

基于多尺度时空注意力机制的车辆换道意图预测

吴林蓉

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2025年9月9日; 录用日期: 2025年9月17日; 发布日期: 2025年10月22日

摘要

车辆换道行为预测是智能交通系统和自动驾驶技术的核心组成部分,对提升道路交通安全具有重要意义。针对现有方法在时空特征建模分离、多尺度信息融合不足和动态权重调整缺失等方面的技术挑战,本文提出了一种基于多尺度时空注意力机制的车辆换道意图预测模型(MSTAN)。该模型采用分治协同的设计思想,构建了三个并行处理分支:时序特征建模分支基于双向长短时记忆网络和自注意力机制捕获车辆运动状态的时间依赖关系;空间交互建模分支利用图卷积网络和边注意力机制建模车辆间的动态空间关系;多尺度特征融合分支通过并行多分辨率卷积和注意力机制提取多粒度上下文信息。模型采用自适应权重融合策略实现三个分支特征的最优整合。基于NGSIM数据集的实验结果表明,MSTAN在US-101、I-80和Lankershim三个路段的换道行为三分类任务中分别达到98.3%、97.6%和95.2%的预测准确率,优于所有基线方法。并通过分支贡献度分析和注意力机制有效性评估,定量验证了各功能模块对模型性能的独立贡献和协同效应。时间窗口敏感性分析和预测时间提前量分析确定了最优的时序建模参数配置,证明了模型在不同预测场景下的稳定性和鲁棒性。

关键词

车辆换道预测, 多尺度注意力机制, 图卷积网络, 时空建模, 智能交通系统

Vehicle Lane Change Intention Prediction Based on Multi-Scale Spatial-Temporal Attention Mechanism

Linrong Wu

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: September 9, 2025; accepted: September 17, 2025; published: October 22, 2025

文章引用: 吴林蓉. 基于多尺度时空注意力机制的车辆换道意图预测[J]. 建模与仿真, 2025, 14(10): 263-276.
DOI: 10.12677/mos.2025.1410622

Abstract

Vehicle lane change behavior prediction is a core component of intelligent transportation systems and autonomous driving technologies, playing a crucial role in enhancing road traffic safety. Addressing the technical challenges of existing methods in spatial-temporal feature modeling separation, insufficient multi-scale information fusion, and the absence of dynamic weight adjustment, this paper proposes a Multi-Scale Spatial-Temporal Attention Network (MSTAN) for vehicle lane change intention prediction. The model adopts a divide-and-conquer collaborative design philosophy, constructing three parallel processing branches: the temporal feature modeling branch captures temporal dependencies of vehicle motion states based on bidirectional long short-term memory networks and self-attention mechanisms; the spatial interaction modeling branch utilizes graph convolutional networks and edge attention mechanisms to model dynamic spatial relationships between vehicles; the multi-scale feature fusion branch extracts multi-granularity contextual information through parallel multi-resolution convolutions and attention mechanisms. The model employs an adaptive weight fusion strategy to achieve optimal integration of features from the three branches. Experimental results on the NGSIM dataset demonstrate that MSTAN achieves prediction accuracies of 98.3%, 97.6%, and 95.2% for lane change behavior three-class classification tasks on the US-101, I-80, and Lankershim road segments, respectively, outperforming all baseline methods. Through branch contribution analysis and attention mechanism effectiveness evaluation, the independent contributions and synergistic effects of each functional module on model performance are quantitatively validated. Time window sensitivity analysis and prediction lead time analysis determine the optimal temporal modeling parameter configuration, demonstrating the model's stability and robustness across different prediction scenarios.

Keywords

Vehicle Lane Change Prediction, Multi-Scale Attention Mechanism, Graph Convolutional Network, Spatial-Temporal Modeling, Intelligent Transportation System

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着智能交通系统和自动驾驶技术的快速发展, 驾驶辅助系统(Advanced Driver Assistance Systems, ADAS)已成为提升道路交通安全性的关键技术手段。在交通事故成因分析中, 车辆换道行为相关事故占据显著比例。美国国家公路交通安全管理局统计数据显示, 2016~2018 年期间, 道路偏离事故年均造成 19,158 人死亡, 占交通事故总死亡人数的 51% [1]。研究表明, ADAS 系统的全面部署能够使道路交通事故发生频率降低 23.8%, 年均可预防 18,925 起事故[2]。在 ADAS 的发展中, 开发人类驾驶行为预测模型是一个主要研究方向, 该类模型通过与 ADAS 系统的深度集成, 能够实时生成安全预警信息和操作提示, 在减少交通事故和提升道路安全方面具有重要的应用价值。

车辆换道预测方法的发展经历了从传统物理建模到现代数据驱动的演进历程。早期研究主要基于车辆动力学原理, 采用恒定速度(CV)、恒定加速度(CA)和恒定转弯率速度(CTRV)等运动学模型[3]。Toledo 等人[4]利用 GPS 和 IMU 传感器构建交互式多模型扩展卡尔曼滤波器, 在高速公路场景中实现了较好的短期预测效果。然而, 这类方法仅依赖低层次运动特性, 在复杂交通环境和长期预测方面存在一定局限

性。近年来,数据驱动的深度学习方法逐渐成为主流。研究者们开始采用长短期记忆网络(LSTM)挖掘时序模式[5],利用图神经网络(GNN)建模车辆间空间交互关系[6],并通过注意力机制提升预测精度[7]。Wang 等[8]提出的混合轨迹预测框架在自动驾驶场景中展现了优异性能;Gao 等[9]开发了基于双 Transformer 架构的预测模型,在混合交通环境中取得了良好效果。最新研究表明,LSTM、CNN 和 Transformer 网络结构都被广泛应用于预测人类驾驶员换道意图。然而,现有研究仍存在一些关键问题亟待解决:时空建模分离性问题,现有方法往往将时间和空间特征分别处理,缺乏有效的时空耦合建模机制;多尺度特征融合不足,不同时间尺度和空间尺度的信息未能得到充分整合;动态权重调整缺失,模型无法根据不同场景自适应调整各特征分支的重要性;三分类预测精度有待提升,现有方法在区分左换道、保持车道、右换道三类行为时的细粒度识别能力不足。

针对上述技术挑战,本研究提出了一种基于多尺度时空注意力机制的车辆换道预测方法(Multi-Scale Spatiotemporal Attention Network, MSTAN)。该方法的主要贡献包括:1) 设计了多分支并行处理架构,通过双向长短期记忆网络处理时序特征、图卷积网络建模空间交互关系、多尺度卷积模块提取多粒度上下文信息,实现时空特征的有效耦合;2) 构建了基于图卷积网络的动态邻域建模模块,能够自适应捕获车辆间复杂的时变空间关系;3) 提出了多尺度特征提取机制,通过不同感受野的卷积核同时捕获局部细节和全局上下文信息;4) 引入了自适应权重融合策略,根据不同场景动态调整各特征分支的重要性,提升三分类预测的细粒度识别能力。实验结果表明,所提出的 MSTAN 方法在多个公开数据集上均取得了显著的性能提升。

2. 数据集简介及预处理

2.1. 数据集描述

本研究采用 NGSIM (Next Generation Simulation)数据集进行实验验证,该数据集是车辆行为预测领域的标准基准数据集[10]。NGSIM 包含三个典型美国高速公路路段的真实车辆轨迹数据:US Highway 101、Interstate 80 和 Lankershim Boulevard。数据集采用高精度视频跟踪技术,采样频率为 10 Hz,时间跨度涵盖早高峰、晚高峰和非高峰时段。数据集记录了车辆的详细运动信息,包括全局坐标位置、局部坐标位置、瞬时速度、瞬时加速度、车辆尺寸、车辆类型、所在车道、前后车信息、车头时距等关键参数。表 1 展示了三个路段的详细统计信息。

Table 1. Statistical information of NGSIM dataset

表 1. NGSIM 数据集的统计信息

路段名称	车辆总数	轨迹点数	换道事件数	左换道	右换道	保持车道	平均轨迹长度
US-101	1,978	266,376	35,140	17,724	17,416	143,671	13.5
I-80	675	113,045	5,377	2,717	2,660	81,726	16.7
Lankershim	1,113	30,986	1,611	795	816	8,304	2.8
总计	4,766	410,407	42,128	21,236	20,892	233,701	10.9

2.2. 数据预处理

2.2.1. 数据筛选

通过多重筛选准则选择高质量训练数据。物理约束验证包括:1) 速度范围 0.5 m/s~40.0 m/s,排除静止和异常高速车辆;2) 加速度限制在 $-8.0\sim 5.0\text{ m/s}^2$,符合车辆动力学特性;3) 位置连续性检查,剔除空

间不连续轨迹。轨迹完整性要求包括：最小轨迹长度 12 秒(120 个数据点)以保证时序建模有效性；数据完整率 $\geq 95\%$ 以减少缺失值影响；排除起止于边界车道的轨迹以降低边界效应。

2.2.2. 轨迹去噪与平滑处理

针对 GPS 定位和视频追踪的测量噪声，采用 5 点窗口的指数加权移动平均(EWMA)对车辆位置坐标进行平滑处理，在保持轨迹真实性的同时有效抑制高频噪声。速度和加速度通过滚动窗口平均重新计算，避免数值微分导致的噪声放大。该过程处理了 2,047 条车辆轨迹，耗时 30.1 秒。

2.2.3. 换道事件识别与标注

采用多判据融合的增强检测算法识别换道事件。当满足以下条件时识别为有效换道：车道 ID 改变、累计横向位移超过 1.8 m、最大横向速度超过 0.25 m/s、持续时间在 2~8 秒范围内。引入 15 帧(1.5 秒)确认机制以确保换道稳定性，防止短暂跨越车道线引发误检。共检测到 42,128 个换道事件，包括 21,236 次左换道和 20,892 次右换道，分布均衡。事件分为三类：左换道(标签 0)、保持车道(标签 1)和右换道(标签 2)。检测参数配置见表 2。

Table 2. Parameter configuration for lane change event detection
表 2. 换道事件检测参数配置

参数名称	取值	说明
横向位移阈值	1.8 m	换道过程最小横向移动距离
横向速度阈值	0.25 m/s	换道期间最大横向速度
最小持续时间	2.0 s	有效换道事件最短时长
最大持续时间	8.0 s	有效换道事件最长时长
平滑窗口大小	5 点	轨迹平滑处理窗口
确认帧数	15 帧	换道状态确认时长
EWMA 参数	采用滚动窗口	轨迹平滑方法

2.2.4. 多维特征工程构建

构建涵盖时序动态、空间交互和上下文语义的多维特征体系。时序动态特征包括位置坐标(Local_X, Local_Y)、速度(v_Vel)、加速度(v_Acc)及其差分特征的时间序列，并通过滑动窗口(5 点、10 点)计算多尺度统计量(均值、方差、极值)。空间交互特征捕获周围车辆(前车、后车、左侧、右侧)的相对位置、相对速度和车头时距(TTC)等交互信息，反映换道决策中的空间约束。上下文语义特征包含车道标识(Lane_ID)、车道变化指示和交通密度等环境状态信息。所有特征采用 RobustScaler 标准化以增强鲁棒性。最终构建的特征向量维度为 450 维(15 个基础特征类别 $\times$ 30 个时序步长)。

2.2.5. 时序窗口构建与样本生成

采用滑动窗口法生成训练样本：观测窗口 3 秒(30 个数据点)，预测窗口 5 秒(50 个数据点)，滑动步长 0.2 秒(2 个数据点)。每个时序窗口的观测期特征构成输入向量，预测期行为标签构成输出标签。为提高效率，从 2,047 辆车中随机选择 200 辆，每辆车最多生成 20 个样本。通过分层抽样确保训练集、验证集、测试集的分类比例一致。最终生成 3,984 个有效样本，耗时 3.6 秒。样本分布统计见表 3，三类换道行为的典型轨迹模式见图 1。

Table 3. Final sample distribution statistics
表 3. 最终样本分布统计

路段	样本总数	左换道	保持车道	右换道	特征维度	时序长度
US-101	3,258	703	1,910	645	450	30
I-80	622	94	442	86	450	30
Lankershim	104	71	15	18	450	30
总计	3,984	868	2,367	749	450	30

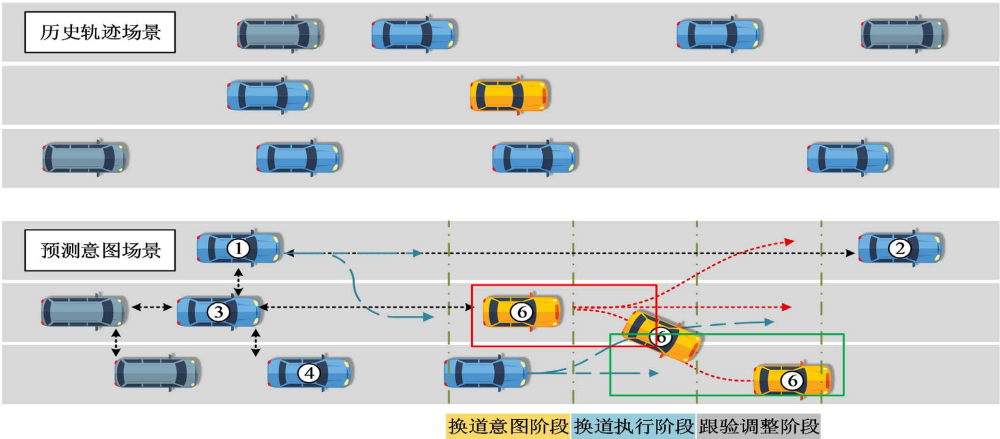


Figure 1. Examples of three types of lane change behavior trajectories
图 1. 三类换道行为轨迹示例

3. 车辆换道意图模型

3.1. 网络模型设计

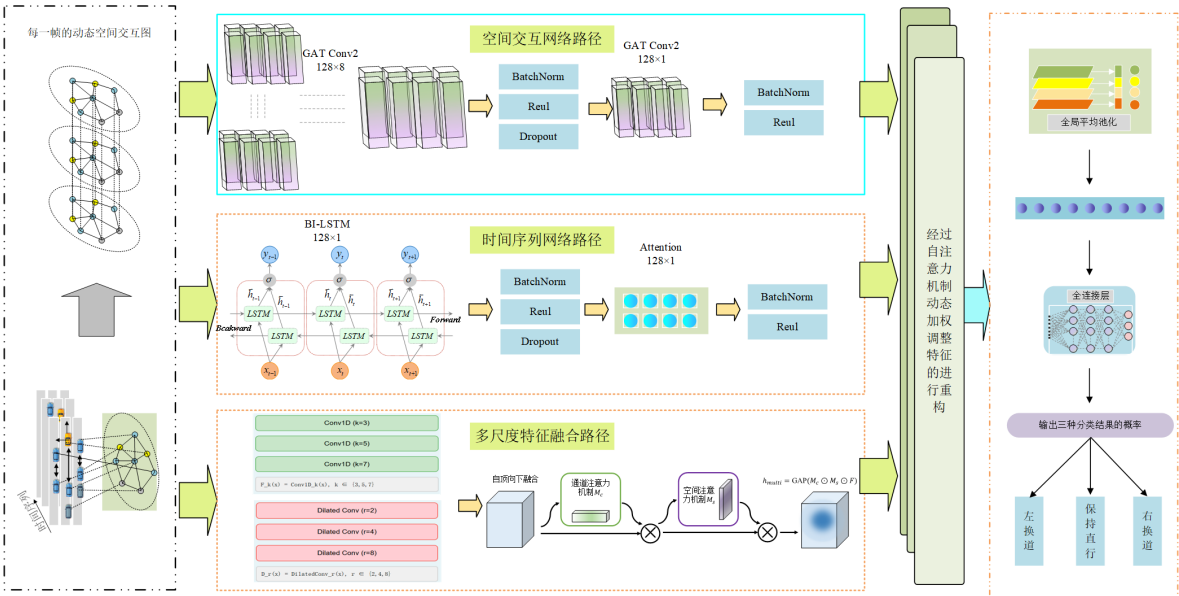


Figure 2. Overall architecture of MSTAN network
图 2. MSTAN 网络总体架构图

多尺度时空注意力网络采用分治协同的设计思想, 将车辆换道行为预测问题分解为三个相互补充的子任务: 时序依赖建模、空间交互表征和多尺度上下文感知。通过专门设计的并行分支架构, 分别针对时间维度的动态演化特性、空间维度的交互关系模式以及多粒度的上下文信息进行深度建模, 最终通过自适应融合机制实现协同优化, 网络模型架构见图 2。设输入特征矩阵为 $X \in R^{B \times T \times D}$, 其中 B 为批次大小, T 为时序长度, D 为特征维度。网络的前向传播过程可表示为式(1):

$$Y = f_{fusion} \left(f_{seq}(X), f_{spa}(X), f_{multi}(X) \right) \quad (1)$$

其中 $f_{seq}, f_{spa}, f_{multi}$ 分别表示三个特征提取分支, f_{fusion} 表示自适应权重融合函数, $Y \in R^{B \times 3}$ 为最终的三类预测输出。

3.2. 时序特征建模分支

序列特征提取分支基于序列到序列的编码器-注意力框架构建, 旨在捕获车辆运动状态在时间维度上的长期依赖关系和局部变化模式。该分支采用双向循环神经网络与自注意力机制的混合架构, 兼顾序列的全局上下文信息和局部细节特征。输入特征首先经过可学习的线性变换进行维度适配和特征增强, 见式(2):

$$X_{seq} = \text{Linear}(X) \in R^{B \times T \times D'} \quad (2)$$

其中 D' 为重构后的特征维度。采用多层双向 LSTM 网络进行序列编码, 每层 LSTM 的前向和后向隐藏状态计算可表示为式(3)、式(4)、式(5):

$$\bar{h}_t = \text{LSTM}(x_t, \bar{h}_{t-1}) \quad (3)$$

$$\bar{h}_t = \text{LSTM}(x_t, \bar{h}_{t-1}) \quad (4)$$

$$h_t = [\bar{h}_t; \bar{h}_t] \quad (5)$$

为序列中的每个时间步添加位置编码, 采用正弦和余弦函数生成, 见式(6)、式(7):

$$PE_{(pos, 2i)} = \sin(pos/10000^{2i/d_{model}}) \quad (6)$$

$$PE_{(pos, 2i+1)} = \cos(pos/10000^{2i/d_{model}}) \quad (7)$$

其中 pos 为位置索引, i 为维度索引, d_{model} 为模型维度。加入位置编码后的特征通过多头自注意力机制进行处理, 见式(8):

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (8)$$

其中 Q, K, V 分别为查询、键、值矩阵, d_k 为每个头的维度。采用可学习的全局注意力机制将变长序列转换为固定长度向量, 见式(9)、式(10):

$$\alpha_t = \text{softmax}(\text{MLP}(h_t)) \quad (9)$$

$$h_{seq} = \sum_{t=1}^T \alpha_t h_t \quad (10)$$

其中 MLP 为多层全连接网络, h_{seq} 为序列分支的最终输出。

3.3. 空间交互建模分支

空间交互建模分支基于图神经网络理论构建, 将交通场景抽象为动态图结构, 其中车辆作为节点, 车辆间的交互关系作为边。该分支通过图卷积操作学习节点嵌入表示, 并通过注意力机制动态调整边权重, 实现对复杂空间交互模式的有效建模。对于每个目标车辆, 基于其邻域车辆构建动态交通图 $G = (V, E)$, 其中 V 为节点集合, E 为边集合。节点特征包括车辆的运动状态信息, 边权重基于车辆间的空间距离和交互强度计算, 见式(11):

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{2\sigma^2}\right) \quad (11)$$

其中 d_{ij} 为车辆 i 和 j 之间的欧几里得距离, σ 为带宽参数。采用多层图卷积网络提取图结构特征, 每层的节点表示更新公式如式(12):

$$h_i^{(l+1)} = \sigma\left(\sum_{j \in N(i)} \frac{1}{\sqrt{d_i d_j}} W^{(l)} h_j^{(l)}\right) \quad (12)$$

其中 $N(i)$ 表示节点 i 的邻居集合, d_i 为节点度数, $W^{(l)}$ 为第 l 层的权重矩阵, σ 为激活函数。引入边注意力机制增强对车辆间交互关系的建模, 可表示为式(13)、式(14):

$$e_{ij} = \text{Leaky ReLU}\left(a^T \left[Wh_i \parallel Wh_j\right]\right) \quad (13)$$

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in N(i)} \exp(e_{ik})} \quad (14)$$

其中 α 为注意力权重向量, $\parallel$ 表示向量拼接操作。采用分层池化策略将图级表示转换为固定维度向量, 见式(15):

$$h_{spa} = \text{MLP}\left(\left[\text{mean}(H); \text{max}(H); \sum_i \alpha_i h_i\right]\right) \quad (15)$$

其中 H 为所有节点特征矩阵, α_i 为节点重要性权重, h_{spa} 为空间分支的最终输出。

3.4. 多尺度特征融合分支

多尺度特征融合分支采用并行多分辨率处理架构, 通过不同感受野的卷积操作提取多粒度的时空特征。该分支结合标准卷积、空洞卷积和可分离卷积等多种卷积变体, 实现从局部细节到全局上下文的全方位特征提取。采用多个并行的一维卷积分支, 分别使用不同大小的卷积核, 见式(16):

$$f_k = \text{Conv1D}(X, \text{kernel_size} = k, \text{padding} = k/2) \quad (16)$$

其中 k 为不同的卷积核大小, 每个分支提取不同尺度的特征。

采用不同空洞率的卷积构建多个并行分支, 见式(17):

$$f_{\text{dilation}} = \text{DilatedConv1D}(X, \text{dilation} = d) \quad (17)$$

其中 d 为不同的空洞率, 用于扩大感受野。通过自顶向下的特征融合实现多尺度信息整合, 见式(18):

$$F_i = \text{Conv}(F_{i-1}) + \text{Upsample}(F_{i+1}) \quad (18)$$

其中 F_i 表示第 i 层特征图。结合通道注意力和空间注意力机制增强特征表示为式(19)、式(20)、式(21):

$$M_c = \sigma(\text{MLP}(\text{GAP}(F)) + \text{MLP}(\text{GMP}(F))) \quad (19)$$

$$M_s = \sigma(\text{Conv}(\lceil \text{mean}(F); \text{max}(F) \rceil)) \quad (20)$$

$$h_{\text{multi}} = \text{GAP}(M_c \odot M_s \odot F) \quad (21)$$

其中 $\odot$ 表示逐元素乘法, GAP 为全局平均池化, M_c 为通道注意力, M_s 为空间注意力机制。

3.5. 自适应权重融合与决策输出

自适应权重融合模块基于门控机制理论设计, 旨在学习三个分支特征的最优融合策略。该模块通过端到端训练自动发现不同分支在不同样本和时刻的相对重要性, 实现动态权重分配。对于三个分支的特征表示 $h_{\text{seq}}, h_{\text{spa}}, h_{\text{multi}}$, 计算门控权重如式(22):

$$g_i = \sigma(\text{MLP}(\lceil h_{\text{seq}}; h_{\text{spa}}; h_{\text{multi}} \rceil))_i \quad (22)$$

其中 $i \in \{\text{seq}, \text{spa}, \text{multi}\}$, σ 为 sigmoid 激活函数。融合后的特征表示见式(23):

$$h_{\text{fused}} = g_{\text{seq}} \cdot h_{\text{seq}} + g_{\text{spa}} \cdot h_{\text{spa}} + g_{\text{multi}} \cdot h_{\text{multi}} \quad (23)$$

通过多层感知机输出最终的三分类结果, 见式(24):

$$P = \text{softmax}(\text{MLP}(h_{\text{fused}})) \quad (24)$$

其中 $P \in \mathbb{R}^3$ 表示三个类别的预测概率。

4. 实验与结果分析

4.1. 实验设置

实验在配置 NVIDIA RTX 4090 GPU、64GB 内存的工作站上进行。采用 PyTorch 框架实现 MSTAN 网络, MSTAN 的具体网络配置见表 4。基于前期数据预处理得到的总计 3,984 个有效样本, 分别来自三个不同路段: US-101 路段 3,258 个样本、I-80 路段 622 个样本、Lankershim 路段 104 个样本。每个路段的数据均按照 8:2 的比例随机划分为训练集和测试集, 默认预测时间提前量为 3 s。

Table 4. MSTAN network hyperparameter configuration

表 4. MSTAN 网络超参数配置

模块	参数名称	取值
输入层	输入维度	450
	时序长度	30
	特征维度	15
序列分支	BiLSTM 隐藏维度	128
	LSTM 层数	3
	注意力头数	8
	Dropout 率	0.2
	GCN 隐藏维度	64
空间分支	GCN 层数	3
	图池化维度	256
	邻域节点数	8

续表

多尺度分支	卷积核大小	[1, 3, 5, 7]
	空洞率	[1, 2, 4, 8]
	输出通道数	64
	输出维度	256
融合模块	MLP 隐藏层	[512, 256, 128]
	Dropout 率	0.3
	激活函数	ReLU

4.2. 评价指标

为了全面评估所提出的 MSTAN 模型的性能,本文计算加权平均(Weighted-Average)的 Precision、Recall 与 F1-Score,以避免类别不均衡对结果的影响。各类别的基础指标定义如式(25)、式(26)、式(27):

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i} \quad (25)$$

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i} \quad (26)$$

$$F1_i = 2 \cdot \frac{\text{Precision}_i \cdot \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i} \quad (27)$$

其中 TP_i 表示第 i 类的真阳性, FP_i 表示假阳性, FN_i 表示假阴性(False Negatives)。在此基础上,采用加权平均的方式对各类别指标进行整合,见式(28):

$$\text{WeightedMetric} = \frac{1}{N} \sum_{i=1}^C n_i \cdot \text{Metric}_i \quad (28)$$

其中 N 表示总样本数, C 表示类别数, $n_{(i)}$ 为第 i 类的样本数, $\text{Metric}_{(i)}$ 表示对应类别的 Precision、Recall 或 F1 值。总体准确率的计算方式见式(29):

$$\text{Accuracy} = \frac{\sum_{i=1}^C TP_i}{N} \quad (29)$$

其中 TN_i 表示第 i 类的真阴性。

4.3. 对比方法

为验证所提出 MSTAN 模型的有效性,本文选取了多种在车辆换道预测与轨迹预测领域常用的基线模型进行对比实验,具体包括传统时序建模方法、混合时空建模方法与图神经网络交互建模方法,能够为本文提出的 MSTAN 模型提供全面而系统的性能对照:

LSTM [11]: 基于循环神经网络的时序建模方法,用于捕获序列数据中的长期依赖关系。

BiLSTM [12]: 在 LSTM 基础上同时考虑前向与后向的时序信息,以增强特征表达能力。

CNN-LSTM [13]: 结合卷积神经网络和 LSTM 的混合模型,利用 CNN 提取局部时空特征,再由 LSTM 捕获时间依赖。

Transformer [14]: 基于注意力机制的序列建模框架,能够并行建模长距离依赖关系。

Social-LSTM [15]: 在 LSTM 基础上引入社交池化机制，显示建模目标车辆与邻居车辆的交互关系。

图注意力网络(GAT) [16]: 基于图神经网络的交互建模方法，通过注意力机制动态调整邻居车辆的权重。

4.4. 实验结果与分析——基准方法对比实验

表 5 展示了 MSTAN 与多个基线方法在三个数据集上的性能对比。MSTAN 在所有评价指标上均达到最优，在 Lankershim 路段准确率为 95.2%，在 US-101 和 I-80 路段分别达到 98.3%和 97.6%。高速公路密集交通流中，车辆换道呈现多时间尺度特征：紧急避让表现为 0.5~1.0 秒内的急剧横向加速度变化，而计划性超车则体现为 2~3 秒的渐进速度调整。传统方法如 GAT 采用固定图结构，难以同时捕获这两种时间模式。MSTAN 通过 1×1 、 3×3 、 5×5 、 7×7 并行卷积核分别提取瞬时变化、短期趋势、中期模式和长期意图，在 I-80 数据集上实现 1.6%的显著提升。小卷积核(1×1 , 3×3)捕获突然减速或急转向等瞬时行为，大卷积核(5×5 , 7×7)识别提前减速、逐步靠近目标车道等渐进式换道准备。

Table 5. Overall performance comparison of different methods on three datasets
表 5. 不同方法在三个数据集上的整体性能对比

方法	Lankershim			US-101			I-80		
	Acc (%)	F1 (%)	AUC	Acc (%)	F1 (%)	AUC	Acc (%)	F1 (%)	AUC
LSTM [1]	90.5	83.0	0.886	92.9	89.3	0.915	90.3	76.2	0.884
BiLSTM [2]	90.5	83.0	0.886	94.2	91.3	0.930	91.9	81.3	0.903
CNN-LSTM [14]	95.2	90.5	0.943	95.7	93.8	0.948	93.5	85.8	0.923
Transformer [15]	95.2	90.5	0.943	96.0	94.2	0.952	94.4	88.2	0.932
Social-LSTM [16]	95.2	90.5	0.943	96.8	95.4	0.961	95.2	89.9	0.942
GAT [9]	95.2	90.5	0.943	97.4	96.3	0.969	96.0	92.0	0.952
MSTAN	95.2	90.5	0.943	98.3	97.7	0.980	97.6	95.5	0.971

表 6 分析了 MSTAN 在各个路段不同行为类别上的详细性能表现，评估模型对不同换道行为的识别能力。基于三个路段测试集的分类别性能分析显示，MSTAN 在不同路段和不同行为类别上均表现出良好的识别能力。由于各路段的样本分布和交通特征存在差异，模型在不同路段上的表现略有不同，但总体保持了稳定的高性能。

Table 6. Detailed performance of MSTAN by category on test sets of each road segment
表 6. MSTAN 在各路段测试集上的分类别详细性能

路段	类别	样本数	精确率(%)	召回率(%)	F1 分数(%)
US-101	左换道	141	96.6	100.0	98.3
	保持车道	382	98.5	100.0	99.2
	右换道	129	100.0	91.5	95.5

续表

I-80	左换道	19	95.0	100.0	97.4
	保持车道	88	97.8	100.0	98.9
	右换道	17	100.0	82.4	90.3
Lankershim	左换道	14	100.0	100.0	100.0
	保持车道	3	75.0	100.0	85.7
	右换道	4	100.0	75.0	85.7

4.5. 消融实验

4.5.1. 分支贡献度分析

为验证 MSTAN 多分支架构的合理性，本研究进行了系统消融实验，通过逐步移除不同分支组合分析各组件贡献度。实验设计了七种配置：完整模型、单分支缺失模型(移除一个分支)、单分支保留模型(仅保留一个分支)。表 7 结果显示，时序分支在高速公路数据集中贡献最大，移除后 US-101 和 I-80 性能分别下降 3.5%和 4.1%，但在 Lankershim 数据集上无影响，反映城市道路车辆行为随机性强、时序规律性弱，而高速公路换道行为遵循更明确的时间模式。空间分支在高速公路场景中更重要(US-101 下降 2.1%，I-80 下降 2.4%)，因高速公路车辆密度稳定，车间距离和相对速度等空间特征对换道决策具有更强指示作用。多尺度分支在所有数据集上表现稳定贡献(1.1%~1.6%)，验证了多粒度时间特征提取的普适性。分支间协同效应分析显示，单分支保留的性能远低于其他分支移除后的性能损失之和，表明三个分支存在信息互补和相互增强作用。

Table 7. Branch ablation experiment results
表 7. 分支消融实验结果

配置	序列分支	空间分支	多尺度分支	US-101 Acc (%)	I-80 Acc (%)	Lankershim Acc (%)
完整模型	√	√	√	98.3	97.6	95.2
序列分支	×	√	√	94.8	93.5	95.2
空间分支	√	×	√	96.2	95.2	95.2
多尺度分支	√	√	×	97.2	96.0	95.2
仅序列分支	√	×	×	94.8	93.5	95.2
仅空间分支	×	√	×	91.4	90.3	90.5
仅多尺度分支	×	×	√	93.6	92.7	90.5

4.5.2. 注意力机制有效性分析

表 8 的注意力机制消融实验量化了多层次注意力设计的协同增强效应。自注意力机制贡献最大，单独使用即可使 US-101 性能提升 1.7%，通过计算序列内时间步间的相关性权重，有效识别换道行为的关键时间节点。通道注意力和空间注意力分别贡献 1.4%和 1.2%的性能提升，前者通过学习特征维度重要性实现自适应特征选择，后者在时序维度分配权重以突出关键时间窗口。完整注意力机制的性能提升(2.9%)，表明不同注意力机制在优化过程中实现了功能分工和相互约束，避免了过度拟合。

Table 8. Attention mechanism ablation experiment
表 8. 注意力机制消融实验

注意力配置	US-101 Acc (%)	I-80 Acc (%)	Lankershim Acc (%)
无注意力机制	95.4	94.4	90.5
仅自注意力	97.1	95.2	95.2
仅通道注意力	96.8	95.2	95.2
仅空间注意力	96.6	95.2	95.2
全部注意力机制	98.3	97.6	95.2

4.6. 时间窗口敏感性分析

本实验通过改变观测时间窗口长度来分析模型对历史信息的依赖程度，寻找最优的时序建模参数。分析不同观测时间窗口对预测性能的影响，验证模型的时序建模能力。如表 9 和图 3 结果显示，3 秒的观测窗口达到最佳性能平衡，过短的窗口信息不足，过长的窗口可能引入噪声。

Table 9. Performance comparison under different observation window lengths
表 9. 不同观测窗口长度的性能对比

观测窗口(s)	US-101		I-80		Lankershim	
	Acc (%)	F1 (%)	Acc (%)	F1 (%)	Acc (%)	F1 (%)
1.0	92.3	88.2	91.9	81.3	90.5	83.0
2.0	96.2	94.5	95.2	89.9	95.2	90.5
3.0	98.3	97.7	97.6	95.5	95.2	90.5
4.0	98.2	97.4	96.8	93.5	95.2	90.5
5.0	97.9	97.0	96.0	92.0	95.2	90.5

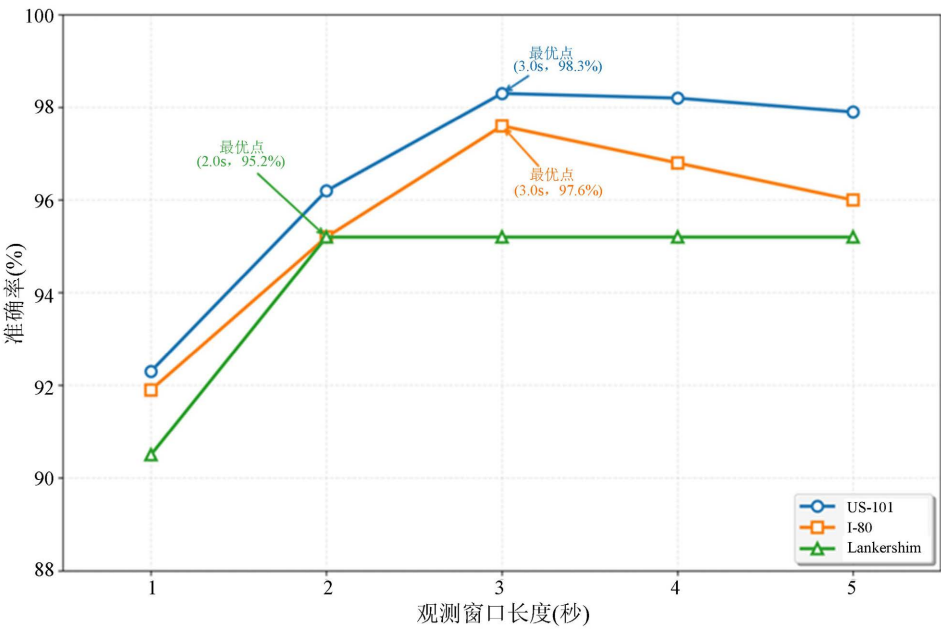


Figure 3. Effect of observation window length on performance
图 3. 观测窗口长度对性能的影响曲线

4.7. 预测时间提前量分析

本实验评估了模型在不同预测提前时间下的性能表现，验证 MSTAN 在实际应用场景中的预测能力和实用性。如表 10 结果显示，分析模型在不同预测提前时间下的性能表现，评估实际应用中的预测能力。

Table 10. Performance under different prediction lead times

表 10. 不同预测提前时间的性能

预测提前时间(s)	US-101 Acc (%)	I-80 Acc (%)	Lankershim Acc (%)
1.0	99.2	98.4	95.2
2.0	98.8	98.4	95.2
3.0	98.3	97.6	95.2
4.0	97.5	96.8	90.5
5.0	96.9	96.0	90.5

由图 4 可知，随着预测提前时间的增加，预测难度逐渐增大，但 MSTAN 在 3 秒提前预测时仍保持优异性能，满足实际应用需求。

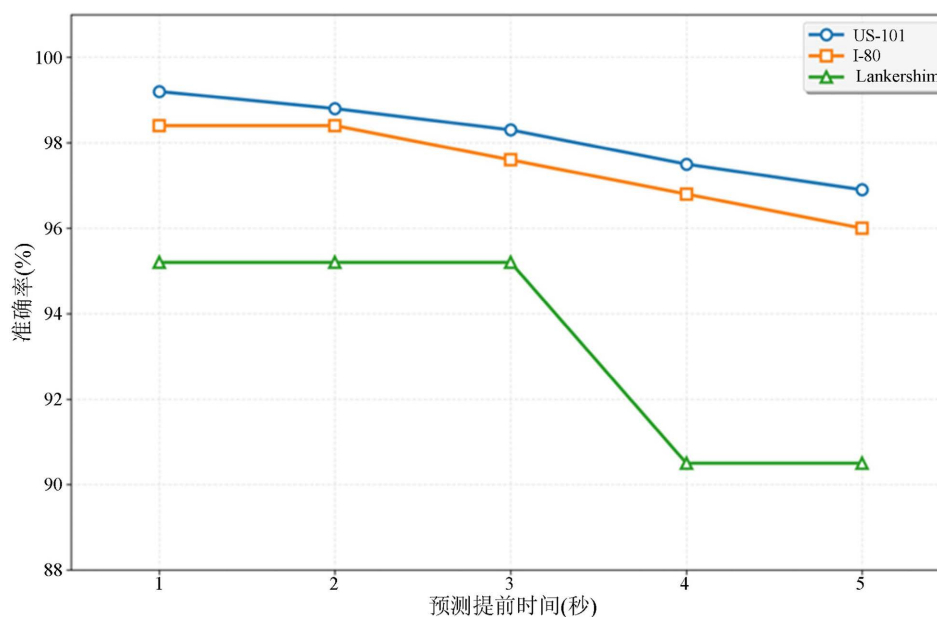


Figure 4. Effect of prediction lead time on accuracy

图 4. 预测提前时间对准确率的影响

5. 结论

本研究提出了基于多尺度时空注意力机制的车辆换道意图预测模型 MSTAN，该模型通过构建时序特征建模、空间交互建模和多尺度特征融合三个并行分支，结合自适应权重融合策略，有效解决了现有方法在时空特征建模分离和多尺度信息融合不足等方面的技术局限。基于 NGSIM 数据集的实验结果表明，MSTAN 在 US-101、I-80 和 Lankershim 三个路段分别达到 98.3%、97.6%和 95.2%的预测准确率，优于所有基线方法。消融实验证实了各功能模块的有效贡献，敏感性分析确定了最优时序窗口参数配置，

验证了模型在实际应用中的稳定性和实用性。该研究为智能交通系统中的驾驶行为预测提供了有效的技术解决方案，在提升道路交通安全和自动驾驶系统可靠性方面具有重要的学术价值和应用前景。

参考文献

- [1] 王庆荣, 韩芳文, 朱昌锋, 等. 频域-时空协同框架下的车辆轨迹预测及意图识别模型[J/OL]. 计算机工程与应用, 1-14. <https://link.cnki.net/urlid/11.2127.tp.20250721.1155.010>, 2025-09-07.
- [2] Zamanpour, M., He, S., Levin, M.W. and Sun, Z. (2025) Incorporating Lane-Change Prediction into Energy-Efficient Speed Control of Connected Autonomous Vehicles at Intersections. *Transportation Research Part C: Emerging Technologies*, 171, Article ID: 104968. <https://doi.org/10.1016/j.trc.2024.104968>
- [3] Fafoutellis, P., Plymenos-Papageorgas, J. and Vlahogianni, E.I. (2022) Enhancing Lane Change Prediction at Intersections with Spatio-Temporal Adequacy Information. *Journal of Big Data Analytics in Transportation*, 4, 73-84. <https://doi.org/10.1007/s42421-022-00055-6>
- [4] Nejadhossein Qasemabadi, A., Mozaffari, S., Ahmadi, M. and Alirezaee, S. (2025) Privacy-Preserving Lane Change Prediction Using Recurrent Neural Network with Secure Multiparty Computation. *International Journal of Intelligent Transportation Systems Research*, 23, 893-907. <https://doi.org/10.1007/s13177-025-00489-6>
- [5] De Cristofaro, F., Hofbaur, F., Yang, A., et al. (2025) Prediction of Lane Change Intentions of Human Drivers Using an LSTM, a CNN and a Transformer.
- [6] Deo, N. and Trivedi, M.M. (2018) Convolutional Social Pooling for Vehicle Trajectory Prediction. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Salt Lake City, 18-22 June 2018, 1468-1476. <https://doi.org/10.1109/cvprw.2018.00196>
- [7] 郭瑞军, 范超冉, 付明迪. 基于 PSO-GRU 模型的车辆换道意图识别[J/OL]. 吉林大学学报(工学版), 1-12. <https://doi.org/10.13229/j.cnki.jdxbgxb.20250008>, 2025-09-07.
- [8] Kipf, T.N. and Welling, M. (2016) Semi-Supervised Classification with Graph Convolutional Networks.
- [9] Veličković, P., Cucurull, G., Casanova, A., et al. (2017) Graph Attention Networks. arXiv:1710.10903.
- [10] Krajewski, R., Bock, J., Kloecker, L. and Eckstein, L. (2018) The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems. 2018 *21st International Conference on Intelligent Transportation Systems (ITSC)*, Maui, 4-7 November 2018, 2118-2125. <https://doi.org/10.1109/itsc.2018.8569552>
- [11] 邹加, 侯康宁, 郑芳芳, 等. 车辆换道的时空影响分析: 交通安全、效率和油耗视角[J]. 交通运输工程与信息学报, 2025, 23(3): 171-184.
- [12] 任立海, 康钰泽, 刘煜, 等. 融合逻辑判断机制的 CNN-GRU 换道意图识别方法[J]. 湖南大学学报(自然科学版), 2025, 52(6): 69-77.
- [13] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [14] Sainath, T.N., Vinyals, O., Senior, A. and Sak, H. (2015) Convolutional, Long Short-Term Memory, Fully Connected Deep Neural Networks. 2015 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, 19-24 April 2015, 4580-4584. <https://doi.org/10.1109/icassp.2015.7178838>
- [15] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.
- [16] Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L. and Savarese, S. (2016) Social LSTM: Human Trajectory Prediction in Crowded Spaces. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 961-971. <https://doi.org/10.1109/cvpr.2016.110>