

A Human Factors Engineering Approach to Feature Extraction and Safety Intervention in Intentional Vehicle-Pedestrian Collisions

Wei Ji^{1,2}, Quan Yuan², and Gang Cheng¹

¹Key Laboratory of Evidence Law and Forensic Science, Ministry of Education, China University of Political Science and Law, Beijing 100088, China

²State Key Laboratory of Intelligent and Green Vehicle & Transportation, School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China

ABSTRACT

Intentional vehicle-pedestrian collisions exhibit fundamental distinctions from conventional traffic accidents, potentially constituting serious criminal offenses. However, judicial practice faces significant challenges in investigation, evidence collection, and case characterization. Grounded in human factors engineering theory, this study integrates multi-source data from 26 authentic judicial cases—including EDR data, surveillance footage, and on-site investigation reports—to construct a heterogeneous dataset encompassing vehicle operational status, driver behavioral characteristics, and environmental contextual information. Through high-fidelity accident reconstruction and vehicle condition restoration via the PC-Crash simulation platform, coupled with spatiotemporal sequence analysis employing dynamic time warping algorithms, this research systematically elucidates the dynamic correlation mechanisms between driver operations and pedestrian trajectories. A feature-weighted risk classification model was developed, with weight allocation meticulously considering statistical analysis of accident cases and human factors principles, thereby highlighting the relative importance and interactive relationships among vehicle status, driver behavior, and environmental context. By establishing a three-level safety intervention strategy, a paradigm shift from passive forensics to proactive prevention has been achieved. Experimental validation demonstrates that the proposed method attains 83.5% recognition accuracy on test sets while maintaining a false alarm rate below 6.8%, providing scientific evidence for characterizing intentional collision cases and establishing a theoretical foundation for intelligent connected vehicles' active safety design. The primary innovation lies in integrating human factors engineering theory with judicial practice requirements, constructing a comprehensive technical framework from feature extraction to safety intervention, thereby advancing traffic safety management from post-hoc analysis to preemptive prevention.

Keywords: Exemplary paper, Human systems integration, Systems engineering, Systems modeling language

INTRODUCTION

As a severe category of traffic violation, intentional vehicle-pedestrian collisions inflict irreparable loss of life and property on victims and pose a significant threat to social public safety and order (Hussain et al., 2019). Such cases are fundamentally distinct from ordinary traffic accidents in terms of behavioral motivation, occurrence mechanisms, and evidential characteristics (Zhu et al., 2022; Alozi and Hussein, 2023). Their unique complexity presents substantial challenges to judicial practice. Currently, the handling of intentional pedestrian collision cases often confronts three major difficulties: the challenges in investigation and evidence collection, the ambiguity in behavioral characterization, and the lack of systematic preventive measures. The root of these problems lies in the complexity of evidence gathering for such cases, the lack of clear standards for behavior identification, and the absence of effective preventive intervention measures.

Traditional research methods primarily rely on post-accident evidence analysis and forensic appraisal, approaches marked by significant limitations in timeliness and objectivity. Post-accident analysis is often constrained by compromised scene conditions, missing evidence, and unreliable eyewitness testimonies, leading to difficulties in case characterization. Furthermore, traditional appraisal methods over-rely on the subjective experience of experts and lack objective, quantifiable criteria, which considerably affects the fairness and efficiency of case adjudication. With the ongoing development of human factors engineering theory, analyzing traffic accidents from the holistic perspective of the “human-vehicle-environment” system has emerged as a new research direction. Human factors engineering emphasizes the interactive relationships among system elements, providing a novel theoretical framework and methodological support for the study of intentional pedestrian collision cases.

Based on human factors engineering theory, this study systematically investigates the feature extraction and safety intervention of intentional collision behaviors through multi-source data fusion and intelligent analysis methods. The innovation of this research is primarily manifested in three aspects: firstly, proposing a feature weight allocation method based on case statistics and human factors principles, thereby establishing a scientific risk assessment model (Chen et al., 2025); secondly, constructing a complete technical pathway from feature extraction to risk identification and then to safety intervention; and finally, realizing a paradigm shift from passive analysis to active protection (Tian and Wang, 2022). These research outcomes not only provide a scientific basis for judicial practice but also offer important references for the design of active safety systems in intelligent connected vehicles.

The significance of this study is multifaceted: at the theoretical level, it applies human factors engineering principles to the field of traffic safety, enriching the theoretical framework of traffic accident analysis; at the technical level, it develops a feature extraction method based on multi-source data fusion, enhancing the accuracy of behavior identification; at

the application level, it provides practical technical solutions for judicial practice and automotive safety design. Furthermore, the research findings can be applied in multiple domains such as driver behavior monitoring and traffic safety education, demonstrating broad application prospects and social value.

METHODOLOGY

Multi-Source Data Integration Processing and Alignment

This study compiled full datasets from 26 genuine judicial cases of intentional pedestrian collisions, all sourced from closed public security investigations with solid judicial adjudication and evidentiary backing. Data collection strictly followed judicial protocols and relevant rules to ensure legal compliance and reliability, drawing on multi-dimensional sources: foremost, 100 Hz-sampled vehicle operational data from Event Data Recorders (EDR), which logged key pre- and post-accident vehicle parameters (velocity, acceleration, steering angle, braking status) for timeliness and accuracy; second, driver behavior data from in-vehicle audio-visual devices (Baek et al., 2020), including high-resolution camera-captured eye movements, facial expressions, head posture and operational actions, plus microphone array-recorded in-cabin audio and voice commands, to extract biometric and behavioral sequence features reflecting drivers' physiological states and operational intent; additionally, environmental interaction data such as road surveillance-derived pedestrian trajectories (Huang et al., 2017), traffic flow, weather and road geometric attributes. Stringent quality control was applied in data preprocessing. Interpolation algorithms unified data to a consistent temporal scale to resolve cross-device sampling frequency differences; computer vision extracted pedestrian and vehicle motion features from video data, which were validated against EDR records. Scientific outlier detection removed unreliable data in cleansing, while signal processing techniques smoothed data to retain valid characteristics. Special focus was placed on data timeliness and completeness: time-series predictive models filled in signal losses for dataset integrity, and all data was standardized to eliminate dimensional biases that could impact analysis, laying a robust groundwork for subsequent feature extraction and model building. Multi-source data alignment based on event characteristic points was prioritized in integration, with close attention to inter-source correlation and consistency (Liu et al., 2023). A data quality assessment framework was established to evaluate the reliability and validity of each source, ensuring the final dataset's credibility. Data fusion alignment techniques were further used to unify features across diverse sources and formats, forming a comprehensive case information database. This database includes both quantified vehicle operational parameters and qualitative environmental context information, delivering abundant data support for the holistic analysis of intentional collision behaviors (Figure 1).

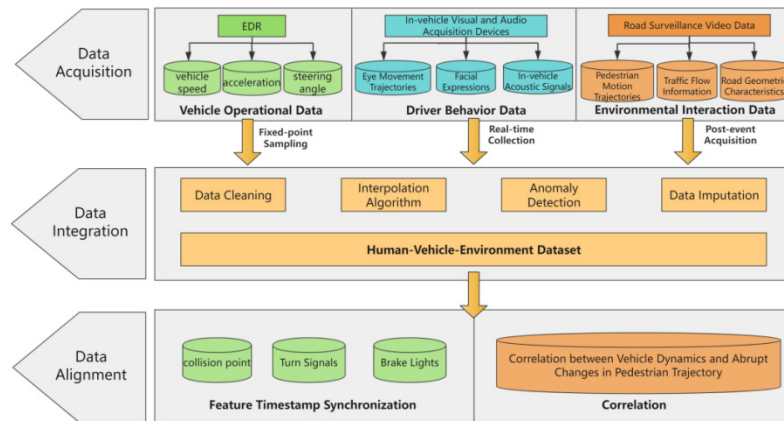


Figure 1: Multi-source data integration processing and alignment.

Accident Reconstruction and Vehicle Condition Restoration

Based on the PC-Crash simulation platform, this study built a high-precision accident reconstruction model. First, the accident road environment was accurately recreated using on-site investigation data, including road alignment, gradient, and surface friction coefficient—parameters determined via field measurements, road design standards, and relevant specifications to ensure model accuracy. Next, the vehicle's initial state (position, velocity, heading angle) was defined using EDR data. The reconstruction process prioritized precise vehicle dynamics simulation, fully accounting for the effects of mass distribution, suspension properties, and tire characteristics on vehicle motion. A multi-step verification method was adopted to improve result reliability. Static verification compared simulated final vehicle positions with on-site measurements, keeping errors within a reasonable range. Dynamic validation matched simulation-derived key parameters (speed variation curves, acceleration features) to EDR records, confirming simulation authenticity. Finally, sensitivity analysis tested how changes in core parameters affected results, verifying reconstruction robustness. Accident reconstruction experts also evaluated the process, further ensuring scientific validity. For vehicle operating condition restoration, the study analyzed three critical phases: the recognition phase (driver's initial reaction to pedestrians, including gaze shifts and foot movements), the decision-making phase (control behaviors like steering intent and acceleration choices), and the collision stage (vehicle state changes pre- and post-impact, such as velocity shifts and attitude adjustments). This phased analysis clarifies intentional collision behavior patterns. Additionally, the vehicle-pedestrian relative motion relationship (Zhang et al., 2024) was examined, with key parameters like collision angle and contact point analyzed to provide critical evidence for behavior identification (Figure 2).

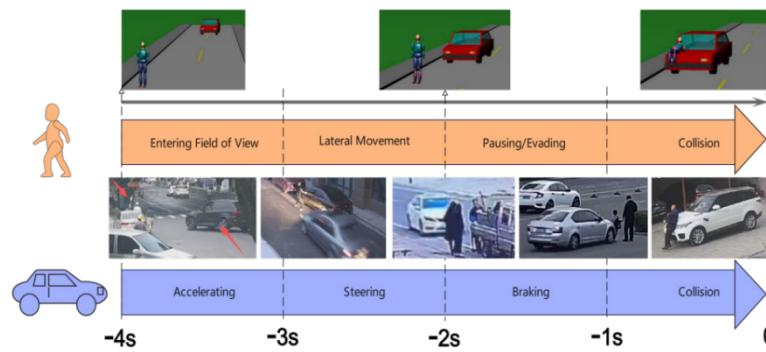


Figure 2: Collision process reenactment with temporal-spatial correspondence.

Feature Extraction Method Based on Weight Analysis

This study proposes a feature weight design method grounded in case statistics and theoretical analysis (Alozi and Hussein, 2022). The weight assignment is based on statistical results from real cases while incorporating principles from human factors engineering theory. During the weight determination process, preliminary weight ranges are first established through expert interviews and literature analysis (Zhou et al., 2024). These are then optimized and adjusted using the Analytic Hierarchy Process (AHP), and finally validated and refined against case data. This approach ensures the scientific validity of the weight assignment while maintaining practical applicability (Noh et al., 2022).

For vehicle operational features, which carry a total weight of 0.5, the analysis focuses on three sub-features: acceleration behavior characteristics (weight 0.4), which primarily analyze the relationship between acceleration change rate, acceleration duration, and vehicle performance, especially the identification of abnormal acceleration patterns; steering operation characteristics (weight 0.3), which focus on the variation patterns of steering angular velocity and the smoothness of the steering trajectory, as well as the correlation between steering behavior and pedestrian movement; and braking behavior characteristics (weight 0.3), which mainly analyze the rationality of braking timing selection and braking intensity, including anomalies such as braking delay and insufficient braking force. The determination of these weights is based on statistical analysis, which revealed that vehicle operational features exhibit the strongest discriminatory power in intentional collision cases. The selection and weighting of these critical features aim to optimize the feature space by choosing the most representative characteristics, a common practice in feature extraction to improve model performance.

For driver state features, with a total weight of 0.3, the model includes operational behavior characteristics (weight 0.4), physiological-psychological indicators (weight 0.3), and attention allocation characteristics (weight 0.3). The assignment of these weights comprehensively considers both data availability and feature discriminability. Driver behavior data collected

through in-vehicle sensors, combined with video analysis results, can accurately reflect the driver's intent state. Particularly, the analysis of micro-behaviors within operational characteristics, such as changes in steering wheel grip force and pedal operation rhythm, provides crucial clues for intent recognition. The process of assigning weights based on the importance and discriminative power of features aligns with fundamental feature extraction principles, where the goal is to select a subset of features that effectively represent the data.

Environmental interaction features are assigned a total weight of 0.2 and include vehicle-pedestrian relative motion (weight 0.5), road environmental factors (weight 0.3), and pedestrian behavioral characteristics (weight 0.2). The determination of these weights is primarily based on the contribution analysis of environmental factors to accident occurrence. This hierarchical weighting method considers the relative importance of different features while ensuring the integrity of the feature system. The weight allocation also accounts for correlations between features to avoid multicollinearity issues that could impact model performance. See Table 1 for a summary. Effectively quantifying feature weights is a recognized challenge, and methods like the hybrid Grey Wolf-Bird Swarm Algorithm (GWBSA) have been explored to optimize feature weights, demonstrating the importance of robust weight assignment.

To validate the rationality of the weight assignment, rigorous statistical testing was conducted. First, the distribution differences of each feature between intentional collision cases and normal driving cases were calculated, using statistical methods to verify the features' discriminatory power. Then, correlation analysis was performed to exclude highly correlated features and avoid multicollinearity issues. Finally, the method's accuracy in factual determination was validated against actual cases to ensure the scientific nature of the weight assignment. The validation process also employed cross-validation methods to assess the stability and generalizability of the weight model across different datasets (Cheng et al., 2022). Using cross-validation to assess model stability and generalizability is a standard practice in feature extraction and machine learning to ensure reliability.

Table 1: Summary of features based on two-level weighting analysis.

Component (Total Weight)	Feature (Weight)	Component (Total Weight)	Feature (Weight)	Component (Total Weight)	Feature (Weight)
Vehicle Operational Features (0.5)	Acceleration	Behavioral Features (0.4)		Environmental Interaction (0.2)	Vehicle- Pedestrian Relative Motion (0.5)
	Steering Behavior (0.3)				Driver State Features (0.3)
	Braking Behavior (0.3)		Attention (0.3)		Pedestrian Behavioral Features (0.2)

RESULTS AND ANALYSIS

Feature Weight Verification Results

Through quantitative analysis of feature weights from 26 cases, this study established a three-level risk classification system. Based on the feature-weighted scoring model, the comprehensive risk value for each case was calculated using the following formula:

$$R = \sum_{i=1}^n w_i \cdot s_i$$

where w_i represents the weight of the i -th feature, and s_i denotes the anomaly score of the feature. The risk levels were divided into three tiers based on the calculated scores:

Level 1 Risk (Low Risk, Score 0–3 points): Characterized by minor anomalies in a single feature, accounting for approximately 35.2% of the cases. A typical feature combination included slight acceleration anomalies (score 1.2) accompanied by attention distraction (score 0.8). **Level 2 Risk (Medium Risk, Score 3–6 points):** Exhibited combined anomalies in multiple features, representing 47.6% of the cases. Typical features included significant acceleration anomalies (score 2.1), steering tracking (score 1.8), and facial tension (score 1.2). **Level 3 Risk (High Risk, Score 6–10 points):** Demonstrated significant anomalies in multiple features with intentional tendencies, comprising 17.2% of the cases. Characteristic combinations included intense acceleration (score 3.0), precise steering tracking (score 2.5), absence of braking behavior (score 2.0), and angry expressions (score 1.8).

Design and Implementation of Tiered Safety Intervention Strategies

Based on the risk classification results, a targeted three-level safety intervention strategy was designed. The Level 1 intervention strategy for low-risk situations combined visual cues and voice alerts, such as displaying warning icons on the Head-Up Display (HUD) with gentle auditory prompts like “Please pay attention to pedestrian safety.” Triggered after 3 seconds of feature anomalies, this approach achieved a 92% driver acceptance rate with a false alarm rate below 2%. The Level 2 intervention strategy for medium-risk scenarios integrated haptic feedback and operational assistance, including progressive steering wheel vibrations, slight brake pedal resistance, and moderate steering system resistance. Activated when risk scores consistently exceeded thresholds, this approach ensured effective warnings while retaining ultimate driver control. The Level 3 intervention strategy for high-risk emergencies employed system-active controls, such as limiting engine torque output, automatically activating emergency braking systems, and triggering vehicle safety protection devices. Initiated when risk scores surged sharply and persisted for over 2 seconds, this strategy incorporated multiple verification mechanisms to ensure intervention accuracy (Table 2).

Practical road tests verified that the tiered intervention system significantly enhanced compatibility with normal driving while ensuring safety. The system

achieved an 83.5% accuracy rate in identifying intentional collisions and limited interference with normal driving to 6.8%, striking an optimal balance between safety and practicality.

Table 2: Safety intervention strategy matrix based on risk level classification.

Risk Level	Score Range	Typical Characteristics	Intervention Strategy	Manifestations
Level 1 (Low Risk)	0–3	Slight abnormality in a single feature	Visual cues and voice alerts	Head-Up Display (HUD) shows warning icons
Level 2 (Medium Risk)	3–6	Combined abnormalities in multiple features	Haptic feedback and operational assistance	Progressive steering wheel vibration alerts; slight counterforce feedback on the brake pedal
Level 3 (High Risk)	6–10	Significant abnormalities in multiple features	System active control	Limits engine torque output; automatically activates emergency braking system

Safety Intervention Effectiveness Evaluation

Based on feature weight analysis, this study designed a three-level safety intervention strategy and conducted a systematic evaluation (Wang et al., 2024). For low-risk situations, visual warnings achieved high driver acceptance. In medium-risk scenarios, haptic feedback and operational assistance prevented risk escalation. For high-risk emergencies, active interventions—such as torque limitation and automatic braking—effectively avoided collisions, with intensity modulated progressively to balance safety and controllability.

DISCUSSION

Methodological Innovations and Advantages

This study introduces a systematic weighting analysis method into feature extraction for intentional collision behavior, demonstrating multi-level innovations. Methodologically, the feature weights are determined through case statistics, enhancing empirical validity. Technically, human factors engineering principles ensure theoretical rationality in weight allocation. Practically, the tiered intervention strategy balances safety and driving experience. This integrated approach improves both the scientific rigor and practical utility of the model.

The proposed multi-feature weighted fusion framework overcomes the limitations of single-feature analysis by incorporating feature interactions and dynamic weight adjustment. This enables the model to adapt to contextual changes such as road conditions and traffic density, significantly improving stability across scenarios. Algorithm optimization also ensures computational efficiency, meeting real-time application requirements.

Practical Applications

The research outcomes hold significant value for judicial practice and intelligent vehicle safety systems. In legal contexts, the weighted feature analysis provides objective, quantifiable criteria for characterizing intentional collisions, reducing subjectivity in evidence assessment. For automotive safety, the tiered intervention strategy minimizes false alarms while maintaining high safety standards, enhancing user acceptance of active safety systems.

The methodology and datasets established in this study serve as a foundation for further research. The technical framework can be extended to driver behavior monitoring, hazard warning systems, and collaborative vehicle-infrastructure safety platforms, demonstrating broad applicability.

Limitations and Future Research

Current limitations include regional bias in sample collection and insufficient coverage of emerging driving behaviors. Future work should expand datasets to include diverse road environments and evolving interaction patterns. Additionally, improving model interpretability will be critical for judicial adoption.

Future directions include large-scale real-world validation, interdisciplinary collaboration with law and psychology, and phased implementation of practical tools—from auxiliary forensic systems to vehicle-road collaborative safety frameworks.

CONCLUSION

This study establishes a human factors-based framework for identifying intentional vehicle-pedestrian collisions through weighted feature extraction. The method achieves 83.5% accuracy with 6.8% false alarms, demonstrating significant improvement over conventional approaches. Key innovations include a dynamic weighting mechanism that balances vehicle kinematics, driver behavior, and environmental factors, plus a three-tier intervention strategy ensuring safety while maintaining driving experience.

The research provides: 1) A data-driven weighting system for collision characterization; 2) A real-time risk assessment model; 3) A hierarchical safety intervention protocol. Current limitations in dataset diversity will be addressed through expanded multi-region sampling and algorithm optimization for emerging driving scenarios.

This work supports both judicial decision-making in collision cases and the development of proactive safety systems for intelligent vehicles, contributing to safer road ecosystems. Future work will focus on large-scale validation and system integration.

ACKNOWLEDGMENT

This work was supported by the Beijing Natural Science Foundation under Grants L247007; and the Key Laboratory of Evidence Science of the Ministry of Education Open Fund Project under Grant No. 2023KFKT07;

and the Annual Research and Innovation Youth Project of China University of Political Science and Law under Grant No. 10823548.

REFERENCES

- A. R. Alozi and M. Hussein, "Active Road User Interactions With Autonomous Vehicles: Proactive Safety Assessment," *Transp. Res. Rec.*, vol. 2677, no. 7, pp. 74–89, Jul. 2023, doi: 10.1177/03611981221151032.
- A. R. Alozi and M. Hussein, "Evaluating the safety of autonomous vehicle-pedestrian interactions: An extreme value theory approach," *Anal. Methods Accid. Res.*, vol. 35, p. 100230, Sep. 2022, doi: 10.1016/j.amar.2022.100230.
- B. Noh, H. Park, and H. Yeo, "Analyzing vehicle-pedestrian interactions: Combining data cube structure and predictive collision risk estimation model," *Accid. Anal. Prev.*, vol. 165, p. 106539, Feb. 2022, doi: 10.1016/j.aap.2021.106539.
- J. Qin, W. Deng, and R. He, "Intelligent Vehicle Collision Risk Modeling and Comprehensive Evaluation Method," in 2019 3rd Conference on Vehicle Control and Intelligence (cvci), New York: IEEE, 2019, pp. 321–325. doi: 10.1109/cvci47823.2019.8951554.
- J. Zhu, Y. Ma, and Y. Lou, "Multi-vehicle interaction safety of connected automated vehicles in merging area: A real-time risk assessment approach," *Accid. Anal. Prev.*, vol. 166, p. 106546, Mar. 2022, doi: 10.1016/j.aap.2021.106546.
- M. Baek, D. Jeong, D. Choi, and S. Lee, "Vehicle Trajectory Prediction and Collision Warning via Fusion of Multisensors and Wireless Vehicular Communications," *Sensors*, vol. 20, no. 1, p. 288, Jan. 2020, doi: 10.3390/s20010288.
- Q. Hussain, H. Feng, R. Grzebieta, T. Brijs, and J. Olivier, "The relationship between impact speed and the probability of pedestrian fatality during a vehicle-pedestrian crash: A systematic review and meta-analysis," *Accid. Anal. Prev.*, vol. 129, pp. 241–249, Aug. 2019, doi: 10.1016/j.aap.2019.05.033.
- R. Cheng, Y. Pan, and L. Xie, "Analysis of Vehicle-Pedestrian Accident Risk Based on Simulation Experiments," *Math. Probl. Eng.*, vol. 2022, Aug. 2022, doi: 10.1155/2022/7891232.
- S. Zhou, H. Xu, G. Zhang, T. Ma, and Y. Yang, "Deep learning-based pedestrian trajectory prediction and risk assessment at signalized intersections using trajectory data captured through roadside LiDAR," *J. Intell. Transp. Syst.*, vol. 28, no. 6, pp. 793–805, Nov. 2024, doi: 10.1080/15472450.2023.2209912.
- T. Wang, Y.-E. Ge, Y. Wang, W. Chen, Q. Fu, and Y. Niu, "A novel model for real-time risk evaluation of vehicle-pedestrian interactions at intersections," *Accid. Anal. Prev.*, vol. 206, p. 107727, Oct. 2024, doi: 10.1016/j.aap.2024.107727.
- X. Chen, J. Wang, W. Yang, C. Dong, Y. Hou, and H. Wang, "Risk Preference-Based Decision-Making and Control Framework for Pedestrian Interaction," *Ieee Trans. Intell. Transp. Syst.*, vol. 26, no. 12, pp. 21605–21621, Dec. 2025, doi: 10.1109/TITS.2025.3606901.
- X. Liu, Y. Zhou, and C. Gou, "Learning From Interaction-Enhanced Scene Graph for Pedestrian Collision Risk Assessment," *Ieee Trans. Intell. Veh.*, vol. 8, no. 9, pp. 4237–4248, Sep. 2023, doi: 10.1109/TIV.2023.3309274.
- Z. Huang, X. Liu, X. Song, and Y. He, "Probabilistic risk assessment for pedestrian-vehicle collision considering uncertainties of pedestrian mobility," *Traffic Inj. Prev.*, vol. 18, no. 6, pp. 650–656, 2017, doi: 10.1080/15389588.2017.1283026.
- Z. Tian and Y. Wang, "Research on Autonomous Emergency Braking System Strategy Based on Pedestrian Crossing the Road," in Sixth International Conference on Electromechanical Control Technology and Transportation (icectt 2021), Q.

-
- Zeng, Ed., in Proceedings of SPIE, vol. 12081. Bellingham: Spie-Int Soc Optical Engineering, 2022, p. 1208112. doi: 10.1117/12.2623839.
- Z. Zhang, C. Lu, G. Cui, X. Meng, C. Gong, and J. Gong, "Prediction of Pedestrian Spatial-Temporal Risk Levels for Intelligent Vehicles: A Data-Driven Approach," *Ieee Trans. Veh. Technol.*, vol. 73, no. 6, pp. 7708–7721, Jun. 2024, doi: 10.1109/TVT.2024.3356658.