A conflict risk graph approach to modeling spatio-temporal dynamics of intersection safety

Tao Wang [a], Ying-En Ge [a], Yongjie Wang [a,b,*], Carlo G. Prato [b], Wenqiang Chen [a], Yuchen Niu [a]

[a] School of Transportation Engineering, Chang'an University, Xi'an 710064, China
[b] School of Civil Engineering, University of Leeds, Leeds LS2 9JT, UK

ARTICLE INFO

Keywords:
Conflict risks
Trajectory data
Deep learning
Surrogate safety measures
Spatio-temporal analysis
Traffic safety

ABSTRACT

Intersections are among the most hazardous roadway spaces due to the complex and conflicting road users' movements. Spatio-temporal modeling of conflict risks among road users can help to identify strategies to mitigate the exacerbation of safety risks and restore hazardous conditions to normal traffic situations. This paper proposes the 'Conflict Risk Graph' as a novel concept to infer real-time conflict risks at intersections at a fine-grained level by mapping conflict-prone locations to nodes within a network characterized by specific topological structures. A significant contribution of this work is the development of a Transformer-based Graph Convolutional Network (Trans-GCN), a model that synergistically combines the Transformer's proficiency in capturing global dependence with the GCN's ability to learn local correlations. The Trans-GCN adeptly models the complex evolution patterns of conflict risks at intersections. The evaluation in this paper against five common state-of-the-art deep learning approaches demonstrates the superior performance of the Trans-GCN in conflict risk inference and adaptability to node changes. Furthermore, extensive experiments with different node configurations reveal a correlation between node setup and model performance, showing that higher spatio-temporal resolution decreases inference accuracy. This insight informs the selection of an optimal node configuration that balances the detailed capture of spatio-temporal dynamics with modeling accuracy, enabling ideal conflict risk inferences at intersections. Ultimately, this work offers significant insights for the enhancement of proactive traffic safety management and the advancement of intelligent traffic systems.

1. Introduction

One of the key aspects of proactive safety management is the real-time inference of traffic risks that allows for the implementation of mitigation strategies to restore hazardous conditions to normal traffic situations. Over the past few decades, considerable effort has been posed on real-time traffic safety evaluation (Arun et al., 2022; Belloumi and Ouni, 2018; Basso et al., 2018; Ferreira et al., 2015; Fu et al., 2023; Ghoul et al., 2023; Howlader et al., 2024; Hussain et al., 2023; Li et al., 2020; Sun et al., 2015; Wu et al., 2023; Xu et al., 2014; Yang et al., 2018; Zheng et al., 2020). A comprehensive literature review by Hossain et al. (2019) identified 78 relevant studies conducted between 2003 and 2018 and revealed that most of these studies (i) utilized loop-detectors or cameras to gather dynamic

* Corresponding author at: School of Transportation Engineering, Chang'an University, Xi'an 710064, China.
E-mail address: wyj@chd.edu.cn (Y. Wang).

traffic data, and (ii) employed statistical based methods such as logistic regression or Bayesian modeling to analyze the data and assess traffic risks. At the time, Hossain et al. (2019) advocated for the development of a proactive safety system capable of timely identifying the evolution of hazardous conditions that can then be mitigated by using various traffic smoothing measures.

Although these aforementioned studies lay the groundwork for realizing real-time intersection traffic safety, existing literature exhibits several limitations. Firstly, these studies rely on crash data, which are limited when inferring future traffic risk at a fine-grained level because of issues such as small samples with low means, underreporting, and inaccuracies (Fu et al., 2023; Wang et al. 2024c; Zheng et al. 2020). Secondly, these studies treat traffic risk modeling as a classification issue yielding a binary outcome (crash or no crash) within a certain period, therefore losing a large amount of information to quantify intersection safety level. Thirdly, most studies investigate correlations between traffic risks and their contributing factors but frequently overlook their intrinsic dependence or have not effectively integrated spatio-temporal information. Finally, they often employ excessively broad spatio-temporal units such as entire cities (Bao et al., 2019), but methods used in coarse-grained studies are not transferable to intersections, as the evolution of conflict risks at the micro-level is rapid and complex.

To fill these gaps, this paper proposes a novel framework to evaluate intersection safety at a fine-grained level. We treat traffic conflicts as surrogates for traffic risk, since the same underlying failure mechanisms contribute to both conflicts and crashes (Sacchi et al., 2013). Then, we introduce a graph-based method to characterize conflict risk at intersections. Lastly, we utilize the inherent spatio-temporal dependence within sequential data to infer future conflict risks, essentially framing it as a sequence forecasting problem in a high-dimensional space. The main challenge in this task stems from the 'curse of local spatial sensitivity' combined with the 'curse of global temporal dependence', which complicates the modeling of conflict risks. The former refers to the challenge of conflict risks propagating within a local zone: the closer the distance between zones, the greater the correlations (Yang et al. 2023; Buddhavarapu et al. 2016). The latter implies that conflict risks are influenced by global historical patterns (Essa and Sayed, 2018; Wang et al. 2024a). However, traditional mathematical, statistical, and deep learning approaches may struggle to capture the complex and dynamic spatio-temporal dependence within conflict risk sequences because of the two aforementioned challenges. Notably, recent advances in spatio-temporal deep learning approaches, in which several learning blocks are used to build up progressively more abstract representations of the multidimensional data (Liu et al., 2023a; Wang et al., 2023; Wang et al., 2024a; Yu et al., 2018; Zhao et al., 2020), have made it possible for neural networks to learn complex spatio-temporal correlations directly from raw sequence data. In this study, we develop one successful architecture, the spatio-temporal deep learning approaches, which combine advantages of multiple deep learning methods, thereby exploiting the global–local correlations present in risk sequences, and building in robustness to inference of intersection safety.

In general, the contributions of this paper can be summarized as follows. The first is the construction of the 'Conflict Risk Graph' for an intersection. Given the challenge of global–local dependence, we leverage computer vision to identify conflict-prone locations at intersections. These locations are then mapped as conflict nodes on a topologically structured risk propagation network. This graph-based network facilitates the dynamic characterization of evolution patterns in conflict risks and enhances our ability to capture the spatio-temporal evolution patterns of risks at intersections. The second is the proposition of a spatio-temporal deep learning approach—Transformer-based Graph Convolutional Network (Trans-GCN)—for modeling the evolution of conflict risks. The Trans-GCN combines the strengths of Graph Convolutional Networks (GCN) for extracting local features with the Transformer's capabilities in capturing global dependence, providing an effective approach to modeling conflict risk evolution. The third step is the identification of the optimal node configuration for ideal conflict risk inferences. Through extensive experiments, we determined the optimal node configuration for the Trans-GCN to model the evolution dynamics of conflict risks while maintaining necessary spatio-temporal details. This configuration ultimately achieves an ideal inference of conflict risk map at intersections. The methodology flowchart for developing the conflict risk inference framework is illustrated in Fig. 1, providing a visual overview of the process.

2. Literature review

2.1. Traffic risk modeling

Developing a proactive traffic safety system requires a dependable traffic safety model that uses real-time traffic data to captute its correlation with traffic risks (Zheng et al., 2020). Traffic risk modeling can generally be categorized into two types based on the data

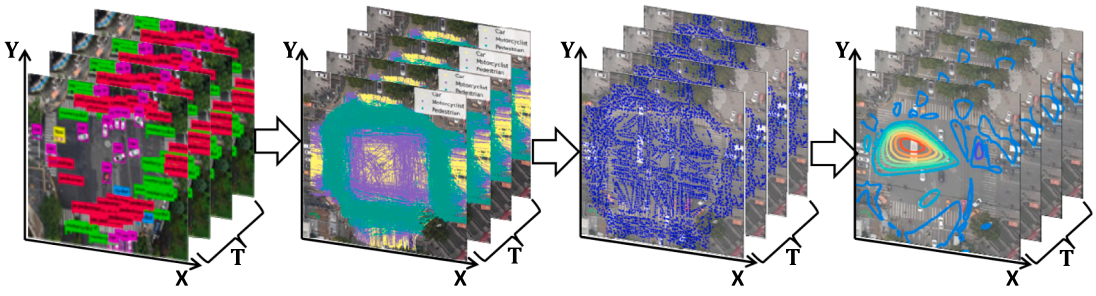


Fig. 1. The methodology flowchart for developing the framework of conflict risk inference.

used, namely crash or conflict data.

Crash-based risk models establish correlations between crash data and their contributing factors, such as weather conditions (Bao et al. 2019; Wu et al., 2013), traffic variables (Stevanovic et al., 2013; Xu et al., 2014; Ferreira and Couto 2015; Fu and Sayed, 2021; Yang et al., 2018; Fu and Sayed, 2023; Zheng et al., 2020), and roadway geometry (Wang et al., 2015; Li et al. 2023). Regression analysis typically helps identify solutions for monitoring prevailing environmental conditions and facilitates proactive traffic safety management. However, the drawbacks of crash-based risk modeling are the limited quantity and high randomness of crash data (Arun et al., 2022; Howlader et al., 2024; Peng et al., 2017). These drawbacks limit the ability of systems based on crash-based models to predict and quantify future traffic safety at a fine-grained level.

Conflict-based risk modeling uses conflict data to explore the correlation between risks and their contributing factors (Ghoul et al., 2023; Guo et al., 2019; Zheng and Sayed, 2019; Zheng and Sayed, 2020). These studies often assess traffic risks via surrogate safety measures (SSMs) such as Time-To-Collision (TTC) or Post Encroachment Time (PET). For instance, Peng et al. (2017) investigated the evolution of crash risks under foggy conditions by utilizing TTC. Similarly, Ghoul et al. (2023) dynamically identified high-risk intersection via a short-term hazardous location identification and ranking framework.

Recently, the rapid development of computer vision has enhanced the accessibility of high-resolution data and hence a surge in related studies from the extraction of trajectory data of road users has been observed (Fu et al., 2023; Wang et al., 2024a; Xie et al., 2019; Yang et al., 2021). For example, Yang et al. (2021) used trajectory data to develop a novel approach that simulated time series of SSMs directly to detect unsafe traffic conditions from trajectory data. St-Aubin et al. (2015) used very rich trajectory data to demonstrate the application of a video-based automated traffic safety analysis that combined SSMs and computer vision to simulate collision potential. In addition, Xie et al. (2019) extracted vehicle trajectories from video recordings and then TTC was calculated for proactive safety analytics. Although the existing literature has significantly enriched the field of traffic risk modeling, there is a lack of fine-grained investigation of the spatio-temporal dependence inherent in risk seuquences.

2.2. Spatio-temporal analysis of traffic risks

Recently, spatio-temporal analysis of traffic risks has been performed extensively in the literature (Essa and Sayed, 2018; Hussain et al., 2023; Kashifi et al., 2022; Li et al., 2020; Liu et al., 2023a; Song et al., 2020; Tran et al., 2023; Wang et al., 2024a; Yu et al., 2020). These studies suggest that risks exhibit significant patterns of aggregation or dissipation in the spatio-temporal space, suggesting that there exists a learning component of dynamic evolution. These studies on modeling risk evolutions can be categorized into two methodologies: mathematical/statistical methods and deep learning techniques.

2.2.1. Mathematical/statistical methods for traffic risk modeling

Statistical-based risk modeling establishes correlations between risks and potential features by taking into account the temporal or spatial dependence of conflict risks (Essa and Sayed, 2018; Ferreira and Couto, 2015). For instance, Essa and Sayed (2018) developed conflict-based Safety Performance Functions (SPFs) for intersections at the signal cycle level and provided insight into how alterations in the signal cycle design can influence the spatial distribution of conflicts. Given that statistical methods inherently lacked the capability to forecast future risks, Hussain et al. (2023) proposed a unified framework that combined extreme value theory with autoregressive integrated moving average models to evaluate in real-time crash risks at intersections. Moreover, Hu et al. (2023) explored the dynamic evolution mechanism of conflict risks by using trajectory data in the spatio-temporal space and employing a two-step approach: (i) applying a spatial Markov model to examine risk level transition probabilities, and (ii) using statistical-based method to quantify the correlation between risks and traffic characteristics.

Mathematical models are also used to describe the dynamic evolution of traffic risks. For example, Yang et al. (2023) introduced the plume diffusion model to capture the non-linear correlations between conflict points while accounting for the spatial evolution of traffic conflict risks at intersections. In addition, Hu et al. (2018) introduced the concept of Global-Local Colocation Quotient to examine the spatial correlation patterns between crashes and intersections while offering an innovative approach to identifying intersections where crashes are more likely to occur. In the study of Guo et al. (2019), their results suggested that conflict incidents cluster in certain zones of intersections, with the frequency decreasing from the center of these zones to the periphery. A two-step calibration method was developed to test the relationships between field-measured and simulated conflicts. Additionally, Xie et al. (2019) suggested that Hidden Markov models outperformed other models when simulating the long-term temporal evolution of conflict risks. Lastly, traffic shockwave theory was often used to model the speed and distance of risk propagation among vehicles (e.g., Wang et al., 2016).

In summary, existing studies in this field reveals that traffic risk at intersections exhibits both ‘local spatial sensitivity’ (Yang et al., 2023) and ’global temporal dependence’ (Hussain et al., 2023). More specifically, ’local spatial sensitivity’ may be attributed to intersection designs that segregate different traffic flows (i.e., right turn, left turn, straight) resulting in localized clustering of conflicts at multiple locations (Chen et al., 2020). In contrast, ‘global temporal dependence’ may be linked to traffic signals, where the presence of signals introduces cyclical long-term dependence in risk at each location (Hussain et al., 2023). One significant gap in previous studies is that they typically model traffic risk based on either temporal or spatial features alone, without effectively integrating spatio-temporal information. This may be due to the limitations of both mathematical and statistical methods in simultaneously capturing this complex dependence.

2.2.2. Deep learning techniques for traffic risk modeling

Deep learning methods have been increasingly used for traffic risk modeling due to their ability to effectively capture complex

correlations within risk sequences (Bao et al., 2019; Huang et al., 2023; Li et al., 2020; Li et al., 2023; Mo et al., 2024; Yu et al., 2020). Convolutional Neural Networks (CNNs) are typically used to capture spatial dependence, while Long Short-Term Memory (LSTM) models are used to capture temporal dependence in sequences. For example, Yu et al. (2020) introduced a CNN to explore multi-dimensional, temporally and spatially correlated pre-crash operational features. Li et al. (2020) and Mo et al. (2024) introduced a LSTM-CNN model for real-time crash risk inference where the LSTM captured temporal dependence and the CNN extracted spatial features. Li et al. (2023) inferred the spatio-temporal locations of secondary crashes by proposing a hybrid deep learning approach that combined a Stacked Sparse Autoencoder (SSAE) with a LSTM network. Huang et al. (2023) presented PL-TARMI, a deep-learning framework leveraging the VGG-16 network for pixel-level traffic crash risk mapping, adeptly extracting multi-level features from both target and surrounding areas using a CNN-based approach. Additionally, Bao et al. (2019) proposed a spatio-temporal CNN-LSTM network (STCL-Net) for inferring citywide short-term crash risks.

Although these deep learning methods have achieved significant outcomes in modeling traffic risks, several issues still require comprehensive attention. Firstly, the scale of study units in previous literature is overly broad, encompassing periods like a day (Huang et al., 2023) or a few minutes (Li et al., 2020), or spaces like entire cities (Bao et al., 2019; Liu et al., 2023), which hinders the deployment of emerging technologies that require higher spatio-temporal resolution. Furthermore, methods used in coarse-grained macro-level studies are not transferable to intersections, as the evolution of conflict risks at the micro-level is rapid and complex. Despite advances like LSTM, training this Recurrent Neural Network (RNN) for learning global dependence remains challenging due to persistent issues with gradient vanishing and exploding (Wang et al., 2024a). Secondly, previous literature typically divided the experimental area into regular grids and then applied convolution operations to risk sequences treated as Euclidean-data (Bao et al., 2019; Hu et al., 2023; Wang et al., 2024a). However, due to the irregular structure of intersection, risk data exhibit non-Euclidean characteristics (Li et al., 2021) and partitioning the experimental area into regular grids and constructing Euclidean-data can inadvertently encompass weakly correlated zons in target area analyses, which may reduce the effectiveness of the approaches (Liu et al., 2023a).

Recently, the Transformer has emerged as a novel approach by utilizing an attention mechanism to process sequential data. Diverging from RNN-based methods, the Transformer enables access to any part of the historical data, regardless of distance, proving potentially more adept at capturing global and recurring patterns (Vaswani et al., 2017). Recent studies suggest that this model excels in simulating freeway conflict risk propagation, for example Wang et al. (2024a) utilized a Transformer to discern spatio-temporal dependence in conflict risks on freeways, with experiments demonstrating its superior performance in identifying complex spatio-temporal correlations. However, the Transformer may not be directly used to model the evolution of conflict risks at intersections on a micro-level due to two inherent limitations of the vanilla Transformer. The first is that the point-wise dot product self-attention, where each element in a sequence independently computes attention scores with every other element through dot products. This mechanism lacks sensitivity to local context, which can lead to anomalies in modeling local evolution of conflict risk at intersections. The second is the Transformer's space complexity that increases quadratically with the number of conflict nodes (Li et al., 2019), rendering it impractical for fine-grained inference tasks involving a large number of conflict nodes. The integration of Graph Neural Networks (GNNs) offers a promising solution for modeling 'local spatial dependence' (Tran et al., 2023; Wang et al., 2023a). Recent studies have shown that combining GNNs with attention mechanisms enhances the understanding of spatio-temporal dynamics and latent causality in risk sequences (Liu et al., 2023a), making the fusion of Transformers and GNNs a significant future direction for effective spatio-temporal modeling of traffic risks.

2.3. Summary

Traffic conflict events, more frequent and visible than crashes, provide crucial insights into crash mechanisms and they are being increasingly recognized as a proactive tool for risk inference and safety analysis (Chen et al., 2022;Essa and Sayed, 2018). Researchers argued that common driving errors underpin both conflicts and crashes, suggesting that mitigating traffic conflicts could reduce crash rates (Sacchi et al., 2013). Consequently, we use traffic conflict as a surrogate for traffic risk, leveraging inherent spatio-temporal dependence within risk sequences to infer future traffic safety at a fine-grained level.

A conflict can initiate a cascade of conflicts, particularly in zones with dense traffic (Xu et al., 2016; Wang et al., 2024a). In other words, conflict risks at intersections exhibit a regular pattern in the spatio-temporal space. For example, a 'space–time conflict heat map' (Essa and Sayed, 2018) or 'risk field' (Yang et al. 2023) are often employed to characterize the dynamic evolution of conflicts distribution at intersections. In this study, we identify intersection conflict-prone zones as conflict nodes, previously termed '**hotspots**' (Ferreira and Couto, 2015) or '**hazardous locations**' (Ghoul et al., 2023). These nodes are characterized by conflict risks which show the dynamic to propagate across the topological node network. This implies the propagation of conflict risks throughout a network of potential conflict nodes, which involves how changes in risk levels at a particular conflict node affect adjacent nodes. Moreover, this also implies the propagation of risk levels at a conflict node over a specific time frame in the future (Wang et al., 2024a). Consequently, modeling the evolution of conflict risk at intersection is a problem of capturing spatio-temporal features in a high-dimensional space:

$$x_i^{t+1} = \gamma(x_i^t, F_{i,j\in Pa(i)}(x_i^t, x_j^t)) \tag{1}$$

where the state of a single conflict node, denoted as x and defined as conflict risk value, is determined using the conflict risk model proposed by Yang et al. (2023) and detailed in Section 3.2. In Eq. (1), γ represents a differentiable function, such as a multilayer perceptron, $j \in Pa(i)$ signifies that the state change in conflict node j influences the state of conflict node i, and F symbolizes an aggregate function, which can take various forms such as sum, mean, or max, depending on the specific requirements of the analysis.

Influenced by physical barriers at intersections, the conflict risk of a node is only related to the risks of its neighboring nodes, with the correlation becoming more pronounced as proximity decreases (Yang et al., 2023).

3. Methodology

The framework is composed of three phases after collecting trajectory data from road users via the YOLO-DeepSort framework (Redmon et al., 2016; Hou et al., 2019; Wang et al., 2021). The first phase is the construction of a 'Conflict Risk Graph.' In the second phase, we develop a novel spatio-temporal deep learning approach to simulate the evolution of conflict risks. In the last phase, we conduct extensive experiments to guide the selection of an optimal node configuration. A graphic representation of the proposed framework, consisting of the three phases, is presented in Fig. 2.

3.1. Framework overview

Phase 1 is designed to identify potential conflict nodes from trajectory data and then create a 'Conflict Risk Graph.' The nodes incorporate both spatial and temporal dimensions, forming what can be described as a spatio-temporal cube, as shown in Fig. 3. These conflict events occurred at each cube are quantified in terms of conflict risks, which depict the state or characteristics of a node at a given time. The risk propagation network formed by these conflict nodes is presented in the form of graph-based data and referred to as the 'Conflict Risk Graph' throughout this study.

The introduce of 'Conflict Risk Graph' facilitates the propagation characterization of intersection conflict risk. Previous studies suggest that the propagation of conflict risks through the graph shows special dependence in the spatio-temporal space. In the spatial dimension, conflicts tend to cluster within certain zones of intersections. Chen et al. (2020) attributed this phenomenon to the intersection design that segregated different traffic flows (i.e., right turn, left turn, straight), resulting in localized clustering of conflicts at multiple locations. In these zones, proximity between conflict nodes intensifies their mutual influence, which then decreases with increasing distance, reflecting the 'local spatial sensitivity.' Recent studies referred to this pattern as Tobler's First Law of Geography (Song et al., 2020; Yang et al., 2023;Tobler, 1970), as shown in Fig. 4(left). In the temporal dimension, the correlation of conflict risks is not diminished over time. The effect of past conflict risks continues to influence present risks, showcasing a 'global temporal dependence,' as shown in Fig. 4(right). This global dependence can also be viewed as a periodic or long-term dependence influenced by

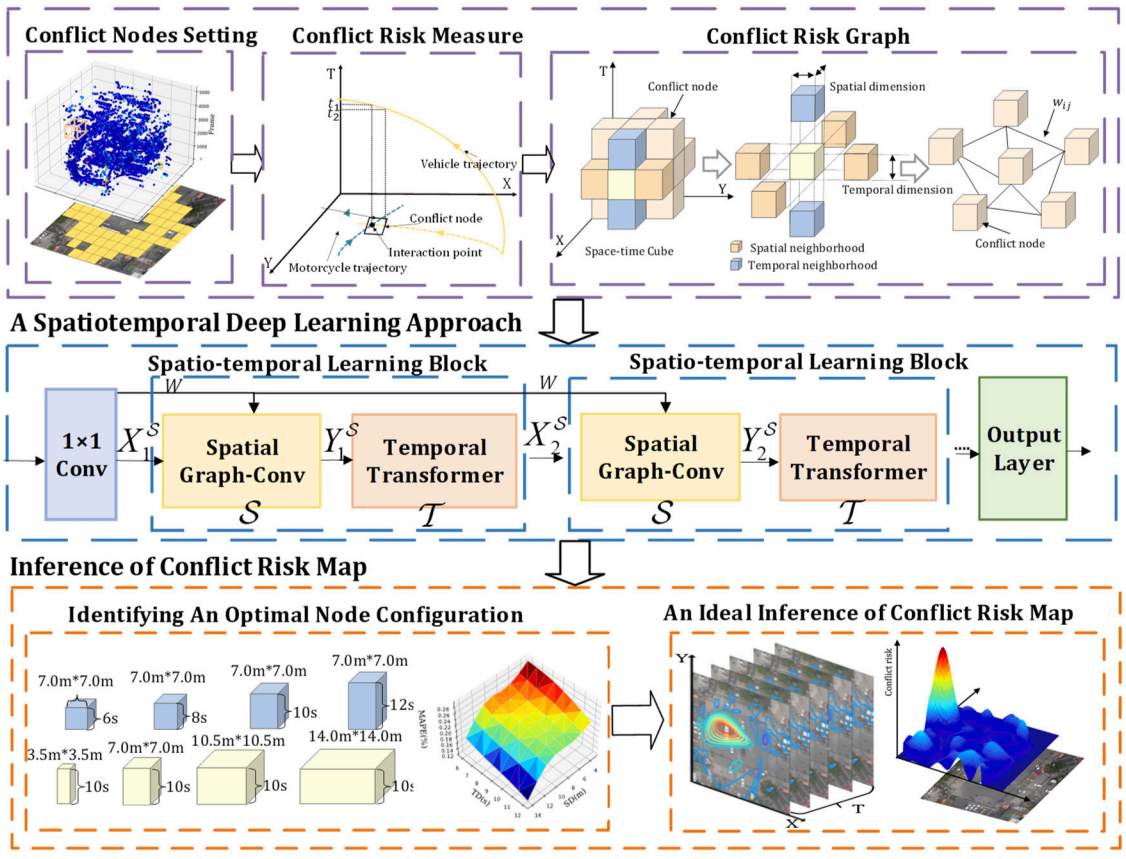


Fig. 2. Methodology flowchart.

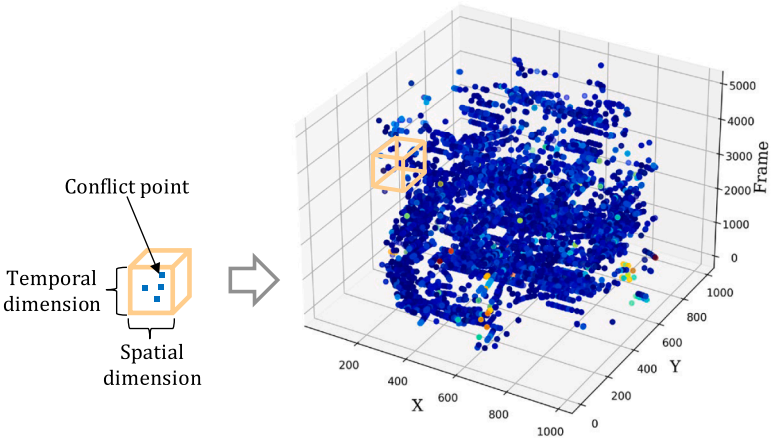


Fig. 3. Identification of conflict nodes from trajectory data.

traffic signals (Hussain et al., 2023). Thus, the essence of conflict risk evolution modeling lies not only in capturing the correlation of sequential data across time and space, but also in thoroughly considering both 'local spatial sensitivity' and 'global temporal dependence.'

Phase 2 is designed to address issues related to the fact that, since the 'Conflict Risk Graph' shows non-Euclidean characteristics, conventional CNN architectures are inadequate for capturing this 'local spatial sensitivity' (Yu et al., 2017). Additionally, RNNs are not well-suited for capturing 'global temporal dependence' due to the problems of gradient vanishing and exploding (Xu et al., 2020). These issues persist despite the development of various adaptations, such as LSTM and Gated Recurrent Unit (Xu et al., 2020). In this Phase, we effectively address these issues by developing a novel spatio-temporal deep learning approach named Trans-GCN. This approach synergically combines the strengths of GCN in extracting 'local spatial features' with the robustness of the Transformer in capturing 'global temporal correlations.'

Phase 3 is designed to reveal a correlation between node setup and model performance. In this Phase, our aim is to uncover the relationship between inference performance and node configuration. This work guides the selection of an optimal node configuration that balances the detailed capture of spatio-temporal dynamics with simulation accuracy, thereby achieving an ideal inference of conflict risk map.

3.2. Construction of 'conflict risk graph'

SSMs are pivotal for assessing traffic safety across specified temporal and spatial domains. Particularly, TTC and PET are prominent safety measures in this field. The former involves continuous evaluations for each conflict event, generating multiple risk values, whereas the latter simplifies the process by providing a single risk value per conflict event. This efficiency makes PET particularly advantageous for evaluating interactions among multiple road users within a given zone (Wu et al., 2020). Thus, PET is used to measure conflict risks at each conflict node in this study. This process starts by identifying potential interacting road users and pinpointing their likely interaction point. It then calculates the time interval between the exit of one user from the interaction point and the entry of the next user, providing a measure of critical safety events, as illustrated in Fig. 5. A PET value of less than a certain threshold is considered indicative of a conflict occurrence, and locations where such conflicts occur are selected as zones where candidate conflict nodes are located.

In Fig. 5, (X_j^t, Y_j^t) refers to the coordinates of road users at the moment they exit the interaction point, (X_i^{t+n}, Y_i^{t+n}) denotes the

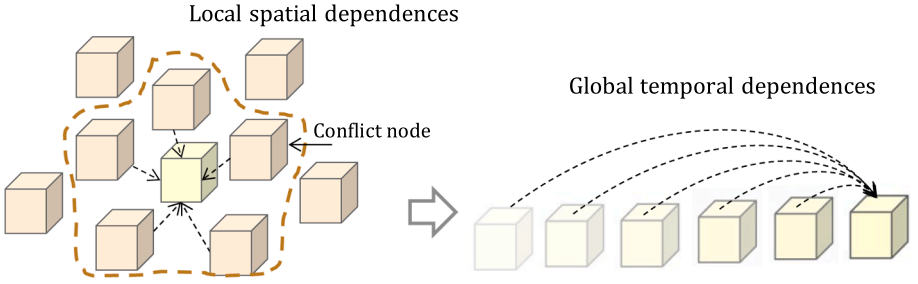


Fig. 4. Attributes illustration of conflict risk propagation.

coordinates of road users at the moment they enter the interaction point, and 'n' represents the PET value generated from this interaction, which can also be depicted as $t_1 - t_2$. The smaller the PET value, the more severe the conflict event, therefore a conflict severity indicator is used to map the time difference to a severity index, as shown in Eq. (2) (Autey et al., 2012; Saunier and Sayed, 2008). Considering that frequent conflicts may occur within a single node (as shown in Fig. 3), a Severity Index (*SI*) has been developed to measure the average severity of conflicts at each node (Yang et al., 2023):

$$SI = \frac{1}{N}\sum_{i=1}^{N} Exp\left(\frac{-PET_i^2}{2\sigma^2}\right) \tag{2}$$

Acknowledging the similar influences of both frequency (*Fr*) and severity (*SI*) of conflicts on traffic risks (Hu et al., 2023; Yang et al., 2023), we incorporate *Fr* and *SI* to identify the conflict risk value using a weighted average approach (Yang et al., 2023):

$$CR = \alpha \times Fr + \beta \times SI \tag{3}$$

An interaction where the PET value falls below a predetermined threshold are classified as conflicts. This threshold is set at 5 s to fulfill the objectives of testing the model based on previous studies (Beitel et al., 2018). The weights of *Fr* and *SI* (α and β) are determined based on their respective value scales. Referring to Yang et al. (2023), α and β are each set to 0.5 each to balance their value scales. Furthermore, after clarifying the impact of *Fr* and *SI*, users have the flexibility to choose different parameter settings (α and β) according to application-specific requirements.

Experimental areas are divided into multiple conflict nodes with specific spatio-temporal space, as shown in Fig. 6. In the spatial dimension, the interconnectedness between conflict nodes is conceptualized as graph-structured data, namely the 'Conflict Risk Graph'. Specifically, the feature of a conflict node, namely the conflict risk value, can be viewed as a graph signal defined on an undirected graph, denoted as $\mathscr{G}$ with weights w_{ij}. At the t-th time-step, in graph $\mathscr{G}_t = (C_t, \mathscr{E}, W)$, C_t represents a finite set of conflict nodes, corresponding to the conflict risk from n conflict nodes in the network; $\mathscr{E}$ refers a set of edges, representing the connectivity between conflict nodes; and $W \in \mathbb{R}^{n\times n}$ signifies the weighted adjacency matrix of $\mathscr{G}_t$. In the dataset, we compute the adjacency matrix of the 'Conflict Risk Graph' based on the distances between conflict nodes in the network, i.e.,

$$w_{ij} = \begin{cases} \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right), i \neq j \text{ and } d_{ij}^2 \leq \epsilon \\ 0, \text{ otherwise.} \end{cases} \tag{4}$$

where w_{ij} represents the edge weight from conflict node i to node j, d_{ij} refers the Euclidean distance from conflict node i to node j, σ is the standard deviation, and ϵ is the threshold manually setting for sparsing the weighted adjacency matrix.

3.3. A novel spatio-temporal deep learning approach

The Trans-GCN model consists of spatio-temporal learning blocks, each composed of a spatial graph convolution layer and a temporal transformer layer, as presented in Fig. 7.

Despite some studies proposing spatio-temporal deep learning approaches that combine GNNs and Transformers for tasks involving the capture of spatio-temporal dependence, such as traffic flow forecasting (Li et al., 2021; Liu et al., 2023b; Xu et al., 2020), methods used in these coarse-grained macro-level studies are not transferable to model intersection situations, as the evolution of conflict risks at the micro-level is rapid and complex.

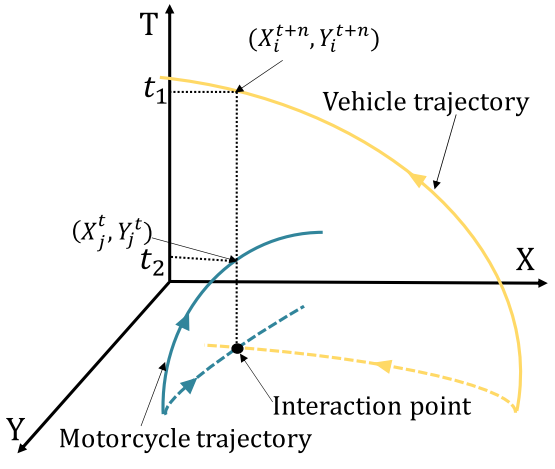


Fig. 5. Measuring conflict events based on PET.

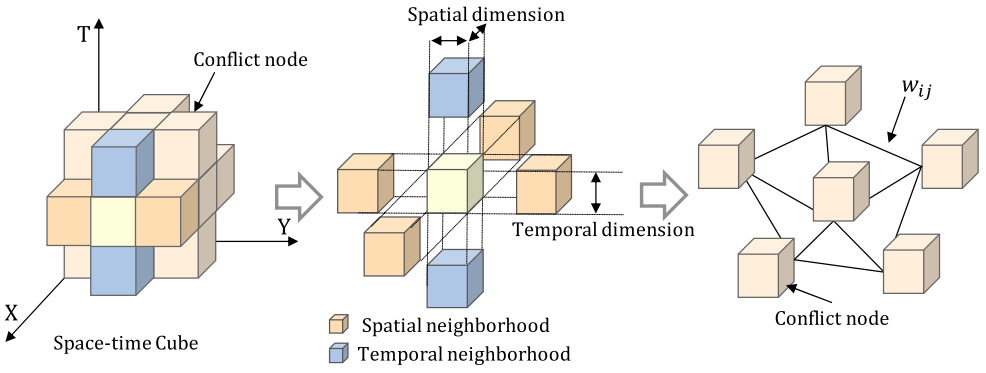


Fig. 6. Schematic diagram of conflict node network.

Our approach specifically targets the unique challenges of modeling conflict risk evolution at intersections, distinguishing it from macro-level traffic flow predictions. Unlike broader traffic systems where 'global spatial dependence' is crucial due to interconnected road segments (Liu et al., 2023b; Xu et al., 2020; Yan et al., 2022), intersection conflict risks rely on 'local spatial dependence' due to traffic signals and design features directing distinct traffic flows (Yang et al., 2023). Previous models like GNNs combined with self-attention mechanisms have effectively addressed global dependence in traffic systems by adapting to irregular road network topologies. However, these methods do not translate well to the micro-level complexities of intersection conflict risks, where local interactions predominate. Recognizing this, we propose a tailored spatio-temporal deep learning approach to accurately simulate conflict risks at intersections, addressing the specific challenges of locality and memory limitations noted in previous studies (Li et al., 2019).

3.3.1. *GCN for modeling 'local spatial sensitivity'*

The problem at hand is based on non-Euclidean data, making traditional CNNs unsuitable for analysis. Presently, a common method is spectral graph convolution which introduces the spectral framework to apply convolutions in spectral domains with graph Fourier transforms (Yu et al., 2018). In this work, we follow the approach and introduce the concept of a graph convolution operator $*\mathcal{G}$ based on the principles of spectral graph convolution. This involves the multiplication of a signal $x \in \mathbb{R}^n$ with a kernel Θ (Shuman et al., 2013),

$$\Theta *\mathcal{G}\mathrm{x} = \Theta(L)\mathrm{x} = \Theta\left(U\Lambda U^T\right)\mathrm{x} = U\Theta(\Lambda)U^T x \tag{5}$$

where the graph Fourier basis $U \in \mathbb{R}^{n\times n}$ represents the matrix of eigenvectors of the normalized graph Laplacian $L = I_n - D^{-1/2}WD^{-1/2} \in \mathbb{R}^{n\times n}$.$\Lambda \in \mathbb{R}^{n\times n}$ stands for the diagonal matrix of eigenvalues of L, and the filter $\Theta(\Lambda)$ is also a diagonal matrix. A graph signal x is filtered by a kernel Θ through the multiplication of Θ and the graph Fourier transform $U^T x$.

In our model, graph convolution is applied to graph-structured data, enabling the extraction of significant local patterns and dependence in the spatial domain. However, computing the kernel Θ in graph convolution as per Eq. (5) can be computationally intensive, involving multiple $\mathcal{O}(n^2)$ multiplications with the graph Fourier basis. To address this, we employ Chebyshev Polynomial approximation methods (Kipf and Welling, 2017) based on the study of Yu et al. (2018). Specifically, the kernel Θ is approximated as a polynomial of Λ in the form $\Theta(\Lambda)x \approx \sum_{h=0}^{H-1}\theta_h\Lambda^h$, where $\theta \in \mathbb{R}^H$ is a vector of polynomial coefficients. H denotes the kernel size of the graph convolution, defining the maximum radius of convolution from the central conflict nodes. Chebyshev Polynomial $T_h(x)$ is used

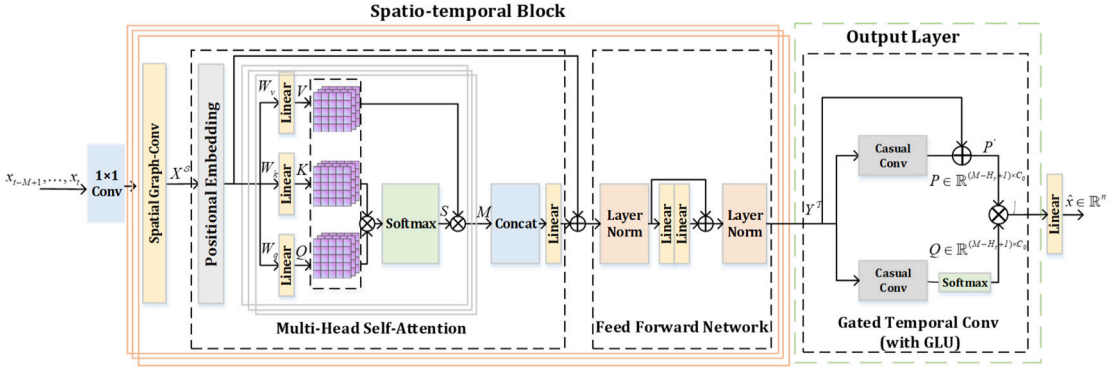


Fig. 7. Schematic illustration of the Trans-GCN.

to approximate kernels as a truncated expansion of order $H-1$ with rescaled $\widetilde{\Lambda} = 2\Lambda/\lambda_{\max} - I_n$ (Hammond et al., 2011). Consequently, the graph convolution can be reformulated as:

$$\Theta * \mathcal{G}x = \Theta(L)x \approx \sum_{k=0}^{H-1} \theta_h T_h(\widetilde{L})x \tag{6}$$

where $T_h(\widetilde{L})x \in \mathbb{R}^{n\times n}$ represents the Chebyshev Polynomial of order h evaluated at the scaled Laplacian $\widetilde{L} = 2L/\lambda_{\max} - I_n$. By applying recursive H-localized convolutions through this polynomial approximation, the computational cost of Eq. (5) can be significantly reduced to $\mathcal{O}(H|\mathcal{E}|)$, as presented in Eq. (6). According to the study of Kipf and Welling (2017), Eq. (5) can be simplified to the Eq. (7), where $\lambda_{\max}$ is taken 2, i.e.,

$$\Theta * \mathcal{G}x \approx \theta_0 x - \theta_1 (D^{-1/2} W D^{-1/2})\, x \tag{7}$$

where θ_0 and θ_1 represent the kernel' parameters. To constrain parameters and stabilize numerical performance, we use the parameter θ to replace θ_0 and θ_1.W and D are renormalized by $\widetilde{W} = W + I_n$ and $\widetilde{D_{ii}} = \sum_j \widetilde{W_{ij}}$ separately. The graph convolution can then be alternatively defined as:

$$\Theta * \mathcal{G}x = \theta(\widetilde{D}^{-1/2} W \widetilde{D}^{-1/2})x \tag{8}$$

Therefore, Eq. (8) achieve a similar effect to $H-$localized convolutions horizontally, all of which utilize features from the $H-1-$order neighborhood of the central conflict nodes.

Furthermore, the graph convolution operator $*\mathcal{G}$ can be adapted for multi-dimensional tensor applications. For signals with multiple channels $\mathrm{X} \in \mathbb{R}^{n\times C_i}$, the generalization of graph convolution is achieved as follows:

$$y_i = \sum_{i=1}^{C_i} \Theta_{ij}(L)\; x_i \in \mathbb{R}^n, 1 \le j \le C_o \tag{9}$$

where C_o and C_i respectively represents the output and input dimension of the feature maps. $\Theta_{ij}(L) \in \mathbb{R}^H$ represents the $C_o \times C_i$ vectors of Chebyshev coefficients. The graph convolution for 2-D variables is represented as $\Theta * \mathcal{G}\mathrm{X}$ with $\Theta \in \mathbb{R}^{H\times C_i \times C_o}$. More specifically, the input of conflict risk inferences consists of M time-steps of 'Conflict Risk Graph'. Each time-step x_t can be viewed as a matrix, with its i-th column representing the C_i-dimensional value of x_t at the i-th conflict node in the 'Conflict Risk Graph' $\mathcal{G}_t$, denoted as $\mathrm{X} \in \mathbb{R}^{n\times C_i}$. For each time-step t of M, an identical graph convolution operation using kernel Θ is applied in parallel to $\mathrm{X}_t \in \mathbb{R}^{n\times C_i}$ Consequently, this graph convolution approach can be extended to 3-D variables, indicated as '$\Theta * \mathcal{G}\mathcal{X}$' with $\mathcal{X} \in \mathbb{R}^{M\times n\times C_i}$. The conflict risk graph first passes through a 1×1 convolutional layer, where the conflict risk is encoded into a d-dimensional representation and then fed into the spatial learning block to extract local spatial features. Finally, the local spatial features extracted by the spatial-GCN are represented as $X^{\mathcal{S}} \in \mathbb{R}^{\mathrm{M}\times n\times d}$.

3.3.2. *Transformer for extracting 'global temporal dependence'*

Temporal patterns in conflict risk sequences can undergo significant changes over time, influenced by various factors like traffic signal cycles (Essa and Sayed, 2018; Hussain et al., 2023). Consequently, the state of a conflict node is globally linked to its surrounding context in the temporal dimension. However, RNN struggles to capture this global dependence due to the challenges of gradient vanishing and exploding (Xu et al., 2020). Even with advancements like LSTM and GRU, these fundamental issues remain unresolved. In comparison to RNNs and their variants, temporal Transformer can be easily scaled to long sequences with parallel processing of global dependence (Li et al., 2019).

After passing through the graph convolution layer, the local spatial features are encoded into a d-dimensional vector for each conflict node at each time-step, represented as $X^{\mathcal{S}} \in \mathbb{R}^{\mathrm{M}\times n\times d}$. In temporal learning block, we parallelize processing across conflict nodes to effectively model global temporal dependence. The 2-D tensor of local spatial features $\widehat{X} \in \mathbb{R}^{M\times d}$ is considered for any given conflict node in graph $\mathcal{G}$. We employ a temporal Transformer to model this long-term temporal dependence among conflict risk sequences. The input to the temporal Transformer comprises a temporal sequence with a sliding window of length M and d channels. In the temporal Transformer, global temporal dependence is dynamically computed within high-dimensional latent subspaces (Liu et al. 2021). These include the query subspace defined by $Q \in \mathbb{R}^{M\times d_A}$, the key subspace by $K \in \mathbb{R}^{M\times d_A}$, and the value subspace by $V \in \mathbb{R}^{M\times d}$:

$$\begin{cases} Q = \widehat{X} W_q \\ K = \widehat{X} W_k \\ V = \widehat{X} W_v \end{cases} \tag{10}$$

where $W_q \in \mathbb{R}^{d\times d_A}$,$W_k \in \mathbb{R}^{d\times d_A}$ and $W_v \in \mathbb{R}^{d\times d}$ are the weight matrices for the query(Q), key(K), and value(V) components, respectively. One-step inferences for x_{t+1} are concurrently derived from historical observations $x_{t-M+1},\ldots,x_t$. We incorporate the scaled dot-product function to account for bi-directional temporal dependence within $x_{t-M+1},\ldots,x_t$, i.e.:

$$S = softmax\left(\frac{QK^T}{\sqrt{d_A}}\right) \tag{11}$$

In Eq. (11), the dot product is used to reduce storage costs and computational complexity in the calculation. *Softmax* is employed to normalize the global temporal dependence, and $\sqrt{d_A}$ prevents saturation caused by the *Softmax* function (Vaswani at al., 2017). The obtained S serves as weights to further capture long-term temporal dependence.

RNN-based models predominantly focus on temporal dependence from preceding time-steps, and as indicated by Devlin et al. (2019), this left-to-right architecture is not ideal for modeling context dependence. To enhance this, we aggregate the V in Eq. (12) using the weights S for the long-term temporal features M (as shown in Fig. 8), i.e.,

$$M = SV \tag{12}$$

To investigate the interactions among latent features, a two-layer feed-forward neural network is developed for M, and the learning process of 'global temporal dependence' is presented in Fig. 8.

$$U = \mathrm{Re}Lu(M'W_0 + b_0)W_1 + b_1 \tag{13}$$

In this model, a residual connection $M' = M + \widehat{X}$ is employed to ensure stable training. W_0 and W_1 are the weight matrices for the three layers, and b_0 and b_1 are the biases. For each conflict node, the output is denoted as $\widehat{Y} = U + M' \in \mathbb{R}^{M\times d}$. Consequently, the output of the temporal Transformer is represented as $Y \in \mathbb{R}^{M\times n\times d}$ by collecting $\widehat{Y}$ across all the conflict nodes.

Overall, each layer of the temporal Transformer effectively captures 'global temporal dependence', as every time-step considers all other time-steps within the sliding window. The Transformer model can be efficiently scaled to accommodate longer sequences simply by increasing M, without significantly impacting computational efficiency (Vaswani et a., 2017; Wang et al., 2024a). In contrast, RNN-based models often struggle with vanishing or exploding gradients, and convolution-based models require an explicit increase in the number of convolutional layers proportional to M (Vaswani et a., 2017).

3.3.3. *Spatio-temporal learning block*

In the Trans-GCN, we develop a spatio-temporal learning block to integrate spatial-GCN and the temporal Transformer to jointly capture the global–local spatio-temporal dependence within conflict risk sequences, as shown in Fig. 2. The input to the l-th spatio-temporal learning block is a 3D tensor $\mathscr{X} \in \mathbb{R}^{M\times n\times d}$ of d-dimensional features for the n conflict nodes at time-steps t-M + 1,..., t. The spatial-GCN $\mathscr{S}$ and temporal Transformer $\mathscr{T}$ are stacked to generate the 3-D output tensor. In the l-th spatial–temporal block, the spatial-GCN $\mathscr{S}$ extracts local spatial features $Y_l^{\mathscr{S}}$ from the input node features $X_l^{\mathscr{S}}$ as well as graph adjacency matrix W, i.e.,

$$Y_l^{\mathscr{S}} = \mathscr{S}(X_l^{\mathscr{S}}, W) \tag{14}$$

$Y_l^{\mathscr{S}}$ is used as the input for the subsequent temporal Transformer, i.e.,

$$Y_l^{\mathscr{T}} = \mathscr{T}(Y_l^{\mathscr{S}}) \tag{15}$$

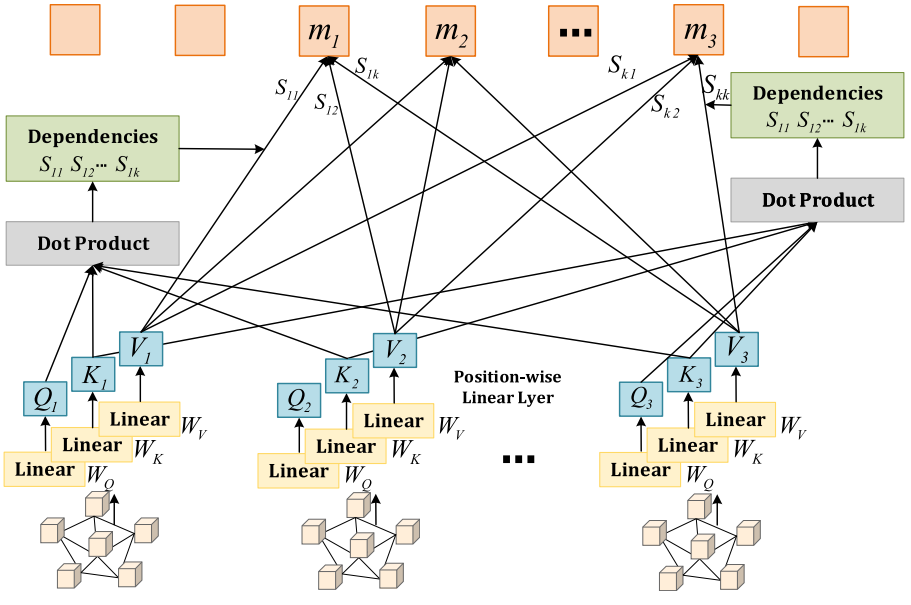


Fig. 8. Schematic illustration of the temporal Transformer.

Consequently, we obtain the output tensor $Y_l^{\mathscr{T}}$, which is also the input of $(l+1)$-th spatial–temporal learning block. Multiple spatio-temporal learning blocks can be stacked to enhance the model's capacity in accordance with the specific requirements of the tasks being addressed.

Following the stacking of spatio-temporal learning blocks, an additional temporal convolution layer coupled with a fully connected layer is appended as the final output layer. This temporal convolution layer transforms the outputs of the last block into a single-step inference. As presented in Fig. 7, the output layer is composed of a causal convolution with a kernel width of H_t, coupled with gated linear units (GLUs) for nonlinear activation. The causal convolution explores the H_t neighbors of input nodes for each conflict node, proceeding without padding to incrementally reduce the sequence length. Consequently, the input to the temporal convolution for each conflict node is a sequence of length M with d channels. The convolution kernel $\Gamma \in \mathbb{R}^{H_t \times d \times 2C_0}$ is specifically designed to transform the input $\widehat{Y}$ to a single output element $[P, Q]$ within $\mathbb{R}^{(M-H_t+1)\times 2C_0}$. Consequently, the operational framework of the temporal gated convolution can be defined as follows:

$$\widehat{J} = P' \odot \sigma(Q) \in \mathbb{R}^{(M-H_t+1)\times C_0} \tag{16}$$

where the symbol $\odot$ denotes the elementwise Hadamard product. In the GLU, the input is split equally into two halves, each with the same number of channels, serving as inputs for the gates in the GLU respectively. The sigmoid gate $\sigma(Q)$ plays a critical role in determining which components of the input P are essential for identifying compositional structures and dynamic variances within the conflict risk sequences. Furthermore, residual connections are implemented among the stacked layers of the temporal convolution to enhance information flow and learning stability. In a similar vein, the temporal convolution can be extended to 3-D variables by applying the same convolution kernel Γ to every conflict node $\widehat{Y}_i \in \mathbb{R}^{M\times d}$ in $\mathscr{G}$ noted as $\Gamma * \mathscr{T} Y$ with $Y \in \mathbb{R}^{M\times n\times d}$.Then, a final output $Z \in \mathbb{R}^{n\times c}$ is derived from the model, and conflict risk evolution modeling for n nodes are calculated by executing a linear transformation across $c-$channels, formulated as $\widehat{x} = Zw + b$. In this expression, $w \in \mathbb{R}^c$ represents a weight vector and b denotes a bias term. The $L2$ loss is used to evaluate the performance of our model, as formulated in Eq. (17). Consequently, the loss function for modeling spatio-temporal dynamics of conflict risks can be expressed as:

$$L(\widehat{x}; W_\theta) = \sum_t \|\widehat{x}(x_{t-M+1}, \ldots, x_t, W_\theta) - x_{t+1}\|^2 \tag{17}$$

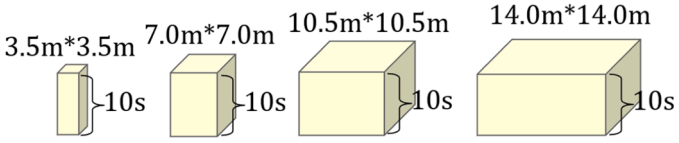
3.4. Conflict node configuration

In this study, experimental areas are segmented into a finite number of conflict nodes, each configured as a square bin with specific spatio-temporal space, as illustrated in Fig. 9. This spatio-temporal cube is often used for traffic safety analysis in previous studies. For example, Song et al. (2020) introduced a spatio-temporal node for modeling pedestrian-injury severities in pedestrian-vehicle crashes considering spatio-temporal patterns. Additionally, Wu et al., (2023) used a spatio-temporal node to investigate the crucial spatio-temporal features of crash risk hot spots on a road network. However, the configuration of the nodes has not yet been explored in terms of its impact on the inference performance. Thus, the correlation between nodes dimensions and model accuracy is investigated to identify the optimal node setup for conflict risk inferences.

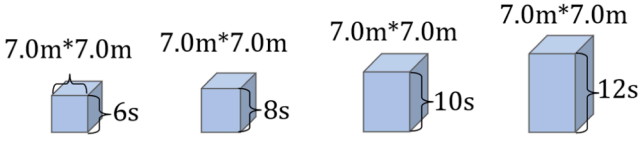
The characteristics of conflict nodes are represented as conflict risks, which consider both the frequency and severity of conflicts occurring within the nodes. The interconnections between these conflict nodes form a risk propagation network with a topological structure. We aim to explore the inference performance of conflict risk evolution under different node configurations, select the optimal configuration of conflict nodes, and achieve the ideal inference of conflict risk map.

4. Experiments

In this section, we present datasets, experiment settings and results.



(a) Conflict node configurations with different spatial dimension



(b) Conflict node configurations with different temporal dimension

Fig. 9. Conflict nodes with different configuration.

4.1. Data acquisition

This study aimed to model the evolution of conflict risk using rich trajectory data captured by drones in aerial videos. The selection of the experimental sites for the analysis required intersections with high traffic volume and frequent conflicts among road users. Thus, following a pre-investigation, the Jixiang (JX) and Qingsong (QS) intersections in Xi'an city, China, are chosen for the case experiment location as they met all the experimental requirements, as shown in Fig. 10. In both locations, the areas within 200 m—which include metro stations, parking lots, and shopping malls—attract a large number of road users who compete for road resources, resulting in frequent traffic conflicts. Moreover, these intersections handle high traffic volumes and host a diverse mix of road users, including motorcycles and other vehicles. According to local traffic authorities, these locations are prone to traffic crashes, particularly chain-reaction crashes.

The next phase of our experimental design aims at determining the optimal timing and duration for filming to ensure a substantial sample size. Given the interest in a timing with numerous conflicts, data collection was conducted during peak traffic hours: morning (7 am to 9 am), noon (11 pm to 1 pm), and evening (6 pm to 7 pm), using a systematic sampling method. For example, considering that the DJI Mavic 2 ZOOM drone's flight time is limited to approximately 25 min, we scheduled each flight cycle in the first 15 min of every half-hour interval to ensure there was enough battery for safe takeoff and landing. Over three days, the drone flew at an altitude of 100 m and records video at a speed of 25 frames per second (FPS), accumulating about 4.5 h of video data at each intersection.

The 'You Only Look Once' (YOLO) algorithm, known for its rapid processing speeds and high accuracy, is highly suitable for real-time trajectory extracting (Wang et al., 2023b;Wang et al., 2024b). In our work, we employ YOLO v3 (Redmon et al., 2016) along with DeepSort (Hou et al., 2019) within the TensorFlow framework to achieve real-time multi-object tracking. Existing object recognition models or datasets, typically tailored for oblique angles of fixed cameras, struggle with the top-down perspective of drone footage, as shown in Fig. 11(a). Thus, we train object recognition models specifically for drone views, with five volunteers annotating images using LabelImg 1.8.0. For more details on the computer vision model, refer to our previous works (Wang et al., 2024b).

The recognition performance of the trained computer vision model is illustrated in Fig. 11(b). In terms of tracking (See Fig. 11(c)), 'Current ['car']' indicates the number of cars in the current frame, while 'Total ['car']' represents the cumulative count of cars tracked up to that frame. In these figures, a coordinate system is established with the upper-left corner as the origin.

Given the frequent traffic conflicts in the interaction of motorcycle-vehicle, motorcycle-motorcycle and vehicle-vehicle at intersections (Chen et al., 2020), we use these interactions as a case study to validate the effectiveness of the proposed framework. Fig. 11 (d) shows the trajectory data of vehicles and motorcycle obtained through these computer vision models.

In this study, any interaction with a PET of less than 12 s was initially considered; those exceeding this threshold were excluded. The process of identifying interaction points among road users is illustrated in Fig. 11(e). Meanwhile, Fig. 11(f) uses darker red hues to highlight areas with a higher frequency of conflicts, indicating increased conflict risk. Notably, critical safety events cluster in certain areas of intersections, with the frequency decreasing from the center of these areas to the periphery, which is consistent with previous studies (Guo et al., 2019). Essentially, our study focuses on how these areas with high risks shift and change over time and space.

The descriptive statistics of interaction points at the two experiment sites are shown in Fig. 12. Specifically, a total of 19,287 vehicles and 48,378 motorcycles were tracked at the JX location, resulting in approximately 340,000 interactions among these road users. Similarly, 84,848 motorcycles and 84,904 vehicles were tracked at the QS location, producing a comparable number of interactions. The spatio-temporal location distributions of these interactions are shown in Fig. 12(a-1), with descriptive statistics presented in Fig. 12(a-3).

In this study, interactions with a PET of less than 5 s are classified as conflicts, according to Beitel et al. (2018). Fig. 12(a-2) and Fig. 12(b-2) respectively show the spatio-temporal location distributions of these conflicts in the two intersections. Overall, the JX location experienced approximately 280,000 conflict events, whereas the QS location saw about 90,000 such events.

(a) JX location (b) QS location

Fig. 10. Experiment sites.

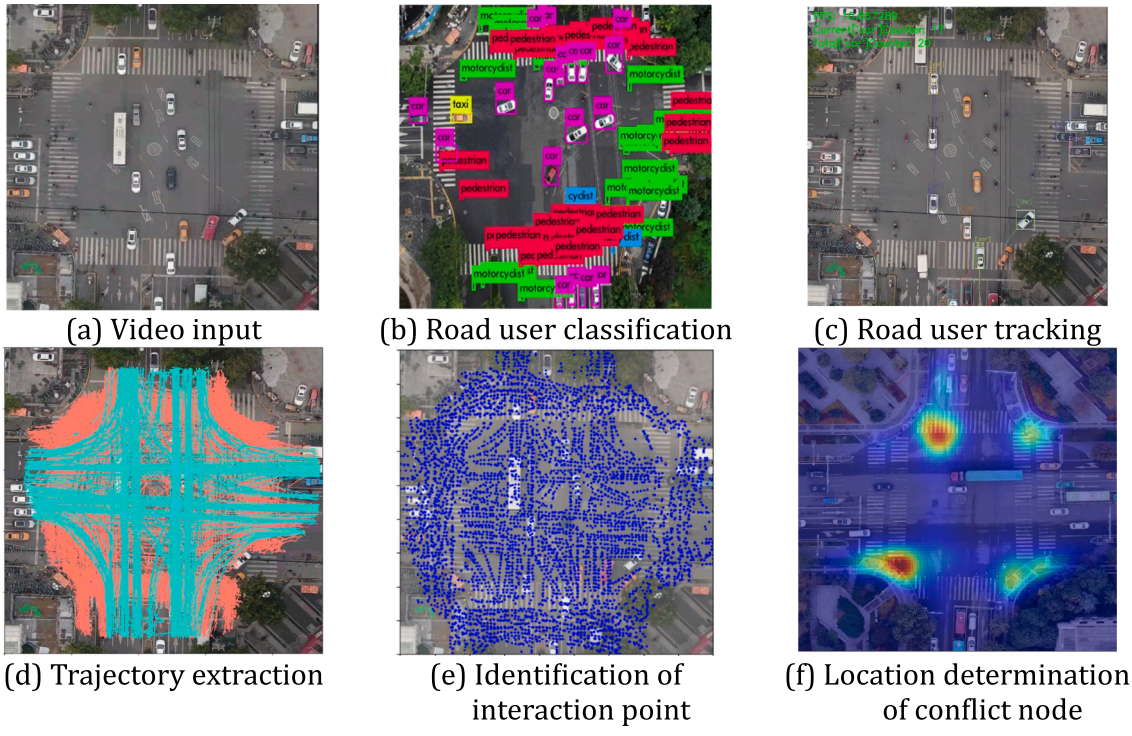


Fig. 11. Schematic diagram of dataset acquisition.

4.2. Model settings

The conflict risk evolution modeling was conducted on a computer equipped with an Intel i7-9700 k processor, NVIDIA GeForce RTX 2080ti graphics card, and 32 GB of memory, running Ubuntu 16.04.6 LTS. The historical time window for all tests was set to ten time-steps, meaning 10 observed time-steps (M = 10) were used to infer conflict risks for the next one-, two-, and three-time-steps (H = 1, 2, 3).

For hyper-parameter determination, we referenced previous studies (Liu et al. 2021; Vaswani et al., 2017; Yu et al., 2018). The graph convolution kernel size was set to 3. The number of channels in the GCN layers was fixed to 16. Additionally, the parameter *d* for the temporal Transformers was set to 128, with the number of heads at 2. The hyperparameters setting in training model are determined by the trial-and-error method (Gemperline et al., 1991). Specifically, we minimized the MSE loss using the Adam optimizer. The optimizer was configured with an initial learning rate of 0.001, a decay parameter of 0.9, and a batch size of 32. The dataset was distributed as follows: 70 % for training, 15 % for validation, and 15 % for testing. Several performance metrics were used to assess our approach's accuracy, including Mean Absolute Errors (MAE), Weighted Mean Absolute Percentage Error (WMAPE), and Root Mean Square Error (RMSE).

4.3. Comparison versus baseline models

We compared the performance of our proposed model with the following deep learning methods for benchmarking purposes (summarized in Table 1).

The Graph Convolutional Network (GCN) (Kipf and Welling, 2017) was designed to overcome traditional CNN limitations by employing spatial or spectral methods to extract features from irregular topological graphs, making it a widely adopted method for learning spatial dependence (Yu et al., 2017; Song et al., 2020). For our purposes, GCN is a representative method for extracting spatial correlations among sequences. This comparison aims to test the performance of merely extracting sequential spatial correlations in modeling the evolution of conflict risk.

The Gated Recurrent Unit (GRU) (Chung et al., 2014) is a deep learning technique for time-series modeling that has shown strengths in various prediction tasks. For our purposes, GRU is a representative method for learning temporal dependence. This comparison aims to test the performance of merely extracting sequential temporal correlations in modeling the evolution of conflict risk.

The Temporal Graph Convolutional Network (TGCN) (Zhao et al., 2020) integrates graph convolutional layers with GRUs, reflecting an excellent structure for spatio-temporal feature learning. For our purposes, TGCN is the benchmark for extracting spatio-temporal correlations among conflict risk sequences, however in isolation with respect to our model.

The Spatio-Temporal Graph Convolutional Network (STGCN) (Yu et al., 2018) is a state-of-the-art traffic forecasting model

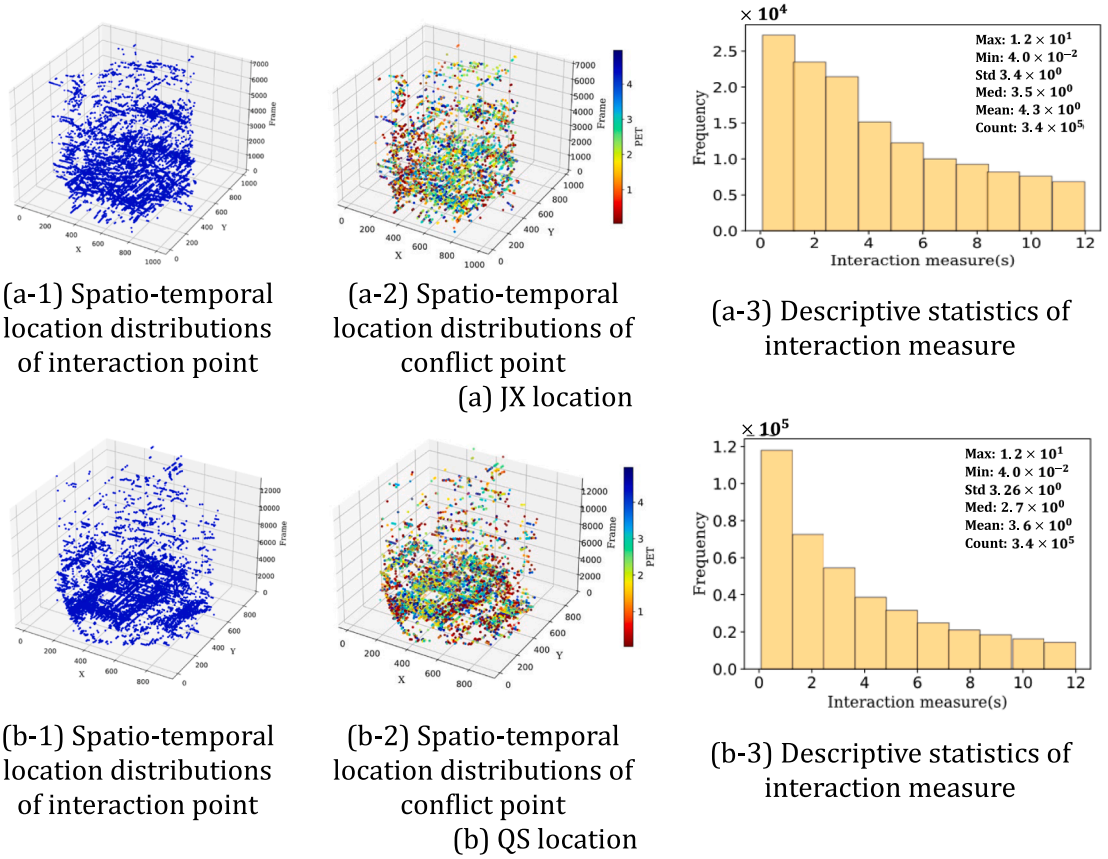


Fig. 12. Extraction and descriptive statistics of interaction events at experiment sites.

Table 1
Model comparison.

	Module of temporal dependence learning	Module of spatial dependence learning
GCN	/	GCN
GRU	GRU	/
TGCN	GRU	GCN
STGCN	CNN	GCN
STTN	Transformer	Transformer
Trans-GCN	Transformer	GCN

combining gated temporal convolution and spatial graph convolution. STGCN was used as a benchmark model and its hyperparameters were based on the study by Yu et al. (2018).

The Spatio-Temporal Transformer Network (STTN) (Wang et al., 2024a) is developed to simulates risk propagation patterns on freeways. In the STTN, the multi-head attention mechanism and stacked layers enable the Transformer to learn complex features in risk sequences. Spatial and temporal learning Transformer are combined to extract and fuse these features, resulting in a fine-grained freeway risk inference. STTN was used as a benchmark model to validate the performance of the proposed model and its hyperparameters were based on the study by Wang et al. (2024a).

In Trans-GCN, the Transformer is used to capture long-term dependencies across different time steps in the sequence, while GCN is employed to capture the local spatial dependencies between conflicting nodes. The architecture and hyperparameter settings of the sub-models are based on the study of Liu et al. (2021), Vaswani et al., (2017) and Yu et al., (2018).

We evaluated the models' inference performance across extensive experiments with various node configurations. Table A and Table B in the Appendix provide details on inference performance of Trans-GCN and benchmark models. Overall, our model outperforms the other models in most of the test cases, particularly on the JX datasets.

In Fig. 13, it summarizes the inference performance of various models across different temporal granularities. The subpar performance of standalone GCN and GRU models in all tasks presents the limitations of relying solely on temporal or spatial feature learning to model the evolution of conflict risks. This emphasizes the need for integrating both temporal and spatial dependence to

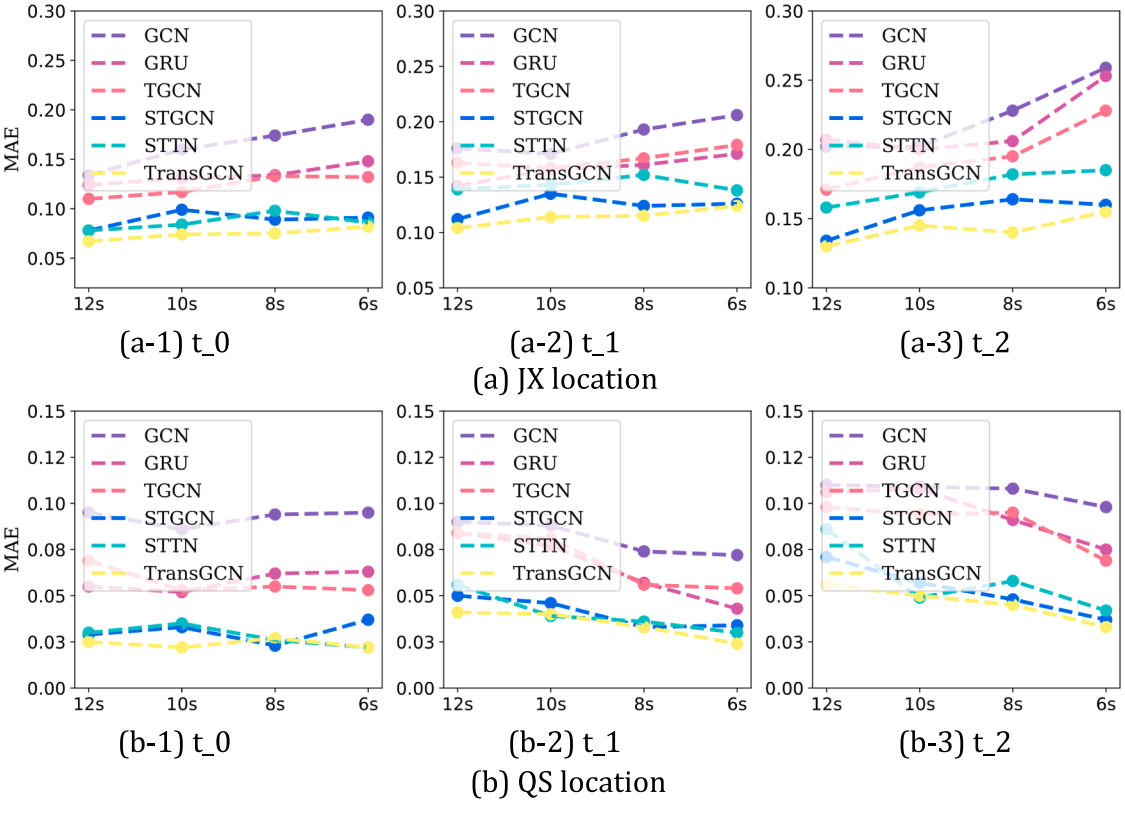


Fig. 13. Comparison of inference performance across different temporal granularity.

accurately model conflict risk evolution at intersections.

Comparative results present that our proposed model consistently outperforms the baseline models, including spatio-temporal deep learning approaches such as TGCN, STGCN, and STNN. This superior performance is evident at every stage of inference, including the first (t_0), second (t_1), and third time-steps (t_2), particularly on the JX dataset where the advantages of Trans-GCN are more pronounced than on the QS dataset. Notably, STTN does not perform as expected in intersection risk modeling, despite its proven advantages in highway conflict risk simulation (Wang et al., 2024a). For example, at the third time-step, Trans-GCN's average modeling accuracy surpasses that of STTN by 21 %. This difference may be attributed to the 'locality-agnostics' of Transformer (Li et al., 2020), where the point-wise dot product self-attention in the Transformer is insensitive to local context, which can make the model prone to anomalies in capturing local spatial correlations.

Additionally, Trans-GCN's average modeling accuracy exceeds that of T-GCN and STGCN by 36 % and 7 %, respectively. This performance gap likely stems from the feature learning capabilities of the temporal modules in models such as T-GCN, STGCN, and Trans-GCN, which respectively use GRU, CNN, and Transformer. These results suggest that Transformer are particularly effective in capturing dependence, outperforming both CNN and GRU. This advantage likely comes from the attention mechanism of Transformer, which focuses on all time-steps of the input sequence, enabling them to learn global temporal dependence more effectively than CNN and GRU. In contrast, CNN may struggle with longer sequences due to their fixed kernel sizes, limiting their ability to handle distant elements simultaneously (Wang et al., 2024a; Vaswani et al., 2017). Similarly, GRU may underperform because they process data sequentially, which restricts their capacity to address distant elements at the same time (Li et al., 2019). Therefore, these factors contribute to the Transformer achieving state-of-the-art (SOTA) results in the comparison of modeling across different temporal granularity.

Furthermore, results show that the inference performance of baseline models is unstable as temporal granularity decreases. In contrast, the Trans-GCN model demonstrates resilience against the negative impacts of reduced temporal granularity, consistently maintaining robust performance across modeling tasks. For example, in the initial modeling tasks, when the temporal granularity reduces from 12 s to 6 s, the MAE standard deviation for STTN is 0.0074, compared to 0.0053 for Trans-GCN. Among the baseline models, GCN exhibits the highest standard deviation of MAE, which is 0.02. This pattern is particularly evident in the first time-step task of the QS dataset, where the MAE standard deviation for Trans-GCN is 0.0024. For STTN, STGCN, TGCN, GRU, and GCN, the MAE standard deviations are 0.0048, 0.0052, 0.0067, 0.0046, and 0.0038, respectively. These results further highlight the Transformer's remarkable resilience in capturing global temporal correlations, despite changes in temporal granularity. This advantage likely comes from its inherent attention mechanism, which enables each time-step to consider all time-steps for its representation, effectively resisting the impacts of changes in temporal granularity. Overall, these findings suggest that our approach skillfully addresses the challenge of learning 'global temporal dependence' in modeling conflict risk evolution with different temporal granularity.

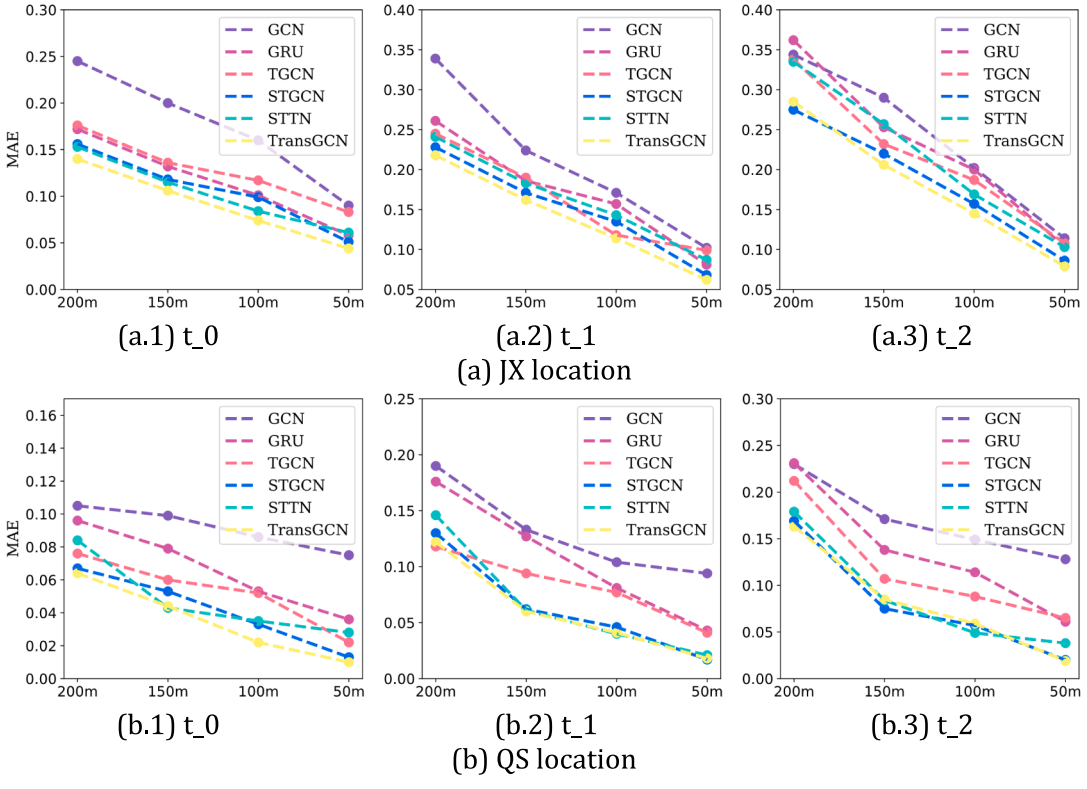


Fig. 14. Comparison of inference performance across different spatial granularity.

As shown in Fig. 14, it shows a comparison of inference performance across different spatial granularities. Similarly, the three spatio-temporal deep learning approaches outperform standalone temporal (GRU) or spatial learning modules (GCN) in all tasks. This further emphasizes the need for integrating spatio-temporal dependence in modeling conflict risk evolution. Fig. 14 also demonstrates that Trans-GCN outperforms these spatio-temporal deep learning approaches across various spatial granularities, with this advantage being more pronounced on the JX dataset. For example, at the first time-step, Trans-GCN's average modeling accuracy surpasses that of TGCN, STGCN, and STTN by 50 %, 18 %, and 17 %, respectively. On the QS dataset, Trans-GCN's average modeling accuracy also exceeds that of TGCN, STGCN, and STTN by 77 %, 26 %, and 66 %, respectively. Besides, in the fine-grained risk inference (node with a configuration of 3.5 m*3.5 m), STGCN also presents relatively superior performance. For example, in the third time-step inference on the QS dataset, STGCN achieves state-of-the-art (SOTA) results with an MAE of 0.019, matching that of Trans-GCN. This outstanding performance is also evident in other time-step inference tasks. However, in the fine-grained risk inference, STTN underperforms compared to STGCN and Trans-GCN. For instance, in the first time-step inference on the QS dataset, the MAE of STGCN and Trans-GCN is 0.14 and 0.18 lower than that of STTN, respectively. This advantage may be due to both models' use of GCN to capture spatial dependence, underscoring GCN's effectiveness in addressing local sensitivity.

In summary, these findings underscore Trans-GCN's proficiency in tackling challenges associated with local spatial sensitivity learning, particularly excelling in tasks requiring fine spatial granularity. The experiment results also demonstrate that the Trans-GCN model skillfully integrates the strengths of Transformer in capturing global temporal correlations with the ability of GCNs to model local spatial sensitivity. This synergy effectively mitigates the adverse effects of node variability, enabling a robust and fine-grained inference of conflict risks.

4.4. Inference performance

4.4.1. Inference performance of trans-GCN

The inference performance of three conflict nodes within the dataset under a node configuration of 10 s by a 7 m*7m is illustrated in Fig. 15. The observed fluctuation in conflict risk at these nodes demonstrates a complex pattern, potentially resulting from the interplay between intersection design and traffic signal timing (Hussain et al., 2023). Notably, there is a pronounced congruence between the modeled conflict risks and the empirical observations, showing the model's adeptness in capturing global temporal correlations. To further scrutinize the inference performance of the Trans-GCN, isopotential maps depicting conflict risks—also known as conflict risk maps—are generated and are displayed in Figs. 16-19. Each map features lines or curves that represent an equipotential surface, indicating uniform potential conflict risk along each contour. These conflict risk maps offer a quantitative assessment of safety levels and effectively delineate the dynamic nature of risk evolution, thereby addressing the limitations identified in previous crash-based studies (Liu et al., 2023). This is beneficial for real-time diagnostics of traffic safety at intersections, allowing for the restoration of normal traffic conditions from potentially hazardous states before crashes occur.

The actual conflict risk map over a spatial granularity of 7 m*7m is depicted in Fig. 16(a-d), while Fig. 17(a-d) presents the inferred risk map at the initial time-step, with darker reds indicating higher risks. In addition, Fig. 18(a-d) displays the actual conflict risk over a temporal granularity of 10 s, while Fig. 19(a-d) shows the inferred risk map at the first time-step. Despite some differences between the inferred and actual risk maps, Trans-GCN effectively identifies spatial detail with elevated conflict risks.

The results clearly show that conflict risk is unevenly distributed at intersections, which reveal some distinct phenomena. Firstly, high-risk areas display a clustering effect, characterized by a concentration of high risk at certain core locations that radiates outward. Yang et al. (2023) attributed this phenomenon to the geographic principle that conflict risks followed the First Law of Geography, spreading from the center to the periphery with the impact diminishing as distance increases. Secondly, high-risk areas consistently occur at intersections where right-turning vehicle flows intersect with pedestrian crosswalks. Chen et al. (2020) explained this occurrence by the stop-and-go nature of right-turning vehicles, which created a domino effect of conflicts. Additionally, the agility of motorcycles exacerbates conflicts at these locations, further highlighting the challenges at these specific intersection points.

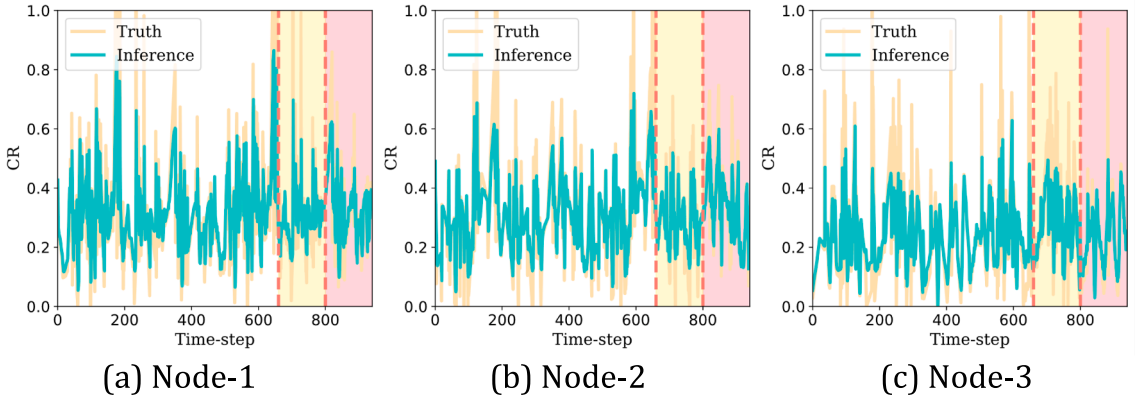


Fig. 15. Temporal inference performance of Trans-GCN on the dataset.

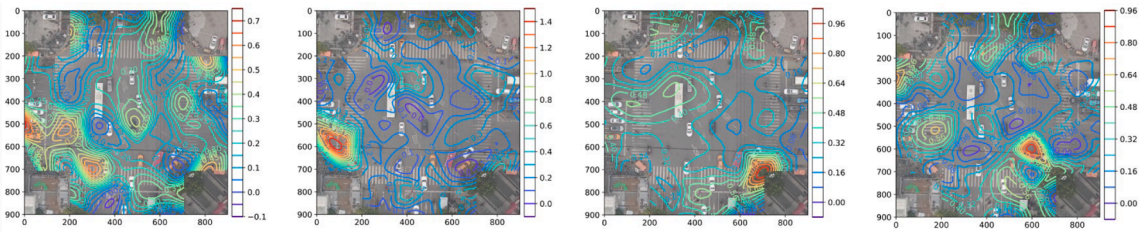


Fig. 16. Real conflict risk map with a node configuration of 7 m*7m.

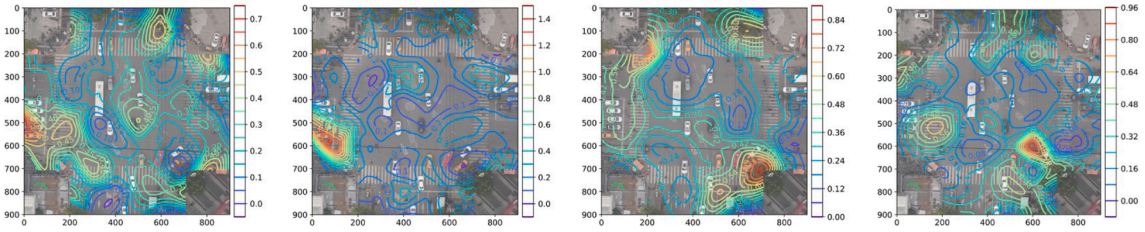


Fig. 17. Inferred conflict risk map with a node configuration of 7 m*7m.

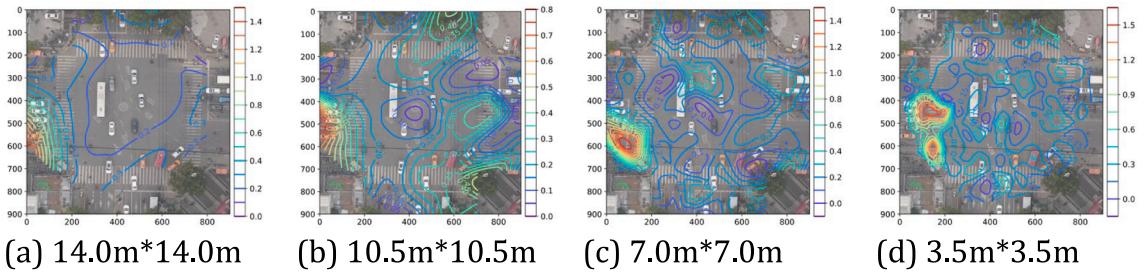


Fig. 18. Real conflict risk map with a node temporal dimension of 10 s.

Specifically, these conflict risk maps consistently show heightened conflict risk at coordinates (100, 600), compared to surrounding areas, as seen in Fig. 20. This pattern may result from expanding vehicle lanes and reducing non-motorized lanes near pedestrian crossings. As a result, as vehicles make right turns, the conflicts between motorized and non-motorized vehicles increase, leading to frequent acceleration and deceleration of vehicles. This contributes to a concentration of conflict points in this zone.

Additional case experiments are conducted to assess the generalizability of our model. Like the evaluations at the JX location, we tested the inference performance of the Trans-GCN on the QS dataset, as depicted in Figs. 21-24. The results suggest that the proposed approach achieves a consistent level of performance as observed on the JX dataset. These additional experiments not only confirm the generalization of our method but also demonstrate its effectiveness across multiple intersections.

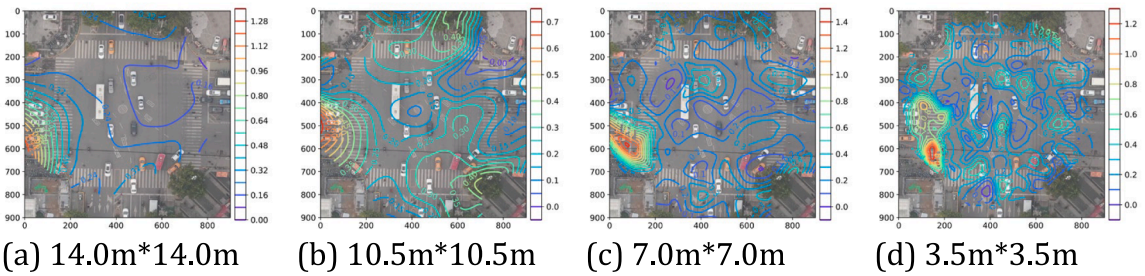


Fig. 19. Inferred conflict risk map with a node temporal dimension of 10 s.

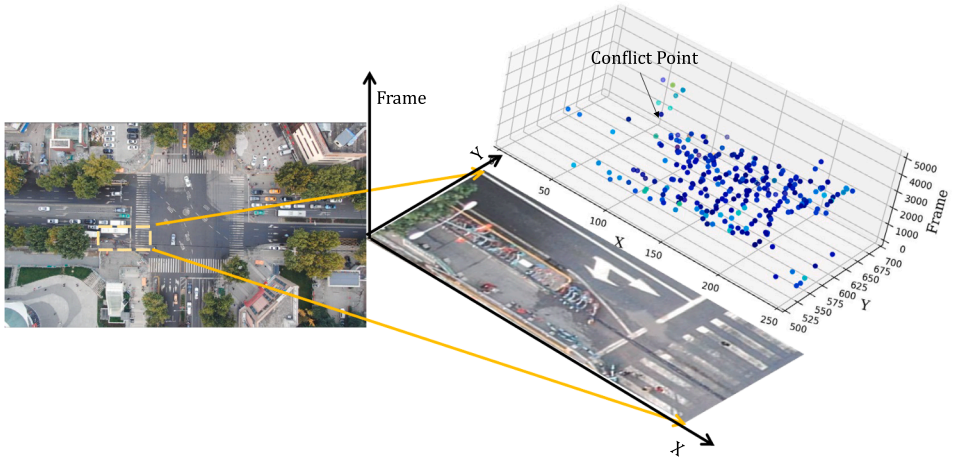


Fig. 20. Illustration of conflict point at the coordinates (0 ∼ 250, 500 ∼ 700).

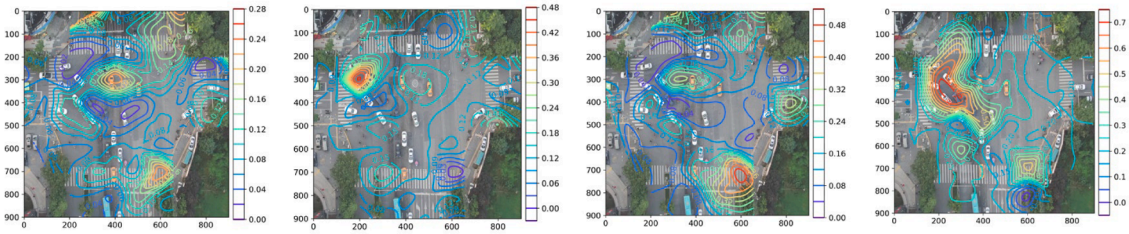


Fig. 21. Real conflict risk map with a node configuration of 7 m*7m.

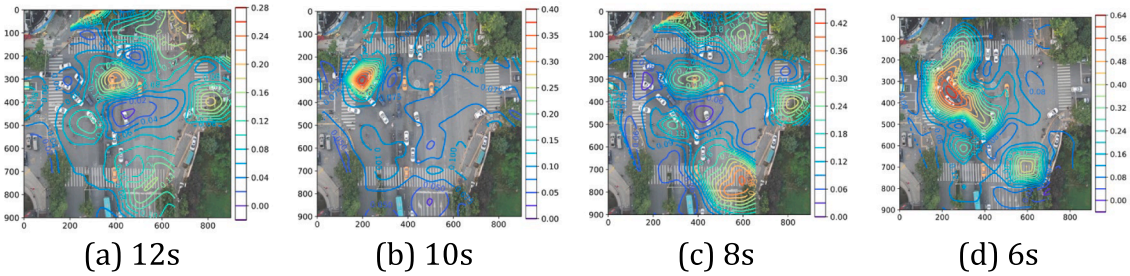


Fig. 22. Inferred conflict risk map with a node configuration of 7 m*7m.

Notably, these observations suggest that the spatio-temporal granularity or node dimensions significantly influence the visualization of conflict risks. For instance, the isopotential maps reveal that a higher spatio-temporal resolution more accurately represents the detail of conflict risks. Still, this enhancement in resolution may lead to a reduction in performance. In Section 4.4.2, we discuss how node configurations affect the performance of conflict risk modeling.

4.4.2. Inference performance with different node configurations

Selecting an appropriate node configuration is crucial for modeling the evolution of conflict risks. In this study, we use MAPE to measure inference performance under different node configurations, as shown in Table 2. A higher MAPE indicates poorer inference performance, and conversely, a lower MAPE suggests better performance. It is observed that as the spatio-temporal granularity increases, the models demonstrate superior performance. In contrast, reducing the spatio-temporal granularity, although providing more accurate information, comes at the cost of sacrificing modeling accuracy. This phenomenon aligns with previous study (Bao et al., 2019) and may be attributed to the fact that data at finer granularities were more susceptible to noise and random variability. These factors have more pronounced effects in smaller nodes, thereby disrupting the accuracy of model inferences. In contrast, larger

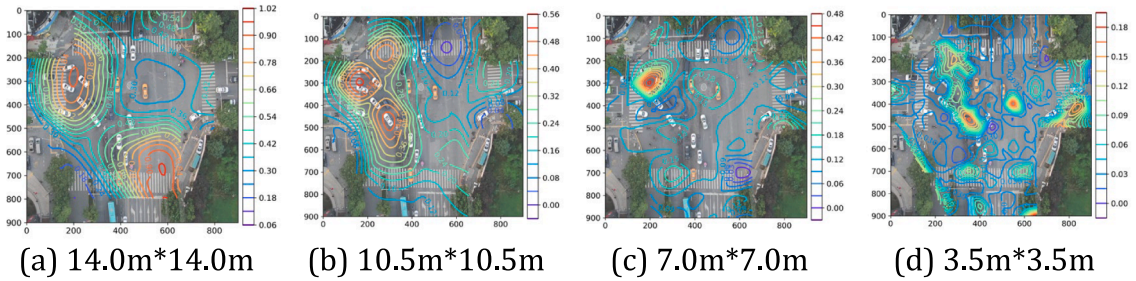


Fig. 23. Real conflict risk map with a node temporal dimension of 10 s.

granularities, by integrating a broader range of data, can smooth out these random variations and anomalies, thereby enhancing the stability and accuracy of the models (Bao et al., 2019). Nevertheless, opting for larger spatio-temporal granularities might lead to the loss of some local details. Therefore, our subsequent discussion will focus on selecting an optimal node configuration for modeling the evolution of conflict risk, balancing both spatio-temporal granularity and inference performance.

The inference performance under various node configurations is shown in Fig. 25. At the '10.5 m*10.5 m' granularity, both JX and QS datasets exhibit relatively low MAPE values across different temporal granularity, indicating optimal inference performance. While smaller spatial granularity (such as 3.5 m*3.5 m) provide higher detail, their relatively high MAPE values indicate greater model inference errors. On the other hand, larger spatial granularity (such as 14.0 m*14.0 m) has lower MAPE values but lack sufficient detail. Hence, a balance must be struck between detail and inference performance.

For the JX group at the initial time-step, the Mean Absolute Percentage Error (MAPE) values for temporal granularities of 6 s, 8 s, 10 s, and 12 s are 0.252, 0.230, 0.220, and 0.204, respectively. Similarly, for the QS group at the initial time-step, the MAPE values for these temporal granularities are 0.297, 0.293, 0.266, and 0.227. These data suggest that the '10.5 m*10.5 m' configuration performs well across various temporal granularity. Compared to the '14.0 m*14.0 m' configuration, it provides higher spatial detail while maintaining low MAPE values, which is beneficial for applications requiring a certain level of detail. Additionally, the '10.5 m*10.5 m' configuration shows consistently optimal performance on both JX and QS datasets, indicating it is a reliable and balanced choice.

When selecting the temporal granularity, a similar balance between temporal detail and inference performance is required. For the

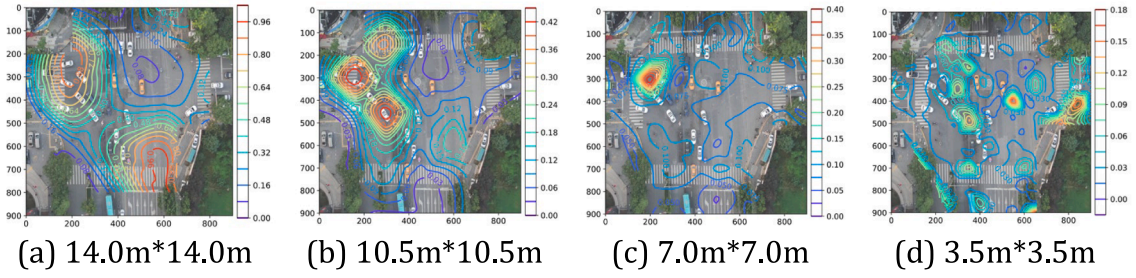


Fig. 24. Inferred conflict risk map with a node temporal dimension of 10 s.

Table 2
Relationship between node configuration and inference performance (MAPE).

TD(s)	SD(m)	t_0(JX/QS)	t_1(JX/QS)	t_2(JX/QS)
6 s	3.5 m*3.5 m	0.295/0.412	0.307/0.419	0.319/0.419
	7.0 m*7.0 m	0.254/0.339	0.273/0.378	0.305/0.378
	10.5 m*10.5 m	0.252/0.297	0.266/0.324	0.271/0.324
	14.0 m*14.0 m	0.177/0.238	0.199/0.275	0.215/0.275
8 s	3.5 m*3.5 m	0.286/0.417	0.304/0.423	0.316/0.423
	7.0 m*7.0 m	0.235/0.328	0.269/0.349	0.271/0.349
	10.5 m*10.5 m	0.230/0.293	0.247/0.322	0.260/0.322
	14.0 m*14.0 m	0.156/0.220	0.198/0.257	0.214/0.257
10 s	3.5 m*3.5 m	0.265/0.400	0.276/0.414	0.287/0.414
	7.0 m*7.0 m	0.224/0.290	0.249/0.321	0.257/0.321
	10.5 m*10.5 m	0.220/0.266	0.238/0.274	0.250/0.274
	14.0 m*14.0 m	0.142/0.196	0.169/0.224	0.188/0.224
12 s	3.5 m*3.5 m	0.233/0.382	0.249/0.399	0.258/0.399
	7.0 m*7.0 m	0.208/0.270	0.228/0.292	0.241/0.292
	10.5 m*10.5 m	0.204/0.227	0.220/0.243	0.230/0.243
	14.0 m*14.0 m	0.120/0.174	0.143/0.186	0.165/0.186

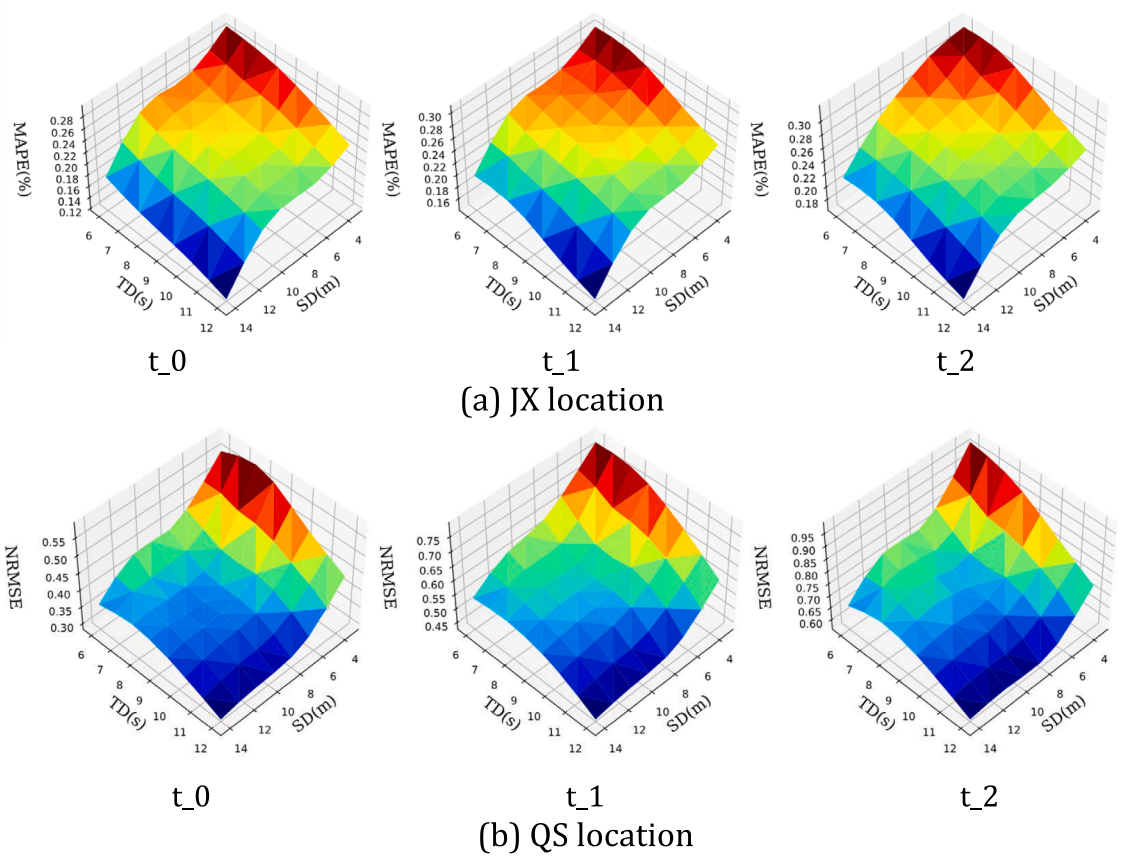


Fig. 25. Modeling performance across different node configuration.

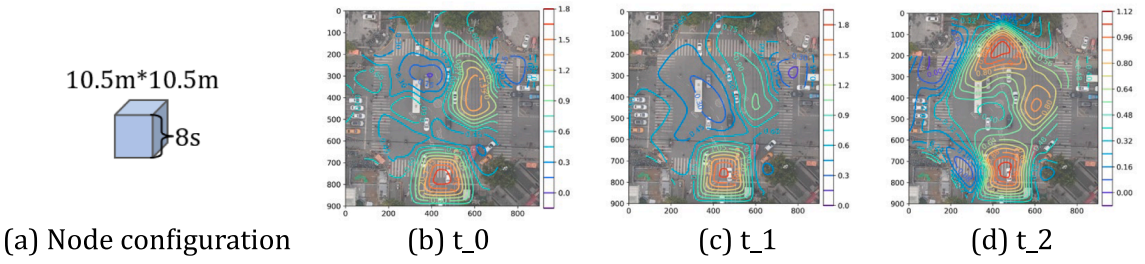


Fig. 26. Real conflict risk map with optimal node configuration at JX location.

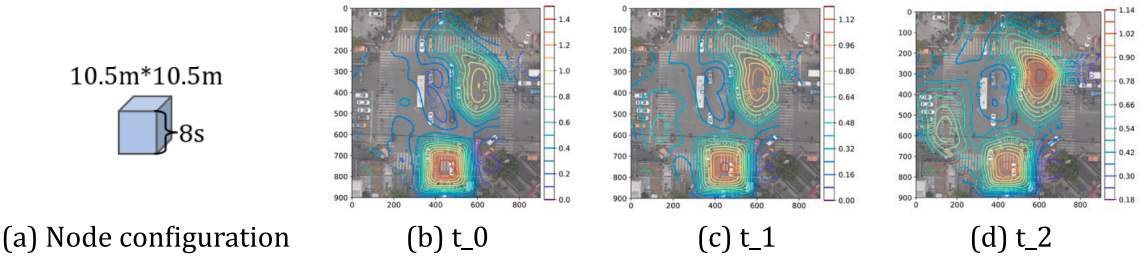


Fig. 27. Inferred conflict risk map with optimal node configuration at JX location.

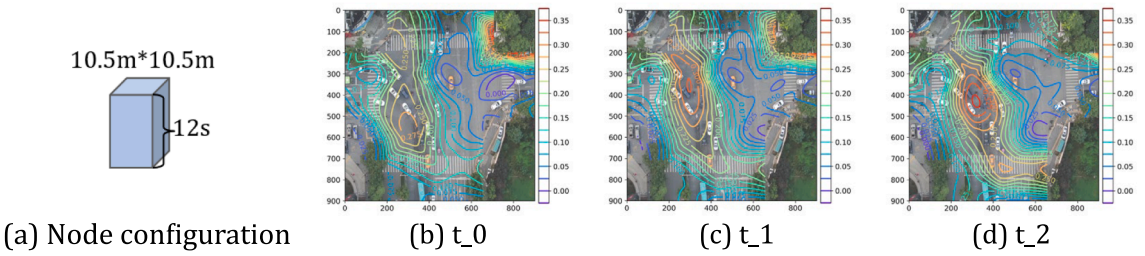


Fig. 28. Real conflict risk map with optimal node configuration at QS location.

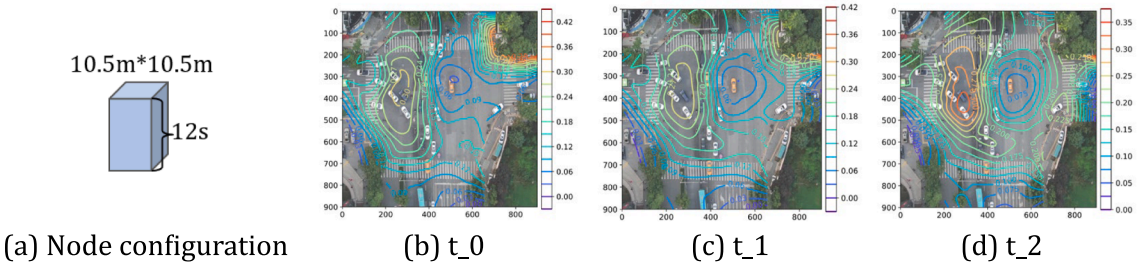


Fig. 29. Inferred conflict risk map with optimal node configuration at QS location.

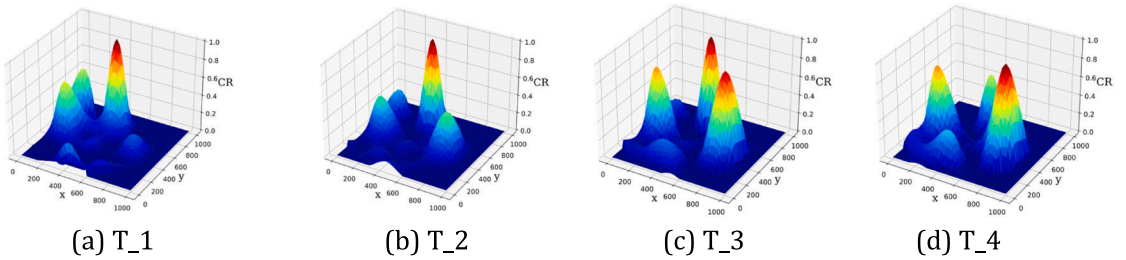


Fig. 30. Real conflict risk map with optimal node configuration at JX location.

JX group, although the 12 s granularity has the lowest MAPE value (0.204), we chose 8 s as the optimal granularity because it provides more temporal detail while maintaining a relatively low MAPE value (0.230). For the QS group, despite the 12 s granularity having the lowest MAPE value (0.227), we selected 10 s as the optimal temporal granularity because it offers more temporal detail with a relatively low MAPE value (0.266). Therefore, the choice of '10.5 m*10.5 m' granularity and appropriate temporal granularity ensures more spatio-temporal detail while maintaining low MAPE values, demonstrating excellent inference performance and representing the best balance between detail and modeling efficiency.

4.4.3. Inference performance with optimal node configuration

The inference results of conflict risk modeling under the optimal node configuration are displayed in Figs. 26-29. This node

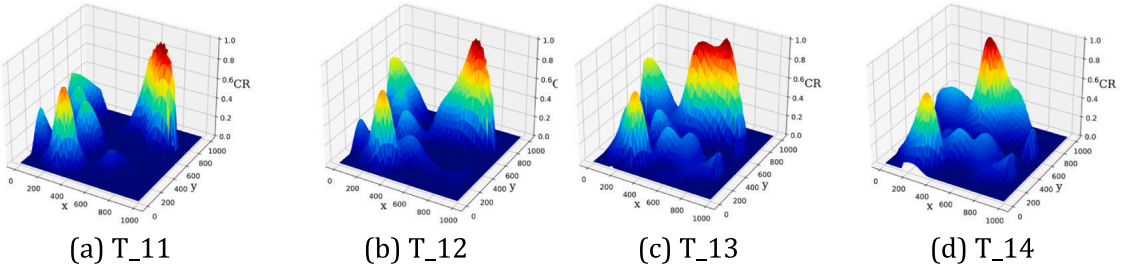


Fig. 31. Real conflict risk map with optimal node configuration at QS location.

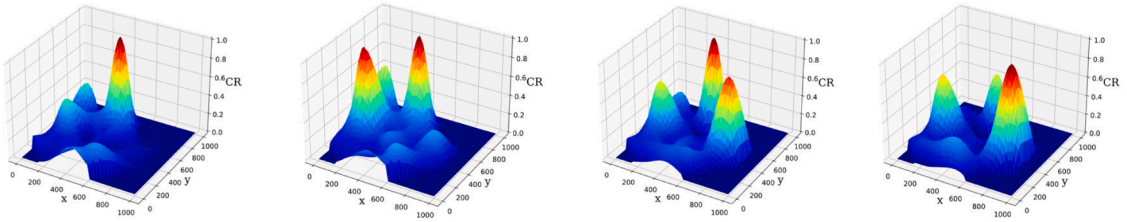


Fig. 32. Inferred conflict risk map with optimal node configuration at JX location.

configuration balances the detailed capture of spatio-temporal dynamics with inference performance. It is noteworthy that, although the performance of risk inference decreases over time, the model still effectively delineates the evolution process of high-risk locations. With the optimal node configuration, the proposed model successfully captured the spatio-temporal patterns of conflict risk evolution, achieving the ideal inferences of conflict risk maps.

The dynamic evolution of conflict risk at the JX location from T = 1 to T = 4, and at the QS location from T = 11 to T = 14 is shown in Figs. 30-31. During this period, several high-risk areas become apparent at the JX location, notably at coordinates (400, 100), (200, 400), (800, 400), and (400, 800). Of these locations, the most significant shift occurs at coordinate (800, 400), where risk values surge from 0.1 to 0.8, and at coordinate (400, 800), where they decrease notably from 0.8 to 0.4. At the QS location, high-risk areas are clearly visible at coordinates (200, 200), (0, 1000), and (800, 800). In these areas, the risk gradually spreads and propagates, which leads to an incremental increase in risk levels in the surrounding regions. Notably, high-risk points are predominantly located near the stop line, a finding consistent with the study of Essa and Sayed (2018). They attributed this to the initial red-light phase, where shorter queues lead to most conflicts occurring just behind the stop line. As the queue lengthens, the focal point of conflict intensity gradually shifts away from the stop line. Then, as the green light begins and the queue disperses, the conflict hotspot progressively decreases. Due to traffic signal control, this pattern of fluctuating conflict risk near the stop line is a recurring phenomenon in all directions of the intersection (Essa and Sayed, 2018).

Overall, these maps illustrate two phenomena: the long-term gradual change of high-risk areas over a period (32 s). The other significant phenomena are that conflict risks exhibit local clustering patterns. This indicates a tendency for risks to accumulate or concentrate in specific areas or under certain conditions, rather than distributing evenly or randomly. These phenomena further substantiate the presence of 'local spatial dependence' and 'global temporal dependence' in the risk evolution process.

We tested the model's performance in inferring long-term risk evolution at the first inference time step, as shown in Figs. 32-33. Results show that the proposed model successfully captured the global–local dependence in the evolution of conflict risk, providing the ideal characterizations and inferences of conflict risk. These risk maps with optimal node configurations effectively infer these nuanced fluctuations, offering valuable insights for the dynamic adjustment of intelligent traffic signals to mitigate the accumulation and spread of conflicts.

4.5. Discussion

Experiments indicate that the evolution of conflict risk at intersections exhibits distinctive spatio-temporal dependence. In the spatial dimension, conflict risks show distinct clustering patterns, signifying that critical safety events are not randomly or evenly distributed but tend to occur frequently in specific areas. Same phenomenon is also showed in the study of (Chen et al. 2020; Essa and Sayed, 2018; Yang et al. 2023). Yang et al. (2023) attributed this phenomenon to the geographic principle that conflict risks followed the First Law of Geography, spreading from the center to the periphery with the impact diminishing as distance increases. Conversely, in the temporal dimension, dynamic evolution patterns of conflict risk at intersections exhibit global dependence. For example, the conflict risk in a particular area gradually changes throughout the study period. This finding aligned with the conclusions drawn by Hussain et al. (2023), who attributed it to the cyclical and regular fluctuations of conflict risk at intersections, induced by traffic signals. Overall, the global–local spatio-temporal dependence in the evolution of intersection conflict risks is intricate and involves complex interactions between space and time, which surpasses the modeling capabilities of traditional mathematical/statistical methods.

We developed the Trans-GCN model, blending the Transformer's prowess in learning global dependence with the GCN's expertise in extracting local dependence. Our comparative evaluations with four advanced deep learning models—GRU, GCN, TGCN, STGCN and STTN—present Trans-GCN's exceptional ability to skillfully capture the intricate evolution patterns of conflict risks. Furthermore, case experiments also show that the proposed model continues to perform excellently on tasks at a fine-grained level, demonstrating

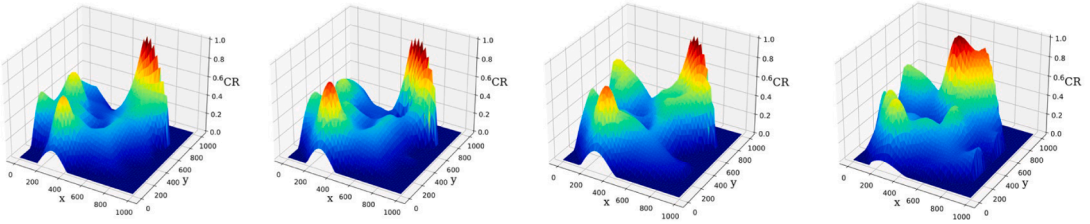


Fig. 33. Inferred conflict risk map with optimal node configuration at QS location.

outstanding resilience. However, experiments suggest that the application of another Transformer-based models (STTN) to capture the spatial dependence of risks has not produced satisfactory outcomes. This may be attributed to the 'locality-agnostics' of Transformer (Li et al., 2020), where the inference of each node's conflict risk utilizes all nodes. Specifically, the point-wise dot product self-attention in the Transformer is insensitive to local context, which can make the model prone to anomalies in the task of conflict risk modeling; furthermore, the space complexity of the canonical Transformer grows quadratically with the number of conflict nodes, rendering the direct modeling of fine-grained tasks infeasible.

In contrast, the success of Trans-GCN can likely be attributed to the distinct strengths of GCN and Transformer in capturing features. Specifically, GCN updates a node's feature representation by aggregating information from its neighboring nodes. This aggregation mechanism allows each node to effectively capture the structural information of its immediate vicinity (local neighborhood). Moreover, in GCN, all nodes in the graph share the same convolutional parameters. This weight-sharing mechanism not only reduces the complexity of the model but also ensures that the model can learn local patterns between nodes, independent of their position in the graph (Kipf et al., 2017). On the other hand, the core of the Transformer is its self-attention mechanism, which allows the model to directly establish dependence between any two positions in a sequence, regardless of the distance between them. This mechanism enables the Transformer to capture long-distance dependence and understand the relationships between elements at far-apart positions in the sequences. Thus, when computing the representation of the current time-step, the model considers information from other time-steps in the sequence, endowing it with a global awareness capability (Vaswani et al., 2017). These factors collectively contribute to the robustness of the Trans-GCN model in modeling tasks with different conflict node configurations.

Extensive experiments tested the inference performance across extensive node configurations. The findings indicate that increasing the spatio-temporal resolution of the modeling comes at the cost of reduced model performance. A plausible explanation for this trend is that data at finer granularities are more susceptible to noise and random variability. These factors significantly impact smaller nodes, disrupting the accuracy of model inferences. In contrast, larger granularities integrate a broader range of data, smoothing out random variations and anomalies, thereby enhancing model stability and accuracy (Bao et al., 2019). These findings suggest a balance between inference performance and node configuration. We selected an optimal node configuration that, while ensuring relatively good model performance, provides additional spatio-temporal details, ultimately achieving an ideal conflict risk map. This work provides a fine-grained and quantitative analysis of intersection traffic safety, addressing gaps identified in previous studies and offering significant potential to enhance the deployment of emerging technologies.

This work can be applied in the field of proactive safety management and monitoring. Several potential applications of the framework can be summarized as follows. First of all, the proposed framework can be viewed as a stepping stone in the process of diagnosing intersection safety risks. In the era of smart intersection, the enhanced capability of obtaining and processing trajectory brings new opportunities to leverage these newly acquired data for proactive safety analysis. From the perspective of detecting unsafe traffic dynamics proactively, the evolution evaluation of this traffic dynamic and the methodological advancement of introducing CRG and Trans-GCN to the field of traffic safety may open new opportunities to leverage largely unexploited trajectory data for improving the current intersection safety conditions and eventually leading to reduce unnecessary injuries and loss of life.

The second application is that one key component of proactive traffic safety management systems is the ability to infer future safety levels and mitigate risks before crashes occur by restoring traffic dynamics to normal through intervention strategies. In this system, an adaptive traffic signal control technology is essential as it adjusts traffic signals in real time to balance efficiency and safety (Stevanovic et al., 2013). The method proposed in this paper enhances the deployment and implementation of this technology as the proposed model captures the spatio-temporal dependence in the conflict risk evolution process and effectively infers future safety levels. This facilitates timely signal adjustments and mitigates risks before they evolve into crashes.

Another application is that a fine-grained traffic safety inference can directly influence driver decisions through intelligent

navigation systems, such as providing recommended nodes that avoid zones with high conflict risks. Specifically, by analyzing immediate trajectory data, the conflict risk at each node is calculated in real-time, followed by evolution simulation. Subsequently, the conflict risk value for each node is displayed in real-time to provide road users with immediate alerts about risk levels, which can help road users choose safer nodes to navigate through intersections. At complex intersections, these warnings and strategic recommendations provided by intelligent navigation systems will become an indispensable part of future traffic safety management, contributing to a more intelligent and safety driving experience.

5. Conclusion

We treat traffic conflicts as surrogates for traffic risks, utilizing the inherent spatio-temporal dependence of risks to model the evolution patterns of conflict risk, ultimately achieving an ideal, fine-grained, and quantified map of conflict risks at intersections, which fills up the significant gaps identified in the literature review and developd a set of proactive traffic safety management systems. The contributions of this work are trifold.

First, the concept of a 'Conflict Risk Graph,' has been proposed to map intersection conflict-prone locations to nodes within a network characterized by specific topological structures. The characteristics defining these nodes, termed as conflict risks, demonstrate a dynamic for propagation throughout the network, revealing a discernible global–local spatio-temporal dependence. This introduced concept is not limited to irregular road structures that can affect traffic safety analyses; instead, it facilitates the modeling of the spatio-temporal evolution patterns of conflict risks at intersections.

Second, the Trans-GCN model has been developed to synergistically combine the Transformer's proficiency in capturing global temporal correlations with the GCN' ability to learn local spatial correlations. This combination enables the Trans-GCN to adeptly model the complex spatio-temporal evolution patterns of conflict risks at intersections. The evaluation in this paper against the five chosen state-of-the-art deep learning models — GRU, GCN, TGCN, STGCN, and STTN — well demonstrates the Trans-GCN's superior performance in risk inference and adaptability to node changes.

Finally, the extensive experiments in this paper reveal a relationship between inference performance and node configurations: Increasing the spatio-temporal resolution of the modeling comes at the cost of reduced inference performance. An optimal node configuration that balances inference performance with spatio-temporal resolution has been identified. This configuration enables the Trans-GCN to effectively model the spatio-temporal dynamics of conflict risks and generate ideal conflict risk maps.

CRediT authorship contribution statement

Tao Wang: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation. **Ying-En Ge:** Formal analysis, Conceptualization, Writing – original draft. **Yongjie Wang:** Investigation, Funding acquisition, Conceptualization, Methodology, Writing – original draft. **Carlo G. Prato:** Writing – original draft, Writing – review & editing. **Wenqiang Chen:** Software, Validation, Writing – review & editing. **Yuchen Niu:** Investigation, Visualization, Writing – review & editing.

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

This work was supported by the National Natural Science Foundation of China (Grant :71801020, 72361137003), National Key Research and Development Plan (Grant:2021YFE0203600), Natural Science Foundation of Shaanxi Province (Grant: 2023-JC-QN-0795), Research Funds for the Interdisciplinary Projects CHU (Grant: 300104240924), the Fundamental Research Funds for the Central Universities, CHD (Grant: 300102344301).

Appendix

Table A
Comparison of inference performance across different temporal granularity.

Method	Location	Metrics											
		6 s*(7.0 m*7.0 m)			8 s *(7.0 m*7.0 m)			10 s*(7.0 m*7.0 m)			12 s *(7.0 m*7.0 m)		
		t_0(MAE /RMSE)	t_1(MAE /RMSE)	t_2(MAE /RMSE)	t_0(MAE /RMSE)	t_1(MAE /RMSE)	t_2(MAE /RMSE)	t_0(MAE /RMSE)	t_1(MAE /RMSE)	t_2(MAE /RMSE)	t_0(MAE /RMSE)	t_1(MAE /RMSE)	t_2(MAE /RMSE)
GCN	JX	0.190/ 0.315	0.206/ 0.336	0.259/ 0.370	0.174/ 0.282	0.193/ 0.307	0.228/ 0.355	0.160/ 0.252	0.171/ 0.266	0.202/ 0.301	0.134/ 0.137	0.176/ 0.259	0.202/ 0.289
	QS	0.095/ 0.078	0.072/ 0.080	0.098/ 0.088	0.094/ 0.081	0.074/ 0.085	0.108/ 0.089	0.086/ 0.081	0.088/ 0.077	0.109/ 0.098	0.095/ 0.116	0.090/ 0.078	0.110/ 0.130
GRU	JX	0.148/ 0.212	0.171/ 0.253	0.253/ 0.350	0.134/ 0.156	0.161/ 0.250	0.206/ 0.317	0.131/ 0.150	0.157/ 0.227	0.200/ 0.288	0.124/ 0.134	0.142/ 0.197	0.207/ 0.279
	QS	0.063/ 0.036	0.043/ 0.060	0.075/ 0.090	0.062/ 0.043	0.057/ 0.072	0.091/ 0.105	0.052/ 0.060	0.077/ 0.088	0.108/ 0.100	0.055/ 0.042	0.084/ 0.093	0.106/ 0.117
TGCN	JX	0.132/ 0.211	0.179/ 0.282	0.228/ 0.344	0.133/ 0.201	0.167/ 0.251	0.195/ 0.301	0.117/ 0.176	0.158/ 0.177	0.187/ 0.280	0.110/ 0.163	0.163/ 0.242	0.171/ 0.219
	QS	0.053/ 0.075	0.054/ 0.080	0.069/ 0.092	0.055/ 0.078	0.056/ 0.081	0.095/ 0.157	0.053/ 0.072	0.081/ 0.096	0.094/ 0.128	0.069/ 0.085	0.084/ 0.100	0.098/ 0.121
STGCN	JX	0.091/ 0.159	0.126/ 0.222	0.160/ 0.285	0.089/ 0.152	0.124/ 0.205	0.164/ 0.273	0.099/ 0.146	0.135/ 0.208	0.156/ 0.246	0.078/ 0.120	0.112/ 0.172	0.134/ 0.210
	QS	0.037/ 0.047	0.034/ 0.049	0.037/ 0.055	0.023/ 0.033	0.033/ 0.047	0.048/ 0.063	0.033/ 0.042	0.046/ 0.059	0.057/ 0.072	0.029/ 0.039	0.050/ 0.064	0.071/ 0.086
STTN	JX	0.083/ 0.154	0.138/ 0.248	0.185/ 0.33	0.098/ 0.175	0.152/ 0.275	0.182/ 0.339	0.084/ 0.148	0.143/ 0.239	0.169/ 0.291	0.078/ 0.131	0.139/ 0.047	0.158/ 0.273
	QS	0.022/ 0.035	0.030/ 0.057	0.042/ 0.069	0.026/ 0.037	0.036/ 0.056	0.058/ 0.076	0.035/ 0.047	0.039/ 0.062	0.049/ 0.070	0.030/ 0.040	0.056/ 0.061	0.086/ 0.099
Trans-GCN	JX	0.082/ 0.151	0.124/ 0.226	0.155/ 0.272	0.075/ 0.128	0.114/ 0.190	0.140/ 0.234	0.074/ 0.123	0.114/ 0.185	0.145/ 0.227	0.067/ 0.105	0.104/ 0.157	0.130/ 0.199
	QS	0.022/ 0.030	0.024/ 0.038	0.033/ 0.052	0.028/ 0.036	0.033/ 0.046	0.045/ 0.061	0.022/ 0.032	0.041/ 0.054	0.066/ 0.084	0.025/ 0.033	0.041/ 0.054	0.056/ 0.073

Table B
Comparison of inference performance across different spatial granularity.

Method	Location	Metrics											
		10s*(3.5 m*3.5 m)			10s*(7.0 m*7.0 m)			10s*(10.5 m*10.5 m)			10s*(14.0 m*14.0 m)		
		t_0(MAE /RMSE)	t_1 (MAE /RMSE)	t_2 (MAE /RMSE)	t_0 (MAE /RMSE)	t_1 (MAE /RMSE)	t_2 (MAE /RMSE)	t_0 (MAE /RMSE)	t_1 (MAE /RMSE)	t_2 (MAE /RMSE)	t_0 (MAE /RMSE)	t_1 (MAE /RMSE)	t_2 (MAE /RMSE)
GCN	JX	0.090/ 0.155	0.102/ 0.166	0.114/ 0.180	0.160/ 0.252	0.171/ 0.266	0.202/ 0.301	0.200/ 0.297	0.224/ 0.335	0.290/ 0.439	0.245/ 0.383	0.339/ 0.534	0.344/ 0.532
	QS	0.075/ 0.063	0.094/ 0.074	0.128/ 0.095	0.086/ 0.081	0.104/ 0.107	0.149/ 0.178	0.099/ 0.133	0.133/ 0.161	0.171/ 0.182	0.105/ 0.128	0.190/ 0.220	0.231/ 0.260
GRU	JX	0.058/ 0.087	0.081/ 0.120	0.104/ 0.155	0.101/ 0.150	0.157/ 0.227	0.200/ 0.288	0.132/ 0.194	0.186/ 0.277	0.253/ 0.367	0.172/ 0.256	0.261/ 0.391	0.362/ 0.527
	QS	0.036/ 0.045	0.043/ 0.051	0.061/ 0.066	0.053/ 0.072	0.081/ 0.096	0.114/ 0.128	0.079/ 0.101	0.127/ 0.150	0.138/ 0.163	0.096/ 0.119	0.176/ 0.207	0.231/ 0.265
T-GCN	JX	0.083/ 0.132	0.099/ 0.153	0.108/ 0.170	0.117/ 0.176	0.118/ 0.177	0.187/ 0.280	0.136/ 0.198	0.190/ 0.277	0.232/ 0.344	0.176/ 0.251	0.245/ 0.372	0.338/ 0.497
	QS	0.022/ 0.026	0.041/ 0.046	0.065/ 0.070	0.052/ 0.060	0.077/ 0.088	0.088/ 0.100	0.060/ 0.074	0.094/ 0.114	0.107/ 0.128	0.076/ 0.092	0.118/ 0.141	0.212/ 0.240
ST-GCN	JX	0.051/ 0.084	0.068/ 0.111	0.086/ 0.140	0.099/ 0.146	0.135/ 0.208	0.157/ 0.246	0.118/ 0.182	0.171/ 0.267	0.220/ 0.344	0.156/ 0.249	0.228/ 0.367	0.275/ 0.455
	QS	0.013/ 0.017	0.017/ 0.023	0.019/ 0.027	0.033/ 0.042	0.046/ 0.059	0.057/ 0.072	0.053/ 0.070	0.062/ 0.085	0.075/ 0.103	0.067/ 0.089	0.130/ 0.165	0.169/ 0.215
STTN	JX	0.061/ 0.109	0.087/ 0.162	0.103/ 0.192	0.084/ 0.148	0.143/ 0.239	0.169/ 0.291	0.115/ 0.177	0.183/ 0.290	0.257/ 0.376	0.153/ 0.246	0.241/ 0.399	0.335/ 0.521
	QS	0.028/ 0.037	0.021/ 0.038	0.038/ 0.049	0.035/ 0.047	0.039/ 0.062	0.049/ 0.070	0.042/ 0.062	0.060/ 0.087	0.084/ 0.114	0.084/ 0.114,	0.146/ 0.179	0.179/ 0.214
Trans-GCN	JX	0.044/ 0.078	0.062/ 0.110	0.079/ 0.134	0.074/ 0.123	0.114/ 0.185	0.145/ 0.227	0.106/ 0.169	0.162/ 0.253	0.206/ 0.320	0.140/ 0.237	0.218/ 0.360	0.285/ 0.465
	QS	0.010/ 0.016	0.018/ 0.023	0.019/ 0.026	0.022/ 0.032	0.041/ 0.054	0.066/ 0.084	0.044/ 0.060	0.060/ 0.084	0.089/ 0.116	0.064/ 0.088	0.122/ 0.164	0.163/ 0.209

References

Arun, A., Haque, M.M., Washington, S., Sayed, T., Mannering, F., 2022. How many are enough?: Investigating the effectiveness of multiple conflict indicators for crash frequency-by-severity estimation by automated traffic conflict analysis. Transp. Res. Part C: Emerging Technol. 138, 103653.

Bao, J., Liu, P., Ukkusuri, S.V., 2019. A spatiotemporal deep learning approach for citywide short-term crash risk prediction with multi-source data. Accid. Anal. Prev. 122, 239–254.

Basso, F., Basso, L.J., Bravo, F., Pezoa, R., 2018. Real-time crash prediction in an urban expressway using disaggregated data. Transp. Res. Part C: Emerging Technol. 86, 202–219.

Beitel, D., Stipancic, J., Manaugh, K., Miranda-Moreno, L., 2018. Assessing safety of shared space using cyclist-pedestrian interactions and automated video conflict analysis. Transp. Res. Part D: Transp. Environ. 65, 710–724.

Buddhavarapu, P., Scott, J.G., Prozzi, J.A., 2016. Modeling unobserved heterogeneity using finite mixture random parameters for spatially correlated discrete count data. Transp. Res. Part B: Methodol. 91, 492–510.

Chen, A.Y., Chiu, Y.-L., Hsieh, M.-H., Lin, P.-W., Angah, O., 2020. Conflict analytics through the vehicle safety space in mixed traffic flows using UAV image sequences. Transp. Res. Part C: Emerging Technol. 119, 102744.

Chen, W., Wang, T., Wang, Y., Li, Q., Xu, Y., Niu, Y., 2022. Lane-based Distance-Velocity model for evaluating pedestrian-vehicle interaction at non-signalized locations. Accid. Anal. Prev. 176, 106810.

Chung, J., Gulcehre, C., Cho, K., Bengio, Y., December 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. In NIPS 2014 Workshop on Deep Learning.

Devlin, J., Chang, M.-W., Lee, K. Toutanova, K. "BERT: Pretraining of deep bidirectional transformers for language understanding," in Proceedings 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT) Minneapolis, MN, USA, June 2018, pp. 4171–4186.".

Essa, M., Sayed, T., 2018. Traffic conflict models to evaluate the safety of signalized intersections at the cycle level. Transp. Res. Part C: Emerging Technol. 89, 289–302.

Ferreira, S., Couto, A., 2015. A probabilistic approach towards a crash risk assessment of urban segments. Transp. Res. Part C: Emerging Technol. 50, 97–105.

Fu, C., Sayed, T., 2023. Identification of adequate sample size for conflict-based crash risk evaluation: An investigation using Bayesian hierarchical extreme value theory models. Analytic Methods in Accident Research 39, 100281.

Gemperline, P.J., Long, J.R., Gregoriou, V.G., 1991. Nonlinear multivariate calibration using principal components regression and artificial neural networks. Anal. Chem. 63 (20), 2313–2323.

Ghoul, T., Sayed, T., Fu, C., 2023. Dynamic identification of short-term and longer-term hazardous locations using a conflict-based real-time extreme value safety model. Analytic Methods in Accident Research 37, 100262.

Guo, Y., Essa, M., Sayed, T., Haque, M.M., Washington, S., 2019. A comparison between simulated and field-measured conflicts for safety assessment of signalized intersections in Australia. Transp. Res. Part C: Emerging Technol. 101, 96–110.

Hossain, M., Abdel-Aty, M., Quddus, M.A., Muromachi, Y., Sadeek, S.N., 2019. Real-time crash prediction models: State-of-the-art, design pathways and ubiquitous requirements. Accident Analysis and Prevention 124, 66–84.

Hou, X., Wang, Y., Chau, L. P., "Vehicle Tracking Using Deep SORT with Low Confidence Track Filtering," in Proceedings of the 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), September 18-21, 2019, Taipei, Taiwan, China, pp. 1–6.

Howlader, M.M., Bhaskar, A., Yasmin, S., Haque, M.M., 2024. A bivariate, non-stationary extreme value model for estimating opposing-through crash frequency by severity by applying artificial intelligence-based video analytics. Transp. Res. Part C: Emerging Technol. 160, 104509.

Hu, Y., Zhang, Y., Shelton, K.S., 2018. Where are the dangerous intersections for pedestrians and cyclists: A colocation-based approach. Transp. Res. Part C: Emerging Technol. 95, 431–441.

Hu, Y., Li, Y., Huang, H., 2023. Spatio-temporal dynamic change mechanism analysis of traffic conflict risk based on trajectory data. Accid. Anal. Prev. 191, 107203.

Huang, Q., Jia, H., Yuan, Z., Wu, R., 2023. PL-TARMI: A deep learning framework for pixel-level traffic crash risk map inference. Accid. Anal. Prev. 191, 107174.

Hussain, F., Ali, Y., Li, Y., Haque, M.M., 2023. Real-time crash risk forecasting using Artificial-Intelligence based video analytics: A unified framework of generalised extreme value theory and autoregressive integrated moving average model. Analytic Methods in Accident Research 40, 100302.

Kashifi, M.T., Al-Sghan, I.Y., Rahman, S.M., Al-Ahmadi, H.M., 2022. Spatiotemporal grid-based crash prediction—application of a transparent deep hybrid modeling framework. Neural Comput. & Applic. 34 (23), 20655–20669.

Kipf, T.N., Welling, M., 2017. Semi-Supervised Classification with Graph Convolutional Networks. In Proceedings of the International Conference on Learning Representations (ICLR).

Li, P., Abdel-Aty, M., Yuan, J., 2020. Real-time crash risk prediction on arterials based on LSTM-CNN. Accid. Anal. Prev. 135, 105371.

Li, S., Jin, X., Xuan, Y., Zhou, X., Chen, W., Wang, Y.-X., Yan, X., "Enhancing the Locality and Breaking the Memory Bottleneck of Transformer on Time Series Forecasting," in Proceedings of the 33rd International Conference on Neural Information Processing Systems (NeurIPS), Vancouver, Canada, December 2019, Article No. 471, pp. 5243–5253.

Li, H., Zhang, S., Li, X., Su, L., Huang, H., Jin, D., Chen, L., Huang, J., Yoo, J., 2021. Detectornet: Transformer-enhanced spatial temporal graph neural network for traffic prediction. In: 29th International Conference on Advances in Geographic Information Systems (SIGSPATIAL), November 2–5, 2021, Beijing. China. ACM, New York, NY, USA, pp. 133–136.

Liu, Q., Li, J., Lu, Z., 2021. ST-Tran: Spatial-Temporal Transformer for Cellular Traffic Prediction. IEEE Communications Letters 25(10), 3325-3329.

Liu, J., Kang, Y., Li, H., Wang, H., Yang, X., 2023b. STGHTN: Spatial-temporal gated hybrid transformer network for traffic flow forecasting. Appl. Intell. 53 (10), 12472–12488.

Liu, X., Lu, J., Chen, X., Fong, Y.H.C., Ma, X., Zhang, F., 2023a. Attention based spatio-temporal graph convolutional network with focal loss for crash risk evaluation on urban road traffic network based on multi-source risks. Accid. Anal. Prev. 192, 107262.

Mo, W., Lee, J., Abdel-Aty, M., Mao, S., Jiang, Q., 2024. Dynamic short-term crash analysis and prediction at toll plazas for proactive safety management. Accid. Anal. Prev. 197, 107456.

Ouni, F., Belloumi, M., 2018. Spatio-temporal pattern of vulnerable road user's collisions hot spots and related risk factors for injury severity in Tunisia. Transport. Res. F: Traffic Psychol. Behav. 56, 477–495.

Peng, Y., Abdel-Aty, M., Shi, Q., Yu, R., 2017. Assessing the impact of reduced visibility on traffic crash risk using microscopic data and surrogate safety measures. Transp. Res. Part C: Emerging Technol. 74, 295–305.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., "You Only Look Once: Unified, Real-Time Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 26 - July 1, 2016, Las Vegas, NV, USA, pp. 779–788.

Sacchi, E., Sayed, T., DeLeur, P., 2013. A comparison of collision-based and conflict-based safety evaluations: the case of right-turn smart channels. Accid. Anal. Prev. 59, 260–266.

Song, L., Li, Y., Fan, W., Wu, P., 2020. Modeling pedestrian-injury severities in pedestrian-vehicle crashes considering spatiotemporal patterns: Insights from different hierarchical Bayesian random-effects models. Analytic Methods in Accident Research 28, 100137.

St-Aubin, P., Saunier, N., Miranda-Moreno, L., 2015. Large-scale automated proactive road safety analysis using video data. Transp. Res. Part C: Emerging Technol. 58, 363–379.

Stevanovic, A., Stevanovic, J., Kergaye, C., 2013. Optimization of traffic signal timings based on surrogate measures of safety. Transp. Res. Part C: Emerging Technol. 32, 159–178.

Sun, J., Sun, J., 2015. A Dynamic Bayesian Network model for real-time crash prediction using traffic speed conditions data. Transp. Res. Part C: Emerging Technol. 54, 176–186.

Tobler, W.R., 1970. A computer movie simulating urban growth in the detroit region. Econ. Geogr 46, 234–240.

Tran, T., He, D., Kim, J., Hickman, M., 2023. MSGNN: A Multi-structured Graph Neural Network model for real-time incident prediction in large traffic networks. Transp. Res. Part C: Emerging Technol. 156, 104354.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. In Advances in neural information processing systems, pp. 5998–6008, 2017.
Wang, L., Abdel-Aty, M., Shi, Q., Park, J., 2015. Real-time crash prediction for expressway weaving segments. Transp. Res. Part C: Emerging Technol. 61, 1–10.
Wang, S., Chen, A., Wang, P., Zhuge, C., 2023a. Predicting electric vehicle charging demand using a heterogeneous spatio-temporal graph convolutional network. Transp. Res. Part C: Emerging Technol. 153, 104205.
Wang, Y., Niu, Y., Zhu, W., Chen, W., Li, Q., Wang, T., 2023b. Predicting Pedestrian Crossing Behavior at Unsignalized Mid-Block Crosswalks Using Maximum Entropy Deep Inverse Reinforcement Learning. IEEE Trans. Intell. Transp. Syst. 1–14.
Wang, T., Ge, Y.-E., Wang, Y., Chen, W., 2024a. A spatio-temporal deep learning approach to simulating conflict risk propagation on freeways with trajectory data. Accid. Anal. Prev. 195, 107377.
Wang, Y., Jia, Y., Chen, W., Wang, T., Zhang, A., 2024b. Examining safe spaces for pedestrians and e-bicyclists at urban crosswalks: An analysis based on drone-captured video. Accid. Anal. Prev. 194, 107365.
Wang, T., Ge, Y.-E., Wang, Y., Chen, W., 2024c. A novel model for real-time risk evaluation of vehicle-pedestrian interactions at intersections. Accid. Anal. Prev. 206, 107727.
Wang, J., Xie, W., Liu, B., Fang, S., Ragland, D.R., 2016. Identification of freeway secondary accidents with traffic shock wave detected by loop detectors. Saf. Sci. 87, 195–201.
Wu, J., Abdel-Aty, M., Yu, R., Gao, Z., 2013. A novel visible network approach for freeway crash analysis. Transp. Res. Part C: Emerging Technol. 36, 72–82.
Wu, Y., Abdel-Aty, M., Zheng, O., Cai, Q., Zhang, S., 2020. Automated safety diagnosis based on unmanned aerial vehicle video and deep learning algorithm. Transp. Res. Rec.: J. Transp. Res. Board 2674 (8), 350–359.
Wu, P., Chen, T., Diew, Y., Wong, M., X., Wang, X., Liu, W., 2023. Exploring key spatio-temporal features of crash risk hot spots on urban road network: A machine learning approach. Transp. Res. A Policy Pract. 173, 103717.
Xie, K., Ozbay, K., Yang, H., Li., C., 2019. Mining automatically extracted vehicle trajectory data for proactive safety analytics. Transp. Res. Part C: Emerging Technol. 106 (2019), 61–72.
Xu, M., Dai, W., Liu, C., Gao, X., Lin, W., Qi, G.-J., Xiong, H., 2020. "Spatial-temporal transformer networks for traffic flow forecasting." arXiv preprint arXiv: 2001.02908.
Xu, C., Wang, W., Liu, P., Guo, R., Li, Z., 2014. Using the Bayesian updating approach to improve the spatial and temporal transferability of real-time crash risk prediction models. Transp. Res. Part C: Emerging Technol. 38, 167–176.
Xu, C., Liu, P., Yang, B., Wang, W., 2016. Real-time estimation of secondary crash likelihood on freeways using high-resolution loop detector data. Transp. Res. Part C: Emerging Technol. 71, 406–418.
Yan, H., Ma, X., Pu, Z., 2022. Learning Dynamic and Hierarchical Traffic Spatiotemporal Features With Transformer. IEEE Trans. Intell. Transp. Syst. 23 (11), 22386–22399.
Yang, L., Luo, X., Zuo, Z., Zhou, S., Huang, T., Luo, S., 2023. A novel approach for fine-grained traffic risk characterization and evaluation of urban road intersections. Accid. Anal. Prev. 181, 106934.
Yang, D., Ozbay, K., Xie, K., Yang, H., Zuo, F., Sha, D., 2021. Proactive safety monitoring: A functional approach to detect safety-related anomalies using unmanned aerial vehicle video data. Transp. Res. Part C: Emerging Technol. 127, 103130.
Yang, K., Wang, X., Yu, R., 2018. A Bayesian dynamic updating approach for urban expressway real-time crash risk evaluation. Transp. Res. Part C: Emerging Technol. 96, 192–207.
Yu, B., Yin, H., Zhu, Z., "Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting," in Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI), July 2018, pp. 3634–3640.
Yu, R., Wang, Y., Zou, Z., Wang, L., 2020. Convolutional neural networks with refined loss functions for the real-time crash risk analysis. Transp. Res. Part C: Emerging Technol. 119, 102740.
Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., Li, H., 2020. T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. IEEE Trans. Intell. Transp. Syst. 21 (9), 3848–3858.
Zheng, L., Sayed, T., 2020. A novel approach for real time crash prediction at signalized intersections. Transp. Res. Part C: Emerging Technol. 117, 102683.