

基于时空图卷积网络的交通事故风险预测研究

王庆荣¹, 魏怡萌¹, 朱昌锋², 田可可¹

(1. 兰州交通大学 电子与信息工程学院, 兰州 730070; 2. 兰州交通大学 交通运输学院, 兰州 730070)

摘要: 交通事故的预测是通过对过去路段发生的交通事故进行分析, 在综合考虑影响交通事故的相关因素后, 对未来路段的交通事故发生状态进行预测。以往的大多数研究通常采用传统机器学习方法或单一深度学习模型预测法, 利用网格化确定预测空间的单位, 忽略了影响交通事故的天气、路况等外部因素, 导致模型的预测性能不佳。提出一种基于时空特性的城市交通事故风险预测模型, 在模型中使用改进的时空图卷积网络, 利用图卷积网络(GCN)提取空间相关特征, 并加入批标准化层解决梯度消失爆炸问题。在时间维度上采用门控线性单元(GLU)实现一维卷积操作, 提取时间相关特征, 并将GCN和GLU组合成时空卷积模块提取时空相关特征, 使用均方误差损失函数解决样本数据零膨胀问题。实验结果表明, 与GLU、SDCAE和ConvLSTM模型相比, 该模型的RMSE指标分别降低了28%、4.87%、4.19%, 能有效捕获时空相关性, 综合性能得到较大提升。

关键词: 深度学习; 城市交通事故; 时空图卷积网络; 时空相关性; 批标准化层

开放科学(资源服务)标志码(OSID):



中文引用格式: 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.

英文引用格式: WANG Q R, WEI Y M, ZHU C F, et al. Research on traffic accident risk prediction based on spatio-temporal graph convolutional network[J]. Computer Engineering, 2022, 48(11): 22-29.

Research on Traffic Accident Risk Prediction Based on Spatio-Temporal Graph Convolutional Network

WANG Qingrong¹, WEI Yimeng¹, ZHU Changfeng², TIAN Keke¹

(1. College of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;

2. College of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China)

[Abstract] Traffic accident prediction is the analysis of past road traffic accident data and other related factors to predict the future state of road traffic accidents. Most of the previous research methods have typically been traditional machine-learning methods or a single deep-learning model prediction method utilizing a grid to determine the prediction space unit, weather, road conditions, and other external factors affecting traffic accidents. However, these methods ignore the problem of zero expansion, which leads to poor prediction performance. Therefore, this paper presents a prediction model for urban traffic accident risk considering spatio-temporal characteristics. The Improved Spatio-Temporal Graph Convolution Network (ISTGCN) is used in the model. First, a Graph Convolution Network (GCN) is used to extract spatial correlation features, and Batch Normalization (BN) layer is added to solve the gradient-disappearing explosion problem. Second, Gated Linear Units (GLU) are used to extract time-dependent features by applying a one-dimensional convolution operation. Finally, the GCN and GLU are combined into a Space-Time Convolution (ST-Conv) module to extract the spatio-temporal correlation features and the Mean Squared Error (MSE) loss function is used to solve the problem of zero expansion of sample data. Experimental results show that compared with the GLU, SDCAE, and ConvLSTM models, the Root Mean Square Error (RMSE) of this model is reduced by 28%, 4.87%, and 4.19% respectively. Thus, this method can effectively capture the spatio-temporal correlation and improve the comprehensive performance.

[Key words] deep learning; urban traffic accident; Spatio-Temporal Graph Convolution (STGCN); spatio-temporal correlation; Batch Normalization (BN) layer

DOI: 10.19678/j.issn.1000-3428.0062961

基金项目: 国家自然科学基金(71961016); 教育部人文社会科学研究规划基金(18YJAZH148); 甘肃省自然科学基金(20JR10RA212, 20JR10RA214)。

作者简介: 王庆荣(1977—), 女, 教授, 主研方向为智能交通、数据挖掘、应急物流; 魏怡萌, 硕士研究生; 朱昌锋, 教授、博士、博士生导师; 田可可, 硕士研究生。

收稿日期: 2021-10-14 修回日期: 2022-01-21 E-mail: 0619685@stu.lzjt.edu.cn

0 概述

随着科技和社会的发展与进步,车辆和道路在增加,交通事故发生的频率也在不断增加,人们的生命和财产受到严重威胁^[1]。预测交通是一项具有挑战性的任务,虽然日常通勤是相对可预测的,但在节假日、休闲活动、恶劣天气等情况下的交通出行均具有很强的随机性,较难预测。目前,通过交通大数据预测交通事故风险成为研究热点,如果能准确预测交通事故,并根据预测结果对交通事故的发生进行重新规划,那么将提高城市交通效率,减少交通事故的发生导致的城市、个人经济损失,同时也能保障人们的人身安全。

目前,研究人员对交通事故预测方面的研究可以分为两大类。第一类是传统预测方法,比如统计回归法、灰色预测法等。传统回归法常用于短期数据变化的预测,但对于随机性较大、干扰因素较多的交通数据来说,其预测的结果比较片面,可靠性难以得到保障。灰色预测法适用于样本数量较少的预测,其模型简单,对中短期预测有较好的预测结果,但对于中长期的预测仍有不足。第二类是神经网络预测法,该方法的自主学习能力较好,非线性映射和高鲁棒性也较强,在交通预测领域的应用前景较广泛。虽然现有对交通事故的预测研究已有较好的成果,但交通事故的发生受很多复杂因素影响,如驾驶员的行为等不可控因素给交通事故的预测增加了许多困难,交通事故的复杂机制依旧不太明确。

经典的机器学习技术大多将交通事故的预测看作分类问题,目的是预测下个时段是否会发生交通事故,未考虑相关复杂因素的影响,因此预测结果相对较差。文献[2]使用逻辑回归方法对美国4号洲际公里的交通事故进行预测,文献[3]提出连续的非凸优化的k-means算法,通过一个等效模型,利用矩阵优化理论分析和解决交通预测模型问题。文献[4]通过动态预测模型建立基于马尔可夫链和云模型预测道路交通事故的数量。文献[5]提出使用多层感知器模型预测事故热点地区的事故情况。文献[6]提出基于反向传播(Back Propagation, BP)神经网络的高速公路事故预测,但BP神经网络的缺点是易陷入局部最小点,收敛速度较慢,训练时间也较长。文献[7]提出一个组合预测优化模型预测交通事故,以提高预测精度,但在如何更好地结合单一模型方面还有待提高。文献[8]提出灰色BP神经网络模型预测发生交通事故后的车型分担率,并结合BP神经网络模型和灰色预测模型进行优势互补,相比单纯的BP神经网络有较高的预测精度和稳定性。文献[9]建立了基于改进深度森林算法的交通事故风险状态等级预测模型,通过对比传统预测模型验证该模型的优势。文献[10]通过引入交通风险概念,建立一个堆叠自编码器模型,研究在城市道路中过路人流量的变化对交通事故风险的影响。文献[11]对长短

期记忆(Long Short-Term Memory, LSTM)模型进行改进后提出双尺度LSTM模型,用以预测交通事故,并得到了很好的效果,但混合模型的训练时间较长,且未考虑复杂的天气、路况等影响因素。

文献[12]提出一种新的堆栈去噪卷积自动编码模型用于预测城市交通事故风险,并使用空间依赖性学习事故中的隐藏因素,实验表明该模型的性能优越,但该模型未考虑天气因素的影响。文献[13]提出一种基于时空注意力网络的区域交通事故风险预测深度学习框架,并综合影响交通事故的外部因素,提高了模型的预测性能。文献[14]提出基于卷积神经网络的车边缘网络,用于交通事故风险预测,经过实验对比发现,该网络的精度较传统机器学习网络较高。文献[15]提出基于路网结构预测短期城市交通事故,采用多层图卷积神经网络和双链LSTM构建预测模型,发现时空特征属性对最终的预测结果影响较大,但该模型存在多源数据和因子分析缺乏的问题。

综上所述,目前交通事故风险预测的研究主要存在以下几点问题:一是未考虑在实际场景下,发生交通事故的多源外部影响因素,如车流量、路况、天气、兴趣点(POI)等;二是未考虑时空特征对交通事故产生的影响;三是现有研究大部分是对交通事故量预测的研究,在交通事故风险预测方面的研究较少。

本文采用神经网络方法构建基于时空特征的交通事故风险预测模型,在考虑时间依赖、空间依赖、时空依赖的情况下,在模型中使用改进的时空图卷积网络(Improved Spatio-Temporal Graph Convolutional Network, ISTGCN),通过添加批标准化(Batch Normalization, BN)层解决梯度消失爆炸问题,并采用均方误差(Mean Squared Error, MSE)损失函数,解决样本零膨胀问题。

1 图卷积网络

图卷积网络(Graph Convolutional Network, GCN)^[16]和卷积神经网络(Convolutional Neural Networks, CNN)均可以作为特征提取器,但GCN的对象是图数据,其先从图数据中提取相关特征,然后用这些特征对图数据进行结点分类、图分类、边预测等。一个图的结构如式(1)所示:

$$G=(V, E) \quad (1)$$

其中: V 是图的结点集合; E 是图边集合。设 W 是图的邻接矩阵, $|V|$ 表示结点的个数, D 是图的度矩阵,其中 $D=\text{diag}(d_1, d_2, \dots, d_N)$, $d_i=\sum_{j=1}^N W_{ij}$,即每个结点的邻结点个数。

GCN也是一个神经网络层,每层之间的传输方式如式(2)所示:

$$H^{(l+1)}=\sigma(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}) \quad (2)$$

其中: $\tilde{A}=A+I$; \tilde{D} 为 \tilde{A} 的度矩阵; I 是单位矩阵; $\tilde{D}_{ii}=\sum_{j=1}^N \tilde{A}_{ij}$

$\sum j\tilde{A}_j H$ 表示每一层的特征;相对输入层来说, H 即为 X , σ 是非线性激活函数。

2 ISTGCN 模型

2.1 时空图卷积网络的结构

时空图卷积网络(Spatio-Temporal Graph Convolutional Network, STGCN)在交通中的应用最开始用于交通事故预测,后来逐渐被学者应用于交通流^[17]的预测。由于交通流和交通事故均在一定的空间和时间背景下发生,因此均会受到时空特性相关复杂因素的影响,且交通事故相比于交通流受时空特性的影响更大,且交通流也是交通事故的一个影响因素。本文将STGCN网络用于交通事故预测中,并验证其预测性能。

STGCN网络的输入是 $N \times C \times V \times T$,先通过尺寸为 1×1 的2d卷积层输出 $N \times (C \times K) \times V \times T$,再与邻接矩阵 $A(K \times V \times V)$ 相乘,这里 K 表示子集的个数,得到空间卷积 $N \times C \times T \times V$,最后通过一个定义的时间长度 $n_t \times 1$,2d卷积层进行时间维度的卷积。其中: N 为个数; C 为通道数,即特征维度; V 为顶点个数; T 为时间维度。STGCN网络的结构如图1所示。

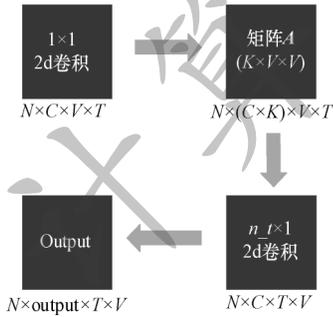


图1 STGCN网络的基本结构
Fig.1 Structure of STGCN network

交通事故的致因主要分为空间相关性、时间相关性和时空相关性三大类,本文对每类因素的不同特征采取不同的网络结构,依次进行特征提取,并将处理后的相关数据放入本文模型中进行训练。

2.2 ISTGCN 模型框架

ISTGCN模型由2个时空卷积模块和1个全连接层组成,时空卷积块的每个模块有2个时间门控序列卷积层和中间1个空间卷积模块,如图2所示为本文模型的结构。首先使用GCN网络提取相关特征的空间属性,然后使用门控线性单元(Gated Linear Units, GLU)来建模事故的相关时间依赖,最后通过时空卷积模块,并利用图形和卷积捕捉时空特征的相关动态变化。此外,以构造好的时空相关属性的图信号权重矩阵作为输入,通过时空卷积模块的时空相关性预测模块,并将前3个模块的输出进行加权融合,输出最终的预测值。

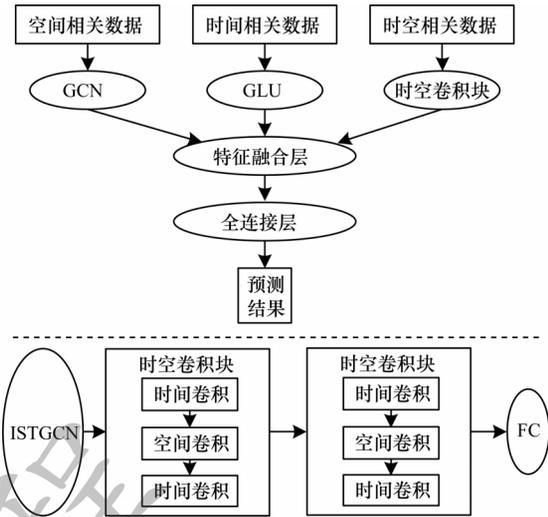


图2 ISTGCN模型结构
Fig.2 Structure of ISTGCN model

空间卷积层由图卷积组成,该卷积将路段及其邻域的空间信息聚合在一起。时间卷积层在图卷积操作上捕获时间维度的标准卷积层,并通过合并连续时隙中的有效信息以更新各个节点的有效信号。使用空间图卷积捕获空间相关属性,利用时间门控卷积捕获时间相关属性,并采用时空卷积块融合空间域和时间域的特征。

为解决梯度爆炸问题,本文在模型中加入BN层,并采用MSE损失函数,在计算损失时,对交通事故风险较大的样本给予较大的权重,避免模型预测结果聚集在0附近。

2.2.1 空间卷积模块

在一般情况下,用图结构表示交通网络,用数学公式表示路网。以往的研究忽略了交通网络的空间属性,交通网络常被分为多个分段或网络,导致网络的连通性和整体性被破坏。本文模型将直接在图结构数据上使用图卷积,用来提取空间域中有意义的特征,并计算图卷积中的核 Θ 。由于图傅里叶基^[18]运算代价较高,因此采用切比雪夫多项式^[19]逼近策略来克服该问题。

切比雪夫多项式在逼近理论中有很重要的作用,其是一种正交多项式序列,以一种递归的方式被定义。通常第1类切比雪夫多项式以符号 T_n 表示,第2类切比雪夫多项式用 U_n 表示。切比雪夫多项式 T_n 或 U_n 代表 n 阶多项式。为了使滤波器局部化并减少参数数量,核 Θ 可以限制为 A 的多项式,例如

$$\Theta(A) = \sum_{k=0}^{K-1} \theta_k A^k, \theta \in \mathbb{R}^K \text{ 是多项式系数的向量。}$$

K 是图卷积的核大小,它决定了中心节点卷积的最大半径。在一般情况下,切比雪夫多项式 $T_k(x)$ 被用作近似核,其 $K-1$ 阶近似核截断展开式为

$$\Theta(\Lambda) \approx \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\Lambda}), \text{ 其中 } \tilde{\Lambda} = 2\Lambda/\lambda_{\max} - I_n, \lambda_{\max} \text{ 表示 } L \text{ 的最大特征值。}$$

将图卷积重写为:

$$\Theta * gx = \Theta(L)x \approx \sum_{k=0}^{K-1} \theta_k T_k(\tilde{L})x \quad (3)$$

其中: $T_k(\tilde{L}) \in \mathbb{R}^{n \times n}$ 是在按比例拉普拉斯算子 $\tilde{L} =$

$2L/\lambda_{\max} - I_n$ 计算的 k 阶切比雪夫多项式,通过多项式逼近递归计算 K -局域卷积。

选择几个空间卷积空间段组成空间分量,选用相应的表示形式,并通过全连接(Fully Connected FC)层略微减小空间特征。然后,将特征输入到堆叠的空间卷积块中,一个空间卷积块由几个空间卷积层组成,聚集了路段与领域的空间信息,通过在该空间卷积层中加入BN层来增加模型初始化的健壮性,提升训练速度与分类效果。BN层对于初始化的要求不高,所以可以使用较大的学习率进行训练,假设第1层的输入和输出分别为 $H^{(l)} \in \mathbb{R}^{N \times F}$ 和 $H^{(l+1)} \in \mathbb{R}^{N \times F}$,则空间卷积层的计算式如式(4)所示:

$$H^{(l+1)} = \text{ReLU}(\text{BN}(H^{(l)} * g\Theta)) \quad (4)$$

其中:BN是标准归一化;ReLU为激活函数; H 是可训练参数。

2.2.2 时间卷积模块

通过门控线性单元(Gated Linear Unit, GLU)提取时间相关属性,相比LSTM来说,GLU模型更加简单,收敛速度更快,需要的op(Operation)更少,这样一来,模型所需非线性的操作也就更少,因此能够有效降低梯度弥散现象,提高模型训练及收敛速度。

在时间维度上采用一维因果卷积对数据进行并行处理,采用GLU实现时间维度上的卷积操作,在每个节点上使用GLU捕获时间上的依赖关系。首先在CNN的卷积层中引入门控线性单元机制,使卷积层的输出变成式(5)所示,表示一个没有非线性函数的卷积层输出,再通过ReLU非线性激活函数的卷积层输出,表达式如式(5)所示:

$$h_i(X) = (X * Q + b) \otimes \text{ReLU}(X * V + c) \quad (5)$$

其中: X 是层 h_i 的输入,即前一层的输出; Q 和 V 表示不同的卷积核; n 表示输出通道数; b 和 c 为偏置参数。式(5)后半部分含激活函数的卷积,即门控机制,其控制了 $X * Q + b$ 当中的哪些有效信息可以传去下一层,然后将该模型进行堆叠以便获取时间相关属性。在获取空间维度上各个节点的相邻信息时,采用基于图形卷积操作上的时间卷积层,通过合并连续时隙中的信息来更新节点信息。

时间特征可以反映每个路段的历史交通状况,第1个时间卷积层的第 v_i 路况的输入信号为 $H_i^{(l)} \in \mathbb{R}^{F \times T}$,输入信号的第 j 通道 $H_{i,j}^{(l+1)} \in \mathbb{R}^{F \times T}$ 的计算式如式(6)所示:

$$H_{i,j}^{(l+1)} = \sigma(b_j^{(l)} + \sum_{k=0}^F W_{f,k}^{(l)} * H_{i,k}^{(l)}), 1 \leq j \leq F \quad (6)$$

其中: $*$ 是有效的互相关运算符; $H_{i,k}^{(l)}$ 表示第1层输入信号的第 k 个通道。卷积核大小是 3×1 ,步长为1,卷积核中输出信号中的特维数与输入信号的特征维数保持一致。

2.2.3 时空卷积模块

本文构造了时空卷积块(Spatio-Temporal Convolutional Block, ST-Conv)以联合处理图结构时间序列,融合时空域的特征。ST-Conv利用图形和卷积捕捉空间和时间相关特征属性的动态变化,时空卷积模块可以根据特定情况的规模和复杂性进行堆叠和扩展。一个空间卷积块由2个时间卷积核和1个空间层组成,空间卷积层是连接2个时间卷积层的桥梁,通过时间卷积层可以实现图卷积到时间卷

积的快速时空传播^[20]。

将GCN网络与加入CNN的门控线性单元GLU组合成ST-Conv,把输入的数据先做时间维度的卷积,将时间维度卷积的输出结果再做图卷积操作,图卷积的输出结果经过1个Relu之后再进行一次时间维度卷积,以上即为整个ST-Conv的输出。

此外,本文使用Relu激活函数,在叠加2个ST-Conv之后增加1个额外的时域卷积层,并和1个全连通层共同作为输出层,以防止过拟合,ST-Conv的输入和输出都是3D的张量。

2.2.4 融合预测模块

堆叠2个时空卷积模块及1个时间卷积模块的输出后,通过拼接操作将空间、时间、时空以及外部信息融合在一起,形成一个融合的表达,将融合后的表示输入,得到输出层,并使用式(7)进行融合:

$$Y = \text{FC}(W_1 * Y_1 + W_2 * Y_2 + W_3 * Y_3) \quad (7)$$

其中: $*$ 表示卷积操作; W_1, W_2, W_3 为卷积核参数; Y_1, Y_2, Y_3 为3个模块的输出;FC为全连接层,输出结果 $Y \in \mathbb{R}^{1 \times J}$ 是下一个阶段发生交通事故的风险。

在交通大数据中,交通事故数据与交通流数据存在较大差别,例如交通事故样本中会出现“零值”,代表伤害程度最低等级,比如未受伤。研究表明,当事故样本中含有大量的零值时,容易导致各影响因素对“零值”的边际概率被高估,若忽略这种零膨胀现象,将导致有偏估计,从而导致错误的结论^[21]。因此,本文采用MSE损失函数解决数据零膨胀问题,在计算损失时给予事故风险样本较大的权重值,以避免预测值在0附近。本文将事故风险分为3个等级(轻度、中度和重度),对应的风险值 i 分别为1、2、3,计算式如式(8)所示:

$$\text{Loss}(Y, Y') = \frac{1}{2} \sum_{i \in I} \lambda_i (Y(i) - Y'(i))^2 \quad (8)$$

其中: Y 表示真实值; Y' 表示预测值; λ_i 代表权重值(事故风险等级为 i 的样本权重)。

本文模型的流程如图3所示。

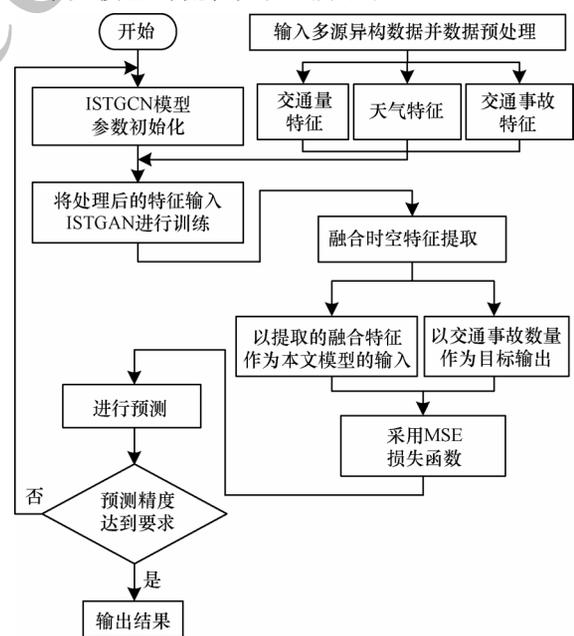


图3 本文模型的流程

Fig.3 Procedure of model in this paper

3 实验结果与分析

3.1 数据来源与分析

选用 Kaggle 上的 2005 ~ 2018 年英国道路安全交通事故和车辆公开数据集,该数据集是实时收集的,使用了多个流量应用程序编程接口(Application Programming Interface, API)。这些 API 广播各种实体捕捉到的交通事件,比如英国和国家运输部门、执法机构、交通摄像头以及道路网络内的交通传感器,目前该数据集大约包含 120 000 起事故记录。本文选取了 2005 ~ 2018 年期间部分事故数据进行验证,并从 2018 年 1 月到 2018 年 12 月的全英国事故中,挑出部分事故严重的时间段进行了可视化,统计分析图如图 4 所示。

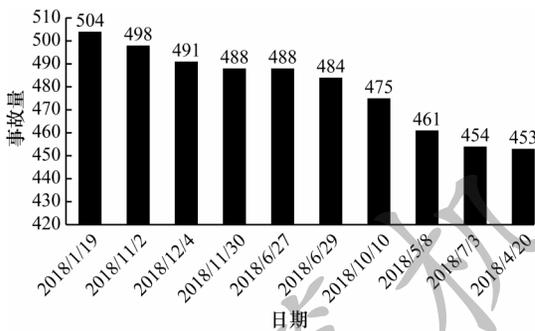


图 4 不同时间下的交通事故量
Fig.4 Traffic accidents at different time

本文收集的数据集均是英国的交通事故量、交通流量、天气数据,此数据集一共包含 32 条属性,包括事故记录 ID、严重程度、开始时间、结束时间、速度、经纬度、风向等。所选取数据的具体分类如下:

1) 交通事故。该数据包含了 2005 ~ 2018 年期间所有记录的交通事故。除了街道相关信息、时间、地点等一些基本信息外,该数据集还包括撞车的原因。本文把数据集按事故严重程度等级划分为 1, 2, 3, 数值越小,表示事故越严重,该数据集不同风险等级下的交通事故量统计如图 5 所示。

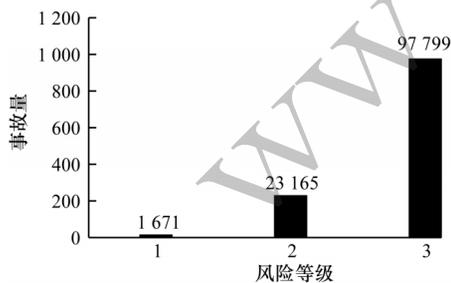


图 5 不同风险等级下的交通事故量

Fig.5 Traffic accidents quantity under different risk levels

2) 交通流量数据。由相关识别设备检测到的行车记录计算而来,行车记录包括设备 ID、车牌号、车辆过车时间、方向、车道号 5 个属性,正常路段车速对交通事故的影响如图 6 所示。

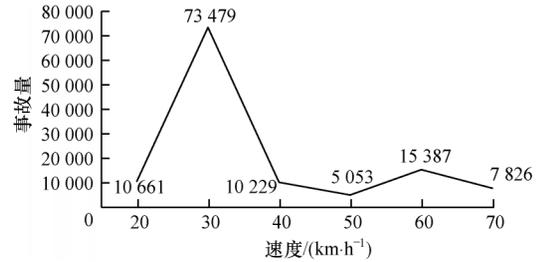


图 6 交通事故量与速度的关系

Fig.6 Relationship between traffic accidents volume and speed

3) 天气数据。本文采集的天气数据来自英国天气数据中心,数据属性有时间、天气、经纬度、能见度、湿度、温度、风向等。天气状况对交通事故的影响如图 7 所示。对天气状况进行标签化,将天气条件恶劣程度等级划分为 -1~9, 数值越小,表示天气越恶劣。由图 7 可以看出天气条件严重恶劣情况下造成的交通事故量最多。

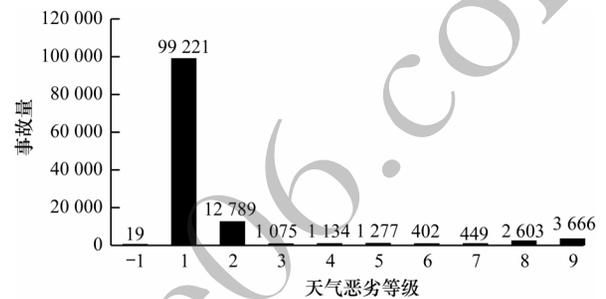


图 7 交通事故量与天气的关系

Fig.7 Relationship between traffic accidents volume and weather

4) POI 数据,POI 数据也来自英国公共数据中心,包含学校、购物中心等的 POI 坐标。

3.2 数据预处理

对收集到的数据进行去除冗余值、补全缺失值等预处理操作之后,将其转化为权重矩阵的形式,并作为模型空间预测单元的输入。数据对象每一个属性权重都是该权重占所有权重的比例,定义一个二元空间对称矩阵 H ,以此来表达 n 个节点位置区域的相邻关系,如式(9)所示:

$$H = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \vdots & \vdots & & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{bmatrix} \quad (9)$$

其中: w_{ij} 指的是第 i 行数据对象的第 j 个属性占所有权重的比例。

图 8 为构造图形结构^[22]的数据权重矩阵,数据点 v_i 为一个图形信号,该图形信号定义在无向图 $G = (v, \varepsilon, w)$ 上,其权重为图 8 的 w_{ij} , v_i 是一组有向的顶点,对应于交通道路网络中监测站的观测值, ε 表示站与站之间连通的一组边。

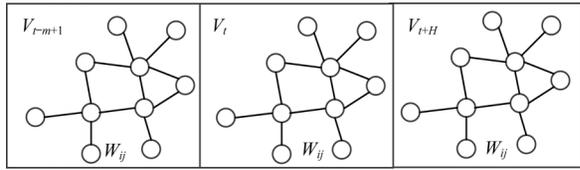


图 8 图形结构的数据权重矩阵

Fig.8 Data weight matrix of the graph structure

3.3 评价指标

本文使用平均相对误差 (Mean Relative Error, MAE)、均方根误差 (Root Mean Square Error, RMSE)、召回率 (Recall) 等评价指标评估预测模型,相关计算式如下所示:

$$M_{MAE} = \frac{1}{N} \sum_{t=1}^N |y(t) - \tilde{y}(t)| \quad (10)$$

$$R_{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (y(t) - \tilde{y}(t))^2} \quad (11)$$

$$R_{Recall} = \frac{T_{TP}}{T_{TP} + F_{FN}} \quad (12)$$

其中: $y(t)$ 为 t 时刻交通事故的实测值; $\tilde{y}(t)$ 为预测值; N 为样本个数; T_{TP} 表示实则为正类且预测为正类的样本数目; F_{FN} 表示实则为正类且预测为负类的样本数目。

3.4 结果分析

本文实验在 Python3.6 和 TensorFlow1.9 的环境下进行,将所选用数据集的交通流量、交通事故量、外部特征属性作为训练输入,2018 年的交通流量、外部特征作为预测模型输入,模型输出交通事故量。在实验中,按照时间顺序以 6:2:2 的比例划分训练集、验证集、测试集。基于 TensorFlow 框架实现 ISTGCN 模型,经过实验比较,使用 3 层图卷积,卷积核大小均为 3; GLU 为 3 层,每层隐藏单元个数为 128,使用 BN 层将数据归一化。每个实验包含 7 个对比交通事故预测模型:历史平均 HA 模型(利用相同时段事故风险的平均值预测下一时段的风险)、基于提升树的机器学习 (XGBoost) 模型^[23]、多层感知机 (Multilayer Perceptron, MLP) 模型、门控循环单元网络 (Gate Linear Unit, GLU) 模型^[24]、基于 CNN 的堆叠降噪自动编码器模型、CNN 和 LSTM 相结合 (ConvLSTM) 模型^[25] 和本文所提出的 ISTGCN 模型。

3.4.1 对比实验分析

将本文模型和对比模型在英国真实数据集上进行预测分析,分别迭代 200 次和 300 次,所有对比模型均在同一数据集下进行调试,结果如表 1 所示。由表 1 可以看出,迭代 300 次的整体预测结果比迭代 200 次的好,HA、XGBoost、MLP 和 GLU 模型的整体预测效果都比较差,这些模型均没有捕获空间特性。SDCAE 模型较 HA、XGBoost、MLP 和 GLU 4 个模型的整体预测效果较好,但却忽略了时间的长短期相关性。ConvLSTM 模型同时考虑了时间和空间相关性依赖,但却忽略了时空特性对交通事故预测的

影响,相比之下,本文 ISTGCN 模型在同时考虑时间、空间、时空相关属性依赖后,取得了较好的预测结果。

表 1 不同模型的结果对比

Table 1 Results comparison of different models

| 模型 | epoch=200 | | | epoch=300 | | |
|----------|-----------|---------|----------|-----------|---------|----------|
| | RMSE | MAE | Recall/% | RMSE | MAE | Recall/% |
| HA | 5.942 3 | 4.816 9 | 55.35 | 4.459 3 | 4.177 6 | 57.54 |
| XGBoost | 6.113 2 | 5.386 7 | 52.84 | 5.015 4 | 4.856 3 | 56.51 |
| MLP | 2.437 9 | 1.997 5 | 63.78 | 2.237 9 | 2.138 5 | 65.01 |
| GLU | 2.235 7 | 1.834 1 | 64.93 | 2.200 6 | 2.072 1 | 65.36 |
| SDCAE | 1.676 4 | 0.690 2 | 66.81 | 1.665 4 | 0.702 4 | 67.72 |
| ConvLSTM | 1.653 1 | 0.647 6 | 67.09 | 1.649 7 | 0.667 2 | 68.50 |
| ISTGCN | 1.586 6 | 0.496 9 | 69.35 | 1.584 3 | 0.523 8 | 70.29 |

3.4.2 模型结构分析

对本文 ISTGCN 模型的模块结构进行对比分析,在去掉空间层和时空层后,对剩下结构进行模型预测训练。由于模型在迭代 300 次时效果最好,因此本文结构分析分别是在迭代 300 次后的预测结果,预测结果如表 2 所示,其中 S、ST 分别代表空间层和时空层。由表 2 可以看出,ISTGCN 模型的每块结构都有不同的作用性能,去除任一模块结构,得到的预测结果都很差,ISTGCN 模型采用空间层获取空间相关属性,使用时间层获取时间相关属性,利用时空卷积层捕获时空相关特征,只有这 3 种结构相互作用,才能得到最佳性能。为验证 ISTGCN 模型性能,使用运行时间消耗这一指标进行性能分析,表 3 所示为不同模块结构训练时间的对比,本文模型 ISTGCN 实现了更快的训练速度,在去除任意结构之后性能表现都不佳,只有 3 种结构彼此相互作用,才能实现最佳性能。

表 2 本文模型结构对预测性能的影响

Table 2 Influence of structure of model in this paper on prediction performance

| 模型结构 | RMSE | MAE | Recall/% |
|-----------|---------|---------|----------|
| ISTGCN-S | 1.611 2 | 0.631 7 | 68.76 |
| ISTGCN-ST | 1.591 0 | 0.610 5 | 69.12 |
| ISTGCN | 1.584 3 | 0.523 8 | 70.29 |

表 3 数据集训练时间消耗

Table 3 Training time consumption of data sets 单位:s

| 模型结构 | 时间消耗 |
|-----------|----------|
| ISTGCN-S | 1 422.51 |
| ISTGCN-ST | 1 208.34 |
| ISTGCN | 1 137.11 |

3.4.3 激活函数分析

分别用 3 种不同激活函数验证模型性能,迭代次数选择 300 次,对比在激活函数前和激活函数后加入 BN 层的变化,激活函数对模型预测性能的影响

结果如表4所示。由表4可以看出,当激活函数为ReLU的时候,预测性能整体较好,将BN层加在激活函数前比加在激活函数之后效果更好,这是因为卷积层在进行BN操作时,会把数据分布在指定的区间,当ReLU函数再进行激活操作时,神经元失活的概率会下降,模型性能得到较大幅度的提升。

表4 激活函数对模型预测性能的影响

Table 4 Impact of activation functions on predictive performance

| 激活函数 | 在BN层之前 | | 在BN层之后 | |
|---------|---------|---------|---------|---------|
| | RMSE | MAE | RMSE | MAE |
| Sigmoid | 1.586 2 | 0.529 4 | 1.588 3 | 0.531 1 |
| Tanh | 1.584 9 | 0.526 6 | 1.587 9 | 0.528 1 |
| Relu | 1.584 3 | 0.523 8 | 1.585 7 | 0.526 9 |

如果先在ReLU函数后进行BN操作,ReLU函数激活后的部分神经元已失活,参与BN操作时的神经元数量就会减少,模型的泛化能力降低。由表4也可以看出ReLU函数在BN层前后的变化并没有Sigmoid和Tanh函数变化大,这是由于BN层在ReLU函数前可以让ReLU函数更好地产生特征选择作用,由于输入值的分布更接近于零均值,此时ReLU函数可以展现单侧抑制作用,且不会出现输入全分布在0的一侧。但如果将BN操作放在ReLU函数之后,在激活函数后进行归一化能更好地保证下一层的输入是零均值,能避免下一层连接权重的梯度同号,模型训练效率更高,收敛更快。由表5结果可知,将ReLU函数放在BN前后的差别不大,但也因不同实验而有区别。

4 结束语

本文提出一种基于时空图卷积的交通事故风险预测模型,以多源异构数据作为输入,分别基于空间卷积层、时间卷积层和时空卷积层处理不同类型的特征,并将处理后的隐含特征拼接成紧凑的表示,馈入全连接层,学习不同特征之间的相互作用,预测下一个阶段的交通事故风险。在空间卷积层中加入BN层,解决梯度爆炸问题,应用MSE损失函数来解决数据零膨胀问题。在英国真实数据集上的实验结果表明,与GLU、SDCAE、ConvLSTM模型相比,该模型的RMSE指标分别降低3.28%、4.87%、4.19%。下一步将优化网络结构与参数,通过收集违章、行人、出租车等数据,训练并优化模型,提高模型的运行速度和预测效率。

参考文献

- [1] FITZPATRICK L. Global status report on road safety 2015: time for action[EB/OL]. [2021-09-04]. <https://www.abe.pl/pl/book/9789241565066/global-status-report-on-road-safety-2015-time-for-action>.
- [2] ABDEL-ATY M, UDDIN N, PANDE A, et al. Predicting freeway crashes from loop detector data by matched case-control logistic regression[J]. Journal of the Transportation Research Board, 2004, 1897(1): 88-95.
- [3] TIAN Z, ZHANG S R. Application of big data optimized clustering algorithm in cloud computing environment in traffic accident forecast[J]. Peer-to-Peer Networking and Applications, 2021, 14(4): 2511-2523.
- [4] MA X B, ZHANG J Y, HUANG P, et al. Traffic accident prediction based on Markov chain cloud model[EB/OL]. [2021-09-04]. https://www.researchgate.net/publication/342785308_Traffic_accident_prediction_based_on_Markov_chain_cloud_model.
- [5] ROLAND J, WAY P D, FIRAT C, et al. Modeling and predicting vehicle accident occurrence in Chattanooga, Tennessee[J]. Accident Analysis & Prevention, 2021, 149: 60-68.
- [6] 邓晓庆, 孟祥海, 郑来. 基于BP神经网络的高速公路事故预测模型[J]. 交通信息与安全, 2016, 34(1): 78-84. DENG X Q, MENG X H, ZHENG L. An accident prediction model for expressway based on BP neural network[J]. Journal of Transport Information and Safety, 2016, 34(1): 78-84. (in Chinese)
- [7] 宋英华, 程灵希, 刘丹, 等. 基于组合预测优化模型的交通事故预测研究[J]. 中国安全科学学报, 2017, 27(5): 31-35. SONG Y H, CHENG L X, LIU D, et al. Research on optimal combination model for forecasting traffic accident [J]. China Safety Science Journal, 2017, 27(5): 31-35. (in Chinese)
- [8] 胡立伟, 张婷, 郭凤香, 等. 基于灰色BP神经网络的道路交通事故车型分担率预测及其预防策略研究[J]. 武汉理工大学学报(交通科学与工程版), 2018, 42(3): 388-392, 397. HU L W, ZHANG T, GUO F X, et al. Traffic accident split rate of vehicle types prediction and prevention strategies study based on gray BP neural network [J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2018, 42(3): 388-392, 397. (in Chinese)
- [9] 李志慧, 孙雅倩, 陶鹏飞, 等. 交通事故后的交通运行风险状态等级预测方法[J]. 吉林大学学报(工学版), 2021, 52(1): 127-135. LI Z H, SUN Y Q, TAO P F, et al. Prediction method of traffic operation risk level after traffic accident[J]. Journal of Jilin University (Engineering and Technology Edition), 2021, 52(1): 127-135. (in Chinese)
- [10] CHEN Q J, SONG X, YAMADA H, et al. Learning deep representation from big and heterogeneous data for traffic accident inference[C]//Proceedings of the 30th Conference on Artificial Intelligence. [S. l.]: AAAI Press, 2016: 338-344.
- [11] 李文书, 邹涛涛, 王洪雁, 等. 基于双尺度长短期记忆网络的交通事故量预测模型[J]. 浙江大学学报(工学版), 2020, 54(8): 1613-1619. LI W S, ZOU T T, WANG H Y, et al. Traffic accident quantity prediction model based on dual-scale long short-term memory network[J]. Journal of Zhejiang University (Engineering Science), 2020, 54(8): 1613-1619. (in Chinese)

- [12] CHEN C, FAN X L, ZHENG C P, et al. SDCAE: stack denoising convolutional autoencoder model for accident risk prediction via traffic big data[C]//Proceedings of the 6th International Conference on Advanced Cloud and Big Data. Washington D. C., USA: IEEE Press, 2018: 328-333.
- [13] ZHU L, LI T R, DU S D. TA-STAN: a deep spatio-temporal attention learning framework for regional traffic accident risk prediction[C]//Proceedings of International Joint Conference on Neural Networks. Washington D. C., USA: IEEE Press, 2019: 1-8.
- [14] 赵海涛,程慧玲,丁仪,等. 基于深度学习的车联边缘网络交通事故风险预测算法研究[J]. 电子与信息学报, 2020, 42(1): 50-57.
ZHAO H T, CHENG H L, DING Y, et al. Research on traffic accident risk prediction algorithm of edge Internet of vehicles based on deep learning[J]. Journal of Electronics & Information Technology, 2020, 42(1): 50-57. (in Chinese)
- [15] 张延孔,卢家品,张帅超,等. 基于路网结构的城市交通事故短期风险预测方法[J]. 智能系统学报, 2020, 15(4): 663-671.
ZHANG Y K, LU J P, ZHANG S C, et al. A short-term risk prediction method for urban traffic accidents based on road network[J]. CAAI Transactions on Intelligent Systems, 2020, 15(4): 663-671. (in Chinese)
- [16] GORI M, MONFARDINI G, SCARSELLI F. A new model for learning in graph domains[C]//Proceedings of 2005 IEEE International Joint Conference on Neural Networks. Washington D. C., USA: IEEE Press, 2005: 729-734.
- [17] 荣斌,武志昊,刘晓辉,等. 基于时空多图卷积网络的交通站点流量预测[J]. 计算机工程, 2020, 46(5): 26-33.
RONG B, WU Z H, LIU X H, et al. Flow prediction of traffic stations based on spatio-temporal multi-graph convolutional network[J]. Computer Engineering, 2020, 46(5): 26-33. (in Chinese)
- [18] 李社蕾,周波,杨博雄,等. 图卷积神经网络基础图结构影响研究[J]. 小型微型计算机系统, 2021, 42(4): 891-896.
LI S L, ZHOU B, YANG B X, et al. Study on the influence of basic graph structure in graph convolutional network[J]. Journal of Chinese Computer Systems, 2021, 42(4): 891-896. (in Chinese)
- [19] MASON J C, HANDSCOMB D C. Chebyshev polynomials[EB/OL]. [2021-09-04]. <https://www.scirp.org/reference/ReferencesPapers.aspx?ReferenceID=575225>.
- [20] YU B, YIN H T, ZHU Z X. Spatio-temporal graph convolutional networks; a deep learning framework for traffic forecasting[EB/OL]. [2021-09-04]. <https://arxiv.org/abs/1709.04875>.
- [21] 李俊辉,汤左淦. 交通事故伤害程度估计中的零膨胀现象研究[J]. 公路工程, 2020, 45(5): 239-245.
LI J H, TANG Z G. An analysis for addressing excess zeros phenomenon in estimating traffic crash injury severity[J]. Highway Engineering, 2020, 45(5): 239-245. (in Chinese)
- [22] 王彦浩. 基于图神经网络的交通事故预测方法研究[D]. 北京:华北电力大学, 2021.
WANG Y H. Research on traffic accident prediction based on graph neural network[D]. Beijing: North China Electric Power University, 2021. (in Chinese)
- [23] CHEN T Q, GUESTIN C. XGBoost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM Press, 2016: 785-794.
- [24] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[EB/OL]. [2021-09-04]. <https://arxiv.org/abs/1412.3555>.
- [25] SHI X J, CHEN Z R, WANG H, et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting[EB/OL]. [2021-09-04]. <https://arxiv.org/abs/1506.04214>.