot be recognized. The frame rate of the video was 25 frames per second.

Video data were collected on a morning peak hour during a weekday (8:00-9:00, 2020 Jan 8, Wednesday) and the same period during weekend as the non-peak hour (8:00-9:00, 2020 Mar 22, Sunday). Manual vehicle matching was conducted to obtain the ground truth.

Using the ground truth data, parameters in the proposed V-ReID method were calibrated for the testing site. The α confidence interval for determining time window size was set as 0.85 in (19-22). The smoothing parameters, β_μ and β_σ, in (16) and (18) were set as 0.6. The normalizing factor γ_{time} in (30) is set as 1. The fusion parameters in (31) were set as $\theta_{lane}=0.4$, $\theta_{time}=0.6$ and $\gamma_{LT}=3.54$to calculate prior probability. The fusion parameters in (35-36) were set as $\theta_{color}=0.4162$, $\theta_{type}=0.3114$, $\theta_{length}=0.2724$, and $\gamma_{CTL}=2.67$ to calculate likelihood functions. The discount parameters in (42-43) were set as $w_{t-1}=0.8$ and $w_{t-1}=0.6$. The termination threshold parameter ε_S in (44) was set as 0.1.

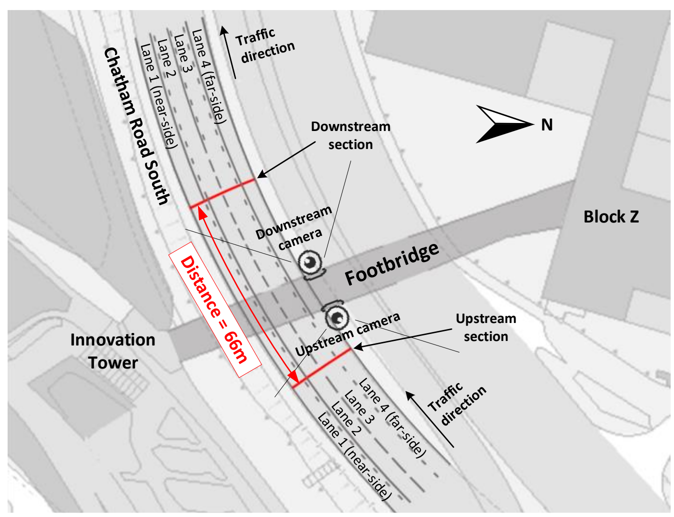


Fig. 8. Road layout of the testing site. The distance between upstream and downstream sections is 66 meters. The complex lane markings in the testing site govern the lane-changing maneuvers, which leads to different traffic conditions at different lanes.

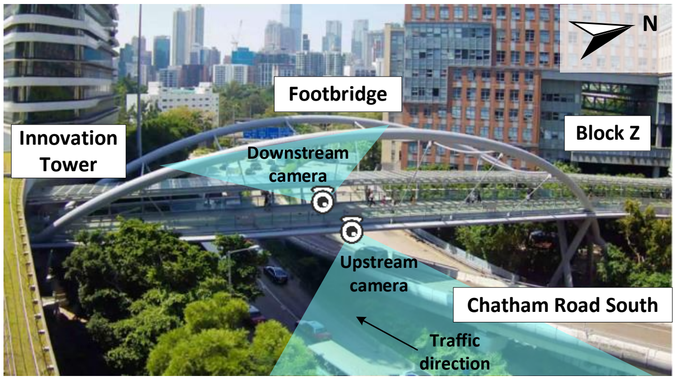


Fig. 9. Photo of the testing site. Two cameras are installed on both sides of the footbridge over the testing road, i.e. Chatham Road South. The upstream camera captures vehicle from the front and the downstream camera from the back.

As for vehicle length measurements, the training dataset with ground truth length was used for calibration. The normalizing factors, η_l, in (15) were set as 0.0288, 0.0247, 0.0237 and 0.0232 for upstream lanes 1 to 4 respectively and 0.0361, 0.0369, 0.03895 and 0.0399 for downstream lanes 1 to 4. An independent dataset with ground truth length was used to validate the calibrated results. The measurement accuracy was evaluated by mean absolute percentage error (MAPE) and root mean square error (RMSE). For upstream and downstream locations, MAPEs were 4.84% and 4.28% respectively and the RMSEs were 0.30 m and 0.46 m respectively. This result suggested a high accuracy of vehicle length measurements achieved in the case study.

B. *Performance Evaluation Metrics*

This section presents the evaluation metrics to assess the V-ReID performance. The vehicle matching accuracy was evaluated by matching error, denoted by ME, as:

$$ME = \frac{N_e^m}{N_{all}^m - N_{null}^m} \times 100\% \tag{45}$$

where N_{all}^m is the total number of vehicles at upstream section, N_{null}^m is the number of vehicles not matching to any downstream vehicle, and N_e^m is the number of wrong matches.

The estimation accuracy of travel time distribution type was evaluated by the type estimation error, denoted by TE_π, as:

$$TE_\pi = \frac{N_e^\pi}{N_T} \times 100\% \tag{46}$$

where N_T is the total number of time periods, and N_e^π is the number of time periods with incorrect estimations of distribution types.

The estimation accuracy of mean travel time was evaluated by the commonly used MAPE and RMSE. They are denoted by $MAPE_\mu$ and $RMSE_\mu$ and calculated by (47) and (48) respectively.

$$MAPE_\mu = \frac{1}{N_T}\sum_{t=1}^{N_T}\left|\frac{\bar{\mu}_{l,t} - \mu_{l,t}}{\mu_{l,t}}\right| \times 100\% \tag{47}$$

$$RMSE_\mu = \sqrt{\frac{1}{N_T}\sum_{t=1}^{N_T}\left(\bar{\mu}_{l,t} - \mu_{l,t}\right)^2} \tag{48}$$

where $\bar{\mu}_{l,t}$ and $\mu_{l,t}$ are respectively the ground truth and estimated mean travel time for lane l during time interval t, and N_T is the total number of time periods.

The estimation accuracy of standard deviation was also evaluated by MAPE and RMSE, denoted by $MAPE_\sigma$ and $RMSE_\sigma$ respectively. Their equations are given by (49) and (50).

$$MAPE_\sigma = \frac{1}{N_T}\sum_{t=1}^{N_T}\left|\frac{\bar{\sigma}_{l,t} - \sigma_{l,t}}{\sigma_{l,t}}\right| \times 100\% \tag{49}$$

$$RMSE_\sigma = \sqrt{\frac{1}{N_T}\sum_{t=1}^{N_T}\left(\bar{\sigma}_{l,t} - \sigma_{l,t}\right)^2} \tag{50}$$

where $\bar{\sigma}_{l,t}$ and $\sigma_{l,t}$ are respectively the ground truth and the estimated standard deviation for lane l during time interval t.

The estimation accuracy of travel time window is evaluated by probability outside the predicted (estimated) time window (POPI) and probability outside the observed time window (POOI) following the previous work [15]. POPI evaluates the percentage of ground truth data outside the estimated time window as:

$$POPI = \frac{100\%}{N_T}\sum_{t=1}^{N_T}\left(1 - \frac{\bar{\Phi}_{l,t}\left(Ub_{l,t}\right) - \bar{\Phi}_{l,t}\left(Lb_{l,t}\right)}{\alpha}\right) \tag{51}$$

where $\bar{\Phi}_{l,t}$ is the CDF of the ground truth of travel time distribution for lane l during time period t, $Lb_{l,t}$ and $Ub_{l,t}$ are lower and upper bounds of the estimated time window for lane l during time period t with respect to α confidence interval. POOI evaluates the percentage of estimated distribution outside the ground truth time window as:

$$POOI = \frac{100\%}{N_T}\sum_{t=1}^{N_T}\left(1 - \frac{\Phi_{l,t}\left(\overline{Ub}_{l,t}\right) - \Phi_{l,t}\left(\overline{Lb}_{l,t}\right)}{\alpha}\right) \tag{52}$$

where $\Phi_{l,t}$ is the CDF of the estimated travel time distribution for lane l during time period t, and $\overline{Lb}_{l,t}$ and $\overline{Ub}_{l,t}$ are lower and upper bounds of the ground truth time window for lane l during time period t. The lower values of POPI and POOI metrics indicate higher accuracy of estimated time window.

C. *Experimental Results*

This section reports case study results. Firstly, the performance of the proposed lane-based V-ReID method was reported during the congested peak hour. Then, the performance of the proposed lane-based V-ReID method was examined by comparing it to the state-of-the-art link-based V-ReID method [26]. Finally, the performance of two V-ReID methods was reported during the uncongested off-peak hour.

1) *Performance of the proposed lane-based V-ReID method*

Fig. 10 reports the results of estimated travel time distributions for Lanes 1 to 4 on every two minutes interval during the congested peak hour. The estimated means, $\mu_{l,t}$, for each lane are given in the figure using red lines. The estimated distribution types, $\pi_{l,t}$, are given on the lines by using either circle or star symbols to respectively represent normal or lognormal distribution. The lower and upper bounds of travel time distributions, $[Lb_{l,t}, Ub_{l,t}]$, are given at every time period by hollow circles. The size of time windows reflects the values of estimated standard deviation $\sigma_{l,t}$. The ground truth data are also given in the figure using circle symbols with different colors. The blue color represents the ground truth data following normal distribution, while the green color representing lognormal distribution.

It can be seen clearly that travel time distributions in terms of all parameters ($\mu_{l,t}$, $\sigma_{l,t}$ and $\pi_{l,t}$) were significantly different among four lanes at the testing road during the congested peak hour. As summarized in Table IV of the ground truth, both mean, $\mu_{l,t}$, and standard deviation, $\sigma_{l,t}$, were dramatically increased from Lane 1 to Lane 4. The average $\mu_{l,t}$ and $\sigma_{l,t}$ during the peak hour of Lane 4 were 57.67 s and 14.39 s, which respectively were 5.6 and 3.6 times as larger as that of Lane 1. The pattern of distribution type, $\pi_{l,t}$, also varied among four lanes. During the peak hour, Lanes 1, 2 and 3 generally followed lognormal distributions, while Lane 4 almost followed normal distributions. Therefore, these observations justify the needs for developing lane-based V-ReID methods to capture distinctive characteristics of lane-level travel time distributions at the same road section.

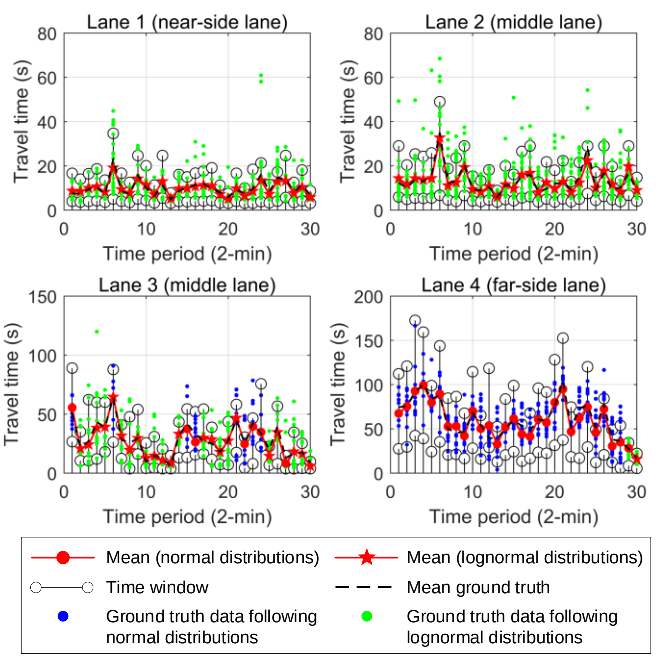


Fig. 10. Travel time estimations of the proposed method during the peak hour. The estimations of travel time distributions are generally close to the ground truth distributions, which were highly dynamic and had significant differences among lanes.

It also can be seen from Fig. 10 that all distribution parameters ($\mu_{l,t}$, $\sigma_{l,t}$ and $\pi_{l,t}$) were highly dynamic during the peak hour, especially on Lanes 3 and 4. For example, the standard deviations $\sigma_{l,t}$ of Lane 3 varied from 6.12 s to 16.64 s; and the distribution types $\pi_{l,t}$ of Lane 3 were lognormal and normal respectively for 73.3% and 26.7% of total time periods. Therefore, these observations highlight the necessities for real-time estimating all parameters of travel time distributions, including not only the mean but also the standard deviation and the distribution type.

TABLE IV

DESCRIPTIVE STATISTICS OF GROUND TRUTH DURING THE PEAK HOUR

Statistics	Link	Lane 1	Lane 2	Lane 3	Lane 4
Average mean (s)	24.93	**10.27**	12.08	25.58	**57.67**
Average standard deviation (s)	21.17	**3.99**	6.22	13.34	**14.39**
Normal distribution ratio	90.0%	0%	0%	**26.7%**	93.3%
Lognormal distribution ratio	10.0%	100%	100%	**73.3%**	6.7%
Average speed (km/h)	9.53	23.14	19.67	9.29	4.12

Table V summarizes the performance of the proposed V-ReID method at the testing road during the peak hour. As summarized, the matching error, ME, for vehicles travelling on Lane 1 was 36.3%. This ME exacerbated to 54.6% for Lane 4, in which traffic conditions are more congested and uncertain. The reason is that heavy traffic flow in congested lanes could increase the number of feasible matching vehicles with similar visual features. Undesirable effects of wrong matching on travel time estimation were mitigated by using the introduced weighting scheme, see (41-42).

TABLE V

PERFORMANCE OF THE PROPOSED METHOD DURING THE PEAK HOUR

Evaluation metrics	Lane 1	Lane 2	Lane 3	Lane 4
ME	**36.3%**	47.2%	52.4%	**54.6%**
$MAPE_{\mu}$	10.3%	10.1%	**11.0%**	8.2%
$RMSE_{\mu}$ (s)	1.3	1.8	3.2	**5.5**
$MAPE_{\sigma}$	13.1%	16.3%	13.6%	**18.6%**
$RMSE_{\sigma}$ (s)	0.5	1.2	2.1	**3.2**
TE_{π}	0.0%	**0.0%**	**10.0%**	3.3%
$POPI$	9.9%	12.8%	**14.2%**	12.2%
$POOI$	22.9%	23.9%	25.1%	**27.9%**

It can be seen from Table V that the proposed V-ReID method can achieve a satisfactory accuracy level of travel time distribution estimations for all lanes. As shown, the mean estimation errors, $MAPE_{\mu}$ was within 11.0% and $RMSE_{\mu}$ ranged from 1.3 s to 5.5 s for all lanes. The standard deviation estimation errors were relatively worse. $MAPE_{\sigma}$ was within 18.6% and $RMSE_{\sigma}$ ranged from 0.5 s to 3.2 s for all lanes. This is due to the estimation of $\sigma_{l,t}$ requiring more samples than that of $\mu_{l,t}$. When sample size was not sufficient, the proposed method tried to utilize the samples from previous time periods, $t-1$ and $t-2$. This treatment may introduce unnecessary temporal variation of travel times, leading to a slight overestimation of $\sigma_{l,t}$. The distribution type estimation errors, TE_{π}, were consistently low for Lanes 1, 2 and 4. For Lane 3, its $TE_{\pi} = 10\%$ was relatively larger than other lanes, since its distribution type varied a lot at different times of the day.

It can also be seen in Table V that the proposed method can produce a reasonable accuracy level of travel time windows. As shown, the $POPI$ metrics for all lanes were within 14.2% indicating that the constructed time windows can capture a high proportion (i.e., over 85.8%) of ground truth data. The $POOI$ metrics for all lanes were within 27.9%, showing that a small proportion (i.e. less than 27.9%) of estimated travel time distribution outside the ground truth time windows.

2) Comparison of Different V-ReID Methods

To further evaluate and benchmark the proposed V-ReID method, the state-of-the-art link-based V-ReID method [26] described in Section III was also implemented. For convenience, this implemented method is hereafter called Wang et al.'s method.

Compared to Wang et al.'s method, the proposed V-ReID method made two major improvements. Firstly, it introduced a lane-based V-ReID framework to capture distinctive traffic conditions among different lanes. Secondly, it estimated full spectrum of distribution parameters including not only the mean $\mu_{l,t}$ but also the standard deviation $\sigma_{l,t}$ and the distribution type $\pi_{l,t}$.

To distinguish the effects of these two improvements, another method was implemented by modifying the proposed V-ReID method. This modified method still followed the lane-based V-ReID framework, but estimated only the $\mu_{l,t}$ on real-time basis. Following Wang et al.'s method, it assumed $\pi_{l,t}$ of all lanes always following the normal distributions, and calculated $\sigma_{l,t} = \phi\mu_{l,t}$, where ϕ is the coefficient of variation which is assumed to be fixed and pre-given based on the historical data. In addition, the modified method adopted the thresholding process of Wang et al.'s method instead of the weighting scheme used in the proposed method.

In this case study, the same ϕ was set for both modified method and Wang et al.'s method by dividing average link mean travel time by average link standard deviation, e.g., ϕ = 21.17 / 24.93 during the peak hour (see Table IV). The same set of parameters given in Section V-A was used for all three methods. For comparisons, link travel time estimations by the modified and proposed methods were also provided.

The estimation performance comparison was firstly made between Wang et al.'s method and the modified method, to investigate the effects of the first improvement, i.e. from link-based V-ReID to lane-based V-ReID. As summarized in Table VI, the performance of Wang et al.'s method was unsatisfactory on the testing urban road during the peak hour. Its mean travel time estimation error, $MAPE_\mu$, was over 54% for the whole link, and even over 80% for Lane 4, i.e., the most congested link. The estimation errors of standard deviation are also quite large. $MAPE_\sigma$ ranged from 45.4% to 147% for all lanes. This is due to the link-based V-ReID framework used by Wang et al.'s method. As shown in Fig.11, Wang et al.'s method produced a single mean travel time estimation to all lanes during any time period and thereby lacked capabilities to capture distinctive traffic conditions among four lanes. Using the estimated link travel times, the constructed link-based travel time windows, however, cannot well cover feasible matches between upstream and downstream vehicles for congested lanes, e.g., $POPI =$ 83.1% and $POOI =$ 94.2% for Lane 4. Consequently, a large proportion of vehicles were mis-matched for congested lanes, e.g., $ME =$ 93.2% for Lane 4. The large vehicle matching errors further caused the exclusion of samples from congested lanes in the mean travel time estimations.

TABLE VI
PERFORMANCE OF THREE TESTING METHODS DURING THE PERK HOUR

	Method	Link	Lane 1	Lane 2	Lane 3	Lane 4
ME	Wang et al.	62.7%	57.8%	54.1%	62.6%	**93.2%**
	Modified	51.8%	45.2%	48.8%	53.9%	**67.5%**
	Proposed	47.6%	36.3%	47.2%	52.4%	**54.6%**
$MAPE_\mu$	Wang et al.	**54.8%**	21.6%	18.7%	53.6%	**80.1%**
	Modified	**10.0%**	15.4%	19.9%	20.0%	**11.8%**
	Proposed	5.6%	10.3%	10.1%	**11.0%**	8.2%
$RMSE_\mu$ (s)	Wang et al.	16.1	3.1	5.4	21.5	**54.2**
	Modified	3.6	2.4	4.3	6.2	**9.8**
	Proposed	1.6	1.3	1.8	**3.2**	5.5
$MAPE_\sigma$	Wang et al.	58.9%	**147%**	89.2%	45.4%	47.1%
	Modified	22.1%	95.6%	79.9%	107%	**289%**
	Proposed	12.3%	13.1%	16.3%	13.6%	**18.6%**
$RMSE_\sigma$ (s)	Wang et al.	13.3	6.3	5.0	7.3	7.6
	Modified	5.0	4.4	5.3	**18.0**	**34.4**
	Proposed	2.8	0.5	1.2	2.1	3.2
TE_π	Wang et al.	10.0%	100%	100%	73.3%	6.7%
	Modified	10.0%	100%	100%	73.3%	6.7%
	Proposed	0.0%	0.0%	0.0%	10.0%	3.3%
$POPI$	Wang et al.	31.1%	4.5%	2.2%	34.6%	**83.1%**
	Modified	6.1%	3.1%	9.5%	5.4%	**6.6%**
	Proposed	12.3%	9.9%	12.8%	**14.2%**	12.2%
$POOI$	Wang et al.	91.4%	94.7%	94.0%	82.8%	94.2%
	Modified	84.2%	94.0%	94.3%	82.6%	**65.9%**
	Proposed	25.0%	22.9%	23.9%	25.1%	**27.9%**

It can be seen in Table VI that the modified method consistently outperformed Wang et al.'s method. By using the proposed lane-based V-ReID framework, the modified method can produce different mean travel time estimations, $\mu_{l,t}$, for four lanes (see Fig. 11). This can reduce $\mu_{l,t}$ estimation error for congested lanes, e.g., $MAPE_\mu$ of Lane 4 was reduced dramatically from 80.1% to 11.8%, and the corresponding $RMSE_\mu$ was reduced from 54.2 s to 9.8 s. With lane-based travel time estimations, lane-level travel time windows were constructed to determine more feasible matches at different lanes. Take Lane 4 for example, $POPI$ was reduced from 83.1% to 6.6% and $POOI$ was reduced from 94.2% to 65.9%. The vehicle matching errors were reduced for all lanes, e.g., ME for Lane 4 was reduced from 93.2% to 67.5%.

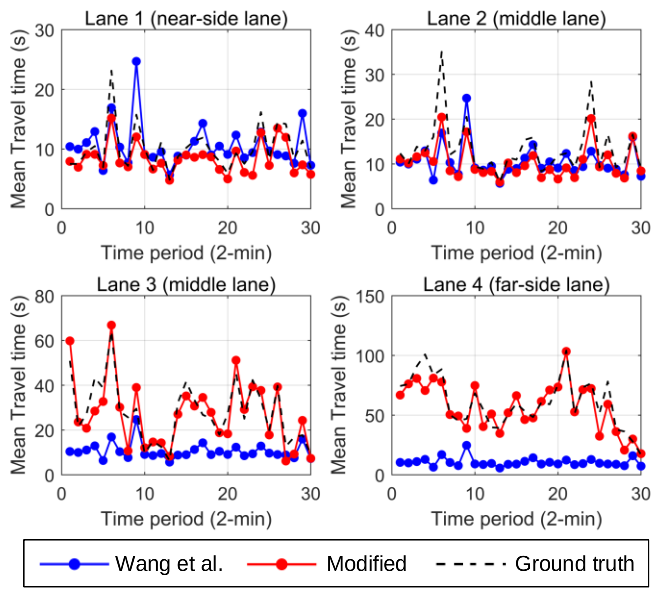


Fig. 11. Mean travel time estimations by Wang et al.'s method and the modified method during the peak hour. Wang et al.'s method produced the same travel time estimation for all lanes, while the modified method can distinguish the differences among lanes and produce more accurate estimations.

The comparison was also made between the modified method and the proposed method to examine the effects of the second improvement, i.e., real-time estimation of all distribution parameters including not only $\mu_{l,t}$ but also $\sigma_{l,t}$ and $\pi_{l,t}$. As summarized in Table VI, the proposed method significantly outperformed the modified method on $\pi_{l,t}$ estimation accuracy. This result is obvious, because the proposed method estimated the distribution type on a real time basis, rather than using the unrealistic assumption of all travel time distributions following normal distributions during the period of interest. As summarized in Table IV, Lanes 1 and 2 followed lognormal distributions, while the distribution types of Lanes 3 and 4 varied at different times of the day.

The proposed method also significantly outperformed the modified method on $\sigma_{l,t}$ estimation accuracy, as shown in Table VI. For example, $MAPE_\sigma$ of Lane 4 generated by the modified method was 289%, which was about 14.5 times larger than that produced by the proposed method. This result is expected, because the proposed method directly estimates standard deviation of travel times instead of indirect estimation adopted by the modified method, in which $\sigma_{l,t}$ was estimated by using a pre-given and fixed coefficient of variation, ϕ. As illustrated in Fig. 12, the actual ϕ values varied significantly on different lanes and at different times of the day. The use of this simple assumption can introduce remarkable bias (i.e., overestimation in this case) on the $\sigma_{l,t}$ estimation.

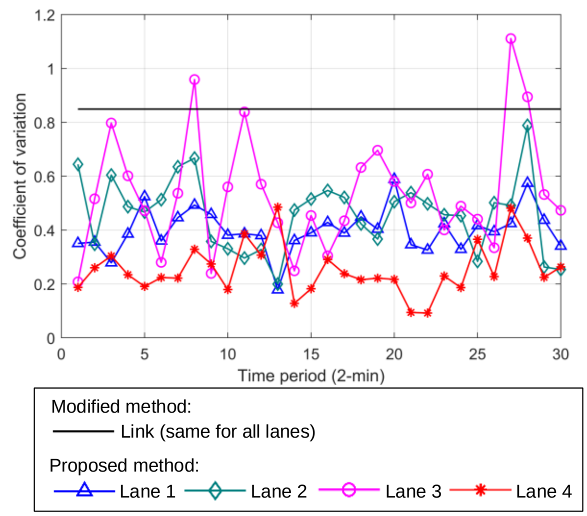


Fig. 12. Coefficients of variation values during the peak hour. The time-varying coefficients of variation adopted by the proposed method are generally smaller and more accurate than the fixed value taken by the modified method.

The improvement of estimations accuracy on the distribution type and the standard deviation led to more accurate time windows. The $POPI$ and $POOI$ metrics of the proposed method were respectively within 14.2% and 27.9% for all lanes. On the contrary, in the modified method, the overestimation of $\sigma_{l,t}$ together with mis-identification of $\pi_{l,t}$ can make the constructed time windows too large. Although such time windows covered a large proportion of vehicle matches (e.g., $POPI = 6.6\%$ for Lane 4), they also included too many infeasible vehicle matches (e.g., $POOI = 65.9\%$ for Lane 4), which could degrade matching accuracy. As a result, the matching errors, ME, of the modified method were consistently worse than the proposed method for all four lanes. For example, $ME = 67.5\%$ of Lane 4 for the modified method was higher than $ME = 54.6\%$ for the proposed method.

As shown in Table VI, the proposed method consistently outperforms the modified method on the $\mu_{l,t}$ estimation accuracy for all four lanes. Take Lane 3 for example, $MAPE_\mu$ was 11% and $RMSE_\mu$ was 3.2 s for the proposed method, which was only half of that for the modified method. This improvement is mainly due to the introduced weighting scheme in the proposed method. To mitigate the effects of wrong matches, the modified method employed Wang et al.'s threshold process by excluding unreliable matches with distinctive index lower than a given threshold. This threshold process was extended in the introduced weighting scheme by explicitly setting a higher weighting to vehicles with more distinctive visual features. In addition, the consequence of wrong matches was considered in the weighting scheme for setting a low weighting to those matches with a large travel time difference to the optimal match.

3) Performance of V-ReID Methods during Off-peak Hour

Table VII summarizes the V-ReID performance for both Wang et al.'s method and the proposed method during the uncongested off-peak hour. The coefficient of variation parameter during the off-peak hour was calibrated using the ground true data as ϕ = 0.63 / 5.43 for Wang et al.'s method (see Table VIII).

TABLE VII
PERFORMANCE OF TWO TESTING METHODS DURING THE OFF-PERK HOUR

	Method	Link	Lane 1	Lane 2	Lane 3	Lane 4
ME	Wang et al.	17.9%	**14.4%**	22.3%	18.9%	15.1%
	Proposed	6.2%	**0.8%**	10.7%	4.8%	6.3%
$MAPE_\mu$	Wang et al.	3.8%	**5.5%**	6.4%	6.3%	6.7%
	Proposed	1.1%	**0.1%**	2.3%	0.7%	1.0%
$RMSE_\mu$ (s)	Wang et al.	0.43	0.44	0.52	0.47	0.52
	Proposed	0.17	0.06	0.21	0.07	0.11
$MAPE_\sigma$	Wang et al.	21.1%	**103%**	60.9%	41.4%	26.5%
	Proposed	1.5%	**0.3%**	3.2%	1.1%	1.6%
$RMSE_\sigma$ (s)	Wang et al.	0.15	0.38	0.28	0.19	0.17
	Proposed	0.01	0.00	0.02	0.01	0.01
TE_π	Wang et al.	0%	30.0%	30.0%	73.3%	56.7%
	Proposed	0%	0%	0%	0%	0%
$POPI$	Wang et al.	8.5%	6.8%	13.4%	10.8%	2.9%
	Proposed	2.9%	1.1%	4.8%	2.2%	3.5%
$POOI$	Wang et al.	43.9%	**62.0%**	62.6%	38.7%	12.4%
	Proposed	4.7%	**3.3%**	7.6%	4.3%	3.6%

As summarized, Wang et al.'s method performed well in terms vehicle matching and mean travel time estimation during the off-peak hour. Its ME, $MAPE_\mu$ and $RMSE_\mu$ for all lanes were within 23%, 7% and 0.6 s respectively. This is because the light traffic volume produced much less candidate matches compared to the congested counterpart during the peak hour, and traffic conditions of all lanes are smooth and homogeneous (see Table VIII). It can also be seen that Wang et al.'s method can correctly identify the distribution type for link travel times. As shown in Table VIII, although travel times may not follow normal distributions for all lanes, aggregating them made the link travel time distribution following normal distributions. However, Wang et al.'s method failed to achieve a satisfactory level of $\sigma_{l,t}$ estimation, e.g., $MAPE_\sigma = 103\%$ for Lane 1. This result highlighted the dynamic nature of coefficient of variation ϕ and the needs for real-time estimation of $\sigma_{l,t}$ during uncongested off-peak hours.

TABLE VIII
DESCRIPTIVE STATISTICS OF GROUND TRUTH DURING THE OFF-PEAK HOUR

Statistics	Link	Lane 1	Lane 2	Lane 3	Lane 4
Average mean (s)	**5.43**	5.55	5.41	5.47	5.34
Average standard deviation (s)	**0.63**	0.32	0.47	0.67	0.60
Normal distribution ratio	100%	70.0%	70.0%	26.7%	43.3%
Lognormal distribution ratio	0%	30.0%	30.0%	73.3%	56.7%
Average speed (km/h)	43.78	42.82	43.94	43.40	44.47

As can be seen, the proposed method still significantly outperformed Wang et al.'s method during the off-peak hour for all evaluation metrics. The $\sigma_{l,t}$ estimation error was dramatically reduced for all lanes, e.g., $MAPE_\sigma$ of Lane 1 was reduced from 103% to 0.3%. More accurate lane-based time windows were constructed, e.g., $POOI$ of Lane 1 was reduced from 62% to 3.3%. Accordingly, a higher accuracy of vehicle matching and mean travel time estimation was achieved for all lanes, e.g., ME for Lane 1 was reduced from 14.4% to 0.8%, and $MAPE_\mu$ for Lane 1 was reduced from 5.5% to 0.1%.

VI. CONCLUSION

Vehicle re-identification (V-ReID) approach is a promising technique to estimate traffic conditions using video images from widely deployed surveillance cameras. However, most existing V-ReID methods focused on the estimation of link travel times in freeways, but failed to capture distinctive traffic conditions among different lanes on congested urban roads. To this end, this paper proposed a new V-ReID method to accurately estimate lane-level travel time distributions on congested urban roads. In the proposed method, adaptive lane-based travel time windows are constructed to delimit feasible matches between downstream and upstream vehicles. A lane-based bipartite graph matching is developed to obtain optimal vehicle matches by explicitly considering visual features, lane changing behaviors and arrival time probability at different lanes. A weighting scheme is introduced to mitigate the effects of wrong vehicle matches. A lane-based travel time estimation technique is developed to real-time estimate lane-level distribution parameters, including not only the mean but also the standard deviation and the distribution type.

To demonstrate the effectiveness of the proposed method, a comprehensive case study was carried out on a congested urban road in Hong Kong. Results of case study justified the significant differences of traffic conditions among lanes of the same road section; and highlighted the dynamics of lane-level distribution parameters (i.e., mean, standard deviation and distribution type) required for real-time estimations. Results of case study showed that the proposed V-ReID method can provide accurate estimation of lane-level distributions on congested urban roads during peak and off-peak hours. The proposed method consistently remarkably outperformed the state-of-the-art link-based V-ReID method [26] with respect to both vehicle matching accuracy and travel time distribution estimation accuracy.

Several further research directions are worth noting. First, travel time distributions in this study were assumed to follow either normal or lognormal distributions. However, previous empirical studies found that travel times can also follow other distribution types, such as Gamma [3] or Burr [34]. Further studies, thereby, should be carried out to incorporate other distribution types into the proposed method. Second, this study only extracted vehicle color, type and length features for vehicle matching. Using video images from specific angles of view, more subtle features, such as vehicles' roof patterns, side views and advertisements, can also be extracted. How to incorporate such subtle features into V-ReID studies is a topic for further study. Third, this study only estimated lane-level travel time distributions for the current time period. How to predict the lane-level travel time distributions in short-term periods is another interesting topic for further study. Last but not the least, the proposed V-ReID method utilizes visual features of video images. However, visual features are sensitive to low illumination conditions such as adverse weather and night time. Further studies are required to extend the proposed method by using thermal cameras that are not sensible to such low illumination conditions.

References

[1] B. Y. Chen, W. H. K. Lam, Q. Li, A. Sumalee, and K. Yan, "Shortest Path Finding Problem in Stochastic Time-Dependent Road Networks With Stochastic First-In-First-Out Property," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1907-1917, 2013.

[2] B. Y. Chen, Q. Li, and W. H. K. Lam, "Finding the k reliable shortest paths under travel time uncertainty," *Transp. Res. Part B,* vol. 94, no. 1, pp. 189-203, 2016.

[3] B. Y. Chen, C. Shi, J. Zhang, W. H. K. Lam, Q. Li, and S. Xiang, "Most reliable path-finding algorithm for maximizing on-time arrival probability," *Transp. B,* vol. 5, no. 3, pp. 248-264, 2017.

[4] L. Lu, Z. He, J. Wang, J. Chen, and W. Wang, "Estimation of lane-level travel time distributions under a connected environment," *J. Intell. Transp. Syst.*, vol. 25, no. 5, pp. 501-512, 2021.

[5] B. Y. Chen, H. Yuan, Q. Li, W. H. K. Lam, S.-L. Shaw, and K. Yan, "Map-matching algorithm for large-scale low-frequency floating car data," *Int. J. Geogr. Inf. Sci.*, vol. 28, no. 1, pp. 22-38, 2014.

[6] J. Zhao, V. L. Knoop, and M. Wang, "Two-dimensional vehicular movement modelling at intersections based on optimal control," *Transp. Res. Part B,* vol. 138, no. 1, pp. 1-22, 2020.

[7] B. Abdulhai and S. M. Tabib, "Spatio-temporal inductance-pattern recognition for vehicle re-identification," *Transp. Res. Part C,* vol. 11, no. 3-4, pp. 223-239, 2003.

[8] X. Liu, W. Liu, T. Mei, and H. Ma, "PROVID: Progressive and Multimodal Vehicle Reidentification for Large-Scale Urban Surveillance," *IEEE Trans. Multimed.,* vol. 20, no. 3, pp. 645-658, 2018.

[9] H. Patel and K. P. Upla, "Night Vision Surveillance: Object Detection using Thermal and Visible Images," in *Proc. 2020 International Conference for Emerging Technology*, June 2020, pp. 1-6.

[10] K. Hyun, A. Tok, and S. G. Ritchie, "Long distance truck tracking from advanced point detectors using a selective weighted Bayesian model," *Transp. Res. Part C,* vol. 82, no. 1, pp. 24-42, 2017.

[11] Q. Yu, B. Wang, and Y. Su, "Object detection-tracking algorithm for unmanned surface vehicles based on a radar-photoelectric system," *IEEE Access,* vol. 9, no. 1, pp. 57529-57541, 2021.

[12] M. P. Muresan, S. Nedevschi, and I. Giosan, "Real-Time Object Detection Using a Sparse 4-Layer LIDAR," in *Proc. 13th IEEE International Conference on Intelligent Computer Communication and Processing*, Sept. 2017, pp. 317-322.

[13] R. Li, G. Rose, and M. Sarvi, "Evaluation of speed-based travel time estimation models," *J. Transp. Eng.,* vol. 132, no. 7, pp. 540-547, 2006.

[14] Z. Jiang, X. Chen, and Y. Ouyang, "Traffic state and emission estimation for urban expressways based on heterogeneous data," *Transp. Res. Part D,* vol. 53, no. 1, pp. 440-453, 2017.

[15] C. Shi, B. Y. Chen, W. H. K. Lam, and Q. Li, "Heterogeneous Data Fusion Method to Estimate Travel Time Distributions in Congested Road Networks," *Sensors,* vol. 17, no. 12, pp. 1-22, 2017.

[16] C. Sun, S. G. Ritchie, K. Tsai, and R. Jayakrishnan, "Use of vehicle signature analysis and lexicographic optimization for vehicle reidentification on freeways," *Transp. Res. Part C,* vol. 7, no. 4, pp. 167-185, 1999.

[17] B. Coifman and M. Cassidy, "Vehicle reidentification and travel time measurement on congested freeways," *Transp. Res. Part A,* vol. 36, no. 10, pp. 899-917, 2002.

[18] C. C. Sun, G. S. Arr, R. P. Ramachandran, and S. G. Ritchie, "Vehicle Reidentification Using Multidetector Fusion," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 3, pp. 155-164, 2004.

[19] K. Kwong, R. Kavaler, R. Rajagopal, and P. Varaiya, "Arterial travel time estimation based on vehicle re-identification using wireless magnetic sensors," *Transp. Res. Part C,* vol. 17, no. 6, pp. 586-606, 2009.

[20] G. Basar, M. Cetin, and A. P. Nichols, "Comparison of vehicle re-identification models for trucks based on axle spacing measurements," *J. Intell. Transp. Syst.,* vol. 22, no. 6, pp. 517-529, 2018.

[21] S. D. Khan and H. Ullah, "A survey of advances in vision-based vehicle re-identification," *Comput. Vis. Image Underst.,* vol. 182, no. 1, pp. 50-63, 2019.

[22] A. Sumalee, J. Wang, K. Jedwanna, and S. Suwansawat, "Probabilistic Fusion of Vehicle Features for Reidentification and Travel Time Estimation Using Video Image Data," *Transp. Res. Rec.,* vol. 2308, no. 1, pp. 73-82, 2012.

[23] F. M. Oliveira-Neto, L. D. Han, and M. K. Jeong, "Online license plate matching procedures using license-plate recognition machines and new weighted edit distance," *Transp. Res. Part C,* vol. 21, no. 1, pp. 306-320, 2012.

[24] Z. Tang, G. Wang, H. Xiao, A. Zheng, and J.-N. Hwang, "Single-camera and inter-camera vehicle tracking and 3D speed estimation based on fusion of visual and semantic features," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops*, June 2018, pp. 108-115.

[25] Y. Bai, Y. Lou, F. Gao, S. Wang, Y. Wu, and L.-Y. Duan, "Group-Sensitive Triplet Embedding for Vehicle Reidentification," *IEEE Trans. Multimed.,* vol. 20, no. 9, pp. 2385-2399, 2018.

[26] J. Wang, N. Indra-Payoong, A. Sumalee, and S. Panwai, "Vehicle Reidentification With Self-Adaptive Time Windows for Real-Time Travel Time Estimation," *IEEE Trans. Intell. Transp. Syst.,* vol. 15, no. 2, pp. 540-552, 2014.

[27] Z. Galil, "Efficient algorithms for finding maximum matching in graphs," *ACM Computing Surveys,* vol. 18, no. 1, pp. 23-38, 1986.

[28] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *ArXiv:1804.02767*. Retrieved October 10, 2020, from http://arxiv.org/abs/1804.02767.

[29] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *Proc. 2017 IEEE Int. Conf. Image Process.*, Sept. 2017, pp. 3645-3649.

[30] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *Proc. 17th Int. Conf. Pattern Recognit.*, Aug. 2004, vol. 2, pp. 28-31.

[31] A. T. G. Thiang and R. Lim, "Type of vehicle recognition using template matching method," in *Proc. Int. Conf. Electr. Electron. Commun. Inf.*, Mar. 2001, pp. 1-5.

[32] B. Y. Chen, W. H. K. Lam, A. Sumalee, Q. Li, and M. L. Tam, "Reliable Shortest Path Problems in Stochastic Time-Dependent Networks," *J. Intell. Transp. Syst.*, vol. 18, no. 2, pp. 177-189, 2014.

[33] M. Yun and W. Qin, "Minimum Sampling Size of Floating Cars for Urban Link Travel Time Distribution Estimation," *Transp. Res. Rec.*, vol. 2673, no. 3, pp. 24-43, 2019.

[34] M. A. P. Taylor, "Fosgerau's travel time reliability ratio and the Burr distribution," *Transp. Res. Part B,* vol. 97, no. 1, pp. 50-63, 2017.

Cheng Zhang received the B.Eng. degree in traffic engineering from Tongji University, Shanghai, China, in 2016, where he is currently pursuing the Ph.D. degree with the Key Laboratory of Road and Traffic Engineering of the Ministry of Education. He was also a Research Assistant with the Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong. His research interests include travel time estimation, traffic management and control, big data analytics in transportation, and computer vision.

Bi Yu Chen (Member, IEEE) received the Bachelor's degree in surveying and mapping engineering and the Master's degree in geographical information science from Wuhan University, Wuhan, China, in 2003 and 2006, respectively, and the Ph.D. degree in transportation engineering from The Hong Kong Polytechnic University, Kowloon, Hong Kong, in 2012.

He is currently a Full Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. He has authored or coauthored more than 40 articles in leading international journals. His research interests are in the areas of intelligent transportation systems, transport geography, GIS for transportation and spatiotemporal big data analytics. He served as an Associate Editor for *Transportmetrica B*.

William H. K. Lam received B.S. and M.S. degrees from the University of Calgary, Calgary, AB, Canada, in 1978 and 1980, respectively, and a Ph.D. degree in transportation engineering from the University of Newcastle upon Tyne, Newcastle, U.K., in 1992.

He has also been an Honorary Professor at the Institute for Transport and Logistics Studies, The University of Sydney, Australia since 2015. He is currently a Chair Professor of Civil and Transportation Engineering with the Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong. His current research interests include transport planning and traffic forecasting, ITS technology and development, smart surveillance and traffic simulation, public transport, and pedestrian studies. He is also the Convenor of the International Advisory Committee of the *International Symposium on Transportation and Traffic Theory (ISTTT)* and a member of the International Scientific Committee of the *International Symposium on Transportation Network Reliability (INSTR)*. He is also the Founding Editor-in-Chief of *Transportmetrica* and is now one of the Co-Editors-in-Chief of *Transportmetrica A: Transport Science*.

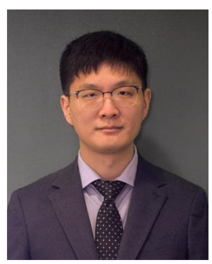
H. W. Ho received his BEng and Ph.D. degree from the Department of Civil Engineering, the University of Hong Kong in 2001 and 2005 respectively. He is currently a Research Fellow in the Department of Civil and Environmental Engineering of The Hong Kong Polytechnic University. He is also a chartered member of The Chartered Institute of Logistics and Transport in Hong Kong (CILTHK). He has authored or coauthored more than 20 journal papers and serve as peer reviewer for several international journals. His current research interests include travel time estimation and prediction, macroscopic models for transportation system, activity-based demand modelling and land-use and transport optimization.

Xiaomeng Shi received the B.Eng. and Ph.D. degrees in Transportation Engineering from School of Transportation, Southeast University, Nanjing, China, in 2013 and 2018, respectively. He is currently an Assistant Professor with School of Transportation, Southeast University. His research interests include the pedestrian flow modeling and analysis, urban mobility analytics and intelligent transportation systems.

Xiaoguang Yang received the B.Eng. and M.S. degrees from Tongji University, Shanghai, China, in 1982 and 1988, respectively, and the Ph.D. degree in intelligent transportation systems under a joint supervision scheme in Kyoto University, Japan, in 1996.

From 2002 to 2014, he was the Founding Head of the Department of Traffic Engineering, Tongji University. He is currently a Tenured Professor with the Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, and the director of Intelligent Transportation Systems (ITS) Research Center, Tongji University. He is also the Director of Traffic Safety Research Laboratory, Shanghai Institute of Disaster Prevention and Relief. His current research interests include transportation engineering, intelligent transportation systems, smart city and intelligent society.

Prof. Yang was awarded more than 10 National or Provincial Prizes for Scientific and Technology Progress.



Wei Ma (Member, IEEE) is currently an Assistant Professor of transportation engineering in the Department of Civil and Environmental Engineering (CEE) at The Hong Kong Polytechnic University (PolyU). His research focuses on the intersection of machine learning, mobility data mining, and transportation network modelling, with applications for smart, efficient, and sustainable mobility systems.

S. C. Wong received the B.Sc. degree in engineering and the M.Phil. degree from The University of Hong Kong (HKU), Pok Fu Lam, Hong Kong, and the Ph.D. degree from University College London, London, U.K. He is currently a Chair Professor with the Department of Civil Engineering, HKU. His research interests include optimization of traffic signal settings, continuum modeling for traffic equilibrium problems, land use and transportation problems, dynamic highway and transit assignment problems, urban taxi services, and road safety.

Andy H. F. Chow (Member, IEEE) received the Ph.D. degree in transport studies from University College London (UCL), London, U.K., in 2007. He was a Senior Lecturer in transport studies with UCL, a Post-Doctoral Researcher with the University of California Berkeley, under the California PATH Program. He is currently an Assistant Professor with the Department of Advanced Design and Systems Engineering, City University of Hong Kong (CityU). His research focuses on modeling and optimizing dynamic transport systems. He received the Gordon Newell Memorial Dissertation Prize in 2008 for his Ph.D. dissertation on optimal control of dynamic transport networks.