Accident Analysis and Prevention 213 (2025) 107942

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Tunnel crash severity and congestion duration joint evaluation based on cross-stitch networks

Chenzhu Wang, Mohamed Abdel-Aty, Lei Han *

Department of Civil, Environmental & Construction Engineering, University of Central Florida, Orlando, FL 32816, United States

ARTICLE INFO

Keywords:
Tunnel crashes
Injury severity
Congestion duration
Joint Evaluation
Cross-stitch networks

ABSTRACT

Tunnels, with limited space and restricted widths/heights, increase the likelihood of crashes and traffic congestion, where the severity and duration of one often exacerbate the other. However, existing studies mainly conducted separate models, which cannot simultaneously analyze the joint impacts of contributing factors on both crash severity and duration. To address such gap, a joint modeling approach was proposed to explore critical features affecting both crash severity and duration and their joint relationships. A total of 2,454 tunnel crashes in Shanxi, China were collected. Five types of characteristics were extracted as inputs: crash, vehicle, road, environment, and temporal features. Then, a joint cross-stitch network model was proposed with two sub-multilayer perceptron (MLP) networks to establish the relationships between input features with crash severity and duration, respectively. Cross-stitch units were introduced between the two sub-networks to share each model parameters with specific weights, enforcing the sub-networks to simultaneously estimate the coupling relationships between variables and two targets (i.e., crash severity and duration). Compared to existing separate models, the joint cross-stitch network models achieved best performance on estimation of both crash severity (7.0%, 10.2% increase in sensitivity and F1 score, respectively) and congestion duration (3.7% reduction in mean squared error). Though the parameter sharing mechanism, it could simultaneously learn the coupling relationships between contributing factors on both crash severity and duration to offer more reasonable interpretations than separate models. The results indicate that congested traffic conditions significantly increase injury severity, while truck-only, two-vehicle, and multi-vehicle crashes notably prolong congestion duration. Moreover, the joint models exhibited some features presenting inverse effects on injury severity in the separate models. The results enhance our understanding of crashes and congestion in tunnels and inform several recommendations for tunnel management to reduce both crash severity and congestion duration.

1. Introduction

Tunnels are vital components of modern transportation infrastructure, enabling the efficient movement of vehicles through otherwise challenging terrains and densely populated urban areas (Goel et al., 2012). Meeting the growing demand for social and economic efficiency, tunnels play an important role in the roadways in China. By the beginning of 2021, the total length of road tunnels in China has exceeded 220,000 km (NBS, 2022). However, these subterranean conduits often present unique challenges, including the potential for traffic congestion and, unfortunately, the occurrence of crashes that vary in terms of their severity and impact on traffic flow.

Traffic congestion is a pervasive issue and along key transportation corridors in China. Within tunnels, congestion can lead to significant delays, economic losses, and increased environmental pollution (Pucher et al., 2006). With the continued increase in demand for tunnel usage, there is a corresponding need to mitigate congestion and reduce the impact of crashes within these vital components of infrastructure (Alderson et al., 2018). Tunnels are inherently confined spaces with limited egress options, making them vulnerable to crashed or secondary collisions that can cause severe consequences (Chang et al., 2024). These crashes can range from minor fender-benders to catastrophic events, with varying degrees of impact on the safety and well-being of tunnel users.

In addition, the tunnel congestion resulting from crashes increases the likelihood of further incidents such as recurrence of crashes or fires, posing a growing risk to drivers and emergency services (Zheng et al., 2020). Such congestion can also hinder emergency response efforts,

* Corresponding author.
E-mail addresses: chenzhu.wang@ucf.edu (C. Wang), m.aty@ucf.edu (M. Abdel-Aty), le966091@ucf.edu (L. Han).

https://doi.org/10.1016/j.aap.2025.107942
Received 11 July 2024; Received in revised form 24 October 2024; Accepted 29 January 2025
Available online 4 February 2025

exacerbating the severity of injuries from tunnel crashes (Sramek et al., 2019). Moreover, stagnant traffic leads to deteriorated air quality inside tunnels, adversely affecting the health of tunnel users and nearby communities (Meng et al, 2011; Yusuf et al., 2014). Furthermore, delays caused by tunnel congestion can result in significant economic losses due to disrupted freight and passenger flows (Anas and Lindsey, 2011). Therefore, mitigating congestion caused by tunnel crashes is a paramount concern to prevent the recurrence of crashes or other hazardous situations.

Analyzing crash severity within tunnels is crucial for developing effective safety measures and emergency response protocols. Understanding the factors that contribute to the severity of tunnel crashes can inform the design of safer tunnels, the implementation of advanced traffic management systems, and the training of tunnel personnel to respond effectively to emergencies.

While both congestion duration and crash severity are important in their own right, they are often treated as separate issues in tunnel management and research. However, there is a clear interplay between these two factors. Congestion can increase the likelihood of crashes, and the severity outcomes of crashes can affect the duration of congestion. Therefore, addressing these challenges in isolation may lead to suboptimal solutions. To the best of the author's knowledge, only a limited number of studies have focused on analyzing both the duration of congestion and the severity of injuries resulted from crashes, particularly those occurring in tunnels.

The introduction of cross-stitch networks to the field of tunnel analysis represents a cutting-edge approach to addressing these challenges (Misra et al., 2016). This joint model allows for the simultaneous estimation of congestion duration and crash severity, effectively capturing the interdependence between these two outcomes (Abay et al., 2013; Wang et al., 2024b). Moreover, such models can produce more efficient and robust parameter estimates (Wang et al., 2024a), with improved predictive accuracy, by accounting for shared unobserved heterogeneity (Wang et al., 2024c; Yasmin and Eluru, 2018). Thus, a joint model could provide a more holistic understanding the crash dynamics by simultaneously analyzing injury severity and congestion duration. Cross-stitch networks is a novel joint modeling approach in the realm of machine learning, which enables various neural networks to exchange information and learn together (Li et al., 2022; Tian and Bai, 2023; Tukra et al., 2020; Zhang et al., 2022). This integration is done using 'cross-stitch' units, which makes them more efficient and improves their ability to perform different tasks more accurately and quickly. The application of cross-stitch networks spans a wide range of domains, including image recognition (Chen et al., 2022), natural language processing (Collobert et al., 2011), and speech recognition (Huang et al., 2013) by combining the strengths, merging the capacities and boosting the accuracy. Given its significant strength, an integrated approach based on cross-stitch networks that combines the analysis of congestion duration and crash severity are developed in current paper.

With the above-mentioned research gaps, this study aims to jointly analyze the injury severity outcomes and consequent congestion duration caused by tunnel crashes. The primary objective of current study is threefold:

1) A cross-stitch network is developed to obtain a comprehensive understanding of the factors influencing joint crash severity and congestion duration within tunnels.
2) A soft parameter sharing layer is proposed to extract the cross-stitch units of the joint frameworks compared to a hard parameter sharing layer, to achieve a more robust and precise evaluation network.
3) The significant factors affecting the two dimesons would be quantified based on SHAP algorithm, so as to provide effective traffic control strategies and measurements.

The paper is structured into six distinct sections. Subsequent to this introductory section, Section 2 delves into the literature review. This is followed by an elaborate description of data preparation in Section 3. Section 4 meticulously details the proposed methodology, while Section 5 is dedicated to showcasing the results of the experiments conducted. The paper culminates with a comprehensive conclusion and discussion in Section 6.

2. Literature review

2.1. Congestion and injury outcomes in tunnels

Recent studies have delved into traffic safety issues related to tunnels. Du et al. (2014) highlighted how rapid and short-term variations in lighting conditions and visibility at tunnel entrances and exits can significantly affect visual conditions and psychological states. Such impacts can lead to changes in driving performance, a phenomenon known as the "black and white hole effects". Furthermore, Calvi et al. (2012) noted that drivers in tunnels are more prone to fatigued driving due to factors like roadway geometry, tunnel walls, and monotonous environments. Ma et al. (2016) observed that summer seems to reduce the probability of fatal injuries, while crashes during the night are less likely to result in injuries. Pervez et al. (2022) found that the zones around tunnel accesses, entrances, and exits are associated with a higher likelihood of severe crashes. Additionally, Wang et al. (2023) established the non-transferability of data across tunnels of different lengths, observing that the "black and white hole effects" are exacerbated in shorter tunnels, increasing the risk of severe injuries.

Regarding tunnel congestion, recent studies have focused on its causes, managements, and control strategies, particularly concerning air pollution, mobility issues, and safety risks. Bari and Naser (2010) conducted simulations to understand airflow patterns and pollution levels caused by emissions from vehicles in severely congested traffic conditions within tunnels. Tan and Gao (2015) proposed optimal traffic control strategies aimed at managing air quality and mitigating congestion in complex urban vehicular tunnels. Liao et al. (2012) evaluated various traffic control strategies, finding that ramp control could reduce average queue lengths in tunnels by about 18 %. Khetwal et al. (2021) developed a stochastic event simulation model to assess the resilience of tunnel infrastructure, focusing on the function of loss in traffic capacity and its duration. Tympakianaki et al. (2019) identified tunnel congestion patterns using a data-driven approach and evaluated the effectiveness of different tunnel management strategies. Sun et al. (2024) performed a simulation analysis to identify congestion patterns and causes in a hypothetical underground loop during morning and evening peak traffic scenarios.

Overall, a limited amount of research has been conducted to focus on analyzing the duration of congestion caused by crashes in tunnels. It is crucial to investigate the factors contributing to both the severity of injuries and the duration of congestion resulting from these crashes. Conducting such analyses could offer new insights into tunnel management strategies aimed at mitigating congestion and reducing the severity of crashes.

2.2. Cross-stitch networks

In this study, the joint evaluation of crash severity (categorical variables) and congestion duration (continuous variables) can be seen as a multi-task modeling. Traditional economic algorithms, limited by their structure and assumptions, struggle to simultaneously model the complex relationships between both categorical and continuous target variables (Asudani et al., 2023; Egger et al., 2021; Karniadaks et al., 2021). Therefore, they exhibit significant limitations for multi-task processing, which requires sophisticated feature extraction (Liu et al., 2022; Tan et al., 2022). To solve such problem, deep learning models have been widely used for their scalability and flexibility in analyzing multiple tasks, often yielding higher prediction performance (Nweke et al., 2018; Wang et al., 2024; Zou et al., 2024). This advantage is especially evident in scenarios involving high-dimension data and complex interdependencies among variables (Schulz et al., 2020).

Multi-task learning is a potent deep learning framework that enhances the performance of predictive models by leveraging shared knowledge across different tasks. For example, Vandenhgende et al. (2020) introduced multi-task learning methods to analyze the estimation of architectural and optimization-based strategies. Additionally, Zhao et al. (2019) developed a multiple relational attention network framework, further advancing the field of multi-task learning. Misra et al. (2016) introduced a novel sharing unit, the cross-stitch unit, offering an optimal blend of shared and task-specific representations. Following this, Beljaards et al. (2020) applied these units to connect networks for image registration and segmentation, achieving performance superior to single-task networks. Numerous studies have validated that the cross-stitch models outperformed other frameworks. For instance, Thanasutives et al. (2021) further enhanced the cross-stitch modules in physics-informed neural networks for shared representation learning, which significantly improves the network accuracy and robustness. Additionally, Luo et al (2023) developed a co-attention learning framework across time and frequency for fault diagnosis in rolling machinery, with experimental results and comprehensive analysis demonstrating its superiority in terms of diagnostic accuracy and adaptability.

To sum up, we propose a cross-stitch network for the joint evaluation of crash severity and congestion duration as a multi-task modeling approach. It can simultaneously analyze the influencing factors for both injury severity and congestion duration from tunnel crashes. Additionally, traditional separate models and a hard parameter sharing strategy is employed to compare the estimations from the developed cross-stitch networks.

3. Data preparation

The crash data occurring in different types of tunnels at highways are collected from specific reports recorded by the expressway management department in Shanxi Province, covering a seven-year period (2012–2018). The data includes 3024 crashes that occurred in a total of 216 tunnels, and the duration of congestion resulted by the crashes are also collected. The congestion duration refers to the interval between the time of crash occurrence and its clearance, as determined by the identification systems of the tunnel monitoring infrastructure. To ensure data accuracy, records with a duration of less than 10 min or more than 480 min (8 h) were excluded, as they may reflect erroneous or incorrect recordings. Additionally, the data was manually reviewed, and any duplicates or entries containing missing variables were excluded to ensure data quality. As a result, the final modeling dataset consists of 2,454 tunnel crashes.

It is noted that the frequencies for property damage only, minor injury, and severe/fatal injury are 2094 (85.33 %), 301 (12.27 %), and 59 (2.40 %), respectively. To mitigate biased estimation due to the lower number of observations for severe/fatal injuries observations, minor and sever/fatal injuries were consolidated into a single injury category: 'Injury', representing 360 cases (16.90 %). And the mean congestion duration is 92.56 min.

Then, the variables extracted from the data are divided into five categories: crash, vehicle, road, environment, and temporal characteristics. The tunnel length is a numeric variable while other variables are binary. The description statistics of these variables are presented in Table 1.

To eliminate the influence of extreme values on the data and regularize the data distribution, the logarithmic values of the crash congestion duration and tunnel length were used. The original and log-transformed distributions of crash congestion duration and tunnel length is shown in Fig. 1. The two original variables show obvious skew distributions. While after the log-transformation, their distribution approaches normal distributions, which would help eliminate the learning bias caused by extreme values.

Table 1
Description statistics of variables in the tunnel crashes.

Category	Variables	Values/Codes	Frequency	Ratio
Target variables	Congestion duration	10–480 (min)		
	Crash severity	1-Property damage only	2094	85.33 %
		2-Injury	360	14.67 %
Crash characteristics	Crash type	1-Rear-end	1347	54.89 %
		2-Sideswipe	120	4.89 %
		3-Roll over	140	5.70 %
		4-Collide with fixtures	232	9.45 %
		5-Fire	78	3.18 %
		6-Others	537	21.88 %
Vehicle characteristics	Involved vehicle types	1-Car only	827	33.70 %
		2-Truck only	1183	48.21 %
		3-Car and Truck	444	18.09 %
	Vehicle number	1-Single vehicle	657	26.77 %
		2-Two vehicles	1434	58.43 %
		3-Multi vehicles	363	14.79 %
Road characteristics	Lane number	2-Two lanes	2237	91.16 %
		3-Three lanes	217	8.84 %
	Pavement	1-Bituminous	2076	84.60 %
		2-Cement	378	15.40 %
	Tunnel length	50–3060 (m)		
Environment characteristics	Weather condition	1-Clear	1321	53.83 %
		2-Cloudy	615	25.06 %
		3-Foggy	23	0.94 %
		4-Rainy	246	10.02 %
		5-Snowy	249	10.15 %
	Traffic flow status	1-Free flow	1465	59.70 %
		2-Slow flow	518	21.11 %
		3-Jam flow	471	19.19 %
	Surface	1-Dry surface	2188	89.16 %
		2-Wet surface	247	10.07 %
		3-Snowy or icy surface	19	0.77 %
Temporal characteristics	Season	1-Spring	508	20.70 %
		2-Summer	725	29.54 %
		3-Autumn	715	29.14 %
		4-Winter	506	20.62 %
	Weekdays	1-Monday	339	13.81 %
		2-Tuesday	356	14.51 %
		3-Wednesday	315	12.84 %
		4-Thursday	349	14.22 %
		5-Friday	350	14.26 %
		6-Saturday	374	15.24 %
		7-Sunday	371	15.12 %
	Day time	1-Early morning	274	11.17 %
		2-Morning	814	33.17 %
		3-Afternoon	939	38.26 %
		4-Evening	427	17.40 %

4. Methodology

In this section, we present a cross-stitch multilayer perceptron model (cross-stitch MLP) for jointly evaluating the tunnel crash severity and duration. As a typical multi-task learning of crash severity classification and duration regression, the proposed cross-stitch units can help the networks sharing their model parameters among the two tasks, therefore extracting the potential common impact factors and representations to crash severity and durations. For comparation, another multi-task model based on hard parameter sharing and existing separate models for only crash severity or duration evaluation were developed as baselines.

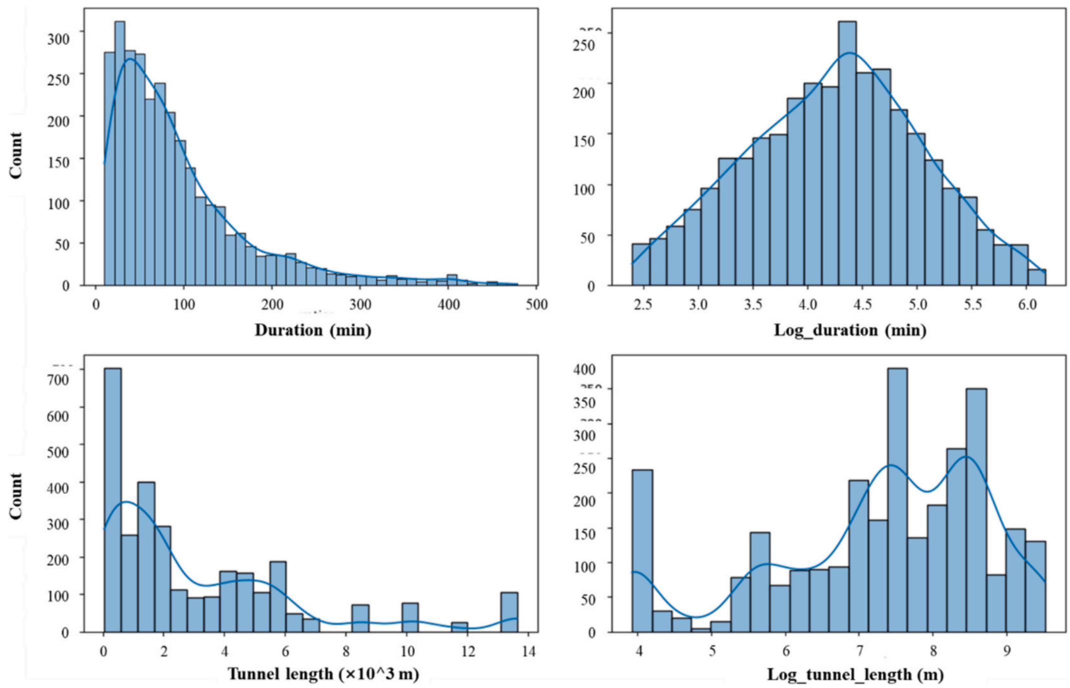

Fig. 1. Original and log-transformed distributions of duration and tunnel length.

4.1. Cross-stitch multilayer perceptron network

4.1.1. Overall structure

Fig. 2 illustrates the structure of the proposed cross-stitch MLP model. First, the categorical variables x_i (e.g., season, weekday, and weather) was transformed into one-hot coding vectors and concatenated together with the numeric variable (tunnel length) as model inputs. The input vector was then fed into two sub-networks for crash severity classification and duration regression, respectively. The two sub-networks have the same networks with three MLP layers, in which the two first layers were connected through two cross-stitch units. The cross-stitch units could help regularize both tasks by learning and enforcing shared representations by combining MLP parameters. Therefore, the sub-network 1 could get both direct parameters from the crash severity classification task and shared parameters (through cross-stitch units) from duration regression task. Also, the sub-network 2 could get the knowledges from both tasks through cross-stitch units. Finally, the last layer features of sub-network 1 need to be passed through a sigmoid function to get the predicted crash severities. The final output of sub-network 2 are directly represent the evaluated crash duration value.

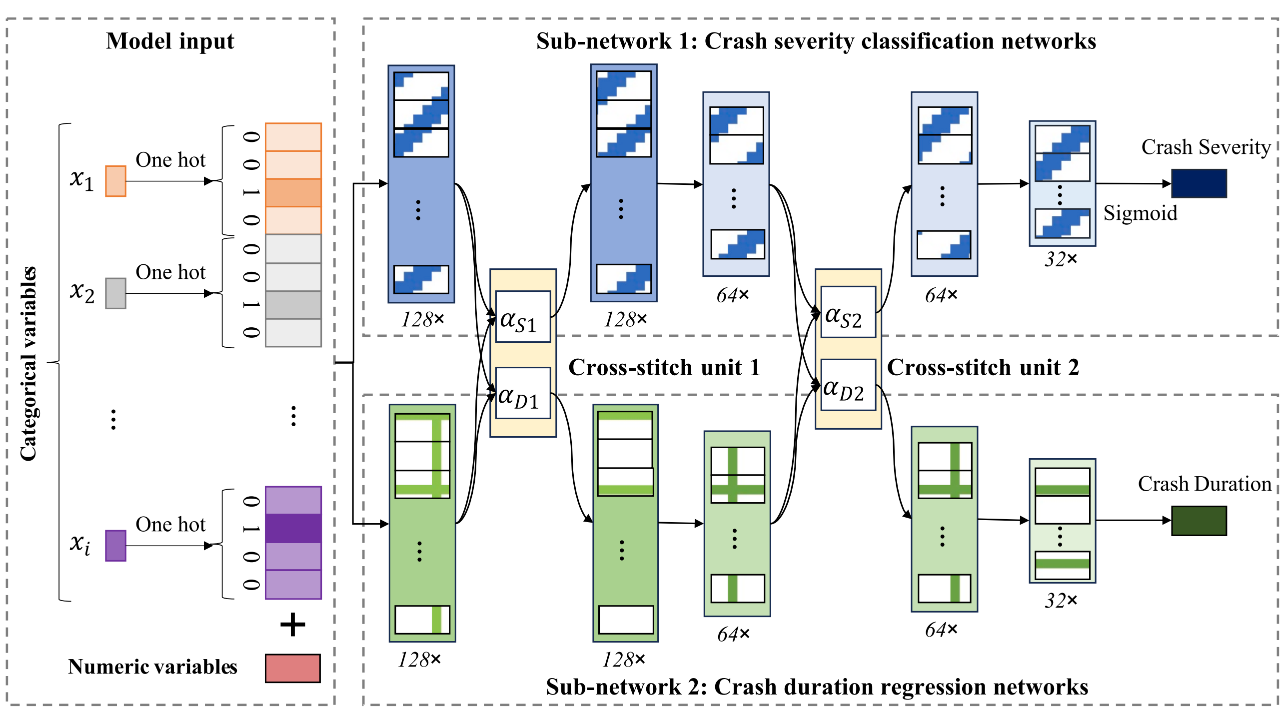

Fig. 2. Structure of the cross-stitch MLP model.

4.1.2. Cross-stitch units

Proposed by Misra et al. (2016), the cross-stitch approach aims to learn shared representations from multiple tasks with learnable weights. In this study, the cross-stitch units are integrated between the MLP layers of the two sub-networks, enhancing them share the common features between the crash severity classification and duration evaluation tasks.

Considering a case of two-task learning with tasks A and B, they share the same inputs and two networks have been trained separately for each task. The cross-stitch unit can combine these two networks into a multi-task network and automatically learn how much sharing is needed for each task. As illustrated in Fig. 3, the cross-stitch unit is adopted to calculate a linear combination of the input shared features/representations. Given two task features x_A, x_B from layer l for both the tasks, the linear combinations $\widetilde{x}_A$, $\widetilde{x}_B$ of both the input task features can be learned, which would be feed into the next layers. This linear combination is parameterized using α. Specifically, at location (i,j) in the networks, the linear combinations $\widetilde{x}_A^{ij}$ and $\widetilde{x}_B^{ij}$ can be calculated by:

$$\begin{bmatrix} \widetilde{x}_A^{ij} \\ \widetilde{x}_B^{ij} \end{bmatrix} = \begin{bmatrix} \alpha_{AA} & \alpha_{AB} \\ \alpha_{BA} & \alpha_{BB} \end{bmatrix} \begin{bmatrix} x_A^{ij} \\ x_B^{ij} \end{bmatrix} \tag{1}$$

Though the cross-stitch operation in each layer l, the network can decide to make certain sharing weight between those tasks by setting α_{AB} or α_{BA}. For instance, if α_{AB} and α_{BA} are zero, it means there is no sharing between task A and B. On the contrary, if α_{AB} and α_{BA} are assigned higher values, a more shared representation is established between them.

Since the cross-stitch unit is a linear combination, their partial derivatives for loss L with tasks A, B are computed as for their back-propagating through model training:

$$\begin{bmatrix} \frac{\partial L}{\partial x_A^{ij}} \\ \frac{\partial L}{\partial x_B^{ij}} \end{bmatrix} = \begin{bmatrix} \alpha_{AA} & \alpha_{AB} \\ \alpha_{BA} & \alpha_{BB} \end{bmatrix} \begin{bmatrix} \frac{\partial L}{\partial \widetilde{x}_A^{ij}} \\ \frac{\partial L}{\partial \widetilde{x}_B^{ij}} \end{bmatrix} \tag{2}$$

$$\frac{\partial L}{\partial \alpha_{AB}} = \frac{\partial L}{\partial \widetilde{x}_B^{ij}} x_A^{ij}, \frac{\partial L}{\partial \alpha_{AA}} = \frac{\partial L}{\partial \widetilde{x}_A^{ij}} x_A^{ij} \tag{3}$$

The α_{AB}, α_{BA} are denoted by α_D, the different task values because they weigh the activations of another task. Likewise, α_{AA}, α_{BB} are denoted by α_S, the same-task values since they weigh the activations of the same task. During the model training, the values of α_D and α_S are trained to be automatically varying. Therefore, the unit can freely move between shared and task-specific representations and choose a middle ground if needed.

4.1.3. Objective function

For model training, the objective is to minimize the combined loss of the two tasks. Given the crash severity is an unbalanced classification task (Injury: Property damage only = 360: 2094), the α-wight focal-loss (Ross and Dollár, 2017) was used by the follows:

$$L_A = -\frac{1}{N}\sum_{k=1}^{N} \alpha y_A^k \log(p_A^k) + (1-\alpha)(1-y_A^k)\log(1-p_A^k) \tag{2}$$

where N is the total number of samples, p_A^k is the output probability of sub-network 1 for the k-th sample, y_A^k is the true crash severity label, α is the weight for the Injury samples. Given the unbalanced ratio is more than 5, so α is set 5/ (5 + 1) = 0.84 to improve the training weights during the fitting of cross-stitch MLP model.

For the regression task of crash duration, the Mean Squared Error (MSE) loss was utilized by the follows:

$$L_B = MSE = \frac{1}{N}\sum_{k=1}^{N} \left(y_B^k - \widehat{y}_B^k\right)^2 \tag{3}$$

where $\widehat{y}_B^k$ is the output of sub-network 2, estimated crash duration, for the k-th sample, y_B^k is the true crash duration value.

Finally, the combined loss function is calculated:

$$L = w_A L_A + w_B L_B \tag{4}$$

where w_A, w_B are the weights for L_A, L_B, respectively. The weights are set as two trainable parameters, initially set to 0.5, allowing the network to balance the two losses to achieve the best model performance.

4.2. Comparation models

4.2.1. Hard parameter sharing MLP

Different from the cross-stitch MLP (soft parameter sharing method), a hard parameter sharing strategy is another widely used multi-task learning method (Vandenhgende et al., 2020). To compare the impact of different feature sharing approaches to the model performance, a MLP with hard parameter sharing layers was developed as shown in Fig. 4. To eliminate the influence of irrelevant factors, the model input, sub-network structures, and loss function are consistent with the cross-stitch MLP model. The difference is that the former cross-stitch units were replaced by a set of hard-parameter sharing layers. Specifically, the model input would be fed into two shared layers to extract the potential common features and representations for both tasks. Then, the sharing features (i.e., output of shared layers) would be concatenated with the outputs of first layers in sub-network 1 and 2 as the inputs to the next MLP layers. Thus, the shared features can be a part of the model parameters to both crash severity classification and duration regression task. Though such hard parameter sharing, the two sub-networks can directly learn the shared features and knowledge to improve their tasks. In addition, the number of neural units in the first layer in both sub-

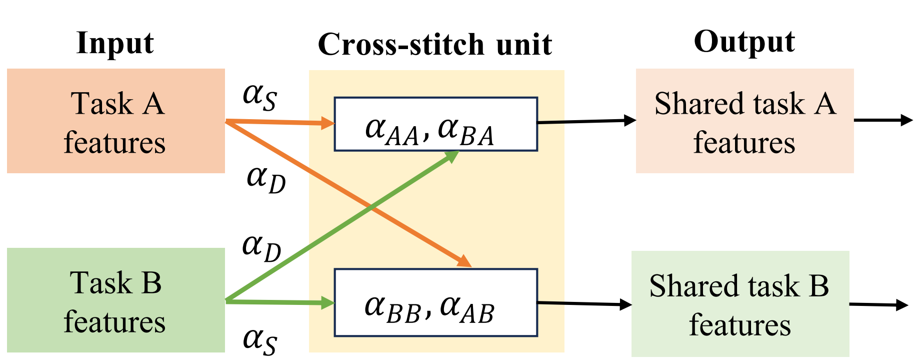

Fig. 3. Illustration of the cross-stitch unit.

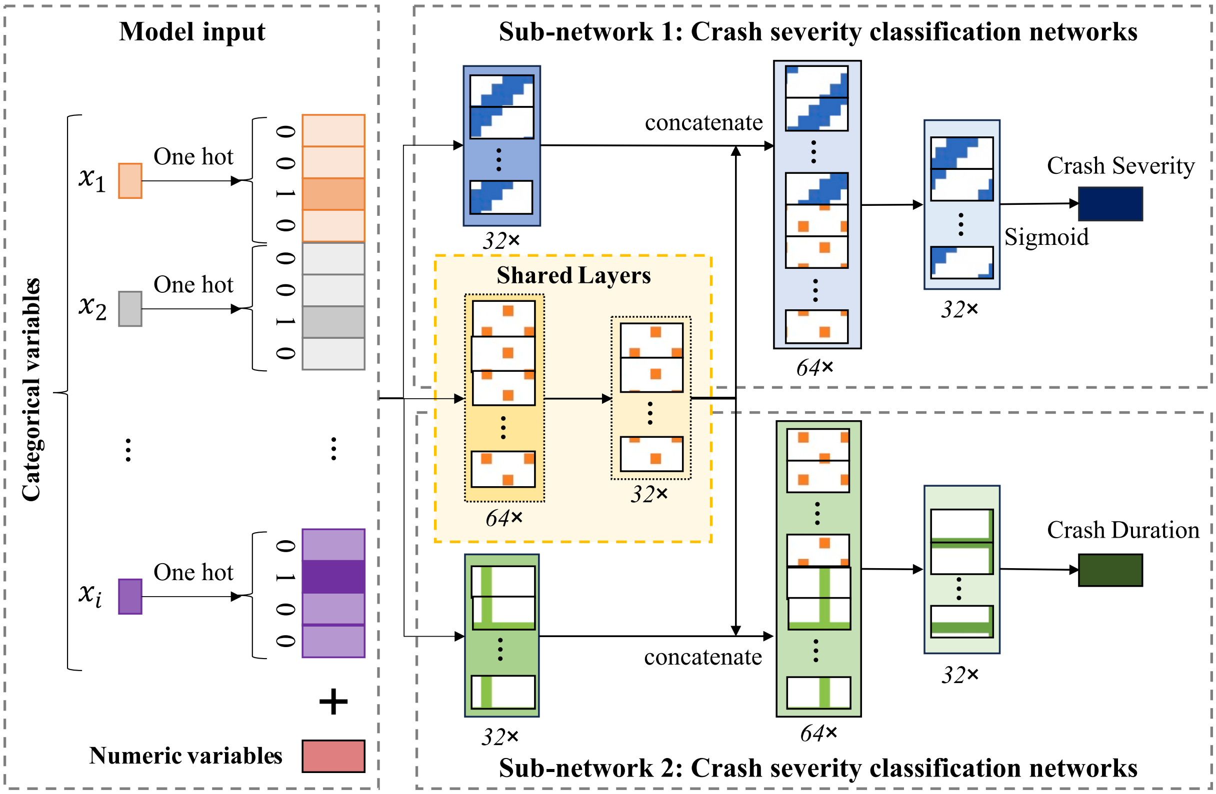

Fig. 4. Structure of the MLP with hard parameter sharing layers.

networks are reduced from 128 to 32 to prevent the model overfitting.

4.2.2. Separate models

To demonstrate the advantages brought by feature sharing of multi-task learning, the separate models for each individual task were constructed as the baselines. For each task, the widely used models in existing crash severity and duration studies were established including the Logistic Regression (LogR, for severity) and Linear Regression (LinR, for duration), Random Forest (RF), Support Vector Machine (SVM), XGBoost, and MLP. The structures of MLPs are the same as the corresponding sub-networks in the cross-stitch MLP model as shown in Fig. 5. Consistently, the cross-entropy loss was used for the crash severity classification MLP and the MSE loss for the crash duration regression MLP.

4.3. Model evaluation metrics

Two kinds of metrics are employed to evaluate the model performances on two tasks:

(1) For the crash severity classification, accuracy, sensitivity, F-1 score, and the Area Under ROC curve (AUC) were used, which can be calculated by Equation (5) – (7) given the classification threshold of 0.5. Accuracy represents the ratio of correctly classified samples to all samples. However, in this unbalanced crash dataset (injury crashes only account for 16.9 %), accuracy is heavily influenced by the majority class classification. Therefore, sensitivity, F-1 score, and AUC are more critical metrics for evaluating classification performance in such cases. Specifically, sensitivity measures the ratio of correctly classified positive samples (i.e., injury crashes), the F1 score is the harmonic mean of sensitivity and precision. AUC measures the area under the Receiver Operating Characteristic curve; a higher AUC indicates better model performance in distinguishing between the two classes.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{5}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{6}$$

$$F1score = \frac{TP}{TP + 0.5(FP + FN)} \tag{7}$$

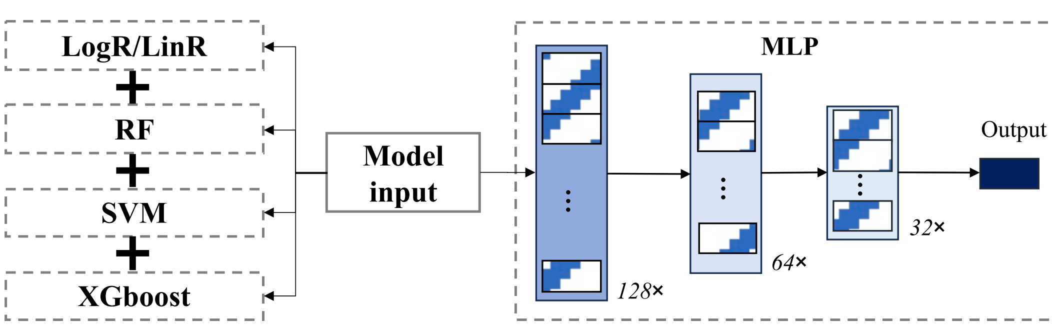

Fig. 5. Structure of the separate MLP models.

where *TP*, *TN*, *FP* and *FN* are given in the following confusion matrix of Table 2.

(2) For the crash duration regression, MSE an Mean Absolute Error (MAE) were the two commonly used metrices, which can be calculated by Equation (3) and (8).

$$MAE = \frac{1}{N}\sum_{k=1}^{N}\left|y_B^k - \hat{y}_B^k\right| \tag{8}$$

5. Results and discussion

This section presents the results of the proposed cross-stitch MLP model and its model metrics comparation with other baselines. To investigate the impacts of different crash factors to the crash severity and duration, the SHapley Additive exPlanations (SHAP) approach (Antwarg et al., 2021) was employed to explain the model results.

5.1. Cross-stitch MLP model results

To develop the cross-stitch MLP, the modeling dataset was divided into the training dataset and test dataset. More specifically, a total of 2118 crash records (70 % of the whole dataset) were randomly selected from the 7-year dataset for model training. The remaining 906 samples (30 % of the whole dataset) was utilized to test the model performance. During the model training, Grid Search method was chosen to find the optimal parameter settings for the models. The final hyperparameters of the model are tuned as shown in Table 3.

Fig. 6 illustrates the model results of the proposed cross-stitch MLP model on the test dataset. In the crash severity classification task, the model can correctly identify 486 non-injury crashes and 69 injury crashes. The accuracy, sensitivity, F1 score are 0.754, 0.566, and 0.421, respectively. Although 75.4 % of samples are correctly classified, the model can successfully identify 56.6 % of injury crashes, which are the focus of crash analysis and prevention. And the AUC of model is 0.741, showing a relatively good model performance. For the crash duration regression task, the predicted values are very close to the true values as shown in the Fig. 6 (c). The overall MSE and MAE are 0.437 and 0.527, respectively, indicating that the model can precisely estimate the duration of a tunnel crash at a relatively small error level. Moreover, it becomes clear that the predicted and true data exhibit a highly correlated trend when these data are restored to the normal scale (Fig. 6(d)). According to the absolute error between the predict and true value, such data can be divided into three level: Good in green (error $\leq$ 30 min); Medium in orange (30 min $<$ error $\leq$ 60 min); and Bad in red (error $>$ 60 min). More than 50 % of samples has been correctly predicted within 30 min error. Approximately 26 % of the samples had errors within 60 min, which is relatively acceptable, as conservative estimates for crash rescue time are often preferred. About 21 % of the samples had errors greater than 60 min. Most of these larger errors occurred in cases of long traffic congestions, possibly influenced by unobserved factors (e.g., rescue efforts, crash location in the tunnel), which were not included in our modeling features. Overall, the above results illustrate that the proposed model exhibits good crash severity evaluation and excellent congestion duration assessment capabilities.

5.2. Model performance comparison

To demonstrate the advantages of the proposed multi-task learning model to the existing separate models, comparative experiments were conducted based on the modeling dataset. The separate models include the RF, SVM, XGBoost and MLP models, which can only achieve one specific task on crash severity classification or duration regression. While the joint model can directly evaluate both the crash severity and duration. In the joint models, the "Hard sharing" is the hard parameter sharing MLP and the "Soft sharing" is the cross-stitch MLP. To ensure the rationality of model comparison, 10-fold cross validation were employed. To be specific, the data were randomly split into 10 groups. For each training iteration, 9 out of 10 the groups were used for training and the remaining one served as the testing set. Therefore, this process allows us to calculate the mean and standard deviation of the model evaluation metrics across the test sets, ensuring a more reliable model comparison and addressing the challenge of unstable results posed by a small sample size. Table 4 presents the results of different models, displaying both the mean and standard deviation of the metrices during 10-fold cross validation. The results highlight several key points:

(1) For the crash severity evaluation, the sample size of non-injury crashes is much larger than the injury crashes (more than 4 times) in this study. Therefore, the sensitivity and F1 score can better reflect the model classification performance under such unbalanced dataset condition. The proposed cross-stitch MLP provided the highest sensitivity of 0.566 and F1 score of 0.421, followed by the hard parameter sharing MLP and MLP. In addition, the cross-stitch MLP shows the highest AUC of 0.741, indicating that it has the best model overall performance.
(2) For the crash duration evaluation task, the proposed cross-stitch MLP provided the lowest MSE of 0.437 and MAE of 0.527. The worst model is the XGBoost model whose MSE and MAE are more than 0.540. The hard parameter sharing MLP and separate MLP model also show a low MSE and MAE, which indicates that the MLP model can better capture the relationships between the crash influence factors and crash congestion durations based on their powerful non-linear fitting capabilities.
(3) Overall, the model performance of the proposed cross-stitch MLP outperformed the separate models in both tasks. Compared to the best metrics in the separate models, the sensitivity and F1 score for crash severity classification have been significantly improved by (0.566–0.529)/0.529 = 7.0 % and (0.421–0.382)/0.382 = 10.2 %, respectively. Meanwhile, the MSE and MAE for the crash congestion duration evaluation dropped by (0.437–0.454)/0.454 = -3.7 % and (0.527–0.540)/0.527 = -2.5 %, respectively.

Table 2
Confusion matrix for crash severity classfication.

True Condition	Prediction Result Injury crash	No injury crash
Injury crash	True Positive (TP)	False Negative (FN)
No injury crash	False Positive (FP)	True Negative (TN)

Table 3
The parameters of cross-stitching MLP model.

Parameter		Tuning range	Selected value
Input	Dimension of inputs	–	49
MLP structure	Number of parameters in 1st layer	32, 64, 128	128
	Number of parameters in 2nd layer	32, 64, 128	64
	Number of parameters in 3rd layer	32, 64, 128	32
Weights for loss*	Weights for severity classification loss	–	0.5
	Weights for duration regression loss	–	0.5
Model training	Optimizer	SGD, Adam, RMSprop	Adam
	Learning rate	1×10^{-4}, 5×10^{-5}, 1×10^{-5}	5×10^{-5}
	Batch size	32, 64, 128, 256	256
	Epoch	50, 100, 150, 200	200

*: These two parameters will be automatically learned and changed during model training.

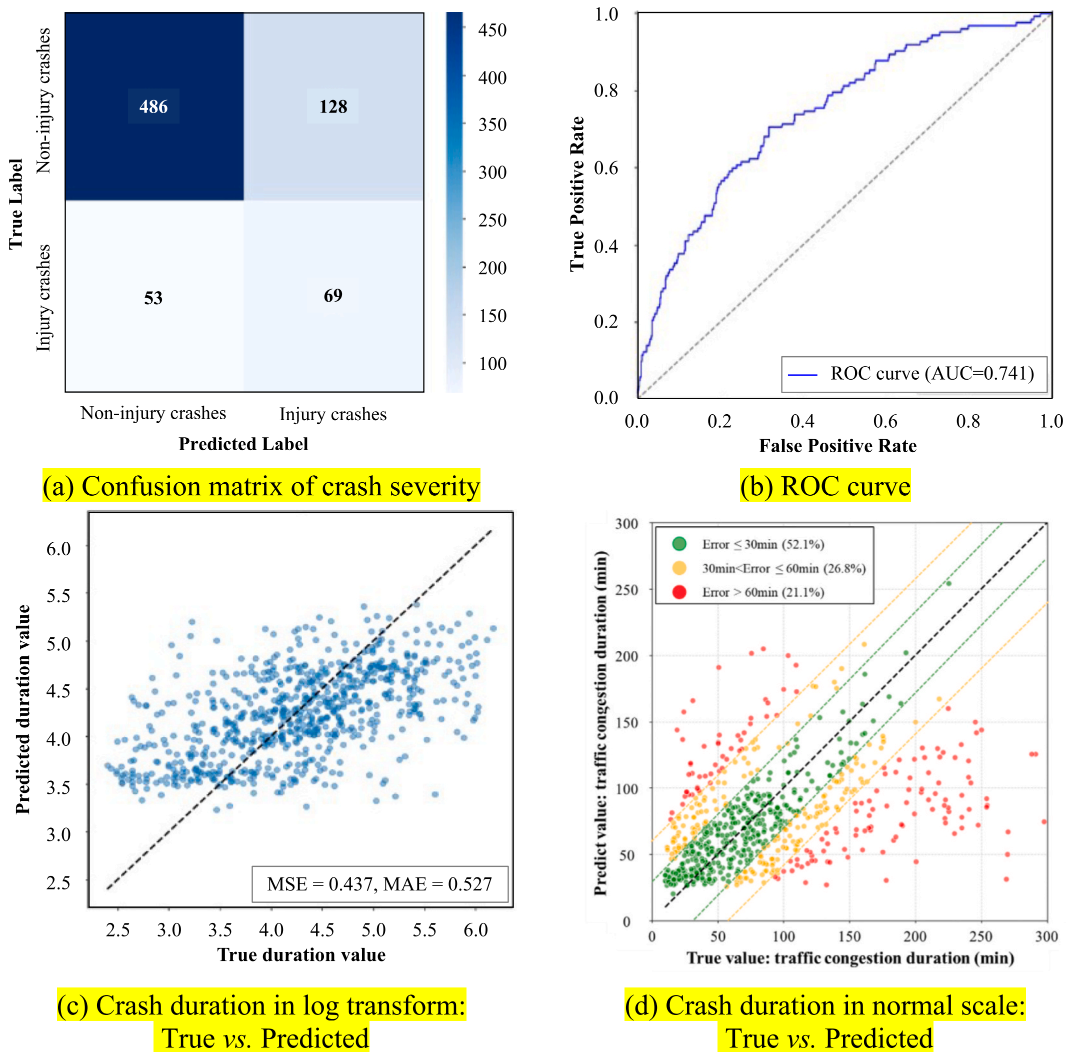

Fig. 6. Model results of the cross-stitch MLP on the test dataset.

The above results show that joint modeling of crash severity and duration can effectively capture the potential common characteristics of crash influences, thereby improving their respective outputs. Among the two joint modeling methods, the cross-stitch MLP achieved the best results. It demonstrates that rather than forcing to apply the same shared features to different tasks, the cross-stitch units could automatically fit the shared weights, thereby extracting the useful features from another task without affecting its own model outputs.

5.3. Model result Explanation

5.3.1. Analysis of the cross-stitch MLP model results

To analyze the relationships between the crash influence factors and crash severity and duration, the widely used SHapley Additive exPlanations (SHAP) approach (Lundberg and Lee, 2017) was employed to explain the results of the proposed Cross-stitch MLP model. It is worth noting that given most of variables (e.g., traffic flow status, crash type) are categorical, they had been transformed to corresponding dummy variables as model inputs. For example, the variable "Crash type" was transformed into six 0/1 variables (i.e., Crash type_1, 2, …, 6), which represent one specific crash type respectively. If the value of variable "Crash type_1" is 1, it means that the crash type is rear-end in this case. Otherwise, the value of variable "Crash type_1" is 0, indicating a non-rear-end crash.

Fig. 7 presents the importance rank of these variables (top-20) for the crash severity and their SHAP values. The means of the absolute SHAP values of each variable were calculated to represent their importance to the crash severity (left side of Fig. 7). While the detailed distributions of SHAP values were also plotted to illustrate their impacts on the crash severity (right side of Fig. 7). From the results, it can be seen that:

(1) The most important variable for crash severity is the "Traffic flow status_1 (free flow)". Its red feature value points concentrate on the negative SHAP value portion, meaning that the contribution to the crash severity being true (i.e., injury crash) reduces when it be 1. The above result indicates that the crashes under free-flow traffic conditions would be more prone to non-injury. It is in line with recent studies (Liu et al., 2023), as increase in traffic volumes tends to increase crash risk in expressway (Sun et al., 2016).
(2) For the second important variable, 'Traffic flow status_3 (jam flow)', the red feature value points consistently appear in the positive SHAP value range, in contrast to the distribution observed under free flow conditions. These results suggest an increased likelihood of injury crash under jam flow conditions. This is consistent with recent studies (Xu et al., 2015), in which the authors claimed that stop-and-go jams can increase the crash risk in standing congested traffic.
(3) "Saturday" is the third important variable, which show positive influences on the crash severity. The similar result is reported by Lu et al. (2016), in which weekends are observed to increase

Table 4
Model results comparation.

Models		**Classification metrics*** **(higher better)**				**Regression metrics*** **(lower better)**	
		Accuracy	Sensitivity	F1 score	AUC	MSE	MAE
Crash severity	LogR	0.653 (0.035)	0.431 (0.058)	0.285 (0.030)	0.604 (0.022)	–	
	RF	**0.857** (0.013)	0.191 (0.072)	0.278 (0.093)	0.575 (0.045)		
	SVM	0.757 (0.023)	0.504 (0.062)	0.382 (0.040)	0.708 (0.039)		
	XGBoost	0.753 (0.009)	0.304 (0.051)	0.368 (0.045)	0.653 (0.039)		
	MLP	0.698 (0.022)	0.529 (0.048)	0.381 (0.033)	0.719 (0.028)		
Crash duration	LinR	–				0.457 (0.024)	0.540 (0.018)
	RF					0.504 (0.050)	0.562 (0.029)
	SVM					0.473 (0.031)	0.562 (0.029)
	XGBoost					0.548 (0.050)	0.584 (0.031)
	MLP					0.454 (0.020)	0.540 (0.023)
Joint model	Hard sharing	0.719 (0.019)	0.557 (0.044)	0.398 (0.019)	0.720 (0.015)	0.443 (0.028)	0.530 (0.012)
	Soft sharing	0.754 (0.011)	**0.566** (0.026)	**0.421** (0.009)	**0.741** (0.022)	**0.437** (0.017)	**0.527** (0.014)

*: The standard deviation of model evaluation metrics of 10-fold cross validation is in parentheses.

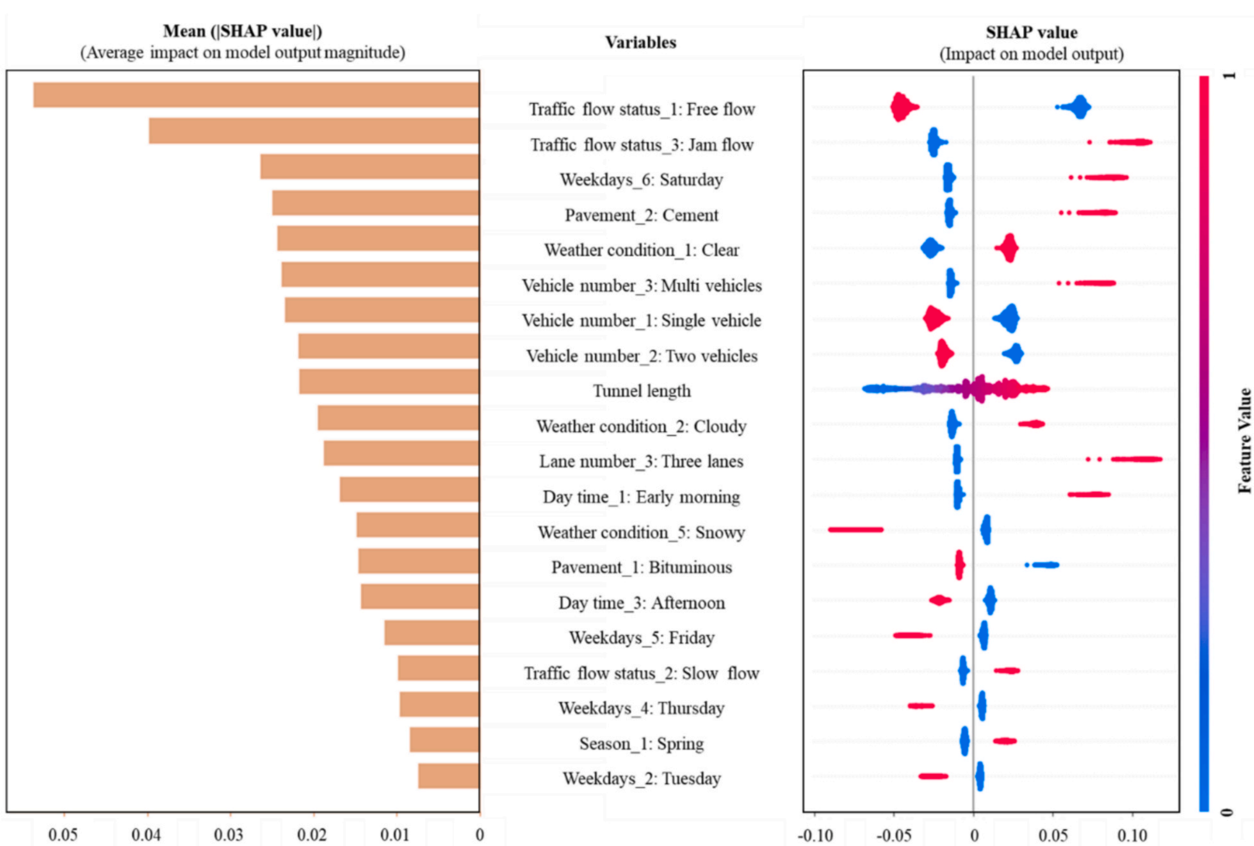

Fig. 7. Top-20 important variables for crash severity and their SHAP values.

injury severity in tunnels. Saturdays frequently experience an increase in recreational or non-work-related travel. Drivers may approach tunnels at higher speeds, either due to unfamiliarity with the routes taken during leisure activities or because they are less restricted by the traffic conditions typical of Saturday rush hours (Morris and Hirsch, 2016; Zhang et al., 2011). Furthermore, non-commuter drivers, such as tourists or occasional travelers, may lack familiarity with tunnel driving conditions, potentially resulting in panic or suboptimal decision-making. Tunnels often serve as bottlenecks, where traffic slows and congestion intensifies (Tympakianaki et al., 2019), particularly on weekends when road usage patterns tend to be more unpredictable (Bhat and Gossen, 2004).

(4) The cement road pavement and cloudy weather condition are other important variables to the crash severity, which all appear to increase the probability of injury crashes. Notably, the reduced visibility associated with cloudy weather is positively correlated with increased drivers' reaction times when entering tunnels. Additionally, the transition from brighter outdoor light to the dimmer environment inside tunnels can cause drivers to react more slowly to obstacles or changes in traffic flow (Domenichini et al., 2017), thereby potentially increasing the likelihood of injuries.

For the crash duration, Fig. 8 presents its top-20 important variables with their SHAP values. From the results, it can be seen that:

(1) The variable "Involved vehicle type_2 (truck only)" emerges as the preeminent variable with a notable concentration of red feature value points in the positive SHAP value domain. It shows a clear trend that crashes involving only trucks are highly associated with extended durations. This could be attributed to the larger size and mass of trucks, which may complicate and prolong the clearance process.

(2) The two-vehicle and multi-vehicle crashes are the second and third important variables for crash duration and have positive SHAP values when their values are 1. It means when the two- and multi-vehicle crashes occur, the duration of crashes would increase. It could be also observed that the multi-vehicle crashes increase the higher duration than the two-vehicle crashes. Two- and multi-vehicle crashes generally lead to longer congestion durations due to their complexity, increased severity, greater traffic disruption, and the more extensive response required from emergency and recovery teams (Zhen et al., 2024).

5.3.2. Comparison of the joint and separate models

To further analyze the advantages of joint modeling over separate modeling, the SHAP distribution of each variable in the two models was compared. For two separate models, their MLP networks were kept consistent with the joint model, but the cross-stitch layers between the networks were deleted, which means that the two networks were trained independently for a single task without parameter sharing. Since experimental results show that the performances of the separate models are lower than that of the joint model, the SHAP value analysis can provide more details in understanding the difference of feature extraction mechanism between such models under the same tasks (i.e., crash severity classification and duration regression).

Fig. 9 demonstrates the SHAP value distributions of variables in joint and separate models for crash severity. For most variables, their SHAP value distribution are consistent in the two models. However, there are significant distribution differences in their SHAP values of some variables, which are divided into two major categories:

(1) **Inversion features:** Their SHAP value distribution in the separate models are opposite to that in the joint model. For the "Surface_2: Wet surface, the SHAP values are positive when its value is 1 in the separate model. In contrast, the corresponding SHAP values become negative in the joint model, indicating that crash severity tends to be lower, with a higher likelihood of non-injury outcomes under wet road surface conditions. This trend can be attributed to drivers generally reducing their speed due to decreased traction and a heightened perception of risk (Theofilatos and Yannis, 2014; Wang et al., 2022), particularly in tunnel environments. Additionally, wet road conditions typically

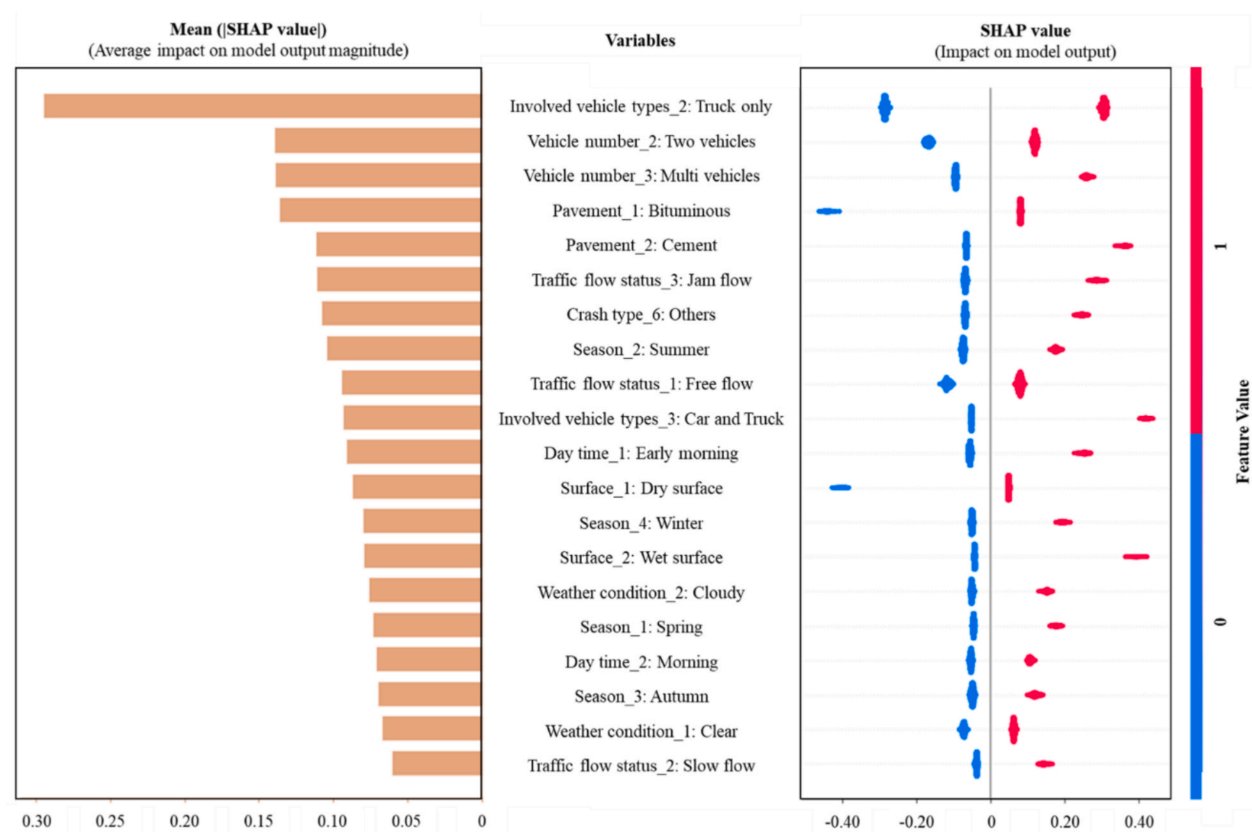

Fig. 8. Top-20 important variables for crash duration and their SHAP values.

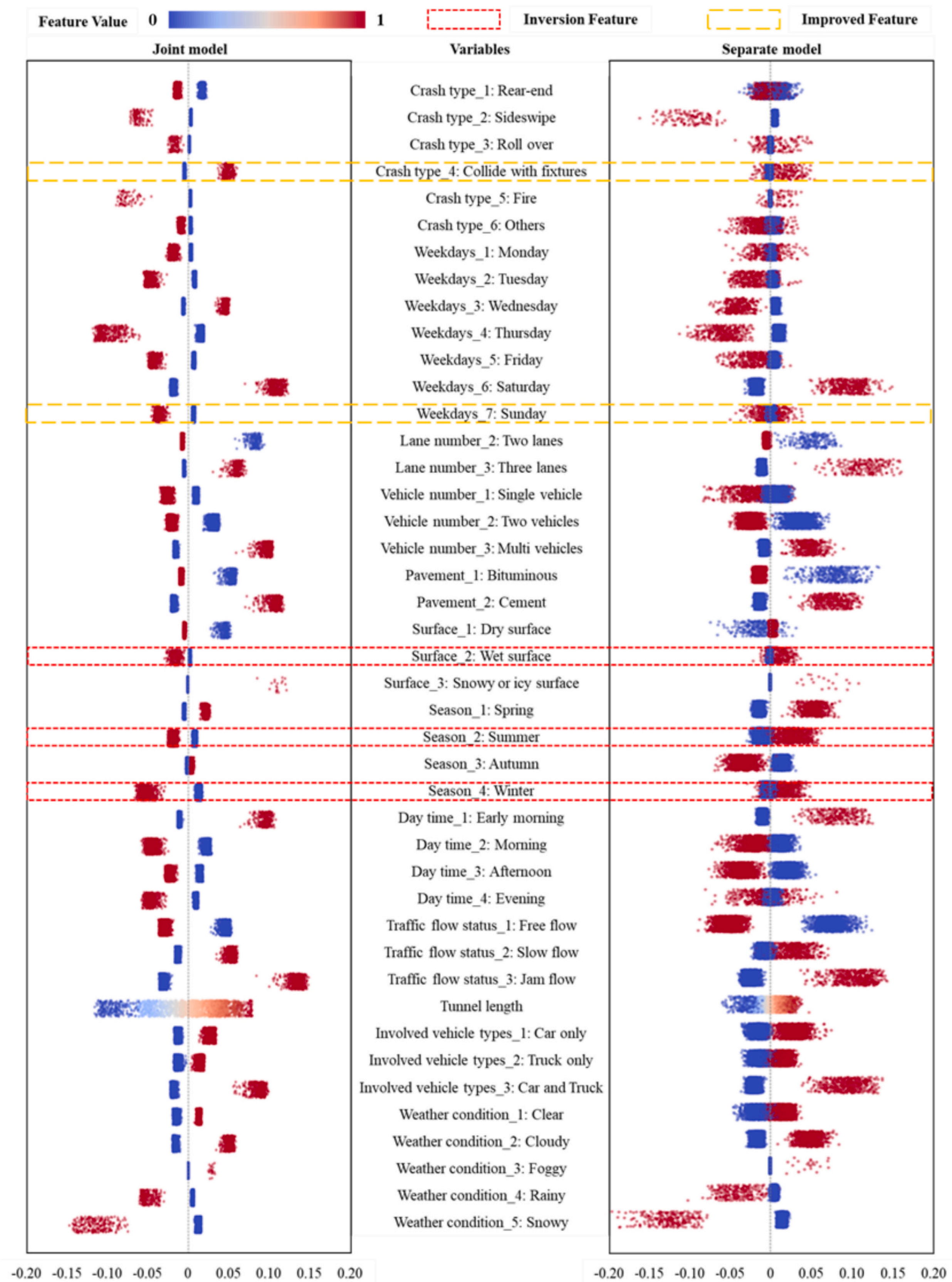

Fig. 9. SHAP value distributions of variables in joint and separate models for crash severity.

encourage more cautious driving behavior, such as increasing following distances and minimizing sudden maneuvers, which collectively help to prevent severe collisions and mitigate the risk of injury (Wu et al., 2020). Another inversion feature are the "Summer" and "Winter". Their SHAP value changes from positive in the separate model to negative in the joint model when its value is 1. Tunnel crashes in summer and winter are more likely to result in property damage-only crashes when both crash severity and duration are considered. In summer, drivers are generally more vigilant and better able to anticipate road

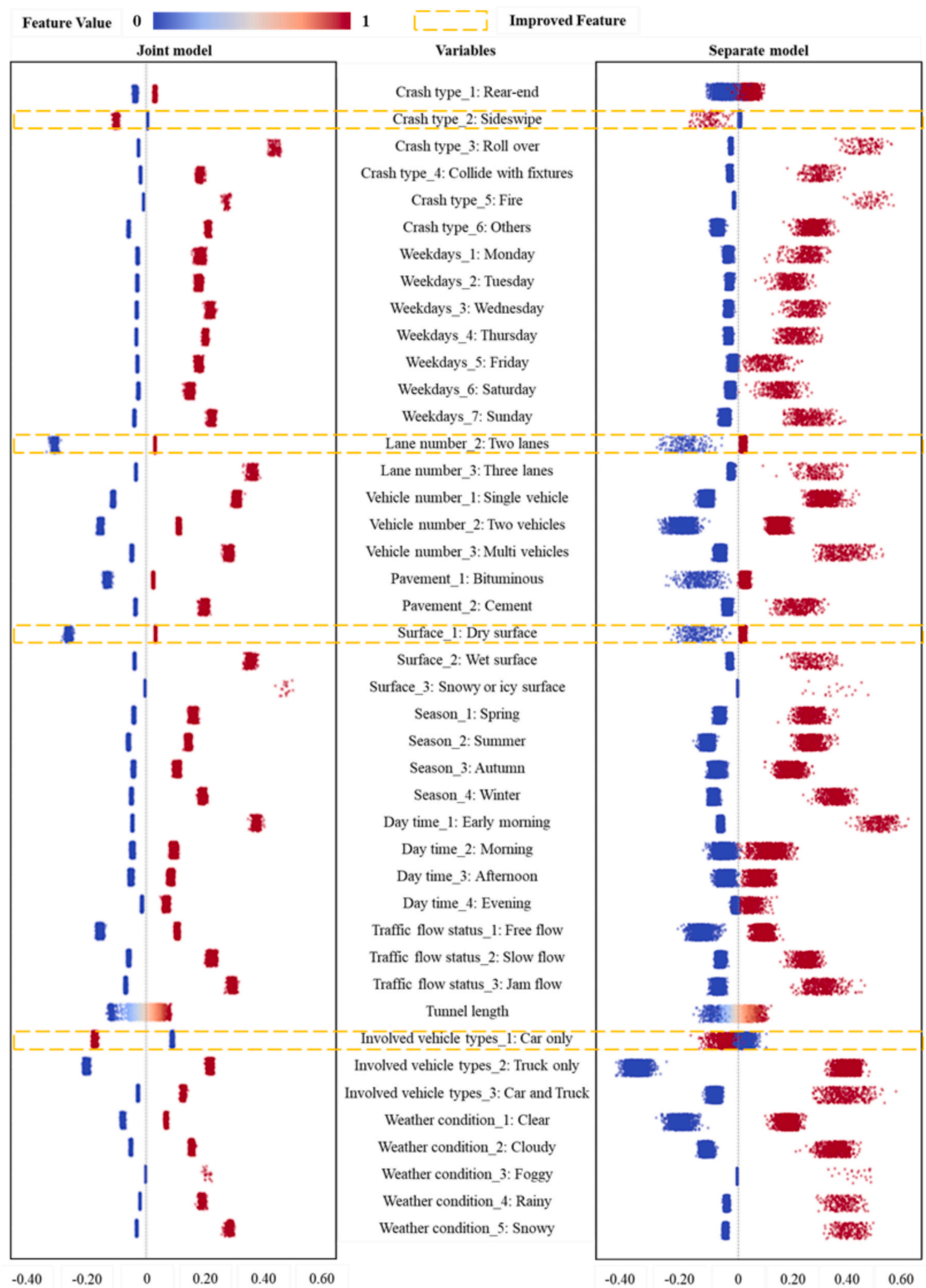

Fig. 10. SHAP value distributions of variables in joint and separate models for crash duration.

conditions, as visibility tends to be clearer and road surfaces are typically dry. In winter, drivers exercise heightened caution in response to potential hazards such as ice, snow, or wet surfaces, which leads to reduced speeds and greater following distances (Hyodo and Hasegawa, 2021). This cautious driving behavior mitigates the impact force during crashes, frequently limiting them to property damage rather than personal injury. Additionally, the tunnel environment encourages lower speeds due to either heavy tourist traffic in the summer (Butler, 2001) or increased caution in winter (Theofilatos and Yannis, 2014). These reduced speeds result in crashes being less severe overall, leading to a higher proportion of PDO crashes. Even when the crash response duration is extended, the severity of the crash often remains low, contributing to a higher incidence of property damage-only outcomes.

(2) **Improved features**: In the separate models, the SHAP value distributions remain close to 0 at different values, but in the joint model, they become significantly distinct, deviating further from 0. This pattern is observed for features such as "Colliding with Fixtures" and "Sunday", where the differences in SHAP values between the separate and joint models are more pronounced.

Otherwise, the variables' SHAP value distributions for crash duration in joint and separate models is compared in Fig. 10. Notably, the SHAP value distribution of all variables are consistent in the two models. The SHAP value distribution for 4 features— "Crash Type 2: Sideswipe", "Lane Number_2: Two Lanes", "Surface: Dry Surface", and "Involved Vehicle Types_1: Car Only"—exhibits significant differences between the joint and separate models. For instance, in the separate model, the variable "Involved Vehicle Types_1: Car Only" shows minimal influence (SHAP values close to 0) regardless of its value (0 or 1). However, in the joint model, the corresponding SHAP values become positive when the value is 1, indicating that crashes involving only cars tend to have shorter durations.

The aforementioned results demonstrate an improvement in the joint model by accounting for the interdependence between injury severity outcomes and congestion durations, as well as addressing shared unobserved heterogeneity. The estimated findings also offer new insights into crash dynamics by simultaneously considering both injury outcomes and congestion duration. To address this, the joint model's cross-stitch layers enhance the feature extraction process by considering their relationships with both crash severity and duration, thereby correcting biased mappings. Consequently, this joint evaluation more accurately discerns the relationship between each crucial feature and crashes, as evidenced by improved model performance in both tasks.

6. Conclusions and Future Direction

This paper embarks on a comprehensive exploration of congestion duration and crash severity within tunnels, leveraging an innovative approach that combines the crash severity and congestion duration evaluation network based on cross-stitch networks.

Based on cross-stitch MLP, the joint models outperformed the separate models, with 7.0 % and 10.2 % in sensitivity and F1 score, respectively for crash severity classification. Moreover, the MSE and MAE for the crash congestion duration evaluation dropped by 3.7 % and 2.5 % respectively. By jointly assessing both crash severity and congestion duration, this study aims to provide a holistic and data-driven perspective on tunnel safety and efficiency. Compared to the separate models, the SHAP value distribution indicated more obvious effects of variables in the joint models. Moreover, some inversion features also exited in the joint models compared to the separate models, while the interpretation in the joint models would provide more reasonable results.

Analyzing the data and findings suggests several recommendations for tunnel management to reduce both congestion duration and crash severity. These include establishing a real-time crash detection and reporting system, enhancing EMS facilities, and deploying more rescue vehicles, particularly in longer tunnels. It is crucial to implement educational programs and strict measures to prevent fatigued driving of truck drivers. Additionally, closing some long tunnels during early morning hours could help prevent severe crashes, particularly for those with high crash frequency.

Nevertheless, some limitations exist in the current study. More datasets will be collected to improve the estimation accuracy. Otherwise, more advanced deep learning methods will be developed to address the unobserved heterogeneity and non-linear relationships in data.

CRediT authorship contribution statement

Chenzhu Wang: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Mohamed Abdel-Aty:** Writing – review & editing, Supervision, Resources, Project administration, Formal analysis. **Lei Han:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Data curation, 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

The authors do not have permission to share data.

References

Abay, K., Paleti, R., Bhat, C., 2013. The joint analysis of injury severity of drivers in two-vehicle crashes accommodating seat belt use endogeneity. Transportation Research Part b: Methodological 50, 74–89.

Anas, A., Lindsey, R., 2011. Reducing Urban Road Transportation Externalities: Road Pricing in Theory and in Practice. Review of Environmental Economics and Policy 5 (1).

Antwarg, L., Miller, R., Shapira, B., Rokacha, L., 2021. Explaining anomalies detected by autoencoders using Shapley Additive Explanations. Expert Systems with Applications 186, 115736.

Asudani, D., Nagwani, N., Singh, P., 2023. Impact of word embedding models on text analytics in deep learning environment: a review. Artificial Intelligence Review 56, 10345–10425.

Bari, S., Naser, J., 2010. Simulation of airflow and pollution levels caused by severe traffic jam in a road tunnel. Tunnelling and Underground Space Technology 25, 70–77.

Beljaards, L., Elmahdy, M., Verbeek, F., Staring, M., 2020. A Cross-Stitch Architecture for Joint Registration and Segmentation in Adaptive Radiotherapy. Proceedings of Machine Learning Research 121, 62–73.

Bhat, C., Gossen, R., 2004. A mixed multinomial logit model analysis of weekend recreational episode type choice. Transportation Research Part b: Methodological 38 (9), 767–787.

Butler, R., 2001. Seasonality in tourism: Issues and implications. In Seasonality in tourism. Routledge, pp. 5–21.

Calvi, A., De Blasiis, M., Guattari, C., 2012. An empirical study of the effects of road tunnel on driving performance. Procedia – Social. Behavioral Science. 53, 1098–1108.

Chang, H., Yu, C., Zhang, L., Xu, X., Dustdar, S., 2024. Safety assessment of tunnel construction based on counterintuitivity detection using multi-profile multi-model ensemble learning. Expert Systems with Applications 240, 122459.

Chen, X., Ren, P., Medhi, J., 2022. Chapter 3.8: Deep Multi-task Learning Approach for Bioelectrical Signal Analysis. In: Computational Intelligence and Image Processing in Medical Applications, pp. 189–212.

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., Kuksa, P., 2011. Natural Language Processing (Almost) from Scratch. Journal of Machine Learning Research 12, 2493–2537.

Domenichini, L., Torre, F., Vangi, D., Virga, A., Branzi, V., 2017. Influence of the lighting system on the driver's behavior in road tunnels: A driving simulator study. Journal of Transportation Safety & Security 9 (2), 216–238.

Du, Z., Huang, F., Ran, B., Pan, X., 2014. Safety evaluation of illuminance transition at highway tunnel portals on basis of visual load. Transportation Research Record 2458 (1), 1–7.

Egger, J., Pepe, A., Gsaxner, C., Jin, Y., Li, J., Kern, R., 2021. Deep learning—a first meta-survey of selected reviews across scientific disciplines, their commonalities, challenges and research impact. PeerJ Computer Science 7, e773.

Goel, R., Singh, B., Zhao, J., 2012. Underground infrastructures: planning, design, and construction. Elsevier, ISBN 978-0-12-397168-5.

Huang, T., Li, J., Yu, D., Deng, L., Gong, Y., 2013. Cross-language knowledge transfer using multilingual deep neural network with shared hidden layers. In: In Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, pp. 7304–7308.

Hyodo, S., Hasegawa, K., 2021. Factors affecting analysis of the severity of accidents in cold and snowy areas using the ordered probit model. Asian Transport Studies 7, 100035.

Khetwal, S., Pei, S., Gutierrez, M., 2021. Stochastic event simulation model for quantitative prediction of road tunnel downtime. Tunnelling and Underground Space Technology 116, 1–20.

Li, X., Ma, X., Xiao, F., Xiao, C., Wang, F., Zhang, S., 2022. Small-Sample Production Prediction of Fractured Wells Using Multitask Learning. SPE Journal 27 (03), 1504–1519.

Liao, T., Hu, T., Ho, W., 2012. Simulation studies of traffic management strategies for a long tunnel. Tunnelling and Underground Space Technology 27, 123–132.

Liu, X., Lu, J., Wang, B., Zhu, M., Zhang, F., Chen, X., 2023. Exploring the main and interaction effects of traffic flow characteristics, roadway design and weather conditions on the real-time crash risk for urban roads with mixed logit model. Journal of Transportation Safety & Security. 2230569.

Liu, Q., Wang, D., Jia, Y., Luo, S., Wang, C., 2022. A multi-task based deep learning approach for intrusion detection. Knowledge-Based Systems 238, 107852.

Lu, J., Xing, Y., Wang, C., Cai, X., 2016. Risk factors affecting the severity of traffic accidents at Shanghai river-crossing tunnel. Traffic Injury Prevention 17 (2), 176–180.

Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, p. 30.

Luo, P., Zhang, X., Meng, R., 2023. Co-attention learning cross time and frequency domains for fault diagnosis. Cognitive Robotics 3, 34–44.

Ma, Z., Chien, S., Dong, C., Hu, D., Xu, T., 2016. Exploring factors affecting injury severity of crashes in freeway tunnels. Tunnelling and Underground Space Technology 59, 100–104.

Meng, Q., Qu, X., Wang, Y., Yuanita, V., Wong, S., 2011. Quantitative Risk Assessment Modeling for Nonhomogeneous Urban Road Tunnels. 31 (3), 1–22.

Misra, I., Shrivastava, A., Gupta, A., Hebert, M., 2016. Cross-Stitch Networks for Multi-Task Learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3994–4003.

Morris, E., Hirsch, J., 2016. Does rush hour see a rush of emotions? Driver mood in conditions likely to exhibit congestion. Travel Behaviour and Society 5, 5–13.

Karniadaks, G., G. Kevrekids, L., Lu, L., Perdikaris, P., Wang, S., Yang, L., 2021. Physics-informed machine learning. Nature Reviews Physics 3, 422–440.

National Bureau of Statistics (NBS). Chinese statistical yearbook (2022). http://www.stats.gov.cn/, 2023-11-22.

Nweke, H., The, Y., Al-garadi, M., Alo, U., 2018. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. Expert Systems with Applications 105 (1), 233–261.

Pervez, A., Lee, J., Huang, H., 2022. Exploring factors affecting the injury severity of freeway tunnel crashes: A random parameters approach with heterogeneity in means and variances. Accident Analysis and Prevention 178, 106835.

Pucher, J., Peng, Z., Mittal, N., Zhu, Y., Korattyswaroopam., 2006. Urban Transport Trends and Policies in China and India: Impacts of Rapid Economic Growth. Transportation Reviews, 27 (4): 379-410.

Ross, T., Dollár, G., 2017. July. Focal loss for dense object detection. In: In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2980–2988.

Schulz, M., Yeo, B., Vogelstein, J., Mourao-Miranada, J., Kather, J., Kording, K., Kording, J., Richards, B., Bzdok, D., 2020. Different scaling of linear models and deep learning in UKBiobank brain images versus machine-learning datasets. Nature Communications 11, 4238.

Sramek, J., Danisovic, P., Hodon, M., 2019. Evaluation and analysis of emergency events in road tunnels. Transportation Research Procedia 40, 1571–1575.

Sun, J., Li, T.N., Li, F., Chen, F., 2016. Analysis of safety factors for urban expressways considering the effect of congestion in Shanghai, China. Accident Analysis and Prevention 95, 503–511.

Sun, P., Yu, Q., You, K., 2024. Intelligent traffic management strategy for traffic congestion in underground loop. Tunnelling and Underground Space Technology 143, 1–16.

Tan, Z., Gao, H., 2015. Traffic control for air quality management and congestion mitigation in complex urban vehicular tunnels. Transportation Research Part c. 58, 13–28.

Tan, M., Hu, C., Chen, J., Wang, L., Li, Z., 2022. Multi-node load forecasting based on multi-task learning with modal feature extraction. Engineering Applications of Artificial Intelligence 112, 104856.

Thanasutives, P., Numao, M., Fukui, K., 2021. Adversarial Multi-task Learning Enhanced Physics-informed Neural Networks for Solving Partial Differential Equations. In: 2021 International Joint Conference on Neural Networks (IJCNN), p. 52387.

Theofilatos, A., Yannis, G., 2014. A review of the effect of traffic and weather characteristics on road safety. Accident Analysis and Prevention 72, 244–256.

Tian, K., Bai, K., 2023. End-to-end multitask learning with vision transformer. IEEE Transactions on Neural Networks and Learning Systems. 3234166, 1–12.

Tukra, S., Lidstromer, N., Ashrafian, H., 2020. Meta learning and the AI learning process. Artificial Intelligence in Medicine 1–15.

Tympakianaki, A., Koutsopoulos, H., Jenelius, E., 2019. Anatomy of tunnel congestion: Causes and implications for tunnel traffic management. Tunnelling and Underground Space Technology 83, 498–508.

Vandenhgende, S., Georgouis, S., Proesmans, M., Dai, D., Gool, L., 2020. Revisiting Multi-Task Learning in the Deep Learning Era. arXiv:2004.13379v1.

Wang, C., Chen, F., Zhang, Y., Cheng, J., 2022. Spatiotemporal instability analysis of injury severities in truck-involved and non-truck-involved crashes. Analytic Methods in Accident Research 34, 100214.

Wang, C., Easa, S., Dong, D., Chen, F., Xiao, F., Cheng, J., 2023. Alternative unobserved heterogeneity models to analyze injury severity of expressway crashes in different tunnel types. Tunnelling and Underground Space Technology 140, 1–16.

Wang, C., Abdel-Aty, M., Han, L., Easa, S., 2024a. Analyzing speed-difference impact on freeway joint injury severities of Leading-Following vehicles using statistical and data-driven models. Accident Analysis and Prevention 206, 107695.

Wang, C., Abdel-Aty, M., Han, L., 2024b. Effects of speed difference on injury severity of freeway rear-end crashes: Insights from correlated joint random parameters bivariate probit models and temporal instability. Analytic Methods in Accident Research 42, 100320.

Wang, C., Abdel-Aty, M., Cui, P., Han, L., 2024c. Effects of helmet usage on moped riders' injury severity in moped-vehicle crashes: Insights from partially temporal constrained random parameters bivariate probit models. Accident Analysis and Prevention 208, 107800.

Wu, Y., Abdel-Aty, M., Wang, L., Rahman, M., 2020. Combined connected vehicles and variable speed limit strategies to reduce rear-end crash risk under fog conditions. Journal of Intelligent Transportation Systems 24 (5), 494–513.

Xu, C., Wang, W., Liu, P., Zhang, F., 2015. Development of a real-time crash risk prediction model incorporating the various crash mechanisms across different traffic states. Traffic Injury Prevention 16 (1), 28–35.

Yasmin, S., Eluru, N., 2018. A joint econometric framework for modeling crash counts by severity. Transportmetrica a: Transport Science 14 (3), 230–255.

Yusuf, J., O`Connell, L., A. Anuar, K., 2014. For whom the tunnel be tolled: A four-factor model for explaining willingness-to-pay tolls. Transportation Research Part A: Policy and Practice. 59: 13-21.

Zhang, K., Batterman, S., Dion, F., 2011. Vehicle emissions in congestion: Comparison of work zone, rush hour and free-flow conditions. Atmospheric Environment 45 (11), 1929–1939.

Zhao, J., Du, B., Sun, L., Zhuang, F., Lv, W., Xiong, H., 2019. Multiple Relational Attention Network for Multi-task Learning. KDD '19, August 4–8, 2019, Anchorage, AK, USA.

Zhen, L., Wu, J., Chen, F., Wang, S., 2024. Traffic emergency vehicle deployment and dispatch under uncertainty. Transportation Research Part e: Logistics and Transportation Review 183, 103449.

Zheng, Z., Wang, Z., Zhu, L., Jiang, H., 2020. Determinants of the congestion caused by a traffic accident in urban road networks. Accident Analysis and Prevention. 136, 105327.

Zou, G., Lai, Z., Wang, T., Liu, Z., Bao, J., Ma, C., Li, Y., Fan, J., 2024. Multi-task-based spatiotemporal generative inference network: A novel framework for predicting the highway traffic speed. Expert Systems with Applications 237.