DOI:10.14188/j.2095-6045.20260440

文章编号:2095-6045(2026)05-0009-07

基于智能体的无人机视频车道级交通流探测

唐炉亮[1] 叶昱彤[1] 施泓羽[1] 齐 恒[1] 谭清华[1] 郭心雨[1] 林 黎[1]

1 武汉大学测绘遥感信息工程全国重点实验室,湖北 武汉,430079

Agent-Based Lane-Level Traffic Flow Detection by Drone Video

TANG Luliang[1] *YE Yutong*[1] *SHI Hongyu*[1] *QI Heng*[1] *TAN Qinghua*[1] *GUO Xinyu*[1] *LIN Li*[1]

1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

摘 要:车道级交通流探测是交通信号控制与道路运行评价的重要基础。针对无人机视频交通流获取依赖人工检测线、相机标定和预设道路拓扑的问题,本文提出一种基于智能体的车道级交通流探测方法。该方法依据道路分支关系划分交汇与非交汇场景,利用视觉智能体定位交汇区域及边界;融合道路分割与车辆运动方向构建方向车道单元,依据轨迹访问次序生成交通流图,实现节点流量与转向流量计算。CAROM Air数据集4个场景实验表明,10个交通方向计数准确率达98.6%,非交汇路段节点流量和中型交叉口边流量相关系数分别为0.996 9和0.988 6,验证了方法的有效性。实现了由视频观测向结构化车道级交通流的自动转换。

关键词:车道级交通流;无人机视频;车辆检测;场景拓扑;道路分割;交通流图

中图分类号:U495;TP391;P208

文献标志码:A

Abstract: Lane-level traffic flow detection provides an important basis for traffic signal control and road operational evaluation. To address the reliance of UAV video-based traffic flow acquisition on manually defined detection lines, camera calibration, and predefined road topology, this study proposes an agent-based method for lane-level traffic flow detection. The method distinguishes junction and non-junction scenarios according to road-branch relationships and employs a visual agent to locate junction regions and determine their boundaries. Road segmentation results are then integrated with vehicle motion directions to construct directional lane units, and a traffic flow graph is generated based on the sequential visits of vehicle trajectories to these units, enabling the computation of node flows and turning flows. Experiments on four scenarios from the CAROM Air dataset show that the counting accuracy across 10 traffic directions reaches 98.6%, while the correlation coefficients for node flows on non-merging/splitting road segments and edge flows at a medium-sized intersection reach 0.996 9 and 0.988 6, respectively, demonstrating the effectiveness of the proposed method. The method enables an automatic transformation from video observations to structured lane-level traffic flows.

Key words: lane-level traffic flow; UAV video; vehicle detection; traffic scene topology; road segmentation; traffic-flow graph

车道级交通流量及其时序变化是交通信号配时、拥堵成因识别和道路运行评价的重要依据,其获取精度直接影响交通状态分析的空间分辨能力[1-3]。当前车道级交通流量的监测方法主要由两个方向构成。

一方面,硬件传感器为主的探测方式,在交通流统计环节,虚拟线圈或检测线仍是常见的车辆计数依据[4];也有研究通过轨迹融合获取连续车辆轨迹及车道级交通参数[5]。上述方法通常预先给定统计位置、传感器布设或道路空间信息,尚未直接解决由视频观测自动生成细粒度交通流的问题。

另一类研究直接以历史交通流序列为输入,采用图神经网络建模道路节点之间的空间关联及其时间演化[6,7],或利用Transformer捕获交通状态的长短期时空依赖[8]。相关方法的预测粒度已经由路段扩展至车道级[9],但其主要面向既有流量序列的状态预测,难以由视频观测实时形成车道级统计单元及其流量事件。

无人机航拍具有覆盖范围大、俯视几何关系清晰等特点,可在单次观测中同时记录道路空间结构

引用格式:唐炉亮,叶昱彤,施泓羽,等. 基于智能体的无人机视频车道级交通流探测[J]. 测绘地理信息,2026,51(5):9-15(TANG Luliang, YE Yutong, SHI Hongyu, et al. Agent-Based Lane-Level Traffic Flow Detection by Drone Video [J]. Journal of Geomatics, 2026, 51(5): 9-15)

基金项目:国家重点研发计划(2024YFB3908603);云南省自然资源智能监测与时空大数据治理重点实验室开放课题(202449CE340023)。

与车辆运动状态。

同时其支持车辆轨迹及车道级交通参数提取，为方向化交通流计算提供了新的数据来源[3,10,11]，但是其仍然难以完善解决如何从视频实时监测到车道轨迹流生成的无缝衔接。

随着计算机科学技术的发展，道路语义分割与视觉空间推理为道路统计单元的自动构建提供了互补的信息来源。通用分割模型已表现出面向不同对象类别的零样本分割能力[12,13]；面向道路与交通基础设施的适配研究则表明，语义分割，道路提取与道路标记提取能够为道路统计单元提供像素级空间支撑[14-18]。自动道路图构建研究进一步表明，车辆运动与道路几何之间的对应关系可用于恢复道路拓扑[19]。ReAct 及相关视觉智能体方法通过图像观测、空间操作和反馈更新组织序贯决策，使复杂道路结构能够在多尺度图像域内逐步定位[20-22]。

基于此，本文提出一种基于智能体的无人机视频车道级交通流探测方法，旨在实现由视频观测向结构化车道级交通流的自动转换。本文的主要贡献包括：①提出了面向不同道路拓扑的交通场景自适应解析方法，根据道路分支连接关系区分交汇与非交汇场景，并引入视觉智能体自动定位交汇区域及确定其空间边界，减少对人工统计区域和预设道路拓扑的依赖；②提出了融合道路空间信息与车辆运动特征的方向车道单元构建方法，通过道路分割、轨迹方向划分与空间包络表达，将连续车辆轨迹组织为具有明确空间归属和通行方向的交通流计算单元；③构建了轨迹访问驱动的车道级交通流图，根据车辆对方向车道单元的时序访问关系自动建立节点及有向连接，实现节点流量与转向流量的统一计算，从而形成从无人机视频感知、道路空间解析到车道级结构化交通流生成的完整计算框架。

1 车道级交通流计算方法

1.1 基于智能体的车道级交通流探测方法

本文以无人机视频帧序列为输入，提出一种智能体辅助的车道级交通流探测方法。该方法以车辆运动轨迹为线索，联合解析交通场景结构、道路空间范围与车辆通行方向，自动构建方向车道单元及其有向连接关系，并输出固定时间窗口内的节点流量和边流量。将视频中的车辆运动转换为具有空间位置、通行方向和时间属性的结构化交通流数据。

本文将道路区域内具有唯一运动方向和空间归属的轨迹集合及其空间包络定义为方向车道单元。若道路掩膜对应单一车道，则其内部轨迹形成一个方向车道单元；若一个连通掩膜覆盖双向道路，则相反方向的轨迹分别形成两个方向车道单元。本文所称“车道级”即指这一由道路空间与车辆运动共同确定的方向化计算单元。

整个框架如图 1 所示。依次包括交通场景判别、车辆检测与轨迹提取、交汇区域估计、道路区域提取、方向车道单元构建、轨迹延拓、交通流图构建和流量聚合。交通流图记为 $G=(V,E)$，其中 V 表示方向车道节点，E 表示车辆在观测时段内实际形成的有向连接。

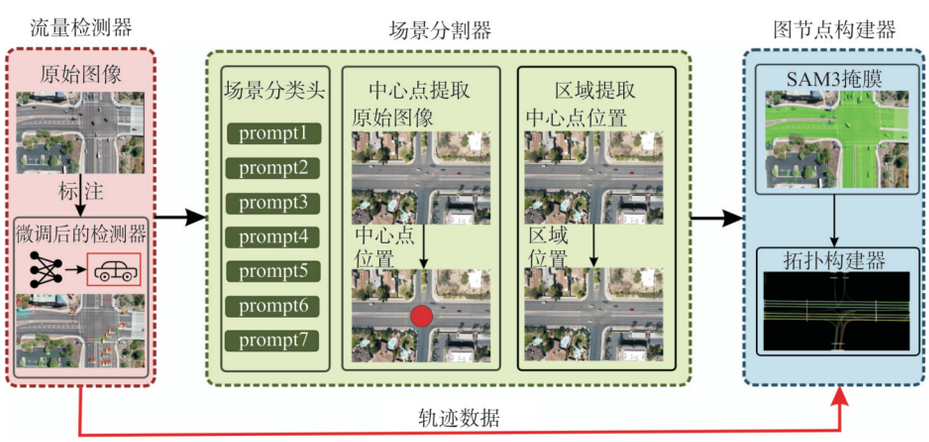


图 1 车道级交通流计算框架模块示意图

Fig. 1 Schematic Diagram of the Lane-Level Traffic Flow Calculation Framework Modules

1.2 交通场景自动构建

本文依据道路分支之间的拓扑关系判别交通场景，并针对交汇与非交汇场景采用不同的道路空间构建方式。设图像域为 Ω，道路区域为 $M\subset\Omega$。对于道路区域内的任意连续子区域 R，以 $\Gamma(R)$ 表示与其边界相接的几何独立道路分支集合；同一条几何道路上的相反行驶方向属于同一分支，不重复计数。

当图像内部存在连接不少于 3 个道路分支且分

支之间形成汇入、分流或转向关系的区域时,判定为交汇场景;否则,判定为非交汇场景。场景状态表示为:

$$z=1\left[\exists R_i\subseteq M:\left|\Gamma(R_i)\right|\geqslant 3\wedge b_i=1\right] \quad (1)$$

式中,R_i 为交汇区域,$z=1$ 和 $z=0$ 分别表示交汇场景和非交汇场景。令 $b_i\in\{0,1\}$ 为布尔标志,表示 $\Gamma(R_i)$ 中的分支在 R_i 内是否形成汇入、分流或转向关系。该定义以道路分支及其连接关系区分两类场景,不以道路外观、车辆数量或车流密度作为判据。

如图 2 所示,对于非交汇场景,直接提取与图像边界相连的道路区域,并沿连续道路走廊建立相邻方向车道单元之间的连接。对于交汇场景,则进一步构建交汇区域,以其边界区分交汇内部空间和外围道路截面;车辆在交汇区域内的运动由入口与出口方向车道单元之间的转移表示。

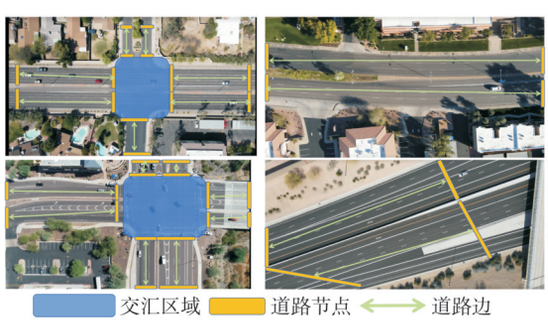


图 2　交会场景与非交会场景的道路空间关系

Fig. 2　Road Spatial Relationships in Merging and Non-merging Scenarios

交汇区域构建包括中心定位和边界确定两个阶段。首先,中心定位过程将当前候选区域划分为规则网格,视觉语言模型根据道路分支的汇入、分流和交叉关系选择包含交汇结构的图块。

所选图块经裁切后继续细分,并保留其相对于原始图像的坐标映射;经过多轮搜索后,以末轮候选图块的几何中心作为交汇中心,如图 3 所示。

随后,以交汇中心为锚点在原始图像上建立矩形候选区域,并通过宽度调整、高度调整和中心平移逐步更新区域的位置与尺度,如图 4 所示。每轮更新后,将候选区域叠加至原始图像并输入视觉语言模型,根据道路分支是否在区域边界处形成稳定的外围道路截面确定后续操作。当候选区域覆盖汇入、分流和转向的核心空间,且其边界保留完整的外围道路横断面时,停止调整,并将最终矩形确定为交汇区域。

1.3　道路区域提取与方向车道单元构建

道路区域为方向车道单元提供像素级空间支

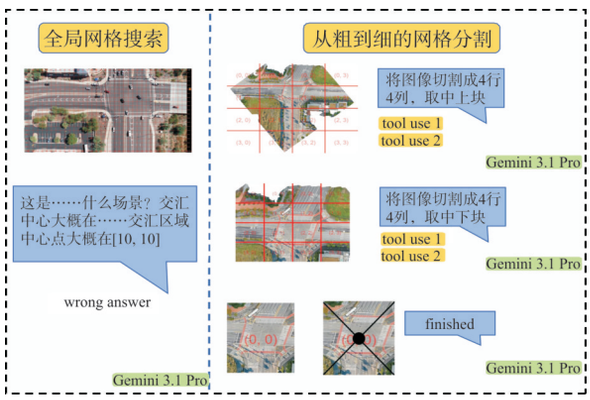


图 3　交汇中心的层级网格定位过程

Fig. 3　Hierarchical Grid Localization Process at the Intersection Center

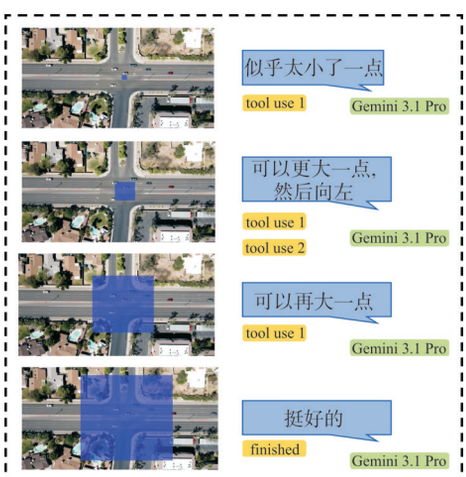


图 4　交汇区域边界的反馈式调整过程

Fig. 4　Feedback-Based Adjustment Process of the Convergence Zone Boundary

撑。当前接收到上一模块提供的道路区域场景划分后,当前阶段需要将二维的栅格影像投影至具体的车道和道路节点。

为了实现这一点,本文采用 Segment Anything Model 3(SAM 3)[23],得到道路的完整掩膜区域。对于每个道路连通区域,本文收集与其相交的车辆轨迹,并依据轨迹在该区域内的位移方向进行分类,从而构建出具体的交通道路节点。单一方向的轨迹保留为一个方向组,双向道路中的相反方向轨迹分别形成两个方向组。随后,以各方向组轨迹点的凸包表示相应方向车道单元的空间范围,凸包形成后,本文对全部相关轨迹进行再次归属。

方向车道单元最终作为交通流图的节点,并记录其空间范围、主方向、中心位置和边界类型,整体结构如图 5 所示。

1.4　交通流探测

在凸包及道路节点生成后,本文根据车辆轨迹

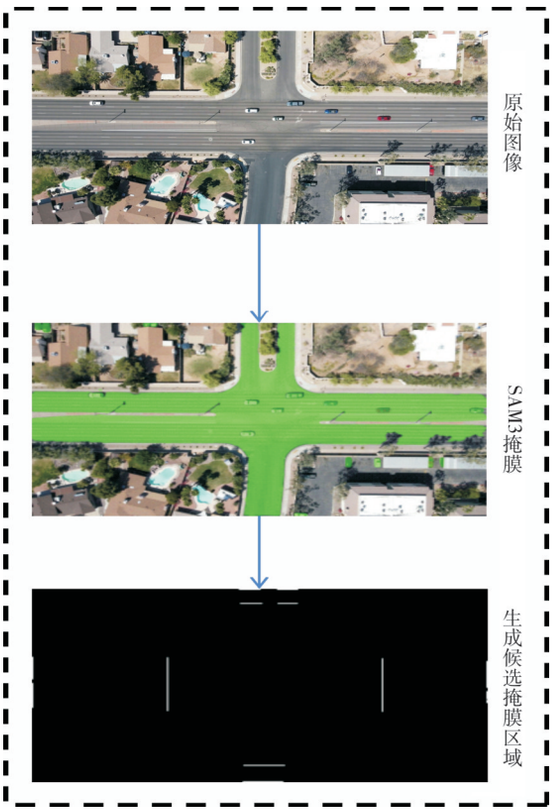


图 5 道路候选掩膜的提取与精修结果

Fig. 5 Extraction and Refinement Results of Road Candidate Masks

的节点访问顺序生成交通流事件。当车辆轨迹中的节点 u 与节点 v 相邻时，将车辆首次完成由节点 u 向节点 v 转移并进入节点 v 的时刻记为有向边 (u,v) 的转移事件时刻。边界延拓轨迹保留原车辆身份，因此同一车辆在同一节点或同一有向边上仅记录一次。设观测时段内共有 N 条边界延拓后的车辆轨迹，其集合表示为：

$$\tilde{T}=\left\{\tilde{T}_i\right\}_{i=1}^{N} \tag{2}$$

式中，N 为观测时段内的车辆轨迹总数。对轨迹 $\tilde{T}_i$，将车辆首次进入节点 v 的时刻记为 $\tau_i(v)$。当节点 u 和节点 v 在该轨迹的节点访问序列中相邻时，将车辆首次完成由节点 u 向节点 v 转移并进入节点 v 的时刻记为 $\tau_i(u,v)$。

为实现交通流事件的时间聚合，以观测时段的起始时刻为时间原点，将观测时段划分为长度相等的时间窗口。第 k 个时间窗口定义为：

$$I_k=[(k-1)\Delta t, k\Delta t\,],\ \Delta t=30\ \mathrm{s},\ k=1,2,\cdots,K \tag{3}$$

式中，K 为时间窗口总数。为判断交通流事件是否发生在指定时间窗口内，定义时间窗口 I_k 的指示函数为：

$$\mathbb{I}_{I_k}(t)=\begin{cases}1, t\in I_k\\ 0, t\notin I_k\end{cases} \tag{4}$$

式中，$\mathbb{I}_{I_k}(t)$ 为时间窗口 I_k 的指示函数。当事件时刻 t 位于 I_k 内时，函数值为 1；否则，函数值为 0。在此基础上，第 k 个时间窗口内节点 v 的节点流量、有向边 (u,v) 的边流量分别定义为：

$$Q_v(v,k)=\sum_{i=1}^{N}\mathbb{I}_{I_k}(\tau_i(v)),\ v\in V \tag{5}$$

$$Q_e((u,v),k)=\sum_{i=1}^{N}\mathbb{I}_{I_k}(\tau_i(u,v)),\ (u,v)\in E \tag{6}$$

式中，V 和 E 分别为道路节点集合和有向边集合；$Q_v(v,k)$ 表示第 k 个时间窗口内首次进入节点 v 的车辆数；$Q_e((u,v),k)$ 表示第 k 个时间窗口内完成由节点 u 向节点 v 转移的车辆数。本文采用 30 s 的时间窗口聚合车辆事件，并在分别统计不同车辆类别后进行求和，得到相应时间窗口内的总交通流量。

2 实验数据与评价方法

2.1 数据来源与实验场景

实验采用 CAROM Air 公开数据集[24-25]。该数据集包含约 100 h 的无人机航拍视频，其提供了高精度的车辆轨迹数据，可用于还原交通场景的完整动态信息并比较航拍视频中的车辆检测轨迹和方向流量。选择部分场景的视频时长，规模等信息见表 1。当前本文选择了该数据集中公开的多个场景数据，自动计算流程以 YOLOv8 和 DeepSORT 从视频中提取的观测车辆轨迹为输入，CAROM Air 提供的参考轨迹用于建立方向计数的参考结果。

表 1 实验场景及交通流图规模

Tab. 1 Experimental Scenarios and Scales of Traffic Flow Graphs

场景	视频时长/min	场景类型	节点数	边数	车辆轨迹数	主要特征
C3_1_0	约 10	非交汇场景	4	2	89	双向直行道路
C_1_2	约 10	交汇场景	16	18	204	中型交叉口
1000_0	约 19	交汇场景	16	20	372	大型交叉口
D0_0_0	约 10	多车道路段	6	2	204	多车道主干道

2.2 实验

YOLOv8 的输入分辨率为 720×720，批次大小为 8，共训练 85 轮。优化器采用 AdamW，初始学习率为 5×10^{-4}，权重衰减为 1×10^{-2}，训练过程中关闭自动混合精度。DeepSORT 的最大轨迹存活帧数为 30，特征库容量为 150，检测框交并比匹配阈值为 0.7。后续交通流分析将不同机动车类别统一为车辆类别。

在交汇区域识别与划分节点,当前的ReAct双智能体采用gemini-3.1-pro,分别负责两个单独的节点,使用LangGraph智能体编排框架,相关具体参数如下所示:道路分割采用road、car、lane、cement floor和shade五类文本提示,车辆区域采用直径10 px的椭圆核扩展,移除面积小于图像面积1%的孤立连通区域,随后采用半径3 px的结构元素,主要参数见表2、表3。

2.3　评价指标

车辆方向计数采用微平均绝对计数准确率。设N_d^{det}和N_d^{ref}分别为第d个交通方向的观测轨迹计数和参考轨迹计数,则有:

$$\mathrm{Acc}_{\mathrm{count}}=1-\frac{\sum_{d=1}^{D}\left|N_d^{\mathrm{det}}-N_d^{\mathrm{ref}}\right|}{\sum_{d=1}^{D}N_d^{\mathrm{ref}}} \tag{7}$$

该指标以各方向绝对计数误差之和衡量总体方向计数偏差,不将不同方向的正负误差相互抵消。节点和边的流量时间序列采用皮尔逊相关系数r、决定系数R^2、平均绝对误差(MAE)评价。式中,r描述两组序列变化趋势的一致性,R^2描述自动计算结果对人工标线序列变异的解释程度,MAE衡量平均误差幅度及对较大误差的敏感程度。

3　结果与分析

3.1　车辆检测与方向计数

YOLOv8训练25轮后,验证集Precision为0.665,Recall为0.585。结果如图6所示。10个交通方向的观测轨迹计数与参考轨迹计数总数均为569辆,各方向绝对误差之和为8辆。微平均绝对计数准确率为98.6%。最大多计出现在1000_0方向2,多计2辆。结果表明,检测结果经过跨帧关联后能够形成较为稳定的方向车辆总数。

从图6可见,当前YOLO+DeepSort的组合在各种场景中总体均表现稳定,尽管存在将静止的小物体误判为车辆的情况,但是大部分小物体并不移动,因此不造成较大影响。在移动目标中,在各个场景下,如十字路口,不规则交叉口,环岛和非交汇场景中均稳定识别到车辆并生成跟踪框,生成跟踪轨迹。

3.2　节点与边流量计算结果

表4反映的是车辆轨迹形成后的方向聚合结果,而验证集Recall和mAP描述的是单帧目标检测性能。

表5给出了4个场景中自动计算结果与人工标线计数结果的比较。图7展示方向车道节点及其边流量。图8展示由轨迹凸包生成交通流图的过程。帧级检测指标与轨迹级方向计数具有不同的评价单位。连续帧中的重复观测为车辆身份关联提供时间冗余,部分帧漏检并不必然导致完整车辆轨迹缺失。

C3_1_0为非交汇双向直线路段,4个节点的总流量差异均不超过5%,节点流量r为0.996 9,边流量r为0.996 1。该结果说明,当道路连通区域与车辆运动方向具有明确对应关系时,方向分类和凸包构造能够稳定表达双向道路的通行流。

C_1_2包含16个节点和18条有向边,节点流量r为0.992 0,边流量r为0.988 6。18条边中有17条的相关系数不低于0.99;剩余1条边的相关系数为0.711 6,主要由于节点凸包覆盖范围与人工横断面位置不一致。1000_0场景的节点流量r为0.985 9,边流量r降至0.917 9、MAE为14.76;偏差集中相邻凸包重叠使方向接近的轨迹在唯一归属时发生节点

表2　主要模型与计算参数

Tab.2　Main Models and Computational Parameters

模块	参数	取值
YOLOv8	输入分辨率	720×720
YOLOv8	批次大小/训练轮数	8/85
YOLOv8	初始学习率/权重衰减	5×10^{-4} / 1×10^{-2}
DeepSORT	最大存活帧数	30
DeepSORT	特征库容量	150
DeepSORT	IoU匹配阈值	0.7
道路分割	每类提示推理次数	3
掩膜处理	车辆区域扩展/道路区域扩展	10 px/40 px
轨迹延拓	步长/最大长度	5 px/300 px
方向划分	方向距离阈值	90°
流量聚合	时间窗口	30 s

表3　智能体相关参数

Tab.3　Agent-Related Parameters

参数	交汇中心定位节点	空间边界分割节点
视觉语言模型	Gemini 3.1 Pro Preview	Gemini 3.1 Pro Preview
编排框架	LangGraph	LangGraph
输入	原始交通场景图像	原始图像、交汇中心
初始状态	完整图像	100×100 px候选矩形
单次模型最大尝试次数	3次	3次



图6　车辆识别图

Fig.6　Vehicle Identification Diagram

切换，并传递至转向边流量。

表 4 观测车辆轨迹与参考轨迹的方向计数结果

Tab. 4 Direction Counting Results of Observed Vehicle Trajectories and Reference Trajectories

场景	方向	观测轨迹计数/辆	参考轨迹计数/辆	差值/辆	方向计数准确率/%
C_1_2	1	44	45	−1	97.8
C_1_2	2	49	49	0	100.0
C_1_2	3	3	3	0	100.0
C_1_2	4	10	10	0	100.0
1000_0	1	92	92	0	100.0
1000_0	2	123	121	+2	98.4
1000_0	3	39	38	+1	97.4
1000_0	4	33	33	0	100.0
C3_1_0	1	86	89	−3	96.6
C3_1_0	2	90	89	+1	98.9

表 5 自动计算方法与人工标线计数法的流量序列比较

Tab. 5 Comparison of Traffic Flow Series Between the Automatic Computation Method and the Manual Line-Marking Counting Method

场景	节点数	节点 r	节点 R^2	节点 MAE	边数	边 r	边 MAE
C3_1_0	4	0.996 9	0.992 7	0.92	2	0.996 1	0.48
C_1_2	16	0.992 0	0.967 1	1.51	18	0.988 6	1.51
1000_0	16	0.985 9	0.970 0	2.44	20	0.917 9	14.76
D0_0_0	6	0.926 5	0.634 8	9.65	2	0.976 2	0.46

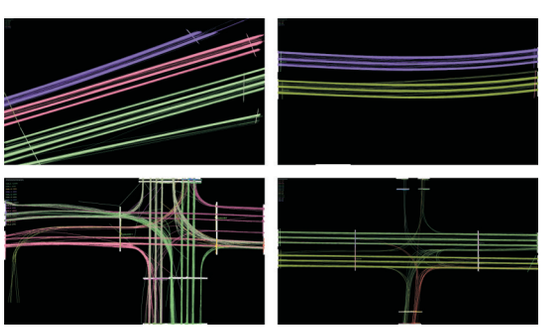


图 7 运行结果示例图

Fig. 7 Example of Running Results

D0_0_0的节点流量 R^2 为 0.634 8，表明宽道路断面内的同向多车道被合并后，凸包范围小于人工横断面，导致节点进入事件系统性低估。

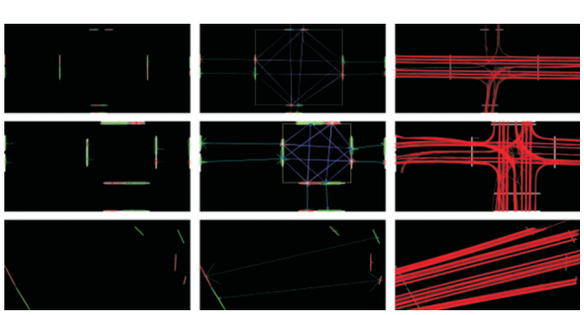


图 8 从凸包到流量处理流程图

Fig. 8 Flow Chart of the Process from Convex Hull to Traffic Processing

上述结果表明，方向车道单元的空间分辨能力取决于道路连通区域与运动方向的对应关系；当同向多车道横向间隔较小或交汇内外道路连续时，仅依靠方向和空间距离难以稳定分离相邻单元。

4 结束语

本文提出一种基于智能体的无人机视频车道级交通流探测方法，实现由视频车辆轨迹向结构化交通流的自动转换。无人机航拍能够提供覆盖范围较大的连续车辆轨迹，为城市交通状态分析提供数据基础[26]；影像与轨迹数据的融合则有助于提高道路网络自动生成的完整性[27]。该方法依据道路分支拓扑关系自适应区分交汇与非交汇场景，利用视觉智能体定位交汇区域并确定空间边界，在此基础上融合道路分割结果与车辆运动方向构建方向车道单元，并根据车辆对各单元的时序访问关系建立有向交通流图，实现节点流量与转向流量的统一计算，从而降低交通流统计对人工检测线、相机标定和预设道路拓扑的依赖。

CAROM Air数据集4个场景的实验结果表明，10个交通方向的微平均绝对计数准确率达到98.6%；非交汇双向路段节点流量相关系数达到0.996 9，中型交叉口边流量相关系数达到0.988 6，表明本文方法能够较准确地恢复不同道路方向的车辆通行及其时序变化。在道路空间与车辆运动方向具有明确对应关系的场景中，自动计算结果与人工标线计数具有较高一致性。

实验同时表明，方法性能受到方向车道单元空间分辨能力的影响。在同向多车道横向间隔较小或交汇区域道路空间连续的复杂场景中，方向相近轨迹可能被合并至同一单元，节点空间重叠也可能引起轨迹归属切换，并进一步影响转向流量计算。后续研究可进一步增强同向相邻车道的精细分离与复杂交汇区域的节点空间约束，以提升复杂交通场景下车道级交通流探测的稳定性与适用性。

参 考 文 献

[1] Klein L A, Mills M K, Gibson D R P. Traffic Detector Handbook: Third Edition—Volume I: FHWA-HRT-06-108[R]. McLean, VA, USA: Federal Highway Administration, 2006, https://www.fhwa.dot.gov/publications/research/operations/its/06108/

[2] Daganzo C F. The Cell Transmission Model, Part II: Network Traffic[J]. Transportation Research Part B: Methodological, 1995, 29(2): 79-93

[3] 黄玲，吴泽荣，洪佩鑫，等. 基于地空信息融合的无

人机交通状态感知方法研究[J]. 中国公路学报, 2021, 34(12): 249-261

[4] Liu Fei, Zeng Zhiyuan, Jiang Rong. A Video-Based Real-Time Adaptive Vehicle-Counting System for Urban Roads[J]. PLoS One, 2017, 12(11): e0186098

[5] 戴喆, 吴宇轩, 董是, 等. 雷达与视觉传感器融合的高速公路全域车辆轨迹与交通参数检测方法[J]. 交通运输工程学报, 2025, 25(1): 197-210

[6] Li Mengzhang, Zhu Zhanxing. Spatial-Temporal Fusion Graph Neural Networks for Traffic Flow Forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(5): 4 189-4 196

[7] 冯宁, 郭晟楠, 宋超, 等. 面向交通流量预测的多组件时空图卷积网络[J]. 软件学报, 2019, 30(3): 759-769

[8] Liu Shipeng, Wang Xingjian. An Improved Transformer Based Traffic Flow Prediction Model[J]. Scientific Reports, 2025, 15: 8284

[9] Shen Guojiang, Yu Kaifeng, Zhang Meiyu, et al. ST-AFN: A Spatial-Temporal Attention Based Fusion Network for Lane-Level Traffic Flow Prediction[J]. PeerJ Computer Science, 2021, 7: e470

[10] Butilă E V, Boboc R G. Urban Traffic Monitoring and Analysis Using Unmanned Aerial Vehicles (UAVs): A Systematic Literature Review[J]. Remote Sensing, 2022, 14(3): 620

[11] 马小龙, 余强, 刘建蓓, 等. 基于无人机视频拍摄的高速公路小型车换道行为特性[J]. 中国公路学报, 2020, 33(6): 95-105

[12] Ravi N, Gabeur V, Hu Y T, et al. SAM 2: Segment Anything in Images and Videos[C]//International Conference on Learning Representations. Singapore, Singapore: International Conference on Learning Representations, 2025

[13] Maquiling V, Byrne S A, Niehorster D C, et al. Zero-Shot Segmentation of Eye Features Using the Segment Anything Model (SAM)[J]. Proceedings of the ACM on Computer Graphics and Interactive Techniques, 2024, 7(2): 1-16

[14] Sultan R I, Li Chengyin, Zhu Hui, et al. GeoSAM: Fine-Tuning SAM with Multi-modal Prompts for Mobility Infrastructure Segmentation [C]//ECAI 2025: 28th European Conference on Artificial Intelligence, Amsterdam, Netherlands: iOS Press, 2025

[15] Feng Wenqing, Guan Fangli, Sun Chenhao, et al. Road-SAM: Adapting the Segment Anything Model to Road Extraction from Large Very-High-Resolution Optical Remote Sensing Images[J]. IEEE Geoscience and Remote Sensing Letters, 2024, 21: 6012605

[16] 戴激光, 王杨, 杜阳, 等. 光学遥感影像道路提取的方法综述[J]. 遥感学报, 2020, 24(7): 804-823

[17] 王国志, 黄玉春. 高分辨率遥感图像的道路自动提取[J]. 测绘地理信息, 2020, 45(3): 34-38

[18] 郑莉, 孙梦, 季铮, 等. 基于车载激光点云的自适应阈值提取道路标记方法[J]. 测绘地理信息, 2019, 44(4): 111-115

[19] Qi Heng, Shi Hongyu, Yang Xue, et al. AutoLRG: A Two-Stage Framework for Automated Lane-Level Road Graph Construction[J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 4413013

[20] Yao S, Zhao J, Yu D, et al. ReAct: Synergizing Reasoning and Acting in Language Models[C]//International Conference on Learning Representations. Kigali, Rwanda: OpenReview.net, 2023

[21] Chowa S S, Alvi R, Rahman S S, et al. From Language to Action: A Review of Large Language Models as Autonomous Agents and Tool Users[J]. Artificial Intelligence Review, 2026, 59(2): 71

[22] Zhang Yunkai, Li Linda, Cui Yingxin, et al. Grid2Matrix: Revealing Digital Agnosia in Vision-Language Models[EB/OL]. 2026: arXiv: 2604.09687. https://arxiv.org/abs/2604.09687

[23] Carion N, Gustafson L, Hu Yuanting, et al. SAM 3: Segment Anything with Concepts [EB/OL]. 2025: arXiv: 2511.16719. https://arxiv.org/abs/2511.16719

[24] Lu D, Eaton E, Weg M, et al. CAROM Air—Vehicle Localization and Traffic Scene Reconstruction from Aerial Videos [C]//2023 IEEE International Conference on Robotics and Automation (ICRA). London, United Kingdom. New York: IEEE Press, 2023

[25] Lu D, Jammula V C, Como S, et al. CAROM — Vehicle Localization and Traffic Scene Reconstruction from Monocular Cameras on Road Infrastructures[J]. IEEE International Conference on Robotics and Automation, 2021:233004562,doi:10.48550/arXiv.2104.00893

[26] Barmpounakis E, Geroliminis N. On the New era of Urban Traffic Monitoring with Massive Drone Data: The pNEUMA Large-Scale Field Experiment [J]. Transportation Research Part C: Emerging Technologies, 2020, 111: 50-71, doi:10.1016/j.trc.2019.11.023

[27] Bai Xiangdong, Feng Xuyu, Yin Yuanyuan, et al. Combining Images and Trajectories Data to Automatically Generate Road Networks[J]. Remote Sensing, 2023, 15(13): 3343, doi:10.3390/rs15133343

修回日期:2026-08-28

第一作者:唐炉亮,博士,教授。主要从事时空大数据挖掘与场景感知。E-mail:tll@whu.edu.cn

通信作者:叶昱彤,硕士,主要从事遥感识别,计算机视觉与智能体应用。E-mail:yeyutong@whu.edu.cn