



A multi-perspective fusion model for operating speed prediction on highways using knowledge-enhanced graph neural networks

Downloaded from: <https://research.chalmers.se>, 2025-11-29 04:09 UTC

Citation for the original published paper (version of record):

Gao, J., Yu, B., Chen, Y. et al (2025). A multi-perspective fusion model for operating speed prediction on highways using knowledge-enhanced graph neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 40(8): 1004-1027. <http://dx.doi.org/10.1111/mice.13382>

N.B. When citing this work, cite the original published paper.



A multi-perspective fusion model for operating speed prediction on highways using knowledge-enhanced graph neural networks

Jianqiang Gao¹ | Bo Yu¹ | Yuren Chen¹ | Kun Gao² | Shan Bao^{3,4}

¹Key Laboratory of Road and Traffic Engineering of the Ministry of Education, College of Transportation Engineering, Tongji University, Shanghai, China

²Department of Architecture and Civil Engineering, Chalmers University of Technology, Gothenburg, Sweden

³Industrial and Manufacturing Systems Engineering Department, University of Michigan, Dearborn, Michigan, USA

⁴Human Factors Group, University of Michigan Transportation Research Institute, Ann Arbor, Michigan, USA

Correspondence

Bo Yu, Key Laboratory of Road and Traffic Engineering of the Ministry of Education, College of Transportation Engineering, Tongji University, 4800 Cao'an Highway, Shanghai 201804, China.

Email: boyu@tongji.edu.cn

Kun Gao, Department of Architecture and Civil Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden.

Email: gkun@chalmers.se

Funding information

National Natural Science Foundation of China, Grant/Award Number: 52102416; National Key R & D Program of China, Grant/Award Number: 2023YFE0202400; Natural Science Foundation of Shanghai, Grant/Award Number: 22ZR1466000

Abstract

This study proposes a multi-perspective fusion model for operating speed prediction based on knowledge-enhanced graph neural networks, named RoadGNN-S. By utilizing message passing and multi-head self-attention mechanisms, RoadGNN-S can effectively capture the coupling impacts of multi-perspective alignment elements (i.e., two-dimensional design, 2.5-dimensional driving, and three-dimensional spatial perspectives). The results of driving simulation data show that root mean squared error, mean absolute error, mean absolute percentage error, and *R*-squared values of RoadGNN-S are superior to those of other classic deep learning algorithms. Then, prior knowledge (i.e., highway geometry supply, driver expectations, and vehicle dynamics) is introduced into RoadGNN-S, and the models' prediction accuracy and transferability are verified by field observation experiments. Compared to the above data-driven models, knowledge-enhanced RoadGNN-S effectively avoids the fundamental errors, improving the *R*-squared value in predicting passenger cars' and trucks' operating speed by 7.9% and 10.7%, respectively. The findings of this study facilitate the intelligent highway geometric design with multi-perspective fusion and knowledge enhancement techniques.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Computer-Aided Civil and Infrastructure Engineering* published by Wiley Periodicals LLC on behalf of Editor.



1 | INTRODUCTION

Operating speed is the most commonly used metric for evaluating the consistency of highway geometric design, closely tied to driving safety (Xuesong Wang et al., 2021). It has been reported that the lack of geometric consistency can lead to unexpected events and high-speed variability, contributing to over 30% of road accidents (Camacho-Torregrosa et al., 2013; Karim & Adeli, 2003a). Operating speed is defined as the speed at which a majority of the drivers are observed operating their vehicles under free-flow conditions, reflecting their chosen speed in response to multi-perspective highway alignment (Al-Sahili & Dwaikat, 2019). The 85th percentile of the observed speed distribution is commonly used as a standard measure to represent the operating speed at specific locations or geometric features (AASHTO, 2018). However, highway alignment exhibits distinct geometric shapes in two-dimensional (2D) design, 2.5-dimensional (2.5D) driving, and three-dimensional (3D) spatial perspectives, and only relying on a single perspective may not adequately capture the variation characteristics of operating speed (F. Wang et al., 2020). Therefore, the fusion of multi-perspective highway alignment elements is of paramount importance in accurately predicting operating speed.

The impact of highway alignment on operating speed varies from different perspectives. The 2D design perspective reflects the most basic shape of highway alignment on the horizontal and vertical projection plane during the geometry design process (Casal et al., 2017). The impact of 2D alignment elements (e.g., tangents, curves, vertical slopes, etc.) on operating speed depends on the combination of horizontal and vertical alignment (Samant & Adeli, 2001). According to Marr's visual theory (Marr, 2010), the 2.5D driving perspective is an observational viewpoint of highway alignment centered on drivers' eyes, depicting how drivers dynamically perceive distance, lane width, and curvature. The 2.5D driving perspective lies between the 2D and 3D perspectives, where the highway alignment exhibits depth information as the road outline gradually narrows from the near to the distant (He et al., 2023). This visual discrepancy between actual and perceived highway alignment prompts drivers to dynamically adjust their speed to match their perceived alignment (B. Yu et al., 2019). In addition, highway alignment is essentially a 3D curve, where spatial curvature directly affects vehicle control and driving comfort (Marinelli et al., 2017). Nevertheless, existing studies mainly focus on a single perspective of highway alignment when predicting operating speed, resulting in a large deviation between the predicted and actual operating speed. For instance, operating speed prediction models based on 2D alignment elements have been demonstrated to be less suitable for combined

and closely spaced horizontal and vertical curve segments (Maji et al., 2020). The operating speed prediction models that are established separately based on 2.5D alignment elements (i.e., the length and curvature of the visual curve) and 3D spatial curvature make efforts to involve the impacts of longitudinal slope and vehicle power performance on operating speed (Adeli & Ghosh-Dastidar, 2004; J. Wang et al., 2022). Therefore, the coupling impacts of multi-perspective highway alignment on operating speed should be further examined and included in the prediction model of operating speed.

Different methods have been employed to predict operating speed, which can be classified into two types, including model-driven and data-driven methods. Model-driven methods, such as simple linear regression (Cardoso et al., 1998), multiple linear regression (Eboli et al., 2017), multiple nonlinear regression (Sil et al., 2020), generalized least squares (Martinelli et al., 2023), simultaneous equations (Himes & Donnell, 2010), and three-stage least-squares estimator (Gong & Stamatiadis, 2008), are used to establish spot-based and continuous operating speed prediction models of multi-lane highways. Among them, the spot-based operating speed prediction models divide the highway into different road segments and estimate the operating speed at the start, middle, and end points of each segment (AbuAddous, 2021). The continuous operating speed models can predict the operating speed at a series of dense and equidistant spots along the highway, using alignment elements obtained from neighboring preceding and subsequent segments along with the driver's current position (Cafiso & Cerni, 2012). Due to inherent assumptions and boundary constraints, model-driven methods perform well on road segments with well-separated horizontal curves and longitudinal slopes, but they often encounter challenges when dealing with complex and diverse combinations of horizontal and vertical alignment. In contrast, data-driven methods stand out in terms of the ability to uncover complex and nonlinear relationships between alignment elements and operating speed, so they can adapt to various alignment combinations and potentially improve accuracy through continuous learning (Jiang & Adeli, 2004; Semeida, 2014). Despite these advantages of data-driven methods, to our best knowledge, only artificial neural networks have been utilized to predict operating speeds of horizontal curve segments on two-lane rural highways (McFadden et al., 2001). Additionally, alternative modeling techniques, such as nonlinear methods, three-stage least squares, and artificial neural networks, do not significantly enhance prediction performance when compared to traditional linear regression models (Medina & Tarko, 2005). The strength of deep learning algorithms lies in powerful feature learning and the modeling of complex nonlinear relationships, thereby enhancing model



adaptability and robustness (Song et al., 2023). There are still some profound challenges in using deep learning algorithms to predict operating speed, such as multi-perspective highway alignment representation and model interpretability, which hinder the practical applications of current data-driven methods. Leveraging the advantages of unstructured data modeling and interpretability, graph neural networks (GNNs) may provide a potential solution to these issues (Zhou et al., 2020).

Integrating prior knowledge into deep learning algorithms has been employed to enhance the prediction accuracy and transferability of data-driven methods. Currently, for neural networks, physical knowledge is integrated by modifying input data (Fei et al., 2022; Karim & Adeli, 2003b), loss functions (Marino et al., 2021), and network architectures (Li et al., 2024; Sheng et al., 2024) for knowledge enhancement. Three physical principles of highway geometry supply, driver expectation, and vehicle dynamics should be adhered to accurately predict operating speed based on prior knowledge. The prior knowledge involves the whole process from drivers' perception of highway alignment to their decision-making of driving behavior. Highway alignment supply, including the superelevation and side friction provided by horizontal curves, ensures vehicle stability within a certain safe margin beyond the design speed (Donnell et al., 2016). Driver expectations are formed through long-term driving experience and memory, which can assist drivers in adjusting the vehicles' speed and acceleration in time (van der El et al., 2019). Vehicle dynamics primarily affect operational speed through centrifugal force and tractive force. The large centrifugal force generated when vehicles rapidly enter curves may lead to driver discomfort, which compels drivers to take self-protective deceleration maneuvers (Dhahir & Hassan, 2016). The traction force can overcome driving resistances (i.e., air resistance, rolling resistance, skidding resistance, and grade resistance) to propel the vehicle forward (Z. Xu et al., 2018). For engineering applications with safety as the main goal, a highly accurate and transferable operating speed prediction is necessary by integrating the above prior knowledge into data-driven methods.

Given the above, although previous studies have established operating speed prediction models from a single perspective of highway alignment, they undermine the coupling impacts of multi-perspective alignment elements on operating speed. In addition, existing data-driven operating speed prediction models still have limitations in the fusion of multi-perspective highway alignment elements and the interpretability and transferability of models. To fill the above research gaps, this study aims to propose a multi-perspective fusion model for predicting operating speed based on knowledge-enhanced GNNs,

named RoadGNN-S. A multi-perspective alignment segment (MPAS) graph is constructed to fuse the 2D, 2.5D, and 3D highway alignment elements. Then, using the driving simulation data, the operating speed prediction model is built by GNNs. Prior knowledge of highway geometry supply, driver expectations, and vehicle dynamics is further introduced into the operating speed prediction model, and the enhanced accuracy and transferability of the model are verified by field observation experiments. The contributions of the study are threefold:

1. By the graph topology and connectivity, the proposed MPAS graph achieves modular representation and the multi-perspective fusion of highway alignment, facilitating a more comprehensive understanding of highway alignment from multiple perspectives.
2. Based on GNNs, this study proposes an interpretable operating speed prediction model, named RoadGNN-S. This model can effectively capture the coupling impacts of multi-perspective alignment elements on operating speed and provide clear and interpretable results during the prediction process.
3. The prior knowledge is considered in GNNs to ensure the theoretical correctness of solutions. This knowledge-enhanced modeling can further improve the model's predication accuracy and transferability, compared with data-driven modeling.

2 | METHODOLOGY

2.1 | Problem definition

Operating speed prediction can be regarded as establishing the dynamic balance relationship among highway geometric supply, driver expectations, vehicle dynamics and operating speed in the traffic system composed of roads, drivers, vehicles, and environment. Therefore, operating speed can be predicted using deep learning algorithms that consider the interactions $\mathcal{I} = (I_g, I_d, I_v)$ of highway geometric supply, driver expectations, and vehicle dynamics to establish a mapping between the input highway alignment element $\mathcal{E} = (v_0, \dots, v_p)$ and output operating speed $S = (V_{85}^0, \dots, V_{85}^Q)$.

2.2 | Overall framework

As illustrated in Figure 1, the whole methodology framework consists of three main parts, including the MPAS graph construction, the GNNs-based operating speed prediction model (RoadGNN-S), and knowledge-enhanced modeling. The purpose of MPAS graph construction is to

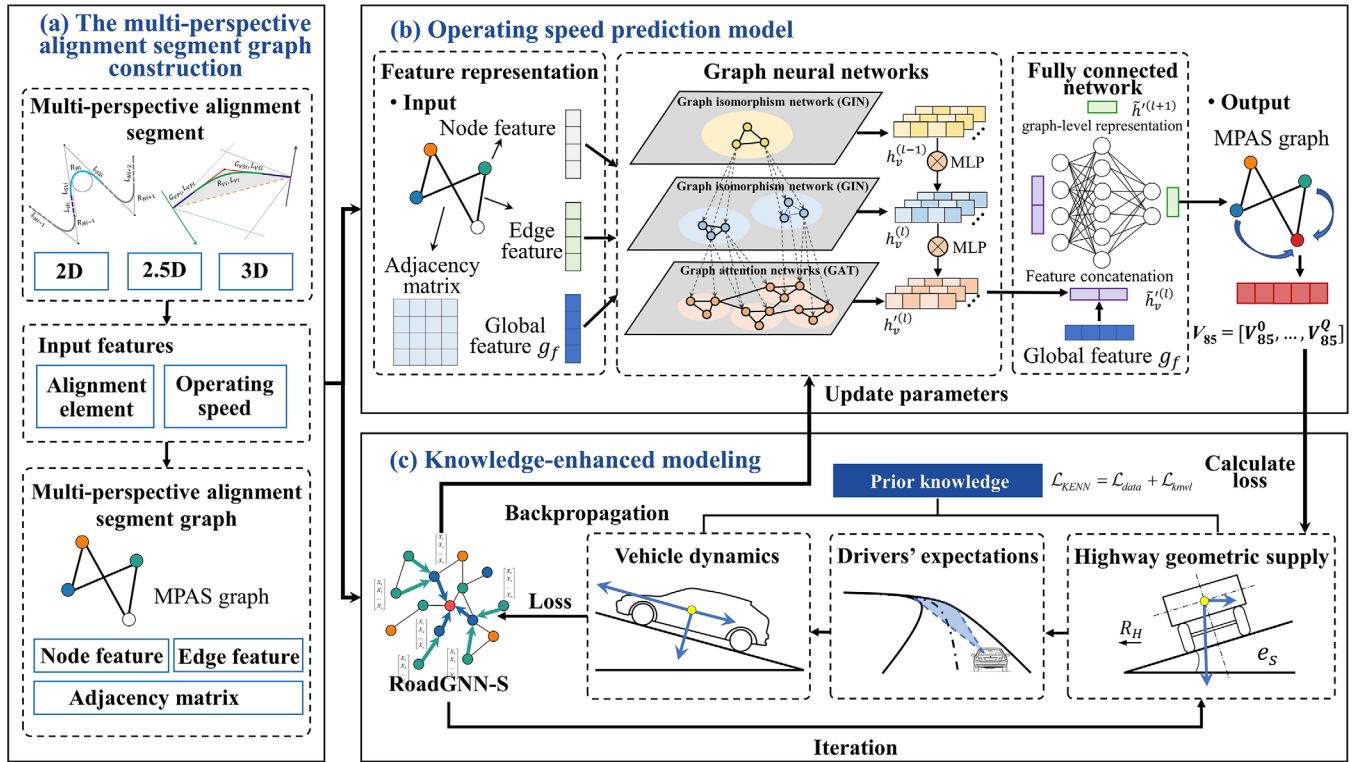


FIGURE 1 The methodology framework of this study. 2D, two-dimensional; 2.5D, 2.5-dimensional; 3D, three-dimensional; MLP, multiple-layer perceptron; MPAS, multi-perspective alignment segment; RoadGNN-S, graph neural networks-based operating speed prediction model.

represent highway alignment and fuse multi-perspective alignment elements (i.e., 2D, 2.5D, and 3D perspectives). Then, MPAS graphs are used to establish an operating speed prediction model based on GNNs. This kind of interpretable deep neural network could effectively capture the coupling impacts of multi-perspective alignment elements on operating speed. The prior knowledge is then integrated into the proposed operating speed prediction model to improve the model's performance and transferability, which considers the interactions of highway geometric supply, driver expectations, and vehicle dynamics.

2.3 | The MPAS graph

2.3.1 | Highway alignment segment modeling

To enhance the applicability of deep learning algorithms in operating speed prediction, this study modularizes the highway alignment. As shown in Figure 2a,b, the highway alignment is divided into a series of adjacent segments (called highway alignment segments), each of which contains a complete horizontal alignment element (i.e., a tangent or a combination of entry spiral, circular curve, and exit spiral) and vertical alignment elements (i.e., a

combination of preceding vertical slope, vertical curve, and succeeding vertical slope). This method not only examines the perfect overlap between horizontal curves and vertical curves but also covers various combinations of horizontal and vertical alignments. Regarding 2D design perspective alignment as the basis and 2.5D and 3D perspective alignment as the supplements, the MPAS model is then established, which can be described by the following formulas. The 2D perspective alignment is fixed within each segment, and the 2.5D and 3D perspective alignments change continuously, providing a more realistic depiction of the dynamic nature of highway alignment.

2D design perspective

Highway alignment design has always employed the traditional 2D design method that separates the horizontal and vertical planes. As illustrated in Figure 2c, horizontal alignment (including tangents, spirals, and circular curves) presents the layout of highways on the horizontal plane, while vertical alignment (including vertical slopes and vertical curves) reflects the vertical undulations of highways. To generically represent the arbitrary combination of horizontal and vertical alignment, this study establishes a highway alignment differential equation provided in Equation (1) based on the formula derivation from

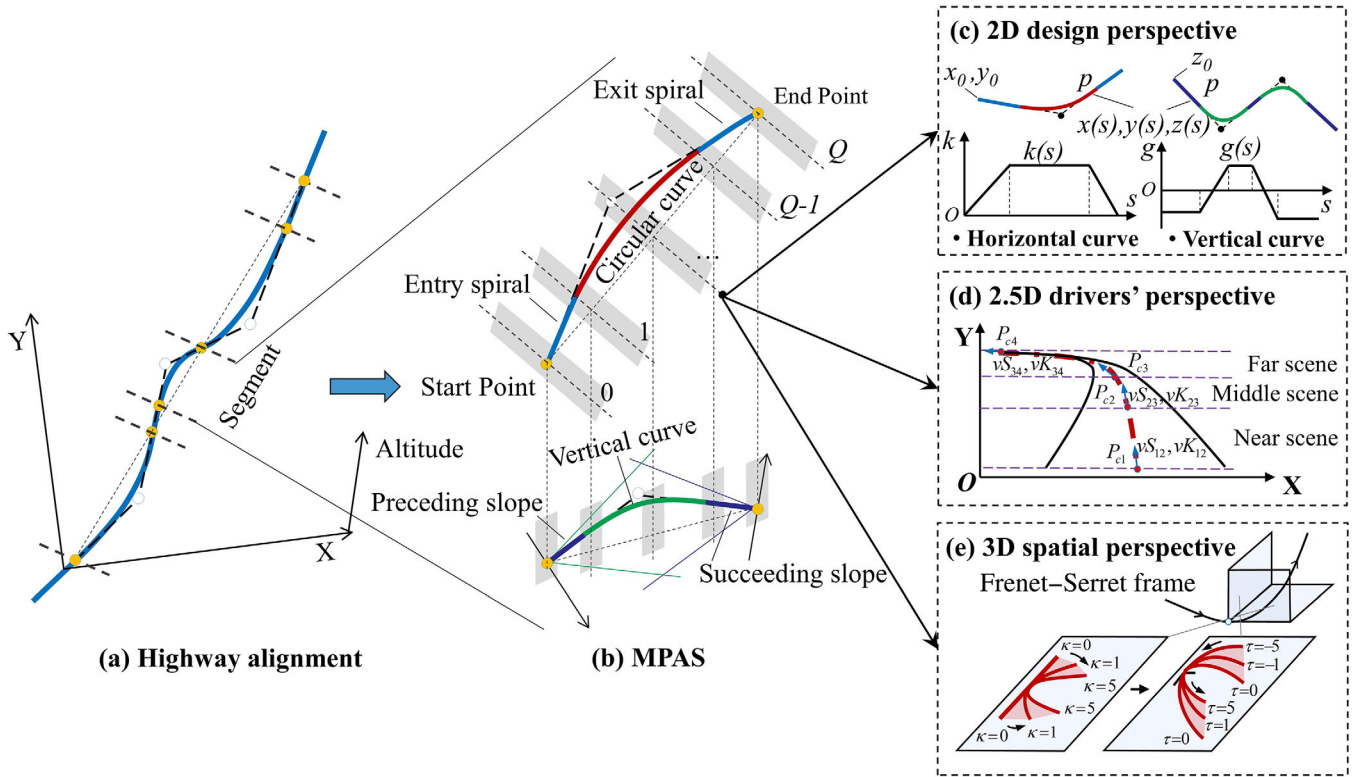


FIGURE 2 Schematic diagram of multi-perspective highway alignment segment segmentation process. 2D, two-dimensional; 2.5D, 2.5-dimensional; 3D, three-dimensional; MPAS, multi-perspective alignment segment.

previous research (Yan & Zhang, 2010).

$$\begin{cases} x(s) = x_0 + \int_0^s \cos \left(k_0 s + \frac{k_e - k_0}{2L} s^2 \right) ds \\ y(s) = y_0 + \int_0^s \sin \left(k_0 s + \frac{k_e - k_0}{2L} s^2 \right) ds \\ z(s) = z_0 + \int_0^s g(s) ds, g(s) = \frac{D_h}{R_V} + G_V \end{cases} \quad (1)$$

where s is the mileage of any point p along the MPAS; L is the length of the horizontal curve; x_0 and y_0 are the initial coordinates; z_0 is the initial elevation; k_0 and k_e are the curvature at the starting and ending points of the horizontal alignment, respectively; $g(s)$ is the grade at the point p , which can be calculated by vertical slope G_V and the radius of vertical curve R_V ; D_h is the horizontal distance along the vertical curve.

2.5D driving perspective

Compared to quadratic curves and other cubic curves, such as cubic Bézier splines and cubic B-splines, the Catmull-Rom spline provides higher accuracy in fitting the highway alignment perceived by drivers' vision (B. Yu et al., 2019). As shown in Figure 2d, the visual lane shape is divided into three distinct regions, including "near scene," "middle scene," and "far scene," based on the locations of four con-

trol points ($P_{c1}, P_{c2}, P_{c3}, P_{c4}$) of the Catmull-Rom spline. The length and curvature of the visual curves within these three distinct regions are used as shape parameters and denoted as $[vS_{i(i+1)}, vK_{i(i+1)}]$. These parameters can be calculated by Equations (2) and (3):

$$P_c(t) = [1 \ t \ t^2 \ t^3] \begin{bmatrix} 0 & 1 & 0 & 0 \\ -0.5 & 0 & 0.5 & 0 \\ 1 & -2.5 & 2 & -0.5 \\ -0.5 & 1.5 & -1.5 & 0.5 \end{bmatrix} \begin{bmatrix} P_{c1} \\ P_{c2} \\ P_{c3} \\ P_{c4} \end{bmatrix} \quad (2)$$

$$vS_{i(i+1)} = S_{i+1} - S_i, vK_{i(i+1)} = \frac{f_{i+1} - f_i}{vS_{i(i+1)}} \quad (3)$$

where $P_c(t)$ denotes the position of a point on the Catmull-Rom spline at the interpolation parameter t , $t \in [0, 1]$; $P_{c1}, P_{c2}, P_{c3}, P_{c4}$ denote the control points of a Catmull-Rom spline, with each point denoted as (S_i, x_i, y_i) ($i = 1, 2, 3, 4$); $vS_{i(i+1)}$ and $vK_{i(i+1)}$ denote the visual curve length and curvature between control points P_{ci} and $P_{c(i+1)}$, respectively; S_i and f_i denote the accumulated curve length and tangent direction angle at the control point P_{ci} , respectively; f_i is calculated by taking the derivative of the Catmull-Rom spline equation with respect to t and evaluating it at the control points;



The derivation process of the standard Catmull–Rom basis matrix is detailed in Y. Wang et al. (2000).

3D spatial perspective

The representation of spatial curves in classical geometry relies on differential equations. In the Frenet–Serret theory (Kim et al., 2012), the curvature and torsion of 3D spatial curves can be described by an orthonormal moving frame along the curve. As shown in Figure 2e, the spatial curvature describes the degree of curvature of the highway alignment, while the spatial torsion represents its rotational properties. According to the derivations by J. Wang et al. (2022), the calculated formulas of spatial curvature and torsion of highway alignment are described in Equation (4):

$$\kappa(s) = \frac{|\dot{r}(s) \times \ddot{r}(s)|}{|\dot{r}(s)|^3}, \tau(s) = \frac{(\dot{r}(s) \times \ddot{r}(s)) \cdot \dddot{r}(s)}{|\dot{r}(s) \times \ddot{r}(s)|^2}$$

$$r(s) = \begin{pmatrix} x(s) \\ y(s) \\ z(s) \end{pmatrix} \quad (4)$$

where $\kappa(s)$ and $\tau(s)$ denote the spatial curvature and torsion at length s of the 3D curve, respectively; s is the mileage of any point p along the MPAS; $r(s)$ denotes the position vector of the mileage s on the road segment.

2.3.2 | Input features

To construct an MPAS graph that can describe the relationship between multi-perspective alignment elements and operating speed, this study utilizes the input features from four sources, including 2D design perspective alignment element features (e.g., circular/tangent, spiral, vertical slope, vertical curve, etc.), 2.5D driving perspective alignment element features (e.g., visual curve length, visual curve curvature, etc.), 3D spatial perspective alignment element features (e.g., spatial curvature, spatial torsion, etc.), and driving behavior features (i.e., operating speed). These input features are described in detail in Table 1. Similar to building blocks, a series of MPASs is able to splice into a complete highway alignment and a continuous operating speed curve.

2.3.3 | MPAS graph construction

Since the proposed MPAS contains different types of features (i.e., different perspective alignment elements and driver behavior), Euclidean distances cannot be used to measure their similarity. Although the input data follow

a fixed structure, the relationships among the multi-perspective features in the MPAS are inherently non-Euclidean and complex. Thus, it is difficult to establish effective relationships between different features using Euclidean data structures (e.g., vectors, sequences, grids, etc.). To solve this issue, the MPAS graph is constructed to organize the above features in the form of a graph structure, which can effectively represent the complex interdependencies and irregular topologies inherent in these features as demonstrated in Figure 3. Graphs are a kind of non-Euclidean data structure consisting of nodes and edges that describe entities and their relationships in geometric space (Bazrafshan et al., 2024). By leveraging the scalability and topological structure of graphs, each MPAS is explicitly represented as a graph, facilitating a modular representation of highway alignment.

As shown in Figure 3, the MPAS graph comprises four types of nodes, including 2D design perspective alignment elements, 2.5D driving perspective alignment elements, 3D spatial perspective alignment elements, and operating speed. The relationships between these nodes are established through edges, which represent three types of relationships: (1) edges between alignment element nodes represent geometric constraints among multi-perspective features (e.g., perspective relations from 2D to 2.5D, compositions from 2D to 3D, and transitions from spirals to circular curves, etc.); (2) edges between operating speed nodes represent acceleration; and (3) edges between alignment element nodes and operating speed nodes represent the influence of alignment elements on operating speed. Additionally, the connectivity in the graph allows nodes to propagate multi-hop information through edges. The 2D, 2.5D, and 3D perspectives of highway alignment act directly on highway geometric design, driver visual perception, and vehicle movement, respectively. Therefore, the MPAS graph could realize the fusion of multi-perspective alignment elements, by establishing a closed-loop within highway geometric supply, driver expectations, vehicle dynamics, and operating speed.

Considering that the geometric constraints between alignment elements and the transmission of operating speed are directional, this study uses directed graphs to represent the MPAS. The MPAS graph can be described as $G = (V, E)$, where $V = \{v_1, \dots, v_m\}$ is the node set and $E = \{e_1, \dots, e_n\}$ is the edge set. Node attributes are denoted as $X^v \in \mathbb{R}^{m \times d}$, where m is the number of nodes, and d is the dimension of node features. Similarly, edge attributes are denoted as $X^e \in \mathbb{R}^{n \times c}$, where n is the number of edges, and c is the dimension of edge features. Table 1 lists the values and dimensions of node and edge attributes. In an MPAS graph, the topological relationship between entities (i.e., nodes) is presented as an adjacency matrix $A = [a_{ij}]_{m \times m}$ that captures the connections between nodes, where if


TABLE 1 Summary description of the multi-perspective alignment segment (MPAS) features.

Categories	Alignment element	Code	Attribute	Description
Two-dimensional design perspective	Circular curve/tangent	Ct	C_H, L_{RH}, e_s	C_H, L_{RH} , and e_s are the curvature, length, and superelevation of the circular curve (or tangent), respectively $C_H = 0$: tangent $C_H > 0$: circular curve
	Spiral	Sp	L_{S1}, L_{S2}, A_1, A_2	L_{S1} and L_{S2} are the length of the entry and exit spirals, respectively. A_1 and A_2 are the parameters of the entry and exit spirals in horizontal alignment, respectively
	Vertical slope	Vs	$G_{VP}, G_{VS}, L_{VP}, L_{VS}$	G_{VP} and L_{VP} are the grade and length of the preceding vertical slopes, respectively. G_{VS} and L_{VS} are the grade and length of the succeeding vertical slopes, respectively $G_V > 0$: upgrade $G_V < 0$: downgrade
	Vertical curve	Ve	R_V, L_{RV}	R_V and L_{RV} are the radius and length of vertical curves, respectively, in accordance with Chinese design specifications for highway alignment (MTPRC, 2017). $R_V > 0$: sag vertical curve $R_V < 0$: crest vertical curve In the Green Book (AASHTO, 2018), R_V represents the rate of vertical curvature
2.5-dimensional driving perspective	Visual curve	Vc	$[vS_{i(i+1)}, vK_{i(i+1)}]$	$vS_{i(i+1)}$ and $vK_{i(i+1)}$ are the length and curvature of the visual curve in the i -th to $(i+1)$ th region, $i = 1, 2, 3$
Three-dimensional spatial perspective	Spatial curve	Sc	κ, τ	κ and τ are the spatial curvature and spatial torsion of the 3D curve, respectively
Driver behavior	Operating speed	V_{85}	V_{85}	The operating speed along the driving direction of vehicles, km/h, $V_{85} = [V_{85}^0, \dots, V_{85}^Q]$. Q represents the number of spots where operating speed is calculated

node i is connected to node j , a_{ij} is 1, and if not, a_{ij} is 0. To balance the efficiency of edge lookups and memory usage, this study employs a sparse adjacency matrix to handle edge connections. While adjacency lists are more efficient for neighbor node exploration, the sparse adjacency matrix provides better performance for frequent edge connectivity queries (M. Xu et al., 2023).

2.4 | RoadGNN-S model architecture

To better consider the coupling impacts of multi-perspective alignment elements on operating speed, this study employs GNNs to establish an operating speed prediction model, named RoadGNN-S. GNNs are emerging deep learning techniques that are capable of processing non-Euclidean data by extending convolution operations (Zhang et al., 2024). The proposed RoadGNN-S utilizes a “graph-in, graph-out” architecture. The independent variables represent MPAS graphs with missing operational

speed node features, and the goal of the model is to learn how to infer the missing features from this incomplete information to generate a complete MPAS graph as the dependent variable. The message-passing mechanism allows nodes to exchange and propagate information from their neighbors, effectively capturing dependencies across multiple hops (Dong & Sun, 2024). Additionally, the multi-head self-attention mechanism assigns varying weights to each edge based on its importance, thereby improving the model’s ability to learn from complex node relationships (Aironi et al., 2024). By utilizing the message-passing and multi-head self-attention mechanisms, this prediction model might effectively leverage the information from neighboring nodes to enhance model generalization and performance.

As demonstrated in Figure 1b, RoadGNN-S consists of four components, including a feature representation layer, two graph isomorphism network (GIN, a variant of GNNs) layers, a graph attention network (GAT) layer, and a fully connected network (FC) layer. The feature

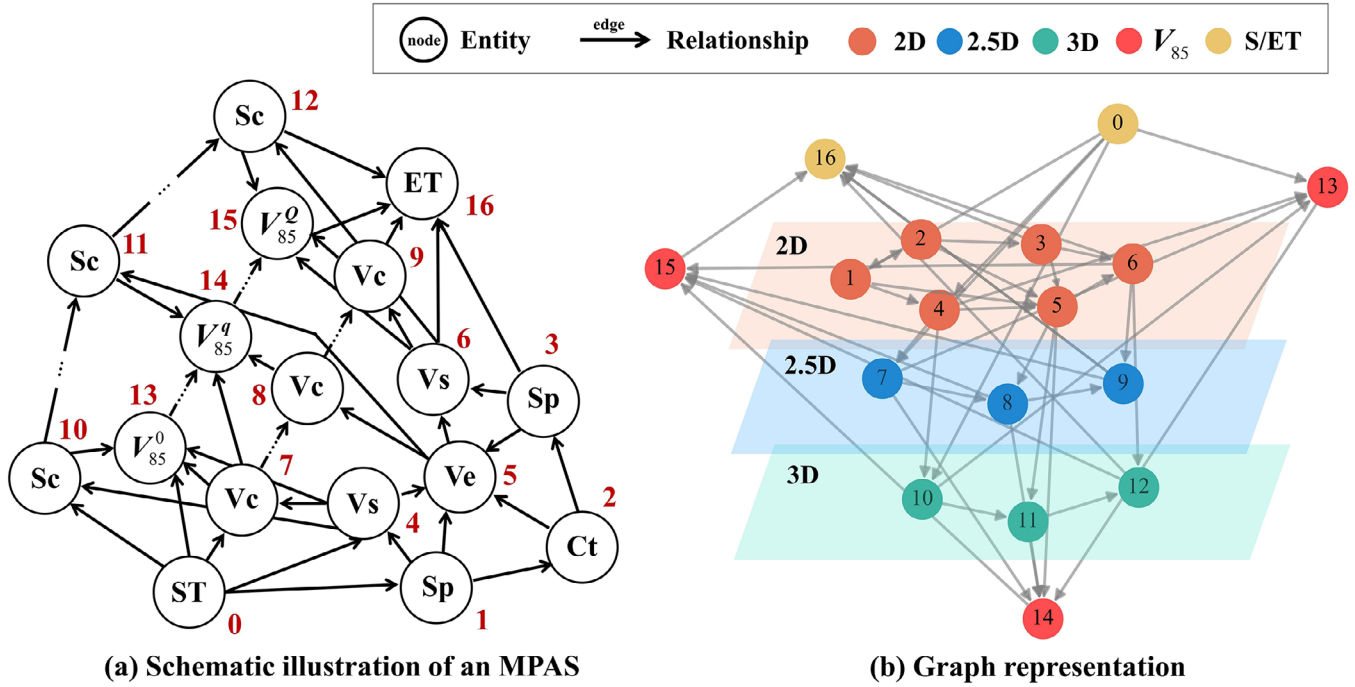


FIGURE 3 Graph representation of an multi-perspective alignment segment (MPAS). *Note:* the numbers in the figure represent node identity (i.e., node ID). ST and ET represent the start point and end point of the MPAS, and these two nodes are used to connect upstream and downstream MPASs. ST and ET nodes contain detailed geometric design information, including coordinates, orientation angles, mileage, and index of the MPAS. The descriptions of all the entity codes are demonstrated in Table 1. 2D, two-dimensional; 2.5D, 2.5-dimensional; 3D, three-dimensional.

representation layer can extract and organize all highway geometric design parameters (e.g., alignment elements, design speed, lane width, etc.) as the input graph. The GIN layer updates local node information by iteratively aggregating neighboring nodes, thereby capturing the individual impact of alignment elements on operating speed from different perspectives. The GAT layer focuses on the global graph structure and captures the intricate relationships among alignment elements from different perspectives, along with their coupling impacts on operational speed. The fully connected network layer maps the learned node feature representations onto graph-level representations that contain continuous operating speed values. The detailed process is described as follows:

Step 1 (feature representation): The highway alignment is first divided into a series of MPASs. Multi-perspective alignment elements are extracted and calculated to construct MPAS graphs $G = (V, E)$. Additionally, design speed V_d , lane width W_l , restricted type R_t (i.e., tunnel, interchange, and speed zone), and initial operating speed V_{85}^0 of the MPAS are incorporated as a global feature $g_f = [V_d, W_l, R_t, V_{85}^0]$ into the model. Each MPAS's initial speed V_{85}^0 is derived from the terminal operating speed of the preceding segment,

thereby ensuring continuity in operating speed transfer throughout the prediction process.

Step 2 (the GIN layer): To address the graph isomorphism problem among different nodes with similar local structures in the MPAS, the GIN model is employed to ensure node distinctiveness and recursively update the feature embeddings of each node to capture both local structural information and contextual relationships within the global graph (Bouritsas et al., 2022). In the GIN, a function $\mathcal{G} : \mathcal{X} \rightarrow \mathbb{R}^{\mathcal{N}}$ is utilized for encoding nodes in the countable multi-set. The node representation $\mathcal{H}(v, X)$ can be calculated by Equation (5):

$$\mathcal{H}(v, X) = (1 + \varepsilon) \cdot \mathcal{G}(v) + \sum_{u \in X} \mathcal{G}(u) \quad (5)$$

where V is a set of nodes of v , $v \in V$; $\mathcal{X}$ is bounded-size multi-set, $X \subset \mathcal{X}$; ε is a learnable parameter; the functions $\mathcal{G}(v)$ and $\mathcal{G}(u)$ map the features of node v and neighboring node u , respectively, through the encoding function $\mathcal{G}$ to a $\mathcal{N}$ -dimensional vector.

GNNs involve three stages in their message-passing process: message computation, aggregation, and update. All nodes in the graph first carry a “message” (i.e., original node features). Then, nodes



aggregate messages from neighboring nodes and use the aggregated information to update their representations as described by Equations (6)–(9):

$$h_v^{(0)} = x_v, \forall v \in V \quad (6)$$

$$m_{uv}^{(l)} = W_m^{(l)} h_u^{(l-1)}, \forall (u, v) \in E \quad (7)$$

$$a_v^{(l)} = \sum_{u \in N(v)} m_{uv}^{(l)}, \forall v \in V \quad (8)$$

$$h_v^{(l)} = \sigma \left(W_u^{(l)} a_v^{(l)} + b^{(l)} \right), \forall v \in V \quad (9)$$

where x_v represents the features (i.e., attributes) of node v ; $m_{uv}^{(l)}$ denotes the message from node u to node v in the l th layer; $a_v^{(l)}$ denotes the aggregated representation of node v in the l th layer; $h_v^{(l)}$ denotes the node representation at the l th layer; $N(v)$ denotes the set of neighboring nodes of node v ; $W_m^{(l)}$ and $W_u^{(l)}$ are learnable weight matrices; $\sigma(\cdot)$ is the rectified linear unit (ReLU) function; $b^{(l)}$ is a bias term for the l th layer.

GIN, a variant of GNNs, employs multiple-layer perceptron (MLP) as update functions, which perform nonlinear transformations on node features and combine the weighted sums of neighboring node features to better capture intricate graph relationships. The MLP used in GIN typically consists of two layers, with ReLU activation applied after each layer. This process can be described by Equation (10).

$$h_v^{(l)} = \text{MLP}^{(l)} \left((1 + \varepsilon^{(l)}) \cdot h_v^{(l-1)} + \sum_{u \in N(v)} h_u^{(l-1)} \right) \quad (10)$$

Step 3 (the graph attention networks layer): the graph attention network layer employs the multi-head self-attention mechanism to dynamically allocate self-attention weights. For the k -th head, the hidden representation $h_v^{(l,k)}$ of node v in layer l is updated by a weighted summary, incorporating both its own and neighboring nodes' information, as specified by Equations (11) and (12):

$$\alpha_{vu}^{(k)} = \text{softmax} \left(a \left(W^{(k)} h_v^{(l-1)}, W^{(k)} h_u^{(l-1)} \right) \right) \quad (11)$$

$$h_v^{(l,k)} = \sigma \left(\sum_{u \in N(v)} \alpha_{vu}^{(k)} W^{(k)} h_u^{(l-1)} \right) \quad (12)$$

where $W^{(k)}$ is the linear transformations matrix for the k th head; $\alpha_{vu}^{(k)}$ is the self-attention weight for the k th attention head between v node and u

node; $a(\cdot, \cdot)$ is the attention coefficient computed by the k th head; $\sigma(\cdot)$ is the ReLU function.

The final representation $h_v^{(l)}$ of node v with multi-head attention is obtained by concatenating the outputs from all attention heads, which can be calculated by Equation (13):

$$h_v^{(l)} = \parallel_{k=1}^K h_v^{(l,k)} \quad (13)$$

where K is the total number of heads; $\parallel$ denotes concatenation.

Step 4 (the fully connected network layer): Before passing the output of the graph attention network layer to the fully connected layer, the global feature g_f can be directly concatenated to the node representations. Then, the concatenated node representations are non-linearly transformed into a graph-level representation containing operating speed information $V_{85} = [V_{85}^0, \dots, V_{85}^Q]$ as given by Equations (14) and (15).

$$\tilde{h}_v^{(l)} = \text{concat} \left(h_v^{(l)}, g_f \right), \forall v \in V \quad (14)$$

$$\tilde{h}^{(l+1)} = \tilde{h}_v^{(l+1)} W_f^{(l+1)} + b^{(l)} \quad (15)$$

where $h_v^{(l)}$ denotes the output node representations from the graph attention networks layer; $\tilde{h}_v^{(l)}$ denotes the concatenated node representations; $\text{concat}(\cdot, \cdot)$ denotes the concatenation operation; $\tilde{h}^{(l+1)}$ denotes the output graph-level representation; $W_f^{(l)}$ and $b^{(l)}$ represent the weight matrix and bias in the fully connected layer, respectively.

2.5 | Knowledge-enhanced modeling

To enhance deep learning models' physical intuition to guarantee the theoretical correctness of the prediction results, this study introduces the prior knowledge involved in the whole process from drivers' perception of the highway alignment to their decision-making of driving behavior into RoadGNN-S. As shown in Figure 1c, the interactions $\mathcal{I} = (I_g, I_d, I_v)$ of highway geometry supply, driver expectations, and vehicle dynamics on operating speed are inserted into the backpropagation process of GNNs, that is, speed and acceleration penalty terms are introduced in the loss function to expedite and optimize the training process. The knowledge-enhanced modeling can better utilize the associated information between nodes and edges in the graph data, enhancing prediction accuracy and generalization performance, especially in situations with scarce training data.



The loss function of knowledge-enhanced neural networks consists of two components: the data-driven loss and the knowledge-enhanced loss. The knowledge-enhanced loss represents the squared residuals between the predicted operating speeds and the speeds that ensure safety and comfort, as well as between the predicted accelerations and the accelerations that satisfy driver expectations and vehicle dynamics, thereby enforcing physical constraints on the RoadGNN-S. Specifically, this loss can be calculated by Equations (16)–(18):

$$\mathcal{L}_{\text{KENN}} = \mathcal{L}_{\text{data}} + \mathcal{L}_{\text{knwl}} \quad (16)$$

$$\mathcal{L}_{\text{data}} = \sum_{q=1}^Q \text{MSE} \left(y_q^{\text{pred}}, y_q^{\text{true}} \right) \quad (17)$$

$$\begin{aligned} \mathcal{L}_{\text{knwl}} = & \lambda_1 \cdot \sum_{q=1}^Q \left(\frac{\left(y_q^{\text{pred}} - \min(V_{ds}, V_{dc}) \right)^2}{\sigma_y^2} \right) \\ & + \lambda_2 \cdot \sum_{q=1}^Q \left(\frac{\left(\frac{y_q^{\text{pred}^2} - y_{q-1}^{\text{pred}^2}}{2\Delta s} - \min(a_{de}, a_{vd}) \right)^2}{\sigma_a^2} \right) \end{aligned} \quad (18)$$

where $\mathcal{L}_{\text{KENN}}$ is the loss function of knowledge-enhanced modeling; $\mathcal{L}_{\text{data}}$ is the loss function of data-driven modeling; $\mathcal{L}_{\text{knwl}}$ is the loss function of introducing prior knowledge; λ_1 and λ_2 are the weights of knowledge loss term, $\lambda_1, \lambda_2 \in [0, 1]$; $\text{MSE}(\cdot)$ is the mean square error loss function; Q is the number of operating speed nodes; y_q^{pred} and y_q^{true} are the predicted value and true value of the operating speed on node q , respectively; Δs is the mileage from the calculated node q to $q - 1$; a_{de} is the acceleration produced by drivers' expectations; a_{vd} is the acceleration produced by vehicle dynamics; V_{ds} is the maximum operating speed that ensures driving safety; V_{dc} is the maximum operating speed that ensures driving comfort; σ_y^2 and σ_a^2 denotes the variance of the predicted speed term and the acceleration term, respectively.

2.5.1 | Highway geometric supply

In the 2D design perspective, traditional highway geometric design theory uses a point mass model to correlate speed and curve radius. This method ensures that vehicles do not experience side skidding and rollover within a certain safety margin of the design speed as given by

Equation (19).

$$\begin{cases} V_{ds} \leq \sqrt{\frac{R_H}{127(f_{\max} + 0.01e_s)}} \quad (\text{non-skidding}) \\ V_{ds} \leq \sqrt{\frac{R_H}{127\left(\frac{b}{2h_g} + 0.01e_s\right)}} \quad (\text{non-rollover}) \end{cases} \quad (19)$$

where V_{ds} is the maximum operating speed that ensures driving safety; R_H is the radius of horizontal curves; $f_{\max}$ is the maximum allowable side friction factor, $f_{\max} \in [0.15, 0.4]$ (AASHTO, 2018); e_s is the rate of superelevation, percent; b is the length of a vehicle's wheelbase; h_g is the height of a vehicle's gravity center.

2.5.2 | Drivers' expectations

In the 2.5D driving perspective, drivers undergo a continuous perception-decision-action cycle process (B. Yu et al., 2019). Drivers' expectations can be categorized into two types: inertial expectations and ad hoc expectations (Vos et al., 2023). Inertial expectations suggest that drivers' speed choice is influenced not only by their cumulative driving experience and memory but also by their ability to maintain a similar expected speed (i.e., the highest safe speed they aim to achieve psychologically) to that of the most recent segment they have driven on. Additionally, ad hoc expectations are formed in real-time as drivers perceive the characteristics and conditions of the road segment they have just navigated. When there is a large discrepancy between the expected speed and the current driving speed, drivers would dynamically change their speed according to ad hoc expectations. Specifically, on visually continuous segments (e.g., tangents, large radius curves, etc.), drivers have higher expectations for operating speed. On visually interrupted segments (e.g., small radius curves and broken back curves, etc.), drivers tend to slow down due to increased attention. Then, once visibility conditions improve, drivers will accelerate to their expected speed. Considering the drivers' visual continuity and their expectations, the above drivers' speed adjustment actions can be described by Equation (20):

$$\begin{cases} a_{de} > 0 \text{ if } \frac{\sum vS(q-1)}{\sum vS(q)} \leq 0.5, V_{85} < V_e \\ a_{de} = 0 \text{ if } 0.5 \leq \frac{\sum vS(q-1)}{\sum vS(q)} < 2, V_{85} = V_e \\ a_{de} < 0 \text{ if } \frac{\sum vS(q-1)}{\sum vS(q)} \geq 2, V_{85} > V_e \end{cases} \quad (20)$$



where a_{de} is the acceleration generated by ad hoc drivers' expectations, $a_{de}^{car} = \pm 0.5 \text{ m/s}^2$, $a_{de}^{truck} = \pm 0.25 \text{ m/s}^2$ (Xuesong Wang, et al., 2019a); $\sum vS(q)$ is the cumulative length of drivers' visual curve at the spot q ; V_e is the expected speed generated by drivers' inertial expectations, with detailed values referenced in the specifications (MTPRC, 2015).

2.5.3 | Vehicle dynamics

In the 3D spatial perspective, vehicles are affected by real physical forces, mainly including normal centrifugal force and tangential traction force along the forward direction. Normal acceleration, determined by both speed and spatial curvature, impacts the driving comfort experienced on curved road segments. The tangential acceleration is primarily determined by vehicles' traction force and driving resistances (i.e., air resistance, rolling resistance, skidding resistance, and grade resistance), reflecting vehicles' acceleration capability. The acceleration generated by vehicle dynamics can be calculated by Equation (21):

$$\left\{ \begin{array}{l} a_{vd} = \sqrt{a_{vn}^2 + a_{vt}^2} \\ a_{vn} = V_{dc}^2 \cdot \kappa, V_{dc} = \sqrt{\frac{a_{vn}}{\kappa}} \\ a_{vt} = \frac{P_e \eta}{\mu V_{85}^0} - \left(\frac{\rho A_v C_a V_{85}^0{}^2}{2\mu^2} + (C_{rr} + \mu C_s) g_r + g_r G_V \right) \end{array} \right. \quad (21)$$

where a_{vd} is the acceleration generated by vehicle dynamics; a_{vn} is the normal acceleration, and the maximum value of a_{vn} ensuring driving comfort is determined to be 1.0 m/s^2 (Jurecki & Stańczyk, 2021); a_{vt} is the tangential acceleration generated by engine traction; V_{dc} is the maximum operating speed that ensures driving comfort; κ is the spatial curvature; V_{85}^0 is the initial operating speed of the MPAS; P_e is the output power of the engine; η is the transmission efficiency, $\eta = 0.90$; μ is the mass of the vehicle; g_r is the gravitational constant (9.81 m/s^2); ρ is the air density (1.207 kg/m^3 at 20°C); A_v is the cross-sectional area of the vehicle, with $A_v^{car} = 2.0 \text{ m}^2$ for cars and $A_v^{truck} = 3.2 \text{ m}^2$ for trucks; C_a is the coefficient of air resistance; C_{rr} is the coefficient of rolling resistance; C_s is the coefficient of skidding resistance; G_V is the vertical grade.

3 | EXPERIMENTS

3.1 | Experimental design and data collection

To investigate various alignment combinations and obtain reliable data, the experiments in this study are divided

TABLE 2 The parameters for vehicle modules.

Parameters definition	Passenger car	Truck	Unit
Total length	6 long	12 long	m
Total width	1.8	2.5	m
Total height	2.0	3.95	m
Total mass	2000	31000	kg
Wheelbase	3800	6500	mm
Max power of engine	90	380	kW
Height of drivers' eye	1.2	2.0	m

into driving simulation and field observation. The data from driving simulation experiments are used to train the proposed RoadGNN-S, while the data from field observations are used to validate the prediction accuracy and transferability of the model.

3.1.1 | Driving simulation experiments

Figure 4a illustrates the Tongji University driving simulator employed in this study. Previous studies have provided detailed explanations of the reliability and effectiveness of this simulator, with t -test results showing that the speeds recorded from the simulator and the comparison roads belong to the same population, with no significant difference between their mean speeds (Chen et al., 2019; Xuesong Wang et al., 2019b). The simulator integrates a fully equipped Renault Mégane III vehicle cab within a dome structure, mounted on an eight-degree-of-freedom motion system. A five-channel projector system offers an immersive front view spanning $250^\circ \times 40^\circ$, at a refresh rate of 60 Hz with a resolution of 1000×1050 . SCANer studio software is used to control the force feedback system and display the road environment, capable of real-time recording of vehicle speed, position, and road mileage (Chen et al., 2022). Considering the significant differences in dynamic performance and driving behavior between passenger cars and trucks, operating speed data for both passenger cars and trucks need to be collected to establish two independent prediction models (Qu & Wang, 2015). As illustrated in Figure 4b,c, the passenger car and truck modules used in the driving simulator experiments are developed by Oktal Co. Ltd., and their parameters are modified according to the Chinese Specifications for Highway Safety Audit (MTPRC, 2015). The parameters of the two types of vehicle modules are listed in Table 2.

In the experiments, two sections of four-lane (two lanes in each direction) highways are constructed with a design speed of 100 km/h and each lane width of 3.75 m, according to the real highway conditions. One is 31.193 km from the DeShang Highway, and the other is 119.150 km from the TianTian Highway. These two highways are closely



FIGURE 4 Driving simulator and vehicle modules.

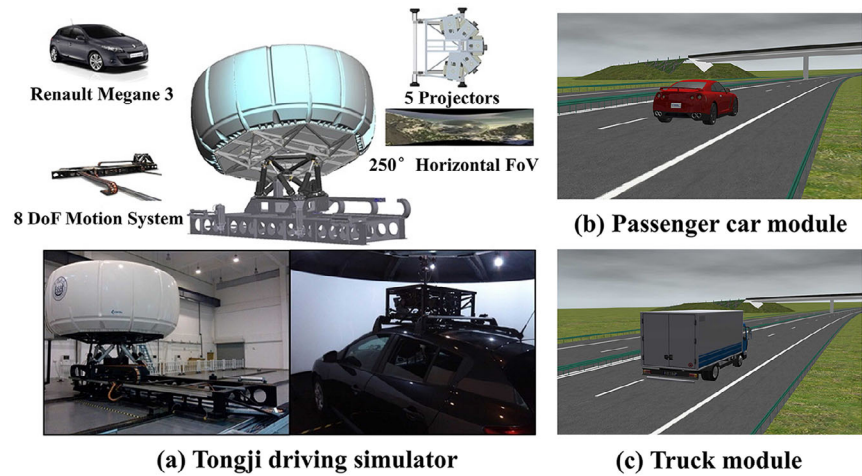


TABLE 3 Descriptive statistics of geometric parameters.

Description	Mean	SD	Min	Max
Curvature of circular curve (m^{-1})	0.0012	0.0005	0.0003	0.0040
Length of circular curve (m)	626.712	276.147	216.446	1948.970
Length of tangent (m)	792.884	1003.697	94.812	6010.057
Length of spiral (m)	136.819	31.760	70.000	185.000
Parameters of spiral (m)	305.455	109.908	167.332	832.050
Grade of vertical slope (%)	0.320	2.842	-5.800	5.800
Length of vertical slope (m)	455.911	501.954	94.812	3190.835
Radius of vertical curve (m)	23050.15	16851.13	6000.00	80000.00
Length of vertical curve (m)	325.772	149.061	126.265	1040.000
Superelevation (%)	2.850	1.250	2.000	6.000

located in Anhui province, China, both traversing mountainous terrains. Table 3 shows the descriptive statistics of the geometric parameters for the sections of highways. The surrounding terrain, horizontal and vertical alignment, superelevation parameters, traffic signs and markings, as well as safety barriers, are accurately replicated from these highway's design documentation, to ensure high consistency with actual driving experience. An overhead perspective of the highways and details of four typical road segments are illustrated in Figure 5.

The basic requirements for operating speed observation conditions include daytime, clear weather, dry road surface, and free-flowing traffic. Considering the distribution of driver gender, age, and vehicle type in China, we employ a stratified sampling method to recruit 58 participants, including 20 male and 18 female passenger car drivers and 20 male truck drivers. Their ages range from 20 to 60 years

TABLE 4 Typical segment division criteria.

Segments	Tangent	Vertical slope	Horizontal curve	Curved slope
Horizontal alignment	$C_H < 0.001$	$C_H < 0.001$	$C_H \geq 0.001$	$C_H \geq 0.001$
Vertical alignment	$G_V < 3\%$	$G_V \geq 3\%$	$G_V < 3\%$	$G_V \geq 3\%$

Note: C_H is the curvature of horizontal curves, and G_V is the grade of vertical slopes.

old (mean = 32.35, SD = 8.76). Overall, participants drive at least 50 km per month and have at least 1 year of driving experience. Research on participant sample size reports 30 to be the minimum size that can accurately collect continuous driving behavior data (F. Wang et al., 2019).

The experimental sessions consist of two stages, preparation and testing. In the preparation stage, participants are informed about the basic operation and the importance of the experiment, and they receive at least a 10-min practical driving session to ensure complete familiarity (Chen et al., 2018). In the testing stage, participants drive through the above two sections of four-lane highways. Due to the potential discomfort caused by prolonged simulation, the experiments are conducted in four separate sessions, with each session not exceeding a half-hour duration. Upon completion of the experiment, each participant is paid 150 RMB for their time. The speed and other motion characteristics of each driver are sampled at a frequency of 20 Hz during the experiment. The 85th percentile speed at each spot (i.e., selected locations in each road segment) along the highway is calculated as the observed value of operating speed.

To improve the applicability of the proposed MPAS, this study divides the MPAS into four types based on the Chinese Specifications for Highway Safety Audit (MTPRC, 2015), including tangent, vertical slope, horizontal curve, and curved slope segments (see Table 4). Using this divided criterion, a total of 532 road segment samples (with 30,856 speed profiles of 58 participants) are finally

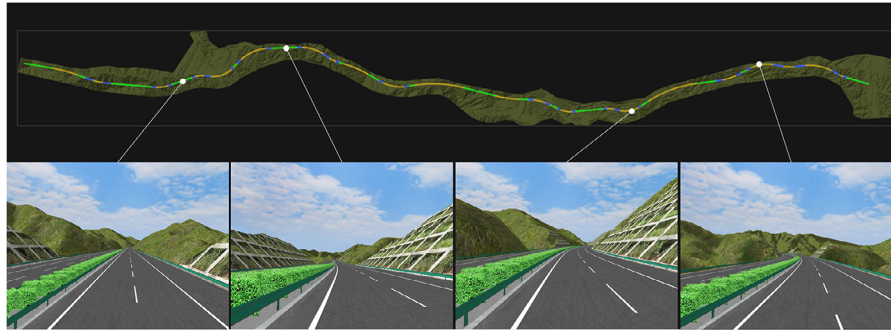


FIGURE 5 Overhead perspectives of the simulated highways.

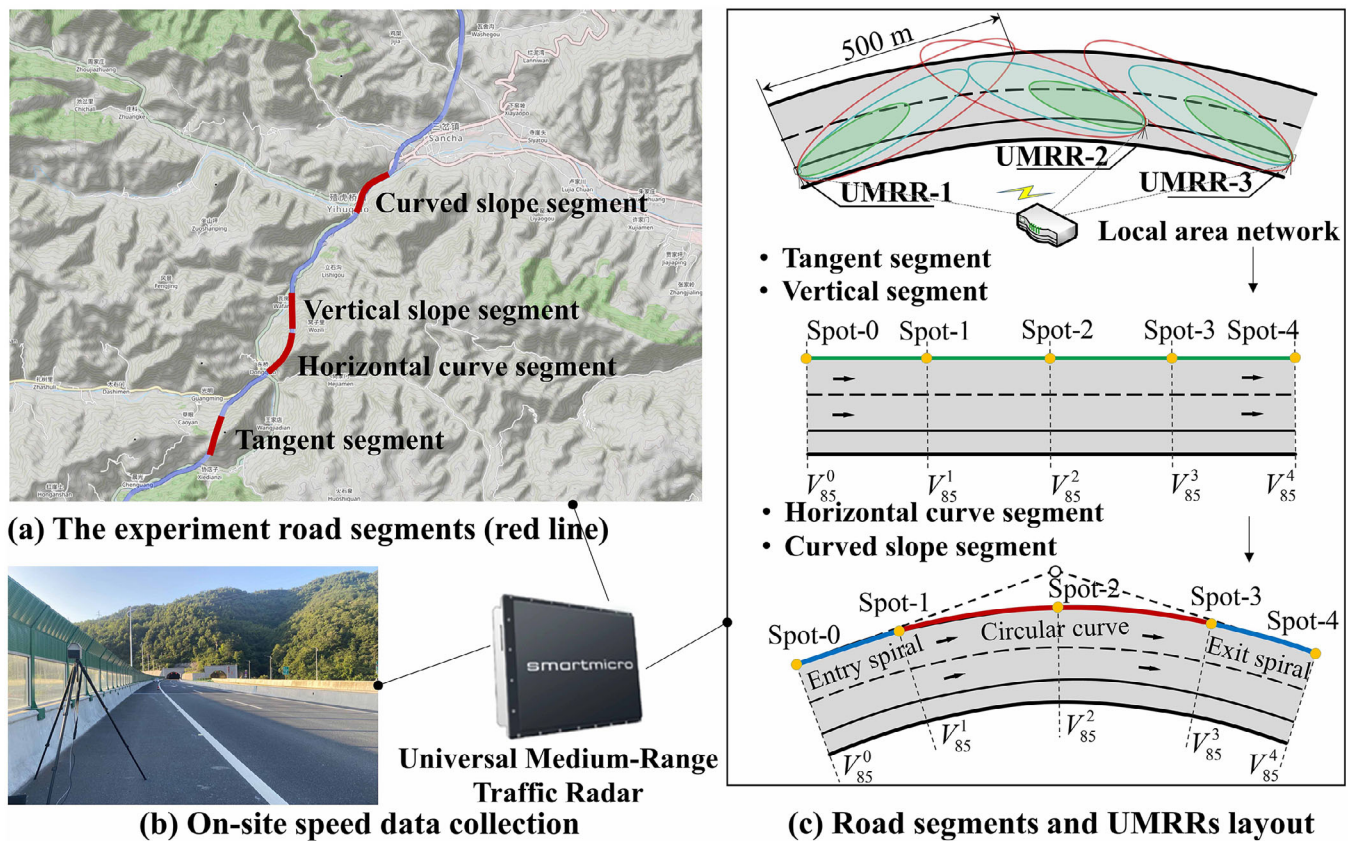


FIGURE 6 Schematic diagram of the experimental road segments and universal medium-range traffic radar (UMRR) layout.

obtained, consisting of 176 tangent segments, 116 vertical slope segments, 128 horizontal curve segments, and 112 curved slope segments.

3.1.2 | Field observation experiments

In this study, the prediction accuracy and transferability of RoadGNN-S are validated by using operating speed data collected from the Lanzhou-Haikou highway K164-K178 in Gansu, China, in 2023. The K164-K178 section is also a four-lane, two-way mountainous highway with a design

speed of 100 km/h. As shown in Figure 6, based on the original design documents of the highway, four typical road segments are selected for continuous speed observations. The universal medium-range traffic radar (UMRR) is applied to detect the speed and type of vehicles in this study (J. Gao et al., 2024). Under line-of-sight conditions, three UMRR units are grouped together in the same local area network and placed on the hard shoulder. The UMRR collects data at a frequency of 20 Hz. Data collection is conducted during periods of daytime, clear weather, good road surface conditions, and free-flow traffic. Free-flow conditions are defined as a headway of not less than 10 s between



leading and following vehicles in the same lane (Silvano et al., 2020). Each selected road segment is continuously observed for 1 h, and over 1200 effective single-vehicle speeds are finally obtained on each segment, of which trucks accounted for about 43.6%. To ensure the accuracy of the field speed observation data, the 3σ principle is applied to identify and eliminate abnormal speed values caused by unusual driving events. Afterward, the 85th percentile speed at each spot is determined to be the actual observed value of operating speed.

3.2 | Training configurations

The proposed RoadGNN-S can perform spot-based and continuous speed prediction by modifying the number of operating speed nodes in each MPAS graph. The more operating speed nodes Q are included, the smoother the operating speed curve becomes. The proposed RoadGNN-S model can predict operating speeds at discrete Q nodes, which correspond to locations where changes occur in the geometric features of the road segment. Using the linear interpolation method, the operating speed at any location between two adjacent Q nodes can be calculated based on the predicted operating speeds and mileages of those nodes and then connected into an operating speed profile (Xuesong Wang et al., 2021). The number of operating speed nodes Q (i.e., spots) in this study is set to 5 (i.e., $V_{85} = [V_{85}^0, V_{85}^1, V_{85}^2, V_{85}^3, V_{85}^4]$) since it can fit various operating speed profiles in all four typical road segments (Xuesong Wang et al., 2020). As shown in Figure 6c, these five spots represent distinct locations along road segments. Specifically, in horizontal curve and curved slope segments, V_{85}^0 and V_{85}^4 denote operating speeds at the start and end points of the road segment, while V_{85}^1 , V_{85}^2 , and V_{85}^3 denote operating speeds at the start, middle, and end points of the circular curve, respectively. In tangent and vertical slope segments, V_{85}^0 , V_{85}^1 , V_{85}^2 , V_{85}^3 , and V_{85}^4 are operating speeds at the four equally divided points along the road segment.

Based on driving simulation experiments, multi-perspective alignment elements and operating speeds are extracted to construct a dataset comprising 532 MPAS graphs. Each MPAS graph consists of 23 nodes and 75 edges. The hyperparameter tuning is conducted using a grid search method, primarily adjusting the learning rate, batch size, hidden layer dimensions, and weights of knowledge loss term. In this study, λ_1 and λ_2 are set to 0.6 and 0.4, respectively. The best hyperparameter combinations are shown in Table 5. Additionally, the proposed RoadGNN-S is implemented by using the PyTorch deep learning framework. The adaptive moment estimation algorithm is employed as the optimizer, with parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 10^{-8}$, where

TABLE 5 Best hyperparameter combination of graph neural networks-based operating speed prediction model (RoadGNN-S).

Hyperparameter	Value
Hidden dimensions of GIN	64
Number of attention heads in GAT	4
Hidden dimensions of GAT	128
Hidden dimensions of FC	512
Learning rate	0.001
Batch size	8
Number of epochs	200

Abbreviations: FC, fully connected network; GAT, graph attention network; GIN, graph isomorphism network.

β_1 and β_2 represent exponential decay rate for the first and second moments of the gradients, and ε is a small value added to the denominator for numerical stability (Pan et al., 2022). It should be noted that RoadGNN-S has low computational requirements and good scalability. The graphics processing unit (GPU) memory used during training is less than 2GB, allowing the model to operate on any consumer-grade central processing unit (CPU).

3.3 | Model validation

Four commonly used metrics are used to estimate the accuracy of the operating speed prediction model, including root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). RMSE, MAE, and MAPE are utilized to measure the deviation between the true value and the predicted value. The lower the RMSE, MAE, and MAPE values, the better the prediction model performance. R^2 represents the ability of the model to explain the data. The closer the R^2 value is to 1, the more accurate the model is. Although R^2 is typically used for linear models, RMSE, MAE, and MAPE are included alongside R^2 to ensure a comprehensive and robust evaluation. These metrics can be calculated using Equations (22)–(25):

$$\text{RMSE} = \sqrt{\frac{\sum_{z=1}^Z (y_z^{\text{pred}} - y_z^{\text{true}})^2}{Z}} \quad (22)$$

$$\text{MAE} = \frac{1}{Z} \sum_{z=1}^Z |y_z^{\text{pred}} - y_z^{\text{true}}| \quad (23)$$

$$\text{MAPE} = \frac{1}{Z} \sum_{z=1}^Z \frac{|y_z^{\text{pred}} - y_z^{\text{true}}|}{y_z^{\text{true}}} \times 100\% \quad (24)$$

$$R^2 = 1 - \frac{\sum_{z=1}^Z (y_z^{\text{true}} - \bar{y}_z)^2}{\sum_{z=1}^Z (y_z^{\text{true}} - \bar{y}_z)^2} \quad (25)$$

**TABLE 6** Comparison between the evaluation indicators of prediction performance.

	Passenger car				Truck			
	Road GNN-S	Feedforward neural networks (FFN)	Recurrent neural networks (RNN)	Convolutional neural networks (CNN)	Road GNN-S	FFN	RNN	CNN
Root mean squared error (RMSE; km/h)	2.7	7.0	5.9	5.4	2.2	5.5	5.9	4.2
Mean absolute error (MAE; km/h)	2.5	4.7	4.6	4.3	2.1	3.4	3.1	2.6
Mean absolute percentage error (MAPE; %)	2.3	4.2	3.7	3.7	3.4	5.0	4.4	3.8
Coefficient of determination (R^2)	0.94	0.81	0.82	0.85	0.96	0.83	0.85	0.87

where y_z^{pred} is predicted operating speed; y_z^{true} is observed operating speed; $\bar{y}_z$ is the mean value of observed operating speed; Z is the number of samples.

4 | RESULTS

4.1 | RoadGNN-S model performance

Using the GNNs architecture as the foundation framework, RoadGNN-S is established to predict operating speed for both passenger cars and trucks, which realizes the fusion of multi-perspective alignment elements by the MPAS graph. For comparison, this study also employs classical deep learning models, including feedforward neural networks (FFN), recurrent neural networks (RNN), and convolutional neural networks (CNN), to build operating speed prediction models. In this study, a total of 532 segment samples collected from the driving simulation experiment are split into a training dataset (80%) and a testing dataset (20%). A five-fold cross-validation is adopted to fine-tune parameters of the models. Different from RoadGNN-S that uses MPAS graphs as input, classical deep learning models can only treat grids or vectors as input.

The operating speed prediction results in the testing dataset are shown in Table 6. Compared with classical deep learning models, RoadGNN-S performs much better in predicting operating speed for both passenger cars and trucks. Specifically, as for passenger cars' operating speed prediction, the R^2 value of RoadGNN-S is 0.94, which is much higher than that of FFN (0.81), RNN (0.82), and CNN (0.85). RoadGNN-S has the lowest MAPE value (2.3%), whereas the values of FFN, RNN, and CNN are 4.2%, 3.7%, and 3.7%, respectively. The RMSE values of FFN, RNN, and CNN all exceed 5.4 km/h, and their MAE values are all greater than 4.3 km/h, while the RMSE and MAE of RoadGNN-S only reach 2.7 and 2.5 km/h, respectively.

Likewise, in terms of trucks' operating speed prediction, RoadGNN-S also have a better performance with R^2 reaching as high as 0.96, while FFN (0.87), RNN (0.85), and CNN (0.83) have relatively lower R^2 values. Meanwhile, the RMSE, MAE, and MAPE values of RoadGNN-S are far less than those values of FFN, RNN, and CNN. The reason why RoadGNN-S is superior to classical deep learning models might be due to the unique message-passing mechanism of GNNs. This kind of mechanism enables efficient information to be exchanged among operating speed nodes and their neighbor nodes (i.e., alignment elements), thereby effectively capturing the impact of multi-perspective alignment elements on operating speed.

Figure 7 demonstrates an example of the observed and predicted operating speed profiles. No matter in what typical type of road segment (i.e., vertical slope, curved slope, tangent, and horizontal curve segments), the operating speed predicted by RoadGNN-S is more consistent with the observed one, when compared with those predicted by FFN, RNN, and CNN. The predicted operating speed profile of RoadGNN-S can reflect the trend of observed operating speed variations more accurately than other classical deep learning models.

The operating speed prediction results for different perspective alignments are illustrated in Table 7. These results demonstrate the substantial advantages of multi-perspective alignment fusion (2D+2.5D+3D) in improving the accuracy of operating speed prediction when compared with using single-perspective (2D) alignment data and using two-perspective (2D+2.5D) alignment data. For passenger cars' operating speed prediction, the value of R^2 rises from 0.76 with the single 2D perspective to 0.88 when the 2.5D perspective is included (2D+2.5D), and finally reaches 0.94 with the addition of the 3D perspective (2D+2.5D+3D). Relative to inputting 2D data, using multi-perspective data leads to remarkable reductions in RMSE (from 9.1 to 2.7 km/h), MAE (from 7.9 to 2.5 km/h),

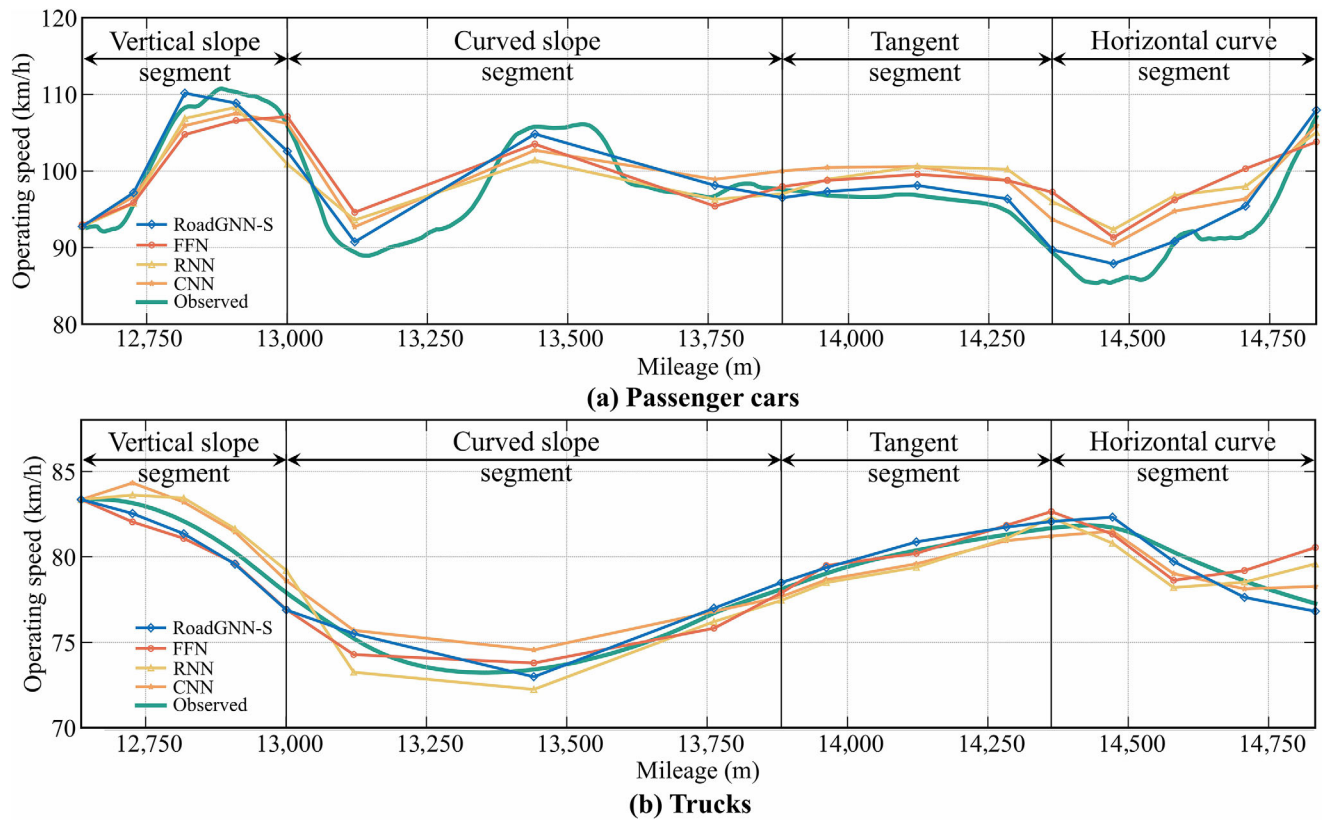


FIGURE 7 Observed and predicted operating speed profiles. CNN, convolutional neural network; FFN, feedforward neural network; RoadGNN-S, graph neural networks-based operating speed prediction model; RNN, recurrent neural network.

TABLE 7 Comparison of prediction performance between the multi-perspective alignments.

	Passenger car				Truck			
	2D	2.5D	2D+2.5D	2D+2.5D+3D	2D	2.5D	2D+2.5D	2D+2.5D+3D
RMSE (km/h)	9.1	4.6	3.4	2.7	5.1	6.0	4.2	2.2
MAE (km/h)	7.9	3.9	3.3	2.5	3.2	5.4	3.9	2.1
MAPE (%)	7.2	3.5	2.9	2.3	4.2	6.6	4.7	3.4
R^2	0.76	0.84	0.88	0.94	0.82	0.73	0.85	0.96

and MAPE (from 7.2% to 2.3%). For trucks' operating speed prediction, the same trend has been observed. The R^2 value similarly rises from 0.82 (2D) to 0.85 (2D+2.5D) and then to 0.96 (2D+2.5D+3D). Utilizing multi-perspective data also markedly decrease the values of RMSE (from 5.1 to 2.2 km/h), MAE (from 3.2 to 2.1 km/h), and MAPE (from 4.2% to 3.4%), when compared with the usage of single 2D perspective data.

4.2 | Comparison of model-driven, data-driven, and knowledge-enhanced modeling performance

To examine the superiority of knowledge-enhanced modeling in improving prediction accuracy and transferability,

this study utilizes all the samples obtained from the driving simulation experiment as the training dataset, and the field observation samples are treated as the testing dataset. Then, a knowledge-enhanced RoadGNN-S and a data-driven RoadGNN-S are established. In each model, the same dimensional graph representation and GNN architecture are used. As for the knowledge-enhanced modeling, this study integrates the prior knowledge that generates the operating speed (i.e., the interaction effects among highway geometry supply, driver expectations, and vehicle dynamics on operating speed) into the backpropagation process of RoadGNN-S. The remaining parameter configurations for both models are consistent to ensure comparability. Additionally, the linear regression model recommended in the Specifications for Highway Safety Audit (MTPRC, 2015) is selected as the model-driven

**TABLE 8** Comparison of prediction performance among different types of modeling methods.

		Model-driven modeling				Data-driven modeling				Knowledge-enhanced modeling			
		RMSE (km/h)	MAE (km/h)	MAPE (%)	R^2	RMSE (km/h)	MAE (km/h)	MAPE (%)	R^2	RMSE (km/h)	MAE (km/h)	MAPE (%)	R^2
Passenger car	Training dataset	9.9	8.8	8.7	0.69	2.2	2.1	1.9	0.95	2.2	2.0	1.9	0.96
	Testing dataset	9.6	8.6	8.5	0.71	7.6	6.9	7.6	0.85	3.3	3.1	3.3	0.92
Truck	Training dataset	6.6	4.8	5.9	0.63	2.1	2.1	3.4	0.96	1.9	1.7	2.2	0.97
	Testing dataset	6.4	4.5	5.4	0.67	5.2	5.0	5.9	0.83	3.0	2.7	3.6	0.92

method for comparison with data-driven and knowledge-enhanced models on the testing dataset, with the detailed calculation formulas specified in the specifications.

The results are demonstrated in Table 8. On the training dataset (from simulation experiments), the R^2 values of the data-driven model achieve 0.95 and 0.96 for predicting the operating speed of passenger cars and trucks, respectively. However, on the testing dataset (from field observations), the data-driven RoadGNN-S exhibits poor generalization with R^2 values only reaching 0.85 and 0.83 passenger cars and trucks, respectively, which markedly decreases, compared with those on the training dataset. In contrast, the knowledge-enhanced RoadGNN-S still shows high accuracies in the testing dataset, with R^2 values of 0.92 for both passenger cars and trucks in operating speed prediction. Thus, the knowledge-enhanced model outperforms the data-driven model, with improving the R^2 values of 7.9% for passenger cars and 10.7% for trucks after transferring to the new dataset. Additionally, in terms of those of the data-driven model, the RMSE and MAE values of the knowledge-enhanced model for predicting passenger cars' operating speed decrease by 4.3 and 3.8 km/h, respectively, while the RMSE and MAE values of the knowledge-enhanced model for trucks' operating speed reduce by 2.2 and 2.3 km/h, respectively. Furthermore, the linear regression model shows lower predictive accuracy on the testing dataset. For passenger cars, the R^2 value is 0.71, with RMSE and MAE of 9.6 and 8.6 km/h, respectively. For trucks, the R^2 value is 0.67, with RMSE and MAE of 6.4 and 4.5 km/h, respectively. These findings suggest that the knowledge-enhanced model achieves excellent performance on the testing dataset (field observation data) in terms of effectiveness and transferability, compared with the model-driven and data-driven methods.

Figure 8 illustrates some cases of trucks' operating speed prediction on four typical road segments using different types of modeling methods. Compared to the linear regression model, the data-driven and knowledge-enhanced models proposed in this study demonstrate a greater accu-

racy in capturing actual speed variations. However, the predicted results of the data-driven RoadGNN-S sometimes have a large deviation from the real values. To be specific, as shown in Figure 8a, without considering driver expectations, the predicted operating speed at the endpoint of a tangent segment is 6.2 km/h lower than the observed value. Due to the neglect of vehicle dynamics, the data-driven model predicts that trucks will accelerate at the endpoint of a vertical steep slope, whereas the observations show a deceleration trend, as illustrated in Figure 8b. In the horizontal curve segment, owing to the lack of consideration of highway geometric supply, the predicted operating speed at the midpoint of the curve is about 7.5 km/h higher than the observed value, which increases the risk of truck rollovers (see Figure 8c). Figure 8d shows that in the curved slope segment, the data-driven model predicts that there is sudden deceleration before entering curves and rapid acceleration after exiting curves, while the actual changes are much gentler. In contrast, as demonstrated in Figure 8, the operating speed prediction results of the knowledge-enhanced RoadGNN-S have better accuracies on four typical highway segments, relative to the model-driven and data-driven methods. This might be because knowledge-enhanced modeling aligns more closely with prior knowledge, which can help effectively avoid the fundamental errors observed in data-driven modeling. In general, the above analyses indicate that knowledge-enhanced modeling outperforms data-driven modeling by integrating prior knowledge to enhance prediction accuracy and transferability.

4.3 | Influence of multi-dimensional perspective alignment elements on operating speed

To further analyze the coupling impacts of multi-perspective alignment elements on operating speed, this study visualizes the self-attention weights (i.e., $\alpha_{vu}^{(k)}$) of the

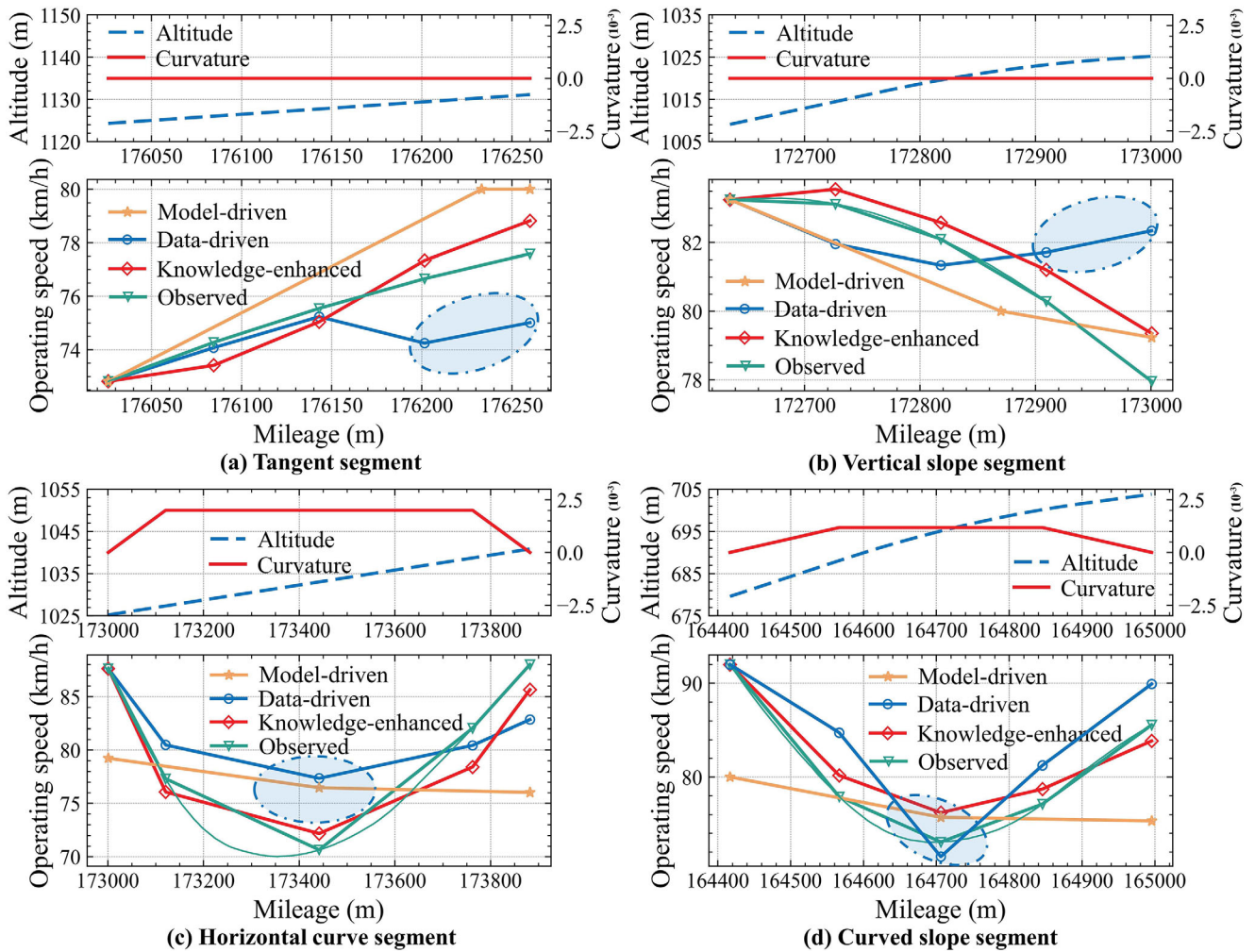


FIGURE 8 Truck operating speed prediction profiles on the testing dataset (field observation data).

GAT layer in the proposed RoadGNN-S. Based on different types of connected nodes, the self-attention weights $\alpha_{vu}^{(k)}$ can be mainly categorized into self-attention weight α_{ee} and self-attention weight α_{ev85} . The self-attention weight α_{ee} between alignment element nodes represents the intensity of the message passing between different alignment element nodes (including in the same perspective and in different perspectives), while the self-attention weight α_{ev85} between alignment element nodes and operating speed nodes denotes the direct impact of alignment elements on operating speed. The indirect impact of alignment elements on operating speed is transferred through messages between alignment element nodes. In the GAT layer self-attention mechanism, the aggregate weight α_a of the alignment elements of different perspectives is obtained through the weighted sum of α_{ee} and α_{ev85} to quantify the coupling impacts of multi-perspective alignment elements on operating speed.

Figure 9 illustrates the self-attention weight distribution of passenger cars' operating speed prediction model. In

four typical road segments, the mean aggregation weights α_a of alignment element nodes across 2D, 2.5D, and 3D perspectives are 0.24, 0.47, and 0.29, respectively, indicating a dominant role of the 2.5D driving perspective virtual curve in predicting passenger cars' operating speed. The thicker the edge, the larger the self-attention weight α_{ev85} from alignment element nodes and operating speed nodes at different spots along the road segment. Within the same type of road segment, the edges connecting 2.5D perspective virtual curve nodes with operating speed nodes at different spots (i.e., V_{85}^0 , V_{85}^1 , V_{85}^2 , V_{85}^3 , and V_{85}^4) exhibit significant variations in thickness. In particular, in the curved slope segments, the self-attention weights α_{ev85} are 0.043, 0.016, 0.081, 0.012, and 0.028. The unbalanced distribution of these self-attention weights α_{ev85} indicates that drivers can selectively adjust their current speed based on specific visual cues at different spots along the highway alignment. Additionally, the lighter the color, the larger the self-attention weights α_{ee} between alignment element nodes. The mean self-attention weight α_{ee} between 2D

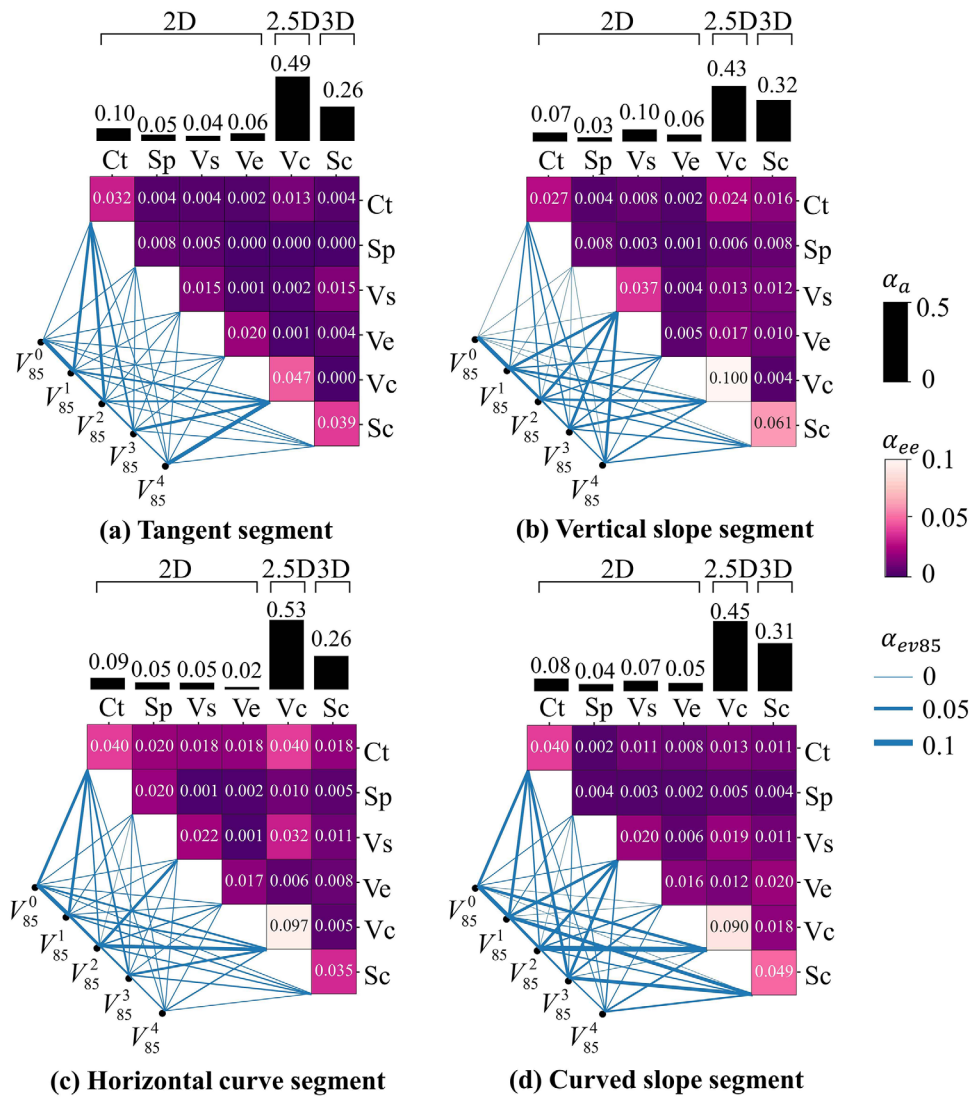


FIGURE 9 Self-attention weight distribution of passenger cars' operating speed prediction model. *Note:* the descriptions of all the variable codes are demonstrated in Table 1. 2D, two-dimensional; 2.5D, 2.5-dimensional; 3D, three-dimensional.

perspective alignment element nodes (i.e., Ct , Sp , Vs , Ve) and 2.5D perspective virtual curve nodes Vc is equal to 0.27, accounting for the largest proportion 61.4%. The following one is the mean self-attention weight between 2D perspective alignment element nodes and 3D perspective spatial curve nodes Sc , with the value 0.11. The lowest mean self-attention weight (0.06) is the one between 2.5D perspective virtual curve nodes and the 3D perspective spatial curve nodes Sc . These results suggest that the impact of the 2D perspective alignment elements on passenger cars' operating speed is mainly transferred through the 2.5D perspective virtual curve rather than direct effects.

Figure 10 demonstrates the self-attention weight distribution of trucks' operating speed prediction model. The mean aggregation weights α_a of alignment element nodes across 2D, 2.5D, and 3D perspectives in four typical road segments are 0.58, 0.16, and 0.26, respectively, suggesting

that the 2D perspective alignment elements play a leading part in predicting trucks' operating speed. Specifically, within the 2D perspective, circular curve nodes Ct and vertical slope nodes Vs exhibit the two highest self-attention weights α_{ev85} toward operating speed nodes in four typical road segments, with mean values of 0.082 and 0.065, respectively. This means that lateral stability and slope resistance are the most critical factors affecting trucks' operating speed. In addition, the impact of the 3D perspective spatial curve on the trucks' operating speed is mainly presented in curved slope segments, with an aggregation weight α_a of 0.36, whereas those weights in other types of road segment types are all below 0.25. In addition, the mean self-attention weight α_{ee} of four typical road segments reaches the largest value (0.18) when it is between 2D perspective alignment element nodes (i.e., Ct , Sp , Vs , Ve) and 3D perspective spatial curve nodes Sc ,

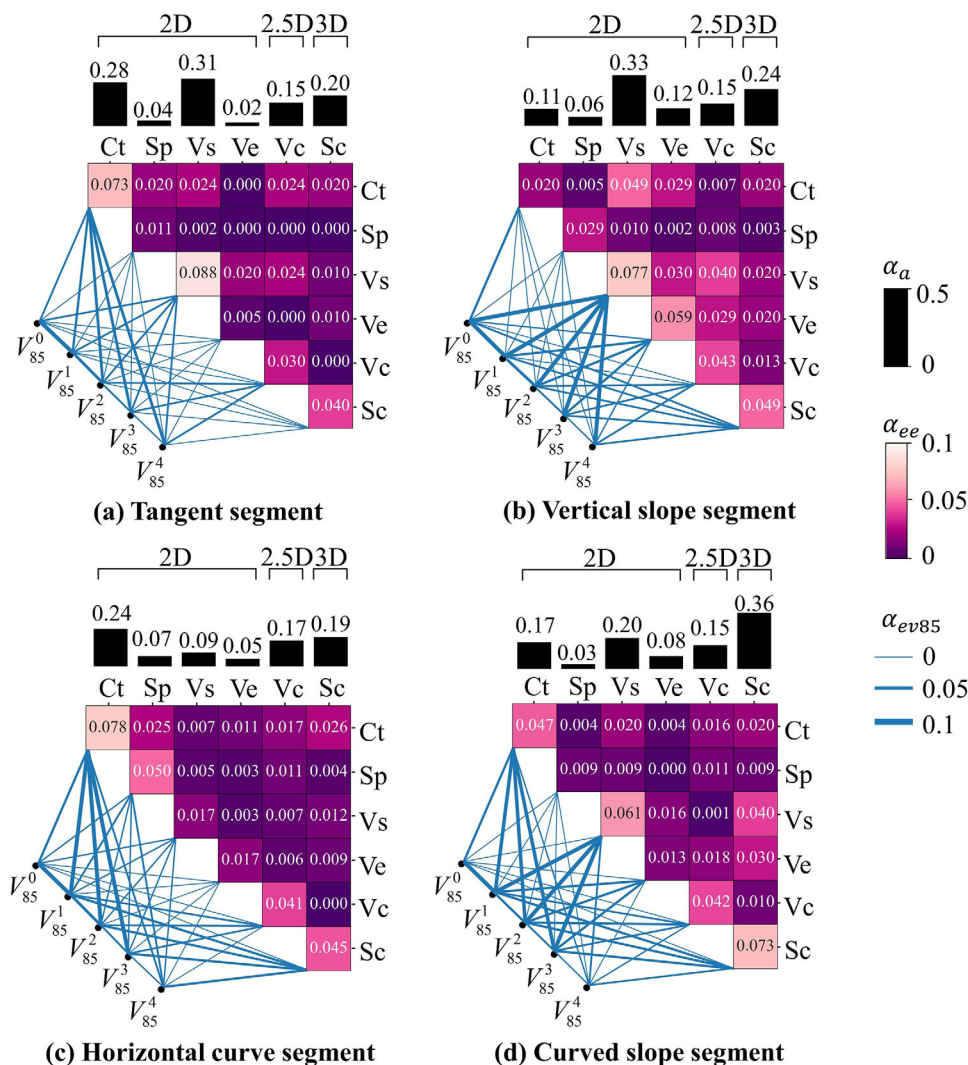


FIGURE 10 Self-attention weight distribution of trucks' operating speed prediction model. *Note:* the descriptions of all the variable codes are demonstrated in Table 1. 2D, two-dimensional; 2.5D, 2.5-dimensional; 3D, three-dimensional.

representing up to 78.3% of the total weights. The mean self-attention weight between 2D perspective alignment element nodes and 2.5D perspective virtual curve nodes Vc is only 0.03. Similarly, the mean self-attention weight between 2.5D perspective virtual curve nodes and 3D perspective spatial curve nodes Sc is as low as 0.02. These results indicate that in the process of predicting trucks' operating speed, 3D perspective spatial curves are closely associated with complex combinations of horizontal and vertical alignment, serving as a complement to 2D perspective alignment elements.

5 | DISCUSSION

This study first uses graph structure to represent the road alignment segments, and then GNNs are applied to predict operating speed. The overall prediction performance of the

proposed RoadGNN-S is better than commonly used methods, with R^2 values of 0.94 and 0.96 for passenger cars' and trucks' operating speeds, respectively. In comparison, most operating speed prediction models for both passenger cars and trucks, based on linear regression considering only horizontal radius and vertical grade, have R^2 values below 0.85 (Martinelli et al., 2023; Sil et al., 2020). When accounting for spatial curvature and vertical grade, operating speed prediction models for passenger cars and trucks using multi-parameter nonlinear regression achieve R^2 values of 0.76 and 0.74, respectively (Xiaofei Wang et al., 2019c). Another model predicting passenger cars' operating speed, which incorporates visual curvature and depth through generalized exponential regression, reaches an R^2 value of 0.66 (F. Wang et al., 2019). Additionally, RoadGNN-S outperforms classical deep learning models (e.g., FFN, RNN, and CNN) in operating speed prediction, with accuracies increasing by over 10% for both passenger cars and trucks.



This is probably because in GNNs, MPAS graphs can better describe the alignment elements information of multiple perspectives, and the message-passing and multi-head self-attention mechanisms are applied to capture both local and global dependencies between alignment element nodes and operating speed nodes. In contrast, complex fully connected structures in FFN can lead to a high divergence value (F. Yu et al., 2021). RNN reliance on historical states poses challenges in handling long-term dependencies and vanishing gradient problems (Kashyap et al., 2022), while CNN only focuses on local features due to kernel size limitations (Ren et al., 2024).

The overall predictive accuracy and transferability of the knowledge-enhanced modeling compare favorably to data-driven modeling. The proposed knowledge-enhanced RoadGNN-S achieves R^2 values of 0.92 for predicting the passenger cars' and trucks' operating speed on real highways, which are 7.9% and 10.7% higher than those of the data-driven RoadGNN-S, respectively. The reason might be that in the backpropagation algorithm of GNNs, speed and acceleration penalty terms are considered a strong regularization, which limits the model search space to a large extent and effectively guarantees the theoretical correctness of the solution (H. Gao et al., 2024). Previous studies have also shown the outstanding predictive performance of knowledge-enhanced GNNs in different fields. For instance, the physical equations of structural mechanics are integrated into the elastic structural analysis model based on GNNs, which improves the prediction accuracy of node forces and moments by 25.8% (Song et al., 2023). Through the introduction of a knowledge inclusion strategy into the loss function and the output layer of a GNNs-based model for building material quantity prediction, the prediction error is significantly reduced by 30.4% (Fei et al., 2023).

Multi-perspective alignment fusion is a new addition, compared with prior research on operating speed prediction. The results of this study indicate that 2D perspective and 2.5D perspective alignment elements are the most important features for trucks' and passenger cars' operating speed prediction, respectively. It has been reported that due to the superior handling and power of passenger cars, the impact of circular curves and vertical slopes in 2D perspective on the speed and acceleration of passenger cars is nearly 22% less than that of trucks (Liu et al., 2020). The impact of 2.5D alignment elements on the passenger cars' operating speed is reflected in two aspects: visual perception deviation and visual cues. Deviations that exist between the actual and perceived curvature and length of the 2.5D perspective visual curve can lead drivers to make incorrect speed choices (Hassan & Sarhan, 2012). Drivers' visual perception precedes the current road segment, so drivers are able to determine their deceleration and accel-

eration behavior in advance based on the visual cues ahead (Vos et al., 2021). Due to the limitation of vehicle power and braking performance, trucks tend to travel at a more uniform speed when approaching horizontal curves compared to passenger cars (Morris & Donnell, 2014). This might be a primary reason why the impact of 2.5D perspective alignment elements on trucks' operating speed is not obvious. Additionally, this study finds that the 3D perspective alignment elements have an advantage in capturing the impact of complex combinations of horizontal and vertical alignment on trucks' operating speed. Another study also observes that trucks' operating speed exponentially declines with spatial curvature (Wang et al., 2019). Therefore, the multi-perspective alignment fusion can better consider the interrelationships among various alignment combinations to ensure a comprehensive understanding of speed dynamics. Overall, increasing evidence confirms the significance of multi-perspective alignment fusion for predicting operating speed.

6 | CONCLUSION

This study fills the research gap by proposing a multi-perspective alignment element fusion model for predicting operating speed from the following aspects. (1) A novel highway alignment segment graph is constructed that can realize modular representation and multi-perspective fusion of highway alignment. (2) GNNs are applied to establish the operating speed prediction model RoadGNN-S, which can effectively capture the coupling impacts of multi-perspective alignment elements on operating speed based on the message passing and multi-head self-attention mechanisms. (3) The prior knowledge of highway geometry supply, driver expectations, and vehicle dynamics are integrated into GNNs to further enhance the prediction accuracy and transferability of RoadGNN-S, ensuring the theoretical correctness of solutions.

The findings of this study have several practical implications. (1) The MPAS graph can effectively fuse multi-perspective alignment elements, which provides a novel and universal solution for comprehensive understanding and representation of highway alignment. This method is helpful to establish a promising bidirectional link between multi-perspective alignment elements and operating speed. The proposed GNNs can achieve a more accurate prediction of operating speed and offer profound insights into quantifying the coupling impacts of multi-perspective highway alignment elements on operating speed. (2) After introducing prior knowledge into the GNN, the theoretical correctness of the prediction result can be guaranteed, which effectively enhances the proposed RoadGNN-S's predication accuracy and



transferability. Accurately measuring the difference between operating speed and design speed, evaluating the variability of operating speeds between consecutive highway segments, as well as comparing the required side friction force and the actual side friction force, are essential for ensuring the consistency of highway geometric design (Camacho-Torregrosa et al., 2013). In addition, this method also has the potential to be applied in other highway design-related deep learning frameworks, paving the way for future advancements toward more intelligent and safer end-to-end highway geometric design. (3) The proposed framework in this study, based on multi-perspective fusion and GNNs, is highly flexible and can be extended to capture the driving speed of different types of travelers, including connected and autonomous vehicles (CAVs) and individual human drivers. For CAVs, by considering vehicle-to-vehicle and vehicle-to-infrastructure interactions, the proposed RoadGNN-S can effectively integrate the real-time multi-modal data (e.g., traffic conditions, vehicle movements, surrounding environments, etc.) to improve the accuracy of driving speed prediction. For human drivers, the proposed RoadGNN-S can predict the driving speed of various driver types by further incorporating behavioral features such as fatigue, age, and driving style (Berghaus et al., 2024).

One of the limitations of this study is that the impact of different geographical regions and highway configurations on operating speed has not been considered. Highways in plains, mountainous areas, and high-altitude areas have great differences in vehicle power performance and driver physiological conditions. More operating speed data and prior knowledge should be collected from different regions, varying design speeds, and different types of roads (e.g., highways, rural roads, and logging roads) to further enhance the broad applicability and accuracy of RoadGNN-S. Additionally, although driving simulation experiments have the advantage of acquiring large amounts of samples and simulating diverse scenarios, field observation experiments can provide irreplaceable scenario details and complex vehicle interactions under real-world conditions. This study has attempted to combine driving simulation with field observation data for analyses. In future work, more field observation data from various types of highways and geographic regions will be collected to further validate and enhance the robustness of the proposed model, in order to better reflect drivers' speed choices under real-world conditions.

ACKNOWLEDGMENTS

This study is jointly supported by the National Natural Science Foundation of China (52102416), the National Key R & D Program of China (2023YFE0202400), and the Natural Science Foundation of Shanghai (22ZR1466000).

REFERENCES

- AbuAddous, M. (2021). Operating speed prediction models for tangent segments: A brief review. *Civil Engineering Journal*, 7(12), 2150–2164.
- Al-Sahili, K., & Dwaikat, M. (2019). Modeling geometric design consistency and road safety for two-lane rural highways in the west bank, Palestine. *Arabian Journal for Science and Engineering*, 44, 4895–4909.
- Adeli, H., & Ghosh-Dastidar, S. (2004). Mesoscopic-wavelet freeway work zone flow and congestion feature extraction model. *Journal of Transportation Engineering*, 130(1), 94–103.
- American Association of State Highway and Transportation Officials. (2018). *A policy on geometric design of highways and streets* (7th ed.). American Association of State Highway and Transportation Officials.
- Aironi, C., Cornell, S., & Squartini, S. (2024). A graph-based neural approach to linear sum assignment problems. *International Journal of Neural Systems*, 34(3), 2450011.
- Bazrafshan, P., On, T., Basereh, S., Okumus, P., & Ebrahimkhanlou, A. (2024). A graph-based method for quantifying crack patterns on reinforced concrete shear walls. *Computer-Aided Civil and Infrastructure Engineering*, 39(4), 498–517.
- Berghaus, M., Lamberty, S., Ehlers, J., Kalló, E., & Oeser, M. (2024). Vehicle trajectory dataset from drone videos including off-ramp and congested traffic—Analysis of data quality, traffic flow, and accident risk. *Communications in Transportation Research*, 4, 100133.
- Bouritsas, G., Frasca, F., Zafeiriou, S., & Bronstein, M. M. (2022). Improving graph neural network expressivity via subgraph isomorphism counting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1), 657–668.
- Cafiso, S., & Cerni, G. (2012). New approach to defining continuous speed profile models for two-lane rural roads. *Transportation Research Record*, 2309(1), 157–167.
- Camacho-Torregrosa, F. J., Pérez-Zuriaga, A. M., Campoy-Ungria, J. M., & García-García, A. (2013). New geometric design consistency model based on operating speed profiles for road safety evaluation. *Accident Analysis & Prevention*, 61, 33–42.
- Casal, G., Santamarina, D., & Vázquez-Méndez, M. E. (2017). Optimization of horizontal alignment geometry in road design and reconstruction. *Transportation Research Part C: Emerging Technologies*, 74, 261–274.
- Cardoso, J., De Macedo, A. L., Kanellaidis, G., Flouda, A., Dimitropoulos, I., Peltola, H., & Dupre, G. (1998). *Improvements of models on the relations between speed and road characteristics* (Task 5.3. SAFESTAR). Laboratório Nacional de Engenharia Civil, Lisboa, Portugal.
- Chen, F., Peng, H., Ma, X., Liang, J., Hao, W., & Pan, X. (2019). Examining the safety of trucks under crosswind at bridge-tunnel section: A driving simulator study. *Tunnelling and Underground Space Technology*, 92, 103034.
- Chen, T., Oviedo-Trespalacios, O., Sze, N. N., & Chen, S. (2022). Distractions by work-related activities: The impact of ride-hailing app and radio system on male taxi drivers. *Accident Analysis & Prevention*, 178, 106849.
- Chen, Y., Quddus, M., & Wang, X. (2018). Impact of combined alignments on lane departure: A simulator study for mountainous freeways. *Transportation Research Part C: Emerging Technologies*, 86, 346–359.



- Dhahir, B., & Hassan, Y. (2016). Reliability-based design of horizontal curves on two-lane rural highways. *Transportation Research Record*, 2588(1), 22–31.
- Donnell, E., Wood, J., Himes, S., & Torbic, D. (2016). Use of side friction in horizontal curve design: A margin of safety assessment. *Transportation Research Record*, 2588(1), 61–70.
- Dong, C., & Sun, D. (2024). Spatial-temporal dynamic hypergraph information bottleneck for brain network classification. *International Journal of Neural Systems*, 34(10), 2450053.
- Eboli, L., Guido, G., Mazzulla, G., & Pungillo, G. (2017). Experimental relationships between operating speeds of successive road design elements in two-lane rural highways. *Transport*, 32(2), 138–145.
- Fei, Y., Liao, W., Lu, X., & Guan, H. (2023). Knowledge-enhanced graph neural networks for construction material quantity estimation of reinforced concrete buildings. *Computer-Aided Civil and Infrastructure Engineering*, 39(4), 518–538.
- Fei, Y., Liao, W., Zhang, S., Yin, P., Han, B., Zhao, P., Chen, X., & Lu, X. (2022). Integrated schematic design method for shear wall structures: A practical application of generative adversarial networks. *Buildings*, 12(9), 1295.
- Gong, H., & Stamatiadis, N. (2008). Operating speed prediction models for horizontal curves on rural four-lane highways. *Transportation Research Record*, 2075(1), 1–7.
- Gao, H., Hu, G., Zhang, D., Jiang, W., Tse, K. T., Kwok, K. C. S., & Kareem, A. (2024). Urban wind field prediction based on sparse sensors and physics-informed graph-assisted auto-encoder. *Computer-Aided Civil and Infrastructure Engineering*, 39(10), 1409–1430.
- Gao, J., Yu, B., Chen, Y., Bao, S., Gao, K., & Zhang, L. (2024). An ADAS with better driver satisfaction under rear-end near-crash scenarios: A spatio-temporal graph transformer-based prediction framework of evasive behavior and collision risk. *Transportation Research Part C: Emerging Technologies*, 159, 104491.
- Hassan, Y., & Sarhan, M. (2012). Operational effects of drivers' misperception of horizontal curvature. *Journal of Transportation Engineering*, 138(11), 1314–1320.
- He, L., Yu, B., Chen, Y., Bao, S., Gao, K., & Kong, Y. (2023). An interpretable prediction model of illegal running into the opposite lane on curve sections of two-lane rural roads from drivers' visual perceptions. *Accident Analysis & Prevention*, 186, 107066.
- Himes, S. C., & Donnell, E. T. (2010). Speed prediction models for multilane highways: Simultaneous equations approach. *Journal of Transportation Engineering*, 136(10), 855–862.
- Jiang, X., & Adeli, H. (2004). Object-oriented model for freeway work zone capacity and queue delay estimation. *Computer-Aided Civil and Infrastructure Engineering*, 19(2), 144–156.
- Jurecki, R. S., & Stańczyk, T. L. (2021). A methodology for evaluating driving styles in various road conditions. *Energies*, 14(12), 3570.
- Karim, A., & Adeli, H. (2003a). Fast automatic incident detection on urban and rural freeways using wavelet energy algorithm. *Journal of Transportation Engineering*, 129(1), 57–68.
- Karim, A., & Adeli, H. (2003b). CBR model for freeway work zone traffic management. *Journal of Transportation Engineering*, 129(2), 134–145.
- Kashyap, A. A., Raviraj, S., Devarakonda, A., Nayak, K. S. R., Santhosh, K. V., & Bhat, S. J. (2022). Traffic flow prediction models—A review of deep learning techniques. *Cogent Engineering*, 9(1), 2010510.
- Kim, K.-R., Kim, P. T., Koo, J.-Y., & Pierrynowski, M. R. (2012). Frenet-Serret and the estimation of curvature and torsion. *IEEE Journal of Selected Topics in Signal Processing*, 7(4), 646–654.
- Li, R., Li, J., Wang, C., Liu, H., Liu, T., Wang, X., Zou, T., Huang, W., Yan, H., & Chen, H. (2024). Multi-semantic decoding of visual perception with graph neural networks. *International Journal of Neural Systems*, 34(4), 2450016.
- Liu, H., Rodgers, M. O., Liu, F. C., & Guensler, R. (2020). Bayesian approach in estimating the road grade impact on vehicle speed and acceleration on freeways. *Transportmetrica A: Transport Science*, 16(3), 602–625.
- Marr, D. (2010). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press.
- Maji, A., Singh, D., Agrawal, N., & Zaman, M. (2020). Operating speed prediction models for tangent sections of two-lane rural highways in Oklahoma State. *Transportation Letters*, 12(2), 130–137.
- McFadden, J., Yang, W. T., & Durrans, S. (2001). Application of artificial neural networks to predict speeds on two-lane rural highways. *Transportation Research Record*, 1751(1), 9–17.
- Marinelli, G., Bassani, M., Piras, M., & Lingua, A. M. (2017). Mobile mapping systems and spatial data collection strategies assessment in the identification of horizontal alignment of highways. *Transportation Research Part C: Emerging Technologies*, 79, 257–273.
- Marino, K., Chen, X., Parikh, D., Gupta, A., & Rohrbach, M. (2021). KRISP: Integrating implicit and symbolic knowledge for open-domain knowledge-based VQA. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, TN (pp. 14111–14121).
- Martinelli, V., Ventura, R., Bonera, M., Barabino, B., & Maternini, G. (2023). Estimating operating speed for county road segments—Evidence from Italy. *International Journal of Transportation Science and Technology*, 12(2), 560–577.
- Medina, A. M. F., & Tarko, A. P. (2005). Speed factors on two-lane rural highways in free-flow conditions. *Transportation Research Record*, 1912(1), 39–46.
- Morris, C. M., & Donnell, E. T. (2014). Passenger car and truck operating speed models on multilane highways with combinations of horizontal curves and steep grades. *Journal of Transportation Engineering*, 140(11), 04014058.
- Pan, J., Lin, H., Dong, Y., Wang, Y., & Ji, Y. (2022). MAMF-GCN: Multi-scale adaptive multi-channel fusion deep graph convolutional network for predicting mental disorder. *Computers in Biology and Medicine*, 148, 105823.
- Qu, X., & Wang, S. (2015). Long-distance-commuter (LDC) lane: A new concept for freeway traffic management. *Computer-Aided Civil and Infrastructure Engineering*, 30(10), 815–823.
- Ren, W., Yu, B., Chen, Y., Gao, K., Bao, S., Wang, Z., & Qin, Y. (2024). An intelligent optimization method for the facility environment on rural roads. *Computer-Aided Civil and Infrastructure Engineering*, 39, 2559–2580.
- Samant, A., & Adeli, H. (2001). Enhancing neural network traffic incident-detection algorithms using wavelets. *Computer-Aided Civil and Infrastructure Engineering*, 16(4), 239–245.
- Semeida, A. M. (2014). Application of artificial neural networks for operating speed prediction at horizontal curves: A case study in Egypt. *Journal of Modern Transportation*, 22, 20–29.



- Sheng, Z., Huang, Z., & Chen, S. (2024). Traffic expertise meets residual RL: Knowledge-informed model-based residual reinforcement learning for CAV trajectory control. *Communications in Transportation Research*, 4, 100142.
- Sil, G., Nama, S., Maji, A., & Maurya, A. K. (2020). Effect of horizontal curve geometry on vehicle speed distribution: A four-lane divided highway study. *Transportation Letters*, 12(10), 713–722.
- Silvano, A. P., Koutsopoulos, H. N., & Farah, H. (2020). Free flow speed estimation: A probabilistic, latent approach. Impact of speed limit changes and road characteristics. *Transportation Research Part A: Policy and Practice*, 138, 283–298.
- Song, L. H., Wang, C., Fan, J. S., & Lu, H. M. (2023). Elastic structural analysis based on graph neural network without labeled data. *Computer-Aided Civil and Infrastructure Engineering*, 38(10), 1307–1323.
- The Ministry of Transport of People's Republic of China. (2015). *Specifications for highway safety audit (JTG B05-2015)*. China Communications Press.
- The Ministry of Transport of the People's Republic of China. (2017). *Design specification for highway alignment (JTG D20-2017)*. China Communications Press.
- Vos, J., de Winter, J., Farah, H., & Hagenzieker, M. (2023). Which visual cues do drivers use to anticipate and slow down in freeway curve approach? An eye-tracking, think-aloud on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 94, 190–211.
- van der El, K., Pool, D. M., & Mulder, M. (2019). Measuring and modeling driver steering behavior: From compensatory tracking to curve driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 61, 337–346.
- Vos, J., Farah, H., & Hagenzieker, M. (2021). How do Dutch drivers perceive horizontal curves on freeway interchanges and which cues influence their speed choice? *IATSS Research*, 45(2), 258–266.
- Wang, X., Wang, X., Cai, B., & Liu, J. (2019a). Combined alignment effects on deceleration and acceleration: A driving simulator study. *Transportation Research Part C: Emerging Technologies*, 104, 172–183.
- Wang, X., Liu, S., Cai, B., Guo, Q., & Wang, X. (2019b). Application of driving simulator for freeway design safety evaluation: A sample size study. In *The 98th Transportation Research Board Annual Meeting*, Washington DC, USA.
- Wang, X., Pu, H., Yao, J., Yan, Y., Li, X., & Zeng, Z. (2019c). A novel speed model for safety evaluation of freeway alignment in Euclidean 3D space. *Traffic Injury Prevention*, 20(4), 392–399.
- Wang, F., Chen, Y., & Lin, K. (2019). Using perception of alignment and short-term memory to understand middle-aged driver's driving speed on mountainous highway. *2019 5th International Conference on Transportation Information and Safety (ICTIS)*, Liverpool, UK (pp. 1098–1104).
- Wang, F., Chen, Y., Wijnands, J. S., & Guo, J. (2020). Modeling and interpreting road geometry from a driver's perspective using variational autoencoders. *Computer-Aided Civil and Infrastructure Engineering*, 35(10), 1148–1159.
- Wang, X., Guo, Q., & Tarko, A. P. (2020). Modeling speed profiles on mountainous freeways using high resolution data. *Transportation Research Part C: Emerging Technologies*, 117, 102679.
- Wang, X., Chen, Z., Guo, Q., Tarko, A., Lizarazo, C., & Wang, X. (2021). Transferability analysis of the freeway continuous speed model. *Accident Analysis & Prevention*, 151, 105944.
- Wang, J., He, S., Zhai, X., Wang, Z., & Fu, X. (2022). Estimating mountainous freeway crash rate: Application of a spatial model with three-dimensional (3D) alignment parameters. *Accident Analysis & Prevention*, 170, 106634.
- Wang, Y., Shen, D., & Teoh, E. K. (2000). Lane detection using spline model. *Pattern Recognition Letters*, 21(8), 677–689.
- Xu, Z., Wei, T., Easa, S., Zhao, X., & Qu, X. (2018). Modeling relationship between truck fuel consumption and driving behavior using data from internet of vehicles. *Computer-Aided Civil and Infrastructure Engineering*, 33(3), 209–219.
- Xu, M., Di, Y., Ding, H., Zhu, Z., Chen, X., & Yang, H. (2023). AGNP: Network-wide short-term probabilistic traffic speed prediction and imputation. *Communications in Transportation Research*, 3, 100099.
- Yan, Y., & Zhang, Y.-H. (2010). Operating speed prediction model and systems of highway based on space compound curvature. *2010 International Conference on Intelligent Computing and Integrated Systems*, Guilin, China (pp. 379–382).
- Yu, B., Chen, Y., & Bao, S. (2019). Quantifying visual road environment to establish a speeding prediction model: An examination using naturalistic driving data. *Accident Analysis & Prevention*, 129, 289–298.
- Yu, F., Bray, S., Wang, D., Shangguan, L., Tang, X., Liu, C., & Chen, X. (2021). Automated runtime-aware scheduling for multi-tenant DNN inference on GPU. *2021 IEEE/ACM International Conference on Computer Aided Design (ICCAD)*, Munich, Germany (pp. 1–9).
- Zhang, C., Tao, M. X., Wang, C., & Fan, J. S. (2024). End-to-end generation of structural topology for complex architectural layouts with graph neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 39(5), 756–775.
- Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., & Sun, M. (2020). Graph neural networks: A review of methods and applications. *AI Open*, 1, 57–81.

How to cite this article: Gao, J., Yu, B., Chen, Y., Gao, K., & Bao, S. (2025). A multi-perspective fusion model for operating speed prediction on highways using knowledge-enhanced graph neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 40, 1004–1027.
<https://doi.org/10.1111/mice.13382>