

顾及高低速信息的道路限速识别方法

张彩丽¹, 向隆刚², 李雅丽³

(1.河南城建学院测绘与城市空间信息学院,河南 平顶山 467000)

(2.武汉大学测绘遥感信息工程国家重点实验室,湖北 武汉 430079)

(3.沈阳建筑大学交通与测绘工程学院,辽宁 沈阳 110000)

[摘要] 随着新兴行业对地图细节丰富度和要素完备性的需求,开展精细可导航路网构建研究是大势所趋,然而基于轨迹数据的道路信息挖掘,多聚焦于路网骨架的识别,较少对道路限速信息进行识别.考虑低速与高速路段的轨迹运行速度分布不同,本文基于轨迹数据对高低速路段分别进行建模,开展轨迹数据限速识别优化.在此基础上融合路网数据限速识别结果,最终完成限速信息的识别.以武汉市二环路区域与成都市二环路区域为例进行实验,验证了方法的有效性.实验结果表明:(1)高低速信息分别建模明显提升了轨迹限速识别准确率;(2)融合轨迹数据限速识别优化结果与路网数据限速识别结果也进一步提升了限速识别结果.

[关键词] 车辆轨迹,道路网信息,邻接路段,高低速信息,道路限速识别

[中图分类号] P208 **[文献标志码]** A **[文章编号]** 1001-4616(2024)04-0031-08

Study on Road Speed Limit Identification Taking into Account High and Low Speed Information

Zhang Caili¹, Xiang Longgang², Li Yali³

(1.School of Surveying and Urban Spatial Information, Henan University of Urban Construction, Pingdingshan 467000, China)

(2.State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China)

(3.School of Transportation and Geomatics Engineering, Shenyang Jianzhu University, Shenyang 110000, China)

Abstract: With the demand of emerging industries for map detail richness and element completeness, it is a general trend to carry out research on the construction of fine navigable road networks. However, road information mining based on trajectory data focuses more on the identification of road network skeletons and less on the identification of road speed limit information. Considering the difference between the trajectory running speed distribution of the low speed and high speed sections, this paper models the high speed and low speed sections separately based on the trajectory data, and carries out the speed limit identification optimization of the trajectory data. Then on this basis, the speed limit identification results based on road network are fused to complete the identification of speed limit information. Experiments are carried out taking the Second Ring Road area of Wuhan City and Chengdu City as examples to verify the effectiveness of the method. Experimental results show that: (1) the separate modeling of high and low speed information significantly improves the speed limit recognition accuracy based on trajectories; (2) the integration of speed limit recognition optimization results based on trajectories with speed limit recognition results based on road network also further improves the speed limit recognition results.

Key words: vehicle trajectories, road network information, adjacent sections, high and low speed information, road speed limit identification

在车辆导航过程中,用户不仅依赖道路网骨架信息,还经常依赖导航中重要的路段限速信息^[1].道路限速信息的自动提取研究主要集中在交通标志识别^[2-4]、限速影响因素分析预测^[5-8],尽管可以得到较好的识别准确率,但数据源获取方式仍为实地采集,成本高且更新周期长.交通管理部门定期对在役道路进

收稿日期:2024-05-23.

基金项目:河南城建学院博士科研启动金资助项目(K-Q2023032)、河南省高等学校重点科研项目(24B420002)、河南城建学院高等教育教学改革研究与实践项目(2024JG158)、辽宁省教育厅青年项目(LJ212410153040)、国家自然科学基金项目(41771474,42071432).

通讯作者:向隆刚,博士,教授,博士生导师,研究方向:轨迹数据分析、时空大数据管理. E-mail:geoxlg@whn.edu.cn

行调研评估,进而调整部分路段及时段的限速,因此道路限速信息获取方法不仅需要能够获取精确的识别结果还要具有时效性,满足道路限速信息的动态变化。

随着互联网的发展和位置定位技术的普及,地理信息的获取方式发生了翻天覆地的变化,得以获取多种多样的道路数据源,如车辆轨迹、遥感影像、激光点云、城市街景等。得益于覆盖范围广、更新频率高、获取成本低等突出优点,车辆轨迹成为生成可导航道路信息的主流数据源。虽然基于车辆轨迹数据的道路网信息提取吸引了大量研究,输出了丰富多样的成果,但是大多数工作聚焦于路网骨架的识别^[9-10],旨在使道路交叉口及路段中心线尽可能准确。目前国内外基于轨迹数据进行道路限速识别方面的研究寥寥且大多较为简单。廖律超等^[11]基于交通轨迹数据,采用 K 近邻算法(K -nearest neighbors, KNN)进行道路限速识别。该方法以天为单位设计特征,对每天的轨迹数据量要求较高。Windén 等^[12]对包括限速在内的八个道路属性进行更新。限速方面基于高速度平均值与整体平均值设计的判别规则进行识别,其依赖于参数设置,无法保证普适性与准确性。现有研究从监督与非监督两个角度开展,都只考虑轨迹数据的速度信息,对轨迹数据质量要求较高,无法适用于数据噪声大、采样间隔不固定的众源轨迹数据。

一方面众源轨迹数据存在固有的测量误差及采样误差,另一方面城市道路网具有非常复杂的空间结构,仅从众源轨迹数据中识别限速信息是一件相当有挑战性的事情。文献[13]在轨迹数据基础上引入道路网信息,考虑上下游邻接路段,开展了道路限速识别研究,在一定程度上缓解了轨迹数据质量不高的问题。但是不同限速类别下驾驶员行驶行为不同,影响轨迹数据对邻近高、低速路段识别。因此在文献[13]基础上,顾及高低速信息,本文进一步开展道路限速识别研究。本文方法一方面基于轨迹特征对高速路段及低速路段分别进行模型训练,优化轨迹特征限速识别;另一方面融合轨迹特征限速识别优化结果及路网特征限速识别结果,突破数据单源区分能力,提升了限速识别准确率,对丰富、完善导航地图具有重要意义。

1 道路限速挖掘方法

本文顾及高低速信息,提出一种新的道路限速识别方法。首先进行道路网及轨迹数据预处理(具体见 2.1),确保数据的完整性和准确性;然后以路段为分析单元,基于路网与轨迹数据设计道路几何特征、道路分布特征、道路拓扑连接特征、道路单双向及道路运行速度、流量、宽度等特征;最后采用不受维数干扰、分类精度较高的随机森林模型(random forest, RF)^[13-16]进行道路限速识别,具体流程如图 1。

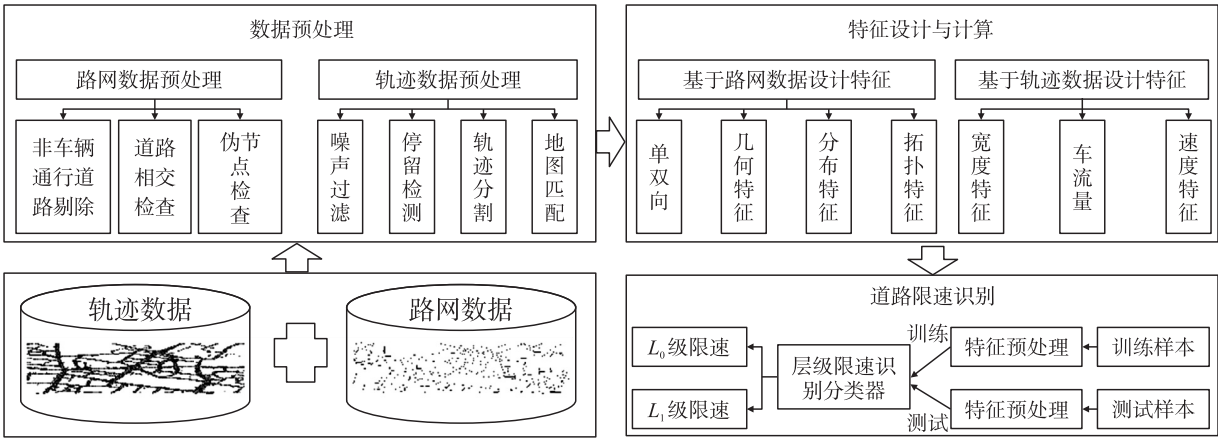


图 1 基于多模特征的道路限速识别框架

Fig. 1 The framework of road speed limit recognition based on multi-mode features

1.1 特征设计与计算

道路网通常由节点边模型 $G = \langle V, E \rangle$ 表示,其中 V 是节点集, E 是两个节点之间的连接弧集,即路段集。节点边模型可以有效地组织和存储路网信息。例如,节点类型、红绿灯信息和转弯关系可以存储为节点属性,而道路等级、道路车道数、道路限速和道路名称等可以存储为路段属性,如图 2 所示。因此,基于该模型,我们以路段为基本单元进行特征设计与计算来识别道路限速信息。

道路特征的设计至关重要,直接影响道路限速识别准确率。经文献[13]验证,道路几何特征、道路分布特征、道路拓扑连接特征、道路单双向特征、道路宽度、道路流量、道路运行速度等可应用于道路限速识

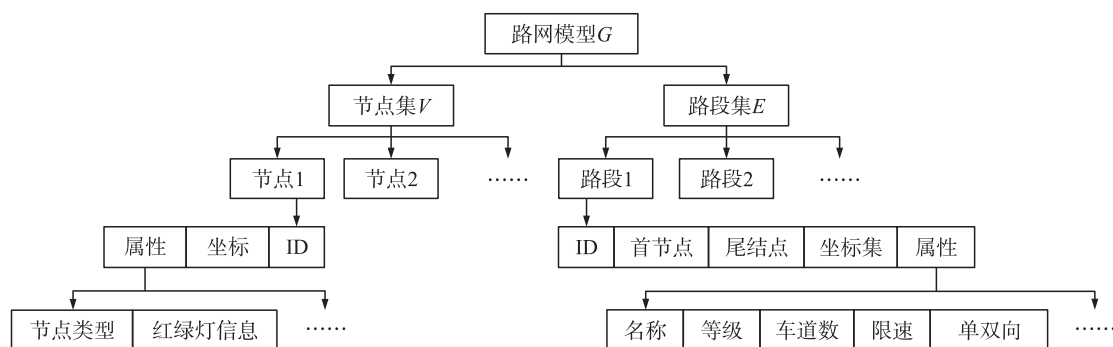


图2 路网模型示意图

Fig. 2 Road network model schematic

别。而这些特征与国家标准^[17]确定的道路功能、运行速度、道路环境(路侧开发程度、道路特征及交通特征)等限速影响因素相对应。因此,本文引入基于路网及轨迹数据设计的道路几何特征、道路分布特征、道路拓扑连接特征、道路单双向特征、道路宽度、道路流量、道路运行速度等提高道路限速识别准确率。另外依据文献[13]和[16]对设计的多模态特征进行了缺失值处理及特征选择。

其中道路几何特征包括道路几何长度和道路弯曲度;道路分布特征包括道路邻接网眼面积、道路上 n 个节点密度平均值和道路上 $n-1$ 个路段线密度平均值;道路拓扑特征包括从组成对象 node 方面度量的节点群聚系数、度中心性、接近中心性、中介中心度,以及从组成对象 link 方面度量的节点群聚系数、度中心性、接近中心性、中介中心度。道路单双向特征包括道路的单双向信息;道路宽度特征包括道路匹配轨迹点到道路中心线距离绝对值的平均值、众数、标准差、中位数、最大值,道路匹配轨迹点到道路中心线距离绝对值的5-95百分位;道路流量特征包括道路所匹配的轨迹条数;道路速度特征包括道路匹配轨迹点速度的平均值、众数、标准差、中位数、最小值、最大值,道路匹配轨迹点速度的5-95百分位,24小时内最大的6个时段平均瞬时速度,高位速特征平均值、众数、标准差、中位数、最小值、最大值。

1.2 道路限速识别

文献[12]认为限速信息可以在层次结构中被划分为两个级别,其中 L_0 级为限速粗略划分,而 L_1 级为限速精细划分。由于低速路段运行速度与高速路段运行速度分布模式不同,本文首先基于路网轨迹特征进行 L_0 级限速分类,初步将限速信息划分为超低速、低速、高速、超高速;然后基于轨迹特征对不同的限速类别进行不同模型训练识别,实现轨迹特征 L_1 级限速识别优化;最后考虑当前路段的属性信息与其左右邻接路段大概率相同^[13],顾及当前路段、左右邻接路段,对轨迹特征 L_1 级限速识别优化结果、路网特征 L_1 级限速识别结果等6个结果进行投票,得到最终 L_1 级限速识别结果。具体流程如图3。

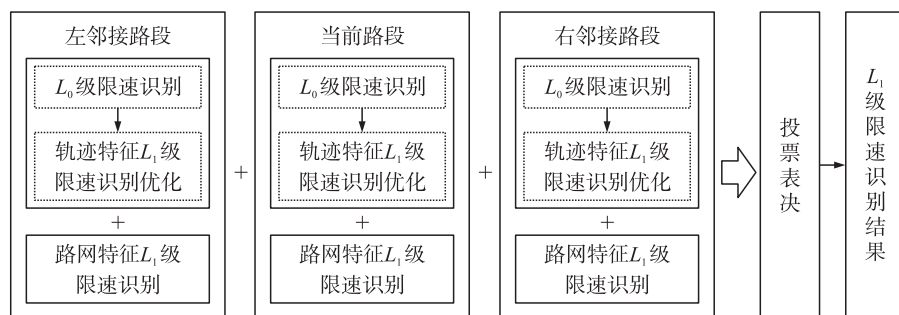


图3 限速识别流程

Fig. 3 Speed limit identification workflow

(1)不同层级限速划分:首先将小于30 km/h限速划分为超低速,大于等于30 km/h小于等于50 km/h限速划分为低速、大于等于60 km/h小于等于80 km/h限速划分为高速、大于80 km/h限速划分为超高速,即为 L_0 级限速划分;然后再对 L_0 级限速进行细化,即超低速划分为10 km/h、15 km/h、20 km/h、25 km/h,低速划分为30 km/h、40 km/h、50 km/h,高速划分为60 km/h、70 km/h、80 km/h,超高速划分为90 km/h、100 km/h、110 km/h、120 km/h,即为 L_1 级限速划分。

(2) L_0 级限速识别:根据文献[18],随机抽取样本的 70%作为训练样本,30%作为测试样本. 对训练样本进行 L_0 级划分,形成新的训练样本,基于路网轨迹特征,对随机森林分类器进行训练,形成 L_0 级限速识别分类器. 考虑左右邻接路段对当前路段限速影响,基于路网轨迹特征,采用 L_0 级限速识别分类器,分别就当前路段及左右邻接路段进行 L_0 级限速识别,然后基于投票表决法,将三个结果融合,实现测试集的 L_0 级限速识别,具体流程如图 4.

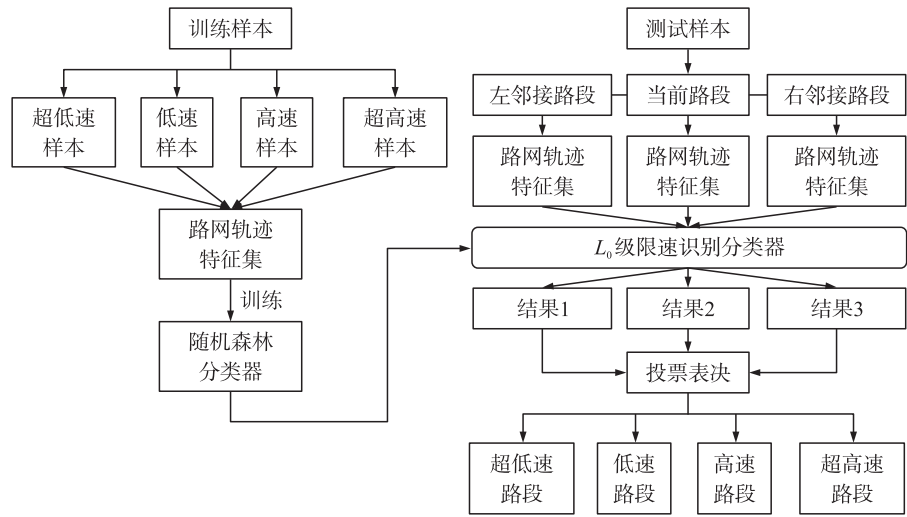


图 4 L_0 级限速识别流程

Fig. 4 The identification workflow of L_0 level speed limit

(3) L_1 级限速识别:基于轨迹特征进行道路限速识别方面,考虑超低限速、超高限速很难精细化识别,将上述两类限速识别到 L_0 级别,而对低速、高速路段分别进行低分类器、高分类器建模,开展 L_1 级限速识别优化,具体见图 5(a)的基于轨迹特征进行 L_1 级限速识别优化流程图. 而基于路网特征进行道路限速识别方面,直接对标记的 L_1 级限速样本进行训练,得到限速分类模型,然后对测试样本进行测试,具体见图 5(b)的基于路网特征进行 L_1 级限速识别流程.

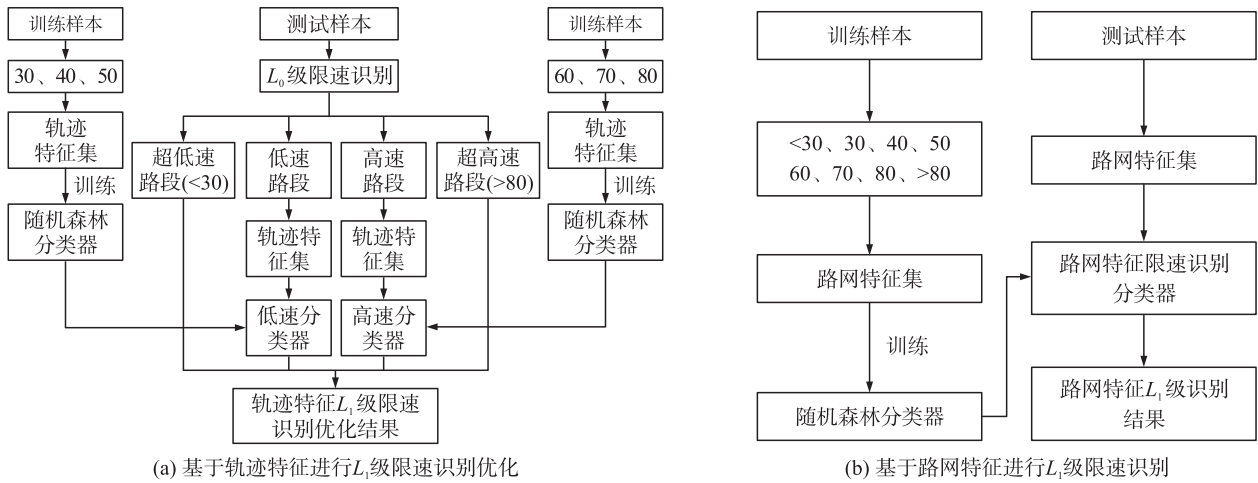


图 5 L_1 级限速识别流程

Fig. 5 The identification workflow of L_1 level speed limit

考虑左右邻接路段影响,本文基于当前路段及左右邻接路段的轨迹特征优化识别结果、当前路段及左右邻接路段的路网特征识别结果,采用投票表决法,将六个结果进行融合,最终得到 L_1 级别限速识别结果. 投票计算公式如下:

$$R(y) = \arg \max(\text{Count}(S_i)),$$

其中 $R(y)$ 为测试样本 y 限速识别结果; $S_i (1 \leq i \leq 6)$ 为当前路段及左右邻接路段的轨迹特征优化识别结果、当前路段及左右邻接路段的路网特征识别结果; $\text{Count}(S_i)$ 为 S_i 识别个数.

2 实验与分析

2.1 研究区域与数据集

本文选取了两个区域进行实验,其中实验区域 1(武汉市二环路区域)采用 2019 年 7 月-9 月众源轨迹数据(图 6(a)),实验区域 2(成都市二环路区域)采用 2018 年 11 月 30 号一天众源轨迹数据(图 6(b)),而基础路网方面两者都采用开放街道地图(OpenStreetMap,OSM)路网,实验区域 1、2 的路网数据分布分别如图 6(c)和(d)所示。

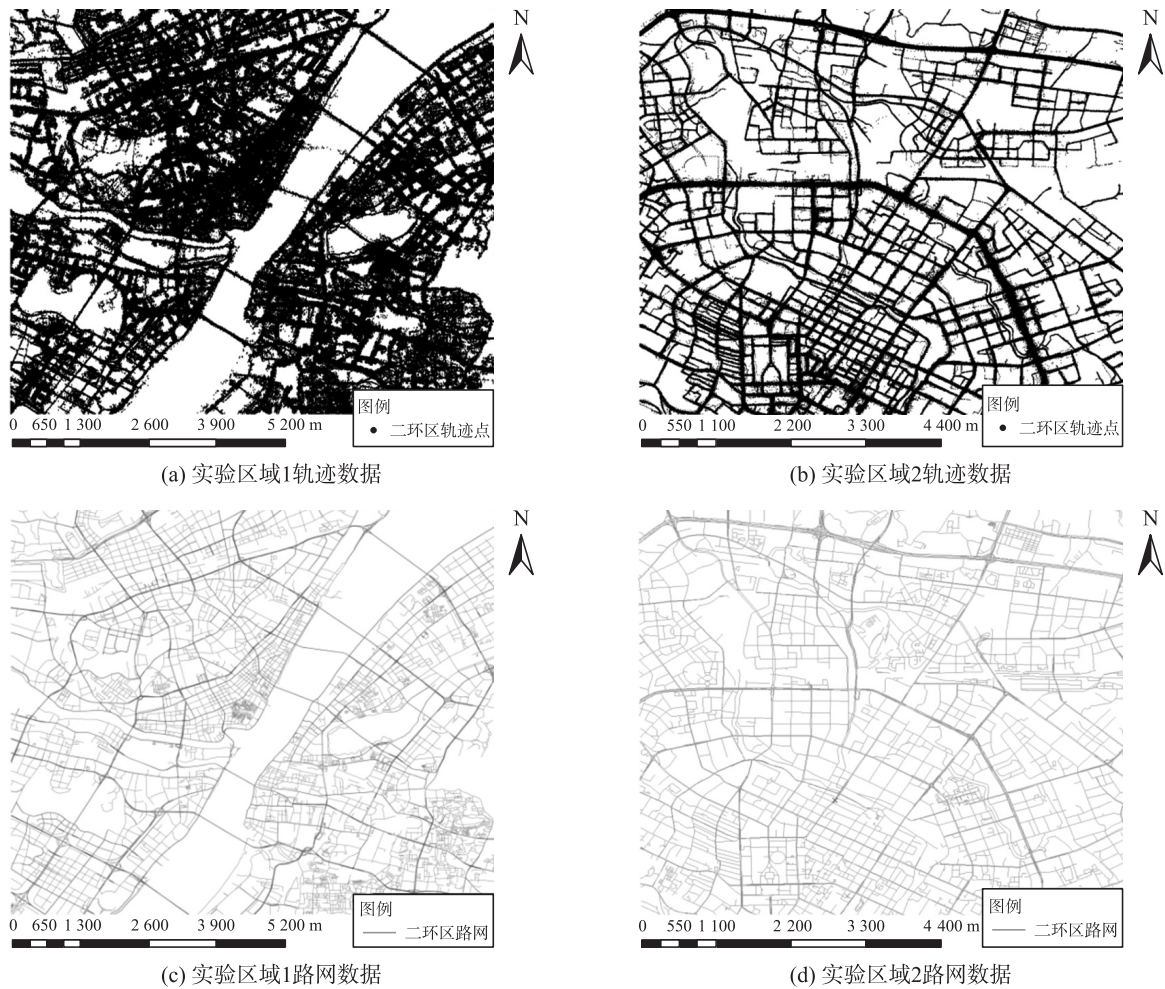


图 6 实验数据

Fig. 6 Experimental data

本文参考文献[16]对实验区域轨迹数据进行了噪声过滤、停留点检测、轨迹分割、地图匹配等预处理,对路网数据进行了非车辆通行道路剔除、道路相交检查及伪节点检查. 处理后路网及轨迹数据具体信息如表 1. 为制作道路限速信息识别的训练及验证样本,从实验区域路网随机选取路段,并进行了限速标注,具体信息如表 2.

表 1 数据集描述信息

Table 1 Description of two datasets

实验区	OSM 路网		轨迹数据	
	原始路段数/条	重构路段数/条	平均采样频率/s	面积(km×km)
实验区域 1	7 733	14 340	30.26	14.8×14.3
实验区域 2	3 318	5 940	5.63	7.4×7.5

表 2 样本数据描述信息
Table 2 Description of road speed limit label datasets

限速 ID	限速类别/(km/h)	实验区域 1/条	实验区域 2/条
0	<30	100	212
1	30	153	56
2	40	353	56
3	50	686	0
4	60	433	197
5	70	111	68
6	80	27	50
标注样本路段总条数		1 863	639

2.2 实验结果与精度评定

考虑到样本数差距大影响结果,我们在进行模型训练前,对训练样本进行了合成少数类过采样技术(SMOTE)处理,在少数类样本之间进行插值来产生额外的样本,使得训练样本能够平衡.在此基础上,进行了不同层级限速识别,其中模型训练时采用了 5 折交叉验证.弱分类器的个数基于精细格网^[19]进行搜索,其他参数设置如下:11 最大特征确定方式:“None”;22 不纯度:“gini”;33 最大树深度:“None”;44 叶子结点包含的最少样本数:“1”;55 类别权值:“balanced”.为确保实验的准确性与客观性,本文所有实验均重复了 10 次,以均值进行结果评定.

观察不同道路限速信息识别情况,本文方法 L_0 级限速分类效果较好,每类识别精度、召回率及 F_1 分数均能达到 70%以上,如表 3.本文方法 L_1 级限速分类结果更细化,整体准确率会有所下降,但除实验区域 1 中的限速 30 外,其他限速识别精度、召回率及 F_1 分数均达到了 65%以上(表 4).实验区域 1 限速 30 识别率低可能由于低速路段运行速度容易偏高.实验区域 1 轨迹数据较杂,包含骑行轨迹、人行轨迹等原因导致.另外由于实验区域 1 轨迹噪声大、采样频率低,该区域限速识别精度整体低于实验区域 2.

表 3 L_0 级不同限速分类精度
Table 3 Classification accuracy of L_0 level speed limit

限速 ID	限速类别	实验区域 1/%			实验区域 2/%		
		精度	召回率	F_1 值	精度	召回率	F_1 值
0	超低速	91.6	75.3	82.7	98.2	98.2	98.2
1	低速	90.1	85.7	87.9	82.6	86.5	84.5
2	高速	73.3	83.9	78.3	95.4	93.9	94.7
3	超高速	—	—	—	—	—	—
总体准确率		84.6			93.9		

表 4 L_1 级不同限速分类精度
Table 4 Classification accuracy of L_1 level speed limit

限速 ID	限速类别/ (km/h)	实验区域 1/%			实验区域 2/%		
		精度	召回率	F_1 值	精度	召回率	F_1 值
0	<30	89.0	86.3	87.6	97.0	98.2	97.6
1	30	69.4	57.9	63.1	80.5	100.0	89.2
2	40	72.7	74.1	73.4	80.6	75.0	77.7
3	50	83.5	74.3	78.7	—	—	—
4	60	67.4	80.5	73.3	94.6	88.5	91.4
5	70	69.4	92.2	79.2	96.5	99.2	97.8
6	80	86.8	59.0	70.2	97.0	100.0	98.5
总体准确率		75.7			93.1		

2.3 实验结果对比分析

为验证本文顾及路网轨迹特征、考虑邻接路段信息及高低速信息的必要性,本文进行了消融实验,结果如表 5 所示.可以看到本文方法比仅考虑单一特征或集成路网轨迹特征识别准确率高.另外本文方法在集成路网轨迹特征,考虑邻接路段基础上,顾及高低速信息后,限速识别准确率也进一步得到了提高,表

明本文方法顾及高低速信息的重要性. 另外实验区域 1 及实验区域 2 中基于轨迹语义特征优化识别的当前路段限速结果分别为 59.7% 及 80.2%, 明显高于基于轨迹语义特征未优化结果 45.9% 及 66.9%, 这应该是本文方法结果有所提升的主要原因.

表 5 消融实验结果

Table 5 Ablation experimental results

类别	道路限速识别准确率/%		类别	道路限速识别准确率/%	
	实验区域 1	实验区域 2		实验区域 1	实验区域 2
轨迹语义特征(优化前)	45.9	66.9	集成特征	66.6	88.3
轨迹语义特征(优化后)	59.7	80.2	考虑邻接路段	74.7	91.8
路网结构特征	60.6	84.3	本文方法	75.7	93.1

本文方法集成路网及轨迹互补特征,考虑左右邻接路段,顾及高低速差异,提出了一种新的道路限速识别方法. 与文献[11]、文献[13]方法识别结果相比,本文方法具有明显的优势,限速识别准确率在实验区域 1 分别提升 150.7%、1.3%,在实验区域 2 分别提升 84.7%、1.4%,如表 6 所示. 需要说明的是对比方法实验采用的训练样本及验证样本跟本文方法一致. 由于文献[11]采用 KNN 算法,因此为进一步验证本文方法集成路网及轨迹特征、考虑邻接路段、顾及高低速差异的必要,本文还基于 KNN 模型进行了消融实验. 从表 6 中可以看到依次集成路网及轨迹特征、考虑邻接路段、顾及高低速信息后,限速识别准确率得到了不同程度的提升. 基于 KNN 算法,集成路网及轨迹特征、考虑邻接路段、顾及高低速信息后,两个实验区域的识别准确率比文献[11]方法识别准确率分别提升了 146.7%、84.7%. KNN 两个重要参数同样进行了精细格网搜索.

表 6 道路限速对比分析

Table 6 Comparative analysis of road speed limits

模型	方法	道路限速识别准确率/%	
		实验区域 1	实验区域 2
RF	本文方法	75.7	93.1
	文献[13]方法	74.7	91.8
KNN	文献[11]方法	30.2	50.4
	轨迹语义特征	37.8	63.5
	路网结构特征	57.1	76.0
	集成特征	63.7	82.7
	顾及邻接路段	72.1	88.1
	顾及高低速	74.5	93.1

3 结论

随着位置获取技术的发展和位置获取设备的普及,众源轨迹数据为可导航路网的更新与维护提供了新的机遇. 因此考虑导航系统中“限速提醒”需求,本文在路网数据基础上,融合轨迹数据信息,考虑邻接路段影响,顾及高低速信息,发展一种新的导航限速信息挖掘方法. 实验结果表明本文方法优于现有限速识别方法. 无论基于现有方法采用的 RF 算法还是 KNN 算法,集成路网及轨迹特征、考虑邻接路段、顾及高低速信息均在一定程度上提升了限速识别准确率.

但本文方法仍存在一些问題,如低速路段识别率依然较低,相邻限速由于特征相近,容易错分等. 另外限速 90 km/h、110 km/h、120 km/h 多设置在高速车道上,需在车道空间识别基础上进行挖掘,而本文限速识别是以路段为基本单元,因此对超高速 90 km/h、110 km/h、120 km/h 统一进行了处理. 未来我们将从车道级限速信息识别方面继续思考,开展研究.

[参考文献]

[1] LI J, QIN Q M, HAN J W, et al. Mining trajectory data and geotagged data in social media for road map inference[J]. Transactions in GIS, 2015, 19(1): 1-18.

- [2] LILLO-CASTELLANO J M, MORA-JIMÉNEZ I, FIGUERA-POZUELO C, et al. Traffic sign segmentation and classification using statistical learning methods[J]. *Neurocomputing*, 2015, 153: 286–299.
- [3] 杨航, 张鑫森, 杨冲. 基于卷积神经网络的公路限速牌识别方法[J]. *地理空间信息*, 2016, 14(1): 31–33.
- [4] 王进, 任小龙, 孙开伟, 等. HSV 颜色空间下用演化超网络识别道路限速标志的研究[J]. *高技术通讯*, 2013, 23(7): 679–684.
- [5] 刘祥敏. 普通公路限速影响因素与预测模型研究[D]. 济南: 山东建筑大学, 2020.
- [6] 宋博. 考虑环境影响的高速公路限速值研究[D]. 西安: 长安大学, 2018.
- [7] GARBER N J, GADIRAJU R. Factors affecting speed variance and its influence on accidents[J]. *Transportation research record*, 1989, 1213(1213): 64–71.
- [8] 孙付勇. 基于神经网络的道路通行时间预估方法的设计与实现[D]. 北京: 北京交通大学, 2019.
- [9] 唐炉亮, 杨雪, 任畅, 等. 轨迹大数据挖掘与高时空精度道路众包测图[M]. 北京: 科学出版社, 2019.
- [10] 张彩丽, 向隆刚, 李雅丽, 等. 基于出租车轨迹的可导航路网构建[J]. *测绘学报*, 2021, 50(12): 1650–1662.
- [11] 廖律超, 蒋新华, 林铭榛, 等. 基于交通轨迹数据挖掘的道路限速信息识别方法[J]. *交通运输工程学报*, 2015, 15(5): 118–126.
- [12] WINDEN K V, BILJECKI F, STEFAN V D S. Automatic update of road attributes by mining GPS tracks[J]. *Transactions in GIS*, 2016, 20(5): 664–683.
- [13] 张彩丽, 向隆刚, 李雅丽, 等. 路段级导航属性信息挖掘[J]. *测绘学报*, 2024, 3(2): 367–378.
- [14] FERNÁNDEZ-DELGADO M, CERNADAS E, BARRO S, et al. Do we need hundreds of classifiers to solve real world classification problems? [J]. *The journal of machine learning research*, 2014, 15(1): 3133–3181.
- [15] 肖湘文, 沈校熠, 柯长青, 等. 基于 Sentinel-1A 数据的多种机器学习算法识别冰山的比较[J]. *测绘学报*, 2020, 49(4): 111–123.
- [16] 张彩丽, 向隆刚, 李雅丽, 等. 顾及路网与轨迹多模特征的道路等级分类研究[J]. *地球信息科学学报*, 2022, 24(10): 1925–1940.
- [17] 全国交通工程设施(公路)标准化技术委员会. 道路交通标志和标线第5部分: 限制速度: GB 5768.5—2017[S]. 北京: 中国标准出版社, 2017.
- [18] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.
- [19] XU Y Y, XIE Z, WU L, et al. Multilane roads extracted from the OpenStreetMap urban road network using random forests[J]. *Transactions in GIS*, 2019, 23(2): 224–240.

[责任编辑: 杜忆忱]