重庆师范大学

学术型硕士研究生文献综述报告

开 题 题 目 图神经网络在车辆轨迹中的应用

所 在 学 院 计算机信息与科学学院

专 业 计算机技术

研 究 方 向 图神经网络

研 究 生 李泽冯

指 导 教 师 杨德刚

入 学 年 月 202109

2023年2月

《图神经网络在车辆轨迹分类的应用》研究综述

李泽冯

重庆师范大学 计算机信息与科学学院

摘要：（中文）

关键词：（中文，5个）

**A Survey of XXXXXX（Title）**

**XXX（Name）**

**XXXXXX(Name of school)**

**Abstract:（English）**.

**Keywords:（English keywords）**

1）论文选题的背景、目的及意义（不少于 1000 字）；

1. **引言（背景、目的及意义）**

地理信息系统与科学的不断发展使得时间地理学从数据采集、存储、 分析到可视化等整个过程都获得了极大的便利［２－５］，其强大的可视化和空间分析功能增强了个体时空轨迹的直观显示及其隐含模式的识别能 力［６－７］。伴随着神经网络与高性能计算技术的不断发展，交叉学科研究的推进（大数据驱动的人类移动模式和模型研究），越来越多的地理信息已经被计算机技术所处理，给地理的学科的发展提供了不同的方向。其中轨迹数据隐含了丰富的城市信息，往往规模巨大且来源广泛。轨迹数据主要包括人类活动轨迹数据、 交通轨迹数据、动物活动轨迹数据和自然现象轨迹数据[1]。这些轨迹数据的时空尺度与粒度各有不同，分别适用于不同研究。其中人类活动轨迹和交通轨迹数据是与人类日常生活联系最紧密的两类轨迹数据，因此有大量的研究围绕着他们。

人类活动轨迹数据，包括微博定位签到数据、照片数据、手机基站定位数据、信用卡消费数据、手机定位数据等，按照时间顺序跟踪这些位置记录就形成关联人在一段时间内的移动轨迹。这类轨迹数据精度不高，通常在200ｍ左右［２］，而且大多涉及个人隐私，并没有得到广泛的应用，一般在研究中作为辅助数据用于丰富语义信息。 交通轨迹数据包括两类：；一类是通过车载GNSS设备采样得到的交通工具移动轨迹，如公交车轨迹数据、出租车轨迹数据、船只轨迹数据等。另一类是通过固定设备采集得到的数据，如城市道路卡口照片、视频监控数据、地铁刷卡数据，高速公路出入站口数据等，城市道路卡口照片数据和视频监控数据太复杂，涉及的信息太多不便于有用信息的提取。高速公路出入站口的数据，数据信息简单明了，能够清楚的表明车辆在那些地区进行流动，也便于实验数据的处理。

高速公路出入站口数据可以反映在特定的时间节点（春节、劳动节、国庆节等）上，人流量的流向。把此类数据处理好，不仅可以推动地区展示自己、帮助游客更好的规划自己的旅途，还可以为旅游监测与分析领域的研究提供工作基础和方法指导。基于大数据和神经网络研究个体或群体轨迹，发现活动中蕴含的空间 认知规律及空间行为和交互模式，建立以人为本的地理信息服务，进而支持个体或群体时空行为决策，已成为地理信息科学研究的前沿问题。此类研究也将为联系“计量”和“行为”两个地理学思想流派以及重新审视人“上—地”关系提供一个全新的视角，从而推动理论地理学的进一步发展。

轨迹聚类是对高速公路出入站口数据进行聚类，得到的结果是具有相似形态且聚集的轨迹簇。例 如，对出发地和目的地之间的轨迹进行聚类， 得到最具有最佳旅行的最优路径［87］。或者，跟踪特定区域间的出行轨迹，从而识别旅客流向规律[88]此外，文献［26］基于相似轨迹聚类的思想提出拥堵同伴的概念和发现算法，筛选出可能发生拥堵的浮动车数据，进而对拥堵区域的变化趋势进行预测。

1）论文选题的背景、目的及意义（不少于 1000 字）；

2）国内外相关研究现状综述（不少于 3000 字）；

3）开题报告正文（研究型课题一般包括研究内容、研究方案、 技术路线、研究方法、拟解决的关键问题、实验条件、特色与创新、 完成预研工作和实验结果等；其他类型课题根据教指委要求撰写）（不 少于 5500 字）；

4）研究计划及预期研究成果（不少于 500 字）

2）国内外相关研究现状综述（不少于 3000 字）；

1. **国内外相关研究现状综述**
   1. 轨迹研究现状
   2. 图神经网络研究现状

最早关于图的概念的提出始于十八世纪柯尼斯堡七桥的论述中，而图论作为数学领域的重要分支，在电子计算发展之前就已经成为了学者的热点研究，像极图理论、拟阵理论、超图理论等在二十世纪中期得到了迅速发展。对于复杂丰富的数据中的实体与实体间的重要信息关系可以通过图的方式对其实现抽象且有效的数据表示。图型数据库成功的解决了传统的关系型数据库在大规模复杂、低结构化的数据中建模困难、计算效率低且不擅于关联查询等缺点。图计算(Graph Computing）的特点包含数据局部性差、规模巨大和高的访存计算比等，其算法主要有路径搜索算法、中心性算法、社群发现算法等三类且在社交网络分析和用户个性化推荐等领域应用广泛411。

与已经取得成熟研究的图计算有所不同,图神经网络(Graph Neural Network, GNN)于2005年 Gori 等人最先将神经网络模型应用到图领域，该方法主要研究的是图结构中节点与节点之间信息关系的传播以及对信息的聚合。之后Scarselli等人【17】于2009年在Gori等人【16】的研究基础上对GNN模型进行了进一步的研究，提出了使用监督学习的算法对GNN模型参数做评估计算。GNN概念首次被提出的时候并没有得到广大研究者的重视。随着Lecun等人严提出的卷积神经网络（CNN)被应用到图数据中并在计算机视觉、图像等方向上得到了显著的成就，使得深度学习迅速崛起。CNN主要作用于网络的局部连接和共享权重，善于对二维图像等具有规则网格的欧几里德结构数据（Euclidean Structure Date）的拓扑结构进行提取处理，但不能直接处理任意图结构数据或者不规则的非网络的卷积，因为该类图比如社交网络拓扑图中的每个节点拥有的邻居节点的个数各不相同。为了能够更好的将卷积扩展到任意图数据结构的处理当中，直接的思想是将CNN 进行泛化应用。主要困难在于如何在非欧式不规则的图结构中对局部的卷积运算和池化算子进行定义，相关的研究中主要分为基于频域和基于空间的两大图卷积类别。

基于频域的图卷积最早源于Shuman等人"基于代数和谱图理论对图信号问题的处理，定义的图谱域的不同方法。其思想是将GNN 的卷积层看作能够对噪声信号进行屏蔽的滤波器而后对获得的输入信号分类的方法。在实际应用中只能对无向的图结构且图边上没有信息的数据进行操作，其中图信号表示为x，滤波器表示为go，图信号卷积表示为g日\*X = Ug(A)UTx，将输入的图信号看作是对拉普拉斯矩阵的特征分解的过程，U表示特征矩阵，A表示对角矩阵。在频域的基础上以卷积定理为出发点对卷积层函数进行定义从而实现图卷积神经网络(GCN)是由 Bruna等人3在2013年首次提出，但该卷积的计算量相对较大。Henaff 等人""基于插值内核引入平滑系数对频谱网络进行了扩展的图卷积神经网络（GCN)，实现了谱滤波用短的局部交互表达远距离交互的空间局部化。在以上研究基础上学者相继提出了变种或者简化的图卷积模型，如 Defferrad等人R提出了 Cheb Net 频域卷积滤波器等对时空的复杂度实现了较大的降低。随后基于空间的图卷积被提出因为它的思想简单又直观引起了研究者的重视，更类似接近于传统的对图的卷积操作运算。主要思想是图结构中相邻节点之间的信息传播和信息聚合依据图中某一中心节点到它的相邻节点间的卷积计算。Scarselli等人在2009年提出了能够分别在结构图的节点与边上定义相关属性通过对已有神经网络的扩展使用等同图卷积结构的函数完成基于空间的图卷积神经网络，并可以直接作用于无环图、无向图和循环图等多数类型图。Atwood等人提出基于图结构的传播卷积神经网络，该模型通过卷积传播的形式对结构图中单个顶点不断往外扩散进行扫描完成特征更新。DCNN (DynamicConvolutional Neural Network)的权重参数与图节点的所在位置无关，主要取决于搜索状态的深度，能够实现按照图的节点、图的边或者图的结构三种分类情况。最新的图卷积神经网络的实现由使用子图进行训练而后组合的 Graph Sage，自适应的图卷积(Adaptive Graph Convolution Network，AGCN)t和基于重要性抽样的图卷积网络快速学习(Fast-GCN)等。以及Kapoor等人在基于空间的卷积方法中面对高阶的卷积效果不佳等问题进行了网络优化，提出 Mix Hop模型通过稀疏正则化来优化领域信息。

随着如何将深度学习更好地应用在图结构中的相关图神经网络研究的进一步发展，Brockschmidt 等人M!W提出的门图序列神经网络和通过半监督方式的图划分神经网络、Krleaa等人的层级模糊图神经网络进行图匹配、以及 Velikovi 等人R提出的图注意力网络等。虽然图神经网络的发展时间相对较短但取得了丰富的科研成果，并有研究者也相继对图神经网络进行了综述。比如，徐冰冰等人.（2017年）着重于图卷积神经网络进行展开，对相关背景、基于谱域的和空间的两大经典方法、如何利用图卷积神经网络对大规模和复杂信息进行建模的最新进展以及在用户推荐、交通预测等相关领域的应用。白铂等人”（2020年)在谱方法和空间方法的基础上提出了一种新的类别基于生成的方法，并对这三类模型进行了阐述。王健宗等人1(2020年)介绍了图技术领域的相关研究历史,对不同的图神经网络技术进行了分析对比以及各种图网络如何与深度学习进行结合的方法同时也对在文本、视频等领域具体的应用做了阐述。

Yunsheng Bai等人，将图形相似度计算转换为一个学习问题来解决传统计算GED时间复杂度高的问题，并提出了一种基于神经网络的方法SIMgnn

当前，深度学习在许多任务上（目标检测[1]，机器翻译[2]和语音识别[3]）的成功部分归因于快速发展的计算资源(如GPU，CPU)、大型训练数据集的可用性以及深度学习从欧几里得数据(如图像、文本和视频)中提取潜在表征的有效性。以图像数据为例，我们可以将图像表示为欧氏空间中的规则网格，卷积神经网络(CNN)能够利用图像数据[4]的移位不变性、局部连通性和组合性。因此，CNN可以提取出与整个数据集共享的局部有意义的特征，用于各种图像分析。

虽然深度学习有效地捕捉了欧几里得空间中的数据的隐藏模式，但越来越多的任务中的数据是以图数据的形式表示。例如，在电子商务中，基于图形的学习系统可以利用用户和产品之间的交互做出高度精准的推荐。在化学中，化合物的分子结构被建模为图数据，它们的生物活性需要在药物发现中被识别。在论文引用文献中，论文通过引用文献连接彼此，根据引用文献的之间的关系，我们可以把论文分为不同的类型。受深度学习技术的启发，Marco Gori等人在2005年提出图神经网络（Graph Neural Networks，GNNs）的概念，使深度学习能够在有图数据的相关场景中得到有效的应用。GNNs的应用领域十分广泛，包括计算机视觉、化学生物、推荐系统、交通网络以及自然语言处理等。

图神经网络是一种基于图数据进行特征转换的神经网络，在深度学习中有很多方法被用来提高分子性质预测的性能，我们将车辆轨迹信息转换为图数据表示，

3）开题报告正文（研究型课题一般包括研究内容、研究方案、 技术路线、研究方法、拟解决的关键问题、实验条件、特色与创新、 完成预研工作和实验结果等；其他类型课题根据教指委要求撰写）（不 少于 5500 字）；

1. **正文**
   1. 研究内容

研究2012、2014、2016三年国庆七天重庆市高速公路收费站关于四川、湖北、广东、贵州车辆的登记信息，根据每个车辆登记信息生成车辆轨迹。

* 1. 研究方案
  2. 技术路线
  3. 研究方法
  4. 拟解决的关键问题
  5. 实验条件
  6. 特色与创新
  7. 完成预研工作和实验结果

4）研究计划及预期研究成果（不少于 500 字）

1. **研究计划与预期研究成果**

**2.1 自然图像描述生成**

自然图像或视频描述生成（Image Captioning）作为计算机视觉与自然语言处理（Natural Language Processing, NLP）的交叉领域其目的是实现图像到文本的映射。得益于机器翻译模型的成功，Vinyals等人[20]首先提出了用于自然描述的基本编码器-解码器（Encoder-Decoder）体系结构，该方法推翻了以前基于模板的方法。在这种方法中，模型前后由两部分组成：编码器和解码器。编码器以一种常用的CNN架构作为主体结构，负责将原始图像映射到全局视觉特征向量中，因此，文献中也称该部分为图像模型。解码器，即语言模型，由一种特定的循环神经网络（Recurrent Neural Network，RNN）架构组成，比如常见的长短期记忆网络（Long Short-Term Memory, LSTM）[21]或门循环单元（Gate Recurrent Unit, GRU）[22]等，负责从全局视觉特征向量中解码生成单词序列。解码过程是以迭代的方式进行的，首先，解码器以特征向量和“启动标记”初始化，预测第一个单词，然后以特征向量和第一个预测单词作为输入，来预测第二个单词……如此往复，直到生成“结束标记”。后来，Xu[24]等人将序列到序列（Seq2Seq）机器翻译模型中的注意力机制（Attention Mechanism）[23]引入到图像内容描述模型中。注意力机制的诞生是为了解决长序列到定长向量转化而造成的信息损失的瓶颈问题，引入注意力机制之后，在翻译生成每个目标语言单词时，模型会关注到所有的源语言单词特征，并赋予不同的关注程度（权重系数不同）。与之相类似，在图像描述任务中，在生成每个描述词汇时，不再仅仅依靠全局视觉特征，而是一个可变的对图像不同区域的加权特征。加权系数反映了对图像不同区域的关注程度，通过一定机制可以学习得到。实验结果表明，基于注意力机制的模型比传统模型获得了更好的性能。

**2.2 医学图像报告生成**

自然图像描述生成的成就极大地促进了医学图像报告自动生成技术的发展。编码器-解码器架构在此领域被广泛应用[25]。然而，好的通用图像描述方法不一定适用于医学图像报告自动生成，主要是因为两个因素。首先，与通用图像描述方法不同，医学图像报告自动生成方法的目的不是简单地描述图像中存在的内容。 其旨在报告与诊断目的相关的临床重要信息。例如，如果只是简单地报告医学图像中显示了哪些器官而没有关于它们的临床重要性描述是不可取的。为了在保证预言流畅性的条件下提升临床准确性，研究者们尝试了诸多方法，并取得了一定成效。一方面，部分方法同通过构建医学概念预测和视觉特征编码共享相同的多任务学习网络。Xue等人[26]图像中提取医学概念作为中间语义特征，馈送到解码器中，从而降低了直接从医学图像映射到报告的难度。Nguyen等人[27]在此基础上更近一步，通过将预测的疾病状态（阴性或阳性）概念返回给模型以生成更为精准的报告。并且将临床历史和主述等病人的临床相关文档作为补充输入以提高疾病状态的准确性。另一方面，部分研究者考虑到医学诊断是一项需要多年专业医学知识支撑的一项工作，因此引入外部知识图来构建知识、数据共同驱动的深度学习方法辅助生成更加精准的诊断描述。Li等人[28]提出了一种新的知识驱动的编码、检索、释义方法。

**2.3 XXXXXX**

……

……

**3. 讨论**

Demner-Fushman等人[33]创建了一个放射学检查数据集（IU X-Ray），包括X射线图像和放射科医师撰写的报告。他们通过开放访问生物医学图像搜索引擎（Open Access Biomedical Image Search Engine，OpenI）公开发布了该数据集的匿名版本。该数据由3955份报告（每位患者一份，全部为英文）和7470张DICOM图像组成。我们发现3851份报告（97.4%）与至少一张图像相关，因此对我们的研究有效。在3851份报告中，599份（15.6%）不包括“COMPARISON”部分，86份不包括“INDICATION”部分，31份（0.1%）不包括“IMPRESSION”部分，而514份（13.3%）不包括“FINDINGS”部分。Jing等人[30]和Srinivasan等人通过将“IMPRESSION”部分和“FINDINGS”部分拼接作为待生成的报告以避免丢掉过多数据。然后按照7：1：2的比例将数据随机划分为训练集、验证集和测试集。且这三个子集中没有重复患者。

**4. 结论**

我们提供了图像报告自动生成方法、公开可用数据集以及其前置任务。就方法而言，目前大多数直流电工作使用编码器-解码器深度学习方法，这主要是因为它们在通用（非医疗）图像字幕中取得了成功。

**参考文献**

[1] 田娟秀,刘国才,谷珊珊,鞠忠建,刘劲光,顾冬冬.医学图像分析深度学习方法研究与挑战[J].自动化学报,2018,44(03):401-424.

[2] 胡颜民. 基于医学图像处理关键技术的骨科疾病计算机辅助诊断系统[D].华南理工大学,2020.

[3] LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series[J]. The handbook of brain theory and neural networks, 1995, 3361(10): 1995.