Sound & Vibration 2026, 60(5), 4536.
doi: https://doi.org/10.59400/sv4536

Article

Rapid prediction of vehicle interior wind noise via a hybrid convolutional neural network-transformer model and geometric styling features

Hongwei Yi[1,2], Penghu Li[3], Jifeng Wang[3], Yuwei Deng[4], Xiaorong Huang[1,2,*], Haibo Huang[3]

[1] Vehicle Measurement Control and Safety Key Laboratory of Sichuan Province, Xihua University, Chengdu 610039, China

[2] Engineering Research Center of Intelligent Control and Simulation Test Technology for New Energy Vehicles of Sichuan Province, Xihua University, Chengdu 610039, China

[3] School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China

[4] Global R&D Center, China FAW Corporation, Limited, Changchun 130013, China

*** Corresponding author:** Xiaorong Huang, 1220190017@xhu.edu.cn

CITATION

Yi H, Li P, Wang J, et al. Rapid prediction of vehicle interior wind noise via a hybrid convolutional neural network-transformer model and geometric styling features. Sound & Vibration. 2026; 6(5): 4536. https://doi.org/10.59400/sv4536

ARTICLE INFO

Received: 19 June 2026
Revised: 15 July 2026
Accepted: 27 July 2026
Available online: 15 August 2026

Abstract: The rapid evaluation of interior aerodynamic noise during the Concept A Surface design stage is important for vehicle acoustic development, but conventional methods are limited by high cost and low efficiency. This study proposes a convolutional neural network transformer-based prediction method for vehicle interior wind noise by integrating vehicle styling features and acoustic technical parameters. An optimal Latin hypercube sampling method was used to generate design combinations, and wind tunnel tests were conducted at 120 km/h. Key vehicle styling parameters, including A-pillar geometry, side mirror dimensions, front windscreen angle, side mirror-to-body spacing, and side window inclination, together with glazing material properties, glass thickness, acoustic transfer function, and interior reverberation time, were selected as input features to predict the driver's left-ear wind noise spectrum. Based on five-fold cross-validation, the proposed model was compared with convolutional neural network (CNN), long short-term memory (LSTM), and transformer models. The CNN-Transformer model achieved the best performance, with mean absolute percentage error (MAPE) and root mean square error (RMSE) values of 2.23% and 0.94 dB, respectively. Compared with the transformer, CNN, and LSTM models, the proposed method reduced MAPE by 24.91%, 36.29%, and 49.59%, and reduced RMSE by 22.95%, 35.17%, and 48.07%, respectively. The model also maintained reliable performance on an independent test set, with MAPE and RMSE values of 4.82% and 1.44 dB. The mean impact value method was further applied to identify the influence of design parameters on interior wind noise, guiding for early vehicle acoustic optimization.

Keywords: interior wind noise; styling features; convolutional neural network; transformer

1. Introduction

Due to a significant shift in the criteria used to evaluate overall vehicle performance, consumer attention, which was previously focused on power performance and fuel economy, has increasingly shifted toward comprehensive driving experiences such as ride comfort and the refinement of wind and road noise [1]. To meet the increasing demand for enhanced comfort, further development of noise, vibration, and harshness (NVH) technologies is required [2, 3]. When a vehicle is in motion, the drivetrain, engine operation, and the interaction between the vehicle body and the surrounding airflow each generate distinct noise sources [4–7]. As conventional noise control technologies continue to improve, the contributions of powertrain and engine noise to interior noise have progressively decreased, while aerodynamic noise under high speed operating conditions has become increasingly prominent [8]. Previous studies have reported that aerodynamic noise becomes a dominant noise source at vehicle speeds above 100 km/h, as intense airflow interaction around the vehicle surface leads to a significant increase in aerodynamic excitation [9, 10]. With the continuous growth of vehicle electrification, conventional powertrain noise contributes less to total vehicle noise, and aerodynamic noise has become a more critical factor in vehicle acoustic performance [11]. Extensive research in related fields has shown that long-term exposure to excessive noise inside the vehicle can lead to a variety of adverse effects. Continuing in such an acoustic environment will affect the physical and mental health of the occupants and cause various health discomforts. In addition, it may also reduce the driver's attention, weaken the judgment and response ability during driving, and further increase the safety risk on the road [12, 13]. Therefore, interior noise reduction has become an important focus in automotive NVH development [14–17].

At present, the evaluation of vehicle wind noise is mainly carried out by wind tunnel test and real vehicle measurement. These methods can provide accurate noise data under real driving conditions and support the development of vehicle NVH [18]. Among them, because the wind tunnel test conditions are controllable and reproducible, it is widely used in vehicle wind noise assessment. Yu and Davies [19] developed and validated a method for simulating unsteady interior wind noise by combining aeroacoustic wind tunnel measurements and multiple yaw angle scanning tests. Their study examined how wind noise changes in the vehicle under dynamic intake conditions and how it affects people's subjective feelings. This provided a basis for evaluating and improving acoustic comfort inside the compartment. Tajima et al. [20] carried out vehicle wind noise experiments under turbulent and fluctuating wind conditions using an aeroacoustic wind tunnel equipped with an active turbulence generation system. By combining the microphone array and surface pressure measurement with the modulation power spectrum analysis, the modulation noise caused by the longitudinal vortex of the A-pillar under unsteady intake conditions is evaluated. The results quantify the contribution of these vortices to the cabin noise, and show that suppressing airflow separation at the A-pillar can reduce the modulation noise. Jamaluddin et al. [18] studied the modulated wind noise in the carriage under unsteady crosswind conditions by using the active crosswind system. The relationship between wind noise modulation and unsteady flow structure is analyzed by microphone array measurement and side window surface pressure data. The results showed that the complex intake conditions will aggravate the broadband wind noise fluctuation inside the carriage. Guleria et al. [21] simulated heavy truck wind noise with the lattice Boltzmann method. They confirmed the results with real road vehicle data. This study analyzed the changes of wind noise under different vehicle speeds and actual driving conditions. In general, these experimental studies have significantly improved the understanding of the generation mechanism of aerodynamic noise under different flow conditions. However, the wind tunnel test is mainly used as a verification and calibration tool in vehicle development. Because it relies on physical prototypes and complex test procedures, it is not suitable as a means to evaluate a large number of design schemes in the early conceptual design stage.

Advances in numerical techniques and computing capability have promoted the application of computational fluid dynamics (CFD) in automotive aeroacoustics. CFD avoids the high cost and limitations of wind tunnel tests, and can evaluate the aerodynamic noise characteristics of the external airflow field of the vehicle without physical experiments. Josefsson et al. [22] used high-resolution numerical simulation and wind tunnel tests to study the external airflow structure of sport utility vehicles and cars. The results show that the vortex structures in the rear and hub regions are closely associated with surface pressure fluctuations and contribute to changes in vehicle drag and aeroacoustic characteristics under high-speed conditions. Nusser and Becker [23] used incompressible large eddy simulation to study the turbulent flow field and extract sound sources and structural loads. This enabled coupled analysis of fluid, structure, and acoustics around the A-pillar and side window. Jia et al. [24] developed a high-fidelity computational aeroacoustic approach by combining the lattice Boltzmann method with large eddy simulation for aerodynamic noise prediction. The proposed method accurately captured unsteady flow structures and their associated acoustic radiation characteristics, demonstrating the effectiveness of CFD-based approaches in resolving complex aerodynamic noise generation mechanisms. These studies indicate that CFD-based methods can improve aerodynamic noise prediction accuracy and provide valuable tools for engineering design optimization and acoustic performance evaluation. However, these methods mainly focus on aerodynamic noise sources and transmission paths. Improving vehicle interior acoustic performance also requires effective noise control strategies, especially the development of advanced acoustic materials and structures. Compared with conventional porous and resonant noise reduction structures, artificially designed acoustic metamaterials can manipulate sub-wavelength acoustic waves through precise structural topology optimization and achieve broadband, low-frequency, and tunable noise reduction within limited structural thickness and weight. In addition, architected materials with non-integer dimensional configurations can realize the coupled improvement of acoustic attenuation, mechanical load capacity, and fluid permeability, providing new solutions to the conflict between lightweight design, ventilation requirements, and broadband noise control in vehicle engineering [25, 26]. These studies indicate that current wind noise simulation techniques have significantly improved the accuracy of aerodynamic noise prediction, while emerging acoustic metamaterial designs offer promising passive control solutions, together providing reliable tools and strategies for early vehicle design optimization and NVH performance assessment. Although CFD-based aeroacoustic approaches provide detailed insights into aerodynamic noise generation and transmission mechanisms, their application in large-scale design exploration during the early vehicle development stage remains challenging due to the high computational cost, complex model preparation, and long simulation cycles required for each design iteration [27].

CFD has been widely used in the field of automotive aerodynamics to study the generation and propagation mechanism of aerodynamic noise. It can provide detailed physical insights into the external flow field and sound source, but its application in large-scale design evaluation is often limited by high computational cost and complex model setting. In order to improve the prediction efficiency, data-driven methods have attracted more and more attention in aerodynamic and acoustic analysis in recent years [28–31]. Yang et al. [32] developed a convolutional neural network (CNN) based aerodynamic noise prediction model for wind turbines based on a large-scale airfoil dataset, demonstrating its great potential in migration applications. Roznowicz et al. [33] proposed a large-scale three-dimensional external flow field approximation model based on graph machine learning. The model uses graph neural networks to understand links in complex flow fields. So it predicts vehicle aerodynamic flow quickly. In addition, Hou et al. [34], Zhao et al. [35] and Dai et al. [36] extended the application of data-driven methods in acoustic research by using a variety of analysis methods based on machine learning and data fusion. However, the existing data-driven research mainly focuses on the prediction of aerodynamic flow field characteristics, external noise sources or acoustic indicators, while there are still few studies on the direct prediction of interior wind noise spectrum for vehicle design parameters. In particular, there is a lack of methods for simultaneously integrating vehicle styling features and acoustic design parameters.

The above research shows that significant progress has been made in the study of vehicle aerodynamic noise through experimental testing, numerical simulation and data-driven prediction. However, there are still some limitations. Although the experimental method can provide accurate noise response data, it usually requires physical prototypes and complex test procedures, which limits its application in rapid evaluation of various design schemes in the early stage of vehicle development. The numerical simulation method can analyze the physical mechanism in detail, but it relies on high-fidelity model construction and complex boundary condition setting, resulting in high computational cost. The existing deep learning methods mainly focus on aerodynamic flow field prediction or general acoustic regression tasks, while the research on predicting vehicle interior wind noise directly from vehicle modeling parameters and acoustic design schemes is still limited. Therefore, it is necessary to develop a data-driven prediction method that can effectively describe the nonlinear relationship between design variables and wind noise response by combining vehicle design features and acoustic technical parameters. These methods support fast wind noise assessment and optimization during early vehicle design. In this study, a vehicle interior wind noise prediction method based on CNN-Transformer is proposed. This method integrates convolutional networks and transformers to capture both local feature patterns and global feature dependencies. It enables direct prediction of vehicle interior wind noise spectra from vehicle design parameters and provides effective support for vehicle styling optimization and acoustic performance improvement. Main contributions of this paper are as follows:

(1) A CNN–Transformer-based vehicle interior noise prediction model is proposed in this study. The model takes vehicle styling features and technical parameters as inputs and predicts the interior wind noise spectrum measured near the driver's

left ear. Unlike existing studies that mainly focus on aerodynamic flow field prediction or acoustic response prediction, this study establishes a direct mapping relationship between early-stage vehicle design variables and the driver's left-ear interior wind noise spectrum.

(2) By integrating vehicle styling features with acoustic technical parameters, an interpretable sensitivity analysis framework based on the mean impact value (MIV) method is developed to investigate the relationship between input parameters and interior wind noise levels at the driver's left-ear position. The obtained sensitivity results provide quantitative guidance for subsequent noise control and structural optimization and help reduce the time required for vehicle acoustic performance development.

This paper is organized as follows. Section 2 explains the Transformer. It also introduces the CNN-Transformer hybrid network built by combining CNN and Transformer. In Section 3, the feature variables involved in this study are first defined. An experimental design is then constructed using optimal Latin hypercube sampling (OLHS), followed by the acquisition of noise samples under different styling configurations through full-scale wind tunnel tests. Section 4 introduces the modeling strategy for driver's left-ear wind noise and develops a CNN–Transformer-based interior wind noise prediction model. The prediction performance of the proposed model is then evaluated through qualitative and quantitative analyses. In addition, uncertainty analysis based on Monte Carlo Dropout and input parameter perturbation is introduced to assess the prediction reliability and robustness of the proposed model. Section 5 compares the proposed model with conventional CNN, transformer, and long short-term memory (LSTM) networks, and further evaluates its generalization ability and prediction stability using an independent test dataset. In addition, the MIV method is introduced to quantitatively analyze the relationships between vehicle styling parameters, engineering configurations, and vehicle wind noise. Section 6 summarizes all the research work presented in the entire text. **Figure 1** shows the framework of this study.

2. The proposed method

2.1. Introduction of transformer model

The transformer, as a pioneering deep learning framework, employs a self-attention mechanism that directly captures global dependencies within sequences, effectively addressing the issues of gradient vanishing and long-range information loss commonly observed in recurrent neural networks [37, 38]. Meanwhile, the parallel computation capability of the transformer improves training efficiency and allows the model to be trained on workstations with limited computational resources. This feature makes the transformer suitable for practical engineering applications. Therefore, the transformer has shown strong performance in complex data modeling tasks in engineering fields [39].

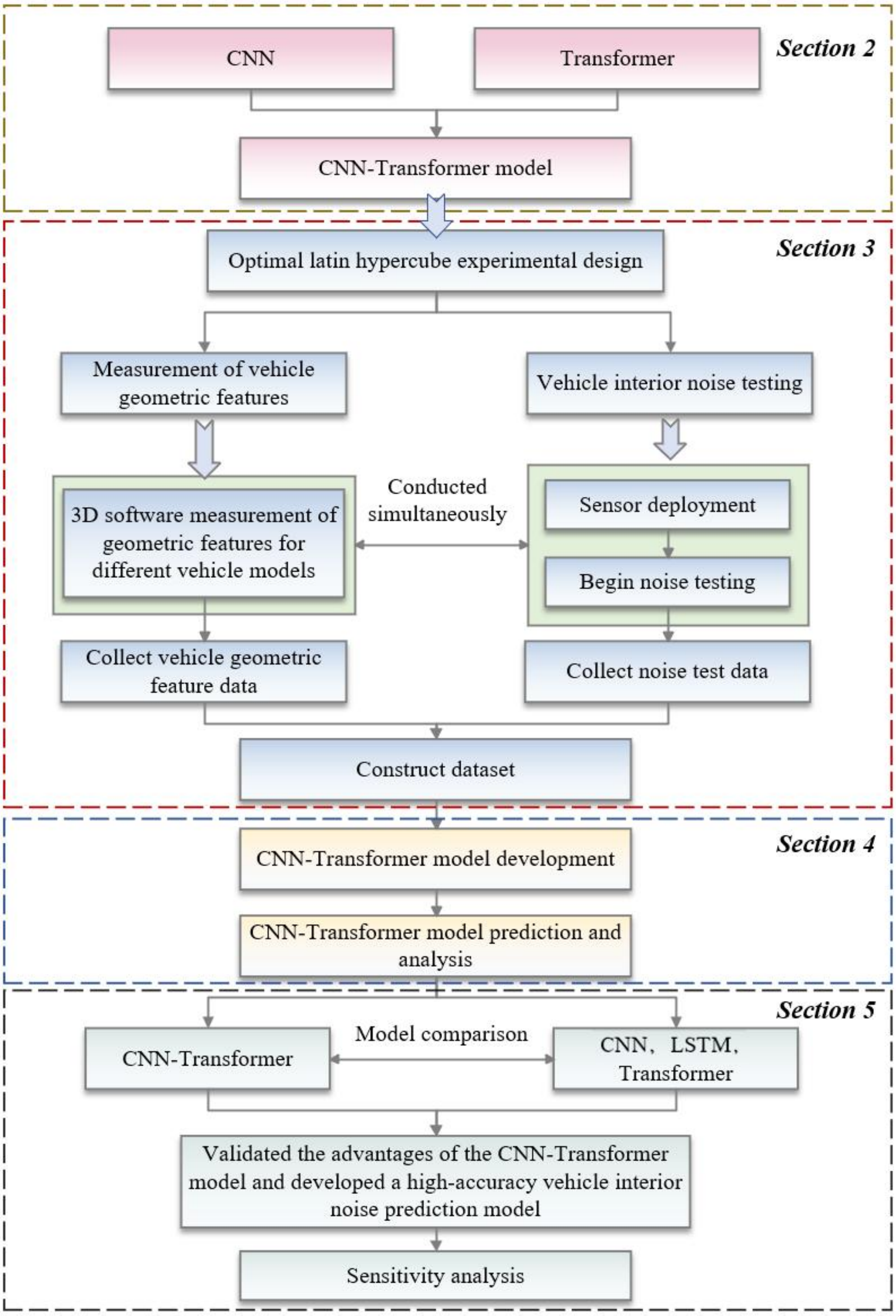


Figure 1. Driver's left ear noise prediction method process.

The transformer model contains two functional parts, including the encoder and decoder. The encoder maps the input information into high-level feature representations, and the decoder generates the output based on these learned features [40, 41]. The complete structure is illustrated in **Figure 2**. Multi-head attention serves as the main mechanism that enables feature interaction within the transformer. It performs parallel attention calculations through multiple independent attention heads. Each attention head learns different feature relationships from the input data, which helps capture various correlations in sequential signals [42]. Based on the attention mechanism, this module enables the network to capture complex interactions among input features and improves its ability to model long-range dependencies. Moreover, stacking numerous scaled dot-product attention layers lets the network carry out concurrent training for multi-dimensional signals and simultaneous attention weight calculation, which greatly boosts the model's power to express complex feature patterns [43].

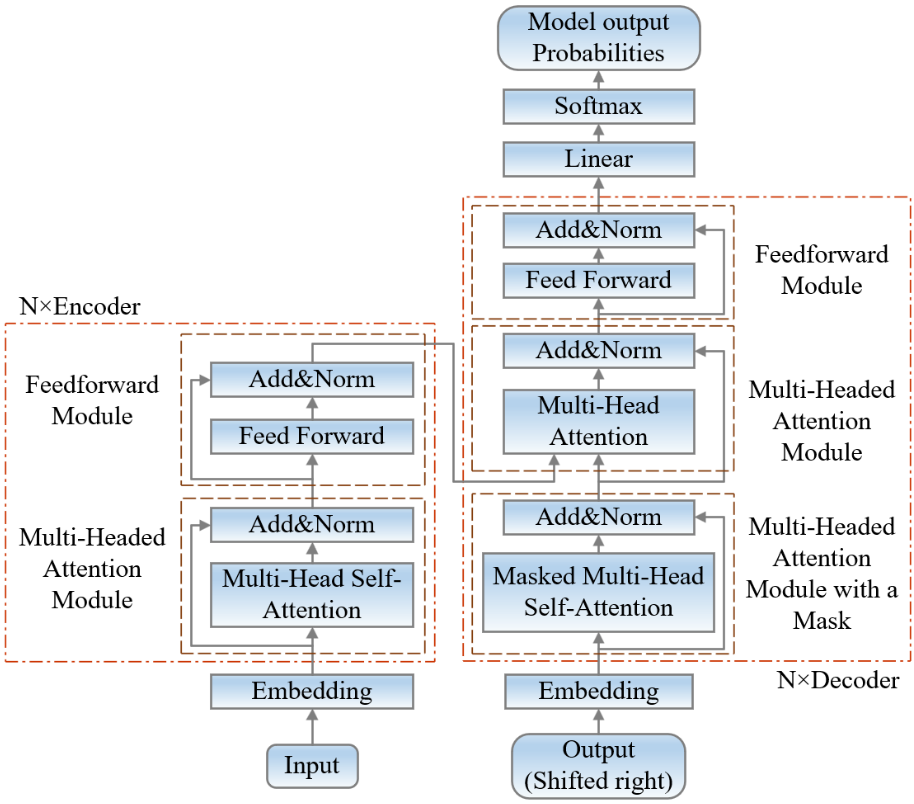


Figure 2. Architecture of a transformer.

To elaborate, linear projection operations are adopted to convert raw input features into three distinct feature vectors for subsequent calculation. Scaled dot-product operations are then applied to Q, K, and V to perform multi-head attention in parallel. The similarity between queries and keys is calculated to obtain attention weights for aggregating value representations. The results generated by all attention heads are then merged and forwarded to the subsequent module for further processing, as shown in Equations (1)–(3) [44].

$$\text{MultiHead}(Q, K, V) = \text{Concat}\left(\text{Head}_1, \text{Head}_2, ..., \text{Head}_h\right) W^O, i = 1, 2, ..., h \quad (1)$$

$$\text{Head}_i = \text{Attention}\left(QW_i{}^Q, KW_i{}^K, VW_i{}^V\right) \quad (2)$$

$$\text{Attention}(Q, K, V) = \text{soft max}\left(\frac{QK^T}{\sqrt{d_k}}\right) V \quad (3)$$

Where, Q, K, and V are obtained by linearly transforming the input vectors through three trainable projection matrices W_i^Q, W_i^K, and W_i^V, respectively. The dimensions of these projection matrices are all $d_{model} \times d_k$. The final output is further linearly transformed using the projection matrix W^O, whose dimension is $h \times d_v \times d_{model}$. Here, d_{model} refers to the size of the feature embedding space, d_k and d_v indicate the feature sizes used for attention calculation, and h represents the number of attention heads. Concat $(\cdot)$ denotes the matrix concatenation operation used to merge the outputs of multiple attention heads while maintaining dimensional consistency. The softmax $(\cdot)$

function normalizes attention weight values into a group of probabilistic weights [45]. **Figure 3** illustrates the complete calculation process.

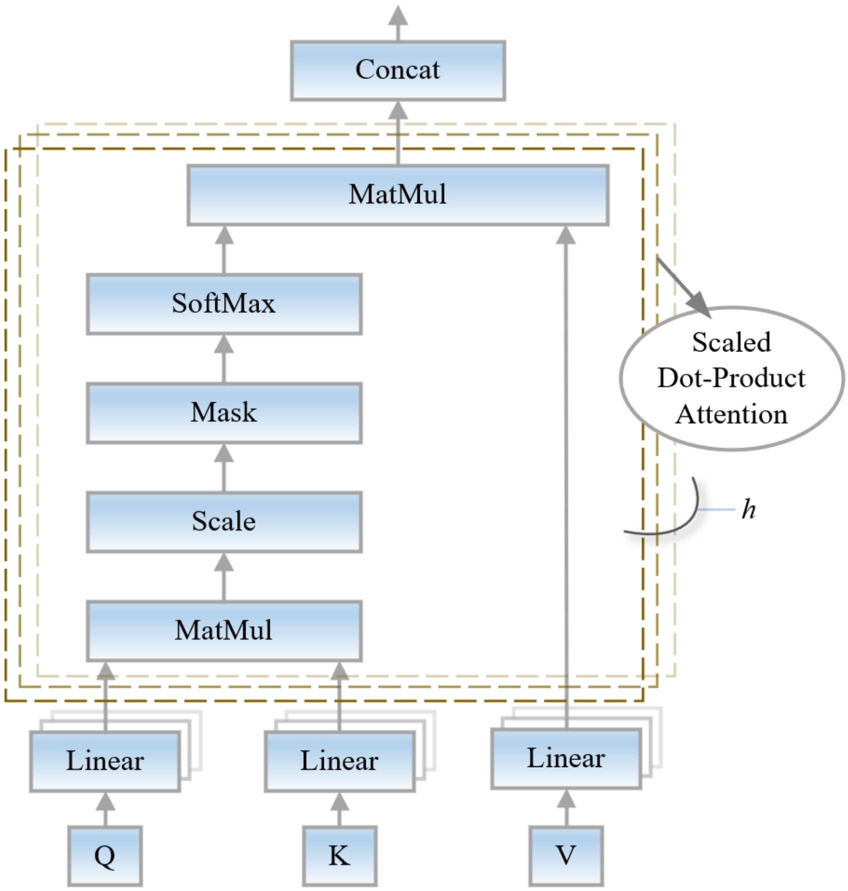


Figure 3. Multihead attention module.

In vehicle wind noise prediction tasks, although the 23-dimensional input feature vector is relatively compact, strong coupling relationships exist among the individual features. A standalone transformer model can directly treat these features as a sequence and capture their global dependencies through the self-attention mechanism [46,47]. However, the model exhibits limited capability in capturing fine-grained local interactions, which provides motivation for introducing a CNN module within a hybrid architecture to enhance performance [48].

2.2. Proposal of the CNN-transformer method

CNNs offer significant advantages in local feature extraction and are particularly well-suited for processing input data with sequential or dimensional correlations [49]. One-dimensional convolutional neural networks (1D-CNN) perform weighted summation within local regions of the input features through sliding convolution kernels. The operation of a single convolutional layer can be expressed as [50]:

$$y_j = \sigma\left(\sum\nolimits_i x_i \cdot w_{ji} + b_j\right) \tag{4}$$

where x_i is the input feature, w_{ji} represents the convolution kernel weight, b_j is the bias, and σ denotes the activation function.

The 1D CNN module applies convolution operations, batch normalization, and nonlinear activation functions to learn local feature representations. It can capture short-range dependencies among neighboring feature dimensions. The feature transformation process is expressed as follows [51]:

$$X_{out} = f_{CNN}\left(X_{in}\right) \in R^{B \times C_{out} \times L} \tag{5}$$

where f_{CNN} represents the mapping function of the 1D CNN module, and X_{in} refers to the feature tensor provided to the network. B represents the sample count within each training batch, C_{out} specifies the dimension of the generated feature maps; and L represents the length of the feature dimension.

The local receptive field of the convolution operation enables the module to capture interactions between neighboring feature dimensions and extract important local information. Batch normalization is also applied to improve training stability and accelerate convergence. Its calculation is defined as:

$$\widehat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \cdot \gamma + \beta \tag{6}$$

where, μ_B and σ_B denote the statistical average and variation of the samples within a batch, respectively. ε is a small value introduced to maintain numerical stability during calculation. γ and β are trainable coefficients that control the scaling and offset of the normalized features. Batch normalization normalizes the feature distribution within each mini-batch, reduces the influence of distribution variation during training, and improves model stability.

In vehicle wind noise prediction tasks, adjacent dimensions of the input vectors, including vehicle styling features and technical parameters, usually show certain correlations. A 1D-CNN can capture these local feature relationships through convolutional layers, batch normalization, and activation functions, and perform initial feature extraction and dimensionality reduction [52]. However, a single CNN has a limited receptive field and cannot fully capture long-range dependencies among input features. So this study builds a CNN-Transformer hybrid network. It combines the strengths of CNN and Transformer for vehicle wind noise prediction. The model first extracts local feature interactions through the CNN module, and then feeds the resulting feature sequence into a transformer encoder for global dependency modeling. Unlike approaches that simply stack complex attention modules, the CNN-Transformer architecture achieves natural synergy between CNN and transformer components, requiring no substantial increase in additional parameters while effectively alleviating gradient vanishing issues in deep network training. Under small and medium-sized datasets, the inductive bias introduced by the CNN module helps reduce the risk of overfitting, while the transformer provides strong nonlinear global modeling capability, leading to improved prediction accuracy and generalization performance [53,54]. The model architecture is shown in **Figure 4**.

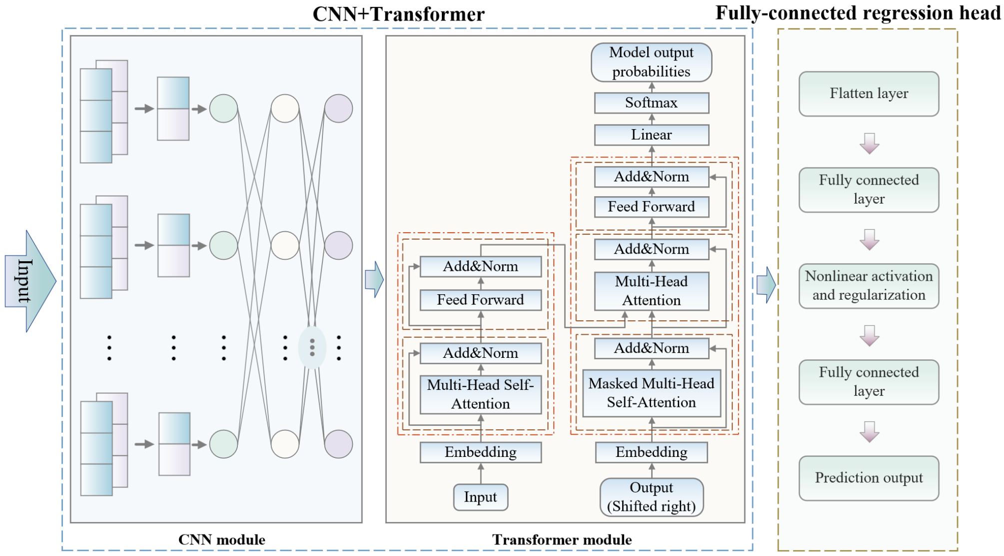


Figure 4. CNN-Transformer overall architecture.

3. Data collection

3.1. Experimental design

To develop a high-precision nonlinear mapping model relating vehicle styling parameters and material acoustic properties to interior wind noise response, a sufficient number of passenger car test samples is required. However, wind noise testing necessitates the replacement of specific glass samples, making large-scale real-vehicle experiments extremely difficult to conduct. Therefore, the OLHS method was adopted for experimental design [55].

The present study did not rely on local parameter refinement of a single vehicle model. Instead, differentiated baseline flow fields were established using 16 baseline sedans representing diverse design languages. On this basis, additional glass material and thickness configurations were generated for each baseline vehicle. The styling parameters of the 16 baseline vehicles, including the A-pillar dimensions in the X and Y directions, the three-dimensional dimensions of the side mirrors, the windscreen inclination angle, spacing between side mirrors and vehicle body panels, and the angle of inclination of the left glass, were defined as global independent variables and simultaneously used as inputs to the prediction model. Consequently, the model learns not only the influence of individual parameter variations but also the aerodynamic acoustic response characteristics associated with different vehicle body contours. For each baseline vehicle, the material properties and thicknesses of the glazing components were parametrically recombined. **Table 1** lists the glass material categories and their corresponding thickness configurations.

Table 1. Types of glass materials and corresponding thicknesses.

Parameters	Design value
Glass material types	tempered glass, acoustic laminated glass
Tempered glass thickness types	3.2 mm, 3.5 mm, 4 mm
Acoustic laminated glass thickness types	4.36 mm, 4.76 mm, 4.96 mm, 5.36 mm, 5.76 mm

Specifically, to achieve maximum coverage of the 23-dimensional feature space within a limited sample size, this study employs OLHS for cross-vehicle scheme allocation. The styling parameters of the 16 benchmark vehicles are treated as mixed inputs, ensuring that each vehicle has a representative distribution of design schemes across the 240 operating conditions. This method reduces the risk of model bias caused by excessive dependence on the styling characteristics of a specific vehicle, and improves the capability of the model to characterize the acoustic insulation behavior of different materials. The final 240 × 23 experimental matrix contains measurement data from 16 vehicles with various glazing configurations. The generated matrix is presented in **Figure 5**, where X1 to X5 correspond to the windshield, front door glass, rear door glass, rear windshield, and roof glass, respectively.

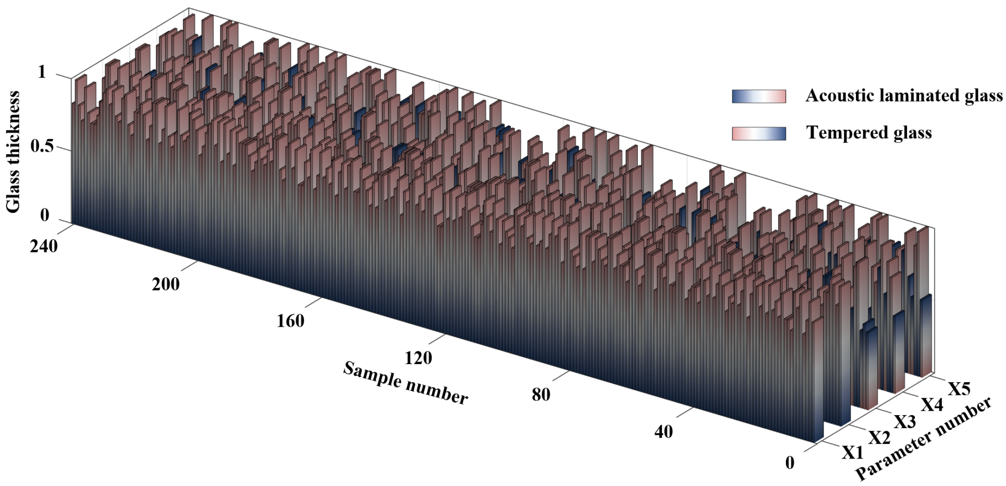


Figure 5. Schematic diagram of the experimental design matrix generated based on OLHS.

3.2. Interior wind noise testing

3.2.1. Test platform and methodology

Automotive wind tunnels are dedicated test devices that replicate real-world driving airflow fields surrounding automobiles, and they find extensive applications in investigations of automotive aerodynamic characteristics and aerodynamic noise behaviors [56]. According to functional positioning, automotive airflow test chambers fall into two main categories: aeroacoustic wind tunnels and environmental wind tunnels. Of the two types, aeroacoustic wind tunnels are purpose-built to quantify and judge internal and external flow-induced noise, acting as a vital test bench for the conceptual design and performance improvement of automotive wind noise [57]. All tests used an aeroacoustic wind tunnel. Its layout appears in **Figure 6**.

The measurement system was based on the testing and analysis platform developed by HEAD Acoustics, Germany. The test object is a passenger vehicle produced by China FAW Group Co., Ltd. The overall test procedure is illustrated in **Figure 7**. After all equipment and wind tunnel operating parameters are properly adjusted, instrument calibration is performed to ensure sufficient measurement accuracy of the acquired data. During the tests, the airflow speed is maintained at 120 km/h without lateral flow interference. Once the in tunnel flow field reaches a stable state, acoustic signals are continuously recorded for 15 s at a sampling frequency of 48 kHz. Each test condition is repeated multiple times, with a minimum of three repetitions, to verify the consistency and stability of the acquired data.

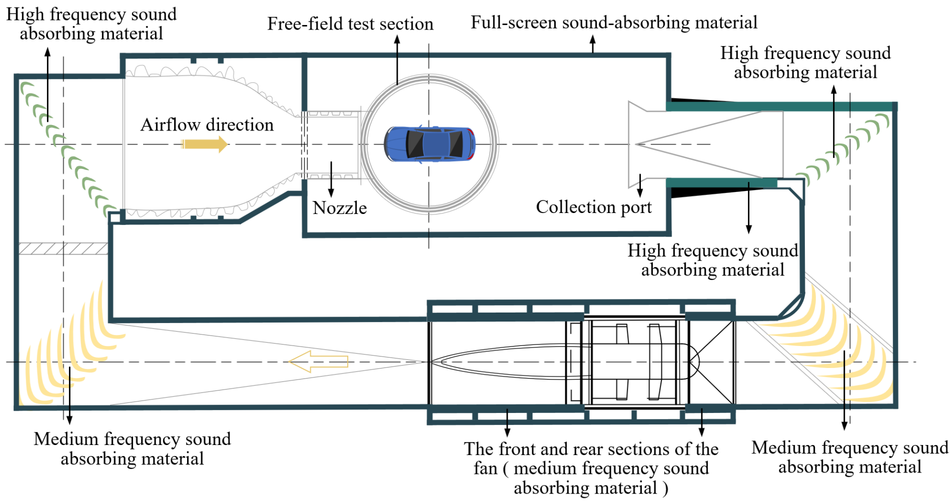


Figure 6. Schematic layout of the aeroacoustic wind tunnel.

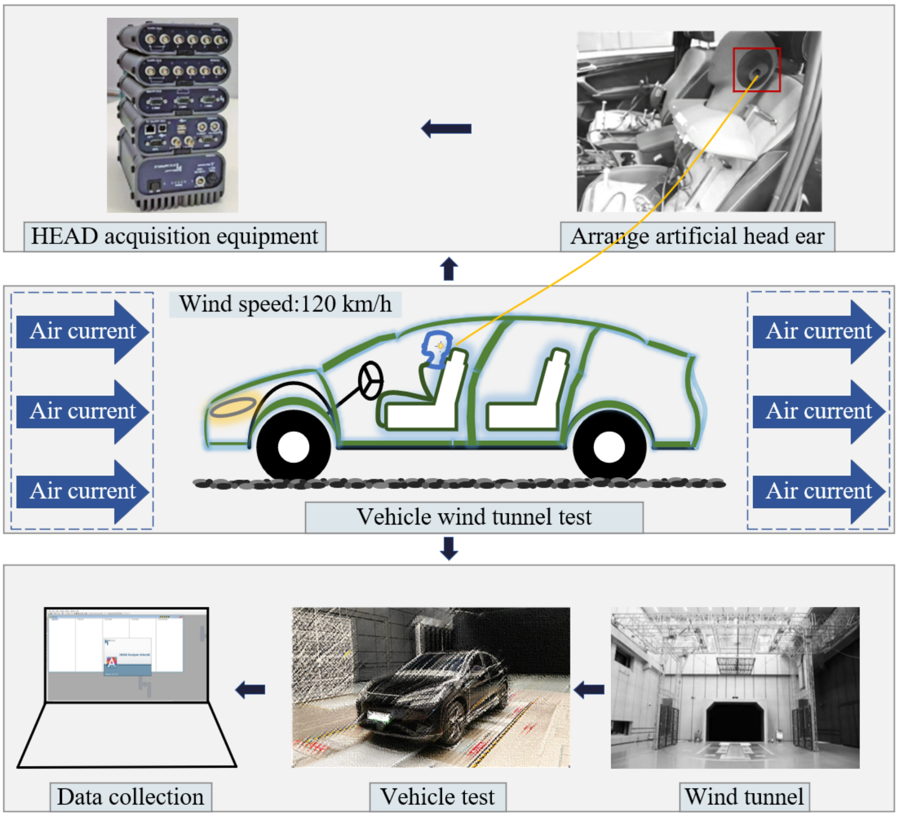


Figure 7. Driver's left ear noise test.

3.2.2. Vehicle noise test results

The data analysis stage considers only acoustic signals within the frequency band from 200 to 8,000 Hz. All spectral curves are processed using A weighting, and loudness is computed following ISO 532-1 [58]. After completing all experimental measurements, noise spectra from multiple randomly selected vehicles are presented, as shown in **Figure 8**. The spectral trends indicate the absence of abnormal peak values, confirming that the tested vehicles were in normal operating condition.

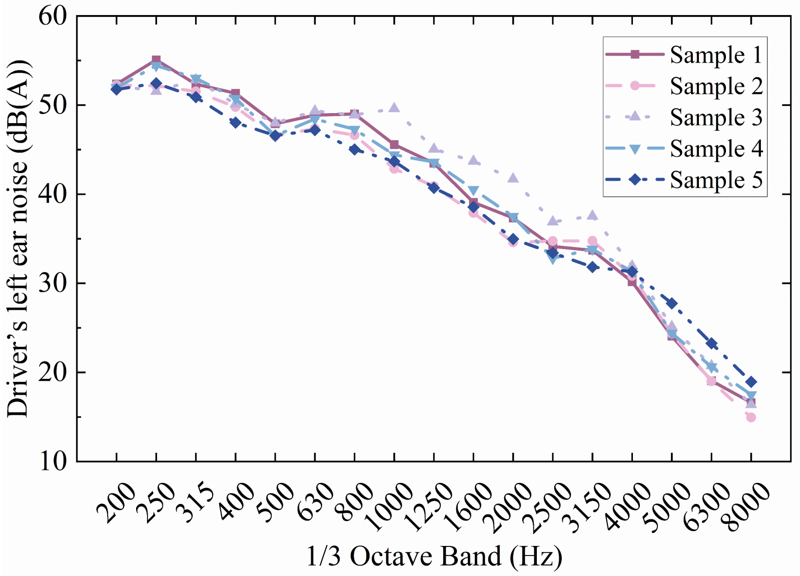


Figure 8. Test results of driver's left ear noise.

To quantitatively evaluate the repeatability of the experimental measurements, three repeated test datasets from a single sample vehicle were randomly selected for analysis. As shown in **Figure 9**, the spectral curves corresponding to the three test conditions exhibit a high degree of overlap, fully confirming that the experiment has excellent repeatability and that the measured acoustic data are stable and reliable.

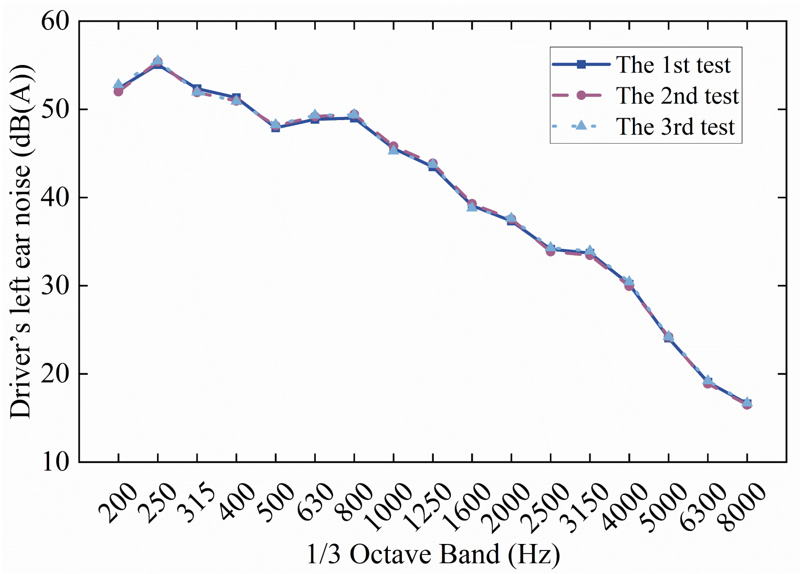


Figure 9. Comparison of driver's left ear noise test results on the same vehicle.

4. Interior wind noise prediction at the driver's position based on CNN-transformer

4.1. Development of the CNN-transformer model

To address the challenges of high cost and long duration in full vehicle wind tunnel testing, as well as the high computational demand of numerical simulations and the discrepancies between simulation results and real-world measurements, this study proposes an efficient method for in-cabin noise prediction. The proposed framework builds a prediction network based on acoustic samples acquired from wind tunnel experiments, where vehicle styling parameters and glazing technical schemes are used as inputs, and the model output corresponds to vehicle wind noise levels. The model input contains two types of variables: vehicle styling parameters and technical parameters. The vehicle styling parameters mainly describe key geometric features that affect aerodynamic noise generation. The technical parameters describe the engineering factors related to wind noise propagation, structural vibration and acoustic attenuation. The model outputs the 1/3 octave band noise pressure level at the driver's left ear.

The input styling features include the vehicle geometric parameters defined in Section 3. The A-pillar dimensions influence the locations of flow detachment and vortex generation, while the side mirror dimensions affect the intensity of wake turbulence. The front windshield angle, the gap between the side mirror and vehicle surface, and the side window inclination further characterize the airflow disturbance and separation behavior [59]. The technical parameters include the material properties and thickness of window glasses, the acoustic transfer function, and the average interior reverberation time. The glazing material and thickness affect acoustic transmission loss and structural vibration response, and they play an important role in controlling low- and medium-frequency noise transmission [60]. The acoustic transfer function and interior reverberation time were introduced to characterize the overall acoustic transmission characteristics and cabin acoustic environment of the vehicle. The acoustic transfer function quantifies the attenuation characteristics of noise transmitted from the vehicle exterior into the cabin through the glazing and body structure. The average interior reverberation time describes the sound diffusion and decay characteristics inside the cabin and mainly affects the perceived high frequency noise level.

The dataset was constructed by obtaining vehicle styling features through three-dimensional software measurements, matching the technical schemes directly according to the design configurations presented in Section 3, and acquiring in-vehicle noise data from wind tunnel tests. The acoustic transfer function and reverberation time were independently measured in a reverberation chamber before the wind tunnel experiments. The acoustic transfer function was obtained by applying a controlled sound excitation outside the vehicle and recording the sound pressure response at the driver's left-ear position, while the reverberation time was calculated based on the sound decay characteristics after excitation. These parameters represent the inherent acoustic transmission and sound field characteristics of the vehicle cabin rather than the aerodynamic noise response itself. A CNN-Transformer hybrid neural network was employed for modeling. The synergistic architecture of 1D-CNN and transformer encoder was utilized to optimize the model, aiming to minimize prediction error and enhance accuracy and generalization performance in vehicle wind noise prediction.

Figure 10 shows the overall network structure of this prediction model. Considering the complex nonlinear coupling relationship of vehicle wind noise prediction, this paper adopts a hybrid network architecture with deep integration of 1D CNN and transformer encoder. The input features are first sent to the 1D-CNN module to complete local feature extraction and noise filtering; then, the features are mapped to a 64-dimensional hidden space through the dimension projection layer, and the dimension arrangement is adjusted to meet the sequence input requirements of transformer. The network core is built by a single-layer transformer encoder module, and the global coupling correlation and nonlinear action weight between various body physical features are mined by means of a four-head multi-head self-attention mechanism. After entering the regression prediction stage, the feature tensor output by the encoder is flattened, and then the end-to-end regression mapping from 23-dimensional input features to 17-dimensional 1/3 octave wind noise results is realized through two layers of fully connected layers. Considering the limited dataset size, the Adam optimizer with a weight decay of 1×10^{-2} and the ReduceLROnPlateau adaptive learning rate strategy were adopted to improve model robustness and achieve stable convergence during training. The main parameter settings of the CNN-Transformer network are shown in **Table 2**.

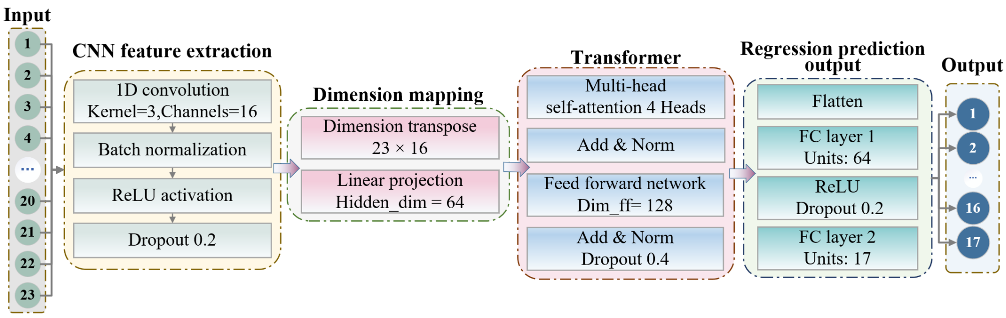


Figure 10. Overall architecture of the prediction model for driver's left ear noise.

Table 2. CNN-Transformer prediction model parameters.

Model structures	Parameter types	Value
Input layer	Input feature dimension	23
CNN layer	Convolutional layer channels Kernel size Activation function	16 3 ReLU
Transformer layer	Hidden layer dimension Number of self-attention heads Number of encoder layers Feed-forward network dimension	64 4 1 128
Output layer	Output dimension	17
Training configuration	Optimization algorithm Learning rate Weight decay Batch size	Adam 5×10^{-5} 1×10^{-2} 64

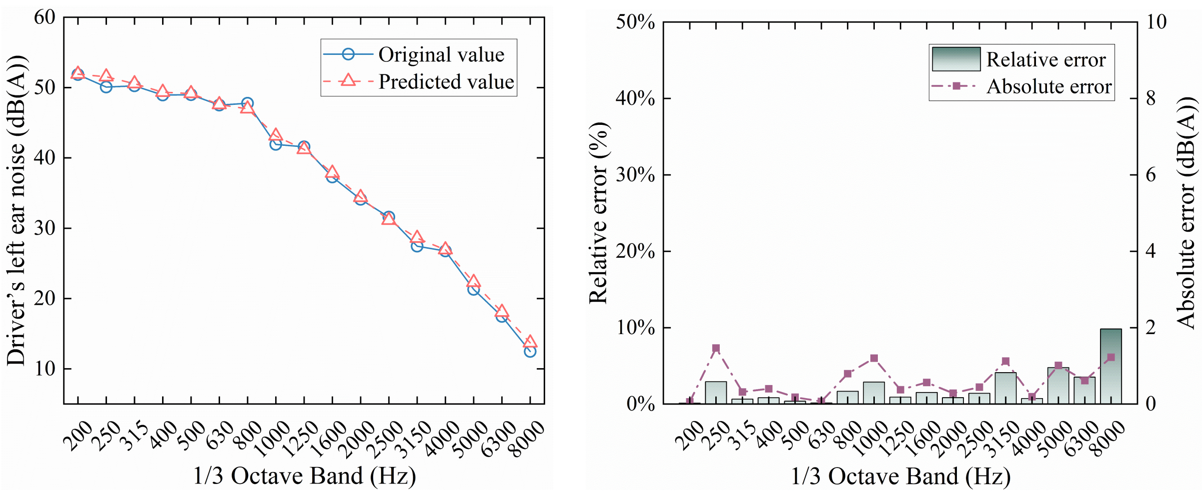
The dataset was constructed from wind tunnel experimental measurements and contains 240 independent vehicle operating condition samples. Each sample is defined by a 23-dimensional input feature space, including 10 vehicle styling geometric parameters and 13 glazing-related technical parameters. The driver's left-ear interior wind noise spectrum was used as the regression output, covering 1/3 octave bands from 200–8,000 Hz. Since the input variables had different magnitudes and units, feature scaling was applied during data preprocessing. Each feature was linearly transformed into the interval of 0–1 to reduce the effect of scale differences on gradient optimization [61].

To obtain an unbiased and reliable evaluation of model performance, a strict five-fold cross-validation strategy was adopted. The complete data set containing 240 sets of samples was randomly disrupted and divided into five non-overlapping subsets. In each iteration, 80% of the samples are randomly chosen for model training, and the other 20% serve as the validation set. The training and validation sets have no common samples in any round. To assess the model, we use the root mean square error (RMSE) and mean absolute percentage error (MAPE) from five independent validation rounds instead of a single leave-out test. This strategy reduces the deviation caused by random data partitioning, thus providing a more reliable model performance evaluation.

All network construction and model training were completed in the PyCharm 2025.3 environment. The computing platform is configured with an Intel Core i5-12600 KF processor and 64 GB of memory. In order to ensure a fair comparison between different models, all networks are trained under the same conditions, and the Adam optimizer is used. Model training uses a learning rate of 1×10^{-4}, a weight decay of 1×10^{-2}, and a batch size of 64. The mean square error (MSE) serves as the loss function. Additionally, the same early stopping strategy optimizes the parameters.

4.2. Prediction analysis using the CNN-transformer model

The trained CNN-Transformer hybrid model was applied to forecast the interior wind noise at the driver's left ear. Independent validation data verified the model. In **Figure 11**, the predicted noise spectrum for a representative vehicle is compared against wind tunnel test results. The high agreement between the predicted and experimental frequency responses proves the accuracy of the model in tracking noise changes. To quantify the errors, MAPE and RMSE values were computed. The validation samples yield a MAPE of 2.23% and an RMSE of 0.94 dB. The model is able to learn the complex nonlinear interactions between vehicle styling parameters, glazing technical schemes, and interior wind noise responses.

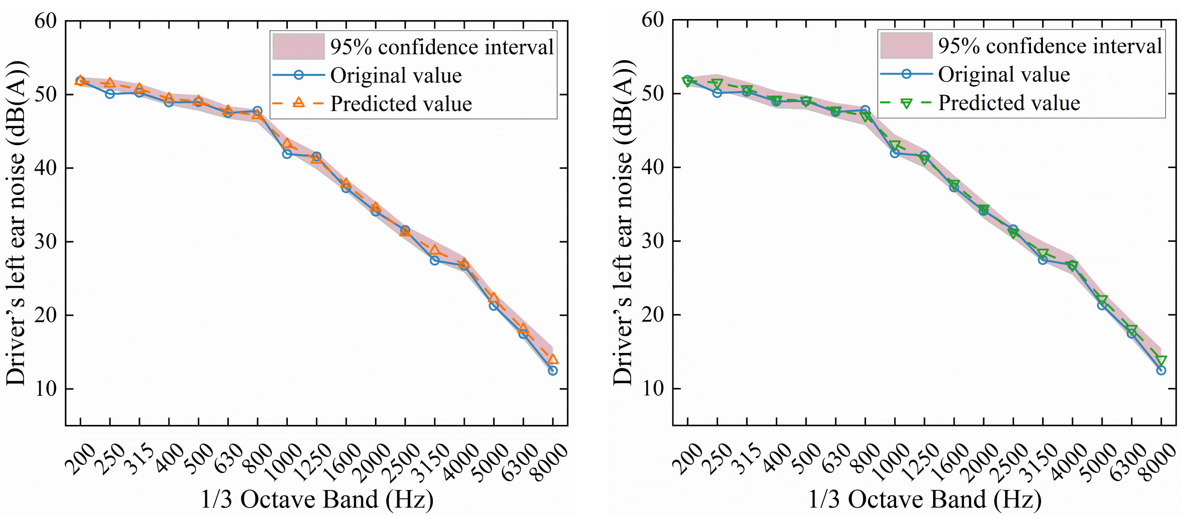


(a) Prediction results of driver's left ear noise. **(b)** Prediction errors.

Figure 11. Prediction results and errors of driver's left ear noise using the CNN-Transformer model.

To further evaluate the reliability of the CNN–Transformer model, prediction uncertainty analysis was conducted from two aspects: model uncertainty and input parameter uncertainty. First, the Monte Carlo Dropout method was applied. With the input sample unchanged, 100 stochastic forward passes were performed to obtain the distribution of prediction results, and the 95% prediction interval was calculated to evaluate the uncertainty caused by model parameters. Second, considering that deviations may exist between the actual values and design targets of vehicle styling parameters and acoustic technical parameters due to manufacturing and assembly processes, random perturbations ranging from 0% to ±5% were applied to the continuous input variables of the selected sample. A total of 100 perturbed samples were generated and predicted using the trained model. The 95% prediction interval was then calculated based on the distribution of the predicted results to evaluate the robustness of the model against input parameter uncertainty.

As shown in **Figure 12a**, the predicted mean value obtained by the Monte Carlo method is in good agreement with the measured internal wind noise spectrum, and the prediction interval is relatively small, indicating that the model has low uncertainty in the current data distribution range. The prediction interval becomes slightly wider in the high-frequency region, which is mainly attributed to the more complex variation and stronger nonlinear characteristics of high-frequency wind noise responses. Furthermore, as shown in **Figure 12b**, the input parameter perturbation analysis shows that when small deviations caused by manufacturing and assembly uncertainties are introduced into the input parameters, the predicted results still maintain a consistent trend with the original noise spectrum, while the variation of the prediction interval remains limited. This indicates that the proposed model has good robustness against inevitable parameter uncertainties in practical engineering applications.



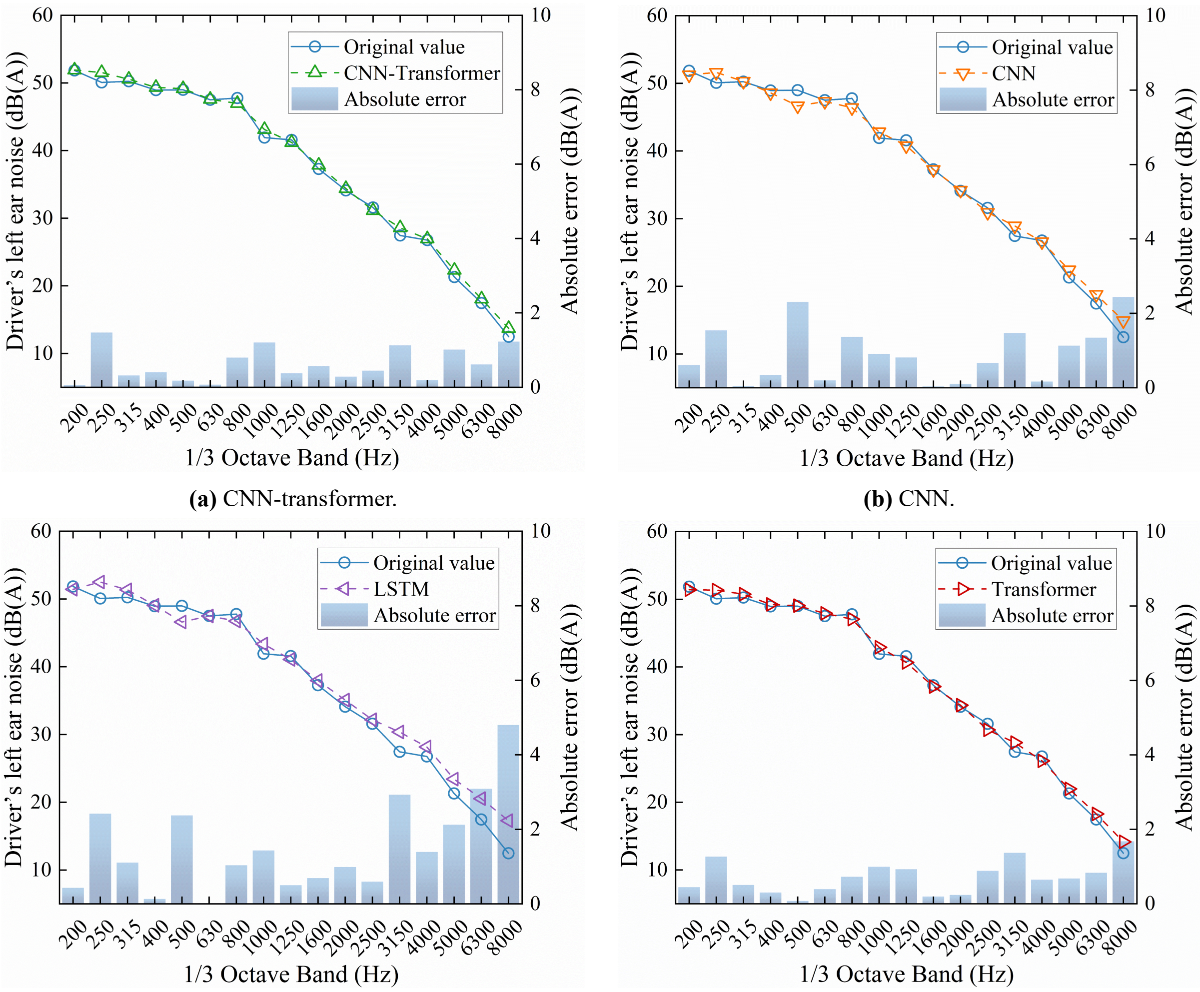
(a) Monte Carlo Dropout-based uncertainty estimation. **(b)** Input parameter perturbation-based uncertainty analysis.

Figure 12. Prediction uncertainty analysis of the CNN–Transformer model.

5. Model comparison and sensitivity analysis

5.1. Model performance comparison

The CNN-Transformer model was compared with three typical deep learning architectures, including CNN, LSTM, and transformer, to examine its capability in predicting driver-position interior wind noise. To ensure a fair comparison, we train and evaluate all models on the identical dataset. Each baseline model is built based on its typical architecture. The CNN extracts local features, the LSTM captures dependencies between them, and the Transformer learns global feature relationships. A validation sample was randomly chosen to test the four models, and **Figure 13** presents their predictions. In addition, the MAPE and RMSE values of each model on the validation set were calculated for quantitative evaluation, as summarized in **Table 3**.



(a) CNN-transformer. **(b)** CNN.

(c) LSTM. **(d)** Transformer.

Figure 13. Original vehicle driver's left-ear noise and prediction results with errors obtained using different models.

Table 3. MAPE and RMSE of the different models.

Model	MAPE	RMSE
CNN-Transformer	2.23%	0.94 dB
CNN	3.50%	1.45 dB
LSTM	4.42%	1.81 dB
Transformer	2.97%	1.22 dB

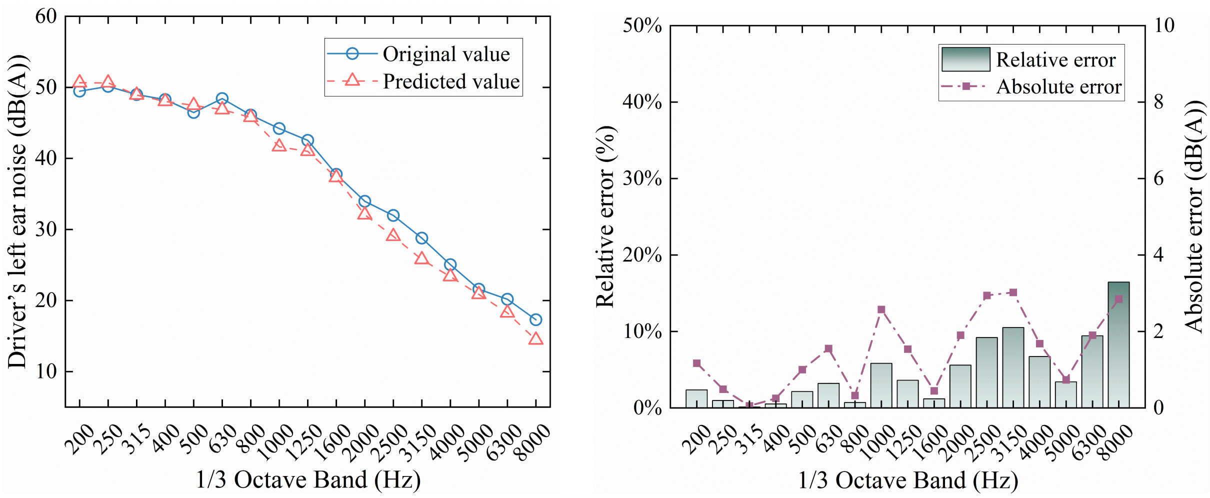
Validation results show that the CNN-Transformer model has the highest accuracy among all tested models, with the best scores in MAPE and RMSE indicators. Specifically, the MAPE of the model is 2.23%, and the RMSE is 0.94 dB. Compared with the CNN model, the MAPE and RMSE were reduced by 1.27% and 0.51 dB, respectively. Compared with the transformer model, these two indicators decreased by

1.28% and 0.28 dB, respectively. Compared with the LSTM model, the performance improvement is the most significant; the MAPE is decreased by 2.19%, and the RMSE is decreased by 0.87 dB.

In comparison, coupling convolution feature extraction with self-attention global modeling is better for modeling the complex links between car parameters and cabin wind noise. The CNN model can extract local features from vehicle modeling parameters and acoustic technical parameters through convolution operation. However, the wind noise inside the vehicle is affected by many factors, including aerodynamic characteristics, structural acoustic conduction and acoustic performance of the compartment. The input parameters also involve complex nonlinear interactions. Therefore, a single CNN structure has limitations in capturing global correlations among different design variables. The LSTM model achieves average MAPE and RMSE values of 4.42% and 1.81 dB, respectively, showing relatively lower prediction performance. This is mainly because LSTM is designed for modeling data with sequential dependencies, while the input variables in this study are vehicle design parameters and the output is frequency-domain noise responses. There is no clear temporal order among these features, which limits the ability of the recurrent structure to represent the nonlinear relationship between vehicle design parameters and noise responses. The transformer model improves the modeling of global feature relationships through the self-attention mechanism, achieving average MAPE and RMSE values of 2.97% and 1.22 dB, respectively. However, without an explicit local feature extraction process, its ability to represent local feature variations is still limited under the condition of a relatively small dataset. In comparison, the CNN–Transformer model combines the local feature extraction capability of CNN and the global relationship modeling capability of transformer. The CNN module first extracts important features from vehicle styling and acoustic parameters, and the transformer module then learns the complex interactions among these features, resulting in more accurate prediction of interior wind noise responses.

Further analysis of **Figure 13** shows that the prediction performance of different models shows a certain frequency dependence, which is mainly related to the complexity of noise response in different frequency bands and the feature representation ability of each model. The CNN model maintains relatively stable prediction performance in most frequency bands. However, because it mainly relies on local convolution feature extraction, its ability to capture global relationships is limited for some frequency bands affected by multi-parameter interactions, resulting in an increase in prediction error. The LSTM model shows more obvious prediction fluctuations in different frequency ranges. This is because its cyclic structure is more suitable for data with time evolution characteristics, and the input variables in this study do not contain intrinsic sequence information. Therefore, LSTM cannot give full play to its sequence modeling ability, resulting in a large difference in prediction performance between different frequency bands. The transformer model can capture the global relationship between input variables through the self-attention mechanism, which is superior to CNN and LSTM in frequency consistency. However, its lack of clear local feature extraction ability limits its ability to characterize some complex spectral changes. In comparison, the CNN-Transformer model gives stable predictions across the full frequency range. It integrates local feature data with global relationships, so it captures the nonlinear mapping between vehicle parameters and broadband wind noise better.

In order to evaluate the reliability and robustness of the proposed model in practical engineering applications, independent data samples are used for verification after training the model under the same experimental conditions. **Figure 14** compares the predicted results with the measured vehicle interior wind noise spectrum. Error analysis shows that the model achieves a MAPE of 4.82% and a RMSE of 1.44 dB on unseen samples. The results show that the CNN-Transformer hybrid model has good generalization performance. The model can still maintain a high prediction accuracy on unseen samples outside the training data set, showing its potential in practical engineering applications.



(a) Prediction results of driver's left ear noise. **(b)** Prediction errors.

Figure 14. Predicted vehicle driver's left-ear noise and error results obtained using the CNN–Transformer model on independent test samples.

5.2. Sensitivity analysis

Different input parameters have different effects on vehicle interior wind noise prediction. In order to guide the optimization of vehicle modeling, this study uses the MIV method to calculate the feature importance of the model [62]. In the original 23-dimensional vehicle styling and technical feature space, glass material type was not included in the MIV analysis. Acoustic laminated glass has better sound insulation performance than conventional tempered glass in practical engineering applications. Glass material is also a binary categorical variable, and the two glass types are encoded as 1 and 2 in the dataset. This does not meet the requirements of the MIV method. The MIV method evaluates feature sensitivity by applying small continuous perturbations to input variables and calculating the changes in model outputs. Therefore, it is suitable for continuous numerical features. Since glass material has only two fixed discrete values, continuous perturbation cannot be applied, and the obtained MIV value would not have a clear physical meaning. Therefore, the glass material parameter was set as a fixed boundary condition and excluded from the sensitivity analysis.

MIV is an effective indicator for evaluating feature importance in neural networks. It applies controlled perturbations to the original input features after model training and measures the changes in model output. The formula is as follows:

$$\mathrm{MIV} = \frac{1}{N_S}\sum\nolimits_{i\,=\,1}^{N_S} y_i\left(X_{X_V\,+\,10\%}\right) - y_i\left(X_{X_V\,-\,10\%}\right) \tag{7}$$

where N_s represents the frequency range of the independent variable, X_V indicates the base value of the independent variable, $y_i\left(X_{X_V\,+\,10\%}\right)$ denotes the predicted result when the variable is increased by 10%, and $y_i\left(X_{X_V\,-\,10\%}\right)$ denotes the predicted result when decreased by 10%.

The absolute value of MIV is used to quantify the sensitivity of each input feature, while the sign of MIV shows whether the feature has a positive or negative effect on the predicted wind noise level. A sensitivity analysis model was established based on vehicle styling features and technical schemes to evaluate the extent to which these parameters are associated with in-cabin noise. The calculated MIV was normalized and transformed into the contribution proportion of each feature, and a corresponding heatmap was generated, as shown in **Figure 15**.

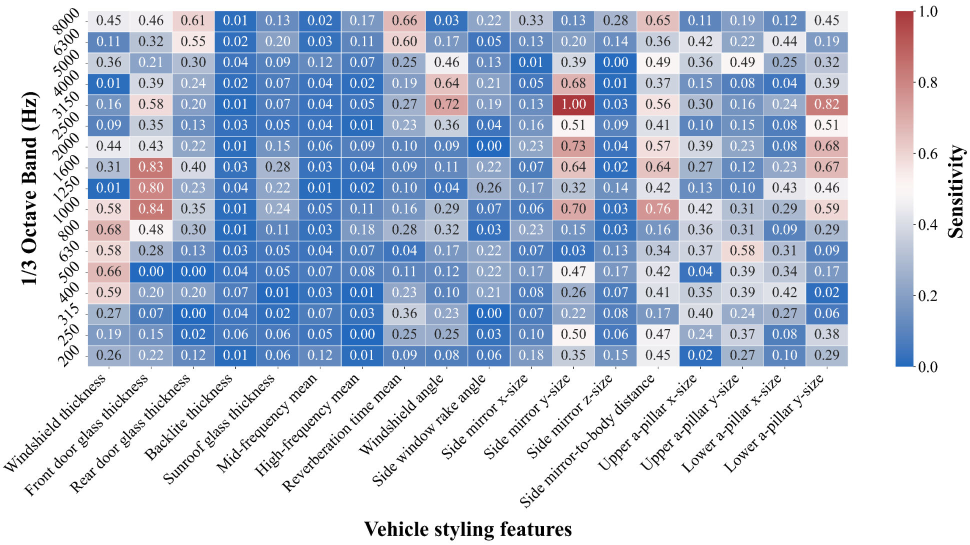


Figure 15. Full-frequency sensitivity analysis.

The following key patterns can be observed from the figure. The MIV sensitivity analysis shows that the absolute MIV values for the Y-direction dimensions of the side mirror and A-pillar are significantly higher than those of other variables, indicating that these two parameters are the dominant drivers of wind noise amplitude. A further comparison across different parameter categories reveals that the influence of side mirror and A-pillar dimensions is mainly distributed within the medium to high frequency region, whereas the sensitivities of front side-window and front windscreen glass thickness are relatively lower and primarily concentrated in the medium to low frequency bands.

These results are consistent with the basic aeroacoustic mechanisms. Under the 120 km/h high-speed condition, turbulent separation around the side mirror and A-pillar generates the main aerodynamic noise sources. Small changes in these geometric parameters directly affect the noise source strength. Glass thickness mainly affects the noise transmission path and changes the interior sound field by modifying the acoustic transfer characteristics. The front side window has higher sensitivity than the front windscreen because it is closer to the main region of A-pillar vortex impingement.

The sensitivity analysis based on the MIV method provides a quantitative foundation for subsequent noise control and structural optimization. In practical engineering design, priority should be given to the key parameters with the highest MIV rankings. Through targeted refinement of these parameters, more substantial noise reduction can be achieved with relatively lower design costs.

6. Conclusion

A CNN–Transformer hybrid prediction framework was developed in this study to predict aerodynamic wind noise at the driver's position in passenger vehicles. The proposed method can be used for driver-side vehicle interior wind noise prediction and provides support for vehicle acoustic optimization. The key findings of this study are as follows. The OLHS approach was applied to design combinations of vehicle styling and glazing acoustic parameters. Wind tunnel experiments were performed at the driver's left-ear position, and the measured spectra were used to build a dataset. Vehicle geometric features and glazing technical parameters were defined as input variables, while the corresponding interior wind noise spectra were treated as prediction targets. A CNN–Transformer hybrid model was developed to establish the nonlinear relationship between input parameters and vehicle interior wind noise responses. The CNN module extracts local geometric features, and the transformer encoder captures global feature relationships. The model can describe the nonlinear relationship among vehicle styling features, glazing configurations, and aerodynamic noise responses. The comparison results show that the proposed model achieves a MAPE of 2.23% and an RMSE of 0.94 dB on the validation dataset. The model also achieves a MAPE of 4.82% and an RMSE of 1.44 dB on independent test samples. Compared with CNN, LSTM and transformer models, the proposed model performs better in prediction accuracy and generalization performance. The MIV method is used to quantify the contribution of each input feature. The sensitivity ranking shows the influence of vehicle shape parameters and glass scheme on the interior wind noise. These results provide guidance for vehicle acoustic optimization and structural design improvement, and shorten the iteration cycle in vehicle NVH development.

The wind noise prediction model of vehicle interior based on CNN-Transformer developed in this study has good prediction accuracy and generalization performance. However, there are still some limitations. Firstly, the model is trained and validated using only 240 passenger car wind tunnel test samples. The size of the data set is limited, which affects the predictive stability of the model and styling configurations that are not included in the data set, and increases the predictive uncertainty. Secondly, the current input features only contain the selected vehicle styling parameters and glass-related technical parameters, and fail to cover all structural factors affecting the aerodynamic noise of vehicle interiors, resulting in prediction errors in some frequency bands. Future research requires more wind tunnel data from diverse car models. A larger dataset will then help the model better predict noise for alternative vehicle designs. In addition, more structural parameters related to aerodynamic noise can be added to the input feature set, such as detailed A-pillar geometric features, local body structure parameters, and sealing clearance. These parameters can provide more design variables for vehicle interior wind noise optimization. Since the added parameters may make the nonlinear mapping relationship more complex, future research should combine uncertainty analysis with physical-based optimization methods to improve the stability and prediction reliability of the models.

Author contributions: Conceptualization, HY and XH; methodology, HY; software, HY; validation, PL, JW and YD; formal analysis, JW; investigation, HY; resources, XH; data curation, HY, PL and JW; writing—original draft preparation, HY; writing—review and editing, HY, HH and XH; visualization, HY and PL; supervision, XH; project administration, HH; funding acquisition, YD. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Science and Technology Project of Sichuan Province (2026YFHZ0230), and the Science and Technology Innovation Platform Plan of Guangxi (LT2504240029).

Institutional review board statement: Not applicable.

Informed consent statement: Not applicable.

Data availability statement: The authors are not authorized to publicly share the original experimental data and source code. For researchers with relevant demands, the corresponding data and codes can be provided upon official authorization from the collaborative enterprise.

Acknowledgment: The authors would like to acknowledge the support from the Institute of Energy and Power Research for the experimental research.

Conflict of interest: The authors declare no conflict of interest. No AI tools were used for data analysis, interpretation, or generation of scientific content. All outputs were critically reviewed and edited by the authors. The authors take full responsibility for the integrity and accuracy of the work.

AI use statement: The authors declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

1. Pang J, Mao T, Jia W, et al. Prediction and Analysis of Vehicle Interior Road Noise Based on Mechanism and Data Series Modeling. Sound & Vibration. 2024; 58(1): 59–80. doi: 10.32604/sv.2024.046247
2. Gad AS, Darakhshan Jabeen S, El-Zomor HM, et al. Optimizing Ride Comfort and Stability in In-Wheel Motor Drive Electric Vehicles: A Correlated Semi-Active Suspension System with Kalman Filter and Spring-Type Analysis. SAE International Journal of Vehicle Dynamics, Stability, and NVH. 2026; 10(4). doi: 10.4271/10-10-04-0030
3. Wang T, Liu Z, Li X, et al. An end-to-end data-driven framework for intelligent evaluation of automotive noise and vibration performance. Engineering Applications of Artificial Intelligence. 2026; 181: 115422. doi: 10.1016/j.engappai.2026.115422

4. Zhang Y, He T, Wang Y, et al. Research on the Influence of Door and Window Sealing on Interior Wind Noise Based on Statistical Energy Analysis. SAE International Journal of Vehicle Dynamics, Stability, and NVH. 2025; 9(3): 423–437. doi: 10.4271/10-09-03-0026
5. Liu K, Liao Y, Wang H, et al. Road Noise Improvement Method for Battery Electric Vehicles Based on Bushing Stiffness Optimization. SAE International Journal of Vehicle Dynamics, Stability, and NVH. 2025; 9(4): 555–569. doi: 10.4271/10-09-04-0032
6. Song D, Hong S, Ha C, et al. Acoustic analysis and data-driven control of vehicle NVH: A framework for manufacturing process optimization. Applied Acoustics. 2025; 233: 110618. doi: 10.1016/j.apacoust.2025.110618
7. Yang M, Dai P, Yin Y, et al. Predicting and optimizing pure electric vehicle road noise via a locality-sensitive hashing transformer and interval analysis. ISA Transactions. 2025; 157: 556–572. doi: 10.1016/j.isatra.2024.11.059
8. Mahmood A, Cockerill T, de Boer G, et al. Heat transfer modeling and optimal thermal management of electric vehicle battery systems. Energies. 2024; 17(18): 4575. doi: 10.3390/en17184575
9. Wang Y, Du M, Su C, et al. Numerical investigation on the contribution of underbody flow-induced noise on vehicle interior noise. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering. 2021; 235(10–11): 2667–2678. doi: 10.1177/09544070211004460
10. Duroyon M, Susini P, Misdariis N, et al. A lexicon to describe specific sounds of the electric car cabin: A verbal approach to comfort improvement. Acta Acustica. 2026; 10: 21. doi: 10.1051/aacus/2026013
11. Masri J, Amer M, Salman S, et al. A survey of modern vehicle noise, vibration, and harshness: A state-of-the-art. Ain Shams Engineering Journal. 2024; 15(10): 102957. doi: 10.1016/j.asej.2024.102957
12. Jin T, Liu X, Chen C, et al. The impact of environmental noise on drivers' cognitive abilities: A case study on in-vehicle voice interaction interfaces. Applied Ergonomics. 2024, 117: 104247. doi: 10.1016/j.apergo.2024.104247
13. Zhu H, Zhao J, Wang Y, et al. Improving of pure electric vehicle sound and vibration comfort using a multi-task learning with task-dependent weighting method. Measurement. 2024; 233: 114752. doi: 10.1016/j.measurement.2024.114752
14. Yang Y, Feng Y, Easa SM, et al. Sound effects on physiological state and behavior of drivers in a highway tunnel. Frontiers in Psychology. 2021; 12: 693005. doi: 10.3389/fpsyg.2021.693005
15. Fan D, Dai P, Yang M, et al. Research on maglev vibration isolation technology for vehicle road noise control. SAE International Journal of Vehicle Dynamics, Stability, and NVH. 2022; 6(3): 233–245. doi: 10.4271/10-06-03-0016
16. Xiang X, Yi H, Hou J, et al. Uncertainty Optimization of Vibration Characteristics in Engineering Machinery Powertrain Mounting Systems Using Monte Carlo and Genetic Algorithms. SAE International Journal of Vehicle Dynamics, Stability, and NVH. 2026; 10(1): 109–128. doi: 10.4271/10-10-01-0006
17. Wan Y, Xia J, Yang L, et al. Clear imaging beamforming method for the noise source identification in car cabin. Sound & Vibration. 2025; 59(1): 1839. doi: 10.59400/sv1839
18. Jamaluddin NS, Oettle N, Staron D. Experimental Investigation into Modulated Aeroacoustic Cabin Noise Arising from Unsteady Flow Conditions. SAE International Journal of Advances and Current Practices in Mobility. 2025; 7(6): 2925–2939. doi: 10.4271/2025-01-0029
19. Yu J, Davies P. Development and verification of a method to simulate non-stationary vehicle interior wind noise. The Journal of the Acoustical Society of America. 2024; 155(3_Supplement): A198. doi: 10.1121/10.0027292
20. Tajima A, Ikeda J, Tsubokura M. Modulated wind noise in a road vehicle due to atmospheric fluctuations and reduction measures. Journal of Wind Engineering and Industrial Aerodynamics. 2025; 266: 106210. doi: 10.1016/j.jweia.2025.106210
21. Guleria A, Novacek J, Ihi R, et al. Validation of Wind Noise for Class-8 Truck Using Lattice Boltzmann Method. SAE International Journal of Commercial Vehicles. 2025; 18(3): 391–408. doi: 10.4271/02-18-03-0021
22. Josefsson E, Sebben S, Urquhart M. Characterisation of the flow around passenger vehicle wheels with varying tyre profiles. International Journal of Heat and Fluid Flow. 2023; 103: 109191. doi: 10.1016/j.ijheatfluidflow.2023.109191
23. Nusser K, Becker S. Numerical investigation of the fluid structure acoustics interaction on a simplified car model. Acta Acustica. 2021; 5: 22. doi: 10.1051/aacus/2021014
24. Jia Q, Zhang J, Liang W, et al. A mathematical-boundary-recognition domain-decomposition Lattice Boltzmann method combined with large eddy simulation applied to airfoil aeroacoustics simulation. International Journal for Numerical Methods in Fluids. 2024; 96(7): 1250–1275. doi: 10.1002/fld.5287
25. Guo Z, Lei Z, Zeng K, et al. Non-integer-dimensional architected materials enabling synergistic acoustic,

mechanical, and fluid coupling. Materials Horizons. 2026; 13(2): 748–762. doi: 10.1039/d5mh01768h
26. Guo Z, Du L, Yang Z, et al. Switchable bidirectional sound absorption via exceptional point modulation in acoustic metastructures with interleaved resonator coupling. Advanced Science. 2025; 12(44): e08951. doi: 10.1002/advs.202508951
27. Cho H, Jeong J, Jang J, et al. VAE-based surrogate model of rotor aerodynamic load distribution for noise prediction. Aerospace Science and Technology. 2026; 178: 112412. doi: 10.1016/j.ast.2026.112412
28. Huang Y, Boerboom M, Wolff K, et al. Find Optimal Suspension Kinematics Targets for Vehicle Dynamics Using Reinforcement Learning. SAE International Journal of Vehicle Dynamics, Stability, and NVH. 2025; 10(1): 25–41. doi: 10.4271/10-10-01-0002
29. Zhang Q, Wang X, Yang D, et al. Data-driven prediction of aerodynamic noise of transonic buffeting over an airfoil. Engineering Analysis with Boundary Elements. 2024; 163: 549–561. doi: 10.1016/j.enganabound.2024.04.006
30. Zhao Y, Liu H, Guo W, et al. Research on sound quality prediction of vehicle interior noise using the human-ear physiological model. The Journal of the Acoustical Society of America. 2024; 156(2): 989–1003. doi: 10.1121/10.0028130
31. Qi Z, Liu P, Duan H, et al. RDAS: A Shared Control Strategy Considering Real-Time Driving Ability in Driver Takeover Tasks. SAE International Journal of Vehicle Dynamics, Stability, and NVH. 2026; 10(2): 303–326. doi: 10.4271/10-10-02-0014
32. Yang H, Yuan W, Zhu W, et al. Wind turbine airfoil noise prediction using dedicated airfoil database and deep learning technology. Applied Energy. 2024; 364: 123165. doi: 10.1016/j.apenergy.2024.123165
33. Roznowicz D, Stabile G, Demo N, et al. Large-scale graph-machine-learning surrogate models for 3D-flowfield prediction in external aerodynamics. Advanced Modeling and Simulation in Engineering Sciences. 2024; 11(1): 6. doi: 10.1186/s40323-024-00259-1
34. Hou J, Yi H, Xiang X, et al. Identification of vehicle suspension shock absorber rattle noise based on wavelet packet feature fusion and GWO-LSTM. Sound & Vibration. 2025; 59(2): 1941. doi: 10.59400/sv1941
35. Zhao J, Yin Y, Chen J, et al. Evaluation and prediction of vibration comfort in engineering machinery cabs using random forest with genetic algorithm. SAE International Journal of Vehicle Dynamics, Stability, and NVH. 2024; 8(4): 491–512. doi: 10.4271/10-08-04-0027
36. Dai R, Zhao J, Zhao W, et al. Exploratory study on sound quality evaluation and prediction for engineering machinery cabins. Measurement. 2025; 253: 117684. doi: 10.1016/j.measurement.2025.117684
37. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. Advances in Neural Information Processing Systems. In: Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017); 4–9 December 2017; Long Beach, CA, USA.
38. Islam S, Elmekki H, Elsebai A, et al. A comprehensive survey on applications of transformers for deep learning tasks. Expert Systems with Applications. 2024; 241: 122666. doi: 10.1016/j.eswa.2023.122666
39. Soydaner D. Attention mechanism in neural networks: Where it comes and where it goes. Neural Computing and Applications. 2022; 34(16): 13371–13385. doi: 10.1007/s00521-022-07366-3
40. Hosseinzadeh R, Sadeghzadeh M. Attention mechanisms in transformers: A general survey. Journal of AI and Data Mining. 2025; 13(3): 359–368. doi: 10.22044/jadm.2025.15584.2679
41. Thirunavukarasu R, Kotei E. A comprehensive review on transformer network for natural and medical image analysis. Computer Science Review. 2024; 53: 100648. doi: 10.1016/j.cosrev.2024.100648
42. Ampazis N, Sakketou F. Diversifying multi-head attention in the transformer model. Machine Learning and Knowledge Extraction. 2024; 6(4): 2618–2638. doi: 10.3390/make6040126
43. Sanger TD. Transformer Models for Signal Processing: Scaled Dot-Product Attention Implements Constrained Filtering. Neural Computation. 2025; 37(10): 1839–1852. doi: 10.1162/neco.a.29
44. Su L, Zuo X, Li R, et al. A systematic review for transformer-based long-term series forecasting. Artificial Intelligence Review. 2025; 58(3): 80. doi: 10.1007/s10462-024-11044-2
45. Cong S, Wang H, Zhou Y, et al. Comprehensive review of Transformer-based models in neuroscience, neurology, and psychiatry. Brain-X. 2024; 2(2): e57. doi: 10.1002/brx2.57
46. Isomura T, Shimizu R, Goto M. Sparse attention is all you need for pre-training on tabular data. Neural Computing and Applications. 2025; 37(3): 1509–1522. doi: 10.1007/s00521-024-10698-x
47. Lee A, Hwang E, Kim D. A Practical CNN–Transformer Hybrid Network for Real-World Image Denoising. Mathematics. 2026; 14(1): 203. doi: 10.3390/math14010203
48. Gao L. CNN-Transformer Hybrid Models for Object Detection: A Comprehensive Review. Science and Technology

of Engineering, Chemistry and Environmental Protection. 2025; 1(4).
49. Wang Z, Liu J, Sun H, et al. A CNN–LSTM–SVM Framework for Mapping Driving Styles to Adaptive Suspension Modes. SAE International Journal of Vehicle Dynamics, Stability, and NVH. 2026; 10(2): 237–258. doi: 10.4271/10-10-02-0011
50. Kim Y. Convolutional neural networks for sentence classification. In: Moschitti A, Pang B, Daelemans W (editors). Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP); Association for Computational Linguistics; 2014. pp. 1746–1751. doi: 10.3115/v1/D14-1181
51. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems. 2012; 60(6): 84–90. doi: 10.1145/3065386
52. Peng C, Cheng S, Sun M, et al. Prediction of sound transmission loss of vehicle floor system based on 1d-convolutional neural networks. Sound & Vibration. 2024; 58(1): 046940. doi: 10.32604/sv.2024.046940
53. Dursun ÖO. Deep Learning-Based Prediction of Commercial Aircraft Noise: A CNN–Transformer Hybrid Model Versus Support Vector Regression and Multi-Layer Perceptron. Aerospace. 2025; 12(11): 1031. doi: 10.3390/aerospace12111031
54. Wu Y, Wang Y, Cai H, et al. An incremental intelligent fault diagnosis method for marine diesel engines based on CNN-Transformer and cosine similarity. Sound & Vibration. 2025; 59(6): 3617. doi: 10.59400/sv3617
55. Jiao D, Zhou H, Huang T, et al. Numerical Study on Aerodynamic Noise Reduction in Passenger Car with Fender Shape Optimization. Symmetry. 2024; 16(6): 651. doi: 10.3390/sym16060651
56. Xu L, Zhu X, Wang Q, et al. The New China Automotive Engineering Research Institute Co., Ltd.: Full-Scale Aero-Acoustic Wind Tunnel. In: Proceedings of the WCX SAE World Congress Experience; 8–10 April 2025; Detroit, MI, USA. doi: 10.4271/2025-01-8779
57. Duell EG, Walter J, Yen J, et al. Progress in aeroacoustic and climatic wind tunnels for automotive wind noise and acoustic testing. SAE International Journal of Passenger Cars-Mechanical Systems. 2013; 6(1): 448–461. doi: 10.4271/2013-01-1352
58. ISO 532-1:2017. Acoustics—Methods for Calculating Loudness—Part 1: Zwicker Method. 2017.
59. Zaareer M, Mourad AH. Effect of vehicle side mirror base position on aerodynamic forces and acoustics. Alexandria Engineering Journal. 2022; 61(2): 1437–1448. doi: 10.1016/j.aej.2021.06.049
60. Xiong Z, Zhang X, Liu P, et al. Optimization of Window Glass Parameters on Vehicle Weight and Wind Noise. SAE International Journal of Passenger Vehicle Systems. 2025; 18(3): 281–298. doi: 10.4271/15-18-03-0018
61. Toer GA, Kim G. An empirical study on the impact of feature scaling and encoding strategies on machine learning regression pipelines. International Journal of Informatics and Information Systems. 2026; 9(1): 242–256. doi: 10.47738/ijiis.v9i1.293
62. Shen X, Shen J, Zhang Y, et al. A multi-source feature engineering-enhanced framework for mid-to-long-term EV charging load forecasting: Integrating self-adaptive optimization and BiLSTM-iTransformer predictor. Energy. 2025; 139020. doi: 10.1016/j.energy.2025.139020