Commun Transp Res

Research Article

Complexity controllable road network generation for virtual testing of autonomous driving

Yu Zhu[1], Jiaxin Wang[1], Shaoxin Yuan[1]✉, Zhigang Xu[1]✉, Xiaobo Qu[2]

Cite this article: https://doi.org/10.26599/COMMTR.2026.9640024

ABSTRACT: Complexity controllable road network generation is pivotal for accelerated virtual simulation testing of autonomous vehicles (AVs), enabling the implementation of progressively challenging test scenarios through incrementally complex road networks, facilitating continuous evaluation of the target autonomous driving algorithms to expose functional flaws and performance boundaries. To this end, this study proposes a complexity controllable road network generation approach via optimized combination of realistic road elements. First, real-world urban road networks are decomposed into diverse road elements. Among these, elements with high potential collision risks (e.g., T-junctions, merging/diverging zones) are abstracted into parameter-configurable graph models defined by node and edge parameters. Then, the instantiation of these road element models is achieved using real-world cartographic data, followed by complexity assessment via a metric quantifying potential collision risk. Concurrently, a complexity evaluation function for road elements is formulated to assign complexity labels to each road instance. A complexity-tunable road network optimization model is developed. Taking the realism and compactness of the generated road networks as optimization objectives, the model selects road elements instance of varying complexity for optimal assembly in a non-intersecting and non-overlapping manner, yielding virtual road networks with controllable complexity while preserving real-world characteristics. Finally, experimental validation was conducted using Xi'an cartographic data, generating 4,500 virtual road networks across 15 distinct complexity levels through combination of road elements with different complexity. To validate the effectiveness of the generated road networks, comparative virtual simulation experiments are conducted on both the generated networks and real-world road networks. Experimental results demonstrate that: (1) The mean relative error between the geometric parameters of road elements in the generated networks and the expected values derived from real-world cartographic data clustering is less than 1.5%, confirming superior environmental reconstruction fidelity; (2) Lane-change test scenarios constructed using the generated road networks successfully identified performance limits of the target autonomous driving algorithm and functional deficiencies under lateral approach conditions; (3) The complexity of the generated road networks shows a strong correlation (absolute correlation coefficients exceeding 0.84) with the uncomfortable driving duration and average vehicle speed in continuous free-driving scenarios, which confirms that higher-complexity road networks correspond to more rigorous testing challenges; (4) Compared to real-world road networks, the generated road networks achieve 32.4% average mileage compression when traversing an equivalent number of road elements, significantly enhancing testing efficiency.

KEYWORDS: autonomous driving; virtual testing; road networks generation; complexity controllable.

1 Introduction

The accelerating development of high-level automated driving functionalities has rendered the replacement of human drivers by automated driving in specific dynamic driving tasks from theoretical possibility to operational reality (Xu et al., 2018; Liu et al., 2024; Xu et al., 2025). Yet the rapid advancements introduce increasingly severe challenges on testing and validation. Recent open-road trials revealed that despite undergoing relatively comprehensive testing, commercially mature autonomous vehicles (AVs) still exhibit significant performance limitations in typical scenarios (Opletal, 2025). Yet the rapid advancements introduce increasingly severe challenges on testing and validation. Recent open-road trials revealed that despite undergoing relatively comprehensive testing, commercially mature autonomous vehicles (AVs) still exhibit significant performance limitations in typical scenarios (Opletal, 2025). The infeasibility of on-road testing that requires millions of kilometers to achieve statistical significance has been proven (Kalra & Paddock, 2016; Xu et al., 2017). Virtual simulation testing substantially enhances testing efficiency (Sang et al, 2024), but it requires the creation of challenging test scenarios to pinpoint functional blind spots and performance limitations.

A test scenario is formally defined as the holistic integration of static road environments and dynamic interactions among traffic participants (Ulbrich et al., 2015; Elrofai et al., 2016; De Gelder & Paardekooper, 2017). The road environment imposes deterministic constraints on movable participants and their spatial-temporal relationships (Weber et al., 2019; Chen et al., 2021). Despite the growing sophistication of dynamic traffic modeling, the road environment construction remains an underdeveloped aspect of test scenario generation (Yang et al., 2024). Traditional approaches relies on labor-intensive manual modeling (Zhu et al., 2022; Zhu et al., 2024), which is incapable of generating road networks for continuous multi-scenario testing, while precluding the requirements for quantitative iterative adjustments of road environments.

[1] School of Information Engineering, Chang'an University, Xi'an 710064, China. [2] School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China.

✉ Corresponding author. E-mail: S. Yuan, shxyuan@chd.edu.cn; Z. Xu, xuzhigang@chd.edu.cn

Received: February 8, 2026; Revised: April 23, 2026; Accepted: April 23, 2026, 2026

Tsinghua University Press | SciOpen

Current mainstream automated road network generations are primarily categorized into the data-driven approach and procedural content generation (PCG) -based approach. Data-driven approaches enable the generation of road networks, achieving alignment with real-world statistics and perceptual realism (Muktadir et al., 2022). It exploits multi-modal real-world information that is extracted from digital maps (Nishida et al., 2016), vehicle GPS trajectories (Li et al., 2022), unmanned aerial vehicle-acquired data (Kong et al., 2023), and satellite remote sensing imagery (Zhang et al., 2019; Zao et al., 2024). Data-driven approaches fall into two generation paradigms. The direct paradigm extracts and geometrically deforms road elements from real world to assemble novel road networks (Aliaga et al., 2008). The alternative paradigm involves an end-to-end road network generation (Hartmann et al., 2017), discovering latent spatial patterns in real world to synthesize topologically coherent road networks.

However, real-world road networks contain abundant non-conflict elements (e.g., repetitive straightaways), which leads to inflation of test mileage, reducing testing efficiency.

Decoupling from data dependencies, PCG-based approaches synthesize road networks through algorithmically defined rules that interconnect manually designed road elements (Muktadir et al., 2022). Related studies (Beneš et al., 2014; Ikram et al., 2023; Tang et al., 2023; Yang et al., 2024) systematically elaborated on generating road elements, and proposed a endpoint connection strategy to ensure geometric continuity between contiguous elements (Aliaga et al., 2008). To mitigate unintended intersections and overlaps during assembly, block incremental generation (BIG) with backtracking mechanism emerges as the dominant paradigm (Gambi et al., 2019a; Li et al., 2021). It executes an iterative workflow "block (element) addition – conflict detection – conflict resolution". Conflict resolution leverages a backtracking mechanism. Upon encountering road elements intersections or overlaps, the conflicting road elements are substituted with freshly generated ones until a valid road is produced.

The backtracking mechanism suffers from inherent scalability constraints. As network scale increases, the probability of unintended intersections and overlaps during expansion rises exponentially, leading to a surge in backtracking iterations and stochastic fluctuations in road network expansion, which necessitate additional effort for subsequent screening in practical applications.

To address scalability challenges, recent studies have incorporated optimization algorithms in PCG-based approaches (Xing et.al., 2021), and conducted comparative analyses of multiple optimization algorithms, evaluating their performance in generating intersection- and overlap-free networks with respect to optimal solution and convergence dynamics(Chen et al., 2018).

PCG-based approach prioritizes the creation of long-tail configurations that are statistically rare in real world (e.g., irregular intersections), but neglects compliance with real-world representativeness and road engineering specifications, thereby limiting the ecological validity of generated scenarios.

Notably, certain studies have assimilated road engineering specifications(Chen et al., 2008; Becker et al., 2020) and AV testing requirements (Chen et al., 2020) into road environment. Nevertheless, these latent constraints remain unincorporated into systematic road network generation approaches and are prone to neglect in practical implementations.

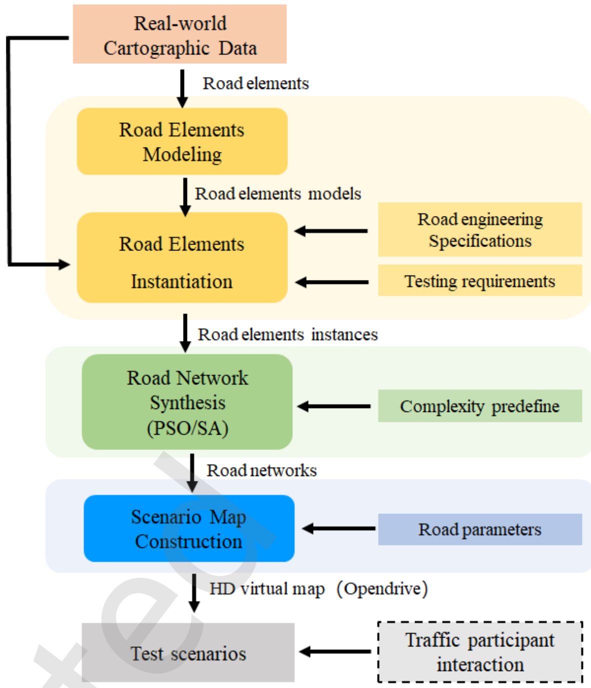


Fig.1. Flowchart of road network generation

To address the aforementioned issues, this study proposes a road network generation approach based on the optimized combination of realistic road elements, which integrates the current mainstream data-driven and PCG-based approaches. Fig.1 illustrates the proposed three-stage workflow. First, real-world urban road networks are decomposed into heterogeneous road elements. Elements with high potential risks (e.g., T-/Cross-junctions, merging/diverging zones) are abstracted into graph models characterized by tunable node and edge parameters.

Subsequently, road element models are instantiated with real-world cartographic data, adhering to road engineering specifications and testing requirements. A complexity quantification metric is established based on potential path conflict counts, enabling complexity labeling of instantiated elements. A network-level complexity evaluation function is then formulated to quantify global complexity.

Then, a road network optimization generation model with tunable complexity is developed. By selecting road element instances with predefined complexity labels, network complexity is systematically controlled. Optimization algorithms are introduced to guide instance interconnection, taking the realism and compactness of the road network as the optimization objectives, avoiding unintended intersections and overlaps between instances and minimizing the spacing among them. Finally, the synthesized networks are automatically compiled into OpenDRIVE-compliant high-definition (HD) maps for AV simulation testing. Specifically, the contributions of this work are summarized as follows:

1. A complexity control method for test road networks is proposed, generating test road networks with expected complexity through combinatorial assembly of road elements with quantitative complexity evaluation labels. It establishes a correlation between test road complexity and scenario challenge levels, addressing the limitations of existing road network generation methods that primarily focus on geometric morphology assembly without the capability to batch-generate multi-level, diversified test road networks according to specific scenario construction requirements, significantly enhancing

both the efficiency and coverage scope of road network generation

2. A road element model instantiation approach is developed based on real-word cartographic data, incorporating road engineering specification and practical testing requirements. This approach ensures that generated road networks maintain high fidelity to real-world conditions, comply with engineering standards, and align precisely with testing objectives. It effectively circumvents the limitations of existing PCG-based approaches that predominantly emphasize rare road structure generation, thereby expanding the applicability and practical utility of generated road networks across diverse testing scenarios.

3. An optimization-based road network synthesis mechanism is constructed, which realizes the non-intersecting and non-overlapping interconnection of road instances while compressing redundant mileage in the road networks. This mechanism eliminates excessive low-risk redundant testing mileage that compromises testing efficiency; and overcomes the inefficiencies inherent in existing road network generation methods that rely on iterative trial-and-error combinations to avoid unintended intersections and overlaps, which typically result in inconsistent network scales and suboptimal computational performance.

The rest of the paper is organized as follows: Section 2 presents the proposed road network generation approach. Section 3 establishes experimental case studies to validate the proposed approach. Section 4 analyses the results of the validation experiments. Finally, Section 5 concludes the study and provides recommendations for future research directions.

2 Methodology

This section presents the complete methodology for generating test road networks. As visualized in the Fig. 2, the proposed approach systematically integrates road elements modeling, quantitative complexity assessment, realistic and specification-compliant instantiation, optimization-driven network synthesis, and scenario map construction. Each step of this unified workflow is elaborated in the subsequent subsections.

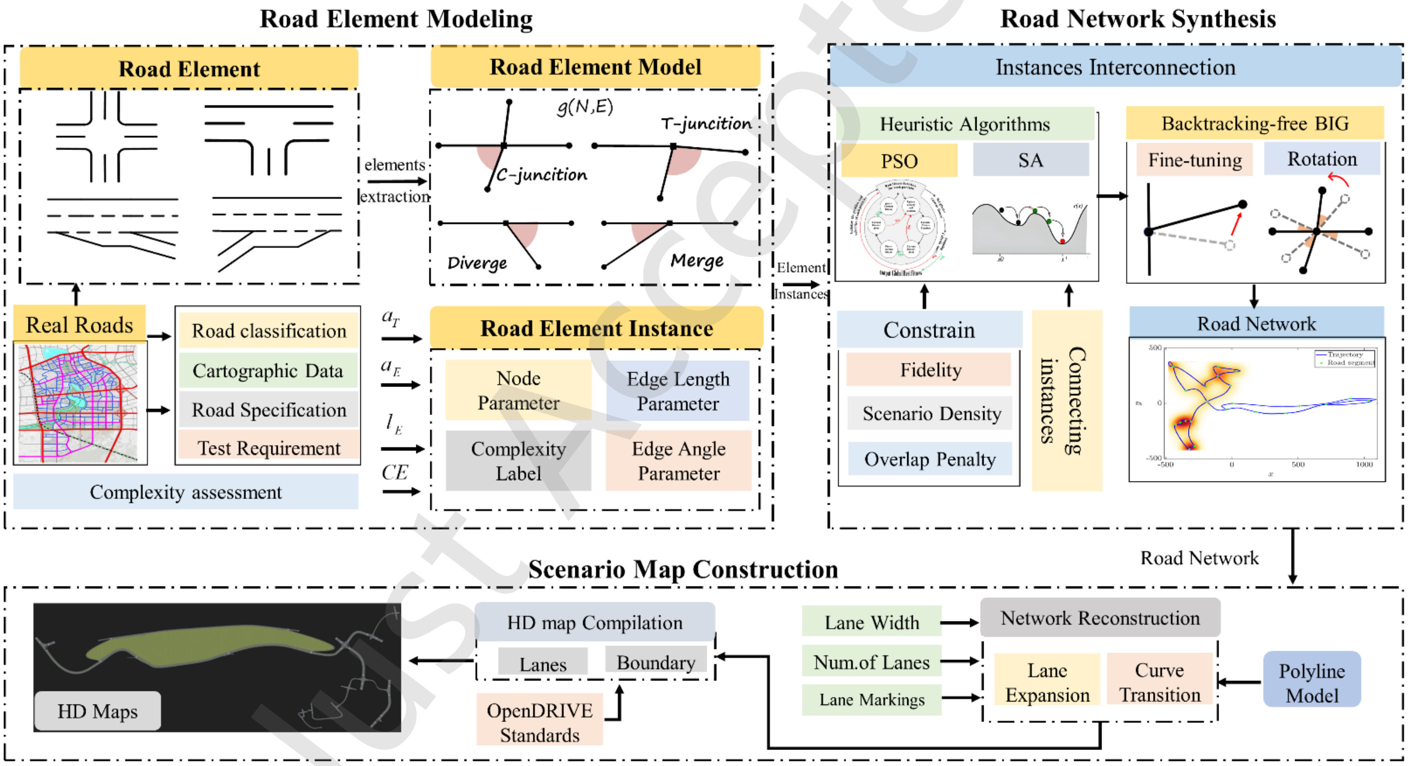
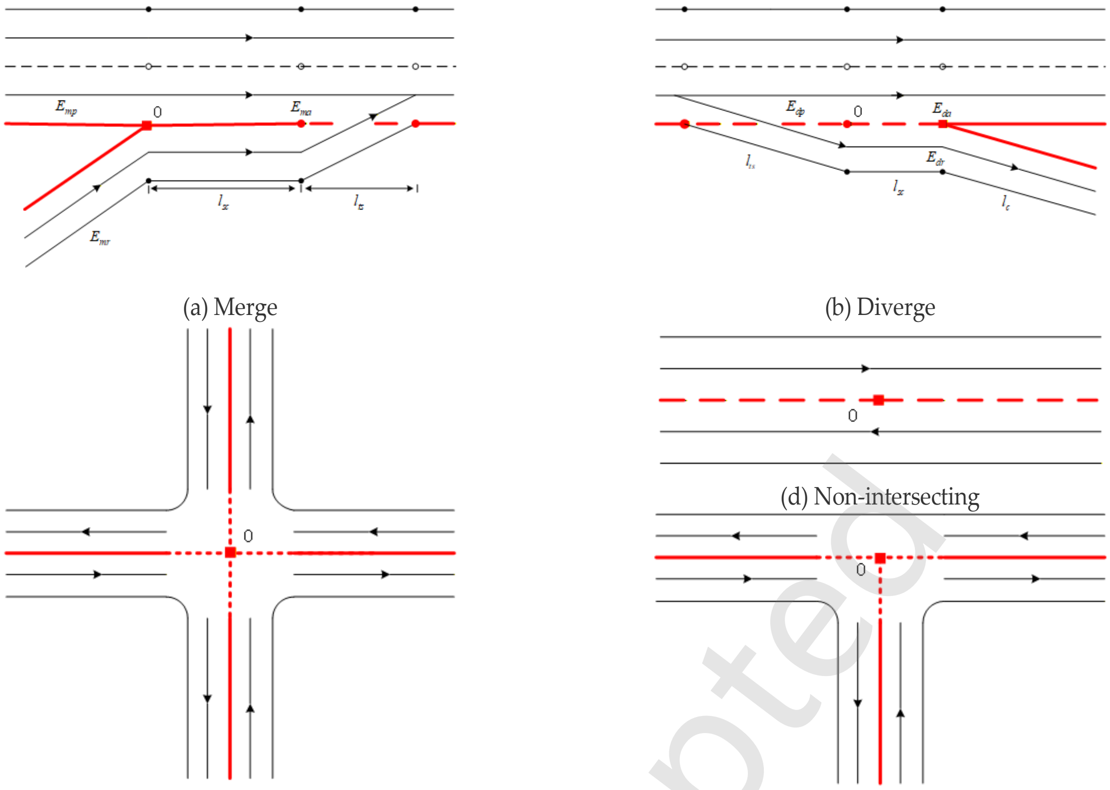


Fig.2. Overview of road network generation

2.1. Road element modeling

A road element constitutes the fundamental building block of test roads, with any complex road amenable to conceptualization as a sequence or network formed by the interconnection of multiple such elements.

The hierarchical paradigm (Parish & Müller, 2001; Li & Shi, 2010; Zhang et al., 2015; Zhao et al., 2023) is adopted for road element modeling. By leveraging the road classification provided in cartographic data, road elements are categorized into four hierarchical levels: expressways, trunk roads, feeder roads, and internal roads. In terms of structural attributes, road elements are further categorized into two categories: intersecting and non-intersecting elements. Intersecting elements constitute the core of road networks in test scenarios, as evidenced by statistical reports indicating that most accidents and high-risk scenarios occur at intersections (Moosavi et al., 2019). Building on prior researches (Kelly & McCabe, 2007; Gambi et al., 2019b), intersecting elements are modeled into four types: T-junctions, cross junctions, merges, and diverges.



(a) Merge

(b) Diverge

(d) Non-intersecting

(c) Cross-junction

(e) T-junction

Fig.3. Schematic diagram of road element models.

To emphasize geometric features, each element is abstracted as a polyline model, as illustrated by the red lines in Fig.3, and modeled as a graph $g(N, E)$. N denotes the centroid and endpoints. Both are defined by positional parameters p_N and attribute parameters a_N. The centroid establishes the spatial reference of the road element, serving as the origin of a local coordinate system. Endpoints are defined relative to this origin, while attribute parameters encode hierarchical levels and connectivity relationships of road elements. E represents the diverging edges from the centroid. Each edge is characterized by a length parameter l_E and an angular parameter $a_E \in [0, 2\pi]$, which is defined as the counterclockwise angle between adjacent edges, capturing the relative rotational positioning of edges.

2.1.1. Road models instantiation

The geometric configuration of road elements is governed by multiple factors. Significant regional heterogeneity arises from geographic environments (Badhrudeen et al., 2022). Road engineering specifications impose explicit regulatory constraints, while practical testing requirements may necessitate specific adaptations. Accordingly, these factors must be explicitly embedded into the instantiation of road element models to ensure environmental realism and testing efficacy.

In the road model instantiation, the edge angular parameters a_E are initialized based on their statistical distribution extracted from real-world cartographic data of predefined regions to ensure geometric fidelity. The edge length parameter l_E are calibrated through a dual-driven formulation, integrating regulatory road engineering specifications with testing tasks requirements.

(1) Merge

A merge element is structured into three components, pre-merging E_{mp}, post-merging E_{ma}, and ramp E_{mr}, as depicted in Fig.3(a). For potential test scenarios around the centroid, the length of E_{mp} should satisfy two critical requirements: (1) enabling velocity transition of the VUTs to the intended test speed, and (2) complying with the minimum sight distance l_1 ahead of the merging nose (China Road and Bridge Corporation, 2021). Similarly, for potential test scenarios within the speed-change lane, the ramp length must fulfill dual constraints: (1) ensuring sufficient distance for acceleration/deceleration, and (2) maintaining the prescribed sight distance l_2 while reserving adequate length l_c to the transition curve required for road element interconnection.

Within the critical zone from the merge nose to the transition section in E_{ma} where high-risk scenarios are spatially concentrated (Ye et al., 2023), VUTs should maintain the intended test speed. Accordingly, while complying with specifications for the lengths of speed-change lanes and transition sections, additional length allocation must be explicitly incorporated to enable speed transitions from the test speed to the predefined terminal velocity.

$$l_{E_{mp}} = max\left(\frac{v_1^2 - v_0^2}{2a} + v_1 t_h, l_1\right) \tag{1}$$

$$l_{E_{mr}} = max\left(\frac{{v'}_1^2 - {v'}_0^2}{2a} + v'_1 t_h, l_2 + l_c\right) \tag{2}$$

$$l_{E_{ma}} = l_{sc} + l_{ts} + \frac{v_2^2 - v_1^2}{2a} \tag{3}$$

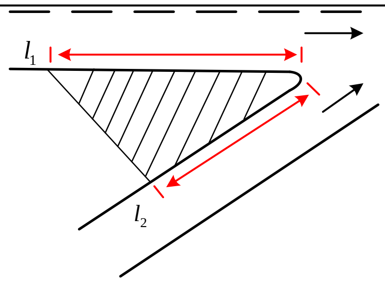


Fig.4. Schematic diagram of merging nose

The initial speed is denoted as v_0, the intended test speedas v_1, and the predefined terminal velocity as v_2. The initial speed in ramp is designated as v_0', with the intended exit speed specified as v_1'. The duration of potential test scenarios is defined as t_h, and the acceleration of VUTs is denoted as a. Specifically, l_{sc} corresponds to the length of the speed-change lane, l_{ts} to the length of the transition section, and l_c to the reserved length for the ramp transition curve.

(2) Diverge

The area preceding the diverging point forms a critical zone with high-risk scenarios (Ye et al., 2023). A diverge element is subdivided into three components: pre-diverging E_{dp}, diverging E_{dr}, and post-diverging component E_{da}. For potential test scenarios before the diverging point, E_{dp} should facilitate speed adjustment to the intended test speed v_1 before the transition section. For non-diverging vehicles, E_{da} should enable deceleration to the predefined terminal velocity v_2. Regarding the diverging component E_{dr}, in addition to the speed-change lane, adequate length l_c should be allocated for the ramp transition curve.

$$l_{E_{dp}} = \frac{v_1^2 - v_0^2}{2a} + v_1 t_h + l_{ts} \tag{4}$$

$$l_{E_{dr}} = l_{sc} + l_c \tag{5}$$

$$l_{E_{da}} = v_1 t_h + \frac{v_2^2 - v_1^2}{2a} \tag{6}$$

Furthermore, the diverging element should comply with the exit recognition sight distance l_{id} requirement stipulated in the specifications, which is mathematically expressed as follows.

$$l_{E_{dp}} + l_{sc} \geq l_{id} \tag{7}$$

(3) T/Cross -Junction

A T- and Cross-junction element is centered at the intersection point, with the component approaching this center denoted as E_{in} and departing from it denoted as E_{out}.

For E_{in}, testing requirements mandate speed adjustment from the initial speed v_0 to the intended test speed v_1, with maintenance of this velocity until entering the intersection. E_{out} should enable VUTs to adjust their speed to predefined terminal velocity v_2 upon completing the test. Additionally, E_{in} should comply with the stopping sight distance l_{st} requirement stipulated in road engineering specifications.

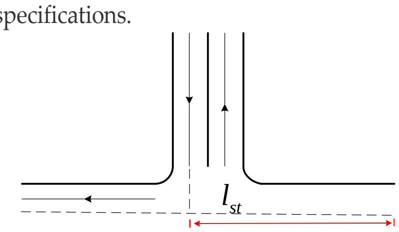


Fig.5. Schematic diagram of intersection sight distance

Furthermore, unlike merge and diverge, junction elements operate bidirectionally. To satisfy testing requirements, the lengths of approach and departure are standardized within the model.

$$l_{E_j} = max\left(\frac{v_1^2 - v_0^2}{2a_{in}} + v_1 t_h, l_{st}, \frac{v_2^2 - v_1^2}{2a_{out}} + v_1 t_h\right) \tag{8}$$

The acceleration of VUTs within E_{in} is defined as a_{in}, within E_{out} as a_{out}.

(4) Non-intersecting

For non-intersecting element, the primary constraint arises from testing requirements that dictate edge-length parameters. Analogous to the intersecting element, the testing workflow is systematically partitioned into three sequential phases, entry, interaction, and termination. During the entry phase, the VUTs undergo velocity regulation from the initial velocity v_0 to the intended test velocity v_1. Throughout the interaction phase, the VUTs sustain velocity stability over the duration of potential scenarios. In the termination phase, it executes velocity adjustment to the predefined terminal velocity v_2.

$$l_n = \frac{v_1^2 - v_0^2}{2a_{en}} + v_1 t_h + \frac{v_2^2 - v_1^2}{2a_{te}} \tag{9}$$

2.1.2. Complexity assessment

As a factor exerting a deterministic influence on test scenarios, the road environment is required to be controllably adjusted to align with the dynamic shifts in test scenario challenge levels. The core hypothesis is that road elements with higher complexity can accommodate test scenarios featuring higher collision risks and more challenging interactions. Specifically, in the absence of consideration for dynamic interactions, such road networks should inherently contain a higher number of potential path conflicts, which can be adopted as a quantitative complexity metric for road elements.

Potential path conflicts stem from merging, diverging, and path intersections. For intersecting road elements, building upon existing studies (Fischerson, 1984) the complexity index CE_J is defined as

$$CE_J = N_M + N_D + N_J \tag{10}$$

where N_M is the number of merging points, N_D is the number of diverging points, and N_J is the number of path crossing points. Obviously, for merge and diverge, potential path conflicts are concentrated at the merging/diverging points themselves., whereas T/Cross-Junction prioritize cross-path interactions among traffic participants.

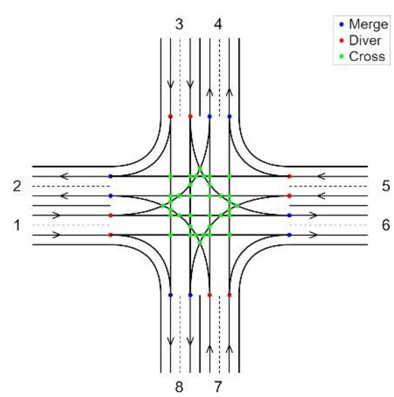


Fig.6. Path crossing points of cross -junction

As shown in Fig.6, in a standard four-way Cross-Junction, the number of same-direction input/output paths at each end is denoted as n_i. T-junction can be treated as a special case where $n_i = 0$ for any pair of input and output ends.

The input path A_i splits into a straight path P_t, a left-turn path P_l, and a right-turn path P_r. Within the cross-junction, the number of diverging points N_D is

$$N_D = \sum_i \left(card\left(P_t^{(i)} \cup P_l^{(i)} \cup P_r^{(i)}\right) - card(A_i) \right) \tag{11}$$

Similarly, paths with different travel directions merge into the output path E_i, the number of merging points N_M is

$$N_M = \sum_i \left(card\left(P_t^{(i)} \cup P_l^{(i)} \cup P_r^{(i)}\right) - card(E_i) \right) \tag{12}$$

Crossing points are classified into straight-path crossing points C_t and left-turn-induced path crossing points C_l, such that the total number of crossing points N_J is defined as follow.

$$N_J = card(C_t) + card(C_l) \tag{13}$$

$$C_t = \bigcup_{i,j \in A, i \neq j} C_{t,t}^{(i,j)} \tag{14}$$

$$C_l = \bigcup_{i,j \in A, i \neq j} C_{l,t}^{(i,j)} \bigcup_{i,j \in A, i \neq j} C_{l,l}^{(i,j)} \tag{15}$$

$C_{t,t}^{(i,j)}$、$C_{l,t}^{(i,j)}$、$C_{l,l}^{(i,j)}$ represent the counts of path crossing points. $C_{t,t}^{(i,j)} = P_t^{(i)} \cap P_t^{(j)}$, $C_{l,t}^{(i,j)} = P_l^{(i)} \cap P_t^{(j)}$, $C_{l,l}^{(i,j)} = P_l^{(i)} \cap P_l^{(j)}$. For the example in Fig. 6, $N_M = 8$, $N_D = 8$, $N_J = 36$.

For non-intersecting elements, in the absence of explicit path crossings, conflicts arise from the merging and diverging during lane-changing maneuvers. The complexity index is dependent with number of lanes. It is assumed that no lane changes occur in a single-lane road, whereas a multi-lane road may exhibit multiple conflicts induced by lane changes. The complexity index for non-intersecting road elements CE_C is defined as

$$CE_C = 2(n_l - 1) \tag{16}$$

n_l denotes the number of lanes available for unidirectional travel.

2.2. Road network synthesis

Continuous scenario testing provides a more accurate reflection of autonomous vehicle performance compared to isolated scenario testing (Zhu et al., 2025). To enable multi-scenario continuity, road elements must be interconnected into cohesive road networks that support sequential scenario testing.

The road network synthesis is realized by sequential connecting two road instances within the set of road instances $G_s = \left\{ g_s^{(1)}, \ldots, g_s^{(N_R)} \right\}$ via connecting instances n_c, gradually forming a road network.

The road instance interconnection employs an enhanced backtracking-free BIG paradigm. After sequentially integrating all road instances into the network, an optimization algorithm is introduced to guide the fine-tuning of their relative positions, involving slight adjustments to endpoint positions and rotations of the road instances themselves, thus avoiding unintended intersections and overlaps within the generated road network.

Connecting instances, akin to non-intersecting road instances, exhibit no inherent cross-path conflicts, with their complexity metric derived from lane-change-induced merging/diverging maneuvers. Consequently, such instances support fewer test scenarios. To minimize the redundant test mileage introduced by connecting instances, road instances should be arranged in a compact spatial configuration.

The optimization objective function f is composed of three distinct components. The first component f_1 is defined as the sum of errors between actual and target parameters l_E and a_E within each road instance, aiming to minimize deviations induced by random positional fine-tuning of endpoints. The second component f_2 is the total length of all connecting instances, which serves to minimize redundant mileage. The third component f_3 is a penalty term arising from the degradation of optimal solutions caused by unintended intersections or overlaps in the network.

To this end, the road instance interconnection can be formulated as an optimization.

$$\underset{P_N, N_c}{minimize} \, f \tag{17}$$

$$f = f_1 + f_2 + f_3 \tag{18}$$

$$f_1 = \sum_i \frac{\delta_l^{(i)}}{l_E^{(i)}} + \sum_{i,j} \frac{\delta_a^{(i,j)}}{a_E^{(i,j)}} \tag{19}$$

$$f_2 = \sum_{k=1}^{N_c} l_{n_c}^{(k)} \tag{20}$$

$$f_3 = \sum_{e,e'} \alpha_0 I_A\left(v(e,e')\right) + \sum_{e,n_c} \alpha_1 I_A\left(v(e,n_c)\right) + \sum_{n_c,n_c'} \alpha_2 I_A\left(v(n_c,n_c')\right) \tag{21}$$

δ_l and δ_a encapsulate errors of length and angular parameters introduced during fine-tuning of endpoints. $\alpha = \{\alpha_0, \alpha_1, \alpha_2\}$ is specified as a set of sufficiently large values, such that the objective function undergoes a significant increase when edges or connecting instances within the network intersect or overlap. I_A is an indicator function, and v is an intersection/overlap check. If edge e and e' are mutually non-intersecting, $v(e,e') > 0$.

For high-dimensional, nonlinear optimization problems such as the interconnection of road instances, heuristic algorithms exhibit distinct advantages in terms of feasibility and computational efficiency. This work explores multiple heuristic optimization algorithms to synthesize road networks, including Particle Swarm Optimization (PSO), Simulated Annealing (SA), and Monte Carlo (MC) stochastic algorithm, to investigate the trade-off between solution quality and computational runtime.

In the Particle Swarm Optimization (PSO) algorithm, each road network composed of road instances is represented as a particle within the search space. Through information exchange among particles, both personal bests *pbest* and global best *gbest* are iteratively updated. A comprehensive description of the algorithm is available in reference (Eberhart & Kennedy, 1995).

Notably, transmission of *gbest* across particles with distinct connection configurations may trigger erratic oscillations in the optimal solution. To address this issue, N_w mutually independent particle swarms are initialized in parallel. Within each swarm, a uniform connection configuration is deployed, and optimization operates independently with a unique *gbest*.

Each swarm has N_p particles, wherein each particle contains instances capable of forming an integrated road network. Particles are parametrically characterized by $\boldsymbol{R}_\tau = \{\boldsymbol{x}_\tau, \boldsymbol{a}_\tau\}$, where $\boldsymbol{x}_\tau$ denotes the position of each instance, and $\boldsymbol{a}_\tau$ represents the cumulative counterclockwise rotation throughout optimization. The subscript τ signifies the uniform connection configuration.

During the optimization, the objective function is designed to decrease iteratively to identify an optimal road network, in which road instances are tightly interconnected, while remaining mutually free of intersections and overlaps. The optimization terminates upon reaching a predefined maximum iteration count.

Given the intrinsic propensity of the PSO to converge to local optima, the Simulated Annealing (SA) algorithm is deployed for road network construction. A key advantage of SA lies in its capacity to probabilistically accept suboptimal solutions, precluding entrapment in local optima (Kirkpatrick et al., 1983).

For SA algorithm, a critical objective in road network generation is to preserve the intrinsic capacity to escape local optima while mitigating convergence impairment caused by accepting suboptimal solutions. To this end, a global optimal solution *nbest* is introduced. Upon stochastic acceptance of suboptimal solutions, an auxiliary decision-making mechanism is integrated to preclude suboptimal solutions from supplanting *nbest*. The retention of *nbest* guarantees a monotonically decreasing objective function throughout the optimization, regardless of whether suboptimal solutions are accepted.

After N_w independent optimizations, the set of optimal road networks encompasses $N'_w (N'_w \leq N_w)$ optimal road networks with distinct connection configurations, from which the global optimal solution r_s is selected. In practice, a larger swarm size and multiple search iterations are employed to enhance the coverage of feasible connectivity configurations and broaden the exploration scope of the optimization search breadth.

Moreover, after interconnection, road instances form a tree topological structure, in which a set of endpoints remains unconnected. Additional direct connections may be instituted between these dead-ends, expanding the tree-like topology into a full-connected road network. The establishment of such connections is governed by the following criteria:

(1) Additional connections are established in accordance with the hierarchy of trunk roads, feeder roads, and internal roads.

(2) Preference is given to establishing connections with other unconnected endpoints in proximate distance.

(3) Additional connections shall neither intersect nor overlap with edges of the extant road network.

The Monte Carlo (MC) stochastic algorithm is adopted as a comparative benchmark against the two algorithms. During optimization, MC exclusively accepts solutions with lower objective function values, without inter-solution information exchange. In the absence of a local optima evasion mechanism, MC is constrained to relying on the premise that when one solution is trapped in a local optimum, multiple concurrent road networks within the same group will endeavor to identify a superior solution.

Given that continuous scenario testing better reflects the real-world performance of VUTs compared to isolated scenarios, it is also imperative to conduct quantitative holistic complexity assessments of synthesized road networks formed through road element selection. This ensures that the synthesized networks enable quantitative adjustment of structural diversity and challenge levels to validate autonomous vehicle performance in multi-scenario continuous testing. The complexity index CE of a generated road network consisting of intersecting instances g_J , non- intersecting instances g_{nJ} and connecting instances g_{nc} is

$$CE = \sum_{g_J} CE_J + \sum_{g_{nJ}} CE_c + \sum_{g_{nc}} CE_n \tag{22}$$

Due to structural similarity, the complexity labels of connecting instances CE_n are computed using the same formulation as defined in Equation (16).

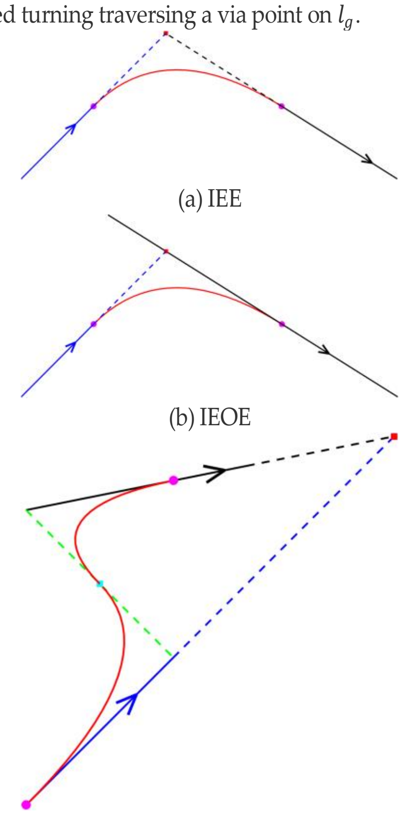
2.3. Scenario map construction

To leverage generated road networks for constructing virtual simulation test scenarios, road networks represented in a topological graph should be converted into high-definition maps compatible with virtual simulation platforms.

Given that simplified polyline-based models of road elements were employed in road network generation, reconstruction into complete models bounded by multi-set boundary and lane lines is prerequisite to high-definition map generation. While preserving the dimensional lengths of all constituent components, supplementary road geometric parameters (e.g., number of lanes, road width) and auxiliary attributes (e.g., directionality, lane marking) must be incorporated.

Connecting instances are reconstructed via the curve model to ensure geometrically smooth transitions between adjacent road elements. As visualized in Fig.7, for two directly interconnected road instances $g_s^{(1)}$ and $g_s^{(2)}$, blue lines represent the edge within $g_s^{(1)}$, black lines represent the edge within $g_s^{(2)}$, and green line denotes the straight-line connection l_g between the edge endpoints of the two road instances. Arrows indicate the driving direction of VUTs. Three distinct planar positional configurations between $g_s^{(1)}$ and $g_s^{(2)}$ are considered: (a)intersection of extended edges (IEE); (b) intersection of single extended edge and opposing edge (IEOE); (c) intersection of extended edge and the opposing reverse extended edge (IEORE).

For configuration (a), the connecting instance is reconstructed as a curved turning linking the endpoints of the two edges. For (b), it is reconstructed as a curved turning between the endpoint of $g_s^{(1)}$'s edge and an internal point on $g_s^{(2)}$'s edge. For (c), it is reconstructed as an S-shaped curved turning traversing a via point on l_g.



(a) IEE

(b) IEOE

(c) IEORE

Fig.7 Positional configurations between adjacent elements.

Through Matlab scripts, the reconstructed road network is able to be translated into OpenDRIVE formats required by autonomous driving systems. This scripting is executed automatically, requiring no additional manual intervention for varying road network inputs.

3 Experiment

We presented a case study-based analysis to demonstrate the feasibility of the proposed road network generation approach and evaluate its performance through virtual simulation testing.

3.1. Road network generation

For road element instantiation, the OpenStreetMap (OSM), an openly accessible global cartographic database, was utilized as the foundation for geometric modeling. Specifically, cartographic data from Xi'an, China, were selected to enable data-driven instantiation. The selected area encompasses the region within the Xi'an Ring Expressway, including 21,078 merges/diverges and T-junctions, and 11,665 Cross-junctions. The angle parameter a_E was designated as the value corresponding to the center of clustering within the value domain (see Table1) to preserve spatial representativeness across the target region.

Table 1. Instantiation of road element model

Element	$a_E(rad)$	$l_E(m)$	CE
Merge			
Exp.	(0.11, 3.03, 3.14)	(143.2, 90, 327.6)	1
Trk.(Exp.r)	(0.21, 2.93, 3.14)	(102.5, 90, 134.4)	3
Trk.	(0.25, 2.90, 3.13)	(102.5, 90, 134.4)	3
Diverge			
Exp.	(0.28, 2.85, 3.15)	(223.2, 140, 122.0)	1
Exp.(Trk.r)	(0.94, 2.25, 3.09)	(142.5, 50, 87.6)	3
Trk.	(0.64, 2.52, 3.12)	(142.5, 50, 87.6)	3
T-Junction			
Trk.	(1.42, 1.76, 3.10)	(102.5, 102.5, 102.5)	9
Trk.(Fed.)	(1.50, 1.65, 3.13)	(102.5, 64.5, 102.5)	11
Trk.(Int.)	(1.51, 1.64, 3.13)	(102.5, 24.8, 102.5)	11
Fed.	(1.52, 1.64, 3.12)	(64.5, 64.5, 64.5)	9
Fed.(Int.)	(1.53, 1.62, 3.13)	(64.5, 24.8, 64.5)	9
Int.	(1.54, 1.62, 3.12)	(24.8, 24.8, 24.8)	9
Cross-Junction			
Trk.	(1.42, 1.52, 1.62, 1.72)	(102.5, 102.5, 102.5, 102.5)	52
Trk.(Fed.)	(1.48, 1.54, 1.60, 1.66)	(102.5, 64.5, 102.5, 64.5)	40
Trk.(Int.)	(1.46, 1.53, 1.62, 1.67)	(102.5, 24.8, 102.5, 24.8)	40
Fed.	(1.50, 1.54, 1.60, 1.64)	(64.5, 64.5. 64.5, 64.5)	32
Fed.(Int.)	(1.52, 1.55, 1.59, 1.62)	(64.5, 24.8, 64.5, 24.8)	32
Int.	(1.52, 1.55, 1.59, 1.62)	(24.8, 24.8, 24.8, 24.8)	32
Non-intersecting			
Exp.	-	354.0	2
Trk.	-	256.5	6
Fed.	-	164.3	2
Int.	-	68.6	2

Exp., Trk., Fed., and Int. correspond respectively to expressways, trunk roads, feeder roads, and internal roads. "Trk.(Exp.r)" of Merge designates the merging of expressway ramps into trunk roads. Length-related parameters were determined in adherence to the road engineering specifications (China Road and Bridge Corporation, 2021)(see Table2).

With respect to testing requirements, a streamlined testing protocol was considered. Within a single road element, the VUT starts the test at a low initial velocity v_0, accelerates to the intended test speed v_1, maintains it for a defined duration, and thereafter decelerates to the terminal velocity v_2. Divergent velocity profiles across road categories were delineated in Table3.

Table 2. Road engineering specification-driven instantiation

Parameter		l_{sc}	l_{ts}	l_1	l_2	l_c	l_{st}
Exp.(m)	Merge	180	70	100	60	30	-
	Diverge	110	80	-			
Trk.(m)	Merge	40	40	100	60	30	75
	Diverge	20	40	-			
Fed.(m)	-	-	-	-		-	40
Int.(m)							20

Table 3. Test requirement-driven instantiation

Parameter(km/h)	Exp.	Trk.	Fed.	Int.
v_0	60	40	20	0
v_1	80	60	40	20
v_2	60	40	20	0
v_0'	20		-	-
v_1'	40		-	-

It was postulated that the test speed persists for $t_h = 2s$ in intersecting elements, whereas non-intersecting elements require a $t_h = 8s$ sustained test speed. Acceleration parameters were delineated based on the 85th percentile of drivers' acceleration and deceleration under conditions that minimize external confounding effects (Liu & Zhang, 2022), with $a_{acc} = 1.1m/s^2$ and $a_{dec} = -1.4m/s^2$.

The lane count used in road element complexity assessment, along with the lane width and inner radius parameters in T-/Cross junctions applied in scenario map construction, were summarized in Table4.

Table 4. Parameters for complexity assessment and scenario map construction

Road element	Number of lanes	Lane width(m)	Radius of inner edge
Exp.	2	3.75	\
Exp.(r)	1	3.75	\
Trk.	4(bi.)	3,75	25
Trk.(r)	1	3.75	\
Ter.	2(bi.)	3.5	20
Res.	2(bi.)	3	15

In Table4, 'r' denotes a ramp, and 'bi.' indicates bidirectional travel.

Table 5. Parameters used for road network synthesis

PSO		SA	
Parameter	Value	Parameter	Value
Inertia Weight(ω)	0.4	Initial temperature (t_s)	1000
Self-learning factor(c_1)	0.3	Final temperature (t_f)	899.86
Global-learning factor(c_2)	0.2	Cooling rate (c_r)	0.99
Maximum iterations(T_{max})	200	Pre-set iterations (N_t)	100
Number of particles in each swarm(N_p)	200	Number of road networks in each group(N_p)	200
Number of particle swarms(N_w)	500	Number of road network groups(N_w)	500

As an example, 15 road networks with varying complexity levels are synthesized by incrementally selecting 8 to 22 distinct road elements. For each complexity levels, Particle Swarm Optimization (PSO) and Simulated Annealing (SA) algorithms were independently executed to generate 100 networks with distinct connection configuration. The Monte Carlo (MC) algorithm was adopted as the baseline for comparison.

The parameters used for road network synthesis are summarized in the Table5.

3.2. Virtual testing in lane-change scenario

The generated networks are then utilized to construct testing scenarios for virtual simulations of autonomous vehicle. Regarding single-scenario testing, we constructed a typical high-risk lane-change testing scenario in the two-lane expressway non-intersecting element.

Both the lane-changing vehicle and the VUT utilized the built-in Tesla Model 3 model from the CARLA virtual simulation software, which incorporates a complete vehicle dynamics model and a high-fidelity appearance model, with dimensions of 4.720 m in length and 1.848 m in width.

The lane-change testing scenario is illustrated in Fig.8, with the lane-changing vehicle serving as the reference vehicle.

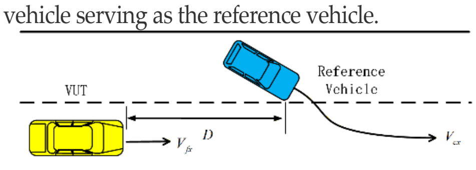


Fig. 8. Schematic diagram of lane-change testing scenarios

Prior to the testing, the reference vehicle travels ahead of the VUT on the left side. During the testing, the reference vehicle gradually approaches the target lane along a predetermined trajectory. At the moment of cut-in, it enters the target lane, creating a path conflict with the VUT following behind. A collision risk arises when the velocity of the reference vehicle is lower than that of the VUT following behind. It is expected that the rear VUT will detect the collision risk and initiate braking response to avoid collision.

Upon entering the target lane, the reference vehicle continues to travel along the predetermined trajectory while maintaining its position within the target lane.

The trajectories of both vehicles were generated using the algorithm proposed in existing research (Zhu et al., 2022), with a sampling frequency of 25 Hz and a trajectory duration of 6 seconds.

The collision risk between two vehicles was characterized by the Time-to-Collision (TTC) at the moment of cut-in.

$$TTC = -\frac{D}{V_{ref}} \tag{23}$$

The distance between two vehicles is denoted as D, while their relative velocity is V_{ref}. Nine categories of test scenarios were selected with TTC within the interval (0s, 30s] and summarized in Table 6.

Table 6. Categories of test scenarios

Test scenario	1	2	3	4	5	6	7	8	9
$TTC(s)$	0.5	1	1.5	2	3	5	10	15	30

As shown in Table7, in each category of test scenarios, the velocity of the reference vehicle was set from 6 m/s to 50 m/s, with a step size of 4 m/s.

Table 7. Velocity of the reference vehicle in test scenarios

Para.	1	2	3	4	5	6	7	8	9	10	11	12
$v_c(m/s)$	6	10	14	18	22	26	30	34	38	42	46	50

A total of 108 test scenarios were selected under the condition of $D \leq 50m$, covering all combinations of TTC and v_c. For each parameter configuration, 100 sets of vehicle trajectories were generated to establish 100 test cases.

Three autonomous driving intervention strategies were implemented during testing: the aggressive strategy, the moderate strategy, and the proactive strategy. Under the aggressive strategy, both vehicles followed their predefined trajectories until the TTC reached the preset threshold specified in the test scenario, at which point the autonomous driving algorithm assumed control of the VUT and responded based on the instantaneous positional relationship between the two vehicles. For the moderate strategy, in the test scenarios with $TTC < 3s$, the autonomous driving algorithm immediately took control when the TTC reached 3 s and executed the appropriate response. The proactive strategy further employed the intervention threshold of $TTC = 5s$.

3.3. Virtual testing in free-driving scenario

For continuous scenario testing, we employed generated road networks to conduct continuous free-driving tests, with the following test protocol:

In section 3.1, 15 road networks with varying complexity levels were synthesized by incrementally selecting 8 to 22 distinct road elements. For each complexity level, 100 road networks with different connection configurations were generated.

During testing, one road network was randomly selected from the 100 generated networks for each complexity level to serve as the test road. The virtual test vehicle randomly selected one endpoint of the test road as the starting point, and then executed continuous free-driving testing within the selected road network, randomly performing maneuvers including merging, diverging, straight driving, turning, and U-turns along the road network until traversing every endpoint of every road element. The test vehicle adopted the following speed configurations for different types of road segments.

(1) Merge: For main line driving, the test vehicle accelerates from a low initial velocity v_0 to the test speed v_1, maintains it until passing the centroid, then decelerates to the to the terminal velocity v_2; for merging maneuvers, the vehicle accelerates from initial speed v_0' on the ramp to the test speed v_1 via the transition lane, and subsequently decelerates to the terminal velocity v_2.

(2) Diverge: For main line driving, the test vehicle adopted the identical velocity profile to mainline driving in merging elements; for diverging maneuvers, the vehicle decelerates to the terminal velocity along the transition lane and then enters the ramp.

(3) T-/Cross junction: According to the branch type, the test vehicle first accelerated from the corresponding initial velocity v_0 to the test speed v_1. Regardless of straight driving or steering, it then adjusts to the initial velocity required for the subsequent branch. After entering the subsequent branch, the vehicle re-accelerates to the test speed and then decelerates to the terminal velocity. For U-turn, both the terminal velocity of current branch and the initial speed of the subsequent branch are set to 0.

(4) Non-intersecting: A acceleration-deceleration profile is applied, mirroring the main line driving in merge/diverge.

(5) Connecting instances: The test vehicle transitions from the terminal velocity of the preceding road instance to the initial velocity of the succeeding one.

An illustrative example of the speed configuration is presented in Table 3. Following the above test procedure and employing the Monte Carlo testing methodology, 100,000 free-driving tests were conducted for each of the 15 complexity levels of generated road networks.

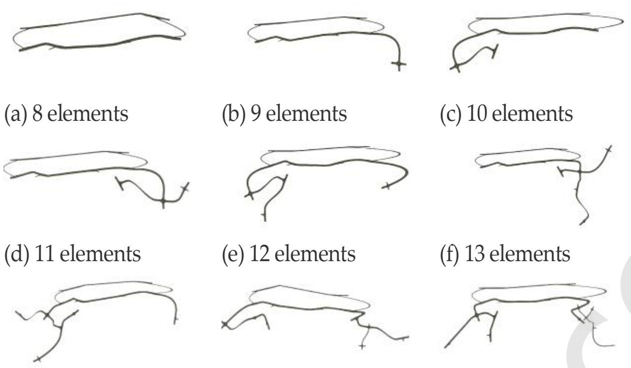
4 Results and analysis

4.1. Generated road networks

The generated road networks are compiled into high-definition (HD) map files for simulation platforms through the scenario map construction. As illustrated in Fig. 9, representative example of road networks synthesized with 8–16 distinct elements demonstrate the progressive increase in complexity. To explicitly visualize complexity evolution, all networks are depicted in tree-like topologies directly derived from optimized interconnection of road elements.

In the experiment, low-complexity elements, such as merge and diverge of expressways and highway and trunk road were prioritized to construct ring-like structures. These configurations minimize network complexity while supporting multi-scenario continuous testing. Subsequent incremental connections were implemented by taking trunk road as road skeleton and integrating T-/Cross junction elements into the road network.

This incremental connection mechanism ensures that road networks at identical complexity levels share similar topological features while exhibiting significant diversity in placement of elements and connectivity patterns, providing flexible options for simulation testing.



(a) 8 elements (b) 9 elements (c) 10 elements

(d) 11 elements (e) 12 elements (f) 13 elements

(g) 14 elements (h) 15 elements (i) 16 elements

Fig.9. Examples of HD maps with distinct road elements

Fig. 10 presents an illustrative example of a full-scale generic road network with 22 distinct road elements, with magnified insets highlighting detailed configurations of selected intersecting elements. Red bounding boxes delineate expressway road elements, blue bounding boxes identify trunk road elements, and green bounding boxes represent feeder road elements.

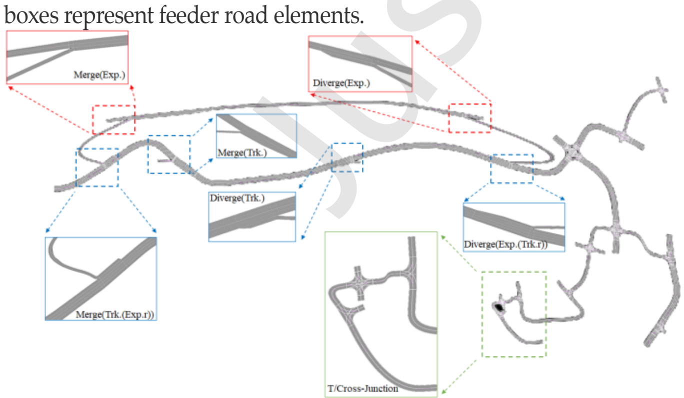
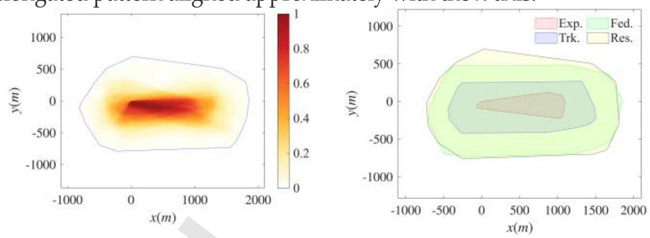


Fig.10. Example of a generated road network with local magnification insets

Furthermore, we analyze the topological similarity and structural diversity across 300 synthesized instances of this full-scale generic road network.

A full-scale generic road network with 22 distinct road elements is selected as a representative case to demonstrate the topological similarity and structural diversity among 300 synthesized test networks.

Convex hulls are employed to characterize the spatial extent of road instances across all three hundred generated networks. As illustrated in Fig.11(a), the blue convex hull delineates the extremal spatial envelope of synthesized networks with distinct topological configurations. The overall distribution of road instances exhibits an elongated pattern aligned approximately with the x-axis.



(a) Convex Hulls of Generated Road Networks (b) Convex Hulls of Heterogeneous Road Elements

Fig.11. Spatial Distribution of Road Instances.

The heatmap within the convex hull highlights the frequency patterns of road element spatial distribution. Regions rendered in deeper red indicate locations consistently covered across a large majority of the generated networks, whereas near-white areas correspond to peripheral zones that appear only under specific topological configurations. The majority of road instances are concentrated within high-frequency zones. In Fig.11(b), nested convex hulls of varying colors represent the extremal spatial extents of different road categories. Expressway and trunk road instances exhibit tightly clustered distributions, forming the structural backbone of the road network. Feeder and internal instances access roads extend outward, expanding the overall spatial coverage of the network.

To quantitatively assess the spatial discrepancy between any two generated road networks, we analyze the distribution of road instances within convex hulls. Specifically, the positional Chamfer Distance between the centroid of road instances is adopted as a metric.

$$d_c(R_A, R_B) = \frac{1}{N_R}\sum_{g_A \in R_A} \min_{g_B \in R_B} \|g_A - g_B\| + \frac{1}{N_R}\sum_{g_B \in R_B} \min_{g_A \in R_A} \|g_A - g_B\| \tag{24}$$

g_A and g_B denote the normalized coordinates of road elements in networks R_A and R_B. N_R represents the number of road instances within a given network.

d_c close to zero indicates high similarity in the spatial distributions of the two point sets, whereas larger values signify greater dissimilarity. Given that all coordinates are normalized, the upper bound approaches $\sqrt{2}$.

Fig.12 presents the distribution of d_c across all generated networks. The d_c among uniformly distributed point sets with equivalent cardinality is illustrated for comparative analysis. The average Chamfer Distance among generated networks was 0.33, comparable to that of the uniformly distributed point sets (0.30). However, the standard deviation reached 0.19, which is four times greater than that of uniform distributions. The d_c distribution exhibited pronounced positive skewness, with an expanded range (0.067–1.38) exceeding that of uniform distributions (0.16–0.51). The coefficient of variation ratio between two distributions was approximately 3.6.

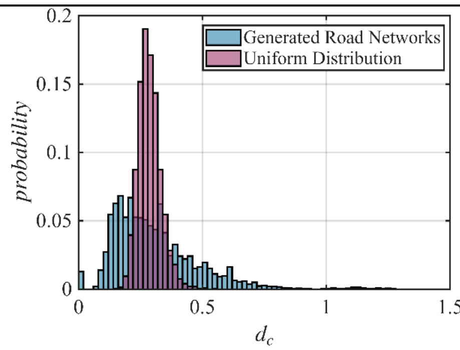


Fig.12. Distribution of chamfer distance

The relatively low average Chamfer Distance indicates structural similarity among networks. However, the presence of outliers with markedly elevated values reveals substantial positional discrepancies of road instances within certain networks.

The Gamma index $\gamma \in [0,1]$ quantifies the density of a road network's topology, defined as the ratio of existing edges to the maximum possible edges.

$$\gamma = \frac{N_c}{N_{cmax}} \tag{25}$$

Where N_c represents the number of connecting elements, and N_{cmax} denotes the maximum number of connecting elements theoretically possible for the same network.

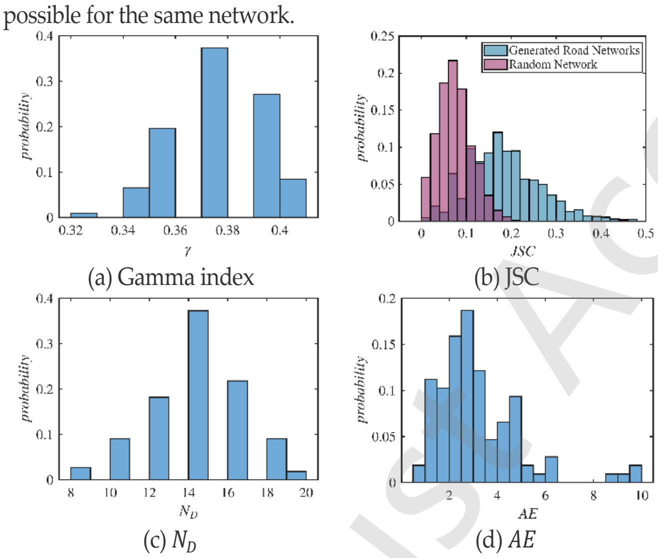


Fig.13. Connectivity metrics of road network.

Fig.13(a) illustrates the Gamma index distribution across all generated networks, showing a narrow normal distribution centered at $\mu = 0.38$, with $\sigma = 0.020$ and a coefficient of variation of 5.3%. This tight clustering suggests minimal topological density variation, with most networks approaching the theoretical upper bound under edge-crossing/overlap constraints.

Despite similar Gamma indices, adjacency matrices reveal subtle topological differences. Let J denote the adjacency matrix where $j_{ik} = 1$ if road instance $g_s^{(i)}$ and $g_s^{(k)}$ are directly connected, and 0 otherwise. The Jaccard Similarity Coefficient $JSC \in [0,1]$ quantifies pairwise adjacency matrix dissimilarity, which is defined as the ratio of the number of identical elements at all corresponding positions in two matrices to the total number of elements in their union. A value closer to 1 signifies higher similarity between the matrices.

$$JSC(A,B) = \frac{|J_A \cap J_B|}{|J_A \cup J_B|} \tag{26}$$

Fig.13(b) shows the JSC distribution of generated networks, with $\mu = 0.18$, $\sigma = 0.082$, significantly higher than random networks with identical Gamma indices ($\mu = 0.051, \sigma = 0.023$). This confirms non-random topological structural consistency, while the large standard deviation indicates inter-individual topological heterogeneity.

Number of dead-end N_D further characterizes connectivity. In Fig.13(c), the approximately normal distribution of N_D shows $\mu = 14.02$ dead-ends per network, 20.6% of total endpoints, with a range of 14. This variability reinforces the presence of divergent topological structure across networks.

An improved reachability matrix $R = [r_{ij}]$ quantifies connectivity between any two road instances, where r_{ij} represents the number of distinct paths between road instances $g_s^{(i)}$ and $g_s^{(j)}$. Undoubtedly, any two instances within a road network are mutually connected. The mean value of elements in R serves as a connectivity metric AE. A higher AE denotes a larger number of inter-element paths, reflecting enhanced connectivity.

$$AE = \frac{\sum_{i \neq j} r_{ij}}{N_R(N_R - 1)} \tag{27}$$

Fig.13(d) reveals bimodal distributions, which attests to strong structural heterogeneity in the topology. Statistical analyses demonstrate that the maximum AE reaches 9.66 that is significantly higher than the mean value of 3.15. This finding indicates that certain road networks feature meshed grids with abundant path redundancy. In contrast, the network associated with the minimum AE exhibits a tree-like topology with limited path options.

The aforementioned analysis demonstrates that the proposed road network generation approach balances structural stability and parametric diversity. While enforcing structural coherence, it permits systematic variations across extended parameter spaces, enabling the generation of diverse yet valid road network configurations.

4.2. Optimization performance

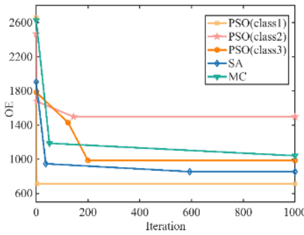
To prevent unintended intersection or overlap, minor stochastic adjustments to endpoints were permitted in optimization. Although the objective function incorporates error constraints, such perturbations risk compromising similarity between road instances and their expected parameter values. A similarity error metric FE was defined as Eq. 28 to quantitatively assess this deviation.

$$FE = \sum_{k \in G_g} \frac{\sum_i \left(1 - \delta_i^{(l)}/l_i\right) + \sum_{i,j}\left(1 - \frac{\delta_{ij}^{(k)}}{a_{ij}^{(k)}}\right)}{N_g^{(k)} N_R} \tag{28}$$

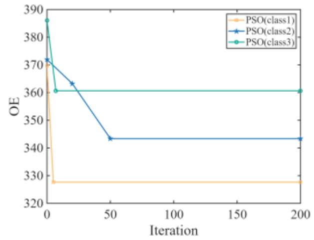
where G_g represents the road instances in network, $N_g^{(k)}$ is the number of endpoints for the k-th road instances, N_R denotes the total number of road instances in G_g.

Validation across PSO, SA, and MC algorithms revealed that the mean similarity error FE across all generated networks remained approximately 1.5%, with maximum deviations bound below 2.4%. This demonstrates that, irrespective of optimization methodology, the generated network preserves high fidelity to expected configurations.

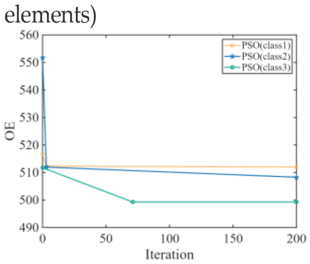
The Latent Class Trajectory Model (LCTM) (Li et al., 2020)was employed to compare the convergence across different algorithms. Figure 14(a) illustrates the evolution of the objective function throughout 1000 iterations during the generation of road networks comprising 22 road elements.



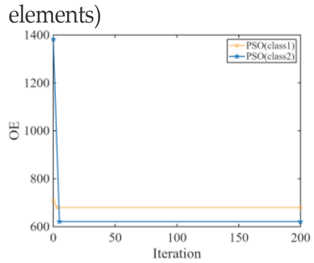
(a)Evolution of OE in optimization convergence (22 elements)



(b)Evolution of OE in optimization convergence (8 elements)



(c)Evolution of OE in optimization convergence (14 elements)



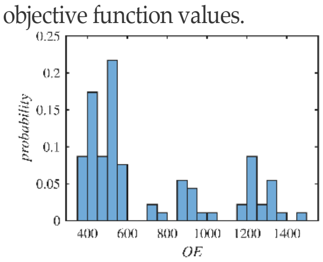
(d)Evolution of OE in optimization convergence (19 elements)

Fig.14. Evolution of OE in optimization convergence

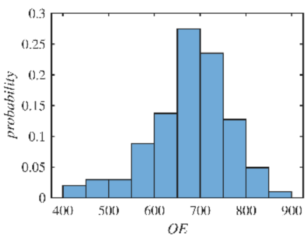
The convergence trajectories of PSO were categorized into 3 latent classes. Class 1 rapidly converges to the optimal solution within minimal iterations and remains stable thereafter. Class 2 and Class 3 exhibit a slower convergence and reach a significantly higher OE than that of Class 1, indicating entrapment in suboptimal local optima. The OE decay of SA exhibited three distinct phases: sharp reduction within the initial phase, decelerated decline in the middle phase, and eventual stabilization at the minimum value after about 600 iterations. In comparison, the OE variation profile of MC resembled that of the SA algorithm with the slowest decay rate. After 1000 iterations, the OE of MC remained significantly higher than SA's final value.

Fig.14(b)-(d) illustrate the evolutionary trajectories during the generation of road networks comprising 8, 14, and 19 road elements, respectively, using the Particle Swarm Optimization (PSO) algorithm. To emphasize convergence patterns, only the results from the initial 200 iterations are displayed.

Comparative analysis reveals that the road element library fundamentally determines the characteristics of the optimized road network. When employing different road element libraries for PSO-based optimized combination, the optimal solution's objective function (OE) values vary due to differences in the number of constituent road elements; however, the convergence remain largely consistent across libraries. Certain latent categories achieve rapid convergence to optimal solutions within a few iterations, while others exhibit slower convergence rates or yield suboptimal solutions with relatively higher objective function values.



(a) PSO



(b) SA

Fig.15. Distribution of OE

Fig.15 illustrates the distribution of optimal objective function values OE with 22 road elements. Corresponding to the latent classes, the OE of PSO formed three distinct clusters with proportions 0.64:0.21:0.15. The solution distribution of SA confirms robustness across diverse network configurations.

Obviously, PSO enables the fastest establishment of an optimal road network. However, it carries a high risk of getting trapped in local optima. In contrast, SA requires more iterations but exhibits robust escape mechanisms from suboptimal basins, ensuring high result consistency. The MC algorithm lacks an optimization-oriented mechanism, resulting in significantly slower convergence rates.

All experiments were conducted on a computing platform equipped with an Intel Core i7-12700 processor and 64 GB of RAM, utilizing the MATLAB environment for road network synthesis. Under these hardware and software configurations, the PSO algorithm required approximately 30 seconds per iteration in the combination of 22 road elements. The PSO algorithm demonstrated rapid convergence, typically achieving the optimal solution within several iterations. Consequently, the optimal road network could be obtained within a few minutes. This computational efficiency is suitable for non-real-time scenario construction applications, such as establishing candidate road network sets for testing scenario development.

In contrast, the SA algorithm consumed less than one minute per iteration, with the complete optimization process requiring approximately 8 hours to execute 600 iterations. The MC algorithm exhibited substantially inferior computational efficiency, requiring more than 24 hours to complete the optimization task.

Despite its relatively longer computational duration, the SA algorithm can be utilized to explore the objective function landscape for optimal solutions and further screen out local optimal solutions yielded by the PSO algorithm.

At the current road network scale, the PSO algorithm demonstrates significantly faster convergence to optimal solutions compared to the SA. This performance disparity arises because SA relies on small translations and rotations of individual road elements within a single candidate network to explore the solution space. In contrast, PSO leverages population-level collaboration among particles, enabling rapid convergence toward the optimal solution within the search space. As network scale expands, the computational complexity of PSO grows approximately linearly with the number of road elements $(O(n))$(Iqbal et al., 2015), while SA exhibits a low-order power-law growth (Kirkpatrick, 1983). This stark difference highlights PSO's superior scalability for large-scale road network synthesis. However, PSO's susceptibility to premature convergence into local optima remains a critical limitation. This issue can be mitigated through further algorithmic improvements (Iqbal et al., 2015), or by a hybrid strategy: first using the SA algorithm to obtain promising objective function values, then filtering the solutions obtained by PSO. This two-stage strategy effectively combines SA's robustness in escaping local optima with PSO's efficiency, achieving both computational scalability and solution quality.

4.3. Virtual testing results

4.3.1 Lane-change scenario

Across the 108 test scenarios, comprising a total of 10,800 test runs, the probability of successful collision avoidance is presented in Table 8.

Table 8. Probability of successful collision avoidance

Test scenario ($TTC(s)$)	aggressive strategy	moderate strategy	proactive strategy
0.5	0.77	0.94	0.94
1	0.82	0.94	0.94
1.5	0.93	0.94	0.94
2	0.95	0.95	0.95
3	0.96	-	0.96
5	0.96	-	-
10	0.97	-	-
15	0.97	-	-
30	0.98	-	-

As indicated in Table8, under the aggressive strategy, the average probability of successful collision avoidance for the two high-risk test scenarios with TTC = 0.5 s and 1.0 s are 77% and 81% respectively. These values are significantly lower than those observed in test scenarios with TTC ≥ 1.5 s (exceeding 93%), revealing the performance limitations of the tested autonomous driving algorithm.

However, even in the low-risk scenario with $TTC = 30s$, the tested autonomous driving algorithm still exhibits a 2% probability of collision occurrence, indicating functional deficiencies in certain specific scenarios.

When employing the moderate and proactive strategies, the probabilities of successful collision avoidance demonstrate substantial improvement, exceeding 94% and maintaining relative stability across all test scenarios. This enhancement is attributed to the earlier intervention of the autonomous driving algorithm compared to the aggressive strategy, demonstrating the algorithm's capability to proactively perceive potential collision risks and execute appropriate avoidance responses.

Table 9 presents the maximum deceleration values and their occurrence times for VUT in collision scenarios.

Table 9. Maximum deceleration and occurrence times in collision scenarios

Test scenario	aggressive strategy		moderate strategy		proactive strategy	
($TTC(s)$)	$a_m(m/s^2)$	μ_m	$a_m(m/s^2)$	μ_m	$a_m(m/s^2)$	μ_m
0.5	-19.59	0.47	-19.61	0.12	-19.61	0.12
1	-19.70	0.17	-19.72	0.12	-19.59	0.13
1.5	-19.69	0.12	-19.67	0.11	-19.67	0.12
2	-19.65	0.14	-19.65	0.14	-19.64	0.14
3	-18.59	0.13	-		-19.77	-0.14
5	-16.93	0.13	-		-	
10	-19.09	0.064	-		-	
15	-16.67	0.084	-		-	
30	-16.39	0.17	-		-	

In Table9, μ_m denotes the normalized timing of maximum deceleration occurrence during the collision avoidance response.

$$\mu_m = \frac{t_m}{t_c - t_a} \tag{29}$$

where t_m represents the absolute time instant when the maximum deceleration occurs, t_c denotes the collision time between the two vehicles, and t_a indicates the time instant when the autonomous driving algorithm assumes control of the VUT.

As observed from Table 9, regardless of the intervention strategy employed, the maximum deceleration value across all test scenarios exhibits remarkable consistency, consistently exceeding $-16m/s^2$ and the maximum deceleration predominantly occurred during the early phase of the collision avoidance response. This observation indicates that the VUT could promptly identify potential collision risks in all scenarios and attempt to minimize vehicle speed as rapidly as possible to prevent collision.

Table 10. Average deceleration and proportion of deceleration duration in collision scenarios

Test scenario	aggressive strategy		moderate strategy		proactive strategy	
($TTC(s)$)	$\bar{a}(m/s^2)$	μ_d	$\bar{a}(m/s^2)$	μ_d	$\bar{a}(m/s^2)$	μ_d
0.5	-6.08	0.96	-8.44	0.96	-8.44	0.97
1	-7.64	0.97	-7,30	0.96	-7.79	0.96
1.5	-7.45	0.96	-8.48	0.,96	-8.66	0.97
2	-6.78	0.95	-7.35	0.95	-7.34	0.96
3	-4.03	0.95	-		-3.88	0.95
5	-2.18	0.94	-		-	
10	-2.84	0.94	-		-	
15	-2.17	0.95	-		-	
30	-1.74	0.87	-		-	

Table 10 presents the average deceleration during collision avoidance responses and the proportion of deceleration duration relative to the total control duration of the autonomous driving algorithm in collision scenarios. The average deceleration generally exhibits a decreasing trend as the preset TTC of the scenarios increases.

The parameter μ_d represents the proportion of deceleration duration relative to the total control duration of the autonomous driving algorithm, defined as:

$$\mu_d = \frac{t_d}{t_c - t_a} \tag{30}$$

where t_d denotes the deceleration duration.

Except for the scenario with $TTC = 30s$, the deceleration accounts for more than 94% of the autonomous driving response duration.

These findings demonstrate that the VUT exhibits greater average deceleration with more aggressive collision avoidance responses in higher-risk scenarios, whereas in lower-risk scenarios, the VUT tends to adopt a more moderate deceleration profile.

Table 11. The relative positions of the two vehicles at the moment of collision

Test scenario	aggressive strategy		moderate strategy		proactive strategy	
($TTC(s)$)	$x(m)$	$y(m)$	$x(m)$	$y(m)$	$x(m)$	$y(m)$
0.5	3.49	1.75	2.48	1.92	2.51	1.92
1	2.96	1.99	2.55	1.99	2.48	1.92
1.5	2.84	1.96	2.53	2.03	2.49	2.04
2	2.95	2.08	2.80	1.99	2.81	1.99
3	2.83	2.02	-		2.06	2.00
5	3.34	1.91	-		-	
10	3.35	1.93	-		-	
15	3.41	1.93	-		-	
30	3.18	1.96	-		-	

As shown in Table 11, regardless of the intervention strategy employed, the average spacing in x-direction between the two vehicles at the moment of collision remains below the reference vehicle's length but exceeds half of its length across all test scenarios. Except for the scenarios with $TTC = 0.5s$, the average spacing in y-direction between the two vehicles at collision consistently slightly exceeds the reference vehicle's width. These results indicate that in the majority of collision, the two vehicles experience lateral collisions, with the collision point located on the rear side of the reference vehicle's lateral surface, approaching the geometric center.

Table 12 The angular difference between the two vehicles at the moment of collision

Test scenario ($TTC(s)$)	aggressive strategy	moderate strategy	proactive strategy
0.5	3.20°	3.37°	3.38°
1	2.70°	2.68°	2.70°
1.5	4.58°	4.49°	4.56°
2	3.72	3.62°	3.59°
3	3.09°	-	2.89°
5	3.77°	-	-
10	2.92°	-	-
15	4.65°	-	-
30	3.96°	-	-

Table 12 presents the angular difference between the travel directions of the two vehicles at the moment of collision. Across all test scenarios and all the intervention strategies, the average angular difference between the two vehicles' travel directions at collision consistently remains below 5°.

The aforementioned analysis demonstrates that, irrespective of the intervention strategy implemented, the tested autonomous driving algorithm exhibits insufficient collision avoidance capability when confronted with the lateral approach condition between the two vehicles, as illustrated in Fig. 16.

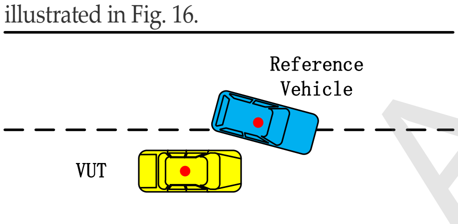


Fig.16. Schematic diagram of the lateral approach condition

We proceed to analyze the collision avoidance responses in non-collision scenarios, thereby exploring the performance boundaries of the autonomous driving algorithm.

Table 13 presents the average TTC margins in non-collision scenarios. Under three distinct intervention strategies, the TTC margins consistently increase with the preset TTC of the scenarios.

Table 13. Average TTC margins in non-collision scenarios

Test scenario $TTC(s)$	aggressive strategy	moderate strategy	proactive strategy
0.5	0.10	1.01	1.29
1	0.46	1.14	1.50
1.5	0.82	1.39	1.80
2	1.21	1.64	2.09
3	1.99	-	2.67
5	3.35	-	-
10	4.98	-	-
15	6.25	-	-
30	8.49	-	-

Under the aggressive intervention strategy, the collision time margin is minimal in the test scenario with $TTC = 0.5s$, nearing the limits of the autonomous driving algorithm's collision avoidance response capability. In lower-risk test scenarios with $TTC \geq 10s$, the autonomous driving algorithm permits the two vehicles to approach within $TTC < 10s$ to maintain a relatively tight spatial relationship between vehicles, ensuring traffic efficiency in realistic driving environments.

When employing the moderate and proactive intervention strategy, the minimum collision time margin in comparable scenarios is significantly higher than that observed under the aggressive strategy, indicating that earlier takeover of vehicle control by the autonomous driving algorithm can enhance safety margins in high-risk scenarios.

Table 14. Maximum deceleration and occurrence times in non-collision scenarios

Test scenario	aggressive strategy		moderate strategy		proactive strategy	
$TTC(s)$	$a_m(m/s^2)$	μ_m	$a_m(m/s^2)$	μ_m	$a_m(m/s^2)$	μ_m
0.5	-19.77	0.18	-20.06	0.15	-19.61	0.14
1	-19.80	0.19	-19.83	0.16	-19.83	0.14
1.5	-19.76	0.21	-19.82	0.17	-19.83	0.15
2	-19.90	0.22	-19.94	0.20	-19.98	0.15
3	-19.78	0.23	-		-19.87	0.16
5	-19.80	0.17	-		-	
10	-19.17	0.14	-		-	
15	-19.14	0.14	-		-	
30	-19.14	0.13	-		-	

Table 14 presents the maximum deceleration and the corresponding occurrence time in non-collision scenarios. Compared to the test results from collision scenarios, the maximum deceleration values remain consistent. However, in non-collision scenarios, the occurrence time of maximum deceleration is delayed relative to collision scenarios. This temporal difference arises because collision scenarios terminate at the moment of impact, resulting in shorter overall scenario durations.

Table 15 displays the average deceleration and the deceleration duration ratio in non-collision scenarios. Relative to collision scenarios, the variation trend of average deceleration remains consistent.

Table15. Average deceleration and proportion of deceleration duration in non-collision scenarios

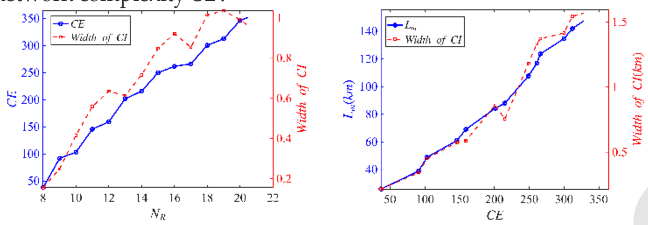
Test scenario	aggressive strategy		moderate strategy		proactive strategy	
$TTC(s)$	$\bar{a}(m/s^2)$	μ_d	$\bar{a}(m/s^2)$	μ_d	$\bar{a}(m/s^2)$	μ_d
0.5	-4.69	0.88	-4.72	0.86	-4.60	0.87
1	-4.89	0.88	-4.59	0.87	-4.48	0.87
1.5	-4.97	0.87	-4.58	0.88	-4.36	0.88
2	-4.79	0.87	-4.38	0.88	-4.05	0.89
3	-4.17	0.90	-		-3.86	0.92
5	-3.79	0.91	-		-	
10	-2.69	0.92	-		-	
15	-2.66	0.91	-		-	
30	-2.06	0.90	-		-	

The comprehensive analysis of both collision and non-collision scenarios demonstrate that the tested autonomous driving algorithm consistently detects collision risks and executes deceleration-based collision avoidance maneuvers across diverse scenarios. Specifically, the algorithm first rapidly reduces vehicle speed through substantial deceleration, followed by fine-tuning the velocity via moderate deceleration to establish a car-following state. The safety limit of the algorithm approaches $TTC = 0.5s$; however, significant functional deficiencies are observed in lateral approach conditions, presenting considerable safety risks that warrant further investigation and algorithmic refinement.

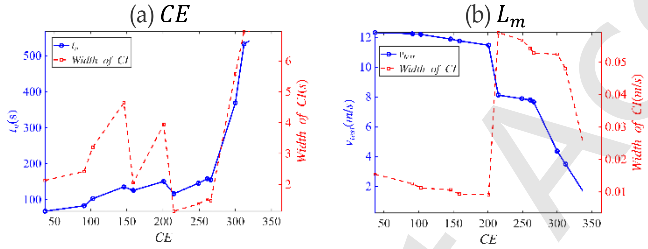
4.3.2 Free-driving scenario

Fig. 17(a) illustrates the variation in complexity across 15 synthesized road networks as the number of selected road elements increases. As the count of road instances escalates from 8 to 22, the mean complexity exhibits a substantial increase from approximately 38 to over 364. Notably, the width of the 95% confidence interval remains significantly smaller than the mean values, indicating high consistency in complexity among the generated road networks. The linear regression and residual analysis results presented in Fig.18(a) and (b) demonstrate that the complexity of generated road networks can be effectively regulated with an approximately linear trend by controlling the number of road elements within the network.

Fig.17(b) reveals that the testing mileage L_m in continuous free-driving scenarios increases because of the growing number of road elements in the network. As evidenced by Fig.18(c) and (d), a linear positive correlation exists between the testing mileage L_m and the network complexity CE.

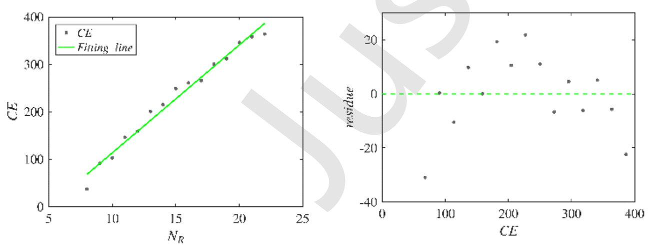


(a) CE (b) L_m

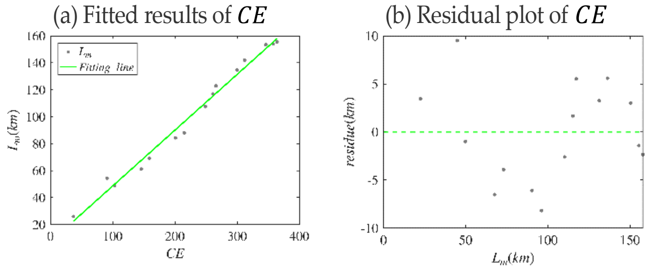


(c) t_u (d) v_{test}

Fig.17. Results of continuous free-driving test



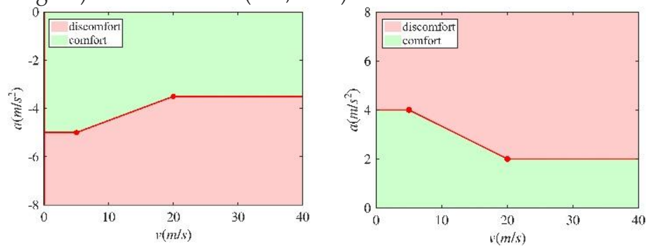
(a) Fitted results of CE (b) Residual plot of CE



(c) Fitted results of L_m (d) Residual plot of L_m

Fig.18. Results of linear fitting and residual analysis of CE and L_m

The duration of uncomfortable driving t_u and the average vehicle speed v_{test} serve as evaluation metrics for assessing the robustness of the test vehicle's control algorithm in continuous free-driving scenarios. Specifically, t_u is defined as the cumulative duration during which the tested vehicle exceeds the comfort threshold (indicated by the red line in Fig. 19) for acceleration (ISO, 2018).



(a) Comfort threshold in deceleration (b) Comfort threshold in acceleration

Fig.19. Comfort threshold in free-driving test (ISO, 2018)

Fig.17(c) and (d) reveal the correlation between network complexity and both the t_u and v_{test}. The narrow confidence interval width indicates high consistency across 100,000 Monte Carlo test results.

A persistent upward trend is observed in the uncomfortable duration as road network complexity increases, indicating a positive correlation between these variables. During the initial phase, t_u exhibits gradual growth, suggesting that the control algorithm demonstrates strong robustness in effectively suppressing abrupt acceleration/deceleration maneuvers. However, when the road network complexity exceeds 260, t_u experiences a sharp increase.

Fig. 20 presents the fitting results between t_u and CE, revealing a cubic polynomial nonlinear relationship. This indicates that the growth rate of t_u significantly exceeds the rate of complexity increase during this phase. This phenomenon occurs because, when confronted with road networks featuring increasingly dense distributions and more diverse types of intersecting road elements, the control algorithm's robustness markedly deteriorates. Consequently, the VUT frequently executes aggressive maneuvers such as abrupt acceleration/deceleration and large-magnitude steering adjustments, leading to a sudden surge in uncomfortable duration.

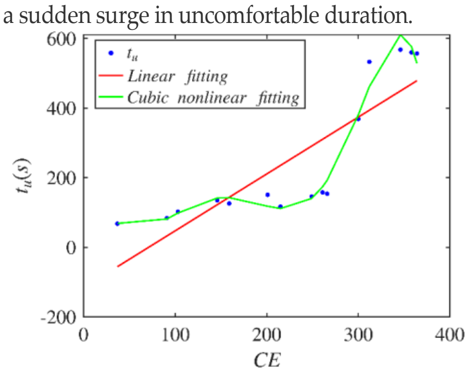


Fig. 20. Results of fitting of t_u

Fig. 17(d) demonstrates a negative correlation between v_{test} and CE. Despite the presence of fluctuations and nonlinear characteristics, the overall trend indicates that test velocity decreases as network complexity increases.

For road networks with fewer road elements, the topological structure is simple with low complexity, comprising ring-like structures formed by highways and trunk roads with high-speed configurations. Consequently, during this phase, the average velocity remains relatively stable at approximately 12 m/s. As additional intersecting road units are incorporated, v_{test} exhibits distinct inflection points characterized by a sharp decline.

Beyond inflection points, the increasing network complexity significantly compromises the robustness of the test vehicle's control algorithm. This degradation manifests through frequent execution of turning and U-turn maneuvers, triggering repeated deceleration behaviors, which substantially reduce the average test velocity.

Fig.21 presents the Pearson correlation coefficients among CE, L_m, t_u, and v_{test}. The correlation coefficient between CE and L_m is 0.992. The significance test confirms $p < 0.001$, indicating extremely high statistical reliability of this correlation. The 95% confidence interval (CI) is [0.976, 0.9973], which falls within the range of high positive values, indicating nearly perfect positive correlation. The correlation coefficient between road network complexity and discomfort duration is 0.838 ($p < 0.001$), with a 95% CI of [0.572, 0.945], indicates a strong statistically significant positive association. For CE and v_{test}, the correlation coefficient of is -0.886, with a 95% CI of [-0.962, -0.685], establishes a strong statistically significant inverse relationship ($p < 0.001$). The significant correlations observed among t_u, v_{test}, and CE further substantiate that road network complexity can serve as a quantitative metric for assessing the challenge level of test scenarios, thereby facilitating the evaluation of control algorithm robustness.

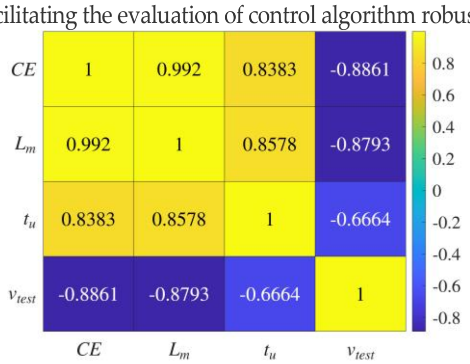


Fig. 21. Pearson correlation coefficient among CE, L_m, t_u, and v_{test}.

To evaluate the testing mileage compression capability of synthesized road networks, real-world road networks in Xi'an are employed for comparative continuous free-driving tests. In real world map, the VUT initiated free-driving testing from randomly selected road element, recording trajectories and travel distances. Experimental conditions matched generated road network testing to ensure comparability. The mileage compression ratio R_c was calculated as:

$$R_c = -\frac{L_g - L_r}{L_r} \times 100\% \tag{31}$$

where L_r and L_g denote average travel distances for homogeneous-element trajectories in real and generated networks, respectively.

Fig.22 presents the average compression ratio R_c for trajectories ranging from 1 to 23 hops, where higher R_c indicate more significant test mileage reduction. Single-hop trajectories achieve an average compression ratio of approximately 80%, demonstrating substantial mileage compression effectiveness. From a statistical perspective, the utilization of generated road networks significantly reduces test mileage, with an average compression ratio of 32.4%. The compression ratio exhibits a trend of initial decline followed by stabilization. When the hop count exceeds 16, the compression ratio stabilizes at approximately 17%, indicating that efficiency optimization can be consistently achieved even for longer trajectories.

Real-world roads are not constrained by the road engineering specifications or testing requirements considered in this study, featuring shorter edge lengths and denser spacing of road elements. This results in negative single-hop compression ratios for certain trajectory samples. As trajectory length increases, the probability of encountering such samples rises, consequently affecting the mileage compression ratio for longer trajectories.

Confidence interval analysis reveals that shorter trajectories exhibit narrower confidence interval widths for R_c with lower data dispersion, indicating stable mileage compression performance. In contrast, longer trajectories demonstrate increased confidence interval widths due to greater trajectory diversity, resulting in elevated statistical dispersion of mileage compression effectiveness.

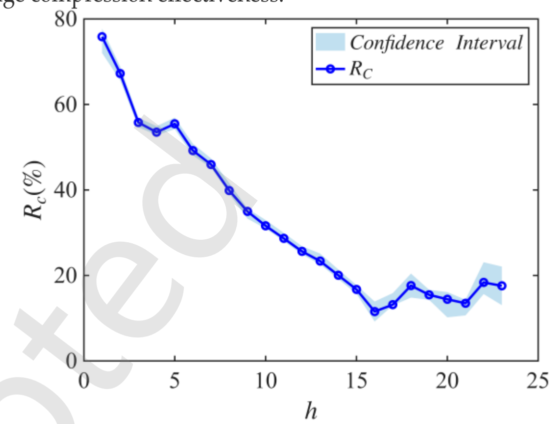


Fig.22. The mileage compression ratio and confidence interval

4.4. Limitations

Despite the encouraging performance demonstrated in this work, several limitations of the proposed method deserve to be discussed.

First, although our constructed intersecting road element models cover 92.99% of actual intersection types within the Xi'an Ring Expressway (inclusive), several specialized road elements remain unaccounted for, such as roundabouts, interchanges, and multi-branch intersections (with more than four branches). This limitation primarily stems from the relatively low occurrence frequency and substantial geometric variations of these road elements, which pose challenges in abstracting them into representative road element models. This gap will be addressed in our future researches.

Second, in the road network generation, complex road elements from real-world networks are simplified into polyline representations, which are subsequently reconstructed into multi-lane roads during scenario map construction. This simplification-reconstruction pipeline inevitably neglects certain geometric details inherent to actual road elements, such as road curvature, abrupt variations in lane count and width, and non-parallel lane configurations. The adoption of average relative error (FE) as the sole evaluation metric enables assessment of geometric similarity between generated and real road elements, but fails to quantify fidelity at the lane scale. Future work will incorporate these factors into the scenario map construction, which is expected to enhance both the geometric fidelity and structural diversity of generated road networks, thereby facilitating the creation of more rare test roads.

Third, the complexity metric is quantified based on the number of potential path conflicts among traffic participants within road elements, under the premise that inherent path conflicts contribute to collision risks. For instance, in a Cross-junction element, left-turning trajectories inherently conflict with opposing through movements, presenting collision risks. However, traffic participant trajectories are defined based on normal traffic flow patterns prescribed by road design standards, without accounting for anomalous movement behaviors observed in real-world traffic, such as wrong-way driving, reversing maneuvers, or illegal pedestrian crossings. Path conflicts resulting from these anomalous trajectories are excluded from the complexity metric computation.

Finally, the proposed road complexity metric quantitatively characterizes one critical factor influencing the challenge level of test scenarios at the road layer. The overall challenge level of test scenarios may also be influenced by road facilities and temporary road modifications (layers 2 and 3 of the 5-layer scenario model) as well as dynamic interactions among traffic participants. These factors and their interrelationships will be investigated in our subsequent researches.

5 Conclusion

This study proposes a complexity controllable road network generation approach based on the optimized combination of realistic road elements for virtual simulation testing. First, real-world urban road networks are decomposed into heterogeneous road elements, in which twenty-two high-risk road elements across five categories abstracted into parameterized graph models. Road element instantiation is performed under engineering specifications and testing requirements using real-cartographic data, ensuring geometric fidelity to real-world counterparts. Subsequently, a complexity-controllable road network generation strategy is introduced. The number of potential path conflicts in each road instance is adopted as the quantitative metric for evaluating its complexity. The complexity of road networks is regulated by selecting road instances with varying complexity. Then, a network optimization model is formulated. With the realism and compactness of the road network set as optimization objectives, the selected road instances are interconnected in a non-intersecting and non-overlapping manner to generate road networks that meet the expected complexity levels. Finally, 4,500 synthetic road networks across 15 complexity levels are generated using Xi'an cartographic data. The effectiveness of the generated road networks is validated through comparative virtual simulation tests performed on both the generated road networks and Xi'an's real-world counterparts. Experimental results demonstrate that (1) The generated road network exhibit high environmental fidelity. The mean relative error between the geometric parameters of road elements in the generated networks and the expected values derived from real-world cartographic data clustering is less than 1.5%, confirming capability of faithfully replicating the geometric features of real-world; (2) Multi-risk-level testing scenarios constructed based on generated road networks reveal functional deficiencies and performance boundaries of the target autonomous driving algorithm. In lane-change scenarios, virtual simulation tests leveraging the synthesized road networks successfully capture both the collision avoidance performance limits of the algorithm and its functional defects under lateral approach conditions; (3) A strong correlation exists between the complexity of the generated road networks and testing challenge intensity. Network complexity shows an absolute Pearson's correlation coefficient r>0.84 with uncomfortable driving duration and the average vehicle speed in free-driving scenarios. This finding confirms that higher-complexity road networks correspond to more rigorous testing challenges, which aligns with the scenario challenge adjustment trend; (4) The generated road networks significantly enhance the efficiency of continuous scenario testing. Compared to real-world road networks, the generated road networks achieve 32.4% average mileage compression when traversing an equivalent number of road elements, significantly improving continuous scenario testing efficiency.

Future research will focus on expanding the road element sets to include complex topologies such as roundabouts, grade-separated interchanges, and bridges.

Nomenclature

PCG	procedural content generation
VUT	vehicle under test
BIG	block incremental generation
LCTM	Latent Class Trajectory Model
$g(N,E)$	Road element model, N denotes the centroid and endpoints, E represents the edges
p_N	positional parameters of centroid and endpoints(m)
a_N	attribute parameters of centroid and endpoints
l_E	length parameter of edges(m)
a_E	angular parameter of edges(rad)
E_{mp}	pre-merging component of merge element
E_{ma}	post-merging component of merge element
E_{mr}	ramp component of merge element
l_1	sight distance ahead of merging nose in E_{mp}(m)
l_2	sight distance ahead of merging nose in E_{mr}(m)
v_0	initial speed(km/h)
v_1	intended test speed(km/h)
v_2	predefined terminal velocity(km/h)
v_0'	initial speed in ramp(km/h)
v_1'	intended exit speed in ramp(km/h)
t_h	duration of potential test scenarios(s)
a	acceleration of VUTs(m/s²)
l_{sc}	length of speed-change lane(m)
l_{ts}	length of transition section(m)
l_c	length for the ramp transition curve(m)
n_c	connecting elements in road network
$l_{E_{mp}}$	length of E_{mp}(m)
$l_{E_{mr}}$	length of E_{mr}(m)
$l_{E_{ma}}$	length of E_{ma}(m)
E_{dp}	pre-diverging component of diverge element
E_{dr}	diverging component of diverge element
E_{da}	post-diverging component of diverge element
$l_{E_{dp}}$	length of E_{dp}(m)
$l_{E_{dr}}$	length of E_{dr}(m)
$l_{E_{da}}$	length of E_{da}(m)
l_{id}	exit recognition sight distance(m)

E_{in}	approaching component of T/Cross - junction
E_{out}	departing component of T/Cross - junction
l_{st}	stopping sight distance in T/Cross - junction(m)
l_{E_j}	length of component of T/Cross - junction(m)
a_{in}	acceleration of VUTs within E_{in}(m/s^2)
a_{out}	acceleration of VUTs within E_{out}(m/s^2)
a_{en}	acceleration of VUTs within entry phase of non-intersection element(m/s^2)
a_{te}	acceleration of VUTs within termination phase of non-intersection element(m/s^2)
g_s	road element instance
δ_l	error of length parameter of edges(m)
δ_a	error of angular parameter of edges(rad)
N_R	number of road elements in road network
G_s	road instances set
e	edge in road network
N_c	number of n_c

Author contributions

The individual contributions of authors to the manuscript should be specified in this section. The Journal recommends authors to use CRediT statements (the Contribution Roles Taxonomy).

Yu Zhu: Writing– original draft, Writing– review & editing, Visualization, Validation, Software, Methodology, Formal analysis, Conceptualization, Data curation.
Jiaxin Wang: Writing– review & editing, Methodology, Formal analysis, Investigation, Conceptualization.
Shaoxin **Yuan:** Writing– review & editing, Methodology, Formal analysis, Conceptualization.
Zhigang Xu: Writing– review & editing, Visualization, Validation, Methodology, Conceptualization, Data curation, Funding acquisition.
Xiaobo Qu: Validation, Supervision.

Replication and data sharing

The source codes and replication package are available on ETS data at https://doi.org/10.26599/ETSD.2026.9190017.

Acknowledgements

This study is partially supported by the National Natural Science Foundation of China (Grant No.52432013 and Grant No.U23A20682).

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article. The author Diange Yang is the Editorial Board Member of this journal. The author Zhigang Xu is the Editor-in-Chief of this journal.

Declaration of generative AI and AI-assisted technologies in the writing process (if applicable)

During the preparation of this work, the authors used Qwen for language enhancement only. After utilizing this tool/service, the authors thoroughly reviewed and edited the content as necessary and take full responsibility for the final content of the published article.

References

Aliaga, D. G., Vanegas, C. A., Beneš, B., 2008. Interactive example-based urban layout synthesis. ACM Transactions on Graphics, **27**, 1–10.

Badhrudeen, M., Derrible, S., Verma, T., Kermanshah, A., Furno, A., 2022. A geometric classification of world urban road networks. Urban Science, **6**, 11.

Becker, D., Ruß, F., Geller, C., Eckstein, L., 2020. Generation of complex road networks using a simplified logical description for the validation of automated vehicles, In: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), 1–7.

Beneš, J., Wilkie, A., Křivánek, J., 2014. Procedural modelling of urban road networks. Computer Graphics Forum, **33**, 132–142.

Chen, G., Esch, G., Wonka, P., Müller, P., Zhang, E., 2008. Interactive procedural street modeling. ACM Transactions on Graphics, **27**, 1–10.

Chen, J., Li, R., Xing, X., Xiong, L., 2020. Road horizontal alignment design method for continuous loop area in autonomous vehicles proving ground based on genetic algorithms. In: CICTP 2020, 1943–1954.

Chen, R., Arief, M., Zhao, D., 2018. An "Xcity" optimization approach to designing proving grounds for connected and autonomous vehicles. arXiv. https://arxiv.org/abs/1808.03089.

Chen, R., Arief, M., Zhang, W., Zhao, D., 2021. How to evaluate proving grounds for self-driving? A quantitative approach. IEEE Transactions on Intelligent Transportation Systems, 22, 5737–5748.

China Road and Bridge Corporation., 2021. Specifications for geometric design of highways (JTG D20—2017). Beijing: China Communication Press.

De Gelder, E., Paardekooper, J. P., 2017. Assessment of automated driving systems using real-life scenarios. In: 2017 IEEE Intelligent Vehicles Symposium (IV), 589–594.

Eberhart, R., Kennedy, J., 1995. A new optimizer using particle swarm theory. In: MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 39–43.

Elrofai, H., Worm, D., Op den Camp, O., 2016. Scenario identification for validation of automated driving functions. In: Advanced Microsystems for Automotive Applications 2016. T. Schulze, B. Müller, G. Meyer (Eds.). Basel, Switzerland: Springer International Publishing, 153–163.

Fischerson M. C., 1984. City traffic. Beijing: China Architecture & Building Press.

Gambi, A., Mueller, M., Fraser, G., 2019a. AsFault: Testing self-driving car software using search-based procedural content generation. In: 2019 IEEE/ACM 41st International Conference on Software Engineering: Companion Proceedings (ICSE-Companion), 27–30.

Gambi, A., Mueller, M., Fraser, G., 2019b. Automatically testing self-driving cars with search-based procedural content generation. In: Proceedings of the 28th ACM SIGSOFT International Symposium on Software Testing and Analysis, 318–328.

Kirkpatrick, S., Gelatt, C. D., Jr., & Vecchi, M. P., 1983. Optimization by simulated annealing. Science, **220**(4598), 671–680.

Li, G., Shi G., 2010. Procedural modeling of urban road network. In: Proceedings of the 2010 International Forum on Information Technology and Applications, 75–79.

Hartmann, S., Weinmann, M., Wessel, R., Klein, R., 2017. StreetGAN: Towards road network synthesis with generative adversarial networks. In: 25th International Conference in Central Europe on Computer Graphics, Visualization and Computer Visionin co-operation with EUROGRAPHICS Association, 133-142.

Ikram, Z., Muktadir, G. M., Whitehead, J., 2023. Procedural generation of complex roundabouts for autonomous vehicle testing. In: Proceedings of the 2023 IEEE Intelligent Vehicles Symposium (IV), 1–6.

International Organization for Standardization, 2018. IS0 15622:2018 Intelligent transport systems — Adaptive cruise control systems — Performance requirements and test procedures. https://www.iso.org/standard/71515.html.

Iqbal, N., Zerguine, A., Al-Dhahir, N., 2015. Decision Feedback Equalization using Particle Swarm Optimization. Signal Processing, **108**, 1–12.

Kalra N., Paddock, S. M., 2016. Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? Transportation Research Part A: Policy and Practice, **94**, 182–193.

Kelly, G., McCabe, H., 2007. Citygen: An interactive system for procedural city generation. In: Fifth International Conference on Game Design and Technology, 8–16.

Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P., 1983. Optimization by Simulated Annealing. Science, **220**, 671–680.

Li, H., Li, C., Wang, A., et al., 2020. Associations between social and intellectual activities with cognitive trajectories in Chinese middle-aged and older adults: A nationally representative cohort study. Alzheimer's Research Therapy, **12**, 115.

Li, J., Lyu, L., Shi, J., Zhao, J., Xu, J., Gao, J., He, R., Sun, Z., 2022. Generating community road network from GPS trajectories via style transfer. In: Proceedings of the 30th International Conference on Advances in Geographic Information Systems, 1–4.

Li, Q., Peng, Z., Zhang, Q., Liu, C., Zhou, B., 2020. Improving the generalization of end-to-end driving through procedural generation. arXiv. https://arxiv.org/abs/2012.13681.

Liu, C., Xu, Z., Liu, Z., Li, X., Zhang, Y., 2024. Decentralized multi-vehicle motion planning for platoon forming in mixed traffic using Monte Carlo tree search. IEEE Transactions on Intelligent Transportation Systems, **25**, 20872–20888.

Liu, C., Zhang, W., 2022. Learning the driver acceleration/deceleration behavior under high-speed environments from naturalistic driving data. IEEE Intelligent Transportation Systems Magazine, **14**, 78–91.

Moosavi, S., Samavatian, M. H., Parthasarathy, S., Ramnath, R., 2019. A countrywide traffic accident dataset. arXiv.

https://arxiv.org/abs/1906.05409.

Muktadir, G.M., Jawad, A., Paranjape, I., Whitehead, J., Shepelev, A., 2022. Procedural generation of high-definition road networks for autonomous vehicle testing and traffic simulations. SAE International Journal of Connected and Automated Vehicles, **6**, 99–120.

Muktadir, G. M., Jawad, A., Shepelev, A., Paranjape, I., Whitehead, J., 2022. Realistic road generation: Intersections. https://doi.org/10.13140/RG.2.2.30541.51683.

Nishida, G., Garcia-Dorado, I., Aliaga, D. G., 2016. Example-driven procedural urban roads. Computer Graphics Forum, **35**, 5–17.

Opletal, J. 2025, July 24. China's massive ADAS test: 36 cars, 15 hazard scenarios, 216 crashes. CarNewsChina.Com. https://carnewschina.com/2025/07/24/chinas-massive-adas-test-36-cars-15-hazard-scenarios-216-crashes/

Parish, Y. I. H., Müller, P., 2001. Procedural modeling of cities. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, 301–308.

Sang M, J. Z., Li, H., 2024. Efficient adversarial scenario test for autonomous vehicles. Journal of Integration Technology, **13**, 15-28.

Tang, S., Zhang, Z., Zhou, J., Zhou, Y., Li, Y.-F., Xue, Y., 2023. EvoScenario: Integrating road structures into critical scenario generation for autonomous driving system testing. In: Proceedings of the 2023 IEEE 34th International Symposium on Software Reliability Engineering (ISSRE), 309–320.

Ulbrich, S., Menzel, T., Reschka, A., Schuldt, F., Maurer, M., 2015. Defining and substantiating the terms scene, situation, and scenario for automated driving. In: IEEE International Conference on Intelligent Transportation Systems (ITSC), 982–988.

Weber, H., Bock, J., Klimke, J., Roesener, C., Hiller, J., Krajewski, R., et al., 2019. A framework for definition of logical scenarios for safety assurance of automated driving. Traffic Injury Prevention, **20**, 65–70.

Xing, X., Li, R., Zhou, R., Chen, J., 2021. Optimization method for the layout design of urban roads for autonomous vehicle tests considering scenario coverage. In: 2021 CHINA SAE CONGRESS PROCEEDINGS 1, 124-130.

Xu Z., Wang, G., Zhai, S., Liu, P., 2025. When automation fails: Examining the effect of a verbal recovery strategy on user experience in automated driving. International Journal of Human–Computer Interaction, **41**, 3772–3782.

Xu Z., Wang, M., Zhang, F., Jin, S., Zhang, J., Zhao, X., 2017. PaTAVTT: A hardware-in-the-loop scaled platform for testing autonomous vehicle trajectory tracking. Journal of Advanced Transportation, 2017, 1–11.

Xu Z., Zhang, K., Min, H., Wang, Z., Zhao, X., Liu, P., 2018. What drives people to accept automated vehicles? Findings from a field experiment. Transportation Research Part C: Emerging Technologies, **95**, 320–334.

Zao Y., Zou, Z., Shi, Z., 2024. Road graph extraction via transformer and topological representation. IEEE Geoscience and Remote Sensing Letters, **21**, 1–5.

Yang F., Lu, Y., Chen, B., Qin, P., Peng, X., 2024. RoadGen: Generating road scenarios for autonomous vehicle testing. arXiv. https://arxiv.org/abs/2411.19577.

Ye, Q., Li, Y., Shen, W., Xuan, Z., 2023. Division and analysis of accident-prone areas near highway ramps based on spatial autocorrelation. Sustainability, **15**, 7942.

Kong, Y., Wu, H., Akram, A., Li, J., Zhao, J., Wang. S., 2023. UAV LiDAR data-based lane-level road network generation for urban scene HD maps. IEEE Geoscience and Remote Sensing Letters, **20**, 1–5.

Zhang, Y., Li, X., Wang, A., Bao, T., Tian, S., 2015. Density and diversity of OpenStreetMap road networks in China. Journal of Urban Management, **4**, 135–146.

Zhang, Y., Xiong, Z., Zang, Y., Wang, C., Li, J., Li, X., 2019. Topology-aware road network extraction via multi-supervised generative adversarial networks. Remote Sensing, **11**, 1017.

Zhao, D., Lam, H., Peng, H., Bao, S., LeBlanc, D.J., Nobukawa, K., et.al, 2017. Accelerated Evaluation of Automated Vehicles Safety in Lane-Change Scenarios Based on Importance Sampling Techniques. IEEE Transactions on Intelligent Transportation Systems, **18**, 595–607

Zhao, Y., Ning, Y., Li, H., Liao, Z., Liu, Y., Li, F., 2023. MSC-DeepFM: OSM road type prediction via integrating spatial context using DeepFM. Sustainability, **15**, 16671.

Zhu, B,. Fan, T., Zhao, W., Li, W., Zhang, P., 2025. Continuous test scenario complexity evaluation method for automated driving vehicles. Journal of Jilin University (Engineering and Technology Edition), **55**, 456-467.

Zhu, Y., Xu, Z., Zhao, X., Wang, R., 2024. TsGAN-based automatic generation algorithm of lane-change cut-in test scenarios on expressways for autonomous vehicles. Journal of South China University of Technology (Natural Science Edition), **52**, 76–88.

Zhu, Y., Zhao, X., Xu, Z., Wang R., 2022. Automatic generation algorithm of lane-change virtual test scenario on highways for automated vehicles using Monte Carlo simulation. China Journal of Highway and Transport, 35, 89–100

Author biography

Yu Zhu received the M.S degree in Electrical Engineering from Dresden University of Technology, Dresden, Germany. He is currently pursuing a doctoral degree in traffic information engineering and control. His research interests include autonomous vehicle testing and intelligent transportation systems.

Jiaxin Wang received the M.S. degree in transportation engineering from Southwest Jiaotong University, Chengdu, China. He is currently pursuing a doctoral degree in traffic information engineering and control. His research interests include intelligent transportation systems and connected automated vehicles.

Shaoxin Yuan received the Ph.D. degree in Intelligent Transportation and Information Systems from Chang'an University, Xi'an. China. As a visiting scholar at the University of Washington (2005). His research focuses on transportation data modeling, traffic analysis, and workflow middleware for intelligent systems.

Zhigang Xu received the M.S. and Ph.D. degrees in traffic information engineering and control from Chang' an University, Xi'an. China. He had worked in the University of California, Davis, as a Visiting Scholar. He is currently a professor, doctoral supervisor, and vice dean of the School of Information Engineering. His researches focus on connected and autonomous vehicles, image processing, and intelligent transportation systems.

Xiaobo Qu is a Changjiang Chair Professor in Intelligent Transportation, Tsinghua University in China. His research is focused on improving large, complex, and interrelated urban mobility systems by integrating with emerging technologies. He has published over 180 journal articles published at top tier journals in the area of transportation. He is an elected member in Academia Europaea-The Academy of Europe since 2020.