IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 1 June 2022, accepted 26 June 2022, date of publication 4 July 2022, date of current version 11 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3188306

APPLIED RESEARCH

Infrastructure-Based Pedestrian Risk Tagging Methodology to Support AV Risk Assessment

WEI MING DAN CHIA[1,2], (Member, IEEE), SYE LOONG KEOH[2,3], (Member, IEEE), ANNA LITO MICHALA[2], AND CINDY GOH[3], (Senior Member, IEEE)

[1]Infocomm Technology (ICT), Singapore Institute of Technology, Singapore 138683
[2]School of Computing Science, University of Glasgow, Glasgow G12 8QQ, U.K.
[3]University of Glasgow Singapore, Singapore 599493

Corresponding author: Wei Ming Dan Chia (dan.chia@singaporetech.edu.sg)

This work was supported in part by the Ignition Grant from the Singapore Institute of Technology. The work of Wei Ming Dan Chia was supported by the University of Glasgow (UofG)-Singapore Institute of Technology (SIT) Ph.D. Scholarship.

ABSTRACT Safety is paramount in AV deployment. Traditionally, AV safety is incorporated during development by identifying the failure of vehicular components using risk and scenarios-based validation approaches with risk indicators. The main challenge of having a comprehensive risk assessment for AV safety is to include all potential complex environments that could occur in real-time, which is more critical for higher AV automation levels. Real-world real-time risk assessment research addresses this shortcoming by providing an advanced warning to the AV during deployment either at the vehicle level or at the infrastructure level. This paper proposes a risk tagging methodology to quantitatively risk tag the severity rating of the real-world environment for real-time risk assessment of AV using the existing roadside infrastructure. The proposed methodology - Spatial-Temporal Risk Estimation Ensemble Technique (STREET), provides advanced risk indicators in the form of pedestrian risk tag figure and time to collision value to the AV. This paper includes the evaluation of STREET, tested on four events over a pre-defined uncontrolled traffic scene from the infrastructure and validated using ground truth and heatmap of pedestrian occurrence. This methodology includes three different algorithms developed to emphasize different events depending on the AV risk and safety management strategy. The STREET reduces the bandwidth needed compared to traditional approaches of streaming video images for lightweight integration of AV risk assessment. The outcome of the pedestrian risk tag from STREET can be used as a severity rating for the existing real-time risk assessment of AV via cooperative mode.

INDEX TERMS Cooperative mode, edge computing, risk algorithm, risk analysis.

I. INTRODUCTION

As Autonomous Vehicles (AV) technology matures and moves gradually from research into industrial deployment, the focus on safety and risk management becomes a pivotal domain. The proven successful adoption of the Advanced Driver Assisted System (ADAS) into the automotive market demonstrates the importance of introducing safety and risk awareness to the drivers at low AV automation levels to prevent accidents in a proactive mode. The increasing demand for higher automation levels of AV [1] due to expanding upcoming market segments (such as mobility as a service) calls for increasing demand for AV safety and risk management. This has resulted in an exponential tightening of safety requirements for AVs' commercialization. The safety requirements include sensing hazards or risks in advance to allow AV sufficient time to react. This time to react requires an extension of the AV sensors to look beyond its operating range limits to prevent collision and provide a higher comfort level to the passengers. Thus, remote support of environment sensing with the help of Road Side Infrastructure (RSI) sensing is needed. Such support with RSI working in cooperative mode, with environment data sent to the AV (Infrastructure to Vehicle) should be lightweight, less resource-demanding,

The associate editor coordinating the review of this manuscript and approving it for publication was Haluk Eren.

and low latency to be feasible for real-world deployment. The data content should also provide information for advanced risk assessment for the operating AV.

Risk assessment for AV started with the adoption of traditional ISO 26262 [2] for vehicles, which focuses on the Automotive Safety Integrity Level (ASIL) rating [1] during the development lifecycle. ASIL rating depends on parameters such as exposure, severity and controllability. This rating is mainly affected by electronic components within the vehicle and is less dependent on environmental conditions. Therefore, when vehicles are driven in real-world settings (different from simulated R&D environment), the driver will need to mitigate unidentified risks in areas such as uncontrolled junctions and dynamic changes in environmental and road conditions. However, with the high level of vehicle automation where ADS is responsible for the AV performance, the control is no longer in the hands of the driver, thus portraying a challenge in using ASIL rating for AV. Taking this into consideration, an additional standard ISO/PAS 21448 [3], also known as Safety of the Intended Functionality (SOTIF), was released to identify risks resulting from functional insufficiencies corresponding to both software and hardware performance limitations. This led to the exponential increase in scenario-based validation [4]–[6] extracted from real-world scenarios during development to form solutions for the identified risks. However, even with the combined scenario-based and traditional validation approach, gaps exist between the real-world and development considerations. As a result, real-time risk assessment frameworks are proposed to AV with advanced warnings beyond the vehicle sensors' limitation [7]–[11] to address and complement these shortcomings.

This paper aims to provide the AV with advanced warnings using RSI in cooperative mode to enhance the real-time risk assessment frameworks. The advanced warning is achieved by providing a severity rating in the form of risk tagging (RT) figures and time to collision with the detected pedestrian. The RT is computed by the RSI utilizing edge computing to convert high bandwidth video images into low bandwidth time-series data. This will greatly reduce the AV's need to be resource-intensive, thus making the overall risk assessment framework more lightweight. RT figures can be used to represent severity ratings for real-time risk assessment frameworks (modified real-time ASIL [9]), while the time to collision proposed in this paper would serve as an advanced warning to the AV in preventing collision with the pedestrian.

This research focuses on uncontrolled traffic areas where pedestrian RT is critical (higher risk) than controlled traffic junctions with clearly defined rules for the AV. The uncontrolled traffic areas typically consist of complex scenarios between pedestrians and AV during real-time occurrence, which may not have been considered during development time.

The contributions of this work are as follows

- Firstly, Spatial-Temporal Risk Estimation Ensemble Technique (STREET) is proposed to provide quantitative RT of the real-world environment as a severity rating for real-time risk assessment of AV. Within STREET, novel environmental risk zoning methods are developed with RT algorithms leveraging the state of the art object recognition approach. Three RT algorithms are developed to emphasize different events depending on the AV risk and safety management strategy. In addition, a time to collision is determined as an advanced warning to prevent pedestrian collision.
- The second domain includes the deployment of STREET in a real-world scene, i.e., tested with a large dataset and demonstrated the detection of pedestrian crossing and occlusion events within the scene. The results include a comparative analysis of the three RT algorithms and time to collision (TTC) figures.

This paper is organized as follows. Section II reviews and compares different AV and cooperative mode risk assessment methodologies, especially between AVs to AVs. Since AV fleets are challenging to deploy in real-world settings, having V2V risk assessment concepts is even more challenging to realize. Therefore, this leads to the interest in having a risk tagging methodology (between AV to pedestrian(s)) in a cooperative mode using I2V remotely and supported by time to collision. This proposed methodology is described in Section III, which illustrates the use of STREET for different scenes with further detailed steps of the process explained. STREET is used within an edge installed in the infrastructure intended for cooperative mode with the AVs. Section IV illustrates the implementation of STREET on a particular scene with four different events and validates the methodology using ground truth and heatmaps. Thereafter, we conclude the paper in Section V with a summary of findings and future work.

II. RELATED WORK

Risk assessment is extensively used in automotive development as part of ISO 26262 [12]. With the development of higher SAE automation levels in vehicles, a new standard such as ISO/PAS 21448 [3], also known as the Safety Of The Intended Functions (SOTIF) [13], has been introduced to identify risks from improper vehicle functions and their operational limits. These standards are defined as regulations to develop risk assessments for AV validation during the development lifecycle [13], [14]. Ensuring the operational AV's safety levels requires the consideration of all potential hazardous (or risk-related) events and devising a corresponding safety goal that allows the AV to be validated during the development lifecycle [15], [16]. This existing process-based risk assessment method uses a qualitative approach based on experts' opinions, and a database of lessons learnt [4], [7], [17]–[19]. However, there is a need for more quantitative risk identification, as explained in [20], [21].

Quantitative approaches include using AV trial operations to scan and gain operational data to learn and improve risk-related issues [21] or using critical real-world scenarios [5], [6], [22]–[24]. Concurrently, a growing interest in using an independent real-time risk assessment [7]–[11] or

monitoring device [25] is pursued to indicate real-time risk at the AV level. These risk assessments use a quantitative approach to assess an environmental scene for hazardous events. Existing methods of risk assessments include AV focusing on risk-related figures and attributions. These methods consist of process-driven [4], [17]–[19], probability [26]–[29], model-based techniques [30]–[32], artificial intelligence methodologies [33]–[35] and lastly cooperative mode based approaches [36]–[44].

The cooperative mode approaches comprises two distinct communication pathways: 1) dedicated short-range communication (DSRC) or 2) Cellular V2X (C-V2X) methodologies. Both communication methods include safety and non-safety vehicle applications in intelligent transportation systems (ITS) and Radio Access Technologies [45]. DSRC is used in the existing real-world context, but with the growing interest in mobile-based connectivity, the mass deployment of DSRC becomes a challenge. Moreover, the increasing interest in 5G deployment fuelled the progress for C-V2X as an ITS communication, which became prevalent compared to DSRC. In the recent development of C-V2X roadmap, real-time ITS environmental perception and control is also included [46]. This recent development adds to the interest of this paper to propose RSI-based, cooperative mode pedestrian RT for the enhancement of real-time risk assessment for AV known as STREET.

Existing risk identification/assessment methods in cooperative mode can be seen in [36]–[43]. At a micro level, [37] proposed a model and probability based approach that implements a human-centric risk assessment algorithm which mimics human driving behaviour in the AV operation (based on radar results and V2V cooperative mode in a DSRC setting). This approach depends on radar results and operational AV fleets, which differs from the proposed approach in this paper. In the case of [38], contextual risk-based decisions are given based on the vehicle perspective in terms of environmental conditions in a vehicular ad-hoc network (VANET). The approach in [41] applied a similar method to [38] and introduced a fuzzy risk-decision model limited to simulation-based analysis. The drawback of these approaches [38], [41] assumes a contextual collection of the environmental events between vehicles, and risk mitigation decisions will be triggered when the risk level exceeds a certain threshold. These approaches are dependent on the available connectivity and data exchanges between AV fleets on the road and with the context limited to lane, road, traffic, weather, speed and time. The concept of [42] uses approaches such as the threat field model to quantify uncertainty with nearby vehicles via V2V communication for a potential collision. In this study, only Monte Carlo sampling was conducted for the vehicle to vehicle communication.

Moving from micro to macro approaches, [36] proposes a model-based approach focusing on fusing all the sensors and radars information from surrounding AVs. This approach depends on V2V cooperative mode to make predictions for latitudinal and longitudinal aspects forming time to collision risk assessment. Similarly to [36], [39] first formulates risk values based on the augmented collision risk between each AV using V2V and then creates a global collision risk at the Infrastructure level. The key focuses of [36] and [39] remain on V2V data and the efficiency of AV cooperative perception.

In summary, the following approaches [36]–[39] [41], [42] perform different risk estimations without including pedestrians and other road users. Moreover, these methods are only applicable on the assumption that all vehicles are AVs.

In [40], Principal Component Analysis (PCA) method is used to describe the feature in the traffic data set and recombined into a set of linearly independent features. Thereafter a Back-Propagation Neural Network is employed to train the predictive model. Its main focus is to predict driving risk based on past traffic accidents. This approach is limited to the quantitative results of the past traffic accident without considering near-misses conditions and/or identifying the potential hazardous events.

Both [43] and [44] demonstrated cooperative perception between the infrastructure and the AV. In the joint perception process, minimally LiDAR is required within the RSI, and in some cases, cameras as well. These methods require low latency and high bandwidth processing to allow joint perception at the AV level. Although the deployment of such techniques will gain more accuracy in terms of precise perception, it is achieved at the expense of LiDAR integration cost. In addition, the use of LiDAR on all RSI might be costly for mass deployment.

This approach focuses on the potential collision of AV with pedestrian(s) using risk tagging methodology in a cooperative I2V environment. This differs from studies investigating and solving the interaction between pedestrians and AV in closer proximities, such as pedestrian behaviour [47] as well as a prediction for pedestrian crosswalk behaviour [48]. These scopes are beyond the intention of this paper. Similarly, the concept of using mobile applications to warn pedestrians using vehicle to the pedestrian is also investigated in [49] with time to collision to the AV as a basis for warning. However, using a mobile device might be a distraction for road users as part of the qualitative findings.

Therefore, this paper aims to identify potential collisions with pedestrian(s) by providing quantitative risking tagging and time to collision in cooperative mode (I2V) with a camera device and edge system within the RSI. The direct outcome of this work can be used as severity rating within the real-time ASIL determination [9] or as RT within the formulated Predictive Risk Number for AV real-time risk assessment [10]. The proposed approach can be used with existing dedicated short-range communication (DSRC) or V2X methodologies. In DSRC, the information from the RSI can be supported as ''Cooperative sensing'' as depicted in [50] or in the case of ''Safety application'' deployment of DSRC in Singapore [51]. In terms of deployment under the C-V2X setup, low latency is achieved under the third stage of C-V2X [46]. Thus RT and time to collision proposed in

this paper can be incorporated under the 5GCAR use case of UCC3: "Cooperative safety".

III. PROPOSED METHODOLOGY – STREET

The proposed methodology, STREET, provides a quantitative RT of the real-world environment. The outcome is used as a severity representation of the scene intended for the AV to assess the risk at that location. STREET is a probability-model based approach that operates on top of deep learning-based detection techniques such as YOLOv5 [52]. This methodology processes the video images captured by the infrastructure and then computes the RT in real-time. These RT data are then sent to the vehicles as lightweight time-series data, thus allowing the vehicles to use the RT to perform real-time risk assessment.

STREET (illustrated in Figure 1) can be deployed in infrastructure or RSI. It uses the infrastructure's camera Field Of View (FOV) to select the Area Of Interest (AOI) and divides them into different risk zonal areas. These risk zonal areas are then assigned an array of risk probability values (known as risk detection matrix) relative to a reference AV's position "X" in the AOI. The assumed AV that provides this relative position is known as the ego vehicle or "X" in this paper. Thereafter, by using deep learning techniques [53] (e.g., YOLOv3 [54], or YOLOV5 [52]), STREET is able to detect and track objects and then map them onto the risk zonal areas in the form of a matrix. Subsequently, a final risk tag figure and time to collision value are derived. We have defined three risk detection models to estimate RT figures of a scene for AV usage in a real-world operation.

This research only focuses on pedestrian(s) crossing in uncontrolled traffic scenarios, which poses a significant risk to the AV. The proposed methodology, STREET, which consists of a **five-step process**is illustrated in Figure 1. STREET is embedded within the edge system of the RSI. The outcome of STREET stored in the local database, will be sent to the AV for risk assessment evaluation via cooperative I2V connections such as 5G.

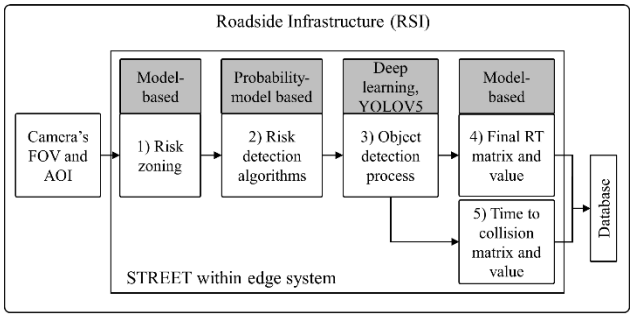

FIGURE 1. Flow of street methodology within the RSI.

A. RISK ZONING OF THE ENVIRONMENT

The risk zoning of the environment is considered the first step of STREET. This can be achieved in a fixed scene from the camera of the RSI.

Three selected scenes A, B and C, are used to demonstrate the process of risk zoning. These scenes are uncontrolled road areas with pre-identified locations. Therefore, the road scenes with fixed camera positions are considered static to the RSI with edge computing capability. Although Figures 2, 3, and 4 illustrate a left-hand driven scenario, the same approach can be applied to a right-hand driven scenario since the measurement of risk considers both directions of traffic. More will be explained in the risk zoning steps.

1) SCENE A (CORNER BEND)

Figure 2 illustrates the scenario when a vehicle or AV at location C is not able to sense the upcoming pedestrian(s) crossing the road before the turn as they are occluded by the building. In addition, the sensors onboard the vehicle or AV do not have sufficient range to sense the pedestrians after the bend. Therefore, the intention is to utilize the RSI to send the RT as a form of severity rating to the AV for its own risk assessment management.

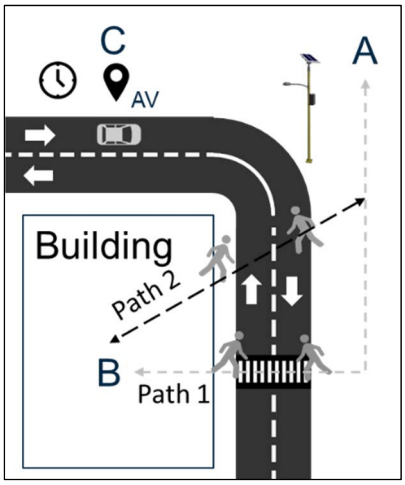

FIGURE 2. Use of roadside infrastructure in corner bend.

2) SCENE B (UNCONTROLLED T-JUNCTION)

Figure 3 illustrates the use of RSI to provide RT as an advanced warning when a pedestrian crosses either path 1 (A to C) or path 2 (B to C) and vice versa. Suppose the AV (at location D) intends to turn right, and the pedestrian crosses path 1 simultaneously, the AV may need to stop in the middle of the road, which incurs risk to the vehicles moving in the opposite direction. Thus, the advance RT warning will be helpful for the AV to detect pedestrian presence on path 1. In another situation, if the AV (at location E) intends to make a right turn, any pedestrian that dashes across the road at path 2 will cause the AV to apply sudden braking. Thus, advanced RT warnings from RSI can trigger the AV to make an informed decision, e.g. the AV will gradually reduce speed before turning.

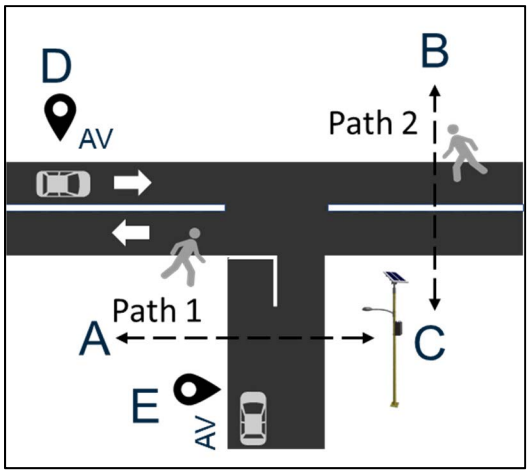

FIGURE 3. Use of roadside infrastructure in uncontrolled T-junction.

3) SCENE C (UNCONTROLLED DOUBLE T-JUNCTIONS)

As shown in Figure 4, which is similar to scene B but with an added inverted T- junction on the right. The AV at location F has the intention to make a right turn. Therefore, there is a need to provide an advance RT warning for path 2 (assuming that detection of path 3 is not an issue for the AV). In addition, the AV at location E has an intention to make a right turn and be exposed to both path 2 and path 3 in worst-case scenarios. Similarly, the advanced RT warning can be helpful in advanced navigation planning of the AVs' movements.

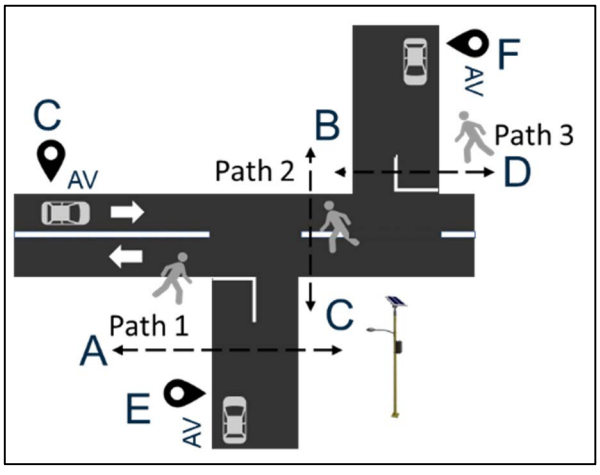

FIGURE 4. Use of roadside infrastructure in uncontrolled double T-junctions.

Based on the identified scenes and the location of the RSI camera, the risk zonal areas can be determined based on the following:

1. Select road section(s) that are exposed to high volume of pedestrian crossing using the Field of View (FOV) of the roadside infrastructure camera.
2. Define selected Area of Interest (AOI) based on the FOV.
3. From the direction of travel, indicate the ego vehicle location as "X" using the corresponding GPS location.
4. Measure or determine the total distance (D_{total}) of the AOI section either from physical measurements or indications from the source of the available map.
5. Determine the number of rows ($rows_{int}$) within the AOI based on (1) (rounded down to the nearest integer). t_{resp} is based on a driver's response time in the real-world traffic [55]. V_{max} is based on the defined speed limit on the selected AOI. V_f is a safety factor that considers if an elevated speed limit is required for the selected AOI.

$$rows_{int} = \frac{D_{total}}{(t_{resp} \cdot V_{max} \cdot V_f)} \tag{1}$$

6. Determine different risk zonal areas based on the scene of the AOI (Figure 5, 6 and 7 illustrates the resulting risk zonal areas of scenes A, B and C, respectively). In STREET, roads are classified as RED zone, while YELLOW zone denotes pavements and GREEN zone is off-road regions. Risk zonal areas can be reduced if they do not exist. For example, if off-road regions do not exist, the GREEN risk zone can be removed. Likewise, if there is no pavement within the AOI, the YELLOW risk zone can be removed. Thus, symmetry is not necessary when defining the risk zonal areas on both sides other than the road, e.g., as shown in Figure 5, where the Green zone only exists on the right side of the road. The direction of travel will determine the starting row, referenced to the ego vehicle marked "X".
   In each scene, there are two-way lanes on the road represented by the RED zone of the risk zonal areas (it covers both left-hand or right-hand driving lane). This is established by using the same AOI for different directions of travel, as indicated by the "white arrow" from the ego vehicle. For example, in Figure 5, scene A, AOI direction A can represent either a left-hand or a right-hand driving configuration so long as the vehicle is moving from top to bottom. At the same time, direction B indicates vehicles driving from the bottom to the top of the same scene A. Therefore, the risk zonal areas can cover both configurations in terms of left or right-hand drive. The intention of this approach is to measure the risk of pedestrians moving in the presence of the ego vehicle at reference "X".
7. Create a 2D matrix (ω_{rz}) with a symbolic representation of the risk zonal areas. Each symbolic representation of the sub-regions has assigned weights to the different regions according to risk zones. The weights will form a basic 2-D matrix (ω_{rz}) as shown in Figure 8 as an example for scene A. The range of weights depends on the selected scene definition in Table 1. Other than the weights, the distance D_x and interval time t_x are distributed uniformly according to (2) and (3).

$$D_{total} = \sum_{x=1}^{rows} D_x \tag{2}$$

$$t_{total} = \frac{D_{total}}{V_{max}} = \sum_{x=1}^{rows} t_x \tag{3}$$

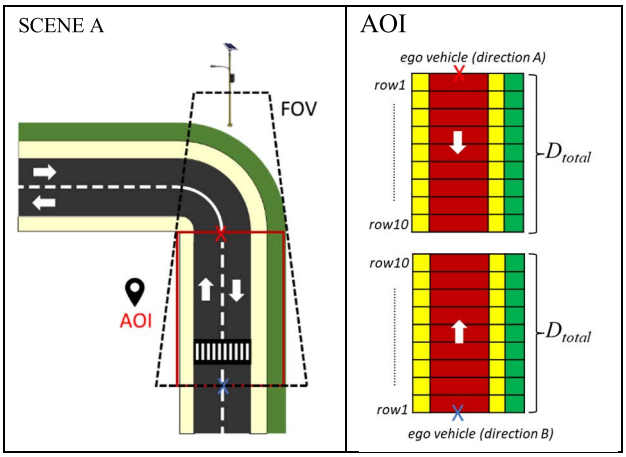

FIGURE 5. Illustration of zoning concept from roadside infrastructure – scene A.

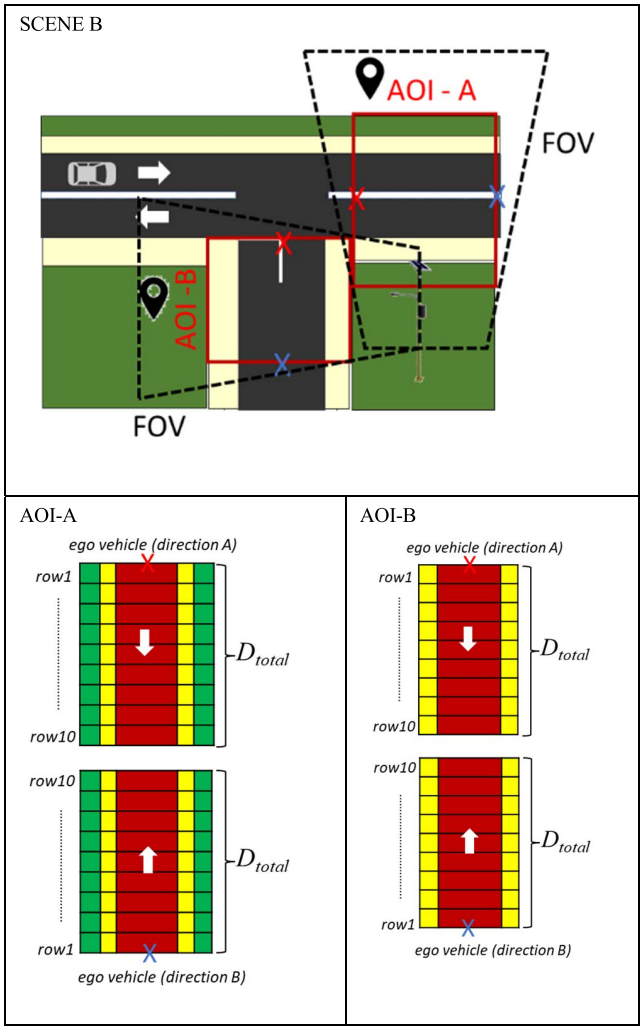

FIGURE 6. Illustration of zoning concept from roadside infrastructure – scene B.

In real-world settings, the selected distance (D_{total}) from each scene ranges from 50-200 meters, depending on the elevation of the camera from the roadside infrastructure.

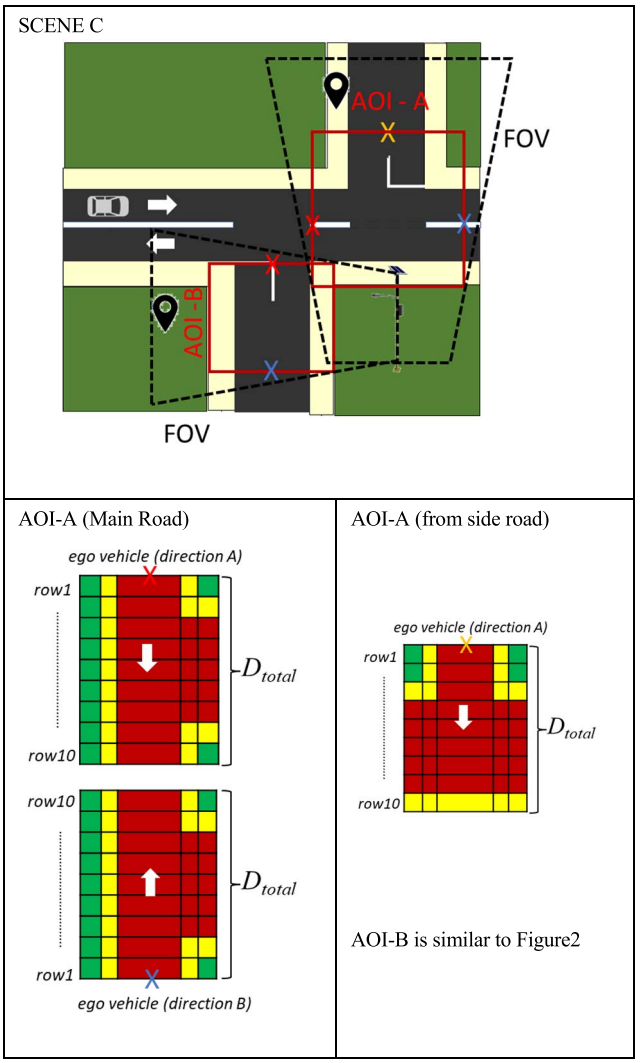

FIGURE 7. Illustration of zoning concept from roadside infrastructure – scene C.

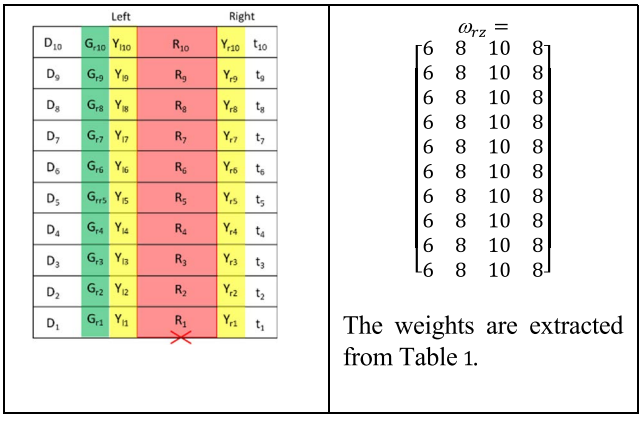

FIGURE 8. Converting AOI sub regions into a 2D matrix (ω_{rz}) with weights – scene A.

B. RISK DETECTION MATRIX

Once the risk zoning of the environment is completed, the probability of collision in each sub-region of the risk zonal

TABLE 1. Weights rating for different sub regions.

Region	Rating	Scene definition
RED	10	Dual lane main road conditions
RED	9	Dual lane small road conditions
YELLOW	8	Dedicated Pavement, shared path or parking lots
YELLOW	7	Single lane pedestrian Pavement
GREEN	6	Accessible pathways away from pavement
GREEN	5	Thin accessible pathways away from the pavement
GREEN	4	Green grass patch away from the pavement
GREEN	3	Tall green grass patch away from the pavement
GREEN	2	Barricaded pathways
GREEN	1	Non-accessible paths with short walls
GREEN	0	Non-accessible paths

areas is computed and represented as a Risk Detection (RD) matrix. This RD matrix is part of the STREET, which uses probability-based modelling consisting of two novel aspects: Zonal Probability Distribution (ZPD) and Inter-zonal Probability Distribution (IPD). ZPD represents the time-stamped layers in Table 2 (e.g. $R_1, R_2 \ldots R_n$, which are time-stamped to $t_1, t_2 \ldots t_n$, where n is the number of rows), while IPD represents the inter-risk zonal areas in Table 2 (e.g. G_{L2}, Y_{L2}, R_2, Y_{R2} and G_{R2} across row 2). All probability figures range from 0 to 1.0, where 1.0 represents a definite collision between the pedestrian and the ego vehicle. We propose three different probability-based models for ZPD: 1) Linear Intervals Distribution (LID), 2) Conservative Exponential Distribution (CED), and 3) Aggressive Exponential Distribution (AED) to react to different events depending on the AV's risk and safety management strategy. The three models provide different focuses on the ZPD layering. LID provides a stepwise incremental RT to determine the movement of the detected pedestrian across the different layers within the ZPD. CED focuses on detecting hazardous events happening on the front and middle layers, while AED focuses mainly on the front layer of the ZPD (closest to "X").

1) LINEAR INTERVALS DISTRIBUTION MODEL (LID)

The probability of collision between a pedestrian (C) and the ego vehicle "X" can be determined as P(C|X). This represents the likelihood of a collision at R_1 (as shown in Figure 8) and thus assigned a probability of 1.0 (4). The probability interval within the defined risk zone (e.g., RED risk zone) can be linearly divided into single layer interval Z_Δ using (6) and eventually finding (5) for each ZPD layer. $R_{n_{max}}$ (e.g. R_{10}) can be the minimum probability figure set at the furthest ZPD layer. The maximum number of rows needed for the scene can be obtained using (1) from the AOI. Using the known $rows_{int}$, the interval (Z_Δ) for each layer can be calculated using (6), and the corresponding risk probability for each derived ZPD layer can be obtained using (5). This approach applies to all other risk zonal areas as well. Therefore, the risk probability is linearly distributed for all layers of the ZPD and effectively used for AV where the stepwise RT can be used to track objects moving across different layers of the ZPD.

$$R_1 = P(C|X) \tag{4}$$

$$R_{row} = R_1 - (row - 1)\, Z_\Delta, \quad row \geq 1 \tag{5}$$

$$Z_\Delta = \frac{R_1 - R_{n_{max}}}{rows_{int} - 1} \tag{6}$$

Using scene A as an example, the total distance of the AOI, $D_{total} = 100$ m, $V_{max} = 40$km/hr and $t_{response} = 0.85$ are used to determine $rows_{int} = 10$ using (1). As defined in (4), the probability of collision at R_1 is 1.0 and hence for the further ZPD layer, $R_{n_{max}}$ the probability of collision is 0.1 and Z_Δ is 0.1. The final RD matrix for Scene A is shown in Table 2.

TABLE 2. LID RD matrix for scene A.

$$RD_{LID} = \begin{bmatrix} G_{l10} & Y_{l10} & R_{10} & Y_{r10} \\ G_{l9} & Y_{l9} & R_9 & Y_{r9} \\ G_{l8} & Y_{l8} & R_8 & Y_{r8} \\ G_{l7} & Y_{l7} & R_7 & Y_{r7} \\ G_{l6} & Y_{l6} & R_6 & Y_{r6} \\ G_{l5} & Y_{l5} & R_5 & Y_{r5} \\ G_{l4} & Y_{l4} & R_4 & Y_{r4} \\ G_{l3} & Y_{l3} & R_3 & Y_{r3} \\ G_{l2} & Y_{l2} & R_2 & Y_{r2} \\ G_{l1} & Y_{l1} & R_1 & Y_{r1} \end{bmatrix}$$

$$RD_{LID} = \begin{bmatrix} 0 & 0 & 0.1 & 0 \\ 0 & 0.1 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.3 & 0.2 \\ 0.2 & 0.3 & 0.4 & 0.3 \\ 0.3 & 0.4 & 0.5 & 0.4 \\ 0.4 & 0.5 & 0.6 & 0.5 \\ 0.5 & 0.6 & 0.7 & 0.6 \\ 0.6 & 0.7 & 0.8 & 0.7 \\ 0.7 & 0.8 & 0.9 & 0.8 \\ 0.8 & 0.9 & 1.0 & 0.9 \end{bmatrix}$$

2) CONSERVATIVE EXPONENTIAL DISTRIBUTION MODEL (CED)

In order to provide different RT figures for different events depending on the AV risk and safety management strategy, another model known as CED is proposed. This model represents an almost consistent high-risk probability for a period of time (window) focusing on the front and middle layers. This is necessary to gain more response time for the AV to focus on RT in the front and middle layers. This response time can be modelled by an exponential decay of the risk probability as shown in (7). Thus, the risk probability should remain relatively high until a time window has passed. The maximum layer for a particular risk zone row_{int} is used in (7).

The CED reduces the risk drastically when the ZPD layers reach the furthest, thus reflecting the focus on the front and middle layers of the ZPD. This model is suitable for scenarios where the middle and front layers of the AOI have a more extensive road coverage (e.g. in the case of scene C, AOI-A). This high probability distribution risk level will capture the AV's attention with higher-RT compared to LID, thus triggering the AV navigation and operations in a conservative manner. Similarly, (4) is used for determining R_1. Therefore, the variable *row* is assumed to be greater or equal to 2 in (7).

$$R_{row} = R_1 - e^{(row - row_{int})}, \quad (row \geq 2) \tag{7}$$

The CED RD matrix is shown in Table 3 using scene A as an example where $\text{rows}_{\text{int}} = 10$ and $R_1 = 1.0$, the values of R_{row} can be determined accordingly.

TABLE 3. CED RD matrix for scene A.

$$RD_{CED} = \begin{bmatrix} 0 & 0 & 0.0 & 0 \\ Y_{l9} - Z_\Delta & R_9 - Z_\Delta & 1 - e^{-1} & R_9 - Z_\Delta \\ Y_{l8} - Z_\Delta & R_8 - Z_\Delta & 1 - e^{-2} & R_8 - Z_\Delta \\ Y_{l7} - Z_\Delta & R_7 - Z_\Delta & 1 - e^{-3} & R_7 - Z_\Delta \\ Y_{l6} - Z_\Delta & R_6 - Z_\Delta & 1 - e^{-4} & R_6 - Z_\Delta \\ Y_{l5} - Z_\Delta & R_5 - Z_\Delta & 1 - e^{-5} & R_5 - Z_\Delta \\ Y_{l4} - Z_\Delta & R_4 - Z_\Delta & 1 - e^{-6} & R_4 - Z_\Delta \\ Y_{l3} - Z_\Delta & R_3 - Z_\Delta & 1 - e^{-7} & R_3 - Z_\Delta \\ Y_{l2} - Z_\Delta & R_2 - Z_\Delta & 1 - e^{-8} & R_2 - Z_\Delta \\ Y_{l1} - Z_\Delta & R_1 - Z_\Delta & 1.0 & R_1 - Z_\Delta \end{bmatrix}$$

$$RD_{CED} = \begin{bmatrix} 0 & 0 & 0.0 & 0 \\ 0.43 & 0.53 & 0.63 & 0.53 \\ 0.66 & 0.76 & 0.86 & 0.76 \\ 0.75 & 0.85 & 0.95 & 0.85 \\ 0.78 & 0.88 & 0.98 & 0.88 \\ 0.79 & 0.89 & 0.99 & 0.89 \\ 0.8 & 0.9 & 1.0 & 0.9 \\ 0.8 & 0.9 & 1.0 & 0.9 \\ 0.8 & 0.9 & 1.0 & 0.9 \\ 0.8 & 0.9 & 1.0 & 0.9 \end{bmatrix}$$

Figures are round to 2 decimal places

3) AGGRESSIVE EXPONENTIAL DISTRIBUTION MODEL (AED)

The last model representing the probability of collision with the ego vehicle is shown in (8). This approach is orthogonal to CED, where the decay of collision probability becomes much faster than LID and CED. One example of this usage is shown in scene B, AOI-B where the AV focuses on pedestrians crossing the front rows. With pedestrians crossing detected in the front rows, it will affect the decision of the AV to delay its intention to turn right. This model also assumes that the AV is only sensitive to front rows. Therefore, the AV is expected to respond fast to the given RT since the RT only appears for an extremely short duration with only front row detections. In this AED model, the variable λ in (8) determines the aggressiveness level of the decay. For illustration purposes, the default value of λ is 6 indicating the sensitivity on the first three rows for RT.

$$R_{\text{row}} = e^{\left(\frac{-\lambda\,(row-1)}{row_{int}}\right)}, \quad row \geq 1 \tag{8}$$

TABLE 4. AED RD matrix for scene A.

$$RD_{CED} = \begin{bmatrix} 0 & 0 & e^{-5.4} & 0 \\ Y_{l9} - Z_\Delta & R_9 - Z_\Delta & e^{-4.8} & R_9 - Z_\Delta \\ Y_{l8} - Z_\Delta & R_8 - Z_\Delta & e^{-4.2} & R_8 - Z_\Delta \\ Y_{l7} - Z_\Delta & R_7 - Z_\Delta & e^{-3.6} & R_7 - Z_\Delta \\ Y_{l6} - Z_\Delta & R_6 - Z_\Delta & e^{-3.0} & R_6 - Z_\Delta \\ Y_{l5} - Z_\Delta & R_5 - Z_\Delta & e^{-2.4} & R_5 - Z_\Delta \\ Y_{l4} - Z_\Delta & R_4 - Z_\Delta & e^{-1.8} & R_4 - Z_\Delta \\ Y_{l3} - Z_\Delta & R_3 - Z_\Delta & e^{-1.2} & R_3 - Z_\Delta \\ Y_{l2} - Z_\Delta & R_2 - Z_\Delta & e^{-0.6} & R_2 - Z_\Delta \\ Y_{l1} - Z_\Delta & R_1 - Z_\Delta & 1.0 & R_1 - Z_\Delta \end{bmatrix}$$

$$RD_{CED} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0.01 & 0 \\ 0 & 0 & 0.02 & 0 \\ 0 & 0 & 0.03 & 0 \\ 0 & 0 & 0.05 & 0 \\ 0 & 0 & 0.09 & 0 \\ 0 & 0.07 & 0.17 & 0.07 \\ 0.10 & 0.20 & 0.30 & 0.20 \\ 0.35 & 0.45 & 0.55 & 0.45 \\ 0.80 & 0.90 & 1.00 & 0.90 \end{bmatrix}$$

Figures are round to 2 decimal places

The AED RD matrix is shown in Table 4 using scene A. Using $\text{rows}_{\text{int}} = 10$ and $R_1 = 1.0$, the values of R_{row} can be determined accordingly. The remaining regions are determined using the same method as described in LID.

4) INTERZONAL PROBABILITY DISTRIBUTION (IPD)

IPD represents the decay in collision probability as the pedestrian moves from one risk zone to another (e.g. RED to YELLOW risk zone). The pedestrian is assumed to be walking at a constant speed. The linear distribution of the risk probability is thus applied across the different risk zonal areas. The amount of linear drop across different regions can be formulated for each region with (9) to (12). One example is to use a similar model like LID, where layer interval Z_Δ can be represented as Z_D for the calculation of the rest of the YELLOW and GREEN regions shown in (9) to (12). In addition, α can be used to control larger step size to differentiate the probability distribution crossing from one region to another.

With the settings of IPD, the probability distribution across the regions (row-wise) will not have the same values.

$$Y_{\text{Ln}} = R_n - aZ_D \tag{9}$$

$$Y_{\text{Rn}} = R_n - \alpha Z_D \tag{10}$$

$$G_{\text{Ln}} = Y_{\text{Ln}} - \alpha Z_D \tag{11}$$

$$G_{\text{Rn}} = Y_{\text{Ln}} - \alpha Z_D \tag{12}$$

C. OBJECT DETECTION (OD) MATRIX

The OD matrix represents the detected objects within the infrastructure's camera AOI. The boundary boxes represent the detected objects using YOLO based deep learning techniques from the camera's real-time video. If the edges of the boundary boxes (bottom edge) overlap with a particular risk zonal area, that risk zonal area will have a weighted representation. This method can account for occluded pedestrians with low overlaps by taking the bottom edge as a reference. However, similar to all 2D computer-vision techniques, accuracy decreases if the overlap of two objects is high.

Multiple objects can be detected in the existing capability of object detections and reside in different or the same risk zonal area(s). Therefore, the OD matrix can have more than one non-zero value. Multiple boundary boxes can be mapped to one zonal risk area, but a single boundary box is limited to only one risk zonal area in a directional many-to-one mapping relationship.

Table 5 illustrates two examples of the OD matrix for scene A (one pedestrian is detected in R_2 and R_3 within example 1, and three pedestrians are detected in Y_{r9} and two pedestrians in R_6 within example 2). The OD matrix will indicate the corresponding detected pedestrian(s) within each sub-region. If no pedestrian detected, the sub-region will be "0".

D. FINAL RT (RT) MATRIX AND FIGURE

The final RT matrix and figure can be computed based on the defined ω_{rz}, RD_{rz} and OD_{rz} matrices. Starting with an element-wise multiplication (also known as Hadamard product) of ω_{rz} and OD_{rz} matrices to obtain OOD_{rz} in (13). Secondly, perform another element-wise multiplication of RD_{rz} *and* OOD_{rz} and lastly, have a summation of all the elements

TABLE 5. OD matrix for scene A.

$$OD_{example\ 1} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & R_3 = 1 & 0 \\ 0 & 0 & R_2 = 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad OD_{example\ 2} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Y_{r9} = 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & R_6 = 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

with the RT matrix to determine the final RT (14).

$$OOD_{rz} = (\omega \circ OD)_{rz} \tag{13}$$

$$RT_{final} = \sum\nolimits_{zn} (RD \circ OOD)_{rz} \tag{14}$$

E. PEDESTRIAN OCCURRENCE HEATMAP MATRIX

The pedestrian occurrence is a post-processing activity used to validate the overall RT over time with the RT results. The pedestrian occurrence can be represented as a heatmap to further analyze its suitability towards LID, CED or AED models. Heatmap refers to the accumulative pedestrian occurrence in each of the sub-regions using OD matrix (OD_{rz}). Frame intervals ($t_{intervals}$) can be set for n number of samples to measure heatmap after each pre-defined duration. The pedestrian occurrence heatmap ($Heatmap_{rz}$) is defined as (15). With the pedestrian occurrence data and the total pedestrian detected, the probability of pedestrian P(occ) occurring in each sub-region can be calculated as (16). Since LID, CED and AED focus on the front, middle and rear layers differently, the heatmap can be separated into the front P(Fr), middle P(Mid) and rear P(Re) probability with separation of the defined rows into front, middle and rear rows. The accumulated P(Fr), P(Mid) and P(Re) will assist in comparison with the RT results over time for a particular scene and decide which risk matrix (LID, CED or AED) is more suitable.

$$Heatmap_{rz} = \sum\nolimits_{N=1}^{t_{interval}} OD_{rzN} \tag{15}$$

$$P(Occ)_{rz} = Heatmap_{rz} \Big/ \sum\nolimits_{r=1}^{rows} \sum\nolimits_{z=1}^{zone} Heatmap_{rz} \tag{16}$$

$$P(Fr) = \sum\nolimits_{r=1}^{front} \sum\nolimits_{z=1}^{zone} P(Occ)_{rz}, \tag{17}$$

$$P(Mid) = \sum\nolimits_{r=1}^{middle} \sum\nolimits_{z=1}^{zone} P(Occ)_{rz} \tag{18}$$

$$P(Re) = \sum\nolimits_{r=1}^{rear} \sum\nolimits_{z=1}^{zone} P(Occ)_{rz} \tag{19}$$

F. TIME TO COLLISION (TTC) MATRIX

In addition to RT computed as part of STREET, it is important to provide a TTC matrix and value based on the minimal time to collision with respect to "X" in the scene. The TTC can then be sent to all AV regardless of their position on the roads. The following steps can determine TTC matrix and value.

1. Obtain a normalized OD matrix known as OD_{binary} (turning the values into binary values only) shown as Table 6 (using Table 5 Scene A example 2 as reference).

TABLE 6. Determine OD_{binary}.

$$OD_{binary} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Y_{r9} = 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & R_6 = 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

2. Obtain the $time_{rz}$ matrix, by using $t_{response}$ and multiply with each row number with the results shown in Table 7.

TABLE 7. Determine $time_{rz}$.

$$time_{rz} = \begin{bmatrix} 0 & 9.0 & 9.0 & 9.0 \\ 0 & 8.1 & 8.1 & 8.1 \\ 0 & 7.2 & 7.2 & 7.2 \\ 0 & 6.3 & 6.3 & 6.3 \\ 0 & 5.4 & 5.4 & 5.4 \\ 0 & 4.5 & 4.5 & 4.5 \\ 0 & 3.6 & 3.6 & 3.6 \\ 0 & 2.7 & 2.7 & 2.7 \\ 0 & 1.8 & 1.8 & 1.8 \\ 0 & 0.9 & 0.9 & 0.9 \end{bmatrix}$$

For example in scene A, $row_{int} = 10$, therefore the $t_{response} = 0.9$ s. The TTC should only be considered for the road and pavement (RED and YELLOW region).

3. Calculate TTC_{rz} by performing a Hadamard product with the OD_{binary} and $time_{rz}$ matrix shown in (20) and then determine the lowest non-zero value as TTC_{value} where $TTC_{value} \in TTC_{rz}$ and $\neq 0$ in (21). The TTC_{value} in this case for scene A can be 5.4 s from "X" based on the given example.

$$TTC_{rz} = (OD_{binary} \circ time_{rz}) \tag{20}$$

$$TTC_{value} = \min(TTC_{rz}) \tag{21}$$

TTC_{value} will provide the estimated time to collision from the ego vehicle with GPS position. This value is sent to the AV in addition to the RT figure. Therefore, other than providing a severity rating for the AV risk assessment, the AV can use the TTC_{value} to calculate an overall time to collision based on its current location through its navigation components as shown in (22) and (23). $D_{(AV-RT)}$ is the distance from the current AV position to "X", while V_{AV} is the current speed of the AV. Therefore, the estimated time to collision for the AV can be calculated using (23) as $TTC_{overall}$.

$$t_{RT} = \frac{D_{(AV-RT)}}{V_{AV}} \tag{22}$$

$$TTC_{overall} = t_{RT} + TTC_{value} \tag{23}$$

IV. IMPLEMENTATION AND RESULTS

The STREET was implemented and tested using scene A. The RSI was set up with the FOV as shown in Figure 9. The RSI is positioned at the corner bend of scene A with a GPS location of 1.300133°, 103.781527°. Figure 9 illustrates the camera's FOV looking at pedestrians crossing after the bend.

FIGURE 9. Real-world scene A, camera's field of view.

A. ROADSIDE INFRASTRUCTURE SYSTEM

The RSI consists of a camera pointing at the FOV. The camera footage is then given to the edge computer (workstation) to perform RT computation. The workstation has a specification of Intel I7 and a GPU of Nvidia 1080 graphics card. The object detection rate occurs between 18-22 milliseconds per frame with YOLOv5. Ensuring that the video frame rate interval is longer than the time needed for object detection, 25 frames per second (fps) or 40 milliseconds per frame is selected for performing RT. The camera is elevated at a height close to 5.5 meters. The connectivity of the RSI can be connected via fibre or cellular connectivity. 5G connections are recommended to lower the latency incurred across the transmission.

B. DATABASE SYSTEM WITH LARGE ENSEMBLE DATA

The overall RT data format includes the date, time, location, RT figure, persons detected, normalized RT, time to collision (TTC_{value}), a total distance of AOI and speed limit for the AOI. A slice of this ensemble time-series data is shown in Table 8. This data can reside firstly within the RSI and be transmitted to the AV. This information ensemble can be backed up to the cloud for future frequentist statistics [56] (with extensive data over time) using RT figures occurrence as a quantitative measure. A local database can store critical data for fast processing or reduce latency together with 5G connectivity. The overall cloud database or server will store and consolidate all the infrastructure information at a regional or country level and broadcast the information to each AV. Normalized RT will be further discussed in section IV-C.

C. STREET IMPLEMENTATION RESULTS

This section explains the five-step process in implementing the STREET onto real-world scene A.

TABLE 8. Illustration of an overall ensemble of RT (LID) format. (* Based on ANSI date/time data types.)

Date* (yyymmdd)	Time* (hh:mm:ss:ffff)	Latitude	Longitude	RT	Detected persons
20200302	11:12:48.0000	1.300056	103.764944	7.0	1
20200302	11:12:48.0400	1.300056	103.764944	12.6	2
20200302	11:12:48.0800	1.300056	103.764944	12.6	2
20200302	11:12:48.1200	1.300056	103.764944	12.6	2
20200302	11:12:48.1600	1.300056	103.764944	7.0	1
Date* (yyymmdd)	Time* (hh:mm:ss:ffff)	Normalized RT	TTC_{value} (s)	D_{total} (m)	V_{max} (km/hr)
20200302	11:12:48.0000	7.0	2.7	100	40
20200302	11:12:48.0400	6.3	2.7	100	40
20200302	11:12:48.0800	6.3	2.7	100	40
20200302	11:12:48.1200	6.3	2.7	100	40
20200302	11:12:48.1600	7.0	2.7	100	40

1) ZONING OF ENVIRONMENT

The AOI with rows and risk zonal areas shown in Figure 10 are based on the FOV of Figure 9. Table 9 shows the parameters determine from the environment and using (1) to determine row_{int} and using Table 1 to determine ω_{rz}. The road is classified as a "RED" zone, while the pavement on the right is classified as a "YELLOW" zone, and lastly, the additional roadside parking is partially classified as a "YELLOW" zone with the remainder as off-road with "GREEN" zone.

TABLE 9. Parameters determined for scene A.

Name	Parameters	Values	Units
Total distance of AOI	D_{total}	50	Meters
Speed limit	V_{max}	20 (5.55)	km/hr (m/s)
Safety factor	S_f	1.0	-
Driver response time	$t_{response}$	0.85	seconds
No. of rows for the AOI	row_{int}	10	Round down

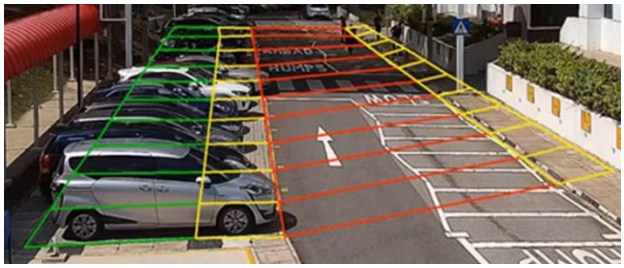

FIGURE 10. Real-world scene A, Area of Interest.

2) RISK DETECTION (RD) MATRIX

The default RD matrix for scene A is LID (Table 2) model, and for comparative analysis, parallel computation of CED (Table 3) and AED (Table 4) models are determined as well. The IPD is set to linear distribution similar to LID model.

3) OBJECT DETECTION (OD) MATRIX

In this methodology, YOLOv5s model was used and tested for scene A. Direct inferencing was performed on scene A with pre-trained models using Cocodatasets. The inferencing was done using 97k samples with 40 ms interval. This was compared with annotated 97k samples as ground truth. Table 10 shows the YOLOv5s accuracy in pedestrian detection in scene A.

TABLE 10. Accuracy of YOLOv5s inferencing with scene A.

Detection	True Positive	True Negative	False Positive	False Negative	Accuracy
YOLOv5s	38.07%	30.28%	16.88%	14.75%	68.35%

Since the purpose is for RT, false positives are also included to ensure that the operating AV is cautious in its operation, thus increasing the accuracy to 85.23%. The RT is performed at every 40 milliseconds per video image frame.

4) FINAL RT MATRIX AND VALUE

A total of 97k samples were risk tagged in sequence. STREET converted the video image information from the AOI into time-series RT figures. For comparative analysis, the RT overview for LID, CED and AED is presented in Figure 11. CED can be seen with the highest RT figure (0 - 60) since the emphasis is a high-risk probability distribution in the front and middle layers, which provides an almost consistent high RT. LID with RT figure can be seen next, ranging from (0 – 40). Generically the purpose of LID is to provide a linear increase or decrease in step when the pedestrian crosses between different sub-regions. Therefore, LID will have a lower overall RT and will have more granular figures for the last row than CED. On the other hand, AED has the lowest RT range (0 – 20). This is true since AED is modelled to detect pedestrians in the first three rows (with λ is set to 6). Therefore, AED gives a sudden surge in RT when front rows detect pedestrian crossings in scene A.

The diverse RT range for LID, CED, and AED makes it difficult to use it as a severity rating for risk assessment. Therefore, a better measurement would be to obtain a normalized RT by dividing the original RT figure by the overall detected pedestrian for each frame. The normalized RT figures are shown in Figure 12, ranging from (0 -10) illustrated as AVG_LID, AVG_CED and AVG_AED.

From the tested 97k samples, the average values of LID, CED, and AED (represented as A_LID, A_CED and A_AED) prove the intended operation of the STREET provides a different representation of scene A. The CED is expected to have the highest average of 4.16 followed by an average LID of 1.86 and finally an average AED of 0.38. The overall variance of the samples is determined with CED 13.224, LID 4.043 and AED 1.532. These readings indicate that scene A has a higher pedestrian occurrence in the middle layers than the front layers (especially with the large variance observed in CED), while the occurrence of the front layer

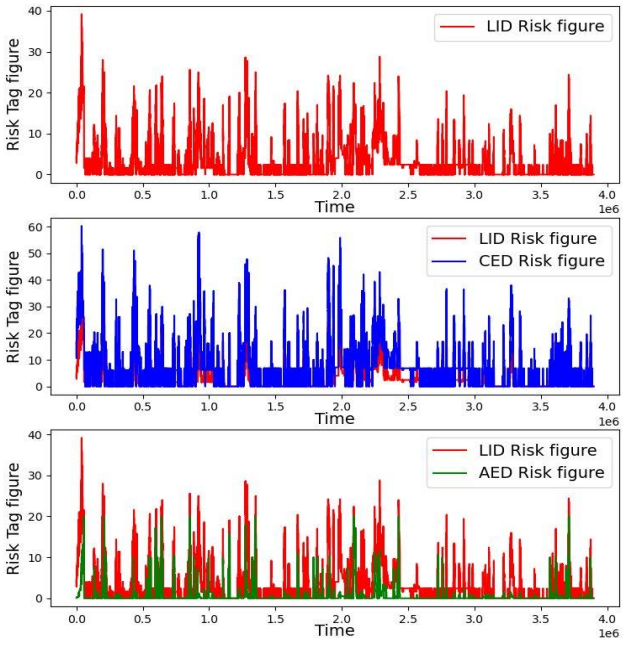

FIGURE 11. Comparison of LID, CED and AED RT figures.

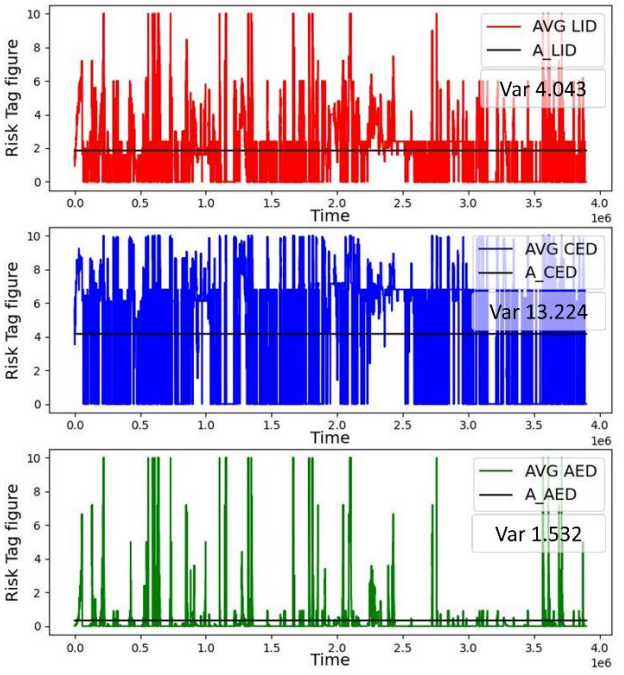

FIGURE 12. Normalized RT figures for LID, CED and AED.

happens to be the least (with the lowest mean and variance). This can be further validated by the heatmap of the overall pedestrian occurrence of scene A shown in Figure 13. This heatmap is obtained from the occurrence of the pedestrian in each sub-region over the tested samples. The probability is then obtained from the occurrence in each sub-region divided by the total occurrence.

The total pedestrian occurrence is 99.4k and with region Y_{L7} being the most populated. The corresponding $Heatmap_{rz}$ and $P(Occ)_{rz}$ is obtained with the P(Fr), P(Mid) and P(Re) in Figure 14. Since the $rows_{int}$ is 10, the rows 1-3 are front, rows 4-7 are middle and rows 8-10 are rear.

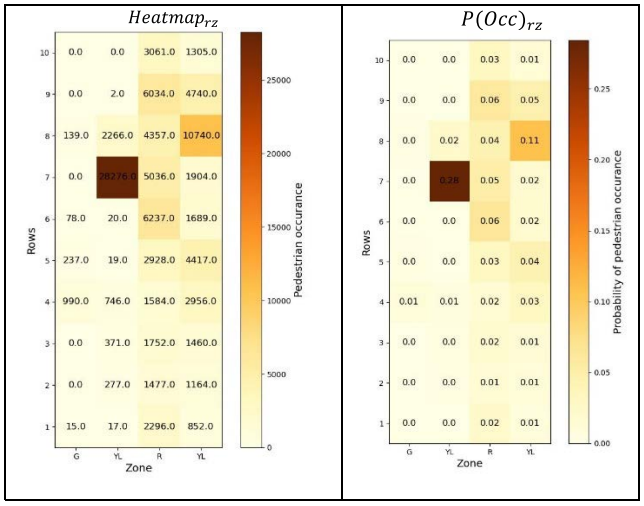

FIGURE 13. Heatmap of overall pedestrian occurrence.

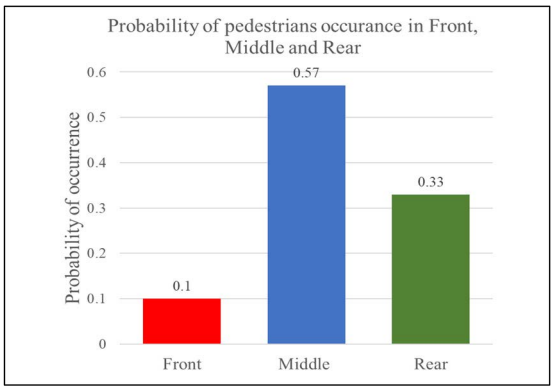

FIGURE 14. Probability of pedestrian within scene A in terms of front, middle and rear rows.

This information shows that middle rows have the most pedestrian occurrence, followed by rear and front rows, which corroborates with the results obtained from the RT in terms of average figures. The AED values are generally low throughout the datasets with low front pedestrian occurrences. While CED is sensitive towards middle and front layers, it is apparent that the middle layers drive the CED with high RT figures. The overall LID RT figures remain lower than CED figures, implying that the rear layers pedestrian occurrences are present, but the impact is not as high as the middle layers.

5) TIME TO COLLISION MATRIX

Besides RT, TTC_{value} is given to the AV to estimate time to collision towards the identified pedestrians. The TTC_{value} is determined at intervals of 40 milliseconds per frame using (20) and (21). Therefore, the overall TTC_{value} for the 97k dataset can be seen in Figure 15. This data aims to illustrate the importance of the proximity of pedestrians towards the ego vehicle in addition to the RT figures.

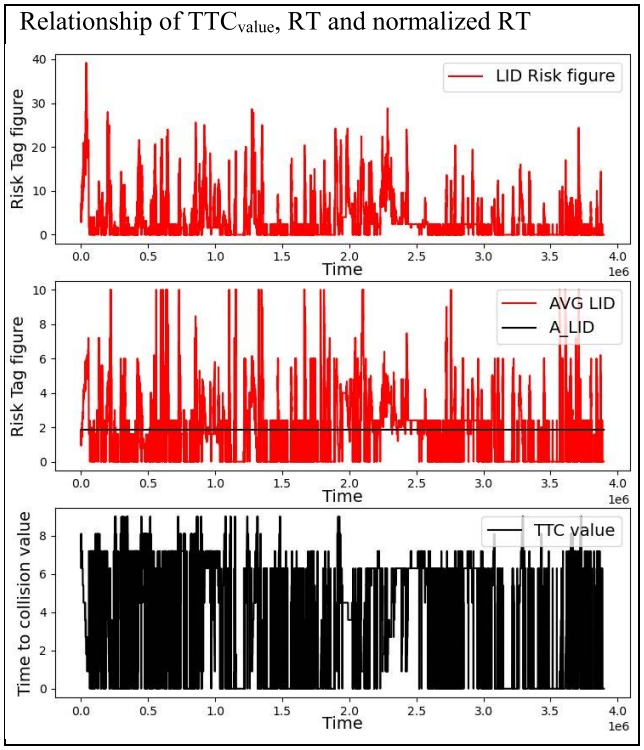

FIGURE 15. Overview of LID RT, normalized RT and TTC value.

The TTC_{value} provides an indication in seconds based on the speed limit of the road and the distance between the reference ego vehicle to the identified pedestrian present in the sub-region. The TTC_{value} is taken as the minimum value from the TTC_{rz} matrix, which relates to the pedestrian with the nearest impact to the ego vehicle. The results of the real-time TTC_{value} is inversely related to the RT figures. As the RT figures increases, the TTC_{value} decreases. Due to the AOI's finite distance per sub-region, the TTC_{value} is determined in sub-second blocks. Examples of detailed TTC_{value} will be discussed in section D.

Further explanation t_{RT} (22) relates to the time the AV would arrive from its current location to "X" of the AOI. This time t_{RT} is determined by obtaining the distance from the AV's current location to the ego vehicle, divided by the speed of the AV. The overall $TTC_{overall}$ (23) is based on the calculated t_{RT} from the AV and TTC_{value}. TTC_{value} are sent together with the rest of the data as part of a dataset shown previously in Table 8. AV will access the database for advanced navigation and routing decisions to perform real-time risk assessments.

D. STREET IMPLEMENTATION EVENTS

Besides validating the overall RT figures, the RT representation of a scene or AOI can also help to identify events

that occurred. LID, CED, and AED are also compared for RT efficiency in these situations. LID will generally be the best default algorithm for these event detections. Pedestrian detections with boxes shown outside the AOI are meant for illustration purposes. The RT does not include those detected pedestrians as part of the RT evaluation. These events can be classified into the following scenarios:

1) PEDESTRIAN MOVEMENT FROM BOTH SIDES AT MIDDLE ROWS

Pedestrian crossing on both sides of the road (images shown in Figure 16) can be detected with the proportional rising and falling, with RT figures varying across zonal risk regions (shown in Figure 17).

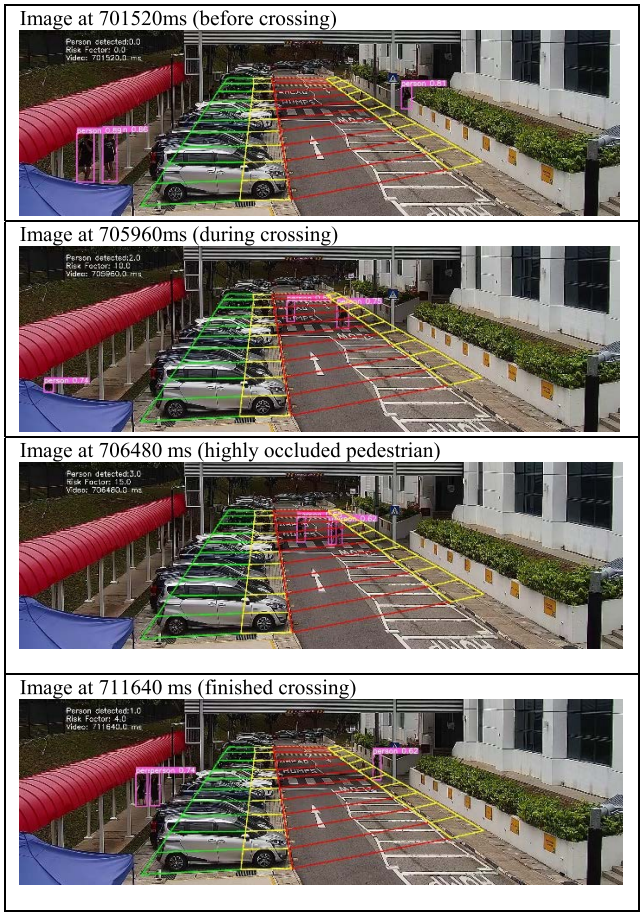

FIGURE 16. Images of pedestrians crossing from both sides.

Here, the highest RT occurs when all pedestrians are on the road region. The challenge in this event is the highly occluded pedestrian with high overlap with another pedestrian, resulting in the difference between the actual RT and the ground truth. However, the variation of RT figures is in the same trend as the ground truth, which accounts for this event's presence. The spikes are due to the detection of the occluded pedestrian, which attempts to bring it close to the ground truth RT. In comparing LID, CED and AED, it is clear that AED is not efficient for this event since the event took place in the middle rows. As for CED, the high RT figures for CED aim to amplify the occlusion and present similar trend characteristics as LID.

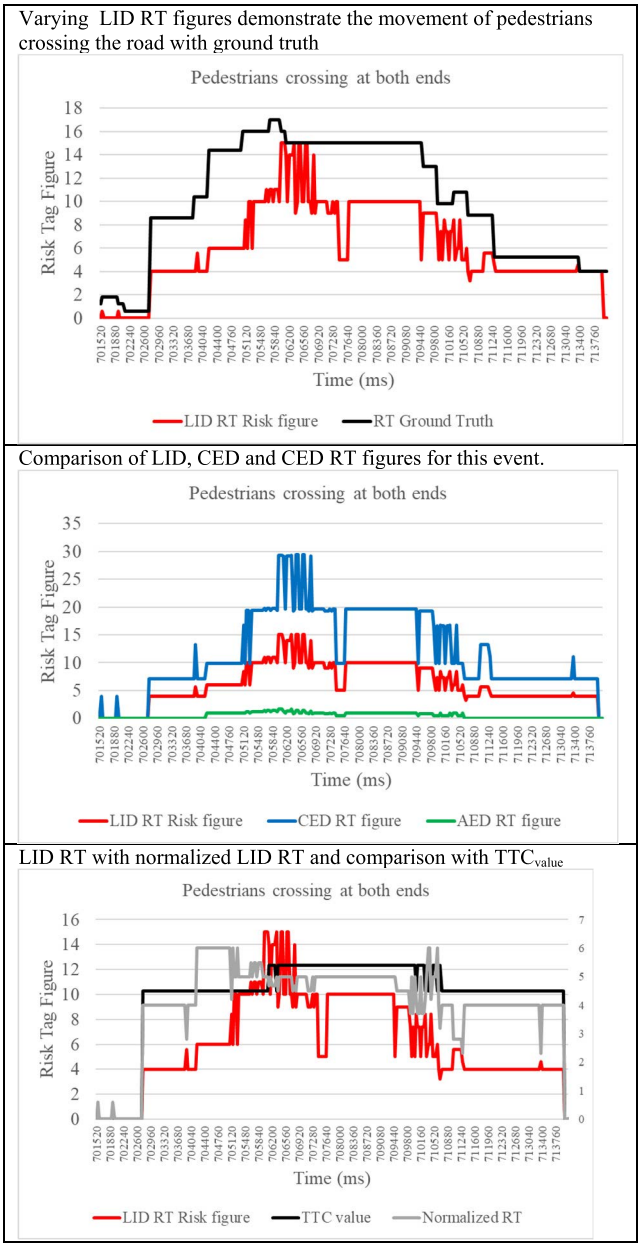

FIGURE 17. Comparative analysis of RT for pedestrian crossing at both ends.

The last image in Figure 17 shows the LID RT and normalized LID RT (secondary axis) compared to TTC_{value} (secondary axis). Since multiple pedestrians cross at both ends of the AOI almost simultaneously, the normalized RT have a different trend line with the LID RT. Normalized RT shows high values when most pedestrians are in the road region, and the values drop when most pedestrians leave the road region. Separately for the time to collision, the TTC_{value} remains between 4.5-5.4 seconds since the crossing mainly occurs at the middle rows. The flat TTC_{values} also indicate mainly

lateral movement, and not much longitudinal movement is detected.

2) PEDESTRIAN DIAGONAL MOVEMENT FROM MIDDLE TO FRONT

This event highlights the different algorithm's sensitivity toward pedestrian movement from middle to front rows. This event represents a scene where the pedestrian crosses diagonally from the middle to the front of the AOI. The objective is to evaluate the sensitivity of AED and CED, in particular to pedestrians moving from middle to front rows. The scene is first demonstrated through the images shown in Figure 18. For comparative analysis, LID, CED and AED figures are shown together with LID ground truth in Figure 19.

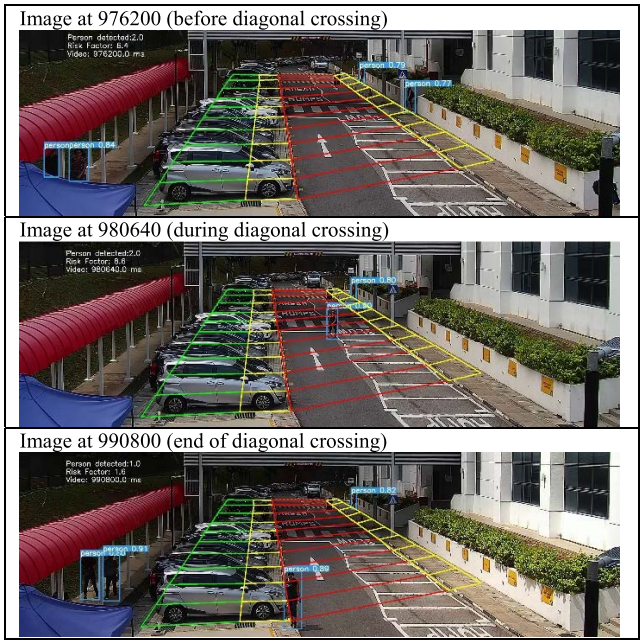

FIGURE 18. Images of pedestrian diagonal crossing.

From Figure 19, LID RT is almost similar to the ground truth with the exception that another pedestrian is seen toggling at the rear boundary of the AOI, thus causing some fluctuations of the RT. A single frame false-negative pedestrian detection causes the dip in LID detection. Other than those, the gradual increase of RT represents the movement of the pedestrian towards the front row. The increase in LID can be seen in a linear form. As for the comparison of LID, CED and AED, it is evident that in this scene, CED is neither efficient nor able to detect the sequence of movement since the risk matrix figures are similar for CED for front and middle rows. Thus only LID and AED are efficient in this event. Moreover, AED performs better if there are pedestrians in the front rows.

The last image in Figure 19 shows the LID RT and normalized LID RT (secondary axis) compared to TTC_{value} (secondary axis). Since the LID RT is similar to the normalized LID RT, the scene only has one pedestrian most of the time. TTC_{value} behaves inversely to the RT figures, which

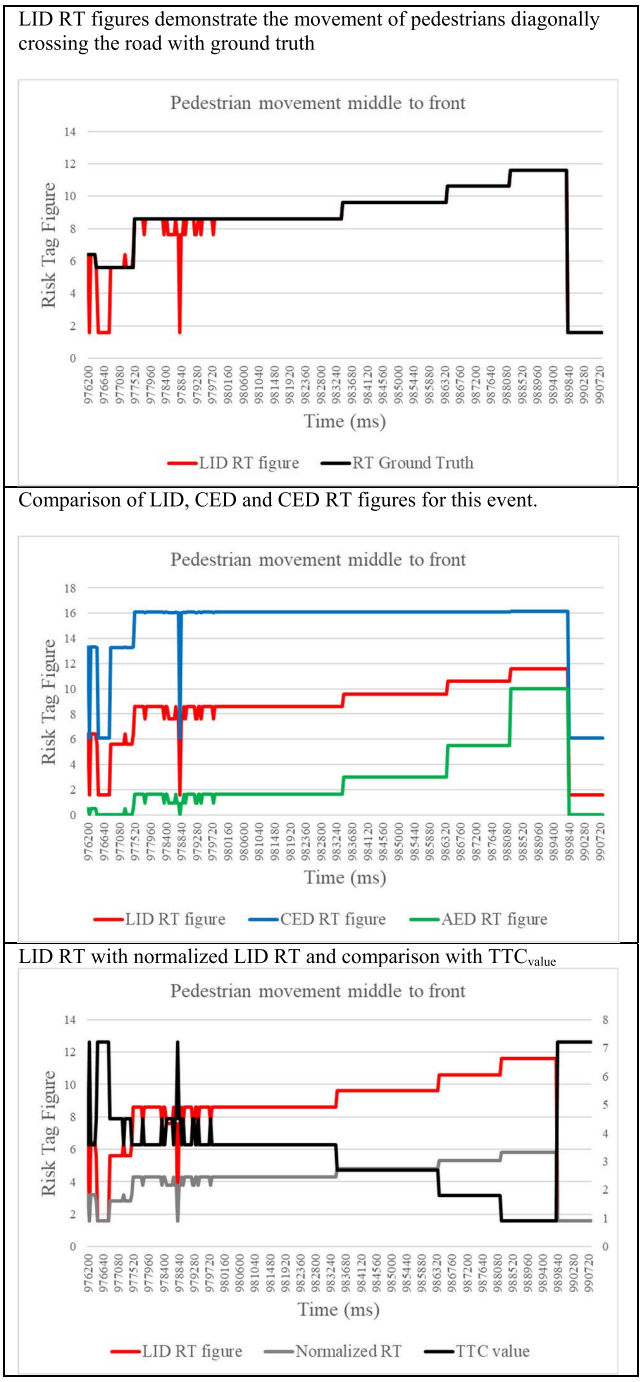

FIGURE 19. Comparative analysis of RT for diagonal crossing.

corroborates that there are pedestrian movements from the middle to the front of the AOI.

3) MULTI-PEDESTRIAN MOVEMENT FROM REAR TO FRONT

STREET is also validated for scenes that consist of multi-pedestrian movements from rear to front rows. From the design of the three algorithms, it is evident that LID and CED will be the most suitable for such events. The images of the multi-pedestrians movement are illustrated in Figure 20.

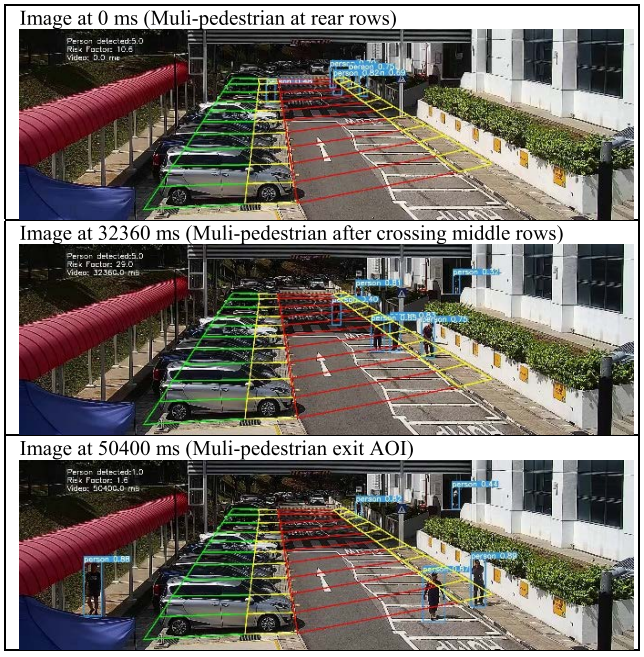

FIGURE 20. Images of multi-pedestrians movement from rear to front.

Figure 21 shows the LID RT ground truth, LID RT and the normalized RT (with secondary Y-axis). These data show that the LID RT trend moves in tandem with the LID RT ground truth with minor spikes due to occlusion among the pedestrians, and some pedestrians are walking near the boundary of the AOI. The normalized RT can be used to observe the movement trend related to the pedestrian's movement. In this event, the multi-pedestrian disperses to other sub-regions as it walks towards the front rows. The rising of normalized RT also reflects the movement of pedestrians from rear to front.

Comparing LID, CED and AED in this event, AED would not be an efficient algorithm to track the movement of the pedestrian, leaving only LID and CED. Following the model design, the startling difference of RT from 0 – 9000 ms demonstrates the linearity of LID RT increment as the pedestrian moves towards the front; in contrast, an exponential increase can be observed in CED RT with higher elevated values. These results verify the intended performance of the three algorithms.

With the last image in Figure 21, the stepwise decreasing TTC_{value} can be observed in tandem to the rising RT. This shows the feasibility of the TTC_{value} and highlights the time to collision as the pedestrians move closer to the front (ego vehicle).

4) MOVING VEHICLE THAT OCCLUDED PEDESTRIAN

The STREET can identify occlusion as shown in Figure 22. The pedestrian on the left second row (Y_{r2}) is occluded by the driven truck. When the occlusion disappears, the pedestrian has reached (Y_{r3}).

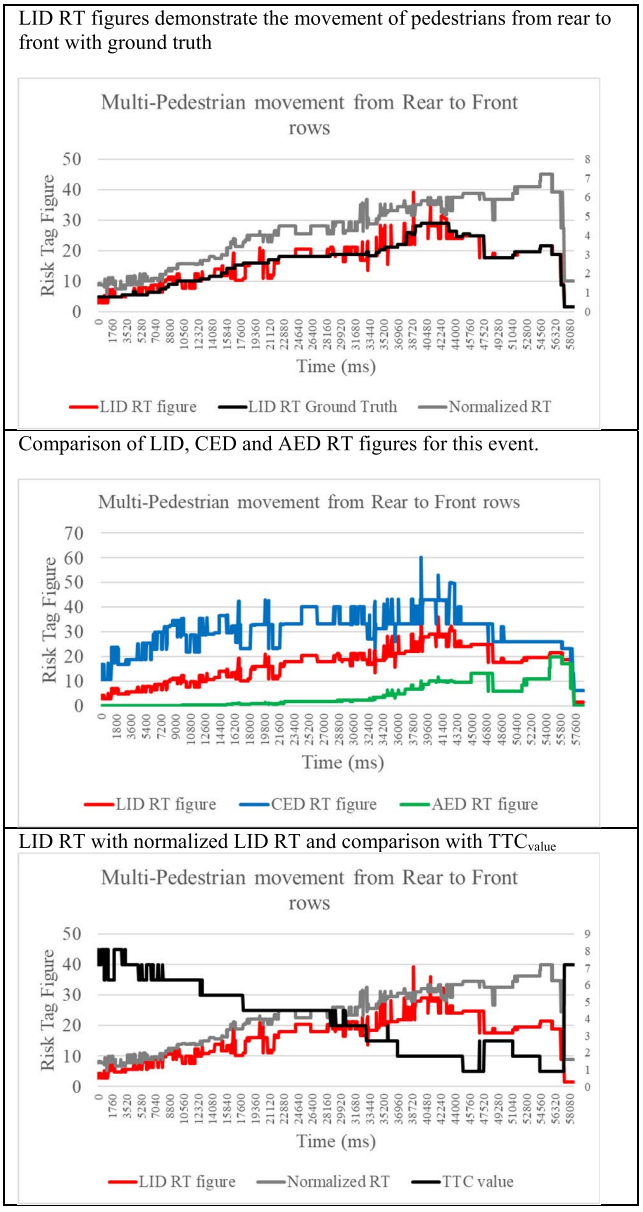

FIGURE 21. Comparative analysis of RT for multi-pedestrian movement from rear to front.

The images in Figure 22 illustrate pedestrians are occluded from moving vehicles based on the sudden drop in LID RT figures for a duration of 56 frames = 2.24 seconds (128640 – 130880 ms) as shown in Figure 23. The LID RT drops instantaneously from 8.4 to 2.0, and returns to 7.6 only after 56 frames (not 8.4, due to the movement of the pedestrian). This event can be detected by inserting an occlusion filter in the STREET. In terms of the LID RT figures, they are close to the LID ground truth. The spikes seen is due to a false positive detection error caused by YOLOv5.

The last image in Figure 23 shows that the TTC_{value} increases sharply once the second-row pedestrian is occluded due to the remaining pedestrian residing at row R_8 of the AOI. In a normal case, if only one pedestrian is occluded in the

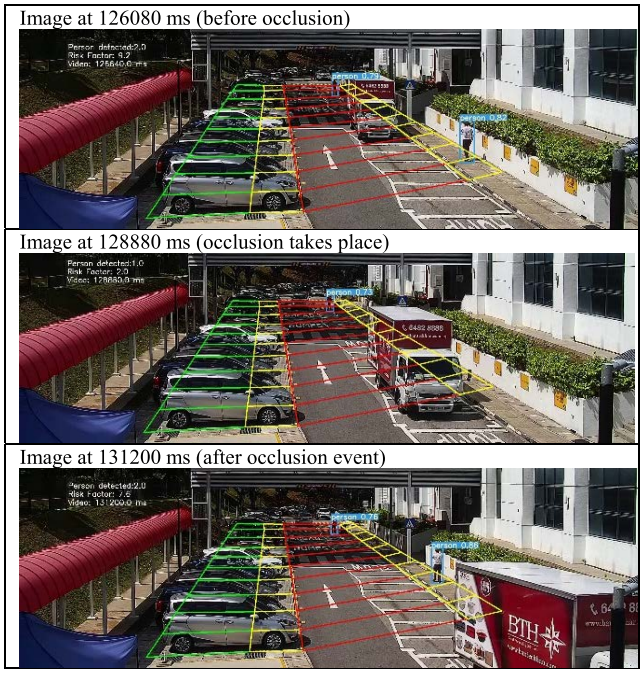

FIGURE 22. Images of moving vehicle that occluded pedestrian and RT results.

scene, the TTC_{value} will drop to zero, which is not a good indicator for occlusion.

Thus, in this event, LID or CED is still a better indicator for occlusion with sharp decrease and increase for an extended period in terms of seconds. Since the pedestrian is moving away from the front row, AED is not obvious for detecting this event.

E. COMPARISON OF LID, CED, AED AND TIME TO COLLISION FINDINGS

With the proposed STREET, risk zoning of scene A was implemented, and the characteristic of the three algorithms LID, CED and AED was verified and validated with the events shown within the scene using ground truth data and TCC_{value} to illustrate its impact. Table 11 summarises the sensitivity for each event, comparing each algorithm based on the dataset of 97k samples. Based on the values, it is evident that if the RT figures are high and TTC_{value} is low, this represents a high-risk scene. If the RT is high and the TTC_{value} is high, this infers that multiple pedestrians are present at middle-rear rows of the AOI. The normalized RT can be used to check for this truth. Lastly, if the RT figures are low with a high TTC_{value}, this represents a low-risk scene.

The utilization of LID, CED, and AED supports the AV real-time risk assessment intentions. The normalized RT ranges from 0 to 10, which acts as a suitable severity rating for the real-time ASIL [9] and an automatic fitting for the real-time recursive risk assessment PRN [10] in terms of automatic risk quantification for the collision of AV with pedestrian(s).

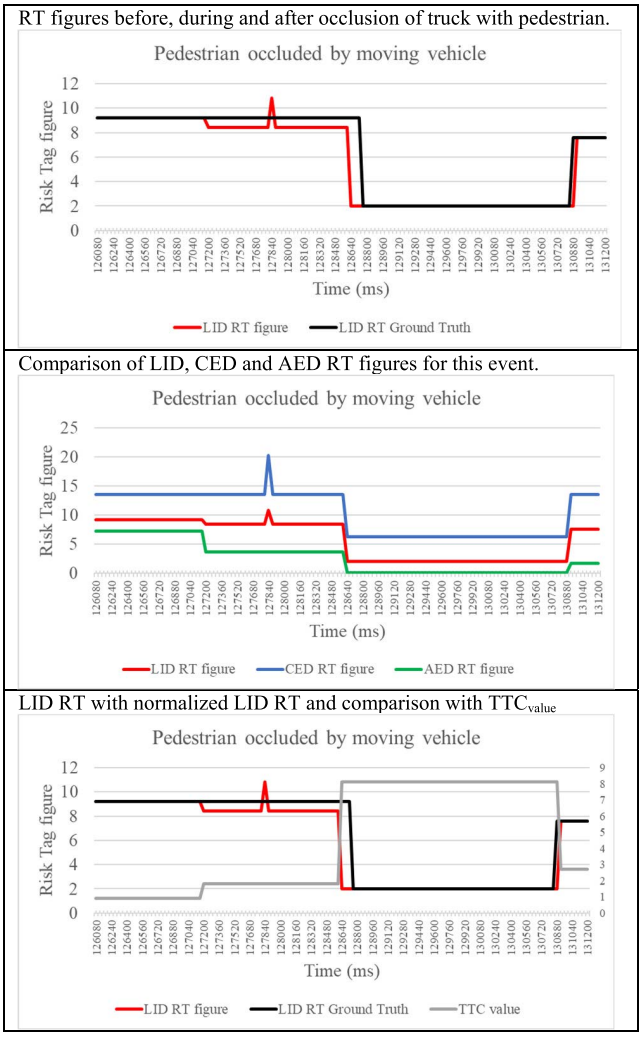

FIGURE 23. Comparative analysis of RT for moving vehicle that occluded pedestrian.

TABLE 11. Comparison table for all three street algorithms and TCC_{value} based on scene A findings.

Scene	Events	Algorithms and TCC_{value}
Corner bend	Pedestrians movement from left to right and vice versa	LID and CED are more effective than AED. If TCC_{value} is flat, the pedestrian crossing is at a similar row.
Corner bend	Pedestrians with diagonal movement from middle to front	LID and AED are more effective than CED, TCC_{value} can be used to validate the same findings with the opposite trend
Corner bend	Multi-pedestrian movement from rear to front	LID and CED are more effective than AED, TCC_{value} can be used to validate the same findings with the opposite trend
Corner bend	Moving vehicle that occludes pedestrian	CED and LID are more effective than AED, TCC_{value} can be used to validate the same findings with the opposite trend

V. CONCLUSION

In this paper, a novel methodology (STREET) is proposed to quantitatively risk tag the severity rating of the

real-world environment for real-time risk assessment via cooperative mode. STREET focuses on pedestrians in uncontrolled traffic areas. Working in cooperative mode using RSI (i.e., infrastructure-based) to overcome AV sensors' limitations, it provides advanced warnings if high-risk scenes are detected in the vicinity of the environment. This method provides a novel risk zoning of the scene and includes three different algorithms, LID, CED and AED of risk detection matrix. The different risk detection matrix focus on different layers of the scene to enhance AV risk and safety management with advanced warning. Furthermore, STREET provides time to collision figures as a form of advanced warning, referencing the ego vehicle to the detected pedestrian. In this paper, STREET is implemented and tested on an uncontrolled traffic scene, comparing normalized RT results based on LID, CED and AED. With the 97k samples of quantitative RT figures, average and variance are determined in identifying the occurrence of pedestrians according to the zone layers. Post-processing validations were conducted using pedestrian occurrence heatmaps and ground truth to validate the findings. In addition, further analyses of LID, CED and AED were conducted with four specific events, which identified their effectiveness in different events using TTC as validation. Thus, based on event(s) selection preference, a specific risk detection matrix can be chosen. As such the normalized RT (ranging from 0 to 10) from STREET can be used as a quantitative severity rating figure for existing real-time risk assessment frameworks, thus resulting in lightweight integration for the AV for risk assessment management. Therefore, in conclusion, STREET provides an enhancement to risk and safety management of AV. It is also recognized that the accuracy of object detection using deep learning methods can be further improved. This will be addressed in future research with increased focus on using this methodology for larger-scale deployment that incorporates multiple cooperative mode exchanges and validation of real-time risk assessment at the AV.

ACKNOWLEDGMENT

This work is performed within the Singapore Institute of Technology (SIT) where the novel algorithm is used and tested. The authors are thankful to MooVita Pte Ltd., for their collaboration in exploring this concept.

REFERENCES

[1] *Surface Vehicle Recommended Practice (R) Considerations for ISO 26262 ASIL Hazard Classification*, SAE, Warrendale, PA, USA, 2018.
[2] *Road Vehicles Functional Safety*, ISO Standard 26262, 2018. [Online]. Available: https://www.iso.org/standard/68383.html
[3] *Road Vehicles—Safety of the Intended Functionality*, ISO/PAS Standard 21448 2019.
[4] S. Ulbrich, T. Menzel, A. Reschka, F. Schuldt, and M. Maurer, "Defining and substantiating the terms scene, situation, and scenario for automated driving," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 982–988, doi: 10.1109/ITSC.2015.164.
[5] E. de Gelder, H. Elrofai, A. K. Saberi, J.-P. Paardekooper, O. Op den Camp, and B. de Schutter, "Risk quantification for automated driving systems in real-world driving scenarios," *IEEE Access*, vol. 9, pp. 168953–168970, 2021, doi: 10.1109/ACCESS.2021.3136585.
[6] B. Kramer, C. Neurohr, M. Büker, E. Böde, M. Fränzle, and W. Damm, "Identification and quantification of hazardous scenarios for automated driving," in *Model-Based Safety and Assessment*, M. Zeller and K. Höfig, Eds. Cham, Switzerland: Springer, 2020, pp. 163–178.
[7] A. Wardziński, "Safety assurance strategies for autonomous vehicles," in *Computer Safety, Reliability, and Security*, M. D. Harrison and M.-A. Sujan, Eds. Berlin, Germany: Springer, 2008, pp. 277–290.
[8] C. Oh, J. Oh, and J. Min, "Real-time detection of hazardous traffic events on freeways: Methodology and prototypical implementation," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2129, no. 1, pp. 35–44, Jan. 2009, doi: 10.3141/2129-05.
[9] S. Khastgir, H. Sivencrona, G. Dhadyalla, P. Billing, S. Birrell, and P. Jennings, "Introducing ASIL inspired dynamic tactical safety decision framework for automated vehicles," in *Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2017, pp. 1–6, doi: 10.1109/ITSC.2017.8317868.
[10] W. M. Dan Chia, S. Loong Keoh, A. L. Michala, and C. Goh, "Real-time recursive risk assessment framework for autonomous vehicle operations," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC-Spring)*, Apr. 2021, pp. 1–7, doi: 10.1109/VTC2021-Spring51267.2021.9448759.
[11] C. M. Hruschka, D. Topfer, and S. Zug, "Risk assessment for integral safety in automated driving," in *Proc. 2nd Int. Conf. Intell. Auto. Syst. (ICoIAS)*, Feb. 2019, pp. 102–109, doi: 10.1109/ICoIAS.2019.00025.
[12] M. A. Gosavi, B. B. Rhoades, and J. M. Conrad, "Application of functional safety in autonomous vehicles using ISO 26262 standard: A survey," in *Proc. SoutheastCon*, Apr. 2018, pp. 1–6, doi: 10.1109/SECON.2018.8479057.
[13] O. M. Kirovskii and V. A. Gorelov, "Driver assistance systems: Analysis, tests and the safety case. ISO 26262 and ISO PAS 21448," in *Proc. IOP Conf., Mater. Sci. Eng.*, 2019, vol. 534, no. 1, Art. no. 012019, doi: 10.1088/1757-899X/534/1/012019.
[14] A. Takacs, D. A. Drexler, P. Galambos, I. J. Rudas, and T. Haidegger, "Assessment and standardization of autonomous vehicles," in *Proc. IEEE 22nd Int. Conf. Intell. Eng. Syst. (INES)*, Jun. 2018, pp. 000185–000192, doi: 10.1109/INES.2018.8523899.
[15] S. Khaiyum, B. Pal, and Y. S. Kumaraswamy, "An approach to utilize FMEA for autonomous vehicles to forecast decision outcome," in *Proc. 3rd Int. Conf. Frontiers Intell. Comput., Theory Appl. (FICTA)*. Cham, Switzerland: Springer, 2015, pp. 701–709.
[16] *Failure modes and effects analysis (FMEA and FMECA)*, IEC Standard 60812, 2018.
[17] G. Bagschik, A. Reschka, T. Stolte, and M. Maurer, "Identification of potential hazardous events for an unmanned protective vehicle," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, 2016, pp. 691–697, doi: 10.1109/IVS.2016.7535462.
[18] F. Warg, M. Gassilewski, J. Tryggvesson, V. Izosimov, A. Werneman, and R. Johansson, "Defining autonomous functions using iterative hazard analysis and requirements refinement," in *Proc. Int. Conf. Comput. Saf., Rel., Secur.*, 2016, pp. 286–297.
[19] T. Stolte, G. Bagschik, A. Reschka, and M. Maurer, "Hazard analysis and risk assessment for an automated unmanned protective vehicle," presented at the IEEE Intell. Vehicles Symp. (IV), Redondo Beach, CA, USA, 2017.
[20] F. Warg, M. Skoglund, A. Thorsen, R. Johansson, M. Brannstrom, M. Gyllenhammar, and M. Sanfridson, "The quantitative risk norm–A proposed tailoring of HARA for ADS," in *Proc. 50th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw. Workshops (DSN-W)*, Jun. 2020, pp. 86–93, doi: 10.1109/DSN-W50199.2020.00026.
[21] W. M. D. Chia, S. L. Keoh, C. Goh, and C. Johnson, "Risk assessment methodologies for autonomous driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, early access, Apr. 13, 2022, doi: 10.1109/TITS.2022.3163747.
[22] T. Menzel, G. Bagschik, and M. Maurer, "Scenarios for development, test and validation of automated vehicles," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 1821–1827, doi: 10.1109/IVS.2018.8500406.
[23] A. Pütz, A. Zlocki, J. Bock, and L. Eckstein, "System validation of highly automated vehicles with a database of relevant traffic scenarios," in *Proc. 12th ITS Eur. Congr.*, Strasbourg, France, 2017.
[24] S. Riedmaier, T. Ponn, D. Ludwig, B. Schick, and F. Diermeyer, "Survey on scenario-based safety assessment of automated vehicles," *IEEE Access*, vol. 8, pp. 87456–87477, 2020, doi: 10.1109/ACCESS.2020.2993730.

[25] C. B. S. T. Molina, J. R. D. Almeida, L. F. Vismari, R. I. R. Gonzalez, J. K. Naufal, and J. B. Camargo, "Assuring fully autonomous vehicles safety by design: The autonomous vehicle control (AVC) module strategy," in *Proc. 47th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw. Workshops (DSN-W)*, Jun. 2017, pp. 16–21, doi: 10.1109/DSN-W.2017.14.

[26] C. Katrakazas, M. Quddus, and W.-H. Chen, "A new integrated collision risk assessment methodology for autonomous vehicles," *Accident Anal. Prevention*, vol. 127, pp. 61–79, Jun. 2019, doi: 10.1016/j.aap.2019.01.029.

[27] P. Ledent, A. Paigwar, A. Renzaglia, R. Mateescu, and C. Laugier, "Formal validation of probabilistic collision risk estimation for autonomous driving," in *Proc. IEEE Int. Conf. Cybern. Intell. Syst. (CIS) IEEE Conf. Robot., Autom. Mechatronics (RAM)*, Nov. 2019, pp. 433–438, doi: 10.1109/CIS-RAM47153.2019.9095806.

[28] M.-Y. Yu, R. Vasudevan, and M. Johnson-Roberson, "Occlusion-aware risk assessment for autonomous driving in urban environments," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 2235–2241, Apr. 2019, doi: 10.1109/LRA.2019.2900453.

[29] M. Lee, M. Sunwoo, and K. Jo, "Collision risk assessment of occluded vehicle based on the motion predictions using the precise road map," *Robot. Auto. Syst.*, vol. 106, pp. 179–191, Aug. 2018, doi: 10.1016/j.robot.2018.05.005.

[30] H. M. Fahmy, M. A. A. El Ghany, and G. Baumann, "Vehicle risk assessment and control for lane-keeping and collision avoidance at low-speed and high-speed scenarios," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 4806–4818, Jun. 2018, doi: 10.1109/TVT.2018.2807796.

[31] K. Beckers, M. Heisel, T. Frese, and D. Hatebur, "A structured and model-based hazard analysis and risk assessment method for automotive systems," in *Proc. IEEE 24th Int. Symp. Softw. Rel. Eng. (ISSRE)*, Nov. 2013, pp. 238–247.

[32] B. Wu, Y. Yan, D. Ni, and L. Li, "A longitudinal car-following risk assessment model based on risk field theory for autonomous vehicles," *Int. J. Transp. Sci. Technol.*, vol. 10, no. 1, pp. 60–68, Mar. 2021, doi: 10.1016/j.ijtst.2020.05.005.

[33] P. Feth, M. N. Akram, R. Schuster, and O. Wasenmüller, "Dynamic risk assessment for vehicles of higher automation levels by deep learning," in *Proc. Int. Conf. Comput. Saf., Rel., Secur.* Cham, Switzerland: Springer, 2018, pp. 535–547.

[34] M. Strickland, G. Fainekos, and H. B. Amor, "Deep predictive models for collision risk assessment in autonomous driving," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2018, pp. 4685–4692, doi: 10.1109/ICRA.2018.8461160.

[35] D. Feng, L. Rosenbaum, and K. Dietmayer, "Towards safe autonomous driving: Capture uncertainty in the deep neural network for lidar 3D vehicle detection," *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 3266–3273.

[36] D. Shin, B. Kim, J. Seo, and K. Yi, "Effects of wireless communication on integrated risk management based automated vehicle," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 1767–1772, doi: 10.1109/ITSC.2015.287.

[37] D. Shin, B. Kim, K. Yi, A. Carvalho, and F. Borrelli, "Human-centered risk assessment of an automated vehicle using vehicular wireless communication," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 667–681, Feb. 2019, doi: 10.1109/TITS.2018.2823744.

[38] V. Thayananthan and R. Ahmed Shaikh, "Contextual risk-based decision modeling for vehicular networks," *Int. J. Comput. Netw. Inf. Secur.*, vol. 8, no. 9, pp. 1–9, Sep. 2016.

[39] S. Demmel, D. Gruyer, J.-M. Burkhardt, S. Glaser, G. Larue, O. Orfila, and A. Rakotonirainy, "Global risk assessment in an autonomous driving context: Impact on both the car and the driver," *IFAC-PapersOnLine*, vol. 51, no. 34, pp. 390–395, 2019, doi: 10.1016/j.ifacol.2019.01.009.

[40] H. Zhao, T. Mao, H. Yu, M. Zhang, and H. Zhu, "A driving risk prediction algorithm based on PCA -BP neural network in vehicular communication," in *Proc. 10th Int. Conf. Intell. Human-Mach. Syst. Cybern. (IHMSC)*, vol. 2, Aug. 2018, pp. 164–169.

[41] R. A. Shaikh and V. Thayananthan, "Risk-based decision methods for vehicular networks," *Electronics*, vol. 8, no. 6, p. 627, Jun. 2019, doi: 10.3390/electronics8060627.

[42] R. V. Cowlagi, R. C. Debski, and A. M. Wyglinski, "Risk quantification for automated driving using information from V2 V basic safety messages," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC-Spring)*, Apr. 2021, pp. 1–5, doi: 10.1109/VTC2021-Spring51267.2021.9448849.

[43] M. Shan, K. Narula, Y. F. Wong, S. Worrall, M. Khan, P. Alexander, and E. Nebot, "Demonstrations of cooperative perception: Safety and robustness in connected and automated vehicle operations," *Sensors*, vol. 21, no. 1, p. 200, Dec. 2020, doi: 10.3390/s21010200.

[44] M. Tsukada, T. Oi, A. Ito, M. Hirata, and H. Esaki, "AutoC2X: Open-source software to realize V2X cooperative perception among autonomous vehicles," in *Proc. IEEE 92nd Veh. Technol. Conf. (VTC-Fall)*, Nov. 2020, pp. 1–6, doi: 10.1109/VTC2020-Fall49728.2020.9348525.

[45] S. A. Abdel Hakeem, A. A. Hady, and H. Kim, "5G-V2X: Standardization, architecture, use cases, network-slicing, and edge-computing," *Wireless Netw.*, vol. 26, no. 8, pp. 6015–6041, Nov. 2020, doi: 10.1007/s11276-020-02419-8.

[46] H. Zhou, W. Xu, J. Chen, and W. Wang, "Evolutionary V2X technologies toward the internet of vehicles: Challenges and opportunities," *Proc. IEEE*, vol. 108, no. 2, pp. 308–323, Feb. 2020, doi: 10.1109/JPROC.2019.2961937.

[47] A. Rasouli and J. K. Tsotsos, "Autonomous vehicles that interact with pedestrians: A survey of theory and practice," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 900–918, 2020, doi: 10.1109/TITS.2019.2901817.

[48] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "Are they going to cross? A benchmark dataset and baseline for pedestrian crosswalk behavior," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2017, pp. 206–213, doi: 10.1109/ICCVW.2017.33.

[49] A. Hussein, F. Garcia, J. M. Armingol, and C. Olaverri-Monreal, "P2 V and V2P communication for pedestrian warning on the basis of autonomous vehicles," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2016, pp. 2034–2039, doi: 10.1109/ITSC.2016.7795885.

[50] L. Hobert, A. Festag, I. Llatser, L. Altomare, F. Visintainer, and A. Kovacs, "Enhancements of V2X communication in support of cooperative autonomous driving," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 64–70, Dec. 2015, doi: 10.1109/MCOM.2015.7355568.

[51] (2017). *IMDA TS- Dedicated ShortRange Communications in Intelligent Transport Systems IMDA, IMDA*. [Online]. Available: https://www.imda.gov.sg/-/media/Imda/Files/Regulation-Licensing-and-Consultations/ICT-Standards/Telecommunication-Standards/Radio-Comms/IMDA-TS-DSRC.pdf?la=en

[52] *Ultralytics/YOLOv5: V6.0—YOLOv5n 'Nano' Models, Roboflow Integration, TensorFlow Export, OpenCV DNN Support*, Zenodo, 2021, doi: 10.5281/zenodo.5563715.

[53] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 11, pp. 3212–3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.

[54] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.

[55] P. Droździel, S. Tarkowski, I. Rybicka, and R. Wrona, "Drivers' reaction time research in the conditions in the real traffic," *Open Eng.*, vol. 10, no. 1, pp. 35–47, 2020, doi: doi:10.1515/eng-2020-0004.

[56] J. Mateu and W. G. Müller, *Spatio–Temporal Design Advances in Efficient Data Acquisition (Statistics in Practice)*. West Sussex, U.K: Wiley, 2013.

WEI MING DAN CHIA (Member, IEEE) is currently pursuing the Ph.D. degree with the University of Glasgow. He has been a Senior Lecturer at the Singapore Institute of Technology, since 2017. Before his academic journey, he has 17 years of industry experience spanning across several automotive companies, such as Delphi, Siemens VDO, and Continental, from 2001 to 2017. His last role in Continental was the Director of the ITS Laboratory and the Head of Advance Development in Asia for infotainment and connectivity products. His main research interest includes applying artificial intelligence for autonomous vehicles safety operations and intelligent transport systems.

SYE LOONG KEOH (Member, IEEE) received the Ph.D. degree in computing science from Imperial College London, in 2005. He is currently an Associate Professor with the School of Computing Science, University of Glasgow Singapore (UGS), and also the Director of Research Programs. Prior to joining Glasgow, he was a Senior Scientist at Philips Research Eindhoven, The Netherlands. His areas of expertise include cyber security for the Internet of Things (IoT), lightweight security systems for cyber-physical systems, and policy-based security management for pervasive and distributed systems. He leads the cyber-security research activities in UofG Singapore where he has designed several lightweight authentication protocols and key management schemes for the IoT, building management, and industrial control systems.

ANNA LITO MICHALA received the M.Sc. degree in biomedical engineering and the Ph.D. degree in edge computing from the University of Strathclyde, in 2008 and 2018, respectively. She is currently an Assistant Professor of understandable autonomous systems at the School of Computing Science, University of Glasgow, U.K. She has expertise in safety-critical systems software engineering, system/sub-system prognostics in extreme and resource constrained environments, and vessel condition monitoring. Her research interests include cyber-physical systems engineering, the Industrial Internet of Things, and machine learning for decision support.

CINDY GOH (Senior Member, IEEE) received the Ph.D. degree from the University of Glasgow, U.K., in 2004. She is currently the Director of the University of Glasgow Singapore (UGS), where she has overall responsibility for its strategy and management. Prior to this, she was the Director of Research Programs at the UGS, from 2013 to 2016. Her research interests include intelligent optimization and data analytics for optimal decision-making and design to advance the state-of-the-art in complex engineering systems, energy and transport networks, and smart manufacturing. She is a fellow of HEA and a Founding Member of the International Union of Radio Science Committee, Singapore.