Heliyon 11 (2025) e41073

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

An autonomous vehicles' test case extraction method: Example of vehicle-to-pedestrian scenarios

Zhengping Tan [a,b,**], Qian Wang [a], Wenhao Hu [c], Pingfei Li [a,b], Liangliang Shi [d], Hao Feng [e,*]

[a] School of Automobile &Transportation, Xihua University, Chengdu, 610039, China
[b] Sichuan Xihua Jiaotong Forensic Science Center, Chengdu, 610039, China
[c] State Administration for Market Regulation Defective Product Recall Technical Center (DPRC), Beijing, 100101, China
[d] China Automotive Engineering Research Institute, Chongqing, 401122, China
[e] Key Lab of Forensic Science, Ministry of Justice, China (Academy of Forensic Science), Shanghai, 200063, China

A R T I C L E I N F O

Keywords:
Autonomous vehicles
Test scenario
Fuzzy comprehensive evaluation
Perception system
Decision-making system
Clustering algorithm

A B S T R A C T

Testing autonomous vehicles (AVs) in hazardous scenarios is a crucial technical approach to ensure their safety. A key aspect of this process is the generation of hazard scenarios. In general, such scenarios are generated through cluster analysis of traffic accident data. However, this approach may not fully capture the criticality of the generated scenarios, as it tends to emphasize the statistical characteristics of the data rather than its real-world applicability. This paper proposes a novel method to enhance scenario adaptation by integrating quantization weights with a new clustering algorithm. These weights, representing the correlation between scenario elements and the AV system, are calculated using fuzzy comprehensive evaluation (FCE). The proposed method is applied to 1044 pedestrian accident cases in China, resulting in the identification of nine categories of typical scenarios and corresponding test schemes for both the perception and decision-making systems of AVs. The results show that the new method increases the proportion of critical scenarios by 17.4 % and 13.6 %, respectively, compared to traditional methods. Overall, the critical scenarios generated in this paper can significantly improve the testing efficiency and safety of AVs.

1. Introduction

Mining more credible and testable typical scenarios has been widely adopted in autonomous driving technology. The scenarios ensure that the test results of AVs are productive. There are two common data sources for the related studies, naturalistic driving data (NDD) and traffic accident data (TAD). The NDD contains rich operating condition information since it was well prepared before data collection. We can be extracted into two-party collision scenarios, single-vehicle-off-road scenarios, lane-changing scenarios, and human-vehicle interaction scenarios from it [1–5]. It greatly improves the path planning capability and user experience. However, The NDD has a significant sparsity feature from the perspective of risk occurrence frequency [6]. On the contrary, the TAD consists of cases

* Corresponding author. Key Lab of Forensic Science, Ministry of Justice, China (Academy of Forensic Science), Shanghai, 200063, China.
** Corresponding author. School of Automobile &Transportation, Xihua University, Chengdu, 610039, China; Sichuan Xihua Jiaotong Forensic Science Center, Chengdu, 610039, China.

E-mail addresses: tanzp@mail.xhu.edu.cn (Z. Tan), fenghao526@163.com (H. Feng).

https://doi.org/10.1016/j.heliyon.2024.e41073
Received 22 July 2024; Received in revised form 21 November 2024; Accepted 6 December 2024
Available online 12 December 2024

where hazards occurred and caused accidents, and using it to extract typical hazard scenarios can solve the sparsity problem to a certain extent. Thus, there are many studies on extracting typical hazard scenarios for TAD, as shown in Table 1. These studies can be categorized into Vehicle-to-Pedestrian(V2P) [7–11], Vehicle-to-Vehicle (V2V) [12,13], and Vehicle-to-Two-wheeler(V2TW) [14–20] according to the type of participants. In terms of clustering calculation methods, Xu et al. [13], Wang et al. [18] applied hierarchical clustering algorithms; Shu et al. [12], Sui et al. [14], Wang et al. [15], and Pan et al. [19] used K-medoids clustering algorithms, which need to specify the initial center of mass. The differences between this series of studies are mainly in the collision objects, the used scenario elements, the data collection regions, the data volume, and the data mining algorithms.

In this paper, there are two innovation points from those studies.

(1) The FCE is used to calculate the weights of scenario elements by targeting the perception and decision-making systems of the tested objects respectively, and then the weight value is put into the calculation of sample distances for cluster analysis. Some new combinations of key scenario elements are discovered for the first time.
(2) The most representative serious pedestrian accident data is selected from the National Automobile Accident In-Depth Investigation System (NAIS) in China. The size of the data is more than 1000 cases. These cases were collected in five regions of China, and all involved pedestrians who were seriously injured or fatal. Most similar studies lack this, which is important to explain the adequacy of the conclusions.

The remaining sections of this paper are organized as follows: Section 2 describes the implementation principle and realization method. Section 3 describes in detail the experimental process and the related comparative analysis process to evaluate the effectiveness of the proposed method. Section 4 discusses the differences between related studies and this paper to assess the novelty of the proposed method. Section 5 summarizes the main conclusions drawn from this paper and provides a prospective view of future research directions.

2. Method

The NAIS database contains a wide range of element categories, each with varying levels of importance depending on the specific test objectives. In the proposed scenario extraction method, the test objective was identified, and accident scenario elements were selected in a targeted manner. These elements were integrated by the characteristics of the test objective, the FCE was performed to calculate weight for each element, which was put into the clustering algorithm. The flow of the method is illustrated in Fig. 1.

Table 1
The scenario studies of recent years.

Author	Type	Scenario Element	Data Resource	Region	Data Volume	Algorithm
Qian et al. [7]	V2P	Road Type, Line of Sight, Lighting Conditions, Pedestrian Movement Type, Pedestrian & Vehicle Movement Direction	NAIS	China	220	Hierarchical Cluster
Tan et al. [8]	V2P	Pedestrian Movement Type, Lighting Conditions, Line of Sight, Weather	NAIS	China	441	Hierarchical Cluster
Lenard et al. [9]	V2P	Lighting Conditions, Weather, Line of Sight, Pedestrian & Vehicle Movement Direction, Vehicle Speed	STATS19 OTS	UK	9360 175	Hierarchical Cluster
Jeppsson et al. [10]	V2P	Line Of Sight, Pedestrian & Vehicle Movement Direction	GIDAS	German	526	Statistical Analyses
Chen et al. [11]	V2P	Line Of Sight, Pedestrian & Vehicle Movement Direction	CIDAS	China	358	Statistical Analyses
Shu et al. [12]	V2V	Collision Time, Post-Encroachment Time, and Maximum Deceleration Rate				K-medoids
Xu et al. [13]	V2V	Weather, Lighting Conditions, Intersection Type, Signal Type, Motion Type, Vehicle Type, Vehicle Speed	NAIS	China	499	Hierarchical Cluster
Sui et al. [14]	V2TW	Collision Time, Line of Sight, Pre-Crash Driving Behavior of The Car Driver, TW Driver Relative Moving Direction	CIDAS	China	672	K-medoids
Wang et al. [15]	V2TW	Time, Road Type, Road Surface, Line of Sight, Motion of Vehicle& PTW, Relative Moving Direction Of PTW& Vehicle	CIMSS-TA	China	239	K-medoids
Han et al. [16]	V2TW	Light Conditions, Weather, Road Type	NAIS	Jiangsu China	116	AHP&FCE
Jiang et al. [17]	V2TW	Weather, Lighting Conditions, Line of Sight, Collision Location, Vehicle Speed.	CIDAS	China	646	Statistical Analyses
Wang et al. [18]	V2TW	Motion Of TW, Motion of Vehicle, Road Type, Weather, Light Conditions	NAIS	Shanghai China	335	Hierarchical Cluster
Pan et al. [19]	V2TW	Collision Time, Line of Sight, Motion of Vehicle & ETW, Relative Motion Direction Between Car& ETW, ETW Type	VRU-TRAVi	Shanghai, Hunan, Fujian China	630	K-medoids
Fan et al. [20]	V2TW	TW Type, Road Conditions, Lighting Conditions, Car & TW Operation Direction	SHUFO	Shanghai China	160	Hierarchical Cluster

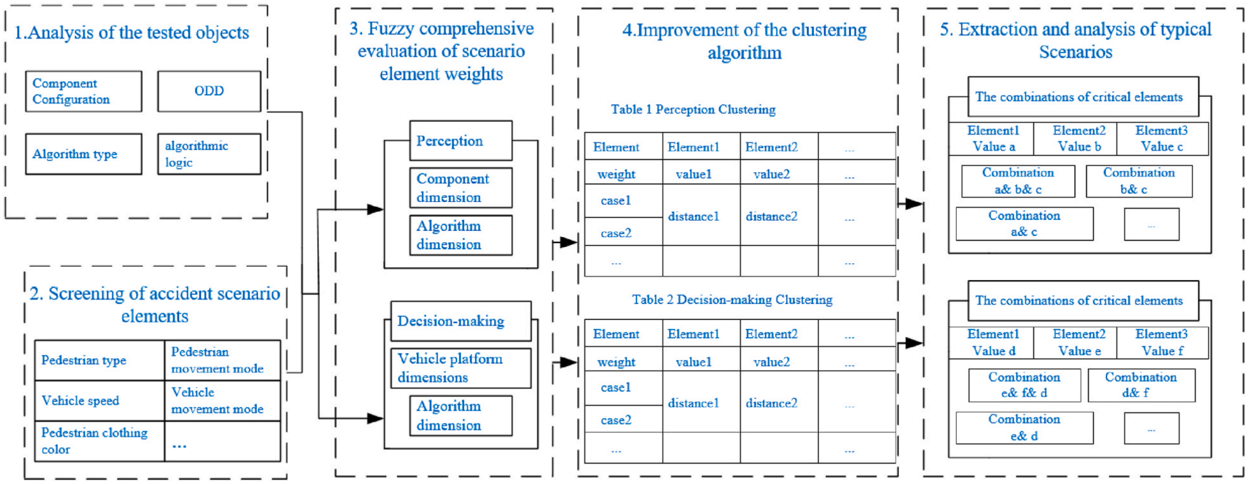


Fig. 1. The flowchart of this method.

2.1. Analysis of the tested objectives

In the research and development phase, each module is typically tested individually before integration testing. In the individual testing phase, the relevant testing scenarios for each module differ because of differences in each module's operational domain, algorithmic logic, or components. The first step of the proposed method is therefore to analyze the test objective (or module) to understand its operational characteristics, including its hardware limitations, algorithmic applications, and information uncertainties as well as the characteristics of the hardware-software combination. For instance, rain significantly impacts the performance of visual sensors compared to millimeter-wave radar sensors [21,22]. Path planning algorithms are generally unaffected by the environment, and vehicle motion control algorithms are influenced by the curvature of the roadway. Thus, identifying such features is key for qualitatively analyzing the elements of a test scenario for element screening.

2.2. Screening of accident scenario elements

Accident scenario elements should be screened on the basis of the test objective. An automated driving system (ADS) has three main systems: perception, decision-making (or planning), and motion control systems. The scenario most affects the performance of the perception and planning systems; thus, the paper focused on these two systems. For perception, the key elements are the environment, road, and target objects. For decision planning, the road, target objects, and other AVs are the most important elements. Thus, these two systems are affected by similar but slightly different elements. The perception system is often tested to identify the sensitivity influence of environmental factors, whereas tests of the decision planning system focus on the road conditions, the state of the target, and the motion of the vehicle. Some scenario elements, such as pedestrian motion, must be considered in the testing of both the aforementioned systems.

2.3. Fuzzy comprehensive evaluation of scenario element weights

After element screening, there are two main methods used in this section to obtain a weight of each element. Weights were obtained separately for the perception and planning systems. The first is the expert scoring method; the second is the FCE [23,24].

The n elements of a scenario were labeled 1, 2, ……, n. The influence of element i on the components and algorithms of the system was first evaluated using the expert scoring method. The expert scored the influence of each element as high, medium, low, or none. The evaluation matrix $R_{element\ i}$ for *element i* was obtained by combining expert opinions as Equation (1):

$$R_{element\ i} = \begin{bmatrix} c_{high} & c_{medium} & c_{low} & c_{none} \\ a_{high} & a_{medium} & a_{low} & a_{none} \end{bmatrix} \tag{1}$$

where $i = 1, 2, \ldots n$ is a subscript denoting the element and c_{high} and a_{high} denote the percentage of experts rating an element as *high* for the components and algorithms, respectively.

The fuzzy comprehensive judgment matrix was then constructed from the weight vector and evaluation matrix as Equation (2):

$$B_{element\ i} = (A_C, A_a) \times R_{element\ i} \tag{2}$$

where A_c and A_a represent the importance weights of components and algorithms in the system, respectively.

Finally, the evaluation benchmarks and corresponding level (the influence of each element as high, medium, low, or none) scores were identified to calculate the combined influence of the element, namely its weight, as Equation(3):

$$F_{element\ i} = B_{element\ i} \times V^T \tag{3}$$

where, $V = \{high\ ,\ medium\ ,\ low\ ,\ none\} = \{0.3\ ,\ 0.2\ ,\ 0.1\ ,\ 0\}$

Thus, the weight for each element i, which indicates its influence on the system, was obtained. A weight vector for the elements for each system was obtained by normalizing the weights as Equation (4):

$$f = \left(F'_{element\ 1}\ ,\ F'_{element\ 2}\ ,\ F'_{element\ 3} \cdots\cdots F'_{element\ n}\right) \tag{4}$$

2.4. Improvement of the clustering algorithm

The selected cluster analysis algorithm is based on system clustering [25,26]. The elements in the traffic accident data set were mostly obtained on nominal scales [27]; thus, the distance between samples was calculated as the Hamming distance [28,29]. The similarity between two samples is represented as Equation (5):

$$H = \left(h_{element\ 1}\ ,\ h_{element\ 2}\ ,\ \ldots\ldots h_{element\ k}\ ,\ \ldots\ldots h_{element\ n}\right) \tag{5}$$

where $h_{element\ k} = \begin{cases} 0, x_{ik} = x_{jk} \\ 1, x_{ik} \neq x_{jk} \end{cases}$, x_{ik} is the value of the kth element for the ith sample, and x_{jk} is the value of the kth element for the jth sample.

The distance between two samples was quantified using Equation (6), which includes the element weights obtained in Section 2.3. This process was looped to obtain the distance between all samples as follows:

$$d_{ij} = \left(H \times f^T\right) / n \tag{6}$$

where n denotes the number of scenario elements.

The inter-class distance was calculated using the group-averaging method [30], as expressed in Equation (7). On the basis of the calculated inter-class distances, the two closest classes were combined into a new class. This process was repeated until the desired number of classes was obtained. Classes G_K and G_L contained n_K and n_L samples, respectively.

$$D_{KL}^2 = \sum_{x_i \in G_K, x_j \in G_L} d_{ij}^2 \Big/ n_K n_L \tag{7}$$

2.5. Extraction and analysis of typical scenarios

In the proposed improved clustering algorithm, elements with larger weights have a greater effect on the calculated distance, thus facilitating the extraction of combinations of scenario elements that are more valuable for the test objective. The typical characteristics of the elements were further extracted using the chi-square test [31,32]. Typical scenarios were extracted with the improved and

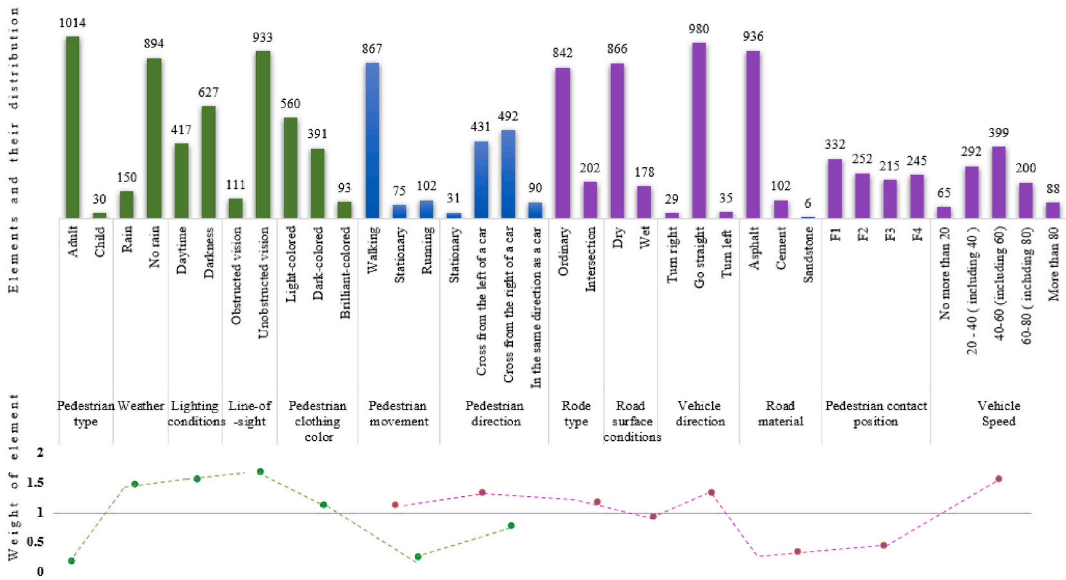


Fig. 2. The distribution and weight of the elements involved in the perception and decision-planning systems.

unimproved clustering methods (i.e., without weights), and the results were compared to verify the effectiveness of the proposed method. Finally, the scenarios identified with the proposed method were used to construct test programs.

3. Results

3.1. Results of typical-scenario extraction

V2P accident data in the NAIS database were analyzed. This database includes 1044 cases of pedestrian fatality from numerous regions in China. Scenario elements were screened for the perception and planning systems, and their weights were evaluated through FCE. A description of the element information is shown in Fig. 2. The scenario elements and their distributions are represented by green bar charts for the perception system, purple bar charts for the decision-planning system, and blue bar charts for both systems. The points in the bottom-line graphs indicate the corresponding element weights.

The seven elements screened for the perception system included pedestrian type (e.g., child or adult), weather, lighting conditions, line of sight, pedestrian clothing color, pedestrian movement (e.g., running or walking), and pedestrian direction. Among these, weather, lighting conditions, pedestrian clothing color, and line of sight were assigned a weight greater than 1. For the planning system, eight elements were considered: pedestrian movement, pedestrian direction, road type (e.g., intersection or standard road), road surface conditions, vehicle speed, road material, pedestrian contact position (*Fi*, where $i = 1, 2, 3, 4$, represents the collision area sorted from the distal to proximal end of the driver's seat), and vehicle direction. Of these elements, vehicle speed, vehicle direction, road type, pedestrian movement, and pedestrian direction were assigned weights greater than 1.

The improved clustering algorithm was implemented using a program written in MATLAB (2018b). The scenario extraction results are presented in Tables 2 and 3.

In Tables 2 and 3, some categories have multiple values for one element; for example, pedestrian may cross the road from the right and left sides, this situation is represented by symmetrically opposite arrows in illustrations (e.g., the illustration of Scenarios 3). This can be broken down into two specific scenarios (Scenario 1, pedestrians cross the road from the right; Scenario 2, pedestrians cross the road from the left). Therefore, the nine scenarios for the perception and planning systems can be subdivided into 16 and 48 specific scenarios, respectively.

In summary, elements with weights of >1 performed better during cluster mining than did elements with weights of <1. Thus, focusing on elements with higher weights is critical.

3.2. Evaluation of the extraction results

This section presents the results for the improved and unimproved clustering algorithms. Many data sets contain imbalanced data, which might result in the clustering method failing to identify some hazardous scenarios. For example, only 14.4 % of the accidents in the examined data set occurred during rain; however, rain is known to be hazardous. High-influence factors, such as rain, are defined as critical factors. Figs. 3 and 4 present comparisons of the statistical results for scenarios involving the critical factors; critical-factor scenarios comprised a higher percentage of all scenario combinations in the results obtained using the improved clustering method than in those obtained using the unimproved clustering method. In this paper, the average is also used to describe the percentage improvement of critical scenarios. For perception, "invisible in the dark" has an increase of 10.7 %, "rain" has an increase of 33.9 %, and "obstructed vision" has an increase of 7.1 %, while the average percentage of the critical scenarios increased by 17.4 %. The decrease in the proportion of "darkness" is in line with reality. From Fig. 2, we know that more traffic accidents occur at night than in the daytime, thus the change in the proportion of the two should be relative after weighting. For decision-making, the share of "vehicle turning" increased by 30 %, "high speed (over 80 km/h)" by 7.2 %, and "intersections" by 3.7 %. The average increase in the percentage of the critical scenarios is 13.6 %. The proportion of "pedestrians running" and "pedestrians crossing the road" scenarios varied similarly to the "darkness" scenario.

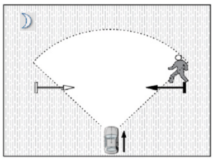
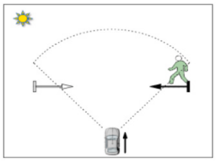
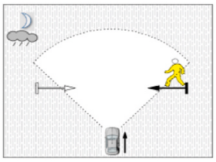
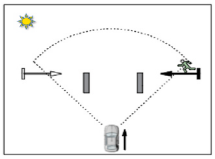
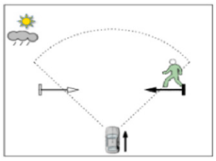
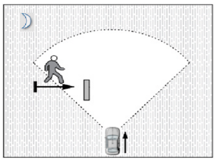
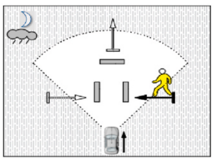
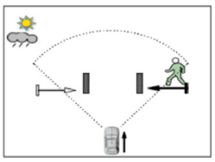
The combinations of critical factors in the scenarios obtained with the improved and unimproved clustering algorithms can be presented as a two-dimensional matrix (Fig. 5). The rows and columns of the matrix for perception are dark-colored clothing, rain, darkness, and obstructed vision; the green and orange bars in Fig. 5 indicate the percentage of each critical-factor combination for the scenarios obtained with both methods. For example, the combination of darkness and dark-colored clothing was 11.6 % more common in the results obtained with the improved method than in those obtained with the unimproved method. Overall, the scenarios extracted using the improved clustering method contained a higher percentage of critical-factor combinations than did those extracted using the unimproved method.

For the decision planning system, the critical factors were vehicle turning, high speed (above 80 km/h), road intersection, pedestrian running, and pedestrian crossing the road (from the left or right). As shown in Fig. 6, compared with the unimproved method, the improved approach obtained a higher percentage of critical-factor combinations across all scenarios, which indicated its higher performance for the decision planning system.

3.3. Test program construction

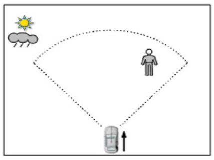
Building on the proposed clustering method, a test program based on a special hierarchical clustering approach was developed to optimize testing resources by significantly enhancing testing efficiency. In this scheme, scenario elements with weights of >1 were placed in the first level of the hierarchy (bold elements in Table 4), and other elements were placed in the second level. The elements

Table 2
Example scenarios for the perception system.

S.	Illustrations	Detailed description
1		Adult walking from the left or right side of a car, no rain, darkness, dark-colored clothing, unobstructed vision, involved in an accident with a car.
2		Adult walking from the left or right side of a car, no rain, daytime, light-colored clothing unobstructed vision, involved in an accident with a car.
3		Adult walking from the left or right side of a car, rain, darkness, brilliant-color clothing, unobstructed vision, involved in an accident with a car.
4		Child running from the left or right side of a car, no rain, daytime, light-colored clothing, obstructed vision, involved in an accident with a car.
5		Adult walking from the left or right side of a car, rain, daytime, light-colored clothing, unobstructed vision, involved in an accident with a car.
6		Adult walking from the left side of a car, no rain, darkness, dark-colored clothing, obstructed vision, involved in an accident with a car.
7		Adult walking from the left or right side of or in the same direction as a car, rain, darkness, brilliant-color clothing, obstructed vision, involved in an accident with a car.
8		Adult walking from the left or right side of a car, no rain, daytime, light-colored clothing, obstructed vision, involved in an accident with a car.

(continued on next page)

Table 2 (*continued*)

S.	Illustrations	Detailed description
9		Adult standing on the side of the road, rain, daytime, dark-colored clothing, unobstructed vision, involved in an accident with a car.

were then ranked on the basis of the difficulty in implementing the relevant condition in a test. Clustering was then used to group test cases with similar element values in each class. For example, for the perception system, environmental elements, such as lighting and weather, were grouped into four categories: daytime with rain, daytime without rain, darkness with rain, and darkness without rain (Table 4). Accordingly, the test program was completed based on the weights of the test elements and the difficulty of arranging the test elements. For further scenario testing, a base test scenario can be identified for testing and other test scenarios in the same group are varied from the base test scenario. The test scenarios for the planning system are listed in Table 5. The Category 3 scenarios were differentiated into Scenarios 3(1) and 3(2) for hierarchical clustering because the road type differed.

4. Discussion

This paper is a continuation of the data mining work which author extracted typical pedestrian accident scenarios in 2021 [8]. There are two improvements being implemented. For one thing, the data size of this paper has increased to more than 1000 cases from in-depth investigation of accidents. For another, the execution process of the clustering algorithm has been adjusted in a directional way from the perspective of the testing needs of objects. There are some important findings in this paper after comparing the findings with similar studies as follows.

4.1. Relation to existing findings

4.1.1. Method innovation

The Hierarchical Ascending [9,13,18] and the K-medoids [14,19,33] are two of the most common algorithms in the study of mining typical scenarios using cluster analysis. The differences between the two methods are mainly reflected in the selection of sample distance and the determination of the number of categories. However, we used the same data and adopted different methods to practice, and the results do not show obvious advantages and disadvantages of the methods. It shows that any traditional algorithm of cluster analysis can reflect the statistical characteristics of data effectively. As a result, there are few researches on how to improve the clustering algorithm to better meet the needs of data mining. Of these few researches, Xia et al. [34] and Shu et al. [12] are the more classic ones. They both used analytic hierarchy process (AHP) to assign weights to scenario parameters and brought the obtained weights into the calculation of sample distances. The main innovation in the method of this paper is reflected in the FCE, which assigns weights to scenario elements from two aspects of perception and decision, which reflects the sensitivity of the system to different scenario elements. In the algorithm, this paper uses the Hamming distance, which makes the sample distance solution simple and reliable when processing data with nominal scale elements. In addition to these, this paper validates the effectiveness of the method.

4.1.2. Important finding

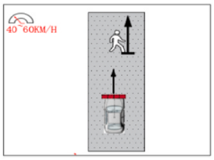
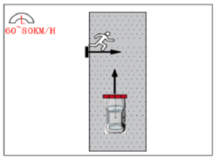
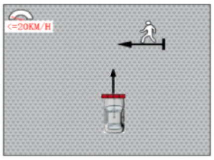
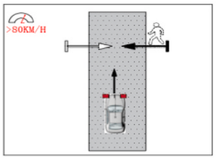
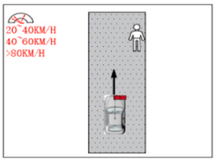
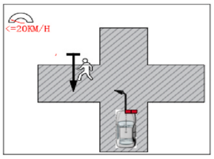
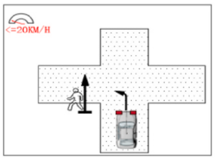
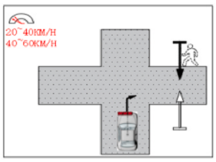
Firstly, in terms of the types of scenarios mined. The important finding of this paper is that some combinations of the critical scenario elements have been extracted for the first time, compared to other researches [9–11]. For perception recognition and decision planning, the combined scenarios of the critical elements are extracted, such as vision obstruction, darkness, and rain; high speed, pedestrians crossing the road, and pedestrians running. There are of course some scenarios mined in this paper that are the same as other similar studies, such as darkness and rain; vehicle turning and pedestrian crossing the road.

Second, in terms of the volume of critical scenarios mined. In this paper, the scenarios that contain critical elements are a larger percentage of the extracted scenarios. For example, the "vision obstruction and vehicle turning" and "pedestrian crossing the road" elements were present in 22.2 % and 44.4 %, respectively, of the extracted scenarios in the study of Jeppsson et al. [10] but 50.0 % and 45.8 %, respectively, of the scenarios extracted in this paper. The "vision obstruction and darkness" and"vehicle turning and pedestrian crossing the road"combinations accounted for 7.1 % and 14.3 % of the scenarios identified by Chen et al. [11]; in our paper, the corresponding proportions were 25.0 % and 45.8 %, respectively. Consequently, the scenarios obtained by this new scenario extraction method are more challenging for the testing of AVs, which can increase the difficulty of testing to a certain extent, and improve the testing efficiency and the safety of vehicles.

4.2. Limitations and future work

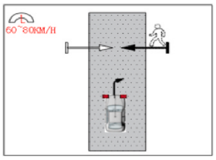
One of the limitations of this paper is the acquisition of scenario element weights mainly relies on experts' subjective evaluations since there is a lack of experimental data on the sensitivity of the perception and decision-making systems to scenario elements. While

Table 3
Example scenarios for the planning system.

S.	Illustrations	Detailed description
1		Dry asphalt surface, ordinary road, straight-moving car, pedestrian walking in the same direction as the car, front-end collision, speed of 40–60 km/h.
2		Dry asphalt surface, ordinary road, straight-moving car, pedestrian running from the left side, front-end collision, speed of 60–80 km/h.
3		Dry sandstone surface, unknown road type, straight-moving car, walking pedestrian crossing from the right side, front-end collision, speed of <20 km/h.
4		Dry asphalt surface, ordinary road, straight-moving car, walking pedestrian crossing from the left or right, collision at F1 or F4, speed of >80 km/h.
5		Dry asphalt surface, ordinary road, straight-moving car, static pedestrian in conflict with the car, collision at F1 or F2, speed of 20–40, 40–60, or 40–80 km/h.
6		Dry cement surface, intersection, left-turning car, walking pedestrian crossing from the right side, collision at F1 or F2, speed of <20 km/h.
7		Asphalt surface of unknown wetness, intersection, left-turning car, walking pedestrian crossing from the left side, collision at F1 or F4, speed of <20 km/h.
8		Dry asphalt surface; intersection; left-turning car; walking pedestrian crossing from the left or right side; collision at F2, F3, or F4; speed of 20–40 or 40–60 km/h.

(*continued on next page*)

Table 3 (*continued*)

S.	Illustrations	Detailed description
9		Dry asphalt surface, ordinary road, right-turning car, walking pedestrian crossing from the left or right side, collision at F1 or F4, speed of 60–80 km/h.

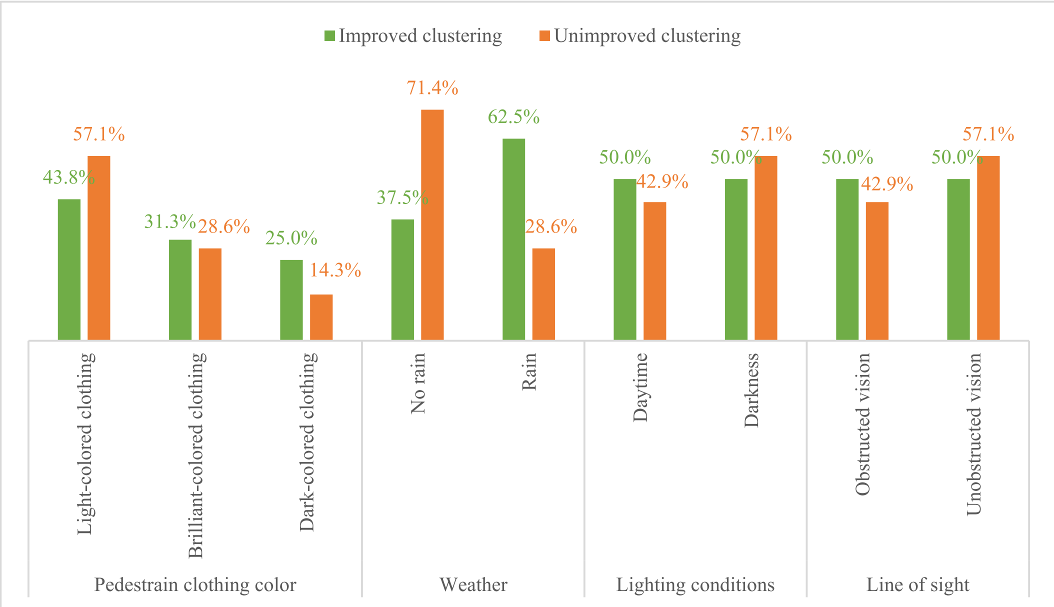


Fig. 3. Distribution of critical elements for the perception system.

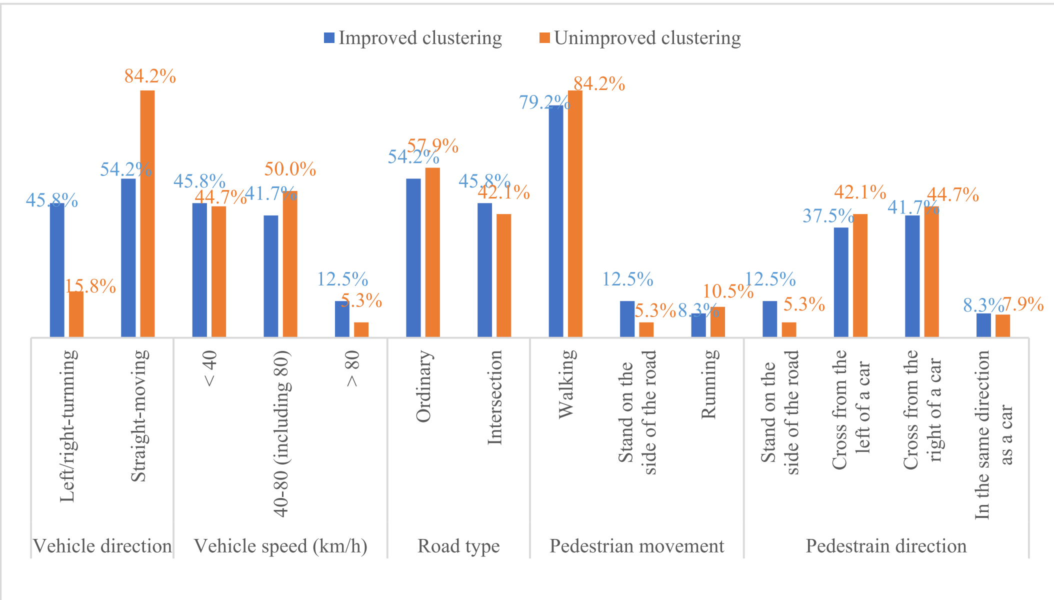


Fig. 4. Distribution of critical elements for the decision planning system.

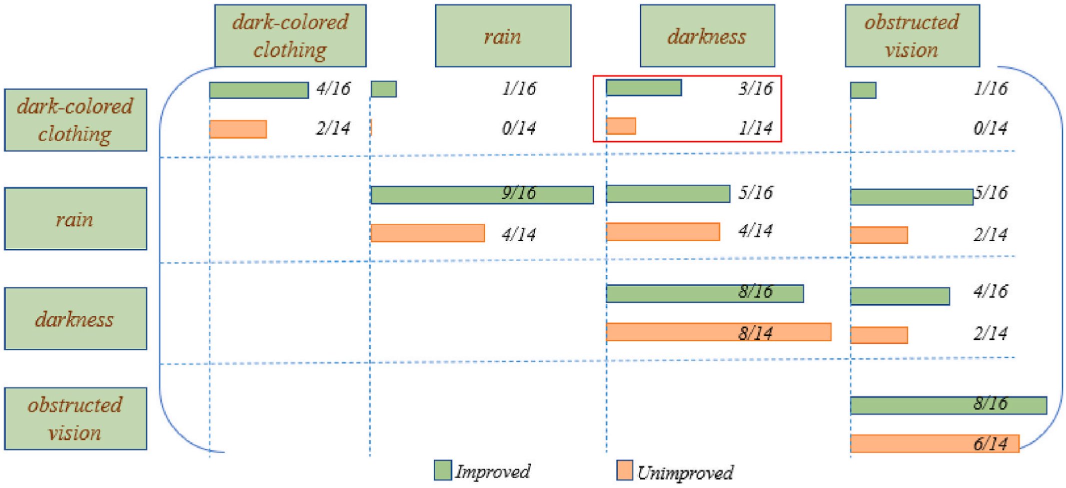


Fig. 5. Frequency of each critical-factor combination for the perception system in the scenarios identified by the improved and unimproved clustering algorithms.

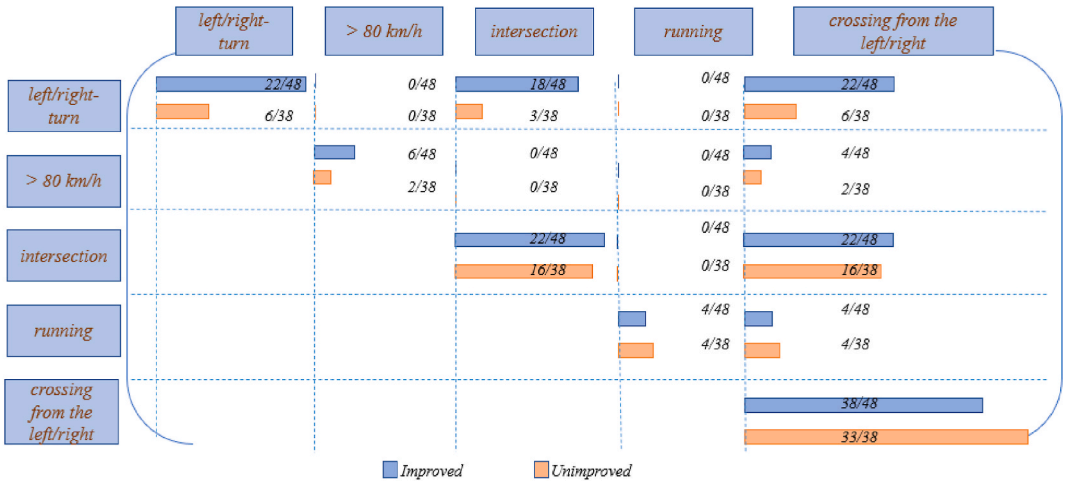


Fig. 6. Frequency of each critical-factor combination for the decision planning system in the scenarios identified by the improved and unimproved algorithms.

Table 4
Proposed test conditions for the perception system[a].

Var. \ S.	5	8	9	2	4	1	6	3	7
lighting condition	daytime	daytime	daytime	daytime	daytime	darkness	darkness	darkness	darkness
weather	rain	rain	rain	no rain	no rain	no rain	no rain	rain	rain
line of sight	unobstructed vision	**obstructed vision**	unobstructed vision	unobstructed vision	**obstructed vision**	unobstructed vision	**obstructed vision**	unobstructed vision	**obstructed vision**
pedestrian clothing color	light-colored	light-colored	**dark-colored**	light-colored	light-colored	dark-colored	dark-colored	brilliant-colored	brilliant-colored
pedestrian direction	left/right	left/ right	***stationary***	left/right	left/right	left/right	***left***	left/right	***left/right /in same direction***
pedestrian movement	walking	walking	***stationary***	walking	***running***	walking	walking	walking	walking
pedestrian type	adult	adult	adult	adult	***child***	adult	adult	adult	adult

[a] The different colors in the tables are used to indicate the different groups of test scenarios.

Table 5
Proposed test conditions for the planning system.

Var. \ S.	1	4	2	3（1）	6	7	8	5	9	3（2）
rode type	ordinary	ordinary	ordinary	ordinary	intersection	intersection	intersection	ordinary	ordinary	intersection
Vehicle direction	straight-moving	straight-moving	straight-moving	straight-moving	left -turn	left -turn	left -turn	right-turn	right-turn	straight-moving
pedestrian movement	walking	walking	**running**	walking	walking	walking	walking	stationary	**walking**	walking
pedestrian direction	in same direction of a car	**left/right**	**left**	**right**	right	**left**	**left/right**	stationary	**left/ right**	right
Vehicle speed (km/h)	40-60	**>80**	**60-80**	**<20**	<20	<20	**20-60**	20-60/ >80	**60-80**	<20
road surface	dry	dry	dry	dry	dry	***dry/wet***	dry	dry	dry	dry
road material	asphalt	asphalt	asphalt	***sandstone***	cement	***asphalt***	***asphalt***	asphalt	asphalt	sandstone

the paper employs FCE from the component and algorithmic dimensions of the system, more meaningful conclusions may be obtained if the subjective and objective evaluations are combined in subsequent studies. Another limitation of this paper is that this method has been validated in pedestrian accident data, and has not been validated for V2TWs, V2V, which is a direction for subsequent generalization of this method. Meanwhile, it is also a direction of subsequent work to fully consider the characteristics of the data such as imbalance and heterogeneity, and then mining scenarios with more targeted and testing value.

5. Conclusion

We thoroughly analyzed the characteristics of AV perception and planning systems to identify relevant scenario elements. FCE was used to assess the influence of each selected element on the test system. This influence was quantified as a weight, which was subsequently incorporated into an improved clustering analysis algorithm. This method was applied to extract typical scenario data from 1044 pedestrian accidents in China, resulting in the identification of nine categories of critical scenarios for both the perception and decision-planning systems. The experimental results indicate that, compared to traditional methods, the proportion of critical scenarios among all acquired scenarios increased by 17.4 % and 13.6 %, respectively, when using the new method. Additionally, the new method extracted critical scenario combinations for perception and decision-planning systems, including factors such as vision obstruction, darkness, and rain; and scenarios involving high-speed vehicles, pedestrians crossing the road, and pedestrians running. Specifically, the proposed method identified more challenging scenarios with combinations of critical factors than the unimproved clustering method. By incorporating these scenarios into AV testing, the vehicle testing efficiency can be significantly enhanced.

However, there are limitations to our paper. One of the limitations of this paper is the acquisition of scenario element weights mainly relies on experts' subjective evaluations since there is a lack of experimental data on the sensitivity of the perception and decision-making systems to scenario elements. Another limitation of this paper is that this method has been validated in pedestrian accident data, and has not been validated for V2TWs, V2V. In response to these limitations, we have several suggestions for improvement in our future work. The first point is to combine expert experience and experimental data for subjective and objective evaluation in the process of calculating element weights. The second point extends the study to other types of accidents. The third point is to fully consider the characteristics of data, such as imbalance and heterogeneity. These improvements may enable us to extract new scenarios that are more valuable and challenging.

CRediT authorship contribution statement

Zhengping Tan: Writing – review & editing, Supervision, Resources, Methodology, Conceptualization. **Qian Wang:** Writing – original draft, Software, Methodology. **Wenhao Hu:** Writing – review & editing. **Pingfei Li:** Investigation. **Liangliang Shi:** Investigation. **Hao Feng:** Writing – review & editing, Methodology.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Consent for publication

All authors have given consent for publication.

Funding statement

This study was a collaborative effort supported by the National Key R&D Program of China (2023YFC3009703), the Science and

Technology Project (Development and Application of 5G+V2X Module and Risk Warning Terminal for Intelligent Transportation (2022YFG0095), Precrash Module Information Extraction and Transformation Based on Vehicle Accident Data (H232270), Accident Precrash Scenario Transformation Toolchain and Database (H231291), Intelligent Vehicle Traffic Accident Scenario Reconstruction and Its Safety Identification Application Research (KF202211)).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J.L. Duan, R.J. Li, L. Hou, W.J. Wang, G.F. Li, S.B. Eben Li, H.B. Gao, Driver braking behavior analysis to improve autonomous emergency braking systems in typical Chinese vehicle-bicycle conflicts, Accid. Anal. Prev. (2017) 74–82, https://doi.org/10.1016/j.aap.2017.08.022.
[2] J.H. Guo, K.Q. Li, J. Wang, Prediction of random motion state of front vehicle based on clustering analysis of dangerous scenarios, Automot. Eng. 42 (7) (2020) 847–853+859, https://doi.org/10.19562/j.chinasae.qcgc.2020.07.001.
[3] J.Y. Ma, Y.F. Tian, X.M. Zhao, Z. Wang, Y.Q. Liu, Second-order vehicle-pedestrian interaction test scenarios based on natural driving data mining, China J. Highw. Transp. 35 (3) (2022) 139–152, https://doi.org/10.3969/j.issn.1001-7372.2022.03.013.
[4] D. Nilsson, M. Lindman, T. Victor, M. Dozza, Definition of run-off-road crash clusters—for safety benefit estimation and driver assistance development, Accid. Anal. Prev. 113 (2018), https://doi.org/10.1016/j.aap.2018.01.011.
[5] L. Zhao, T. Xu, Z.S. Zhang, Y.J. Hao, Lane-changing recognition of urban expressway exit using natural driving data, Appl. Sci. 12 (19) (2022), https://doi.org/10.3390/app12199762.
[6] S. Feng, Y.H. Feng, C.H. Yu, Y. Zhang, H.X. Liu, Testing scenario library generation for connected and automated vehicles, Part I: methodology, IEEE Trans. Intell. Transport. Syst. 3 (2020) 1–10, https://doi.org/10.1109/TITS.2020.2972211.
[7] Y.B. Qian, Y.C. Qiu, L.Y. Xiao, W.H. Hu, H.L. Dong, Research on test scenarios of AEB pedestrian system based on knowledge and accident data, Int. J. Veh. Saf. 12 (3–4) (2022), https://doi.org/10.1504/IJVS.2022.129627.
[8] Z.P. Tan, Y.Y. Che, L.Y. Xiao, W.H. Wu, P.F. Li, J. Xu, Research of fatal car-to-pedestrian precrash scenarios for the testing of the active safety system in China, Accid. Anal. Prev. 150 (2021), https://doi.org/10.1016/j.aap.2020.105857.
[9] J. Lenard, A. Badea-Romero, R. Danton, Typical pedestrian accident scenarios for the development of autonomous emergency braking test protocols, Accid. Anal. Prev. 73 (2014) 73–80, 0.1016/j.aap.2014.08.012.
[10] H. Jeppsson, M. Östling, N. Lubbe, Real life safety benefits of increasing brake deceleration in car-to-pedestrian accidents: simulation of Vacuum Emergency Braking, Accid. Anal. Prev. 111 (2018) 311–320, 0.1016/j.aap.2017.12.001.
[11] Q. Chen, M. Lin, B. Dai, J.G. Chen, Typical pedestrian accident scenarios in China and crash severity mitigation by autonomous emergency braking systems, in: SAE 2015 World Congress & Exhibition, Society of Automotive Engineers, Michigan, USA, 2015, pp. 10495–10504, https://doi.org/10.4271/2015-01-1464.
[12] H. Shu, Z.Y. Wang, S. Li, Construction of key scenarios for autonomous vehicles going straight at intersections without signal lights, China J. Highw. Transp. 35 (7) (2022) 328–338, https://doi.org/10.3969/j.issn.1001-7372.2022.07.028.
[13] X.Y. Xu, Z.H. Zhou, W.H. Hu, AEB intersection test scenario based on accident data mining, J. Beijing Univ. Aeronaut. Astronaut. 46 (10) (2020) 1817–1825, https://doi.org/10.13700/j.bh.1001-5965.2019.0548.
[14] B. Sui, N. Lubbe, J. Bärgman, A clustering approach to developing car-to-two-wheeler test scenarios for the assessment of Automated Emergency Braking in China using in-depth Chinese crash data, Accid. Anal. Prev. (2019) 105242, https://doi.org/10.1016/j.aap.2019.07.018.
[15] X.H. Wang, Y. Peng, T. Xu, Q. Xu, X.H. Wu, G.L. Xiang, S.G. Yi, H.G. Wang, Autonomous driving testing scenario generation based on in-depth vehicle-to-powered two-wheeler crash data in China, Accid. Anal. Prev. 176 (2022), https://doi.org/10.1016/j.aap.2022.106812.
[16] D.S. Han, Z.X. Ma, X.C. Zhu, Y.F. Zeng, Car-cyclist accident scenario analysis for autonomous vehicles, Journal of Automotive Safety and Energy 11 (2) (2020) 220–226, https://doi.org/10.3969/j.issn.1674-8484.2020.02.009.
[17] L.J. Jiang, J.Y. Liu, Y.R. Zhang, L. Wang, D.Z. Wang, Research on typical scenarios of two-wheeler accidents in China, in: Proceedings of the 19th Asia-Pacific Annual Conference on Automotive Engineering and the 2017 Annual Conference of the China Society of Automotive Engineering, Technical Center of Shanghai Automotive Group Corporation, 2017, pp. 825–829.
[18] R. Wang, Y.B. Qian, H.L. Dong, W.P.F. Yu, Data mining-based collision scenarios of vehicles and two wheelers for the safety assessment of intelligent driving functions, World Electric Vehicle Journal 14 (10) (2023), https://doi.org/10.3390/wevj14100284.
[19] D. Pan, Y. Han, Q.Q. Jin, H. Wu, H.W. Huang, Study of typical electric two-wheelers pre-crash scenarios using K-medoids clustering methodology based on video recordings in China, Accid. Anal. Prev. 160 (2021), https://doi.org/10.1016/j.aap.2021.106320.
[20] T.C. Fan, H.Y. Wang, Z. Yang, Based on the typical scenario extraction and analysis of two-wheeler accidents in China, Traffic & Transportation 35 (2) (2019) 41–45, https://doi.org/10.3969/j.issn.1671-3400.2019.02.010.
[21] T.C. Meng, J. Huang, C.M. Chew, D.G. Yang, Z.H. Zhong, Configuration and design schemes of environmental sensing and vehicle computing systems for automated driving: a review, IEEE Sensor. J. 23 (14) (2023) 15305–15320, https://doi.org/10.1109/jsen.2023.3260193.
[22] D. Watzenig, M. Horn, Automated Driving: Safer and More Efficient Future Driving, Springer, 2016, https://doi.org/10.1007/978-3-319-31895-0.
[23] S.Z. Hu, J.Q. Gu, Fuzzy comprehensive evaluation method and its application, J. Donghua Univ. 21 (1) (1995) 74–80. DOI: cnki: sun: dhdz.0.1995-01-010.
[24] Z. Xue, X. Zeng, L. Koehl, Development of a method based on fuzzy comprehensive evaluation and genetic algorithm to study relations between tactile properties and total preference of textile products, J. Text. Inst. 108 (7) (2017) 1085–1094, https://doi.org/10.1080/00405000.2016.1263142.
[25] F. Murtagh, P. Contreras, Algorithms for hierarchical clustering: an overview, II, Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery 7 (6) (2017), https://doi.org/10.1002/widm.1219.
[26] H. Sun, Business data analysis based on hierarchical clustering algorithm in the context of big data, J. Phys. Conf. 1744 (4) (2021), https://doi.org/10.1088/1742-6596/1744/4/042135.
[27] H.B. Xu, D.B. Bo, S. Bai, An optimized clustering algorithm for nominal scale variables, Microelectron. Comput. (12) (2003), https://doi.org/10.3969/j.issn.1000-7180.2003.12.003.
[28] A. Harshitha, S. Nayak, S. D'Souza, P.G. Bhat, Hamming index of the product of two graphs, Eng. Lett. 30 (3) (2022).
[29] V. Singh, C. Kumar, Improving hamming-distance computation for adaptive similarity search approach, Int. J. Intell. Inf. Technol. 18 (2) (2022), https://doi.org/10.4018/ijiit.296270.
[30] N. Muda, A.R. Othman, Modified distance in average linkage based on M-estimator and MADn criteria in hierarchical cluster analysis, AIP Conf. Proc. 1682 (1) (2015), https://doi.org/10.1063/1.4932506.
[31] C.C. Shen, P. Sambit, T. J. V. The chi-square test of distance correlation, J. Comput. Graph Stat. 31 (1) (2022), https://doi.org/10.1080/10618600.2021.1938585.

[32] H. Wang, C. Chen, Comparison of chi-square test and representative decision tree in features that influence vehicle style, International Journal of Machine Learning and Computing 11 (5) (2021), 10. 18178/ijmlc.2021. 11.5.1057.
[33] P. Nitsche, P. Thomas, R. Stuetz, R. Welsh, Pre-crash scenarios at road junctions: a clustering method for car crash data, Accid. Anal. Prev. 107 (2017), https://doi.org/10.1016/j.aap.2017.07.011.
[34] Q. Xia, J.L. Duan, F. Gao, T. Chen, C. Yang, Automatic generation method of test scenario for ADAS based on complexity, in: SAE 2017 Intelligent and Connected Vehicles Symposium, Society of Automotive Engineers, Suzhou, China, 2017, pp. 194–202, https://doi.org/10.4271/2017-01-1992.