

城市区域功能感知的细粒度疫情风险评估模型

邱鸣杰, 谭智一, 鲍秉坤

(南京邮电大学通信与信息工程学院, 南京 210000)

摘要:针对疫情期间实施精准疫情防控的需求,达到兼顾疫情防护和社会经济发展的目的,实现了细粒度条件下的疫情风险评估深度学习模型,用于探索城市场景下街道、社区级别的疫情风险实时预测。提出一种城市区域功能感知的细粒度疫情风险评估模型,主要包括区域功能特征学习和区域关联挖掘 2 个模块。首先,区域功能特征学习模块将 POI(point of interest)分布和疫情风险特征相融合,用于表达不同城市功能对于风险传播的影响,以引入风险扩散的先验知识。然后,区域关联挖掘模块将低层的城市网格按照功能分类映射至更高层的功能区域空间,并直接捕捉功能区域间的空间依赖,避免了低效的堆叠卷积计算。所提模型在 2022 年南京市新冠疫情数据集上进行了验证,相比于传统的时间序列和时空序列预测方法,所提模型相较于现有方法,在平均绝对百分比误差指标上降低了 8%~23%,在均方根误差指标上降低了 0.3~1.2。同时所提模型的可学习参数量大幅降低,模型计算效率远优于现有方法。

关键词:疫情风险评估;细粒度;城市区域功能;图卷积网络;城市计算

中图分类号:TP183

文献标志码:A

文章编号:2095-2783(2023)11-1165-07

开放科学(资源服务)标识码(OSID):



Fine grained epidemic risk assessment model of urban region functionality aware

QIU Mingjie, TAN Zhiyi, BAO Bingkun

(School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210000, China)

Abstract: In response to the need for precise epidemic prevention and control during the pandemic, a deep learning model for fine grained epidemic risk assessment has been developed, aiming to achieve a balance between epidemic prevention and socio-economic development. This model focused on exploring real-time prediction of epidemic risks at the street and community level in urban scenarios. It contained two main modules: region functional feature learning and region correlation mining. Firstly, the region functional feature learning module integrated the distribution of points of interest (POI) with epidemic risk to express the impact of different urban functions. Secondly, the region correlation mining module mapped lower-level grids into higher-level functional regions based on classifications. The proposed model was validated on the Nanjing 2022 COVID-19 dataset. Compared with traditional time series methods and spatio-temporal sequence methods, the proposed method achieved a reduction of 8%-23% in mean absolute percentage error and a reduction of 0.3-1.2 in root mean square error. Additionally, the proposed model also significantly reduced the learnable parameter number.

Keywords: epidemic risk assessment; fine grained; urban region functionality; graph convolution network; urban computing

疫情风险评估任务(epidemic risk assessment)是根据地区各类时空特征,对当地疫情风险进行评估。在疫情流行期间,风险评估模型能够提前对涉疫地区发出风险警告,因此受到了众多的关注。然而,现有的疫情风险评估方法主要聚焦于省、市级别的粗粒度建模^[1-4],缺少街道、社区级别的细粒度建模。基于细粒度建模的疫情风险评估有着同样重要的应用价值。例如,在 COVID-19 新型冠状病毒流行期间,若能在病例大规模爆发之前准确地预测城市中各个街道的潜在传播风险,将有助于实现疫情的灵活防控,保证经济、社会的正常运行。

尽管先前的工作在粗粒度上已经获得了初步的成功,但由于以下挑战,这些工作的架构并不能直接

迁移至细粒度上:

1) 长距离空间依赖难以捕捉。从粗粒度转换至细粒度的过程中,由于分辨率的提高,城市网格数量同样也会大大增加。由于城市中发达的交通体系,人们可以轻易地通过公共交通或私家车在短时间内到达城市中的任何地方,形成了数量众多的长距离跨网格疫情传播路径。因此,在细粒度条件下,长距离空间依赖在疫情风险评估中起到了重要作用。在经典时空预测任务^[4-7]中,计算长距离空间依赖的普遍做法是堆叠金字塔卷积层,通过每一层中的短距离卷积,一步一步地提高感受野。然而,这样的迭代计算效率低下,同时会引入大量参数,导致模型难以优化的问题。另一些方法使用空洞卷积代替普通卷

收稿日期:2023-06-11

基金项目:国家科技创新 2030-“新一代人工智能”重大项目(2020AAA0106200);国家自然科学基金资助项目(61936005);江苏省自然科学基金资助项目(SBK2021043792)

第一作者:邱鸣杰(1999—),男,博士研究生,主要研究方向为数据挖掘、时空序列预测

通信作者:鲍秉坤,教授,主要研究方向为多媒体计算、社交多媒体、计算机视觉、人工智能,bingkunbao@njupt.edu.cn

积,缓解了参数量爆炸的问题^[8],但并未从根源上改变运算效率不佳的金字塔型卷积结构。因此,在捕捉长距离空间依赖上,现有的方法并未给出很好的解决思路。

2) 地区功能对于疫情的影响被忽视。病毒的载体是人群,而在现实生活中,人群的流动在很大程度上受到地区功能的影响。市中心区域由于商业区众多,人群来往接触密切,导致该区域更容易爆发疫情。因此,地区的功能中包含了有关人群流动和病毒传播的潜在规律,为风险评估任务提供了重要的先验知识。然而在其他时空序列预测任务中,区域功能仅作为一个外部特征来控制区域流量的权重^[9-10],并没有显式建模地区特征和预测目标的对应关系。因此,学习不同功能区域对城市疫情传播的影响仍是一项有挑战性的工作。

为了解决上述问题,本文提出一种城市区域功能感知的细粒度疫情风险评估模型,以实现街道、社区级别的疫情风险实时预测。模型大致分为 2 个模块:首先,针对城市功能特征,设计新的区域功能特征学习模块,将时变的风险特征和时不变的 POI 分布等进行融合,学习城市功能对疫情传播的影响。然后,提出新的区域关联挖掘模块来捕获长距离空间依赖。该模块将所有低层次网格按照功能分类映射至更高层次的功能区域空间,并在区域空间中直接捕捉长距离空间依赖。相比于现有方法,本文模型的可训练参数量大大减小,同时较好地保持了功能区域的空间结构,更适合细粒度城市市场下的大规模数据集。

综上所述,本文的主要贡献如下:

1) 所提出的方法综合考虑了风险和区域功能特征,实现了在街道、社区级别上的细粒度疫情风险评估模型;

2) 提出了新的区域关联挖掘模块,将低层次城市网格映射至更高层次区域空间进行空间依赖计算,在保持区域空间结构的同时大幅降低参数计算量;

3) 提出了新的区域功能特征学习模块,融合时变的风险特征和时不变的 POI 分布,用于学习不同区域功能对于疫情传播的影响;

4) 所提出的模型在 2022 年南京市新冠疫情数据集上进行了验证,相比于现有的时空预测方法,本文提出的模型平均百分比误差降低了 8%~23%,均方根误差降低了 0.3~1.2。

1 相关工作

1.1 疫情风险评估

为了对抗新型冠状病毒肺炎疫情,研究人员从多个角度进行了大量研究。从任务的角度出发,与本文最相关的研究是疫情风险评估任务。目前,在粗粒度疫情风险评估任务上的相关工作已经取得了初步的成功,例如:Ye 等^[10]将美国的州、市视作不同

层次的节点,将地区间的空间依赖视作边,构建了一张层次元路径图进行风险信息的聚合;Xiao 等^[11]将对抗学习引入风险评估,通过对不随城市而变化的疫情特征进行学习,实现了风险评估知识的跨城市迁移;Pluchino 等^[3]采用数据驱动的方法,从疫情危害、暴露地区和易感居民数量 3 个方面对意大利的各大区进行了风险等级拟合。尽管这些工作都取得了初步的成功,但鲜有工作将疫情风险建模的思路迁移至城市计算领域,导致了细粒度疫情风险评估工作的缺失。

1.2 时间序列与时空序列预测

从方法角度出发,与本文最相关的工作是时间序列和时空序列预测任务。与细粒度疫情风险评估类似,其他以城市计算为背景的时间、时空序列预测任务(如交通预测、城市功能区识别等)同样通过城市各区域的时空序列进行分析^[11-14]。然而在细粒度建模方面,现有工作大多数聚焦于通过扩大卷积网络的感受野来处理高分辨率带来的巨量网格数据,常见的方法有堆叠卷积层^[4-5]、使用空洞卷积^[8]、捕捉全局依赖^[10]等,这些做法虽然取得了一定的效果,但在计算效率上有缺陷。因此,对于疫情风险评估这类时效性要求较高的任务,直接迁移现有的时间、时空序列预测模型并不可行。基于计算效率的考虑,本文没有采用传统的时空卷积方法,而是通过将网格映射至区域的方法,显著降低了模型参数量,提高了计算效率。

2 定义

2.1 城市网格

如图 1 所示,所提模型对目标城市地图按照一定大小的网格进行切分,从而获得城市网格集合。城市网格是构成城市结构的基本单位,因此一个大型规模城市的网格数量非常庞大,往往超过 10 万。以分割长度数量(H)和分割宽度数量(W)为例,模型通过均匀分割,从而产生 $N = HW$ 个网格。通过增大网格分辨率 H 和 W 可以获得更加细粒度的城市网格。

2.2 功能区域

功能区域是城市中具有不同土地功能的区域,例如商业区、办公区、住宅区等。功能区域从整体描述了地区的总体功能职责,相比于城市网格,功能区域提供了一种更自然、更高层语义的城市分割方法。另外,由于功能区域包含多个城市网格,因此空间形状并不固定,且功能区域的数量(M)将远小于城市网格数(N)。通过映射矩阵(P)可以将网格集合从网格空间映射至区域空间。

2.3 任务定义

针对细粒度疫情风险评估任务,作如下定义,已知细粒度疫情风险历史观测序列(X_i),记作 $\{X_i | i = 1, 2, \dots, t-1\}$,预测任务的目标为预测未来时间步的风险(X_t)。

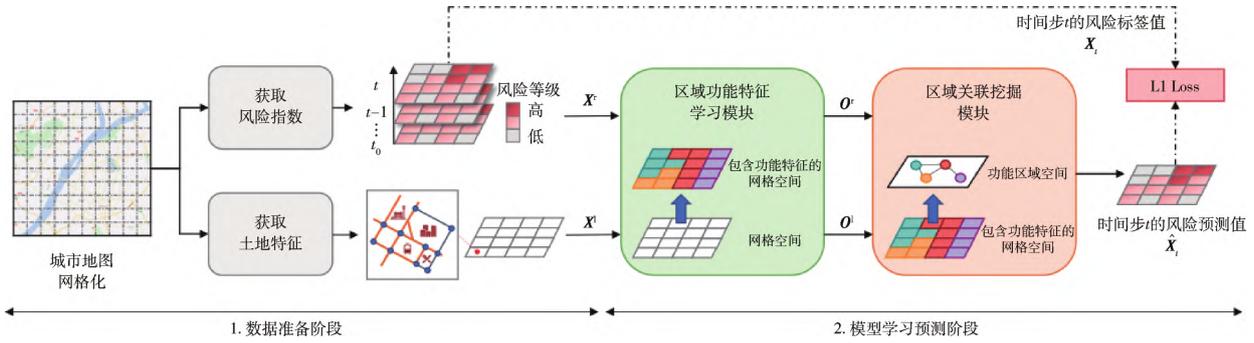


图 1 本文模型的整体框架

Fig. 1 Overall architecture of the model proposed herein

3 方法

本文方法主要包括 2 个阶段:数据准备阶段和模型学习预测阶段。数据准备阶段主要包括城市网格的生成和网格数据的获取。模型学习和预测阶段则包含地区功能特征学习模块和区域关联挖掘模块,功能特征学习模块以数据准备阶段输出的风险时间序列(X_t)和土地特征(e)为输入,学习网格的高层功能表征;区域关联挖掘模块以包含高层功能特征的城市网格为输入,将网格空间映射至区域空间,并在区域空间中进行空间依赖的计算,最后通过卷积神经网络聚合领域的风险信息,并进行疫情风险的评估。方法部分使用的相关数据与符号定义见表 1。

表 1 参数符号定义

Table 1 Symbol definition of parameters

符号	符号定义
N	城市网格总数
M	城市功能区域总数
δ_i	编号为 i 的网格与城市几何中心的距离
n_i	编号为 i 的网格中 POI 的总数
s_t^i	编号为 i 的网格在时间步 t 的新增感染人数
c_t	在时间步 t 的城市新增感染总数
X_t^i	在时间步 t , 网格 i 的风险指数标签值
\hat{X}_t^i	在时间步 t , 网格 i 的风险指数预测值
G, I, C	网格内地区活跃因素、易感因素和城市因素
O^f	区域功能特征学习模块输出
O^r	区域关联挖掘模块输出
Y, \hat{Y}	模型风险标签值、预测值

3.1 数据准备阶段

在数据准备阶段,本文采用在城市交通预测中常用的方法^[15-16],将城市地图依照固定大小的网格进行划分,同时对网格的长宽进行限制,以保证细粒度的要求。在本文实践中,采用分辨率为 $500\text{ m} \times 500\text{ m}$ 的网格对南京地区进行了初步的划分和清洗,总共筛选出约 5 万个有效的城市网格。

在获得了城市网格之后,模型依照网格的位置信息对网格内的风险和土地特征进行自动获取。首

先针对网格内的疫情风险,本文通过一系列的时空特征构造了新的风险时间序列。对于 t 时刻的网格 $i \in (1, N)$, 风险指数由地区活跃因素 (G)、易感因素 (I) 和城市因素 (C) 3 个部分组成,具体计算方式如下:

$$\begin{aligned}
 X &= G \times I + C, \\
 G &= \frac{1}{\sqrt{\delta}} + \alpha \times |\lg(n)|, \\
 I &= \sum_{t=7}^t s_t^i + \beta, \\
 C &= \frac{\exp(s_t^i)}{\sum_{i \in (1, N)} \exp(s_t^i)}.
 \end{aligned} \tag{1}$$

式中: α, β 为可训练的参数;活跃因素 G 代表该地区在最近的活跃程度,由地理位置因子和 POI 数量因子控制;易感因素 I 代表该网格在最近 7 d 内的疫情状况,由 7 d 内的疫情新增感染数控制;城市因素 C 则反映了城市总体的疫情状况,由所有网格的疫情状况控制。

最后生成的风险时空序列为 $X^r \in \mathbf{R}^{H \times W \times I}$ 。本文构造的风险指数综合考虑了地理、活跃、疫情因素,能够全面表示细粒度地区的疫情传播风险;同时,该风险指数经过归一化处理后在数值上连续,更有利于模型训练的稳定。风险指数的构造能更加全面地反映疫情发展趋势,且有利于训练的稳定。

同样,因为城市网格中 POI 的类别及密度表明该网格的土地功能和人口流动模式,对于学习城市中潜在的人口流动规律有所帮助,所以在该阶段中会获取每个网格中包含的 POI 列表,作为土地特征输出。

3.2 模型学习预测阶段

在模型学习预测阶段,模型将数据准备阶段的数据作为输入,输出风险指数的预测值。本阶段按照从局部到全局的模式,将低层网格空间映射至高层区域空间,并在区域空间进行信息聚合与预测。

3.2.1 功能特征学习模块

由于粗粒度与细粒度的风险建模思路不同,在先前的疫情风险评估工作中,很少有工作将地区功能作为建模依据来学习城市潜在的人口流动规律。

而在交通预测等其他城市计算任务中,POI 仅作为评判地区交通权重的因素^[9],目前并未从 POI 特征出发探索区域功能联系的工作。因此,本文提出了全新的区域功能特征学习模块,根据网格内部存在的各类 POI,学习包含城市功能职责的高层语义,从而学习细粒度下城市风险传播模式。

区域功能特征学习模块的细节如图 2 所示。一般来说,城市区域的功能职责往往由域内的 POI 决

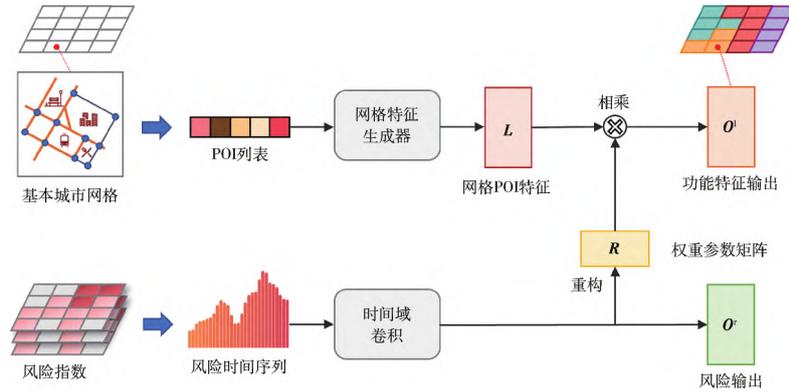


图 2 区域功能特征学习模块
Fig. 2 Region function learning module

本文中采用了长短时记忆网络(long short-term memory network, LSTM)作为时间卷积方法以提取时间域特征。同时,为了更好地发掘区域特征在疫情风险传播时的作用,模块由时间域表征(O^r)通过维度重构额外提取了一个权重参数矩阵(R)帮助网络功能特征的学习。

同时,针对网格中包含的 POI,模块首先按照类别对这些 POI 进行分类并放入列表。然后,网格特征生成器通过一个映射矩阵对列表进行映射生成网格的 POI 特征,其中包含各类 POI 的密度、POI 的热门度等。接着,根据之前得到的权重参数矩阵(R)对 POI 特征进行调整,如式(3)所示。

$$O^f = f_1(L, R) \in \mathbf{R}^{d \times N}. \quad (3)$$

在实践中,本文将融合函数 f_1 设置为矩阵乘法,从而避免引入过量参数。在此过程中, L 可以被视作网格中的 POI 表征,而 R 可以视作控制表征的参数矩阵。这样的做法可以使模型学习到各个区域特征在疫情风险传播时的作用,便于之后的空间映射和关联捕捉。

通过输入基本的城市网格,本模块通过提取网格内的风险和 POI 特征,输出了融合高层功能特征的城市网格,相比于其他将城市 POI 特征引入城市计算模型思路^[8],本文方法更加轻量化,更适合在细粒度条件下进行建模。

3.2.2 区域关联挖掘模块

计算区域之间的风险关联是建立疫情风险评估模型的重要步骤。考虑到疫情风险评估任务对时效性的要求,在提取了城市网格的高层功能特征之后,

定,例如,在住宅区域中居民住宅相关 POI 的占比会更高,商业区中商场和娱乐场所相关 POI 的占比更高。按照这样的建模思路,区域功能特征学习模块综合利用了网格内的风险时间序列和 POI 列表。首先,针对数据准备阶段输出的风险指数,模块按照网格降维为时间序列,并通过时间卷积获取时间表征,如式(2)所示。

$$O^r = \text{MaxPool}(\text{Conv}_t(X_{t-l_b:t})) \in \mathbf{R}^{h \times N}. \quad (2)$$

本文进一步提出了全新的区域关联挖掘模块,将网格映射至区域空间,并在区域空间上进行信息聚合后,再映射回网格空间进行风险预测。考虑到本任务城市网格的数量很大,采用这样的做法避免了对每个城市网格进行空间依赖计算,显著降低了计算复杂度;同时,映射至区域空间的的做法保持了功能区域的空间结构,保证了城市拓扑的完整性,有助于学习城市内流动规律。

图 3 展示了区域关联挖掘模块的相关细节。首先,该模块以功能特征模块的输出(O^r 和 O^f)为输入,通过线性变换后转换为输入网格特征(X^G)。接着,本文以 k -Means 算法为基础设计了新的聚类算法,用于对网格内的功能和地理特征进行聚类,基于城市网格特征的聚类目标如式(4)所示。

$$J = \sum_{i=1}^M \sum_{j=1}^N \tau_{ij} \nu(x_j, \mu_i), \tau_{ij} = \begin{cases} 1, & x_j \in i; \\ 0, & x_j \notin i; \end{cases} \quad (4)$$

$$\nu(x_j, \mu_i) = \|x_j - \mu_i\|^2,$$

$$[\text{Tran}]_{ij} = \text{Argmin}(J) \in \mathbf{R}^{N \times M}.$$

该聚类算法的目标为在地理位置 μ 的约束下,最小化函数 J ,从而计算出在区域功能特征上最相似的节点簇。在计算节点簇之后,生成对应的映射矩阵 $[\text{Tran}]_{ij} \in \mathbf{R}^{N \times M}$,实现从网格区域至区域空间的映射。

在获得了空间映射矩阵之后,区域关联挖掘模块在功能区域空间进行依赖捕捉和领域聚合等操作。区域空间图生成模块首先根据映射矩阵生成对应的节点表征和区域邻接矩阵,其中节点表征由同属一个区域的网格表征的平均值计算得到,区域邻

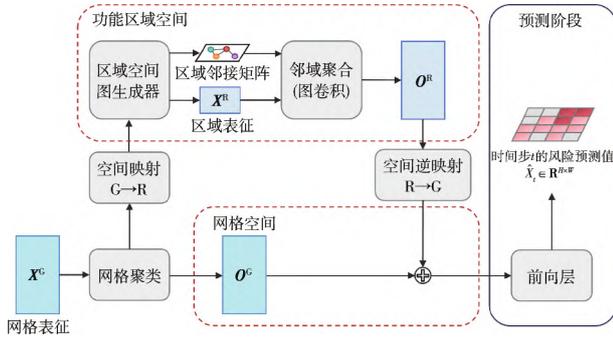


图 3 区域关联挖掘模块细节

Fig. 3 Details of the regional relation mining module

接矩阵则由距离和区域相关性这 2 个主要变量进行控制,如式(5)所示。

$$\lambda_{i,j} = f(\theta_s^T \cdot \text{Concat}(X_i^R, X_j^R)) + \frac{1}{\sqrt{\zeta_{i,j}}}. \quad (5)$$

式中: $i, j \in (0, M)$ 为功能区域; f 为激活函数; θ_s 为可训练的参数矩阵; $\zeta_{i,j}$ 为 2 个区域几何中心的距离。

这样构建邻接矩阵首先能充分考虑不同功能区在空间距离上的联系,使得 2 个距离较近的区域在疫情上的空间依赖也更加密切;其次,这样的构建方法同样考虑了不同功能区之间潜在的关联,例如,办公区往往与住宅区相关联、商业区与公共交通区域相关联等。针对这些依赖,本模块采用了可训练的参数进行学习。

由于区域邻接矩阵和区域表征形成了自然的非欧几里得空间,受到图卷积神经网络^[16]的启发,本模块采用 GCN(graph convolution network)进行消息的领域聚合,卷积过程如式(6)所示。

$$O^R = \text{Aggregate}(\{\delta_{j,i} \cdot M_j^O, \forall j \in \mathcal{N}(i)\}). \quad (6)$$

式中: $\mathcal{N}(i)$ 为与节点 i 相邻的节点集合; $\delta_{j,i}$ 为 2 个节点之间边的权重。

模块按照这样的卷积过程实现邻域信息的聚合,之后通过对空间映射矩阵的逆变换得到 $[\text{Tran}]^{-1}$,将区域空间重新映射回网格空间,最后与网格空间表征拼接后进行预测。

本模型基于细粒度的疫情数据,采用了端到端的训练方式,因此可以迁移到其他不同的城市。模型采用了均方根误差函数进行训练:

$$\mathcal{L}^Y = \hat{Y} - Y. \quad (7)$$

4 实验结果与分析

4.1 数据集构造

由于现阶段没有公开的细粒度疫情风险数据集,因此本文构造了新的以南京市 2022 年疫情为背景的风险数据集。本文数据集的来源主要为南京市政府在 2022 年度与疫情相关的通告,包括划定风险区、新增感染与无症状感染、疫情相关人员流调等。本文通过提取关键字等方法将非结构化文本整理为

时间序列,为数据准备阶段提供输入。同时,针对土地特征,本文收集了南京市总计 15 万个活跃 POI 的相关数据,并获取每个 POI 的经纬度位置标签,便于城市网格的分割。

经整理后的风险时间序列总共包含 269 个时间步,并将训练集、测试集和验证集按照 0.7:0.2:0.1 的比例进行分割。POI 数据集已做脱敏处理,其中的隐私和敏感信息均已消去。

4.2 基线方法

由于本文是第一个在城市计算背景下探讨疫情风险评估的工作,所以并没有能直接与本文模型比较的基线方法。因此,本文选取了经典的时间序列预测与时空序列预测模型,包括时间序列模型 ARIMA、Nbeats^[17]、LSTM^[18]、GRU^[19],时空序列模型 ConvLSTM^[20]。基线模型上的超参数设置与本文模型基本保持一致。

4.3 评价方法、超参数设置与实验环境

本文主要从 2 个方面进行模型的评价。第一,从计算效率角度,采用模型可训练参数量 (# Trainable Param) 作为指标。一个模型所使用的参数量越大,意味着计算复杂度越高,模型训练及预测所需的时间也会越长;在疫情风险评估这类时效性要求较高的任务中,往往要求压缩模型的可训练参数量以达到节省硬件资源和减少计算耗时的效果,因此可训练参数量是一项衡量模型计算效率的重要指标。第二,从预测准确率角度,本文采用平均绝对百分比误差 (mean absolute percent error, MAPE) 和均方根误差 (root mean squared error, RMSE) 作为评价指标。在时间与时空序列预测中,计算误差率是常见的衡量模型预测准确率的做法,因此本文选取了最常用的 MAPE 和 RMSE 用作对比,其中 MAPE 衡量了预测风险值与实际风险值偏离的百分比, RMSE 衡量了预测风险值与实际风险值偏离的绝对大小情况。2 个指标的数值均是越小越好,计算公式如下:

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|,$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}.$$

关于模型使用的优化和超参数设置,本文设置了 Adam 优化器进行参数优化。本文所有模型的批训练大小为 4,学习率固定为 0.001,最大迭代次数设置为 1 万次,并启用提早停止的训练策略。对于每个模型,本文均分别进行了 5 次实验并求平均误差,以减小随机初始化带来的影响。

在本文所涉及的实验均在计算服务器上进行,其操作系统版本为 Ubuntu 22.04,中央处理器为 Intel Xeon Platinum 8369B,图形处理器为 NVIDIA Titan V 12 GB,内存为 256 GB。在软件环境方面,深度学习框

架版本为 Pytorch 1.7.1, CUDA 版本为 10.1。

4.4 实验结果与分析

为了全面对比所提模型与其他模型在细粒度风险预测任务上的性能和效率,本文从模型参数量和准确率 2 个方面进行了综合评估,评估结果见表 2。总体上来看,本文模型在参数量和准确率上均取得了较大的提升。例如,相比于在性能上表现最好的 ConvLSTM 模型,本文模型在每次实验中都降低了 8% 的平均绝对百分比误差和 0.2~0.3 的均方根误差,但在参数使用量上却仅为前者的 0.02%;而相比于模型参数量最少的 ARIMA、Nbeats 等模型,本文模型在参数量上差异不大,但误差率获得了最高 20% 的提升。

表 2 预测性能对比结果
Table 2 Comparison of predicting performance

模型	参数量/ 10^6	误差		
		MAPE/%	RMSE	MAPE 差值/%
ARIMA	<0.01	50.6	2.095	23.4
Nbeats	0.05	46.5	1.142	19.3
LSTM	0.42	42.4	1.545	15.2
GRU	0.34	41.0	1.604	13.8
ConvLSTM	3.69	35.2	1.031	8.0
本文模型	0.07	27.2	0.834	

注: MAPE 差值以本文模型为基线;加粗数字为最优结果。

经典模型在本任务上的失效主要来源于 2 个方面。第一,缺少针对疫情风险特征的利用。在时间序列预测模型中,模型仅关注目标时间序列本身的趋势、周期、抖动等统计学上的性质,使得模型的参数量较小;然而,疫情风险是由多个因素共同作用而形成的复杂多变量时间序列,不重视风险相关的特征而强行提取统计性质的做法不会奏效,从而导致了时间序列模型较高的误差率。第二,缺少有效的在细粒度条件下的建模方式。时空序列模型在一定程度上考虑了风险相关的特征,例如空间相邻特征等,因此在误差率上有所提升;但是由于缺少有效的捕捉长距离空间依赖的方法,这些模型只能通过堆叠卷积层来扩大卷积感受野,在细粒度的条件下导致了参数量的爆炸,更引发了模型难以训练和计算效率低下等一系列问题。

本文模型在这 2 个方面做出了改进。首先,针对长距离空间依赖捕捉的挑战,本文模型采用了将城市网格从低层空间映射至区域高层空间的办法,在功能区的层面捕捉空间依赖,由于区域数量(M)远远小于网格数量(N),这样的做法事实上避免了大量的计算,因此大大降低了计算复杂度。同时,本文提出使用功能区域特征来引导捕捉空间关联和学习风险特征的方法更好地利用了疫情风险的相关特征,从更深层次学习城市内在的流动规律,因而能更好地模拟风险在城市中的传播,因此在准确率指标上取得了领先。

4.5 消融实验

为了进一步了解所提模型各个部分在预测时的作用,本文设计了 2 个消融实验来验证区域功能特征学习模块和区域关联挖掘模块在疫情风险评估时的作用,从而验证其有效性。

4.5.1 消融实验 1: 区域功能特征学习模块的作用

本节实验删除了功能特征学习模块,用来阻止模型学习到与地区土地特征相关的知识。同时,为了让模型能正常运行,使用了一个简单的循环神经网络来提取风险特征,作为区域关联挖掘模块的输入。在本节消融实验中,模型无法正常获取到区域所在的功能特征,只能依靠风险特征进行推理。

图 4 展示了去除该模块的可视化过程,图中每个点都代表一个城市网格。在功能特征学习模块的帮助下,模型能较好地学习到功能区域的空间结构,并将同类的网格映射至一个区域内。而去除功能特征学习模块的模型无法正常地建立形状规则的区域,致使无法正常预测。

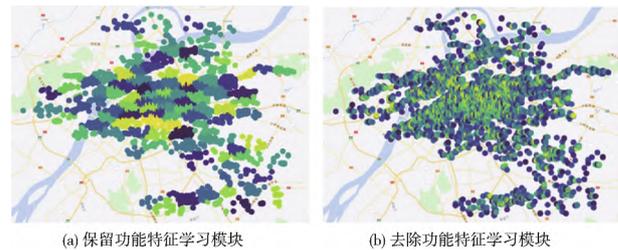


图 4 消融实验 1 中 2 种情况下的空间映射可视化

Fig. 4 Visualization of spatial mapping in two scenarios in ablation experiment 1

4.5.2 消融实验 2: 区域关联挖掘模块的作用

本节实验修改了区域关联挖掘模块,用于探究区域关联在建模疫情风险传播时的作用。在这种情况下,模型实际上退化为了普通的时空序列预测模型,即城市网格之间两两进行关联的计算。由于在细粒度条件下,计算两两之间的空间关联开销巨大,本节实验限制了每个网格的最大关联个数为 50,即每个网格最多与 50 个其他网格进行领域聚合操作。

表 3 展示了消融实验 2 的实验结果,在阻止模型学习高层的区域语义之后,模型仅能从网格空间进行推理,这极大地增大了模型的参数量,加重了计算

表 3 消融实验结果对比
Table 3 comparison of the ablation study result from the proposed model

方法	参数量/ 10^6	误差	
		MAPE/%	RMSE
本文模型 (无功能特征学习模块)	0.04	53.4	1.995
本文模型 (无区域关联挖掘模块)	20.00	36.5	1.792
本文模型	0.07	27.2	0.564

负担。同时,由于缺失了关键的区域功能特征,预测准确率也相应下降。实验证明,区域关联模块所提供的高层语义对于准确预测起到了重要作用。

5 结 语

本文研究了在细粒度条件下建模疫情风险传播的问题。首先,不同于现有的时空序列预测模型堆叠金字塔式卷积层来扩大感受野的做法,本文针对细粒度条件做了改进,通过将城市网格映射至更高语义空间进行计算,从而大大降低了模型使用参数量。其次,针对城市区域的功能特征,本文通过轻量的区域功能特征学习模块,使模型学习到城市内潜在的人口流动规律。通过广泛的对比实验,在南京市 2022 年疫情数据集上的实验结果验证了本文模型优越的预测准确率和计算效率。

(由于印刷关系,查阅本文电子版请登录:<http://www.paper.edu.cn/journal/zgkjlw.shtml>)

[参考文献] (References)

- [1] XIAO C X, ZHOU J B, HUANG J Z, et al. C-Watcher: a framework for early detection of high-risk neighborhoods ahead of COVID-19 outbreak [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(6): 4892-4900.
- [2] FU Z X, WU Y, ZHANG H L, et al. Be aware of the hot zone: a warning system of hazard area prediction to intervene novel coronavirus COVID-19 outbreak [C]// Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2020: 2241-2250.
- [3] PLUCHINO A, BIONDO A E, GIUFFRIDA N, et al. A novel methodology for epidemic risk assessment of COVID-19 outbreak [J]. Scientific Reports, 2021, 11: 5304.
- [4] GUO S N, LIN Y F, LI S J, et al. Deep spatial-temporal 3D convolutional neural networks for traffic data forecasting [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(10): 3913-3926.
- [5] ZHANG J B, ZHENG Y, QI D K. Deep spatio-temporal residual networks for citywide crowd flows prediction [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2017, 31(1): 2145-2160.
- [6] ZHANG J B, ZHENG Y, QI D K, et al. DNN-based prediction model for spatio-temporal data [C]// Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. New York: ACM, 2016: 1-4.
- [7] PAN Z Y, WANG Z Y, WANG W F, et al. Matrix factorization for spatio-temporal neural networks with applications to urban flow prediction [C]// Proceedings of the 28th ACM International Conference on Information and Knowledge Management. New York: ACM, 2019: 2683-2691.
- [8] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [EB/OL]. [2023-05-21]. <https://arxiv.org/abs/1511.07122>.
- [9] LIN Z Q, FENG J, LU Z Y, et al. DeepSTN+: context-aware spatial-temporal neural network for crowd flow prediction in metropolis [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 1020-1027.
- [10] YE Y F, FAN Y J, HOU S F, et al. Community mitigation: a data-driven system for COVID-19 risk assessment in a hierarchical manner [C]// Proceedings of the 29th ACM International Conference on Information & Knowledge Management. New York: ACM, 2020: 2909-2916.
- [11] LIU L B, QIU Z L, LI G B, et al. Contextualized spatial-temporal network for taxi origin-destination demand prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(10): 3875-3887.
- [12] LIU L B, ZHANG R M, PENG J F, et al. Attentive crowd flow machines [C]// Proceedings of the 26th ACM International Conference on Multimedia. New York: ACM, 2018: 1553-1561.
- [13] LIU L B, ZHEN J J, LI G B, et al. Dynamic spatial-temporal representation learning for traffic flow prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(11): 7169-7183.
- [14] PAN Z Y, LIANG Y X, WANG W F, et al. Urban traffic prediction from spatio-temporal data using deep meta learning [C]// Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2019: 1720-1730.
- [15] YAO H X, TANG X F, WEI H, et al. Revisiting spatial-temporal similarity: a deep learning framework for traffic prediction [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 5668-5675.
- [16] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks [EB/OL]. [2023-05-21]. <https://arxiv.org/abs/1609.02907>.
- [17] ORESHKIN B N, CARPO D, CHAPADOS N, et al. N-BEATS: neural basis expansion analysis for interpretable time series forecasting [EB/OL]. [2023-05-21]. <https://arxiv.org/abs/1905.10437>.
- [18] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780.
- [19] DEY R, SALEM F M. Gate-variants of gated recurrent unit (GRU) neural networks [C]// 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS). New York: IEEE, 2017: 1597-1600.
- [20] SHI X, CHEN Z, WANG H, et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting [C]// Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2015: 802-810.