

面向高速公路事故风险预测的深度学习方法

阮鸿柱¹, 黄小弟¹, 王金宝¹, 杜梦辉²

(1. 云南省综合交通发展中心, 云南 昆明 650031;

2. 北京交通大学 计算机与信息技术学院, 北京 100044)

摘要:高速公路的交通事故风险预测对智能交通和公共安全具有重要意义。现有方法通过挖掘历史事故的时空特征预测交通事故风险。但是,在高速公路事故风险预测中仍存在以下两个挑战。首先,事故具有不均衡的空间分布,相邻路段的事故分布差异可能较大,而相隔较远却具有相似拓扑连接关系路段的事故分布可能较相似。另外,由于事故的偶发性,其在时间维的分布非常稀疏,因此在捕获事故影响因素时缺乏足够的样本。针对第一个挑战,使用自适应图卷积网络以数据驱动的方式学习路段间的空间相关性;此外,根据 Mixup 策略进行数据增广以生成足够多的事故风险样本解决事故数量稀疏的问题,然后用对比学习方法以更好地区分风险与非风险样本,以实现更准确的事故风险预测。基于桂林市高速公路网真实交通数据集的实验结果表明,相比于最优方法,该方法的平均绝对误差指标降低了 18.3%,平均准确率、召回率指标分别提升了 8.1%、6.9%,因此,该方法可以更准确地预测高速公路事故风险。

关键词:智能交通;交通事故风险预测;对比学习;自适应图神经网络;数据增广

中图分类号:TP391.4

文献标识码:A

文章编号:1673-629X(2023)11-0189-07

doi:10.3969/j.issn.1673-629X.2023.11.028

A Deep Learning Approach for Highway Accident Risk Prediction

RUAN Hong-zhu¹, HUANG Xiao-di¹, WANG Jin-bao¹, DU Meng-hui²

(1. Yunnan Transportation Development Center, Kunming 650031, China;

2. School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China)

Abstract: Highway traffic accident risk prediction is vital to intelligent transportation and public safety. Existing approaches predict traffic accident risk by mining the spatio-temporal characteristics of historical accidents. However, there are still two challenges in highway accident risk prediction. Firstly, the accidents have uneven spatial distribution. The difference in accident distribution between adjacent roads may be large, while the accident distribution of distant roads with similar topological connection relationships may be similar. In addition, due to the contingency of the accident, its distribution in the time dimension is quite sparse, so there is not enough sample to capture the influence factors of traffic accidents. For the first challenge, we use an adaptive graph convolutional network to learn the spatial correlation between roads in a data-driven way. In addition, we adopt data augmentation based on a Mixup strategy to generate enough accident risk samples to solve the problem of data sparsity and then use contrastive learning to better distinguish risk and non-risk samples so as to achieve more accurate accident risk prediction. The experimental results based on the real traffic dataset of the Guilin expressway network show that compared with other optimal methods, the mean absolute error of the proposed method is reduced by 18.3%, and the average accuracy and the recall are increased by 8.1% and 6.9%, respectively. Therefore, the proposed method can predict highway accident risk more accurately.

Key words: intelligent transportation; traffic accident risk prediction; contrastive learning; adaptive graph convolutional network; data augmentation

0 引言

交通事故对公共安全造成了严重威胁。如果能够提前准确预测高速公路的交通事故风险,管理部门可以更好地进行交通规划,减少交通事故,并发布交通事

故风险警告,驾驶员可以选择更安全的路线,避免交通危险。

交通事故风险预测是指在观测到的历史交通事故基础上,预测目标位置在未来时间可能发生潜在交通

收稿日期:2023-01-13

修回日期:2023-05-17

基金项目:中央高校基本科研业务费(2019JBM023)

作者简介:阮鸿柱(1993-),男,高级工程师,硕士研究生,研究方向为智慧高速和交通信息工程;通讯作者:杜梦辉(1999-),男,硕士研究生,研究方向为智慧交通。

事故的风险值^[1]。除历史事故风险值之外,还可以分析导致事故发生的外部因素,如天气、时间、车流量等。早期的一些研究使用统计方法或线性机器学习模型,例如支持向量机^[2],差分整合移动平均自回归模型^[3]挖掘交通事故的历史发生模式。但是,这些方法难以捕获到交通事故数据中复杂且非线性的时空相关性。因此,一些研究基于深度学习方法,利用非线性模型提升交通事故风险预测的准确性^[4-9]。

尽管现有的很多深度学习方法已被有效地应用于交通事故风险预测,但是仍存在以下两个挑战,从而影响了现有方法的预测精度。

第一个挑战是交通事故具有不均衡的空间分布^[10],比如在交通量较大或者路段交汇情况较复杂的路段,交通事故数量可能较多,而一些路况较好的路段上交通事故数量可能较少。图卷积网络已被广泛应用于捕获路网中交通数据的空间相关性,但是传统的图卷积网络需要预定义邻接关系图以学习节点间的空间相关性。由于有限的先验知识,预定义图难以全面地反映空间相关性。一方面,由于不同的路段结构特征,相邻路段的事故分布差异可能较大;另一方面,相隔较远的路段,由于具有相似的拓扑连接关系,可能具有相似的事故分布。

第二个挑战在于交通事故在时间维的分布非常稀疏^[11]。由于事故的偶发性,一些路段在大部分时间的事故风险值为零。由于缺乏事故风险样本,训练好深度神经网络具有较大的挑战。此外,由于非风险样本的数量远大于风险样本的数量,交通事故预测将偏向于输出零事故风险的结果,而不是发现潜在的事故风险。交通事故数量在时间维稀疏的问题,并没能被之前的工作很好地解决。

针对以上两个挑战,该文提出一种基于对比学习的自适应时空图卷积网络(Contrastive Learning Based Adaptive Spatio-Temporal Graph Convolutional Network, CASTGCN)用于交通事故风险预测。对于第一个挑战,基于矩阵分解的思想,该文使用自适应图卷积网络^[12]以数据驱动的方式学习节点间的空间相关性,解决交通事故空间分布不均衡的问题。对于第二个挑战,该文设计了一种对比学习方法,该方法基于 Mixup 策略^[13]生成足够多的事故风险样本。设计的 Mixup 策略利用事故数据中固有的周期模式,通过聚合两个相邻周的具有相同小时片的样本来实现数据增强,如果任一样本是风险样本,生成的样本将被标记为具有混合得分的风险样本,否则被标记为非风险样本。经过数据增强,非风险样本和风险样本的比例差将显著减小。然后利用风险和非风险样本构建正例和负例进行有监督的对比学习。通过对比学习使得模型更好

地区分风险样本和非风险样本的嵌入表示,从而提升事故预测的准确度。

贡献总结为:首先,基于自适应图卷积网络以数据驱动的方式挖掘交通数据中复杂的空间相关性;其次,基于交通事故数据的周期性设计了一种 Mixup 策略并将其整合到对比学习方法中,解决事故数据在时间维稀疏的问题,使得模型可以更好地地区分风险样本和非风险样本;最后,基于桂林市高速公路数据,将提出方法与多种预测方法进行对比,验证了提出方法的有效性和先进性。

1 前置定义

1.1 高速公路路网

将高速公路路网定义为一个无向图 $G = (L, E)$, 其中 $L = \{l_0, l_1, \dots, l_M\}$ 是 $|L| = M$ 个路段的集合, E 是 $|E| = Q$ 条边的集合。如果车辆可以从路段 l_i 通向路段 l_j , 则路段 l_i 和路段 l_j 存在连接关系, (l_i, l_j) 表示连接路段 l_i 和路段 l_j 的边。

1.2 交通事故风险

首先,根据交通事故的具体情况将其分为 3 种类型:轻型事故、中型事故、严重事故。对于第 t 个时间间隔中的一个路段,给定第 k 个类型的交通事故数量 n_k ,该路段的风险指数(risk)定义为不同类别交通事故的加权求和:

$$\text{risk} = \sum_{k=1}^3 w_k \times n_k \quad (1)$$

其中, w_k 表示不同类型交通事故的权重: $w_1 = 1$ 表示轻型事故, $w_2 = 2$ 表示中型事故, $w_3 = 3$ 表示严重事故。

1.3 图信号矩阵

在第 t 个时间间隔内,观测到高速公路路网 G 上的图信号用矩阵 $\mathbf{X}_t = (\mathbf{x}_{t,1}, \mathbf{x}_{t,2}, \dots, \mathbf{x}_{t,M})^T \in \mathbb{R}^{M \times C}$ 表示,其中 $\mathbf{x}_{t,i} \in \mathbb{R}^C$ 表示第 t 个时间间隔内路段 l_j 的 C 个观测量的取值。

1.4 交通事故风险预测

在交通事故风险预测任务中,给定历史 T 个时间间隔的时空图信号矩阵 $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_T)$, 预测所有路段在第 $T+1$ 个时间间隔的交通风险指数:

$$\mathbf{Y}_{T+1} = \mathcal{F}_{\theta}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_T) \quad (2)$$

其中, θ 为模型参数, $\mathbf{Y}_{T+1} \in \mathbb{R}^M$ 为预测得到的所有路段的事故风险指数值。

2 方法介绍

该文提出一种基于对比学习的自适应时空图卷积网络(CASTGCN)用于交通事故风险预测,如图 1 所示。输入历史多个时间间隔的时空图信号矩阵,通过

自适应空间相关性建模、自适应时间相关性建模、卷积层捕获时空图信号矩阵中的时空相关性,然后通过预测层得到各个路段在未来的交通风险指数,采用均方误差损失函数指导模型训练。另一方面,通过 Mixup 数据增广方法生成事故风险样本,通过自适应时空相关性建模得到风险样本和非风险样本的嵌入表示,利用增广样本的嵌入表示进行有监督的对比学习,使模型更好地区分风险样本和非风险样本。

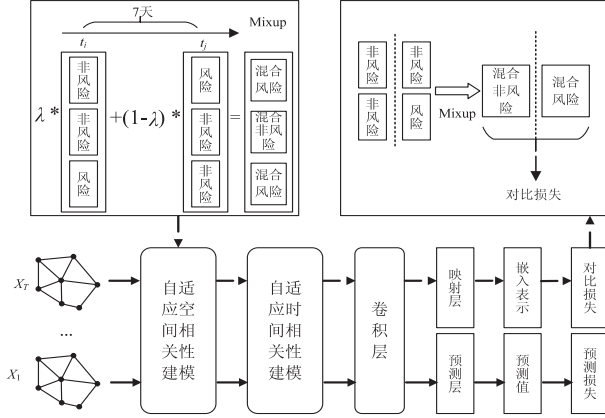


图 1 基于对比学习的自适应图卷积网络

2.1 输入层

经过数据预处理得到路网 G 中所有节点在各个时间步的特征作为输入。具体输入特征构造为:在第 t 个时间间隔内,观测到高速公路路网 G 上的信号用矩阵 $X_t = (x_{t,1}, x_{t,2}, \dots, x_{t,N})^T \in \mathbb{R}^{N \times C}$ 表示,其中 $x_{t,v} = (x_{t,v}^{\text{risk}}, x_{t,v}^{\text{event}}, x_{t,v}^{\text{flow}}, x_{t,v}^{\text{prec}}, x_{t,v}^{\text{vis}}, x_{t,v}^{\text{weekday}}, x_{t,v}^{\text{is}}) \in \mathbb{R}^{1 \times C}$ 表示第 t 个时间间隔内节点 v 的 C 个观测量的取值。其中 $x_{t,v}^{\text{risk}}$ 为该时间间隔内节点的风险指数, $x_{t,v}^{\text{event}}$ 为该时间段内节点处交通事件数量, $x_{t,v}^{\text{flow}}$ 为该时间间隔内节点流量值, $x_{t,v}^{\text{prec}}$ 为该时间段内节点处平均降水量, $x_{t,v}^{\text{vis}}$ 为该时间段内节点处可见度, $x_{t,v}^{\text{weekday}}$ 为该时间段对应的周中天数, $x_{t,v}^{\text{is}}$ 为该时间段对应的时片。

2.2 自适应空间相关性建模

图卷积网络因其具有学习图结构数据中的空间相关性的能力被广泛用于交通流预测。目前普遍使用的图卷积计算公式如下:

$$H = (I_N + D^{-\frac{1}{2}} A D^{-\frac{1}{2}}) X W + b \quad (3)$$

其中, $A \in \mathbb{R}^{N \times N}$ 为邻接矩阵, D 为度矩阵, $X \in \mathbb{R}^{N \times C}$ 为图信号输入, $H \in \mathbb{R}^{N \times F}$ 为图卷积计算的输出, $W \in \mathbb{R}^{C \times F}$ 和 $b \in \mathbb{R}^F$ 为可学习参数。图卷积通过一组共享的参数 W 和 b 聚合节点自身和邻接节点信息,实现空间特征提取。

传统图卷积计算通常使用预定义的邻接矩阵 A , 邻接矩阵 A 通常基于距离计算得到。这种方法的局限性在于节点特征在图上的传播将完全依赖于预定义的图结构,而预定义的图结构难以保证其能完全包含

路网交通数据中复杂的空间相关性。一方面,由于不同的路段结构特征,相邻路段的事故分布差异可能较大;另一方面,相隔较远的路段,由于具有相似的拓扑连接关系,可能具有相似的事故分布,如图 2 所示。

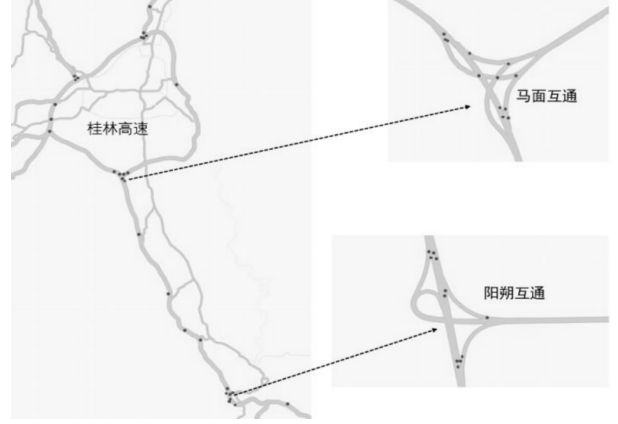


图 2 桂林高速公路事故(圆点)的空间分布
(2022 年 3 月 5 日 8 时至 2022 年 3 月 12 日 8 时)

针对预定义图的局限性,该文使用自适应图卷积以数据驱动的方式学习图节点的嵌入表示,计算自适应邻接矩阵用于图卷积运算。具体地,首先为每个图节点随机初始化一个节点嵌入表示矩阵 $E_v \in \mathbb{R}^{N \times d_e}$, 该矩阵中的每一行表示一个节点的嵌入表示, d_e 为嵌入表示维度。通过 E_v 与其转置 E_v^T 的乘积计算节点相关性:

$$D^{-\frac{1}{2}} A D^{-\frac{1}{2}} = \text{Softmax}(\text{ReLU}(E_v \cdot E_v^T)) \quad (4)$$

自适应邻接矩阵不需要任何先验知识,在训练过程中通过反向传播机制学习自适应邻接矩阵的参数,以数据驱动的方式学习路网交通数据中的空间相关性。

该文将式(4)中通过自适应学习方法计算得到的 $D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$ 应用到式(3),对各个时间片的图信号矩阵进行自适应图卷积运算,得到各个时间步空间图的嵌入表示 H_t :

$$H_t = (I + \text{softmax}(\text{ReLU}(E_v \cdot E_v^T))) X_t W + b \quad (5)$$

其中 X_t 为第 t 个时间片下的图信号矩阵。

2.3 自适应时间相关性建模

在时间维度上,不同时间段的路网信息之间存在显著的相关性,在不同情况下的时间相关性也不同。因此基于注意机制^[14]对历史不同时间步的交通数据赋予不同相关性。使用如下公式计算出时间注意力矩阵:

$$C = V \cdot \sigma((\mathcal{H})^T U_1) U_2 (U_3 \mathcal{H}) + e \quad (6)$$

$$C'_{ij} = \frac{\exp(C_{ij})}{\sum_{j=1}^T \exp(C_{ij})} \quad (7)$$

其中, $\mathbf{V}, \mathbf{e} \in \mathbb{R}^{T \times T}$, $\mathbf{U}_1 \in \mathbb{R}^N$, $\mathbf{U}_2 \in \mathbb{R}^{F \times M}$, $\mathbf{U}_3 \in \mathbb{R}^F$ 是可学习参数。时间相关性矩阵 \mathbf{C} 的值根据输入的图信号矩阵自适应学习得到, 其中元素 $C_{i,j}$ 表示时间片 i 和 j 下路网信息的相关性。最后, 使用 Softmax 函数对矩阵进行归一化。用计算得到时间注意力矩阵与空间相关性建模得到的嵌入表示相乘, 得到新的时空图表示:

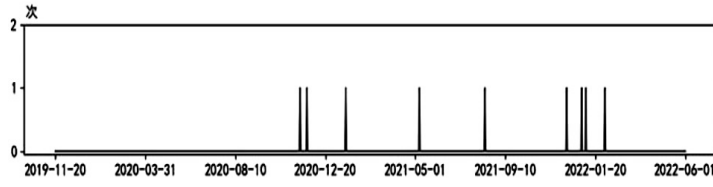
$$\mathcal{H}' = \mathcal{H}\mathbf{C}' = (\mathbf{H}_{S-T+1}, \mathbf{H}_{S-T+2}, \dots, \mathbf{H}_S) \quad (8)$$

此时, 对 $\mathbf{H}'_t, \forall t \in [0, T]$ 中任意元素 $\mathbf{H}'_{t,i,j}$, 已包含在空间、时间维度聚合后的信息。在上述基础上, 在时间维度上应用 R 个标准卷积层进一步聚合相邻时间片间的信息。对其中第 r 步卷积, 可表示为:

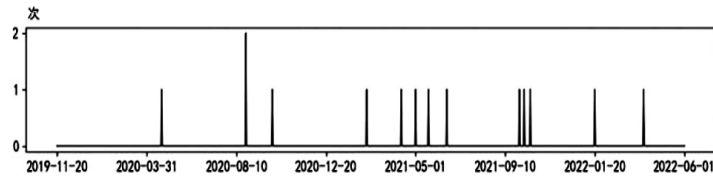
$$\mathcal{H}^{(r)} = \text{ReLU}(\Phi * \mathcal{H}^{(r-1)}) \in \mathbb{R}^{F_r \times M \times T_r} \quad (9)$$

其中, $*$ 表示标准卷积计算, Φ 为时间维卷积核参数, F_r 表示第 r 步卷积后图节点的特征维度, T_r 表示第 r 步卷积后时间维度大小。最终输出为:

$$\mathbf{Z} = \text{ReLU}(\Phi * \mathcal{H}^{(R)}) \in \mathbb{R}^{F_m \times M} \quad (10)$$



(a) 严关塘堡互通-溶江至溶江-灵川西路段



(b) 桂林东-尧山至尧山-灵川路段

图 3 桂林高速公路事故数量随时间的分布(2021 年 11 月 20 日 0 时至 2022 年 6 月 1 日 0 时)

为减轻因事故风险样本稀疏带来的影响, 该文基于 Mixup 数据增广方法生成足够多的事故风险样本。通过对交通事故的时空分析可知, 其发生时间具有以周为单位的周期性, 即不同周的相同周中天发生交通事故的概率相比于不同周中天发生交通事故的概率更接近。基于这种周期性, 通过构造两个连续周具有相同时间戳样本 x_i 和 x_j 的凸组合生成增广样本:

$$\mathbf{X}_{\text{mix}} = \{\lambda x_i + (1 - \lambda)x_j \mid t_i - t_j = 7 \text{ 天}\} \quad (11)$$

$$\mathbf{Y}_{\text{mix}} = \{\lambda y_i + (1 - \lambda)y_j \mid t_i - t_j = 7 \text{ 天}\} \quad (12)$$

其中, λ 为混合参数, 用于调节两个数据源的影响。如果 x_i 或 x_j 为事故风险样本, 则将生成的增广样本标记为事故风险样本。

2.5 对比学习方法

在实现基于 Mixup 的数据增广后, 通过对比学习方法利用增广后的数据解决事故风险数据在时间维度稀疏的问题。

在得到聚合前 T 个时间段时空相关性的路段嵌入表示 \mathbf{Z} 后, 通过一个多层感知机作为预测器, 得到所有路段在第 $T + 1$ 个时间段的交通风险指数 $\mathbf{Y}_{T+1} = \text{MLP}(\mathbf{Z})$ 。

2.4 基于 Mixup 的数据增广

传统的对比学习方法^[15]通过对样本进行数据增强得到正样本对和负样本对, 然后利用正、负样本对计算对比损失以实现表示学习。表示学习的目的是使正样本对在嵌入空间中的差异性更小, 负样本对在嵌入空间中的差异性更大, 从而使得模型更好地区分正样本对和负样本对。尽管传统的对比学习方法在很多领域广泛应用, 但是, 由于交通事故数据的稀疏性(如图 3 所示), 引入对比学习难以直接使模型更好地区分事故风险样本和非事故风险样本。相比于大量的非事故风险样本, 少量的事故风险样本使得模型难以学习到事故风险数据的分布, 从而难以调整模型参数实现事故风险的预测。

首先, 根据数据增广后的风险值对生成的事故风险/非风险样本进行分组, 得到正例集合和负例集合。具体的, 风险值大于 0 的样本构成正例集合, 风险值等于 0 的样本构成负例集合。

然后, 将正例集合和负例集合输入由自适应空间相关性建模和自适应时间相关性建模构成的编码器 $f(\cdot)$, 通过映射函数 $g(\cdot)$ 得到增广样本的嵌入表示 \mathbf{Z}_{pos} 和 \mathbf{Z}_{neg} :

$$\mathbf{Z}_{\text{pos}} = \{g(f(x_i)) \mid x_i \in \mathbf{X}_{\text{mix}}, y_i > 0\} \quad (13)$$

$$\mathbf{Z}_{\text{neg}} = \{g(f(x_i)) \mid x_i \in \mathbf{X}_{\text{mix}}, y_i = 0\} \quad (14)$$

最后, 利用增广样本的嵌入表示在潜在空间进行有监督的对比学习, 使得模型更好地区分事故风险样本和非事故风险样本。受益于基于 Mixup 的数据增广, 更多事故风险样本的特征可以被模型充分利用, 从而防止对比学习在非事故风险样本上发生过拟合, 通过学习到更好的事故风险嵌入表示实现更准确的交

通事故风险预测。

在交通事故风险预测中,对比损失的目的是使得事故风险样本的嵌入表示构成的正样本的相似度更高,非事故风险样本的嵌入表示构成的负样本对的相似度更低:

$$L_C = - \sum_{z_i \in z_{pos}} \log \frac{\exp(\text{sim}(z_i, z_j))}{\sum_{z_i \in z_{pos} \cup z_{neg}} \exp(\text{sim}(z_i, z_j))} \quad (15)$$

其中, $\text{sim}(\cdot)$ 表示余弦相似度。

2.6 损失函数

训练过程中,损失函数包括通过原始数据计算的事故风险预测损失 \mathcal{L}_F 和增广数据计算的对比损失 \mathcal{L}_C 。 \mathcal{L}_F 由预测值和真实值的均方误差计算得到:

$$\mathcal{L}_F = \frac{1}{M} \sum_{l=1}^M (Y[l] - \hat{Y}[l])^2 \quad (16)$$

其中, Y 和 \hat{Y} 分别表示预测时间间隔 $T+1$ 中交通事故风险的真实值和预测值, M 为路段数量。

总的损失函数 \mathcal{L} 为两项损失函数的组合,参数 α 控制对比学习的影响:

$$\mathcal{L} = \mathcal{L}_F + \alpha \mathcal{L}_C \quad (17)$$

2.7 模型训练

模型训练的损失函数使用均方误差函数,优化算法选用 Adam 算法^[16],Adam 算法是一种基于“动量”思想的随机梯度下降算法,在迭代前计算梯度的二阶动量并计算滑动平均值,并用于更新参数。

3 实验与结果

实验使用桂林市路网区域的交通数据,共有 36 个路段,从 2019 年 11 月 20 日到 2022 年 6 月 1 日。按 6:2:2 的比例划分训练集、验证集和测试集,使用训练集的均值和标准差对所有数据进行归一化,输入 $T=10$ 个时间步的数据,预测未来 1 个时间步的交通事故风险指数。

3.1 评价指标

从回归和排序两方面验证提出的模型的预测性能。该文用平均绝对误差 (Mean Absolute Error, MAE) 从回归的角度评价预测准确率。通过 Recall 和 MAP 以排序的角度评测事故风险区域的预测准确度。在时间间隔 $t (1 \leq t \leq T)$,如果在 k_t 个路段存在事故风险,Recall 表示真实具有事故风险的 k_t 个路段与预测得到事故风险值最高的 k_t 个路段的重合率。MAP 表示预测路段的平均准确率。MAE 越低,表示模型预测的交通事故风险指数越准确,Recall 和 MAP 越高则表示模型识别事故风险的能力越强。MAE, Recall, MAP 的计算公式为:

$$\text{MAE} = |Y_t - \hat{Y}_t| \quad (18)$$

$$\text{Recall} = \frac{1}{T} \sum_{t=1}^T \frac{|S_t \cap R_t|}{|R_t|} \quad (19)$$

$$\text{MAP} = \frac{1}{T} \sum_{t=1}^T \frac{\sum_{j=1}^{|R_t|} \text{pre}(j) \times \text{rel}(j)}{|R_t|} \quad (20)$$

其中, Y_t, \hat{Y}_t 分别为时间间隔 t 中所有路段的真实事故风险指数和预测事故风险指数。 R_t 为时间间隔 t 中真实发生交通事故的路段, S_t 为时间间隔 t 中预测得到的交通事故风险指数最高的 $|R_t|$ 个路段。 $\text{pre}(j)$ 表示预测得到的路段风险指数排序列表中从 1 至 j 个路段的准确率, $\text{rel}(j) = 1$ 表示该路段有事故风险发生, $\text{rel}(j) = 0$ 表示该路段没有事故风险发生。

3.2 基线模型

将提出模型 (CASTGCN) 与其他 6 种现有可用于交通事故风险指数预测的模型进行了对比。

MLP^[17]: 多层感知机,通过设置隐藏层捕获高阶函数关系。

LSTM^[18]: 一种经典的循环神经网络模型,广泛用于处理序列数据。

SVM^[19]: 一种基于交通流稳定性系数和支持向量机的交通风险预测方法。

ISTGCN^[11]: 一种使用改进的时空图卷积网络预测交通风险的方法。

ASTGCN^[20]: 一种用于交通指数预测的基于注意力的时空图卷积模型,其在时间和空间为都应用了集合注意力机制的卷积操作。

AGCRN^[12]: 一种应用了自适应图卷积和循环神经网络的交通风险预测模型。

3.3 实验与环境设置

文中模型基于 Pytorch 框架实现,自适应图神经网络的节点嵌入维度为 32,图卷积、时间注意力卷积、隐藏层大小均设置为 32,对比损失权重 α 设置为 0.2,混合参数 λ 设置为 0.1。时间注意力卷积模块在时间维度堆叠了 3 层卷积得到最终预测结果。模型训练采用了 Adam 优化算法,初始学习率为 0.000 1,训练批次大小为 128。实验使用 Python 3.6 编写,Pytorch 框架版本 1.10.2,硬件配置如表 1 所示。

表 1 实验硬件配置

配置项目	配置内容
操作系统	Linux Ubuntu5.15.0
CPU	Intel Xeon Gold 6150 2.70 GHz
内存	256 GB
GPU	NVIDIA Geforce RTX 3090
CUDA 加速	Cuda 11.4

3.4 实验结果分析

按上述使用不同的方法在相同场景下进行预测,

结果如表 2 所示。

表 2 不同方法在数据集上的实验结果比较

模型	MAE	MAP/%	Recall/%
MLP	0.76	11.13	16.05
LSTM	0.58	12.31	20.65
SVM	0.63	12.05	17.09
ISTGCN	0.50	12.48	21.27
ASTGCN	0.52	12.37	20.92
AGCRN	0.49	12.65	22.35
CASTGCN	0.40	13.68	23.91

从表 2 可以看出,文中模型在 3 种评价指标中均达到了最佳性能。其中,MLP 是经典深度学习结构,可以用于几乎所有预测任务,但由于其结构单一,在预测事故风险时缺少相应的模块建模路网信息的时空相关性,导致效果并不理想。SVM 基于交通流稳定性系数为输入,并且不需要搜索最优参数,效果优于 MLP,但这种方法将各个路段的特征及风险当作独立的样本,无法建模路网信息的时空相关性,因此预测能力有限。

LSTM 有能力捕获时间序列上长短期依赖,但是其难以捕获空间维度上路网信息的相关性,其效果仍然有限。ISTGCN、ASTGCN 和 AGCRN 作为时空数据预测模型,都集成了空间和时间两个维度的处理模块。然而,ISTGCN 和 ASTGCN 都基于预定义的邻接矩阵捕获特征的空间属性,其效果稍弱于使用了自适应图卷积的 AGCRN,因为基于预定义的图难以建模路网信息中完整的空间相关性。该文基于 Mixup 策略设计数据增广方法,使用对比学习解决事故数据在时间维稀疏的问题,获得了更好的事故风险样本的嵌入表示。此外,该文将自适应图学习和注意力机制结合,以数据驱动的方式挖掘交通数据中复杂的时空相关性,从而实现了最佳效果。

3.5 消融实验分析

为进一步分析模型各个组件的作用,设计了原模型的 3 个变体进行消融实验,将这 3 个变体与原模型进行比较。变体的描述如下:

wo_S:在原 CASTGCN 基础上,不使用自适应图卷积建模路网信息的空间相关性;

wo_T:在 CASTGCN 基础上,不使用时间注意力机制建模路网信息的时间相关性;

wo_Cta:在 CASTGCN 基础上,不使用对比学习方法,将对损失权重 α 设置为 0。

消融实验的结果如表 3 所示。

表 3 消融实验结果

模型	MAE	MAP/%	Recall/%
wo_S	0.53	12.18	20.65
wo_T	0.64	11.17	19.13
wo_Cta	0.46	12.62	22.57
CASTGCN	0.40	13.68	23.91

首先,在不建模路网信息的空间相关性的情况下(wo_S),模型在各项指标上都有所降低,由于具有相似拓扑连接关系,相似流量等交通量变化情况路段间的事故风险存在相关性,因此该文通过自适应图卷积捕获路网信息的空间相关性是有效的。此外,在不建模路网信息的时间相关性的情况下(wo_T),模型表现显著降低,因此该文基于注意力机制自适应地学习路网信息的时间相关性对于准确地事故风险预测是有效的。最后,在不使用对比损失的情况下(wo_Cta),模型效果有所降低,这说明基于 Mixup 策略的数据增广缓解了事故数量稀疏的影响,引入对比损失可以使模型更好地区分风险样本和非风险样本,因此实现了更准确的事故风险预测。

4 结束语

为更准确地预测高速公路的事故风险,该文提出了一种基于自适应时空相关性建模和对比学习的深度学习模型。使用自适应图卷积网络捕获路网信息的空间相关性,解决基于预定义图的图卷积运算难以包含路网信息完整空间相关性的问题。此外,基于注意力机制自适应建模路网信息的时间相关性,然后通过卷积层进一步聚合相邻时间片间的信息。接着,基于 Mixup 策略实现数据增广,解决交通事故在时间维度分布稀疏的问题,并利用对比学习方法使模型更好地区分风险样本和非风险样本,以实现更准确的事故风险预测。最后,通过实验证明了提出方法的有效性和先进性。今后将进一步挖掘高速公路静态特征数据,比如是否存在弯道、匝道、下坡等,以更全面地建模影响高速公路事故风险的因素。

参考文献:

- [1] 宁 静,余红艳,赵 东,等.一种路网级交通事故风险预测方法[J].北京邮电大学学报,2022,45(2):72-78.
- [2] SHARMA B,KATYAR V K,KUMAR K. Traffic accident prediction model using support vector machines with Gaussian kernel[C]//Proceedings of fifth international conference on soft computing for problem solving. Singapore:SocProS, 2016:1-10.
- [3] WILLIAMS B M,HOEL L A. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process;theoretical

- basis and empirical results[J]. *Journal of Transportation Engineering*, 2003, 129(6): 664–672.
- [4] CHEN Q, SONG X, YAMADA H, et al. Learning deep representation from big and heterogeneous data for traffic accident inference[C]//Thirtieth AAAI conference on artificial intelligence. Arizona: AAAI, 2016: 338–344.
- [5] HUANG C, ZHANG C, DAI P, et al. Deep dynamic fusion network for traffic accident forecasting[C]//Proceedings of the 28th ACM international conference on information and knowledge management. Beijing: CIKM, 2019: 2673–2681.
- [6] LIANG Y, KE S, ZHANG J, et al. Geoman: multi-level attention networks for geo-sensory time series prediction[C]//International joint conferences on artificial intelligence organization. Sweden: IJCAI, 2018: 3428–3434.
- [7] ZHAO L, GAO M, WANG Z. ST-GSP: spatial-temporal global semantic representation learning for urban flow prediction[C]//Proceedings of the fifteenth ACM international conference on web search and data mining. Arizona: WSDM, 2022: 1443–1451.
- [8] WANG P, GE C, ZHOU Z, et al. Joint gated co-attention based multi-modal networks for subregion house price prediction[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021, 35(2): 1667–1680.
- [9] YUAN Z, ZHOU X, YANG T. Hetero-convlstm: a deep learning approach to traffic accident prediction on heterogeneous spatio-temporal data[C]//Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining. London: SIGKDD, 2018: 984–992.
- [10] WANG B, LIN Y, GUO S, et al. GSNet: learning spatial-temporal correlations from geographical and semantic aspects for traffic accident risk forecasting[C]//Proceedings of the AAAI conference on artificial intelligence. Vancouver: AAAI, 2021: 4402–4409.
- [11] 王庆荣, 魏怡萌, 朱昌锋, 等. 基于时空图卷积网络的交通事故风险预测研究[J]. *计算机工程*, 2022, 48(11): 22–29.
- [12] BAI L, YAO L, LI C, et al. Adaptive graph convolutional recurrent network for traffic forecasting[J]. *Advances in Neural Information Processing Systems*, 2020, 33: 17804–17815.
- [13] ZHANG H, CISSE M, DAUPHIN Y N, et al. Mixup: beyond empirical risk minimization[C]//International conference on learning representations. Vancouver: ICLR, 2018.
- [14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in neural information processing systems. Cambridge: MIT Press, 2016: 5998–6008.
- [15] HE K, FAN H, WU Y, et al. Momentum contrast for unsupervised visual representation learning[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Seattle: IEEE, 2020: 9729–9738.
- [16] KINGMA D P, BA J. Adam: a method for stochastic optimization[C]//International conference on learning representations. San Diego: ICLR, 2015.
- [17] GARDNER M W, DORLING S R. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences[J]. *Atmospheric Environment*, 1998, 32(14–15): 2627–2636.
- [18] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735–1780.
- [19] 刘星良, 单珏, 刘唐志, 等. 基于交通流稳定性系数的高速公路交通事故实时风险预测[J]. *交通信息与安全*, 2022, 40(4): 71–81.
- [20] GUO S, LIN Y, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting[C]//Proceedings of the AAAI conference on artificial intelligence. Hawaii: AAAI, 2019: 922–929.