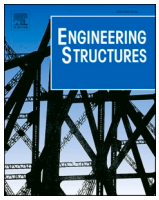
Engineering Structures 314 (2024) 118415

Contents lists available at ScienceDirect

Engineering Structures

journal homepage: www.elsevier.com/locate/engstruct



Prediction of BLEVE-induced response of road tunnel using Transformer network with modified self-attention (SAMT)

Ruishan Cheng [a], Wensu Chen [a,*], Hong Hao [b,a,**], Qilin Li [c]

[a] Center for Infrastructural Monitoring and Protection, School of Civil and Mechanical Engineering, Curtin University, Australia
[b] Earthquake Engineering Research & Test Center, Guangzhou University, China
[c] School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Australia

ARTICLE INFO

Keywords:
Road tunnels
BLEVE
Transformer network
Modified self-attention
Support rotation

ABSTRACT

Road tunnels might be exposed to Boiling Liquid Expansion Vapor Explosion (BLEVE) due to the transportation of liquified gas tankers passing through road tunnels. An efficient and accurate prediction of the response of road tunnels under internal BLEVEs can facilitate the reliable BLEVE-resistant design and risk assessment of road tunnels. This study introduces an advanced deep-learning model that employs a Transformer-based architecture with a modified self-attention mechanism, termed as Self-Attention Modified Transformer (SAMT), to predict BLEVE-induced support rotation of tunnel structure, which is a common criterion in assessing reinforced concrete structure damage to blast loads. Unlike the Transformer with the traditional self-attention mechanism, the proposed SAMT effectively aggregates global information across all variables while mitigating undue dependencies among uncorrelated variables. Consequently, the proposed SAMT is better suited for processing tabular data with uncorrelated variables. The feasibility and advantages of the proposed SAMT are verified by extensive data generated using calibrated numerical models of box-shaped road tunnels subjected to internal BLEVEs. By comparing the performance of the proposed SAMT with the non-modified Transformer network (FT-Transformer) as well as two other typical deep learning networks, i.e., Multi-layer Perceptron (MLP) and Residual Network (ResNet), it is found that the SAMT offers higher prediction accuracy and robustness than the other three models in predicting BLEVE-induced support rotations of box-shaped road tunnels. The study demonstrates that the proposed SAMT is an effective tool for the prediction of BLEVE-induced support rotations of road tunnels.

1. Introduction

Road tunnels are important components in traffic networks for time and cost savings in the transportation of goods and personnel between two locations [1]. With the growing demand in liquefied petroleum gas (LPG) for residential and commercial use, the number of LPG tankers transported through road tunnels has been increasing rapidly [2], leading to an increasing probability of Boiling Liquid Expanding Vapors Explosion (BLEVE) inside road tunnels [3]. Accidental BLEVEs inside tunnels are very likely to cause significant damage to road tunnels, inevitable interruption of traffic, and great economic loss and injuries [4,5]. Therefore, it is essential to conduct the blast-resistant design to ensure the safety of road tunnels when exposed to internal BLEVE incidents.

Current designs of road tunnels against blast loads, given in the existing codes and guidelines (e.g., GB 50038–2005 [6]), typically rely on simplified assumptions regarding blast loads and tunnel structure mechanical properties, as well as their surrounding conditions. These simplifications of blast loads and tunnel properties may lead to overly conservative or unsafe design of blast resistance of tunnels. To accurately predict the response of road tunnels under blast loads, experimental tests and numerical simulations have been widely employed in

Abbreviations: LPG, Liquefied petroleum gas; BLEVE, Boiling Liquid Expanding Vapors Explosion; SAMT, Transformer with modified self-attention; TPE, Tree-Structured Parzen Estimator; MHSA, multi-head self-attention; FNN, Feedforward neural network; MLP, Multi-layer perceptron; RC, Reinforced concrete; SA, Self-attention; ResNet, Residual Network; MSE, Mean Squared Error; FT, Feature-Tokenizer.

* Corresponding author.
** Corresponding author at: Earthquake Engineering Research & Test Center, Guangzhou University, China
E-mail addresses: wensu.chen@curtin.edu.au (W. Chen), hong.hao@curtin.edu.au (H. Hao).

https://doi.org/10.1016/j.engstruct.2024.118415
Received 14 September 2023; Received in revised form 6 June 2024; Accepted 8 June 2024
Available online 18 June 2024

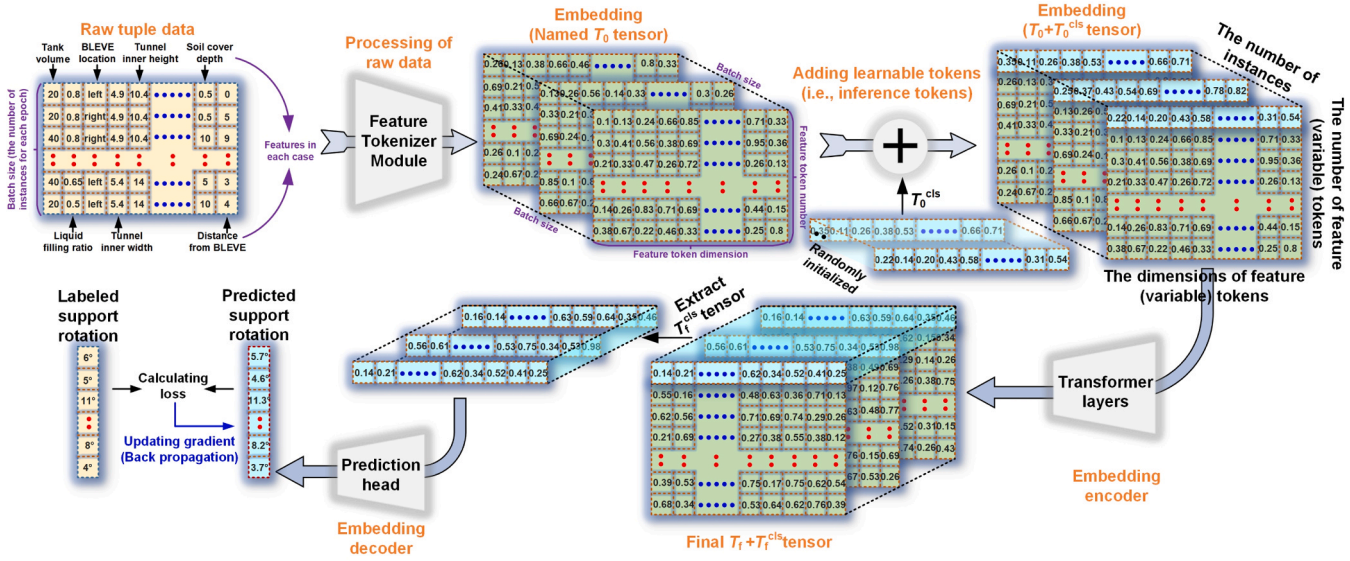


Fig. 1. Flowchart of trainable FT-Transformer model for the prediction of BLEVE-induced support rotations of tunnels.

current practice and research [7–10]. However, the tests of road tunnels under blast loads are expensive, time-consuming, and highly sensitive to testing setup. On the other hand, the numerical simulation of road tunnels under blast loads requires profound modeling knowledge, fine mesh, and careful model calibrations. The execution of finely-meshed numerical models needs advanced computational resources and professional software and facilities, which hinders its application in engineering practice. Therefore, more efficient approaches are needed for engineers to carry out the blast-resistant analysis and design of road tunnels to resist blast loads.

In recent years, deep learning-based methods have made great progress and are applied to solve various civil engineering problems [11–15] due to their efficiency and ease of use once the model is established. The popular application of deep learning methods in predicting structural response is partially encouraged by their excellent performance in processing audio, texts, and images [16]. However, the nature of the data involved in predicting structural responses to blast loads is fundamentally different. Typically, such data is presented in tabular format where each column represents a variable, and each row is an instance [17]. Unlike image or text data, where local variables exhibit global dependencies and correlations among local variables significantly influence final outputs, the parameters for structural response prediction are generally independent. This presents a distinct challenge for the applicability of many deep learning models designed for data with inherent inter-variable relationships.

Several studies have tried to use typical deep learning method such as Multi-layer perceptron (MLP) to process the tabular dataset for the prediction of structural response under blast loads. For instance, Zhou et al. [18] adopted MLP to evaluate the damage levels of reinforced concrete (RC) columns subjected to blast loads. The parameters of the RC columns and blast loads are used as the variables to train MLP to predict structural damage levels. In addition, Liu et al. [19] employed MLP to predict the final plastic deformations of stiffened plates subjected to close-in explosions. The charge weight, the stand-off distance between the charge and stiffened plates, and the thicknesses of the plate and stiffener were used as the variables in tabular datasets. It should be noted that many advanced deep learning methods [20] have been developed and proven to outperform MLP in computer vision and natural language processing. The performance of MLP, although it yielded good predictions, was not compared with other emerging deep learning methods for the prediction of structural response under blast loads. Li et al. [21] performed a comparative study of different deep learning methods in BLEVE load predictions and concluded Transformer method outperformed other commonly used deep learning methods in explosion load predictions. This is attributed to the self-attention (SA) mechanism in Transformer architectures, which excels at aggregating features from different variables [22]. However, the SA mechanism, by design, captures inter-variable dependencies, potentially reducing the model's performance when applied to engineering tasks with independent variables, such as predicting the responses of engineering structures and civil infrastructures (e.g., road tunnels) under blast loads.

The aim of this study is to develop an improved Transformer-based architecture specifically tailored for the response prediction of road tunnels subjected to internal BLEVEs. The proposed model, namely the Transformer with modified self-attention (SAMT), refines the conventional self-attention mechanism of the Transformer-based network to better process the tabular data with uncorrelated variables. Many advanced techniques, such as hyperparameter tuning based on Tree-Structured Parzen Estimator (TPE) approach and early stopping, have been incorporated into the proposed SAMT model to obtain optimal performances. A total of 115 cases and 3450 instances generated from the validated numerical simulations of box-shaped road tunnels subjected to internal BLEVEs are utilized to train the proposed model and evaluate its performance.

It is noted that the support rotation of tunnel structures under blast loads is used as the indicator of tunnel damage level under internal BLEVEs, since many existing codes (e.g., UFC-3–340-02 [23]) adopt this indicator to establish the structural damage criteria. Its detailed calculation has been given in the authors' previous study [27]. To demonstrate the advantage of the proposed SAMT to predict the support rotations of road tunnels subjected to internal BLEVEs, the non-modified Transformer-based network (Feature-Tokenizer Transformer) and two typical deep learning models, including Multi-layer perceptron (MLP) and Residual Network (ResNet) are also employed to process these data for comparative analysis. In addition, the datasets with different numbers of variables and instances are generated based on the given data to test the robustness of the SAMT in predicting the support rotations of road tunnels. The methodology of the SAMT is given in Section 2, followed by the data preparation of BLEVE-induced support rotations of road tunnels in Section 3. The implementation details and performance evaluation of the SAMT in predicting the support rotations of road tunnels under internal BLEVEs are presented in Section 4.

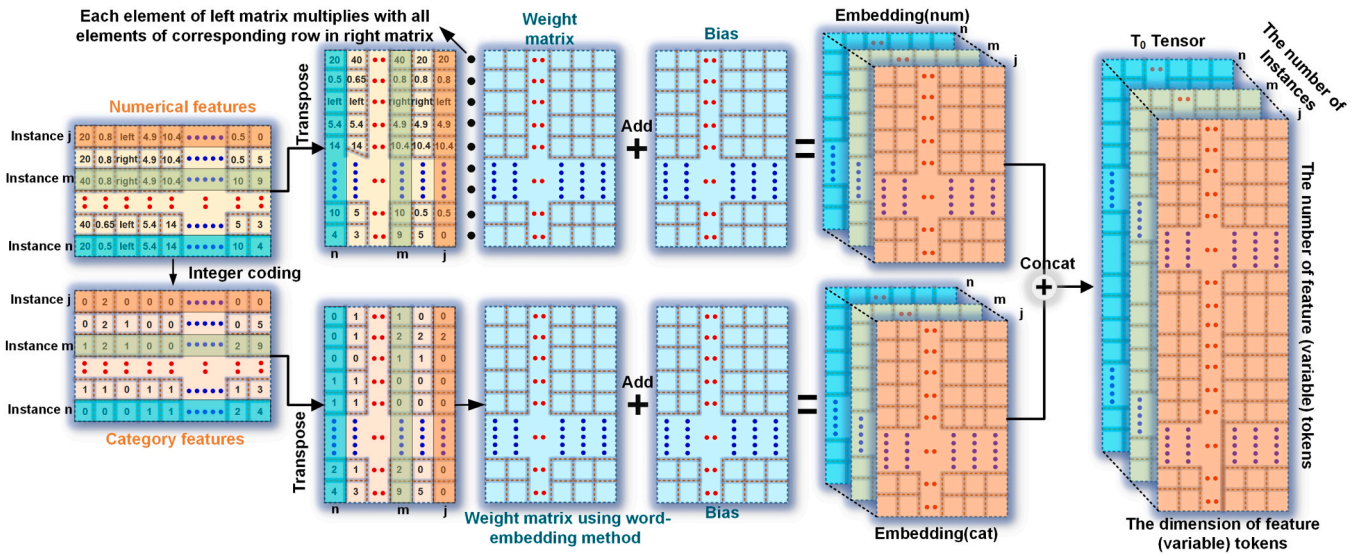


Fig. 2. Details of Feature Tokenizer module for feature extraction of original data.

2. Methodology of Transformer with modified self-attention (SAMT)

Transformer-based networks, as one type of advanced deep learning models, have been developed and widely used to treat complex texts, audio, and images [24]. They have achieved remarkable accomplishments on various natural language processing and computer vision tasks [25]. To extend this success to the domain of tabular data with instances given in rows and variables given in columns, Feature-Tokenizer Transformer (FT-Transformer) has been proposed by Gorishniy et al. [22] to address the non-sequential nature of tabular data by generating multi-dimensional embeddings for each variable. However, the self-attention mechanism utilized in Transformer networks establishes the dependencies between different variables, which is not appropriate for the inference of variables without interdependence. In this section, the FT-Transformer model is refined to process tabular data with no cross-correlation among variables, which is common in civil engineering and is also consistent with the characteristics of the collected data to predict the BLEVE-induced support rotations of road tunnels in this study. The details of modifications are elaborated in the subsequent sections.

2.1. Feature-Tokenizer Transformer model

The Feature-Tokenizer Transformer (FT-Transformer) model is an extension of the original Transformer architecture developed by Vaswani et al. [26]. Designed to accommodate tabular data, the FT-Transformer converts variables into multi-dimensional embeddings, subsequently processed by a stack of Transformer layers. Three main parts, i.e., Feature Tokenizer module, Transformer layers, and Prediction head, are included in the FT-Transformer model. Fig. 1 shows its overall architecture and procedures to establish a trainable network for the prediction of tunnel support rotations under internal BLEVEs.

The original data involving the parameters of tunnels and BLEVEs

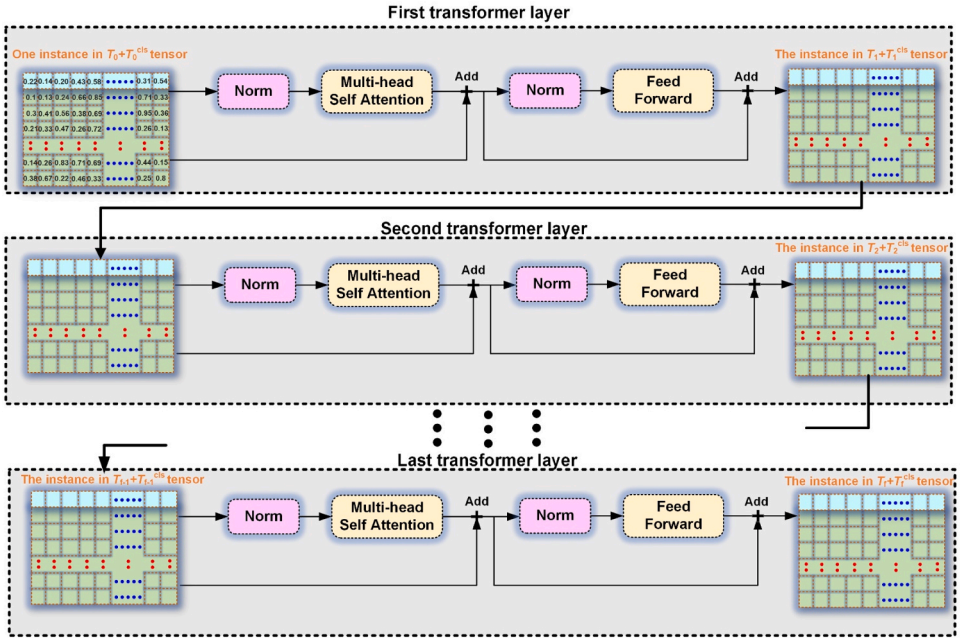


Fig. 3. The architecture of Transformer layers to process an input instance.

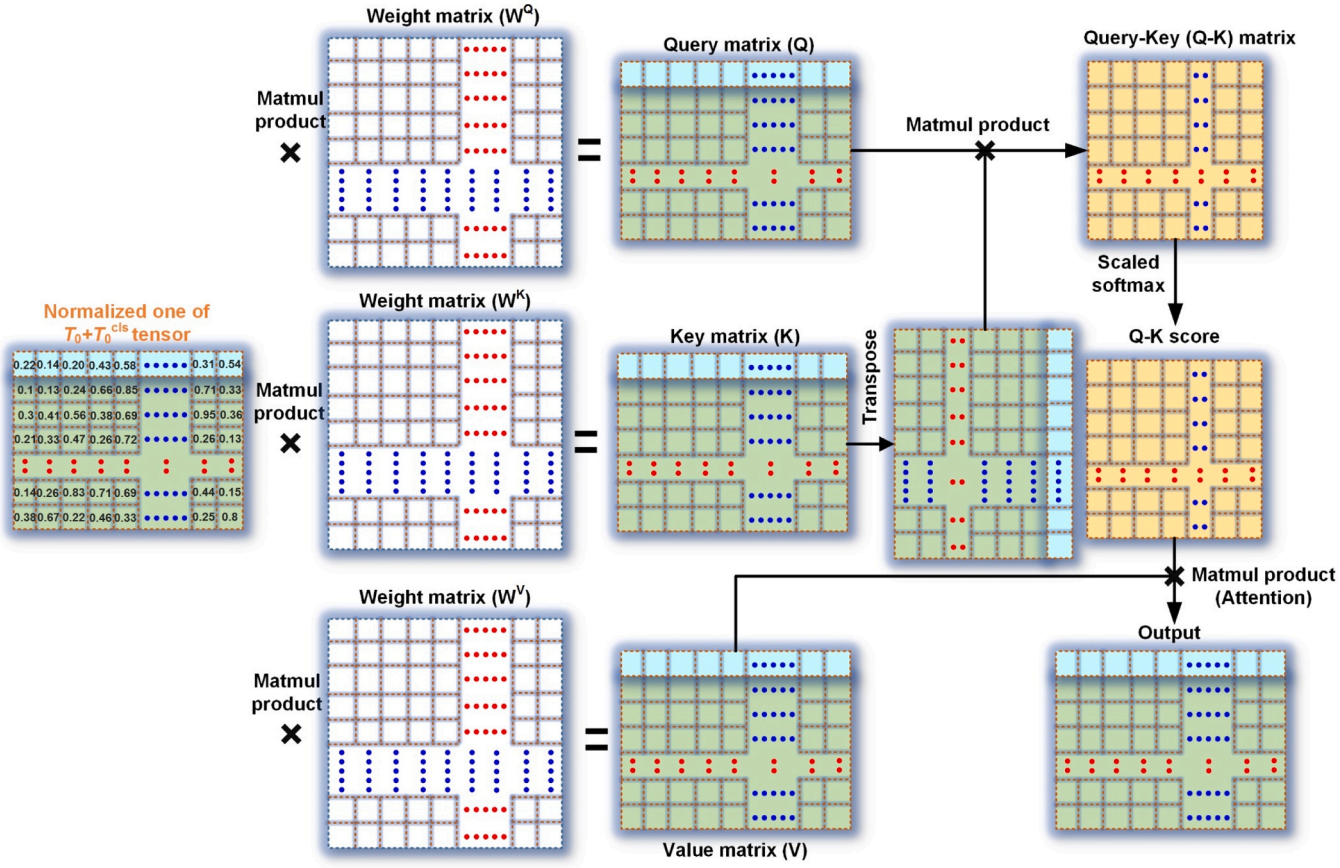


Fig. 4. The architecture of self-attention mechanism.

are first converted by the Feature Tokenizer module into a three-dimensional tensor (the embedding). The generated embedding is then augmented with a randomly initialized, learnable token (i.e., the inference tokens) and processed by the following stacked Transformer layers. Next, the Prediction Head module is employed to decode the enriched inference token to obtain the predicted support rotations. The losses between the labeled (prior) and predicted support rotations are calculated to update the gradients of weights and biases of the FT-Transformer via back-propagation. A reliable network to predict the support rotations of tunnels under BLEVEs can be finally established through iterative refinement. Detailed explanations of each major component of the FT-Transformer model are provided in the subsequent sections.

2.1.1. Feature Tokenizer module

The objective of the Feature Tokenizer module is to convert each scalar feature, including numerical and categorical features, to a tensor embedding on which the subsequent Transformer acts. To generate the numerical embedding, each numerical feature in any row of the numerical feature transpose matrix multiplies element-wisely with all elements in the corresponding row of the weight matrix $W^{(num)}$ and is then added with all elements in the corresponding row of the bias matrix $B^{(num)}$, as shown in Fig. 2. The categorical features are converted into a categorical embedding by using the word-embedding method and then adding it with the bias matrix $B^{(cat)}$. The numerical embedding and categorical embedding are finally concatenated into an embedding (i.e., T_0 tensor) for subsequent processing.

2.1.2. Transformer layers

Before feeding to the Transformer layers, the embedding (T_0 tensor) output from the Feature Tokenizer module is augmented with an inference token (i.e., T_0^{cls}) (see Fig. 1), where the sizes of the columns, rows, and pages are the dimension of each feature token, the number of feature tokens, and the number of instances, respectively. The goal of the Transformer layers is to refine the feature token, including the inference token, with global contextual information. In particular, the inference token will aggregate information from all other features so that it can be then used to make the final prediction.

Fig. 3 presents the overall procedure of Transformer layers encoding an instance extracted from the three-dimensional tensor. The feature tokens of the instance are firstly normalized using the *layer-normalization* method and are then passed into the multi-head self-attention (MHSA) sub-module. To reduce the information degradation caused by data propagation in the network, the residual connection is added, i.e., the output from the MHSA is added with feature tokens of the extracted instance (i.e., unnormalized input) as the input of the next sub-module, i.e., the feedforward neural network (FNN) sub-module. Similarly, the output from the FNN is added with the unnormalized input for FNN as the input of the following Transformer layer, which has the same architecture as the first transformer layer. The final encoded tensor is obtained by repeating the above process from the second transformer layer to the last transformer layer. The FNN in the Transformer layers is a well-known multi-layer perceptron, which is not detailed herein. The MHSA is the core block of the Transformer layers, which is therefore described subsequently in detail.

The self-attention mechanism in the MHSA module can aggregate the normalized feature tokens of T_0 and T_0^{cls} as the output. Fig. 4 shows the whole procedure of the self-attention mechanism to process a normalized instance originating from the stacked tensor of T_0 and T_0^{cls}. Three matrices i.e., Query (Q) matrix, Key (K) matrix, and Value (V) matrix are firstly generated by multiplying the normalized instance with three trainable weight matrices (W^Q, W^K, W^V). The correlations between

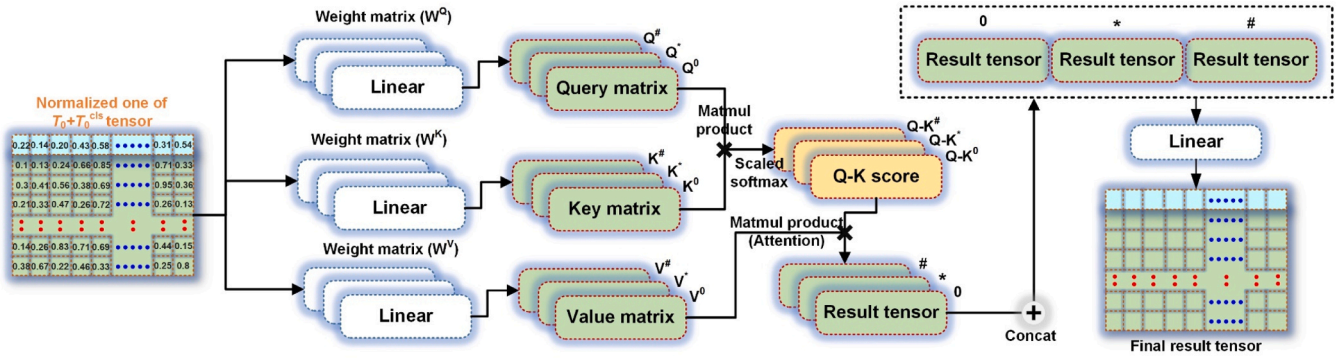


Fig. 5. The architecture of multi-head self-attention (MHSA) sub-module to process an instance.

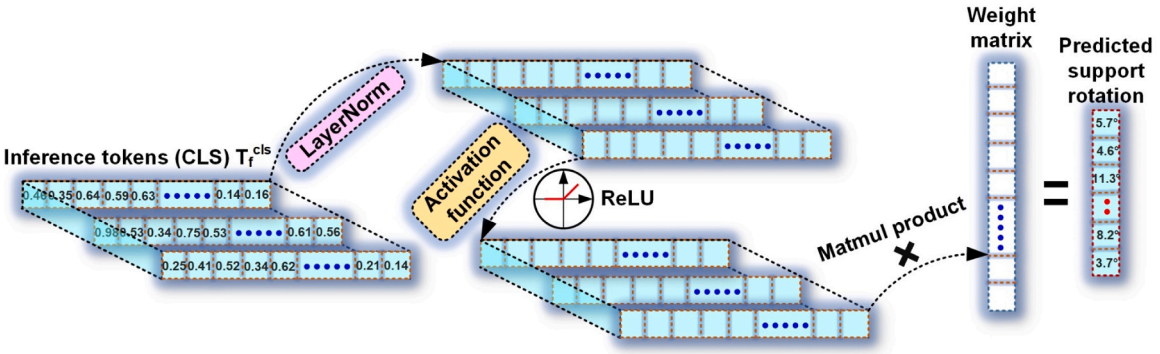


Fig. 6. The flowchart of prediction head module to decode inference tokens.

different features of the normalized instance are computed as the product of the *Q* matrix and transposed *K* matrix, which is divided by the scaling factor $\sqrt{d_k}$(i.e., the square root of the dimension of the row in the *K* matrix) and is operated by the Softmax function to generate *Q-K* scores, i.e., the weights on the *V* matrix. The final output is given by multiplying the *Q-K* scores with the *V* matrix. The expression is given below.

$$Attention\left(Q,K,V\right)=\text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

The multi-head self-attention sub-module is an extension of the self-attention mechanism and allows the normalized instance to represent more characteristics in different subspaces. The multi-head self-attention sub-module employs multiple parallel weight matrices W^Q, W^K, W^V to linearly project the normalized instance into multiple groups of matrices of *Q*, *K*, and *V*, as shown in Fig. 5. The self-attention process is performed on each of these projected groups of *Q*, *K*, and *V* in parallel. The tensors generated from all projected groups are concatenated (see Eq.$MultiHead(Q,K,V)=Concat(Attention(Q^1,K^1,V^1),\ldots,Attention\quad(Q^h,K^h,V^h))$ (2) and projected linearly into a final output tensor.

$$MultiHead\left(Q,K,V\right)=Concat\left(Attention\left(Q^1,K^1,V^1\right),\ldots,Attention\left(Q^h,K^h,V^h\right)\right) \tag{2}$$

2.1.3. Prediction head

The prediction head module is employed to decode the inference tokens output from the module of Transformer layers. The inference tokens (T_f^{cls}) are firstly normalized using the *layer-normalization* method and are then activated using the *ReLU* function, as shown in Fig. 6. The final prediction of support rotations is achieved by linearly projecting the activated inference tokens with a trainable weight matrix (see Eq. $\hat{y}=\text{Linear}(\text{ReLU}(Layernorm(T_f^{cls})))$ (3)). The weight matrix is iteratively updated by minimizing the loss between the predicted and actual support rotations, as shown in Fig. 1. The loss is expressed using the function of Mean Squared Error (MSE) in this study, as given in Eq. $MSE=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)$ (4).

$$\hat{y}=\text{Linear}\left(\text{ReLU}\left(Layernorm\left(T_f^{cls}\right)\right)\right) \tag{3}$$

$$MSE=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i) \tag{4}$$

where *n* is the total number of instances, y_i and $\hat{y}_i$are the actual and predicted support rotations of the i^{th} instance, respectively.

2.2. Modified self-attention mechanism

As discussed in Section 2.1.2, the output from the MHSA of Transformer layers is calculated by the product of the *Q-K* score matrix (i.e., weights on *V* matrix) and *V* matrix. That is, any element in each token of the output matrix is the dot product of the *Q-K* scores in the same row as the element and the vector in the *V* matrix in the same column as the element. For example, the first element in the inference token (i.e., *CLS* token) of the output matrix (i.e., S_1V_1 given in Fig. 7) is the sum of the products of the scores in the first row of the *Q-K* score matrix and the

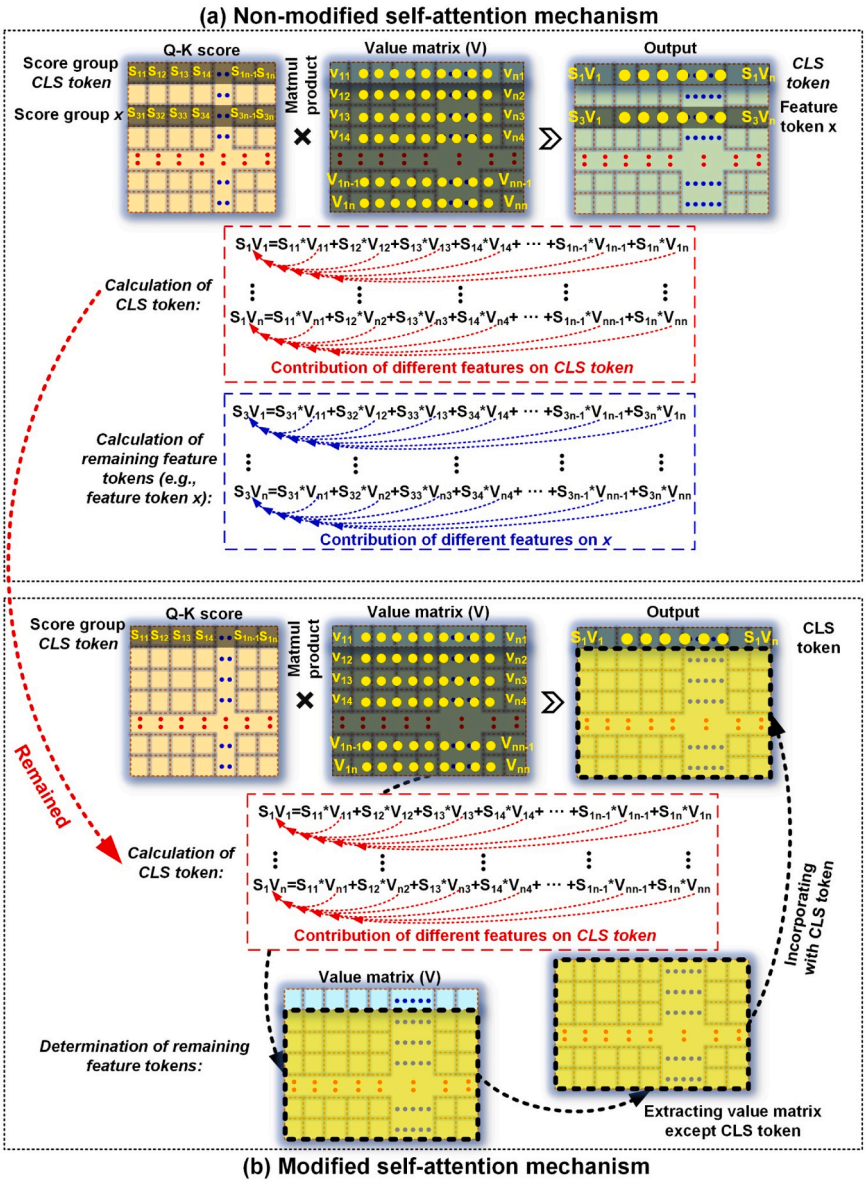


Fig. 7. Comparison of original and modified self-attention mechanisms.

values in the first column of the V matrix. Therefore, all local feature tokens of the V matrix contribute to the inference token in the output matrix. This mechanism facilitates the aggregation of relevant information from all local feature tokens to the inference token, which can unbiasedly reflect all features of the normalized instance. In addition, this mechanism can associate each feature token with other local feature tokens. For example, the element S_3V_1 in the output matrix given in Fig. 7 is the dot product of the score group x and the vector at the first column of the V matrix. Each feature in the V matrix has a certain contribution to S_3V_1. In summary, the inference token first aggregates the details of all local feature tokens and then shares the global summary among local feature tokens by using self-attention mechanism, which can be regarded as two procedures, i.e., the aggregation and propagation of global information. During the propagation phase, each local feature token establishes the correlation with other local feature tokens through the action of self-attention mechanism. Since the module of Transformer layers stacks multiple Transformer layers as shown in Fig. 3, the correlation information between different local feature tokens, established in the preceding Transformer layer, is aggregated into the inference tokens in subsequent Transformer layers via the self-attention mechanism.

Consequently, the final inference token output from the module of Transformer layers is significantly affected by the cross-correlation of local feature tokens. This becomes problematic when predicting structural responses under blast loads, as the data for these predictions often consists of uncorrelated variables. In other words, there is no inherent correlation among different local feature tokens. Therefore, using the inference token to aggregate data with uncorrelated variables may lead to the inclusion of inappropriate global information. This occurs because the self-attention mechanism still enforces correlation between different local feature tokens that are not intrinsically dependent on each other. Such inappropriate aggregation adversely affects the prediction accuracy for structural responses under blast loads.

To overcome the aforementioned issue, a modification is introduced to the self-attention mechanism in the MHSA sub-module. The modification keeps the process of inference tokens aggregating the information of all local feature tokens but discards the propagation process of global information. The specific details are shown in Fig. 7(b). The calculation of the inference token of the output matrix in each Transform layer remains the same as the non-modified self-attention mechanism to aggregate information from all local feature tokens. However, the local

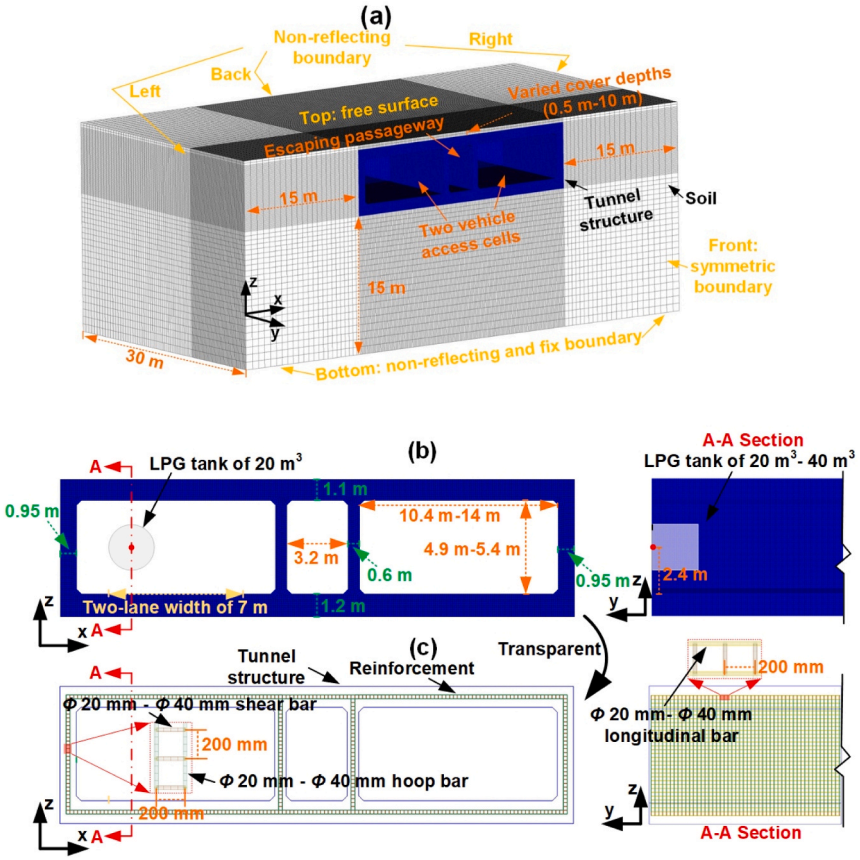


Fig. 8. Numerical models to generate the data of BLEVE-induced support rotations of road tunnel structures [27].

feature tokens of the output matrix in each Transformer layer are determined by extracting the elements in V matrix except for the inference tokens instead of generating them using the self-attention mechanism. The final output is obtained by concatenating the inference token and the extracted element tokens. This modification preserves the effective aggregation of relevant information from all local feature tokens, while eliminating the interference of forced correlations between different local feature tokens, thereby mitigating their undue influence on the inference tokens.

2.3. Auto-hyperparameter tuning using TPE approach

The performance of Transformer-based networks as well as many other machine learning approaches depends heavily on the selection of hyperparameters, e.g., the number of Transformer layers. In this study, the Tree-Structured Parzen Estimator (TPE) approach [27] is employed to automatically tune the hyperparameters of the self-attention-modified Transformer network. Its details are given below.

The TPE approach is a Bayesian-based optimization algorithm, which has achieved outstanding performances in optimizing the hyperparameters of various machine learning methods. It firstly assumes the prior distributions and search ranges of all hyperparameters to be tuned (e.g., the number of Transformer layers), which are stored in the tree-structured memory. A certain number of observations, i.e., several groups of hyperparameters are randomly picked from the hyperparameter space as the prior set D at the first iteration. The TPE defines the probability density $p(x|y, D)$ using the following assumption:

$$p\left(x|y,D\right)=\begin{cases} p\left(x|D^{(l)}\right) & \left(\mathrm{y}\le \mathrm{y}^{\gamma}\right) \\ p(x|D^{(g)}) & \left(\mathrm{y}>\mathrm{y}^{\gamma}\right) \end{cases} \tag{5}$$

where $p(x|D^{(l)})$ is the density of x in the better hyperparameter set $D^{(l)}$ that the corresponding loss is not greater than $\mathrm{y}^{\gamma}$and $p(x|D^{(g)})$ is the density of the remaining observations in D (i.e., $D(g)$), in which y^{γ} is the top-γ-quantile objective value in the set of observations D.

Parzen Estimators, i.e., kernel density estimators $p(x|D^{(l)})$ and $p(x|D^{(g)})$ are estimated via:

$$p\left(x|D^{(l)}\right)=\omega_0^{(l)}p_0(x)+\sum_{n=1}^{N^{(l)}}\omega_n k\left(x,x_n\middle|b^{(l)}\right) \tag{6}$$

$$p\left(x|D^{(g)}\right)=\omega_0^{(g)}p_0(x)+\sum_{n=N^{(l)}+1}^{N}\omega_n k(x,x_n|b^{(g)}) \tag{7}$$

where p_0 is a non-informative prior, ω_n is the weight set, k is the kernel function, $b^{(l)}$ and $b^{(g)}$ are the bandwidths of k for $p(x|D^{(l)})$ and $p(x|D^{(g)})$. The details of the above parameters can refer to Watanabe [27]. It is noted that the TPE algorithm utilizes the expected improvement (EI) as an acquisition function to update the optimal hyperparameter candidates. EI can be derived as follows:

$$\begin{aligned} EI_{\mathrm{y}^{\gamma}}\left(x|D\right)&=\int_{-\infty}^{\mathrm{y}^{\gamma}}(\mathrm{y}^{\gamma}-y)p(y|x,D)dy \\ &=\int_{-\infty}^{\mathrm{y}^{\gamma}}\left(\mathrm{y}^{\gamma}-y\right)\frac{p(x|y,D)p(y|D)}{p(x|D)}dy \\ &=\frac{p\left(x|D^{(l)}\right)}{\gamma p\left(x|D^{(l)}\right)+\left(1-\gamma\right)p\left(x|D^{(g)}\right)}\int_{-\infty}^{\mathrm{y}^{\gamma}}(\mathrm{y}^{\gamma}-y)p(y|D)dy \end{aligned} \tag{8}$$

Subjected to:

$$p\left(x|D\right) = \int_{-\infty}^{+\infty} p(x|y,D)p(y|D)dy$$
$$= p\left(x|D^{(l)}\right)\int_{-\infty}^{y^{\gamma}} p(y|D)dy + p\left(x|D^{(g)}\right)\int_{y^{\gamma}}^{+\infty} p(y|D)dy \quad (9)$$
$$= \gamma p\left(x|D^{(l)}\right) + \left(1-\gamma\right)p\left(x|D^{(g)}\right)$$

Based on Eq. (8), EI can be expressed as

$$EI_{y^{\gamma}}\left(x|D\right) \propto \left(\gamma + \left(1-\gamma\right)\frac{p(x|D^{(g)})}{p\left(x|D^{(l)}\right)}\right)^{-1} \quad (10)$$

Based on Eq. $EI_{y^{\gamma}}(x|D) \propto \left(\gamma + (1-\gamma)\frac{p(x|D^{(g)})}{p(x|D^{(l)})}\right)^{-1}$ (10), it can be found that to maximize improvement, the optimal hyperparameter set should have high probability under $p(x|D^{(l)})$ and low probability under $p(x|D^{(g)})$.

2.4. Evaluation metrics

Many metrics, such as the Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and coefficient of determination (R^2), can be adopted to evaluate the performance of supervised machine learning methods. In this study, RMSE and R^2 are employed as the evaluation metrics of the self-attention-modified Transformer network, as given in Eqs. (11) and (12). RMSE is sensitive to large local errors between predicted and labeled data, which can reflect the adaptability of the modified Transformer network to different instances. R^2 quantifies the historical distribution relation between predicted and labeled data, which can measure the overall robustness of the modified Transformer network. The trained model is better when the smaller RMSE is achieved and the R^2 is closer to 1.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2} \quad (11)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \widetilde{y})^2} \quad (12)$$

where y_i and $\widehat{y}_i$ are the labeled and predicted support rotations of the i^{th} instance, respectively, $\widetilde{y}$ is the mean support rotations of all labeled instances.

3. Data of BLEVE-induced tunnel support rotations

Since it is not practicable to carry out massive experimental and field tests to obtain sufficient data of BLEVE-induced support rotations of road tunnel structures for training the model, validated numerical models of road tunnels subjected to internal BLEVEs are established in this section to generate the training data. The details are given below.

3.1. Numerical model and data generation

The numerical model of box-shaped road tunnels subjected to internal BLEVEs has been built in the author's previous study [28]. In this section, the numerical model is again employed and parameterized to generate the data of BLEVE-induced support rotations of road tunnels. The overall view of the numerical model is given in Fig. 8. The tunnel structures consist of two traffic cells and a middle escape passageway. The thicknesses of tunnel structures are 1.1 m in the top wall, 1.2 m in the bottom wall, 0.95 m in the side wall, and 0.6 m in the middle wall. The tunnel is set to be 15 m from the left, right and bottom surfaces of the model and has a length of 30 m along its longitudinal direction. The size of solid elements to mesh tunnel structures and soil mass near the tunnel is 100 mm. Other areas of the numerical model are meshed using the gradually increased sizes of solid elements. Two-layer of steel cages are arranged inside the tunnel structures and meshed by 50 mm beam elements. The spacing between adjacent hoop rebars, neighboring longitudinal rebars, and adjoining shear rebars is 200 mm. The diameters and strengths of steel reinforcements as well as the concrete grades of tunnel walls are varied to generate the data of BLEVE-induced support rotations of road tunnels with different structural configurations. The inner widths and heights of the traffic cells and the properties and configurations of the soil mass around the tunnel are also changed to consider their effects on BLEVE-induced support rotations of road tunnel structures.

Table 1
Summary of varied parameters for the simulation to generate the data of support rotations.

Variable (feature)	Notation	Ranges
Tank volume	V_{tank}	20 m^3–40 m^3
Liquid filling ratio	R_{lf}	50 %–80 %
BLEVE location	L_B	Left and right lanes
Inner height of tunnel	TH	4.9 m–5.4 m
Inner width of tunnel	TW	10.4 m–14 m
Structural component of tunnel	LC	Roof, side wall, and floor
Tunnel structural thickness	CT	0.95 m (Side), 1.1 m (Top), 1.2 m (Bottom)
Concrete grades	CS	C30-C60
Diameter of steel reinforcement	R_d	20 mm–40 mm
Strength of steel reinforcement	R_s	400 MPa-600 MPa
Soil cover depth	S_d	0.5 m–10 m
Soil shear modulus	S_s	16 MPa-96 MPa
Soil-cement ratio	S_{cr}	0 %–7 %
Grouted soil thickness	S_{ct}	0–1.5 m
Distance from BLEVE center	D_B	0–9 m

In addition, BLEVEs on the left and right lanes of the left traffic cell from different volumes of Liquified Petroleum Gas (LPG) tanks with varied liquid-filling ratios inside tanks are simulated. Table 1 summarizes varied parameters (i.e., variables) for the simulation to generate the data of support rotations. A total of 115 cases are simulated by adjusting different variables given in Table 1. It is noted that the support rotations of the top wall, side wall, and bottom wall at the first 9 m segments along the longitudinal direction of the tunnel in each case are measured with an interval of 1 m, resulting in a total of 3450 data instances.

*MAT_72REL3, *MAT_24, and *MAT_147 in LS-DYNA are employed as the material models of tunnel concrete, steel rebars inside concrete, and soil mass around the tunnel, respectively. To ensure the accuracy of the built numerical model in predicting the response of road tunnels under internal blast loads, the authors in the previous studies [28,29] have calibrated the models to simulate responses of tunnel structures and the soil mass around the tunnel by using the test of a reinforced concrete (RC) slab subjected to blast loads [30] and the test of soil mass subjected to internal blast loads [31], respectively. To avoid repetition, the tests [30] and numerical models and the partial calibration results are briefly presented in Fig. 9.

3.2. Validation of data correlation

The data quality of support rotations generated by using the validated numerical models of road tunnels subjected to internal BLEVEs is further evaluated in this section by calculating linear relations of the data corresponding to different variables. The Pearson correlation coefficient is selected as the indicator to illustrate the relationships between different variables including 15 input variables given in Table 1 and the output variable (i.e., the support rotation). It is defined as:

$$\rho_{X,Y} = \frac{E[XY] - E[X]E[Y]}{\sqrt{E\left[X^2\right] - (E[X])^2}\sqrt{E\left[Y^2\right] - (E[Y])^2}} \quad (13)$$

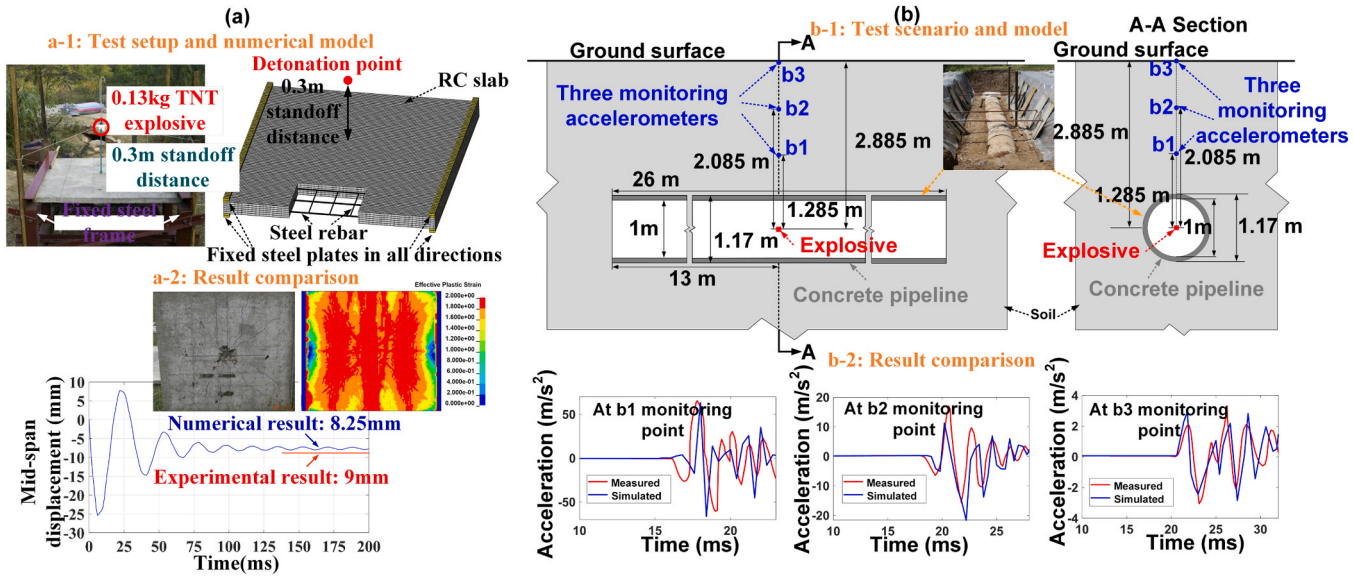


Fig. 9. Model calibration and the corresponding results for (a) tunnel structures [29] and (b) soil mass around the tunnel [28].

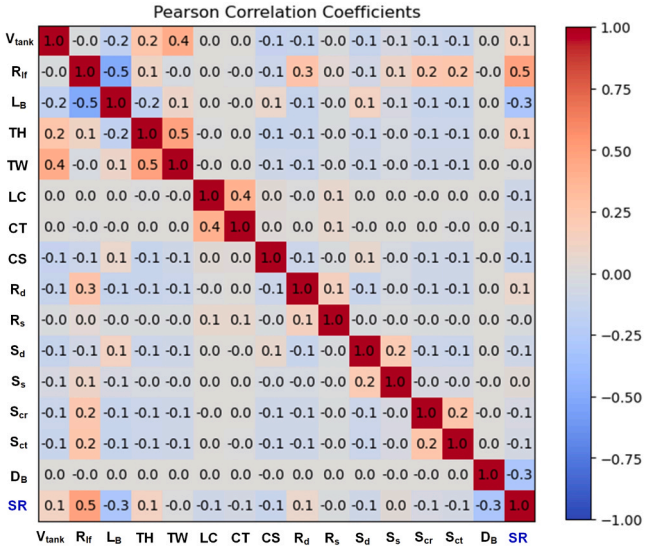


Fig. 10. Pearson correlation coefficients between different variables in given data.

where $\rho_{X,Y}$ is the Pearson correlation coefficient, X and Y are data vectors corresponding to the variables selected from 16 variables (i.e., 15 input variables and one output variable); $E[\bullet]$ is the symbol to calculate the expectation of the data given in the bracket.

Fig. 10 shows the matrix of Pearson correlation coefficients between the data of all variables. It is noted that the variables given in this study are first normalized and then their Pearson coefficients are calculated. That is, the Pearson coefficients of normalized variables are calculated in this study to identify the correlation among different variables. It can be found that the values of Pearson correlation coefficients between the same variable are equal to 1, while the ones between different input variables vary between −0.5 and 0.5, indicating no high correlation between the data of different input variables in the prepared dataset. That is, the selected variables are independent of each other and are not redundant to express the data features, which further validates the reliability of the given data. In addition, it can be seen that the most correlated input parameters with the output variable, i.e., support rotation are the liquid-filled ratios of LPG tank, BLEVE locations, and standoff distances. However, their Pearson correlation coefficients are not greater than 0.5, which indicates the support rotation is not governed by any individual input variable, but by their combination.

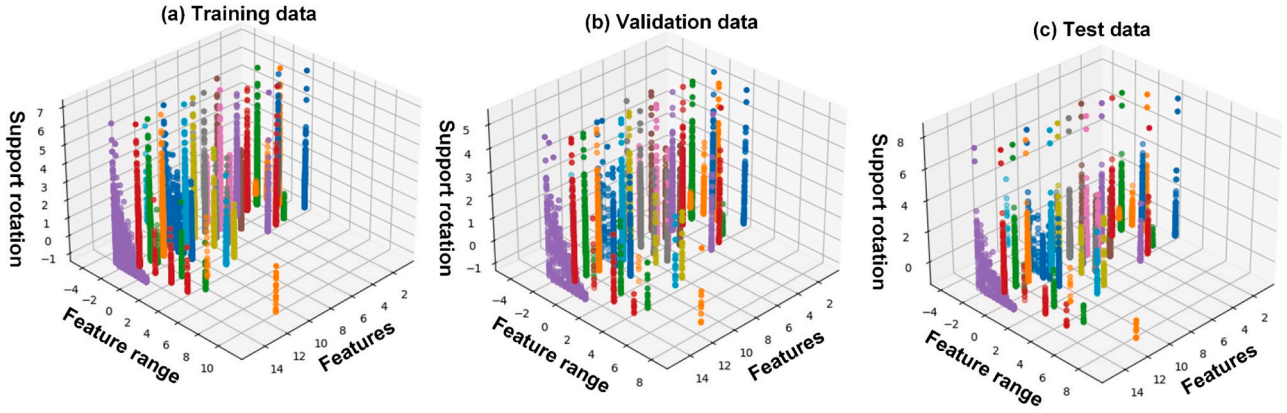


Fig. 11. Division of normalized data of support rotations versus different variables.

Table 2
Search space of hyperparameters for all models and corresponding optimal values.

Model	Hyperparameter	Search space	Distribution	Optimal value
SAMT	Learning rate	[1e^{-5}, 1e^{-2}]	Log uniform	1e^{-4}
	Weight decay	[1e^{-6}, 1e^{-3}]	Log uniform	6.87e^{-6}
	Dropout rate	[0.1, 0.5]	Uniform	0.16
	Token dimensions	[64, 512]	Int uniform	296
	Number of Transformer layers	[1, 10]	Int uniform	5
	Dimension of FFN hidden layer	[1, 32]	Int uniform	9
Non-modified Transformer	Learning rate	[1e^{-5}, 1e^{-2}]	Log uniform	2e^{-3}
	Weight decay	[1e^{-6}, 1e^{-3}]	Log uniform	2e^{-4}
	Dropout rate	[0.1, 0.5]	Uniform	0.26
	Token dimensions	[64, 512]	Int uniform	256
	Number of Transformer layers	[1, 10]	Int uniform	2
	Dimension of FFN hidden layer	[1, 32]	Int uniform	3
ResNet	Learning rate	[1e^{-5}, 1e^{-2}]	Log uniform	1.24e^{-5}
	Weight decay	[1e^{-6}, 1e^{-3}]	Log uniform	1.24e^{-5}
	Dropout rate	[0.1, 0.5]	Uniform	0.25
	Dimension of main block	[1, 10]	Int uniform	1
	Number of ResNet block	[1, 5]	Int uniform	3
	Dimension of hidden layer in each block	[1, 32]	Int uniform	9
MLP	Learning rate	[1e^{-5}, 1e^{-2}]	Log uniform	3e^{-3}
	Weight decay	[1e^{-6}, 1e^{-3}]	Log uniform	4.1e^{-5}
	Dropout rate	[0.1, 0.5]	Uniform	0.42
	Number of MLP blocks	[1, 5]	Int uniform	3
	Dimension of hidden layer in each block	[1, 32]	Int uniform	10

4. Prediction of support rotation of tunnel structure using SAMT

The simulated data in Section 3 are divided into training dataset, validation dataset, and test dataset to build the self-attention-modified Transformer (SAMT) model and evaluate its performance. Three common neural network models, i.e., Multi-layer perceptron (MLP), Residual Network (ResNet) and the non-modified FT-Transformer network are also applied to compare their prediction performance with SAMT. In addition, the datasets with different numbers of variables and instances are generated to evaluate the robustness and capacities of the SAMT in processing different characteristics of datasets. The details are given below.

4.1. Implementation details

To clearly demonstrate the procedure of building trainable machine learning models to predict the support rotations of road tunnels subjected to internal BLEVEs, the detailed implementation of the above-mentioned four deep learning models and their final prediction performance for one run on the collected data are given in this section.

4.1.1. Data division and pre-processing

A total of 115 cases and 3450 instances are simulated using the numerical models described in Section 3.1. The instances are randomly divided, with 60 % for the training dataset, 20 % for the validation dataset, and another 20 % for the test dataset. To minimize the effects of dimensional differences between the data of different variables in all instances on the model performance, all instances are standardized along the variable dimension by subtracting their mean values and dividing their standard deviations. Fig. 11 visualizes the normalized data by showing the distribution of normalized support rotations versus the normalized variable ranges. It can be seen that all normalized variables compactly distribute within certain ranges (i.e., between −4 and 10), which indicates that the data pre-processing technique using the normalization method can eliminate the dimensional difference between the data of different variables.

4.1.2. Hyperparameter tuning and model training

To build a reliable SAMT model, the optimal hyperparameter set for the model needs to be determined based on the given normalized data. In this section, the TPE approach given in Section 2.3 is employed to automatically tune the hyperparameters of the model. The search space of all hyperparameters of the model is given in Table 2. Different hyperparameter sets can be generated by extracting different values from the search space of each hyperparameter. The loss between the labeled and predicted support rotations is calculated as the objective

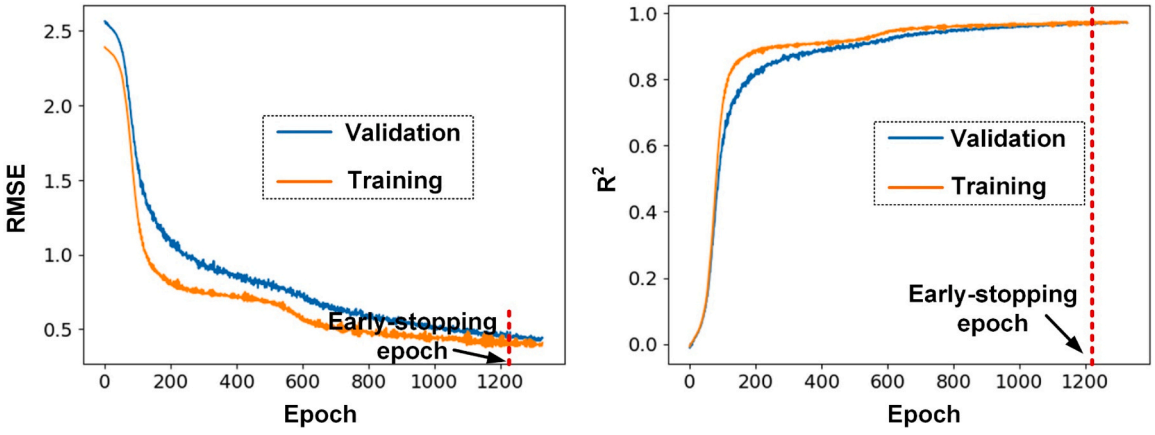


Fig. 12. Variations of RMSE and R^2 with the increased epochs in the process of model training and validation.

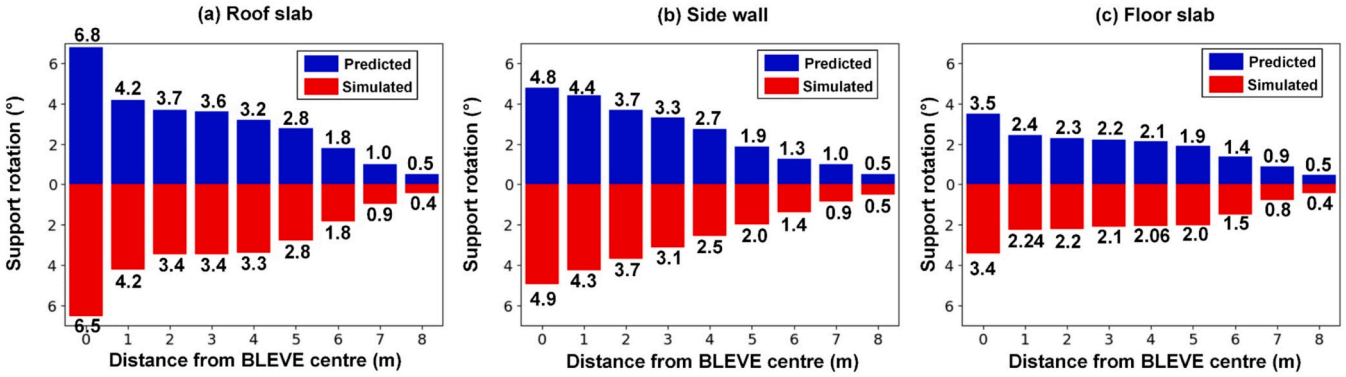


Fig. 13. Comparison of simulated and SAMT-predicted support rotations of different tunnel structures at varied distances from the BLEVE center in a given case.

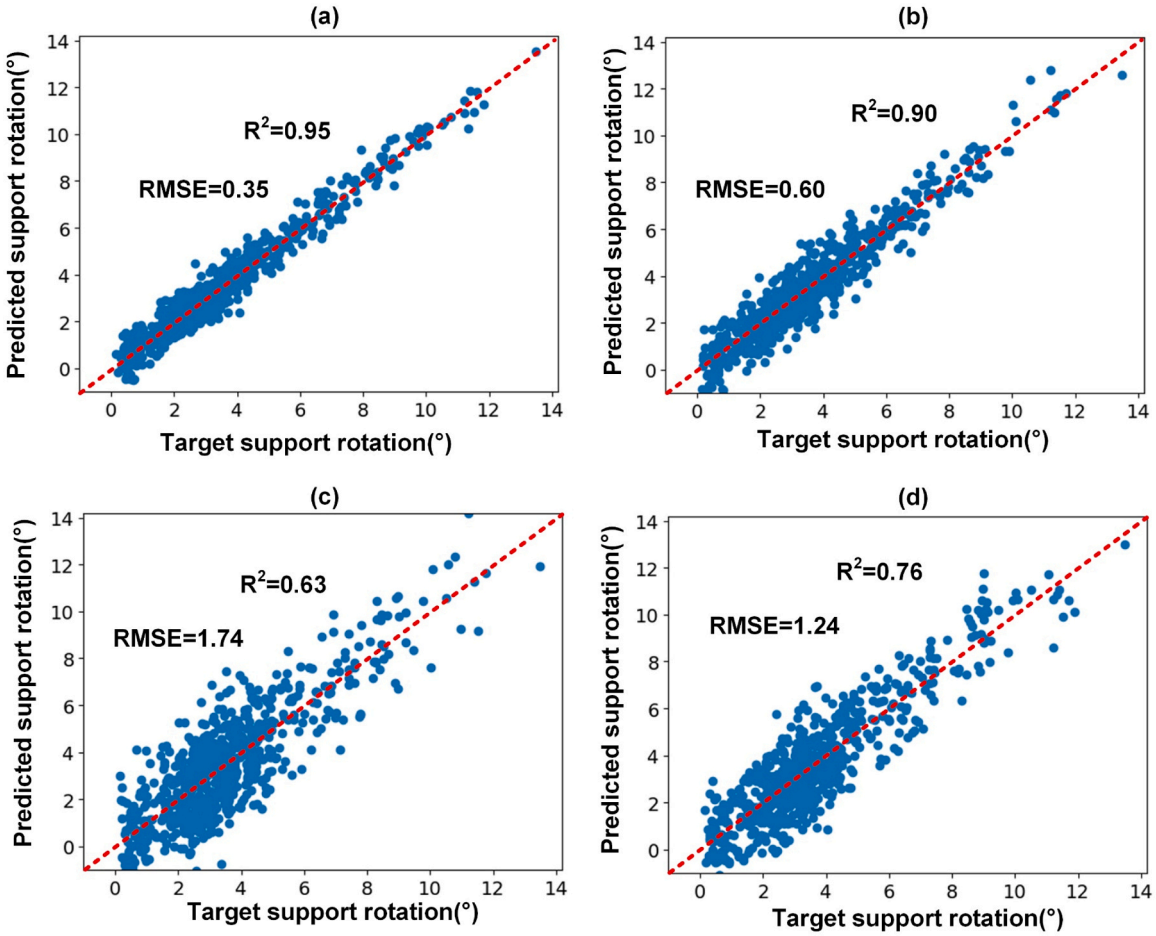


Fig. 14. The scatter plots of simulated support rotations on full test set and predicted support rotations based on (a) SAMT, (b) FT-Transformer, (c) ResNet, and (d) MLP.

function for each hyperparameter set. 100 trials are conducted to determine the hyperparameter set, giving the best result on the validation dataset. It is noted that the search space and optimal hyperparameter sets for the other three networks (i.e., non-modified Transformer, ResNet, and MLP) are also summarized in Table 2.

With the determined optimal hyperparameter set, the training dataset with 60 % of all instances is inputted into the SAMT model to continuously adjust the model parameters with a maximum epoch of 2000. It is noted that the early-stopping technique is adopted in the training process by setting early-stopping patience of 100 epochs. That is, the training is stopped when the performance of the model is not improved in the continuous 100 epochs. The trainable model parameters are picked at the 100^{th} epoch before stopping to generate the final SAMT model for the prediction of BLEVE-induced support rotations of road tunnels. Fig. 12 shows the variation of two indicators (i.e., RMSE and R^2) to evaluate the model performance with the increased epochs. It can be found that variation rates of the values of two indicators on the training dataset and validation dataset significantly reduce at the early stage of the training and gradually become stable with the increased epochs. The optimal epoch is at the 1236^{th} epoch, i.e., the iteration corresponding to the minimum RMRE and the largest R^2. Therefore, the SAMT is built by fixing the weights and biases at the 1236^{th} epoch as the final parameter

values. The training of the other three models (i.e., Non-modified Transformer, ResNet, and MLP) is processed with the same procedure and the same datasets as the SAMT and thus is not given herein to avoid repetition.

4.1.3. Prediction results on test dataset

To evaluate the performance of the trained SAMT model, the data of a case of box-shaped road tunnel subjected to an internal BLEVE extracted from the test dataset is first employed to examine the accuracy of the model. The road tunnel in this case has an inner width of 10.4 m and an inner height of 4.9 m. The tunnel structures consist of a 1.1 m thick roof slab, a 1.2 m thick floor slab, a 0.95 m thick side wall, and a 0.6 m thick mid-wall. The tunnel components in this case are composed of C40 concrete (i.e., the concrete with a compressive strength of 40 MPa) and HRB400 steel reinforcement (i.e., steel with a yield strength of 400 MPa) inside the concrete. A BLEVE due to the rupture of a 40 m^3 liquified petroleum gas tank with a liquid filling ratio of 50 % is assumed to occur on the right lane of the left cell of the box-shaped road tunnel. The simulated support rotations of the roof slab, side wall, and floor slab at the first 8 m segments along the longitudinal direction of the tunnel are compared with the predicted ones using the SAMT, as shown in Fig. 13. It can be found that the maximum difference between the predicted and simulated support rotations is not greater than 0.3 degrees, which indicates that the predicted support rotations match well with the simulated ones.

To further demonstrate the performance of the SAMT, all instances on the test dataset are predicted using the SAMT, and the results are compared with the simulated support rotations. The other three models are also employed to predict support rotations of all instances on the test dataset. Fig. 14 shows the prediction results of the four models on the test dataset and the corresponding RMSE and R^2 between the predicted and simulated support rotations. It can be found that R^2 of the SAMT and non-modified Transformer are respectively 0.95 and 0.9, which are obviously higher than those of MLP and ResNet. The results show that Transformer-based networks are more accurate in predicting the tabular data due to the action of the self-attention mechanism in aggregating variable features to the final inference. It is noted that R^2 of the SAMT is closer to 1 than that of the non-modified Transformer. The RMSE of the SAMT is smaller than that of the non-modified Transformer. The results indicate that the modified self-attention mechanism in the SAMT model is better than the original self-attention mechanism in the FT-Transformer to predict the support rotations of road tunnel structures subjected to internal BLEVEs based on the tabular data.

The proposed model with modified self-attention is developed to predict structural responses based on tabular data consisting of uncorrelated variables. It may be less appropriate to process the data with correlated variables since the modified self-attention mechanism does not capture the cross-correlations of inter-dependent variables. In addition, the computational cost of the Transformer-based neural networks is usually higher than some typical deep learning models, such as Multi-layer Perceptron (MLP) and Residual Network (ResNet), despite the higher prediction accuracy of Transformer-based neural networks. It is worth mentioning that the modification in SAMT is tailored for scenarios with uncorrelated variables, which are commonly encountered in civil and structural engineering data. Therefore, SAMT is expected to be applicable for predicting the response of various civil infrastructures under specific loads. In this study, SAMT is used for predicting tunnel response under BLEVE loads. It can be similarly applied to other types of structures, except that it needs more data and large computer power for application to different large-scale real-world structures.

4.2. Performance evaluation on different characteristics of datasets

To evaluate the adaptability of the SAMT in analyzing tabular data with different characteristics, the model is employed to process the datasets with different numbers of variables and instances. Three other models (i.e., FT-Transformer, MLP, and ResNet) are also utilized on these datasets and their results are compared with those of the SAMT. The details are given below.

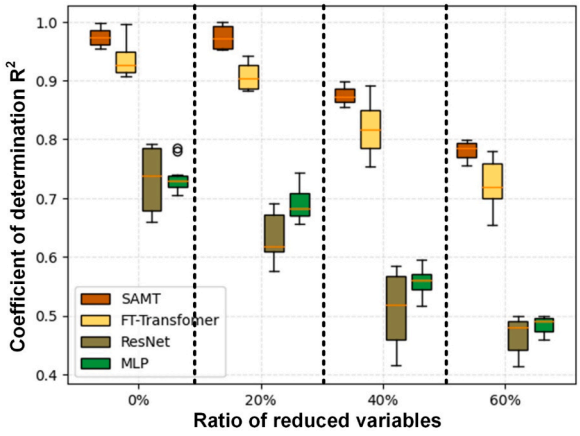


Fig. 15. Coefficients of determination between the simulated support rotations and the support rotations predicted based on the four models obtained from 10 predictions for each number of variables.

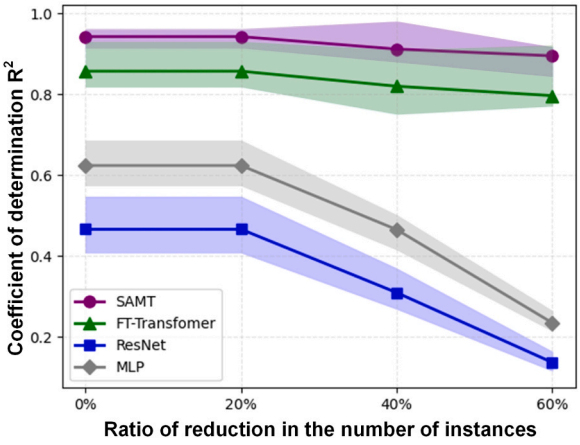


Fig. 16. Coefficients of determination (R^2) between the simulated support rotations and the support rotations predicted by the four models for datasets with reduced number of instances.

4.2.1. Performance on datasets with reduced variables

The original dataset used in the above discussion has a total of 15 variables as illustrated in Section 3.2. In practice, the data with different numbers of variables might be obtained based on various site conditions and explosion scenarios. To test the adaptability of the SAMT model in processing the datasets with different numbers of variables, the datasets with the reduction of 0 %, 20 %, 40 %, and 60 % in the number of original variables are obtained by randomly removing variables from the original dataset. It is noted that in the case of each reduction ratio, 10 trials are conducted to generate 10 datasets, which are obtained by randomly removing the corresponding ratio of variables in each trial. The SAMT is used to process these datasets to generate respective predictions. In addition, the other three models, i.e., non-modified Transformer, MLP, and ResNet are employed to process the datasets with different variable numbers for comparative analysis with the SAMT.

Fig. 15 shows the coefficients of determination (R^2) between the simulated support rotations and the ones predicted based on the four models for the cases with different variable reduction ratios. It can be found that R^2 of the four models all gradually reduces with the reduced variable numbers. The results indicate that more non-redundant

Table 3
The novelty of the proposed SAMT and its comparison with three other models.

Performance		Proposed SAMT	FT-Transformer	ResNet	MLP
Novelty		(1)Mitigating undue dependencies among uncorrelated variables (2) Effectively aggregating global information across all variables	/	/	/
Accuracy (R^2) on full dataset		0.95	0.90	0.63	0.76
Robustness	Average R^2 on dataset with reduced variables (10 trials)	0.98, 0.97, 0.87, and 0.78 for 0 %, 20 %, 40 %, and 60 % reduction	0.92, 0.91, 0.82, and 0.72 for four ratios of reduction	0.74, 0.61, 0.52, and 0.48 for four ratios of reduction	0.73, 0.68, 0.55, and 0.48 for four ratios of reduction
	Average R^2 on dataset with reduced instances (10 trials)	0.95, 0.95, 0.90, and 0.89 for 0 %, 20 %, 40 %, and 60 % reduction	0.85, 0.85, 0.82, and 0.79 for four ratios of reduction	0.46, 0.46, 0.31, and 0.15 for four ratios of reduction	0.62, 0.62, 0.46, and 0.23 for four ratios of reduction

variables included in datasets are advantageous to enhance the accuracies of the four models in predicting the BLEVE-induced support rotations of road tunnels. In addition, as the number of variables reduces, the SAMT and non-modified Transformer (i.e., FT-Transformer) always have larger R^2 values than the other two models, i.e., MLP and ResNet. It is also found that the range of R^2 (i.e., the maximum difference of R^2 for 10 datasets) of the SAMT is smaller than that of the non-modified Transformer model in the case of each reduction ratio. Therefore, it can be concluded that the SAMT model among the four models can achieve the best prediction accuracy of BLEVE-induced support rotations of road tunnels by using the same datasets with different variable numbers.

4.2.2. Performance on datasets with reduced instances

The original dataset with a total of 3450 instances is used in the above discussion. To evaluate the robustness of the SAMT model when faced with varying numbers of instances, additional datasets are generated with reduced instance counts. Specifically, the instance numbers are reduced by 0 %, 20 %, 40 %, and 60 % relative to the full set of 3450 instances, which are achieved by randomly selecting the corresponding ratios of instances from the full dataset. 60 % of each additional dataset is then randomly selected to generate the corresponding training datasets for different reduced instance cases. To ensure statistical validity, 10 trials for each level of instance reduction are conducted, resulting in 10 datasets for each case that have the same number of removed instances but differ in the specific instances that are omitted. The other three models (i.e., non-modified Transformer, MLP, and ResNet) are also employed to process these datasets. The predicted support rotations based on the four models are compared with the simulated support rotations.

Fig. 16 shows the coefficients of determination (R^2) for four models across varying levels of instance reduction, ranging from 0 % to 60 %. It can be found that the SAMT consistently achieves the highest R^2 across all four cases with different reduction ratios, highlighting its superior accuracy in predicting BLEVE-induced support rotations of road tunnel structures compared to the other three models. In addition, Transformer-based models (i.e., SAMT and Non-modified Transformer) maintain a stable prediction performance, yielding R^2 values between 0.9 and 0.98 even as the ratio of instance reduction increases from 0 % to 60 %. In contrast, the MLP and ResNet models experience significant declines in R^2 when the reduction ratios of instances reach or exceed 40 %. Therefore, it can be concluded that SAMT not only offers higher prediction accuracy but also demonstrates greater robustness to reductions in the dataset size when compared to the other models.

5. Conclusion

The present study proposes a self-attention-modified Transformer (SAMT) model designed to predict the support rotations of box-shaped road tunnels subjected to internal BLEVEs. The SAMT model successfully mitigates the issue of forced correlations among uncorrelated variables commonly presented in tabular data, while retaining the beneficial aspects of Transformer-based networks in aggregating relevant information across all variables. The model's effectiveness is validated through 115 cases comprising a total of 3450 instances, generated from numerical simulations of box-shaped road tunnels exposed to internal BLEVEs. The main conclusions are summarized in Table 3 and as follows.

(1) The proposed SAMT can accurately predict the support rotations of box-shaped road tunnels subjected to internal BLEVEs by yielding a coefficient of determination no less than 0.95 in 10 trials for all instances given in this study.
(2) By comparing the coefficients of determination obtained by the SAMT and three other deep-learning models (i.e., FT-Transformer, Multi-layer perceptron (MLP), and Residual network (ResNet)), the SAMT model has higher accuracy in predicting BLEVE-induced support rotations of road tunnels based on the datasets with different numbers of variables.
(3) The SAMT model exhibits exceptional robustness in processing datasets with varying numbers of instances for the prediction of BLEVE-induced support rotations of road tunnels, which further underlines its capability to handle tabular data with uncorrelated variables.

CRediT authorship contribution statement

Hong Hao: Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Qilin Li:** Writing – review & editing, Validation. **Ruishan Cheng:** Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Wensu Chen:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Conceptualization.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgments

The authors acknowledge the financial support from the Australian Research Council (ARC) via Australian Laureate Fellowship (FL180100196).

References

[1] Ministry of Transport of the People's Republic of China. Specifications for design of highway tunnels Section 1 Civil engineering. JTG 2018;3370:1–2018.
[2] Cheng R, Chen W, Hao H, Li J. A state-of-the-art review of road tunnel subjected to blast loads. Tunn Space Technol 2021;112:103911.

[3] van den Berg AC, Weerheijm J. Blast phenomena in urban tunnel systems. J Loss Prev Process Ind 2006;19:598–603.
[4] To CW, Chow WK, Cheng FM. Numerical studies on explosion hazards of vehicles using clean fuel in short vehicular tunnels. Tunn Space Technol 2021;107.
[5] Cheng R, Chen W, Hao H, Li J. Effect of internal explosion on tunnel secondary and adjacent structures: a review. Tunn Space Technol 2022;126:104536.
[6] Ministry of Housing and Urban-Rural Development of the People's Republic of China. Code for design of civil air defence basement; 2005; GB 50038-52005.
[7] Xu L, Chen L, Fang Q, Dong Y. Blast resistance of a folded arch cross-section immersed tunnel subjected to internal explosion. Tunn Space Technol 2022;125.
[8] Liu Z, Wu J, Cao C, Li S, Yan Q. Dynamic performance and damage assessment of a shallow buried tunnel under internal explosion. Tunn Space Technol 2023;133.
[9] De A, Morgante AN, Zimmie TF. Numerical and physical modeling of geofoam barriers as protection against effects of surface blast on underground tunnels. Geotext Geomembr 2016;44:1–12.
[10] Zhou L, Li X, Yan Q, Li S. Blast test and probabilistic vulnerability assessment of a shallow buried RC tunnel considering uncertainty. Int J Impact Eng 2023:104717.
[11] Pan X, Yang TY, Xiao Y, Yao H, Adeli H. Vision-based real-time structural vibration measurement through deep-learning-based detection and tracking methods. Eng Struct 2023;281.
[12] Liu F, Li J, Wang L. PI-LSTM: Physics-informed long short-term memory network for structural response modeling. Eng Struct 2023;292.
[13] Fernandez-Navamuel A, Zamora-Sánchez D, Omella ÁJ, Pardo D, Garcia-Sanchez D, Magalhães F. Supervised deep learning with finite element simulations for damage identification in bridges. Eng Struct 2022;257.
[14] Pathirage CSN, Li J, Li L, Hao H, Liu W, Ni P. Structural damage identification based on autoencoder neural networks and deep learning. Eng Struct 2018;172:13–28.
[15] Hao H, Bi K, Chen W, Pham TM, Li J. Towards next generation design of sustainable, durable, multi-hazard resistant, resilient, and smart civil engineering structures. Eng Struct 2023;277:115477.
[16] Sengupta S, Basak S, Saikia P, Paul S, Tsalavoutis V, Atiah F, et al. A review of deep learning with special emphasis on architectures, applications and recent trends. Knowl Based Syst 2020;194:105596.
[17] Almustafa MK, Nehdi ML. Machine learning prediction of structural response for FRP retrofitted RC slabs subjected to blast loading. Eng Struct 2021;244:112752.
[18] Zhou XQ, Huang BG, Wang XY, Xia Y. Deep learning-based rapid damage assessment of RC columns under blast loading. Eng Struct 2022;271:114949.
[19] Liu YZ, Ren SF, Zhao PF. Application of the deep neural network to predict dynamic responses of stiffened plates subjected to near-field underwater explosion. Ocean Eng 2022;247:110537.
[20] Wang A, Singh A, Michael J, Hill F, Levy O, Bowman SR. GLUE: a multi-task benchmark and analysis platform for natural language understanding; 2018; arXiv preprint arXiv:1804.07461.
[21] Li Q, Wang Y, Shao Y, Li L, Hao H. A comparative study on the most effective machine learning model for blast loading prediction: from GBDT to transformer. Eng Struct 2023;276.
[22] Gorishniy Y, Rubachev I, Khrulkov V, Babenko A. Revisiting deep learning models for tabular data. Adv Neural Inf Process Syst 2021;34:18932–43.
[23] US Department of Defense. Structures to resist the effects of accidental explosions; 2008; UFC 3-340-02.
[24] Goodfellow I, Bengio Y, Courville A. Deep learning. MIT Press; 2016.
[25] Li Q, Wang Y, Li L, Hao H, Wang R, Li J. Prediction of BLEVE loads on structures using machine learning and CFD. Process Saf Environ Prot 2023;171:914–25.
[26] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN. et al. Attention is all you need. Adv Neural Inf Process Syst; 2017: p. 30.
[27] Watanabe S. Tree-structured Parzen estimator: understanding its algorithm components and their roles for better empirical performance; 2023; arXiv preprint arXiv:2304.11127.
[28] Cheng R, Chen W, Hao H, Li J. Response of buried box-shaped road tunnel against internal BLEVE and its damage mitigation. Tunn Space Technol 2023;138.
[29] Cheng R, Chen W, Hao H, Li J. Dynamic response of road tunnel subjected to internal Boiling liquid expansion vapour explosion (BLEVE). Tunn Space Technol 2022;123:104363.
[30] Wang W, Zhang D, Lu F, Wang S-C, Tang F. Experimental study on scaling the explosion resistance of a one-way square reinforced concrete slab under a close-in blast loading. Int J Impact Eng 2012;49:158–64.
[31] Bonalumi P, Colombo M, Comina C, di Prisco M, Foti S, Galli A. Characterization of blast effects on surrounding soil: internal detonations in underground pipes. Appl Mech Mater 2011;82:302–7.